



(12) 发明专利申请

(10) 申请公布号 CN 101966337 A

(43) 申请公布日 2011. 02. 09

(21) 申请号 201010279118. 4

C12N 15/63 (2006. 01)

(22) 申请日 2003. 11. 21

G01N 33/53 (2006. 01)

(30) 优先权数据

G01N 33/534 (2006. 01)

10254601. 0 2002. 11. 22 DE

G01N 33/535 (2006. 01)

(62) 分案原申请数据

G01N 33/574 (2006. 01)

200380103782. 1 2003. 11. 21

G01N 33/577 (2006. 01)

(71) 申请人 加尼梅德医药品有限公司

C12Q 1/68 (2006. 01)

地址 德国美因茨

C07K 14/47 (2006. 01)

(72) 发明人 U·萨海恩 O·图雷茨

C07K 16/30 (2006. 01)

M·柯斯勒司克

A61P 35/00 (2006. 01)

(74) 专利代理机构 中国国际贸易促进委员会专

利商标事务所 11038

代理人 罗菊华

(51) Int. Cl.

A61K 45/00 (2006. 01)

A61K 48/00 (2006. 01)

A61K 39/395 (2006. 01)

A61K 39/39 (2006. 01)

A61K 38/17 (2006. 01)

C12N 5/10 (2006. 01)

C12N 15/11 (2006. 01)

C12N 15/12 (2006. 01)

权利要求书 8 页 说明书 48 页 序列表 76 页

附图 48 页

(54) 发明名称

肿瘤内差异表达的基因产物及其应用

(57) 摘要

本发明涉及与肿瘤相关的基因产物的鉴别, 涉及编码所述产物的核酸。本发明也涉及如下疾病的治疗和诊断, 所述疾病中基因产物的异常表达与肿瘤相关, 蛋白质、多肽和肽的表达与肿瘤相关。本发明还涉及编码所述多肽、肽和蛋白质的核酸。

1. 一种药物组合物,包含抑制肿瘤相关抗原的表达或活性之药剂,所述肿瘤相关抗原具有由选自如下的核酸编码的序列:

(a) 包含选自 SEQ ID NO:1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分或其衍生物,

(b) 在严格条件下与 (a) 所述核酸杂交的核酸,

(c) 与 (a) 或 (b) 的核酸简并的核酸,以及

(d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。

2. 一种药物组合物,包含有肿瘤抑制活性的药剂,其对表达或异常表达肿瘤相关抗原的细胞有选择性,所述肿瘤相关抗原具有由选自如下的核酸编码的序列:

(a) 包含选自 SEQ ID NO:1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分或其衍生物,

(b) 在严格条件下与 (a) 所述核酸杂交的核酸,

(c) 与 (a) 或 (b) 的核酸简并的核酸,以及

(d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。

3. 按照权利要求 2 的药物组合物,其中所述药剂造成细胞死亡的诱导,细胞生长的减少,细胞膜的损伤或细胞因子的分泌。

4. 按照权利要求 1 或 2 的药物组合物,其中所述药剂是与编码肿瘤相关抗原的核酸选择性杂交的反义核酸。

5. 按照权利要求 1 或 2 的药物组合物,其中所述药剂是选择性结合肿瘤相关抗原的抗体。

6. 按照权利要求 2 的药物组合物,其中所述药剂是选择性结合肿瘤相关抗原的补体激活抗体。

7. 一种药物组合物,其包含给予时选择性地增加 HLA 分子和肿瘤相关抗原或其部分之间的复合物的量之药剂,所述肿瘤相关抗原具有由选自如下的核酸编码的序列:

(a) 包含选自 SEQ ID NO:1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分或其衍生物,

(b) 在严格条件下与 (a) 所述核酸杂交的核酸,

(c) 与 (a) 或 (b) 的核酸简并的核酸,以及

(d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。

8. 按照权利要求 7 的药物组合物,其中所述药剂包含选自如下的一个或多个成分:

(i) 肿瘤相关抗原或其部分,

(ii) 编码肿瘤相关抗原或其部分的核酸,

(iii) 表达肿瘤相关抗原或其部分的宿主细胞,以及

(iv) 肿瘤相关抗原或其部分和 HLA 分子之间的分离的复合物。

9. 按照权利要求 1,2 或 7 的药物组合物,其中所述药剂包含两个或多个如下药剂,在每种情况下所述药剂选择性抑制不同肿瘤相关抗原的表达或活性,在每种情况下所述药剂对表达不同肿瘤相关抗原的细胞有选择性或者所述药剂增加 HLA 分子和不同肿瘤相关抗原或其部分之间的复合物的量。其中有至少一个所述肿瘤相关抗原具有由选自如下的核酸编码的序列:

(a) 包含选自 SEQ ID NO :1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分或其衍生物,

(b) 在严格条件下与 (a) 所述核酸杂交的核酸,

(c) 与 (a) 或 (b) 的核酸简并的核酸,以及

(d) 与 (a)、(b) 或 (c) 的核酸互补的核酸,

10. 一种药物组合物,包含一个或多个选自如下的成分:

(i) 肿瘤相关抗原或其部分,

(ii) 编码肿瘤相关抗原或其部分的核酸,

(iii) 结合肿瘤相关抗原或其部分的抗体,

(iv) 与编码肿瘤相关抗原的核酸特异性杂交的反义核酸,

(v) 表达肿瘤相关抗原或其部分的宿主细胞,以及

(vi) 肿瘤相关抗原或其部分和 HLA 分子之间的分离的复合物,

所述肿瘤相关抗原具有由选自如下的核酸编码的序列:

(a) 包含选自 SEQ ID NO :1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分或其衍生物,

(b) 在严格条件下与 (a) 所述核酸杂交的核酸,

(c) 与 (a) 或 (b) 的核酸简并的核酸,以及

(d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。

11. 按照权利要求 8 或 10 的药物组合物,其中 (ii) 的核酸存在于表达载体中。

12. 按照权利要求 8 或 10 的药物组合物,其中 (ii) 的核酸功能性地连接于启动子。

13. 按照权利要求 8 或 10 的药物组合物,其中所述宿主细胞分泌肿瘤相关抗原或其部分。

14. 按照权利要求 8 或 10 的药物组合物,其中所述宿主细胞还表达结合肿瘤相关抗原或其部分的 HLA 分子。

15. 按照权利要求 14 的药物组合物,其中所述宿主细胞以重组的方式表达 HLA 分子和 / 或肿瘤相关抗原或其部分。

16. 按照权利要求 14 的药物组合物,其中所述宿主细胞内源性表达 HLA 分子。

17. 按照权利要求 8、10、14 或 16 的药物组合物,其中所述宿主细胞是抗原递呈细胞。

18. 按照权利要求 17 的药物组合物,其中所述抗原递呈细胞是树突细胞或巨噬细胞。

19. 按照权利要求 8、10 和 13-18 中任一项的药物组合物,其中所述宿主细胞是非增殖性的。

20. 按照权利要求 5 或 10 的药物组合物,其中所述抗体是单克隆抗体。

21. 按照权利要求 5 或 10 的药物组合物,其中所述抗体是嵌合或人源化抗体。

22. 按照权利要求 5 或 10 的药物组合物,其中所述抗体是天然抗体的片段。

23. 按照权利要求 5 或 10 的药物组合物,其中所述抗体偶联于治疗或诊断药剂。

24. 按照权利要求 4 或 10 的药物组合物,其中所述反义核酸包含编码肿瘤相关抗原的核酸之 6-50 个连续核苷酸的序列。

25. 按照权利要求 8 和 10-13 中任一项的药物组合物,其中由所述药物组合物提供的肿瘤相关抗原或其部分结合表达异常量的所述肿瘤相关抗原或其部分的细胞之表面的 MHC

分子。

26. 按照权利要求 25 的药物组合物,其中所述结合造成细胞溶解反应和 / 或诱导细胞因子的释放。

27. 按照权利要求 1-26 中任一项的药物组合物,进一步包含药物上可接受的载体和 / 或辅佐剂。

28. 按照权利要求 27 的药物组合物,其中所述辅佐剂是皂甙、GM-CSF、CpG、细胞因子或趋化因子。

29. 按照权利要求 1-28 中任一项的药物组合物,其可被用于治疗特征在于肿瘤相关抗原表达或异常表达的疾病。

30. 按照权利要求 29 的药物组合物,其中所述疾病是癌症。

31. 按照权利要求 29 的药物组合物,其中所述疾病是肺肿瘤、乳腺肿瘤、前列腺肿瘤、黑素瘤、结肠肿瘤、胃肿瘤、胰腺肿瘤、耳鼻咽喉肿瘤、肾细胞癌或子宫颈癌、结肠癌或乳腺癌。

32. 按照权利要求 1-31 中任一项的药物组合物,其中所述肿瘤相关抗原包含选自 SEQ ID NO :9-19,45-48,60-66,85,90-97,100-102,105,106,111-116,118,120,123,124 和 135-137 的氨基酸序列,其部分或其衍生物。

33. 一种诊断特征在于肿瘤相关抗原表达或异常表达的疾病之方法,所述方法包括:用具有由选自如下的核酸编码的序列之肿瘤相关抗原:(a) 包含选自 SEQ ID NO :1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分或其衍生物,(b) 在严格条件下与 (a) 所述核酸杂交的核酸,(c) 与 (a) 或 (b) 的核酸简并的核酸,以及 (d) 与 (a)、(b) 或 (c) 的核酸互补的核酸,

进行分离自患者的生物样品中

(i) 编码肿瘤相关抗原或其部分的核酸的检测,和 / 或

(ii) 肿瘤相关抗原或其部分的检测,和 / 或

(iii) 肿瘤相关抗原或其部分的抗体的检测,和 / 或

(iv) 特异于肿瘤相关抗原或其部分的细胞毒性或 T 辅助淋巴细胞的检测。

34. 按照权利要求 33 的方法,其中所述检测包括

(i) 将生物样品与药剂接触,所述药剂特异性结合编码肿瘤相关抗原或其部分的核酸,结合肿瘤相关抗原或其部分,结合抗体或细胞毒性或 T 辅助淋巴细胞,以及

(ii) 检测药剂与核酸或其部分、与肿瘤相关抗原或其部分、与抗体或与细胞毒性或 T 辅助淋巴细胞之间复合物的形成。

35. 按照权利要求 33 或 34 的方法,其中所述检测与相当的正常生物样品中的检测相比较。

36. 按照权利要求 33-35 中任一项的方法,其中所述疾病以两个或多个不同肿瘤相关抗原的表达或异常表达为特征,并且其中所述检测包括编码所述两个或多个不同的肿瘤相关抗原或其部分的两个或多个核酸的检测,所述两个或多个不同的肿瘤相关抗原或其部分的检测,结合所述两个或多个不同的肿瘤相关抗原或其部分的两个或多个抗体的检测,或特异于所述两个或多个不同的肿瘤相关抗原的两个或多个细胞毒性或 T 辅助淋巴细胞的检测。

37. 按照权利要求 33-36 中任一项的方法,其中所述核酸或其部分使用特异性杂交所

述核酸或其部分的多聚核苷酸探针进行检测。

38. 按照权利要求 37 的方法,其中所述多聚核苷酸探针包含编码肿瘤相关抗原的核酸之 6-50 个连续核苷酸序列。

39. 按照权利要求 33-36 中任一项的方法,其中所述核酸或其部分通过选择性扩增所述核酸或其部分进行检测。

40. 按照权利要求 33-36 中任一项的方法,其中待检测的肿瘤相关抗原或其部分在与 MHC 分子的复合物中。

41. 按照权利要求 40 的方法,其中所述 MHC 分子是 HLA 分子。

42. 按照权利要求 33-36 和 40-41 中任一项的方法,其中所述肿瘤相关抗原或其部分使用特异性结合所述肿瘤相关抗原或所述其部分的抗体进行检测。

43. 按照权利要求 33-36 中任一项的方法,其中所述抗体使用特异性结合所述抗体的蛋白质或肽进行检测。

44. 一种用于测定特征在于肿瘤相关抗原表达或异常表达之疾病的缓解、过程或开始的方法,所述方法包含监测来自患有所述疾病或被怀疑患有所述疾病的患者之样品的选自如下的一个或多个参数:

(i) 编码肿瘤相关抗原或其部分的核酸的量,

(ii) 肿瘤相关抗原或其部分的量,

(iii) 结合肿瘤相关抗原或其部分的抗体的量,

(iv) 特异于肿瘤相关抗原或其部分和 MHC 分子之间复合物的溶细胞性 T 细胞或细胞因子释放 T 细胞的量,所述肿瘤相关抗原具有由选自如下的核酸编码的序列:

(a) 包含选自 SEQ ID NO:1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分或其衍生物,

(b) 在严格条件下与 (a) 所述核酸杂交的核酸,

(c) 与 (a) 或 (b) 的核酸简并的核酸,以及

(d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。

45. 按照权利要求 44 的方法,所述方法包括测定第一个样品中第一个时间点和另一个样品中第二个时间点的参数,并且其中通过比较两个样品测定疾病的过程。

46. 按照权利要求 44 或 45 的方法,其中所述疾病特征在于两个或多个不同肿瘤相关抗原表达或异常表达,并且其中监测包含监测:

(i) 编码所述两个或多个不同肿瘤相关抗原或其部分的两个或多个核酸的量,

(ii) 所述两个或多个不同的肿瘤相关抗原或其部分的量,

(iii) 结合所述两个或多个肿瘤相关抗原或其部分的两个或多个抗体的量,和 / 或

(iv) 特异于所述两个或多个不同肿瘤相关抗原或其部分和 MHC 分子之间复合物的两个或多个溶细胞性 T 细胞或细胞因子释放 T 细胞的量。

47. 按照权利要求 44-46 中任一项的方法,其中所述核酸或其部分的量使用特异性杂交所述核酸或其部分的多聚核苷酸探针进行监测。

48. 按照权利要求 47 的方法,其中所述多聚核苷酸包含编码肿瘤相关抗原的核酸之 6-50 个连续核苷酸的序列。

49. 按照权利要求 44-46 中任一项的方法,其中所述核酸或其部分的量通过选择性扩

增所述核酸或其部分进行监测。

50. 按照权利要求 44-46 中任一项的方法,其中所述肿瘤相关抗原或其部分的量使用特异性结合所述肿瘤相关抗原或其部分的抗体进行监测。

51. 按照权利要求 44-46 中任一项的方法,其中所述抗体的量使用特异性结合抗体的蛋白质或肽进行监测。

52. 按照权利要求 44-46 中任一项的方法,其中所述溶细胞性 T 细胞或细胞因子释放 T 细胞的量使用递呈肿瘤相关抗原或其部分和 MHC 分子之间复合物的细胞进行监测。

53. 按照权利要求 37-38,42-43,47-48 和 50-52 中任一项的方法,其中多聚核苷酸探针、抗体、蛋白质或肽或细胞用可检测的方式进行标记。

54. 按照权利要求 53 的方法,其中所述的可检测的标志物是放射活性的标志物或酶标志物。

55. 按照权利要求 33-54 中任一项的方法,其中所述样品包含体液和 / 或机体组织。

56. 一种治疗特征在于肿瘤相关抗原表达或异常表达的疾病的方法,所述方法包括施用按照权利要求 1-32 中任一项的药物组合物,所述肿瘤相关抗原具有由选自如下的核酸编码的序列:

(a) 包含选自 SEQ ID NO:1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分或其衍生物,

(b) 在严格条件下与 (a) 所述核酸杂交的核酸,

(c) 与 (a) 或 (b) 的核酸简并的核酸,以及

(d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。

57. 一种治疗、诊断或监测特征在于肿瘤相关抗原表达或异常表达的疾病,所述方法包括施用结合所述肿瘤相关抗原或其部分并偶联于治疗或诊断药剂的抗体,所述肿瘤相关抗原具有由选自如下的核酸编码的序列:

(a) 包含选自 SEQ ID NO:1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分或其衍生物,

(b) 在严格条件下与 (a) 所述核酸杂交的核酸,

(c) 与 (a) 或 (b) 的核酸简并的核酸,以及

(d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。

58. 按照权利要求 42,50 或 57 的方法,其中所述抗体是单克隆抗体。

59. 按照权利要求 42,50 或 57 的方法,其中所述抗体是嵌合或人源化抗体。

60. 按照权利要求 42,50 或 57 的方法,其中所述抗体是天然抗体的片段。

61. 一种治疗具有特征在于肿瘤相关抗原表达或异常表达的疾病的患者之方法,所述方法包括:

(i) 取出来自所述患者的含有免疫活性细胞的样品,

(ii) 在有利于针对抗所述肿瘤相关抗原或其部分的溶细胞性 T 细胞或细胞因子释放 T 细胞产生的条件下,使所述样品与表达所述肿瘤相关抗原或其部分的宿主细胞接触,以及

(iii) 将溶细胞性 T 细胞或细胞因子释放 T 细胞以适合溶解表达肿瘤相关抗原或其部分的细胞的量引入患者,所述肿瘤相关抗原具有由选自如下的核酸编码的序列:

(a) 包含选自 SEQ ID NO:1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分

或其衍生物，

- (b) 在严格条件下与 (a) 所述核酸杂交的核酸，
- (c) 与 (a) 或 (b) 的核酸简并的核酸，以及
- (d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。

62. 按照权利要求 61 的方法，其中所述宿主细胞重组性地表达结合肿瘤相关抗原或其部分的 HLA 分子。

63. 按照权利要求 62 的方法，其中所述宿主细胞内源性表达结合肿瘤相关抗原或其部分的 HLA 分子。

64. 一种治疗具有特征在于肿瘤相关抗原表达或异常表达的疾病的患者之方法，所述方法包括：

- (i) 鉴别由与所述疾病相关的细胞表达的核酸，所述核酸选自：
 - (a) 包含选自 SEQ ID NO :1-8、41-44、51-59、84、117 和 119 的核酸序列之核酸，其部分或其衍生物，
 - (b) 在严格条件下与 (a) 所述核酸杂交的核酸，
 - (c) 与 (a) 或 (b) 的核酸简并的核酸，以及
 - (d) 与 (a)、(b) 或 (c) 的核酸互补的核酸，
- (ii) 用所述核酸或其部分转染宿主细胞，
- (iii) 为表达所述核酸培养被转染的宿主细胞，以及
- (iv) 将宿主细胞或其提取物以适合增加对疾病相关的患者的细胞之免疫反应的量引入患者。

65. 按照权利要求 64 的方法，所述方法进一步包含用表达被鉴别的 MHC 分子并递呈肿瘤相关抗原或其部分的宿主细胞鉴别递呈肿瘤相关抗原或其部分的 MHC 分子，

66. 按照权利要求 64 或 65 的方法，其中所述免疫反应包括 B 细胞反应或 T 细胞反应。

67. 按照权利要求 66 的方法，其中所述免疫反应是 T 细胞反应，包括特异于递呈肿瘤相关抗原或其部分的宿主细胞或特异于表达肿瘤相关抗原或其部分的患者的细胞之溶细胞性 T 细胞或细胞因子释放 T 细胞的产生。

68. 按照权利要求 61-67 中任一项的方法，其中所述宿主细胞是非增殖性的。

69. 一种治疗特征在于肿瘤相关抗原表达或异常表达的疾病的方法，所述方法包括：

- (i) 鉴别来自表达异常量肿瘤相关抗原的患者的细胞，
- (ii) 分离所述细胞样品，
- (iii) 培养所述细胞，以及
- (iv) 将所述细胞以适合引发对细胞的免疫反应的量引入患者，所述肿瘤相关抗原具有由选自如下的核酸编码的序列：
 - (a) 包含选自 SEQ ID NO :1-8、41-44、51-59、84、117 和 119 的核酸序列之核酸，其部分或其衍生物，

- (b) 在严格条件下与 (a) 所述核酸杂交的核酸，
- (c) 与 (a) 或 (b) 的核酸简并的核酸，以及
- (d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。

70. 按照权利要求 33-69 中任一项的方法，其中所述疾病是癌症。

71. 一种抑制患者癌症发展的方法,所述方法包含给予按照权利要求 1-32 中任一项的有效量的药物组合物。

72. 按照权利要求 33-71 中任一项的方法,其中肿瘤相关抗原包含选自 SEQ ID NO: 9-19、45-48、60-66、85、90-97、100-102、105、106、111-116、118、120、123、124 和 135-137 的氨基酸序列,其部分或其衍生物。

73. 一种核酸,选自

(a) 包含选自 SEQ ID NO:3-5 的核酸序列之核酸,其部分或其衍生物,

(b) 在严格条件下与 (a) 所述核酸杂交的核酸,

(c) 与 (a) 或 (b) 的核酸简并的核酸,以及

(d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。

74. 一种核酸,其编码包含选自 SEQ ID NO:10 和 12-14 的氨基酸序列的蛋白质或多肽,其部分或其衍生物。

75. 一种重组 DNA 或 RNA 分子,其包含按照权利要求 73 或 74 的核酸。

76. 按照权利要求 75 的重组 DNA 分子,其是载体。

77. 按照权利要求 76 的重组 DNA 分子,其中所述载体是病毒载体或噬菌体。

78. 按照权利要求 75-77 中任一项的重组 DNA 分子,所述分子进一步包含控制核酸表达的表达控制序列。

79. 按照权利要求 78 的重组 DNA 分子,其中所述表达控制序列与所述核酸同源或异源。

80. 一种宿主细胞,所述宿主细胞包含按照权利要求 73 或 74 的核酸或者按照权利要求 75-79 中任一项的重组 DNA 分子。

81. 按照权利要求 80 的宿主细胞,其进一步包含编码 HLA 分子的核酸。

82. 一种蛋白质或多肽,其由按照权利要求 73 的核酸编码。

83. 一种蛋白质或多肽,其包含选自 SEQ ID NO:10 和 12-14 的氨基酸序列,其部分或其衍生物。

84. 按照权利要求 82 或 83 的蛋白质或多肽的免疫原性片段。

85. 按照权利要求 82 或 83 的蛋白质或多肽的片段,所述片段结合人类受体或人类抗体。

86. 一种药剂,其特异性地结合蛋白质或多肽或其部分,所述蛋白质或多肽由选自如下的核酸编码:

(a) 包含选自 SEQ ID NO:1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分或其衍生物,

(b) 在严格条件下与 (a) 所述核酸杂交的核酸,

(c) 与 (a) 或 (b) 的核酸简并的核酸,以及

(d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。

87. 按照权利要求 86 的药剂,其中所述蛋白质或多肽包含选自 SEQ ID NO:9-19、45-48、60-66、85、90-97、100-102、105、106、111-116、118、120、123、124 和 135-137 的氨基酸序列,其部分或其衍生物。

88. 按照权利要求 86 或 87 的药剂,其是抗体。

89. 按照权利要求 88 的药剂,其中所述抗体是单克隆、嵌合或人源化抗体或者抗体的

片段。

90. 一种抗体,其选择性结合 (i) 蛋白质或多肽或其部分和 (ii) 结合所述蛋白质或多肽或其部分的 MHC 分子之间的复合物,

所述抗体不与 (i) 或 (ii) 单独结合并且所述蛋白质或多肽由选自如下的核酸编码:

(a) 包含选自 SEQ ID NO :1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分或其衍生物,

(b) 在严格条件下与 (a) 所述核酸杂交的核酸,

(c) 与 (a) 或 (b) 的核酸简并的核酸,以及

(d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。

91. 按照权利要求 90 的抗体,其中所述蛋白质或多肽包含选自 SEQID NO :9-19、45-48、60-66、85、90-97、100-102、105、106、111-116、118、120、123、124 和 135-137 的氨基酸序列,其部分或其衍生物。

92. 按照权利要求 90 或 91 的抗体,所述抗体是单克隆、嵌合或人源化抗体或抗体的片段。

93. 按照权利要求 86-89 中任一项的药剂或按照权利要求 90-92 中任一项的抗体和治疗或诊断药剂之间的结合物。

94. 按照权利要求 93 的结合物,其中所述治疗或诊断药剂是毒素。

95. 一种用于检测肿瘤相关抗原表达或异常表达的试剂盒,所述试剂盒包含用于如下检测的试剂:

(i) 编码肿瘤相关抗原或其部分的核酸的检测,

(ii) 肿瘤相关抗原或其部分的检测,

(iii) 结合肿瘤相关抗原或其部分的抗体的检测,和 / 或

(iv) 特异于肿瘤相关抗原或其部分和 MHC 分子之间的复合物的 T 细胞的检测,

所述肿瘤相关抗原具有由选自如下的核酸编码的序列:

(a) 含有选自 SEQ ID NO :1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分或其衍生物,

(b) 在严格条件下与 (a) 所述核酸杂交的核酸,

(c) 与 (a) 或 (b) 的核酸简并的核酸,以及

(d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。

96. 权利要求 95 的试剂盒,其中所述用于检测编码肿瘤相关抗原或其部分的核酸之试剂是用于选择性扩增所述核酸的核酸分子。

97. 权利要求 96 的试剂盒,其中用于选择性扩增核酸的所述核酸分子包含编码肿瘤相关抗原的核酸之 6-50 个连续核苷酸的序列。

98. 一种重组 DNA 分子,包含源自选自 SEQ ID NO :1-8,41-44,51-59,84,117 和 119 的核酸序列的启动子区域。

肿瘤内差异表达的基因产物及其应用

[0001] 本申请是申请日为 2003 年 11 月 21 日、申请号为 200380103782.1、发明名称为“肿瘤内差异表达的基因产物及其应用”的发明专利申请的分案申请。

[0002] 尽管各学科间的方法和经典治疗程序的充分使用,癌症仍然是死亡的主要原因之一。较新治疗概念旨在通过使用重组肿瘤疫苗和其他特异性措施例如抗体治疗,将患者的免疫系统整合进整体治疗概念之中。这样一个策略成功的先决条件是肿瘤特异性或肿瘤相关抗原或抗原决定簇被患者的免疫系统识别,其免疫系统的效应物功能被干预性地增强。肿瘤细胞在生物学上与其同起源的非恶性细胞有生物特性的本质性不同。这些差异是因为肿瘤发展中获得的基因改变,导致尤其是癌症细胞内定性或定量改变的分子结构的形成。此类被肿瘤宿主特异性的免疫系统识别的肿瘤相关结构,被称为肿瘤相关抗原。肿瘤相关抗原的特异性识别涉及细胞和激素机制,它们是两个功能性相互连接的单元:CD4⁺ 和 CD8⁺T 淋巴细胞分别识别被递呈在 MHC(主要组织相容性复合物)II 型和 I 型分子上的加工过的抗原,而 B 淋巴细胞生产直接与未加工过的抗原结合的循环抗体分子。免疫系统对肿瘤细胞上抗原的识别导致细胞毒性效应物(effector)装置的启动以及在有 T 辅助细胞的情况下造成癌症细胞的消除,上述事实产生了肿瘤相关抗原潜在的临床治疗重要性(Pardoll, Nat. Med. 4:525-31, 1998)。相应地,肿瘤免疫学的中心目标是在分子水平定义这些结构。这些抗原的分子性质在很长时间都是迷。只有在发展合适的克隆技术之后,才有可能为了肿瘤相关抗原而通过分析细胞毒性 T 淋巴细胞(CTL)的靶结构(van der Bruggen 等人, Science 254:1643-7, 1991)或者通过将循环自身抗体用作探针(Sahin 等人, Curr. Opin. Immunol. 9:709-16, 1997)系统地筛选肿瘤的 cDNA 表达文库。为了这个目标, cDNA 表达文库从新鲜的肿瘤组织制备并且在合适的系统中重组表达为蛋白质。分离自患者的免疫效应物,就是有肿瘤特异性溶解模式的 CTL 克隆,或循环自身抗体被用来克隆各自的抗原。

[0003] 近年来,运用这些方法在多种肿瘤中定义了抗原的多重性。然而,上面举例说明的经典方法中用于抗原鉴别的探针是来自通常已经有晚期癌症的患者的免疫效应物(循环自身抗体或 CTL 克隆)。许多数据指出肿瘤会导致,例如,耐受和 T 细胞的无反应性,以及疾病的过程中,尤其是那些能引起有效免疫识别的特异性从免疫效应分子的全部特性(repertoire)中丢失。当前的患者研究并未得出任何可靠的证据来证实以前发现和使用的肿瘤相关抗原的真实作用。相应地,不能排除引起自发免疫反应的蛋白质是错误的靶结构。

[0004] 本发明的目的是提供用于癌症的诊断和治疗的靶结构。

[0005] 根据本发明,这个目的通过权利要求的主题内容可以达到。

[0006] 根据本发明,采用了用于鉴别和提供表达与肿瘤相关的抗原以及为此编码该抗原的核酸的策略。这个策略基于一个事实,即以器官特异性方式(例如仅在结肠、肺或肾组织中)表达的特定基因,在各个器官的肿瘤细胞中被激活,并且在其它组织的肿瘤细胞中以异位的和禁止的方式也被激活。首先,数据采集提出所有已知器官特异性基因尽可能完整的清单,然后依靠特异性 RT-PCR 通过表达分析来评价不同肿瘤内它们的异常激活。然而,在传统策略中,正常组织文库的转录产物组(transcriptoms)通常用电子手段从肿瘤组织库中减去,假定剩下的基因是肿瘤特异性的(Schmitt 等人, Nucleic Acids Res. 27:

4251-60,1999 ;Vasmatzis 等 人, Proc. Natl. Acad. Sci. USA. 95 :300-4,1998 ;Scheurle 等人, Cancer Res. 60 :4037-43,2000)。

[0007] 然而,被证明更加成功的本发明的概念是基于利用数据采集用电子手段提取所有器官特异性的基因,然后评价所述基因在肿瘤中的表达。

[0008] 因而本发明一方面涉及用于鉴别肿瘤内差异表达的组织特异性基因的策略。所述的策略将公用序列库的数据采集(“在芯片上”)(“in silico”)与其后的实验室试验研究的评价(“湿试验台”)(“wet bench”)联合起来。

[0009] 根据本发明,基于两个不同的生物信息程序的联合策略使得新的肿瘤基因能够被鉴别。这些基因以前被分类为纯粹器官特异性的。这些基因在肿瘤细胞内异常激活的发现允许它们被赋予有功能意义的实质上新的性质。根据本发明,这些肿瘤相关基因及其编码的基因产物被鉴别并给予独立的免疫原性作用。

[0010] 根据本发明鉴别的肿瘤相关抗原具有由选自如下的核酸编码的氨基酸序列,(a) 包含选自 SEQ ID NO :1-8,41-44,51-59,84,117 和 119 的核酸序列之核酸,其部分或其衍生物,(b) 在严格条件下与 (a) 的核酸杂交的核酸,(c) 与 (a) 或 (b) 的核酸简并的核酸,(d) 与 (a)、(b) 或 (c) 的核酸互补的核酸。在优选的实施方案中,根据本发明鉴别的肿瘤相关抗原具有由选自 SEQ ID NO :1-8,41-44,51-59,84,117 和 119 的核酸编码的氨基酸序列。在进一步优选的实施方案中,根据本发明鉴别的肿瘤相关抗原包含选自 SEQ ID NO :9-19、45-48,60-66,85,90-97,100-102,105,106,111-116,118,120,123,124 和 135-137 的氨基酸序列,其部分或其衍生物。

[0011] 本发明一般涉及根据本发明鉴别的肿瘤相关抗原或其部分或其衍生物用途,为其编码的核酸或抗所述编码核酸的核酸的用途,抗根据本发明鉴别的肿瘤相关抗原或部分或其衍生物的抗体的用途。这个应用既涉及这些抗原、功能片段、核苷酸、抗体等等的单个体也涉及两个或多个的联合,在一个实施方案中,也与其它肿瘤相关基因和抗原联合,作为诊断、治疗和进展的对照。

[0012] 治疗和 / 或诊断的优选疾病是那些选择性表达或异常表达根据本发明鉴别的一个或多个肿瘤相关抗原的疾病。

[0013] 本发明还涉及表达与肿瘤细胞相关的核酸和基因产物。

[0014] 此外,本发明涉及基因产物,即核酸和蛋白质或肽,其由已知基因剪接的改变(剪接变异体)或使用选择性开放阅读框的翻译的改变而产生。在这方面,本发明涉及核酸,所述核酸包含选自根据序列列表的 SEQ ID NO :3-5 的序列之核酸序列。而且,在这方面,本发明涉及蛋白质或肽,所述蛋白质或肽含有选自根据序列列表的 SEQ ID NO :10 和 12-14 的序列之氨基酸序列。本发明的剪接变异体根据本发明能用作肿瘤疾病诊断和治疗的靶点。

[0015] 尤其是,本发明涉及根据序列列表的 SEQ ID NO :10 的氨基酸序列,所述序列由根据本发明鉴别的一个选择性开放阅读框编码,并且在蛋白质的氨基端上有额外的 85 个氨基酸而与前述的蛋白质序列 (SEQ ID NO :9) 不同。

[0016] 非常不同的机制可导致剪接变异体的产生,例如:

[0017] - 可变的转录起始位点的利用

[0018] - 额外的外显子的利用

[0019] - 来自单个或两个或多个外显子的完全或不完全剪接

[0020] - 经由突变改变的剪接调节子序列（新的供体 / 受体序列的缺失或产生）

[0021] - 内含子序列的不完全去除

[0022] 基因改变的剪接导致改变的转录序列（剪接变异体）。剪接变异体在其改变的序列区域中的翻译导致发生改变的蛋白质，其与原先的蛋白质在结构和功能上截然不同。肿瘤相关剪接变异体可产生肿瘤相关转录产物和肿瘤相关蛋白质 / 抗原。这些可被利用作为肿瘤细胞检测和肿瘤治疗靶点的分子标志物 (marker)。例如根据本发明，血液、血浆、骨髓、唾液、支气管灌洗液、体内分泌物和组织活检中肿瘤细胞的检测，例如，在提取核酸之后用剪接变异体特异性的寡聚核苷酸通过 PCR 扩增来进行。尤其是，成对引物适合作为寡聚核苷酸，至少其中的一个在严格条件下与肿瘤相关剪接变异体的区域结合。根据本发明，为此目的在实施例中描述的寡聚核苷酸是合适的，尤其是寡聚核苷酸具有或包含选自序列列表的 SEQ ID NO :34-36、39、40 和 107-110 的序列。根据本发明，所有序列依赖性的检测系统适合用于检测。除了 PCR，这些检测系统是，例如，基因芯片 / 微阵列系统、Northern blot、RNAse 保护试验 (RDA) 及其它。所有检测系统共同具有以特异性杂交于至少一个剪接变异体特异性的核酸序列为基础的检测。然而，肿瘤细胞也可，根据本发明，通过识别剪接变异体编码的特异性抗原决定簇的抗体进行检测。在这方面，本发明涉及尤其是具有或包含选自序列列表的 SEQ ID NO :17-19、111-115、120 和 137 之序列的肽以及抗这些肽的特异性抗体。适合免疫的尤其是其抗原决定簇与基因产物的变异体截然不同的氨基酸，所述氨基酸优选地在健康细胞中产生。用抗体检测肿瘤细胞此处可以在分离自患者的样品上或者静脉给予抗体后成像而进行。除诊断的用途以外，具有新的或改变的抗原决定簇的剪接变异体是免疫治疗很有吸引力的靶点。本发明的抗原决定簇可用于靶向治疗活性的单克隆抗体或 T 淋巴细胞。在被动免疫治疗中，识别剪接变异体特异性的抗原决定簇的抗体或 T 淋巴细胞在这里被过继性转移。在其他抗原的情况下，抗体也可以通过使用标准技术（动物的免疫，用于分离重组抗体的淘选策略 (panning strategies)）利用包括这些抗原决定簇的多肽来生产。或者，可能为了免疫使用编码含有所述抗原决定簇的寡聚肽或多肽的核酸。用于体外或体内产生抗原决定簇特异性的 T 淋巴细胞的各种技术是已知的，并且已有详细描述（例如 Kessler JH, 等人, 2001, Sahin 等人, 1997），并且同样地以利用含有剪接变异体特异性的抗原决定簇或编码所述寡聚肽或多肽的核酸的寡聚肽或多肽为基础。含有剪接变异体特异性的抗原决定簇或编码所述多肽的核酸的寡聚肽或多肽也可用作主动免疫治疗（疫苗接种和疫苗治疗）中的药物活性物质。

[0023] 本发明还描述了性质和二级修饰的程度在正常和肿瘤组织中不同的蛋白质（例如 Durand&Seta, 2000 ; Clin. Chem. 46 :795-805 ; Hakomori, 1996 ; Cancer Res. 56 : 5309-18）。

[0024] 蛋白质修饰的分析可在 Western 印迹中进行。尤其是，通常具有几千道尔顿分子量的糖基化导致更大总质量的靶蛋白质，所述靶蛋白质能在 SDS-PAGE 中被分离出来。为了检测特异性 O- 和 N- 糖苷键，在使用 SDS 变性之前蛋白质溶解产物与 O- 和 N- 糖基化酶孵育（根据各自生产厂商的说明书，例如，PNGase、内切糖苷酶 F、内切糖苷酶 H, Roche 诊断）。之后，进行 Western 印迹。如果靶蛋白质的分子量变小了，特异性的糖基化在与糖苷酶孵育后可照这样进行检测，并且因而，对修饰的肿瘤特异性也能进行分析。肿瘤细胞和健康细胞中差异糖基化的蛋白质区域尤其引人注目。然而，至今这样的糖基化的差异仅在一些细胞

表面蛋白质有描述（例如，Muc1）。

[0025] 根据本发明，有可能检测肿瘤中 Claudin-18 的差异糖基化。胃肠道癌症、胰腺癌、食道肿瘤、前列腺肿瘤以及肺肿瘤具有较低水平的糖基化的 Claudin-18。健康组织中的糖基化掩盖了 Claudin-18 的蛋白质抗原决定簇，而在肿瘤细胞上因为缺乏糖基化蛋白质，所述抗原决定簇未被掩盖。对应地，有可能根据本发明选择结合于这些结构域的配体和抗体。这样的根据本发明的配体和抗体不与健康细胞上的 Claudin-18 结合，因为这里的抗原决定簇由于糖基化被掩盖了。

[0026] 按照上面对源自肿瘤相关剪接变异体的蛋白质抗原决定簇的描述，因而有可能出于诊断和治疗的目的，使用差异糖基化来区别正常细胞和肿瘤细胞。

[0027] 一方面，本发明涉及一种药物组合物，所述药物组合物包含识别根据本发明鉴别的肿瘤相关抗原并且对具有根据本发明鉴别的肿瘤相关抗原表达或异常表达的细胞有优先选择性的药剂。在特定的实施方案中，所述药剂可造成细胞死亡的诱导、细胞生长的减少、对细胞膜的损害或者细胞因子的分泌并且优选地具有肿瘤抑制活性。在一个实施方案中，所述药剂是与编码肿瘤相关抗原的核酸选择性地杂交的反义核酸。在进一步的实施方案中，所述药剂是选择性结合肿瘤相关抗原的抗体，尤其是选择性结合肿瘤相关抗原的补体激活或毒素偶联抗体。在进一步的实施方案中，所述药剂包括两个或多个药剂，每一个所述药剂选择性识别不同肿瘤相关抗原，其中至少一个是根据本发明鉴别的肿瘤相关抗原。识别并不需要直接伴有抗原活性和表达的抑制。在本发明的这个方面，选择性地局限于肿瘤的抗原优选地作为标记以招募效应物装置到该特异性位点。在优选的实施方案中，所述药剂是识别 HLA 分子上的抗原并且溶解这样标记的细胞之细胞毒性 T 淋巴细胞。在进一步的实施方案中，药剂是选择性结合肿瘤相关抗原并且因而招募天然或人工效应物装置到所述细胞的抗体。在进一步的实施方案中，所述药剂是增强特异性识别所述抗原的其它细胞的效应物功能之 T 辅助淋巴细胞。

[0028] 一方面，本发明涉及一种药物组合物，所述药物组合物包含抑制根据本发明鉴别的肿瘤相关抗原的表达或活性的药剂。在优选的实施方案中，所述药剂是与编码肿瘤相关抗原的核酸选择性杂交的反义核酸。在进一步的实施方案中，所述药剂是选择性结合肿瘤相关抗原的抗体。在进一步的实施方案中，所述试剂包含两个或多个药剂，每一个所述药剂选择性抑制不同肿瘤相关抗原的表达或活性，其中至少一个是根据本发明鉴别的肿瘤相关抗原。

[0029] 本发明此外还涉及一种药物组合物，所述药物组合物包含在给予时可选择性增加 HLA 分子和根据本发明鉴别的肿瘤相关抗原的肽表位之间的复合物的量的药剂。在一个实施方案中，所述药剂包含选自如下的一个或多个组分，(1) 肿瘤相关抗原或其部分，(2) 编码所述肿瘤相关抗原或其部分的核酸，(3) 表达所述肿瘤相关抗原或其部分的宿主细胞，以及 (4) 所述肿瘤相关抗原的肽表位和 MHC 分子之间的分离的复合物。在一个实施方案中，药剂包含两个或多个药剂，每一个所述药剂选择性增加 MHC 分子和不同肿瘤相关抗原的肽表位之间的复合物的量，其中至少一个是根据本发明鉴别的肿瘤相关抗原。

[0030] 本发明还涉及包含一个或多个选自如下组分的药物组合物，(1) 根据本发明鉴别的肿瘤相关抗原或其部分，(2) 编码根据本发明鉴别的肿瘤相关抗原或其部分的核酸，(3) 结合根据本发明鉴别的肿瘤相关抗原或其部分的抗体，(4) 与编码根据本发明鉴别的肿瘤

相关抗原的核酸特异性杂交的反义核酸, (5) 表达根据本发明鉴别的肿瘤相关抗原或其部分的宿主细胞, 以及 (6) 根据本发明鉴别的肿瘤相关抗原或其部分和 HLA 分子之间的分离的复合物。

[0031] 编码根据本发明鉴别的肿瘤相关抗原或其部分的核酸可存在于药用组合物中, 所述核酸在表达载体中并功能性连接于启动子。

[0032] 存在于本发明的药物组合物中的宿主细胞可以分泌肿瘤相关抗原或其部分, 将其表达于表面或额外地表达结合所述肿瘤相关抗原或所述其部分的 HLA 分子。在一个实施方案中, 所述宿主细胞内源性表达 HLA 分子。在进一步的实施方案中, 所述宿主细胞以重组方式表达 HLA 分子和 / 或肿瘤相关抗原或其部分。所述宿主细胞优选是非增殖性的。在优选的实施方案中, 所述宿主细胞是抗原递呈细胞, 尤其是树突细胞、单核细胞或巨噬细胞。

[0033] 存在于本发明的药物组合物中的抗体可以是单克隆抗体。在进一步的实施方案中, 所述抗体是嵌合或人源化抗体, 天然抗体或合成抗体的片段, 所有的所述抗体可以通过组合技术生产。抗体可以偶联于在治疗或诊断上有用的药剂。

[0034] 存在于本发明的药物组合物中的反义核酸可以包含编码根据本发明鉴别的肿瘤相关抗原之核酸的 6-50 个, 尤其是 10-30 个、15-30 个和 20-30 个连续核苷酸的序列。

[0035] 在进一步的实施方案中, 由本发明的药物组合物直接或通过核酸表达提供的肿瘤相关抗原或其部分与细胞表面的 MHC 分子结合, 所述的结合优选地造成细胞溶解效应和 / 或诱导细胞因子释放。

[0036] 本发明的药物组合物可以包含药物相容性载体和 / 或佐剂。所述佐剂可以选自皂甙、粒细胞巨噬细胞 - 克隆刺激因子、CpG 核苷酸、RNA、细胞因子或化学因子。本发明的药物组合物优选的用于治疗特征在于肿瘤相关抗原选择性表达或异常表达的疾病。在优选的实施方案中, 所述疾病是癌症。

[0037] 本发明此外还涉及治疗或诊断特征在于一个或多个肿瘤相关抗原表达或异常表达的疾病的方法。在一个实施方案中, 治疗包括给予本发明的药物组合物。

[0038] 一方面, 本发明涉及诊断特征在于根据本发明鉴别的肿瘤相关抗原表达或异常表达的疾病的方法。方法包括 (1) 编码肿瘤相关抗原或其部分的核酸的检测和 / 或 (2) 肿瘤相关抗原或其部分的检测, 和 / 或 (3) 抗肿瘤相关抗原或其部分的抗体的检测和 / 或 (4) 特异于分离自患者的生物样品中肿瘤相关抗原或其部分的细胞毒性或 T 辅助淋巴细胞的检测。在特定的实施方案中, 所述检测包括 (1) 生物样品与试剂接触, 所述试剂特异性结合编码肿瘤相关抗原的核酸或其部分、所述肿瘤相关抗原或所述其部分、对肿瘤相关抗原或其部分有特异性的抗体或细胞毒性或 T 辅助淋巴细胞, 以及 (2) 检测所述试剂和核酸或其部分、和肿瘤相关抗原或其部分、和抗体或细胞毒性或 T 辅助淋巴细胞之间的复合物的形成。在一个实施方案中, 所述疾病以两个或多个不同肿瘤相关抗原的表达或异常表达为特征, 并且所述检测包括编码所述两个或多个不同肿瘤相关抗原或其部分的两个或多个核酸的检测, 两个或多个不同肿瘤相关抗原或其部分的检测, 两个或多个结合所述两个或多个不同肿瘤相关抗原或其部分的抗体的检测, 特异于所述两个或多个不同肿瘤相关抗原的两个或多个细胞毒性或 T 辅助淋巴细胞的检测。在进一步的实施方案中, 分离自患者的生物样品与相当的正常生物样品进行比较。

[0039] 在另外一方面, 本发明涉及测定特征在于根据本发明鉴别的肿瘤相关抗原表达或

异常表达的疾病的缓解、进程或开始的方法,所述方法包括监测来自具有所述疾病或怀疑患有所述疾病之患者的样品的参数,一个或多个所述参数选自(1)编码肿瘤相关抗原或其部分的核酸的量,(2)肿瘤相关抗原或其部分的量,(3)结合肿瘤相关抗原或其部分的抗体的量,以及(4)特异于肿瘤相关抗原或其部分和 MHC 分子之间的复合物的溶细胞性 T 细胞或 T 辅助细胞的量。所述方法优选地包括测定第一个样品在第一个时间点以及另外一个样品在第二个时间点的所述参数,以及通过比较两个样品来测定疾病的进程。在特定的实施方案中,所述疾病以两个或多个不同肿瘤相关抗原的表达或异常表达为特征,并且所述监测包含监测(1)编码所述两个或多个不同肿瘤相关抗原或其部分的核酸的量,和/或(2)所述两个或多个不同肿瘤相关抗原或其部分的量,和/或(3)两个或多个结合所述两个或多个不同肿瘤相关抗原或其部分的抗体的量,和/或(4)两个或多个特异于所述两个或多个不同肿瘤相关抗原或其部分和 MHC 分子之间的复合物的溶细胞性 T 细胞或 T 辅助细胞的量。

[0040] 根据本发明,核酸或其部分的检测或监测核酸或其部分的量可以使用特异性杂交所述核酸或所述其部分的多聚核苷酸探针来进行,或者可以通过所述核酸或所述其部分的选择性扩增来进行。在一个实施方案中,多聚核苷酸探针包含所述核酸的 6-50 个,尤其是 10-30 个、15-30 个和 20-30 个连续核苷酸的序列。

[0041] 在特定的实施方案中,待检测的肿瘤相关抗原或其部分存在于细胞内或细胞表面上。根据本发明,肿瘤相关抗原或其部分的检测或监测肿瘤相关抗原或其部分的量可以通过使用特异性结合所述肿瘤相关抗原或其部分的抗体来进行。

[0042] 在进一步的实施方案中,待检测的肿瘤相关抗原或其部分存在于与 MHC 分子,尤其是 HLA 分子的复合物中。

[0043] 根据本发明,抗体的检测或监测抗体的量可以通过使用特异性结合所述抗体的蛋白质或肽来进行。

[0044] 根据本发明,溶细胞性 T 细胞或 T 辅助细胞的检测或者监测特异于抗原或其部分和 MHC 分子之间的复合物之溶细胞性 T 细胞或 T 辅助细胞可以使用递呈所述抗原或其部分和 MHC 分子之间的复合物之细胞来进行。

[0045] 用于检测或监测的多聚核苷酸探针、抗体、蛋白质或肽或细胞优选地以可检测的方式被标记。在特定的实施方案中,可检测的标志物是放射活性标记或酶标记。T 淋巴细胞可额外地通过检测由 MHC 和肿瘤相关抗原或其部分之间的复合物的特异性刺激诱发它们的增殖、它们的细胞因子生成和他们的细胞毒活性来进行检测。T 淋巴细胞也可通过重组 MHC 分子或者两个或多个 MHC 分子的复合物来进行检测,所述 MHC 分子荷载了一个或多个肿瘤相关抗原的特定的免疫原性片段,并且能通过与特异性 T 细胞受体接触来鉴别特异性 T 淋巴细胞。

[0046] 在另外一个方面,本发明涉及治疗、诊断或监测特征在于根据本发明鉴别的肿瘤相关抗原表达或异常表达的疾病的方法,所述方法包括给予结合于所述肿瘤相关抗原或其部分并且偶联于治疗或诊断药剂的抗体。所述抗体可以是单克隆抗体。在进一步的实施方案中,所述抗体是嵌合或人源化抗体或天然抗体的片段。

[0047] 本发明还涉及治疗具有特征在于根据本发明鉴别的肿瘤相关抗原表达或异常表达的疾病的患者之方法,所述方法包括(1)取出含有来自所述患者的免疫反应活性细胞的

样品, (2) 在有利于抗所述肿瘤相关抗原或其部分的溶细胞性 T 细胞产生的条件下, 将所述样品与表达所述肿瘤相关抗原或其部分的宿主细胞接触, 以及 (3) 将溶细胞性 T 细胞以适合溶解表达肿瘤相关抗原或其部分的细胞的量引入患者。本发明同样地涉及克隆抗肿瘤相关抗原的溶细胞性 T 细胞之 T 细胞受体。所述受体可以被转移进其它的 T 细胞, 所述 T 细胞因而获得期望的特异性, 并且按照方法 (3) 还可以被引入患者。

[0048] 在一个实施方案中, 所述宿主细胞内源性表达 HLA 分子。在另外一个实施方案中, 所述宿主细胞重组表达 HLA 分子和 / 或肿瘤相关抗原或其部分。所述宿主细胞优选是非增殖性的。在优选的实施方案中, 所述宿主细胞是抗原递呈细胞, 尤其是树突细胞、单核细胞或巨噬细胞。

[0049] 在另外一个方面, 本发明涉及治疗具有特征在于肿瘤相关抗原表达或异常表达的疾病的患者之方法, 所述方法包含 (1) 鉴别编码根据本发明鉴别的肿瘤相关抗原并由所述疾病相关的细胞表达的核酸, (2) 用所述核酸或其部分转染宿主细胞, (3) 为所述核酸的表达培养被转染的宿主细胞 (当得到高效率的转染时, 此方法就不是必须的), 以及 (4) 将宿主细胞或其提取物以适合增加对患者的疾病相关细胞之免疫反应的量引入患者。方法可还包括用表达鉴别的 MHC 分子以及递呈所述肿瘤相关抗原或其部分的宿主细胞, 鉴别递呈肿瘤相关抗原或其部分的 MHC 分子。免疫反应可以包含 B 细胞反应或 T 细胞反应。此外, T 细胞反应可以包含溶细胞性 T 细胞和 / 或 T 辅助细胞的产生, 所述溶细胞性 T 细胞和 / 或 T 辅助细胞特异于递呈肿瘤相关抗原或其部分的宿主细胞, 或特异于表达所述肿瘤相关抗原或其部分的患者的细胞。

[0050] 本发明也涉及治疗特征在于根据本发明鉴别的肿瘤相关抗原表达或异常表达的疾病的方法, 所述方法包括 (1) 鉴别来自表达异常量的肿瘤相关抗原的患者的细胞, (2) 分离所述细胞的样品, (3) 培养所述细胞, 以及 (4) 将所述细胞以适合诱发对所述细胞的免疫反应的量引入患者。

[0051] 优选地, 根据本发明使用的宿主细胞为非增殖性的, 或者使其为非增殖性的。特征在于肿瘤相关抗原表达或异常表达的疾病尤其是癌症。

[0052] 本发明还涉及选自如下的核酸: (a) 包含选自 SEQ ID NO:3-5 核酸序列的核酸, 其部分或其衍生物, (b) 与 (a) 的核酸在严格条件下杂交的核酸, (c) 与 (a) 或 (b) 的核酸简并的核酸, 以及 (d) 与 (a)、(b) 或 (c) 互补的核酸。本发明此外还涉及编码包含选自 SEQ ID NO:10 和 12-14 的氨基酸序列的蛋白质或多肽之核酸, 其部分或其衍生物。

[0053] 在另外一个方面, 本发明涉及本发明的核酸的启动子序列。这些序列可以功能性地与另一个基因相连接, 优选地在表达载体中, 并且因而保证所述基因在合适细胞中的选择性表达。

[0054] 在另外一个方面, 本发明涉及重组核酸分子, 尤其是 DNA 或 RNA 分子, 其包含本发明的核酸。

[0055] 本发明还涉及含有本发明的核酸或包含本发明的核酸的重组核酸分子之宿主细胞。

[0056] 所述宿主细胞还可以包含编码 HLA 分子的核酸。在一个实施方案中, 所述宿主细胞内源地表达 HLA 分子。在进一步的实施方案中, 所述宿主细胞重组表达 HLA 分子和 / 或本发明的核酸或其部分。优选地, 所述宿主细胞是非增殖性的。在优选的实施方案中, 所

述宿主细胞是抗原递呈细胞,尤其是树突细胞、单核细胞或巨噬细胞。

[0057] 在进一步的实施方案中,本发明涉及与根据本发明鉴别的核酸杂交并可以用作基因探针或“反义”分子的寡聚核苷酸。与根据本发明鉴别的核酸或其部分杂交并呈寡聚核苷酸引物或有效的样品 (competent samples) 形式的核酸分子,可以用于寻找与根据本发明鉴别的所述核酸同源的核酸。PCR 扩增、Southern 和 Northern 杂交可以被用来寻找同源核酸。杂交可以在低严格条件下进行,更加优选地在中等严格以及最优选地在高度严格条件下进行。术语“严格条件”根据本发明指的是允许多聚核苷酸之间特异性杂交的条件。

[0058] 在另外一个方面,本发明涉及由选自如下的核酸编码的蛋白质、多肽或肽:(a) 包含选自 SEQ ID NO:3-5 的核酸序列之核酸,其部分或其衍生物,(b) 与 (a) 的核酸在严格条件下杂交的核酸,(c) 与 (a) 或 (b) 的核酸简并的核酸,以及 (d) 与 (a)、(b) 或 (c) 互补的核酸。在优选的实施方案中,本发明涉及包含选自 SEQ ID NO:10 和 12-14 的氨基酸序列的蛋白质或多肽或肽,或其部分或其衍生物。

[0059] 在另外一个方面,本发明涉及根据本发明鉴别的肿瘤相关抗原的免疫原性片段。所述的片段优选地与人类 HLA 受体或人类抗体结合。本发明的片段优选地包含至少 6 个,尤其是至少 8 个,至少 10 个,至少 12 个,至少 15 个,至少 20 个,至少 30 个或至少 50 个氨基酸的序列。

[0060] 在这方面,本发明尤其是涉及具有或包含选自 SEQ ID NO:17-19、90-97、100-102、105、106、111-116、120、123、124 和 135-137 的肽,其部分或其衍生物。

[0061] 在另外一个方面,本发明涉及结合根据本发明鉴别的肿瘤相关抗原或其部分的试剂。在优选的实施方案中,所述试剂是抗体。在进一步的实施方案中,所述抗体是嵌合或人源化抗体或重组技术生产的抗体或是抗体的片段。此外,本发明涉及一种抗体,所述抗体选择性结合 (1) 根据本发明鉴别的肿瘤相关抗原或其部分与 (2) 根据本发明鉴别的肿瘤相关抗原或所述其部分结合的 MHC 分子之间的复合物,所述抗体不单独与 (1) 或 (2) 结合。本发明的抗体可以是单克隆抗体。在进一步的实施例中,所述抗体是嵌合或人源化的抗体或者天然抗体的片段。

[0062] 尤其是,本发明涉及尤其是抗体这样的试剂,其特异性结合具有或包含选自 SEQ ID NO:17-19、90-97、100-102、105、106、111-116、120、123、124 和 135-137 的序列的肽,其部分或其衍生物。

[0063] 本发明此外还涉及结合根据本发明鉴别的肿瘤相关抗原或其部分的本发明的试剂或本发明的抗体与治疗或诊断试剂之间形成的结合物 (conjugate)。在一个实施方案中,所述治疗或诊断试剂是毒素。

[0064] 在另外一个方面,本发明涉及用于检测根据本发明鉴别的肿瘤相关抗原的表达或异常表达的试剂盒,所述试剂盒包含用于检测 (1) 编码肿瘤相关抗原或其部分的核酸,(2) 肿瘤相关抗原或其部分,(3) 结合肿瘤相关抗原或其部分的抗体,和 / 或 (4) 特异于肿瘤相关抗原或其部分和 MHC 分子之间的复合物的 T 细胞之试剂。在一个实施方案中,用于检测核酸或其部分的试剂是用于选择性扩增所述核酸的核酸分子,其包含尤其是所述核酸的 10-30 个,15-30 个和 20-30 个连续核苷酸的序列。

[0065] 发明详述

[0066] 根据本发明,基因被描述为在肿瘤细胞中选择性表达或异常表达的并且是肿瘤相

关的抗原。

[0067] 根据本发明,这些基因和 / 或它们的基因产物和 / 或它们的衍生物和 / 或部分优选的治疗手段的靶结构。从概念上说,所述治疗手段旨在抑制选择性表达的肿瘤相关基因产物的活性。如果所述异常的各自选择性的表达对于肿瘤的致病性有功能上的重要性,并且如果其连接同时伴有相应细胞的选择性损伤,这种治疗手段就很有用。其它的治疗概念企图将肿瘤相关抗原作为标记,以招募具有选择性损伤肿瘤细胞的潜能的效应物装置。这里,靶分子自己的功能及其在肿瘤发展中的作用完全无关。

[0068] 核酸的“衍生物”根据本发明意思是存在于所述核酸中的一个或多个核苷酸的替代、缺失和 / 或增加。此外,术语“衍生物”也包含在核苷酸碱基上、在糖基上或在磷酸基上核酸的化学衍生化反应。术语“衍生物”也包含含有核苷酸和天然并不存在的核苷酸类似物的核酸。

[0069] 根据本发明,核酸优选地是脱氧核糖核酸 (DNA) 或核糖核酸 (RNA)。所述核酸根据本发明包含基因组 DNA、cDNA、mRNA、重组产生的和化学合成的分子。根据本发明,所述核酸可以作为单链或双链以及线性或共价环形封闭分子而存在。

[0070] 根据本发明描述的核酸优选地已经被分离。术语“分离的核酸”根据本发明意思是指核酸是 (1) 例如通过多聚酶链式反应 (PCR) 而体外扩增, (2) 通过克隆而重组产生, (3) 例如通过切割和凝胶电泳而分离纯化,或 (4) 例如通过化学合成法而合成。分离的核酸是可通过重组 DNA 技术进行操作的核酸。

[0071] 如果两个核酸序列能够杂交并且互相形成稳定的双链体,那么一个核酸互补于另一个核酸,所述杂交优选地在允许多聚核苷酸之间特异性杂交的条件下 (严格条件) 进行。严格条件在例如 Molecular Cloning :A Laboratory Manual, J. Sambrook 等人,编,第 2 版, Cold Spring Harbor Laboratory press, Cold Spring Harbor, New York, 1989 或者 Current Protocols in Molecular Biology, F. M. Ausubel 等人,编, John Wiley & Sons, Inc., New York 中有描述,并且指的是,例如,在 65℃ 杂交缓冲液 (3.5xSSC, 0.02% Ficoll, 0.02% 聚乙烯吡咯烷酮, 0.02% 牛血清白蛋白, 2.5mM NaH₂PO₄ (pH 7), 0.5% SDS, 2mM EDTA) 中的杂交。SSC 是 0.15M 氯化钠 / 0.15M 柠檬酸钠, pH 7。杂交后,已转入 DNA 的膜,例如在室温用 2×SSC 然后在高达 68℃ 的温度下用 0.1–0.5×SSC/0.1×SDS 进行清洗。

[0072] 根据本发明,互补的核酸具有至少 40%, 尤其是至少 50%、至少 60%、至少 70%、至少 80%、至少 90% 以及优选地至少 95%、至少 98% 或至少 99% 相同的核苷酸。

[0073] 编码肿瘤相关抗原的核酸根据本发明可以单独存在或与其它核酸,尤其是异源性核酸联合存在。在优选的实施方案中,所述核酸功能性地连接于与所述核酸同源或异源的表达控制序列或调节序列。如果编码序列和调节序列以所述编码序列的表达或转录在所述调节序列的控制之下或影响之下的方式互相共价连接,那么它们就“功能性”地互相连接。如果编码序列被翻译成功能蛋白质,那么由于功能性连接于所述编码序列的调节序列,所述调节序列的诱导导致所述编码序列的转录,而不会造成编码序列的移码或者所述编码序列不能被翻译成期望的蛋白质或肽。

[0074] 术语“表达控制序列”或者“调节序列”根据本发明包含启动子、增强子和其他调节基因表达的控制元素。在本发明特定的实施方案中,表达控制序列能被调节。调节序列的精确结构因物种或细胞类型可不同,但一般地包含分别参与转录和翻译的起始的 5' 端非

转录和 5' 端非翻译序列,例如 TATA 盒、帽序列、CAAT 序列等等。更加特异性地,5' 端非转录调节序列包含启动子区域,所述启动子区域包括为了控制功能性连接的基因转录之启动子序列。

[0075] 因而,一方面,此处阐述的肿瘤相关抗原可以与任何表达控制序列和启动子相组合。另一方面,然而,此处阐述的肿瘤相关基因产物的启动子,根据本发明,可以与任何其它的基因相组合。这允许这些启动子的选择性活性被利用。

[0076] 根据本发明,核酸此外还可以与编码多肽的另一个核酸联合存在,所述多肽控制蛋白质或来自宿主细胞的所述核酸编码多肽的分泌。根据本发明,核酸也可与编码多肽的另一个核酸联合存在,所述多肽造成编码的蛋白质或多肽被锚定在宿主细胞膜上或被分隔在所述细胞特定的细胞器内。相似地,与如下核酸联合是可能的,其代表了报告基因或任何“标签”。

[0077] 在优选的实施方案中,重组 DNA 分子根据本发明是载体,在合适的情况下有启动子,所述载体控制核酸的表达,例如编码本发明的肿瘤相关抗原的核酸。这里使用的术语“载体”是指最一般意义的载体,并且包含任何的核酸的中间媒介物,使得所述核酸,例如可被引入原核和 / 或真核细胞,并且在合适的情况下被整合进基因组中。此类载体优选地可被复制和 / 或表达在细胞内。中间媒介物适合,例如用于电穿孔,用于微粒轰击,用于脂质体给予,用于借助于农杆菌的转移或用于通过 DNA 或 RNA 病毒的插入。载体包含质粒、噬菌粒或病毒基因组。

[0078] 编码根据本发明鉴别的肿瘤相关抗原的核酸可以用于宿主细胞的转染。核酸这里意思是重组 DNA 和 RNA。重组 RNA 可以通过 DNA 模板的体外转录进行制备。此外,它在应用前可以通过稳定化序列、加帽和多乙酰化进行修饰。

[0079] 根据本发明,术语“宿主细胞”涉及任何能被外源性核酸转化或转染的细胞。术语“宿主细胞”根据本发明包含原核(例如大肠杆菌)或真核细胞(例如树突细胞、B 细胞、CHO 细胞、COS 细胞、K562 细胞、酵母细胞和昆虫细胞)。特别优选哺乳动物细胞,例如来自人类、小鼠、仓鼠、猪、山羊、猿类的细胞。细胞可以来源于多种多样的组织细胞,并且包含原代细胞和细胞系。特异性样品包含角质细胞、外周血白细胞、骨髓的干细胞和胚胎干细胞。在进一步的实施方案中,宿主细胞是抗原递呈细胞,尤其是树突细胞、单核细胞和巨噬细胞。核酸可以以单拷贝或两个或两个以上拷贝的形式存在于宿主细胞中,并且在一个实施方案中,所述核酸表达在宿主细胞中。

[0080] 根据本发明,术语“表达”是指最一般意义的表达,并且包含 RNA 或者 RNA 和蛋白质的生产。它还包含核酸的部分表达。此外,可以进行瞬间或稳定表达。哺乳动物细胞中优选的表达系统包含 pcDNA3.1 和 pRc/CMV (Invitrogen, Carlsbad, CA),所述表达系统含有选择性标志物例如授予 G418 抗性的基因(并且因而使得稳定转染的细胞系被选择)和巨细胞病毒(CMV)的增强子-启动子序列。

[0081] 在本发明的 HLA 分子递呈肿瘤相关抗原或其部分的那些情况下,表达载体还可以包含编码所述 HLA 分子的核酸序列。编码 HLA 分子的核酸序列和编码肿瘤相关抗原或其部分的核酸可以存在于同一个表达载体上,或两个核酸可以存在于不同的表达载体上。在后一种情况下,这两个表达载体可以被共转染进细胞内。如果宿主细胞既不表达肿瘤相关抗原或其部分也不表达 HLA 分子,编码它们的两个核酸或者在同一个表达载体或者在不同的

表达载体上而被转染入细胞。如果细胞已经表达 HLA 分子,那么只有编码肿瘤相关抗原或其部分的核酸序列被转染进细胞。

[0082] 本发明还包含用于扩增编码肿瘤相关抗原的核酸之试剂盒。这样的试剂盒包含,例如杂交编码肿瘤相关抗原的核酸的一对扩增引物。为了避免引物二聚体的形成,引物优选地包含核酸的 6-50 个,尤其是 10-30 个,15-30 个以及 20-30 个连续核苷酸之序列,并且是不相互重叠的。一个引物将与编码肿瘤相关抗原的核酸中的一条链杂交,另一个引物则与互补链杂交,这样的设计允许编码肿瘤相关抗原的核酸的扩增。

[0083] “反义”分子或“反义”核酸可以用于调节,尤其是降低,核酸的表达。术语“反义分子”或“反义核酸”根据本发明指的是寡聚核苷酸,所述寡聚核苷酸是寡聚核糖核苷酸、寡聚脱氧核糖核苷酸、被修饰的寡聚核糖核苷酸或被修饰的寡聚脱氧核糖核苷酸,并且所述寡聚核苷酸在生理条件下杂交于包含特定基因的 DNA 或者杂交于所述基因的 mRNA,因此抑制所述基因的转录和 / 或所述 mRNA 的翻译。根据本发明,“反义分子”还包含构建体,所述构建体含有与其天然启动子方向相反的核酸或其部分。核酸或其部分的反义转录产物可以和天然发生的指定酶的 mRNA 形成二联体,并且因而阻止 mRNA 的蓄积或 mRNA 翻译成活性酶。另一个可能性是核酶用于失活核酸。根据本发明优选的反义寡聚核苷酸具有靶核酸的 6-50 个,尤其是 10-30 个,15-30 个以及 20-30 个连续的核苷酸之序列,并且优选地是与靶核酸或其部分完全互补的。

[0084] 在优选的实施方案中,反义寡聚核苷酸杂交于氨基端位点或 5' 端上游位点例如翻译起始位点、转录起始位点或启动子位点。在进一步的实施方案中,反义寡聚核苷酸杂交于 3' 端非翻译区域或 mRNA 剪接位点。

[0085] 在一个实施方案中,本发明的寡聚核苷酸由核糖核苷酸、脱氧核糖核苷酸或它们的组合组成,一个核苷酸的 5' 末端与另一个核苷酸的 3' 末端通过磷酸二酯键互相连接。这些寡聚核苷酸可以用传统方法进行合成或重组生产。

[0086] 在优选的实施方案中,本发明的寡聚核苷酸是“被修饰的”寡聚核苷酸。这里,为了增加寡聚核苷酸的例如稳定性或治疗效果,寡聚核苷酸可以用非常不同的方法来修饰,而并不削弱结合于其靶点的能力。根据本发明,术语“被修饰的寡聚核苷酸”意思是一种寡聚核苷酸,所述寡聚核苷酸中 (1) 至少它的两个核苷酸通过合成的核苷酸之间的键互相连接 (例如不是磷酸二酯键的核苷酸之间的键) 和 / 或 (2) 通常在核酸中找不到的化学基团共价连接于寡聚核苷酸。优选的合成的核苷酸之间的键是硫代磷酸酯、烷基磷酸酯、二硫代磷酸酯、磷酸酯基、烷基硫代磷酸酯、氨基磷酸酯、氨基甲酸酯、碳酸酯、磷酸三酯、氨基乙酸酯、羧甲基酯和肽。

[0087] 术语“被修饰的寡聚核苷酸”还包含具有共价修饰的碱基和 / 或糖的寡聚核苷酸。“被修饰的寡聚核苷酸”包含,例如带有共价结合于 3' 位置低分子量的有机基团 (不是羟基) 和 5' 位置磷酸酯基团的糖残基的寡聚核苷酸。被修饰的寡聚核苷酸可以包含,例如 2' - 氧 - 烷基化核糖残基,或非核糖的另一种糖,例如阿拉伯糖。

[0088] 优选地,根据本发明描述的蛋白质和多肽已被分离。术语“分离的蛋白质”或“分离的多肽”意思是蛋白质或多肽已经从其天然环境中分离。已被分离的蛋白质或多肽可以是基本上已经纯化的状态。术语“基本上已经纯化”意思是蛋白质或多肽基本上没有与之在自然界或体内相关的其它物质。

[0089] 这样的蛋白质和多肽可以,例如,用于生产抗体和用于免疫学或诊断学试验或用作治疗剂。根据本发明描述的蛋白质和多肽可以从生物样品中分离出来,例如组织或细胞匀浆,并且也可以重组表达于多种多样的原核或真核表达系统。

[0090] 为了本发明的目的,蛋白质或多肽的衍生物或者氨基酸序列的衍生物包含氨基酸插入变异体,氨基酸缺失变异体和 / 或氨基酸替代变异体。

[0091] 氨基酸插入变异体包含氨基端和 / 或羧基端融合以及特定的氨基酸序列中单个或两个或两个以上氨基酸的插入。在具有插入的氨基酸变异体的情况下,一个或多个氨基酸残基被插入氨基酸序列中特定的位点,尽管随机的插入和合适筛选得到的产物也是可能的。氨基酸缺失变异体特征在于一个或多个氨基酸从序列中的去除。氨基酸取代变异体特征在于序列中至少一个残基被去除并且在它的位置上插入另外一个残基。优先选择修饰在同源蛋白质或多肽之间非保守的氨基酸序列的位置。优选选择用其它具有相同特性例如疏水性、亲水性、电负性、侧链的体积等等的氨基酸去取代(保守替代)的氨基酸。保守替代,例如,涉及一个氨基酸与另一个如下列出的与被替代的氨基酸一组的氨基酸进行交换:

[0092] 1. 小脂肪族、非极性或者有点极性的残基:丙氨酸、丝氨酸、苏氨酸(脯氨酸、甘氨酸)

[0093] 2. 带负电的残基和它们的酰胺:天冬酰胺、天冬氨酸、谷氨酸、谷氨酰胺

[0094] 3. 带正电的残基:组氨酸、精氨酸、赖氨酸

[0095] 4. 大脂肪族、非极性残基:甲硫氨酸、亮氨酸、异亮氨酸、缬氨酸(半胱氨酸)

[0096] 5. 大芳香族残基:苯丙氨酸、酪氨酸、色氨酸。

[0097] 由于它们在蛋白质体系结构中的特定角色,三个残基显示在括号内。甘氨酸是唯一没有侧链的残基并且因而授予碳链柔性。脯氨酸具有与众不同的大大限制碳链的几何结构。半胱氨酸能形成二硫键桥。

[0098] 上述氨基酸变异体借助于已知肽合成技术可以很容易地进行制备,例如通过固相合成(Merrifield,1964)以及相似的方法或者通过重组DNA操作。用于将预先确定位点的替代突变引入具有已知或部分已知序列的DNA是众所周知的,并且包含例如M13突变成。用于制备具有替代、插入或缺失突变的蛋白质之DNA序列的操作在例如Sambrook等人,(1989)中有详细描述。

[0099] 根据本发明,蛋白质、多肽或肽的“衍生物”也包含与酶相关的任何分子的单个或多个替代、缺失和 / 或添加,例如碳水化合物、脂肪和 / 或蛋白质、多肽或肽。术语“衍生物”也延伸到所述蛋白质、多肽或肽的所有功能性化学等同物。

[0100] 根据本发明,肿瘤相关抗原的一部分或片段从多肽衍生而来,具有多肽的功能特性。这样的功能特性包含与抗体的相互作用,与其它多肽或蛋白质的相互作用,与核酸的选择性结合以及酶活性。一个特定的特性是与HLA形成复合物的能力,并且在合适的情况下产生免疫反应。此免疫反应以刺激细胞毒性或T辅助细胞为基础。本发明的肿瘤相关抗原的一部分或片段优选地包含肿瘤相关抗原的至少6个,尤其是至少8个、至少10个、至少12个、至少15个、至少20个、至少30个或至少50个连续氨基酸。

[0101] 编码肿瘤相关抗原的核酸的一部分或片段根据本发明涉及核酸的一部分,所述核酸的一部分如上定义地,编码至少肿瘤相关抗原和 / 或肿瘤相关抗原的一部分或片段。

[0102] 编码肿瘤相关抗原的基因的分离和鉴别也使特征在于一个或多个肿瘤相关抗原

的表达之疾病的诊断成为可能。这些方法包含测定编码肿瘤相关抗原的一个或多个核酸和 / 或测定编码的肿瘤相关抗原和 / 或从它们那里衍生的肽。核酸还可以以传统的方法进行测定,包括通过多聚酶链式反应或用标记探针的杂交。肿瘤相关抗原或从它们那里衍生来的肽也可以通过筛选识别抗原和 / 或肽的患者抗血清进行测定。它们也可以通过筛选患者的 T 细胞对相应的肿瘤相关抗原的特异性进行测定。

[0103] 本发明还使得可结合此处描述的肿瘤相关抗原的蛋白质被分离,包括抗体和所述肿瘤相关抗原的细胞内结合伴侣。

[0104] 根据本发明,特定的实施方案应该涉及提供源自肿瘤相关抗原的“显性失活”多肽。显性失活多肽是非活性的蛋白质变异体,所述显性失活多肽经由与细胞内机器相互作用,将活性蛋白质从其与细胞内机器 (cellular machinery) 的相互作用取代,或者所述显性失活多肽与活性蛋白质竞争,因此减少所述活性蛋白质的作用。例如,结合于配体但对结合于配体不产生任何信号作为应答的显性失活受体能够减少所述配体的生物学作用。相似地,通常与靶蛋白质相互作用但不磷酸化所述靶蛋白质的显性失活催化的非活性激酶减少所述靶蛋白质的磷酸化,作为对细胞信号的反应。相似地,结合于基因控制区域中的启动子位点但不增加所述基因的转录的显性失活转录因子可通过占据启动子位点但不增加转录来减少正常转录因子的作用。

[0105] 细胞中显性失活多肽表达的结果是活性蛋白质功能的降低。技术工人可制备蛋白质的显性失活变异体,例如通过传统的突变方法以及通过评价变异体多肽的显性失活作用。

[0106] 本发明还包含如下物质:例如结合肿瘤相关抗原的多肽。这样的结合物质可用于,例如,为了检测肿瘤相关抗原以及肿瘤相关抗原和其结合伴侣的复合物的筛选试验,以及用于肿瘤相关抗原的纯化和其与结合伴侣的复合物的纯化。这样的物质还用于抑制肿瘤相关抗原的活性,例如通过与这样的抗原结合。

[0107] 所以本发明包含结合的物质,例如,抗体或抗体片段,所述的结合物质能选择性地结合于肿瘤相关抗原。抗体包含用传统方式生产的多克隆和单克隆抗体。

[0108] 这样的抗体能够识别天然和 / 或变性状态的蛋白质 (Anderson 等人, J. Immunol. 143 :1899-1904,1989 ;Gardsvoll, J. Immunol. Methods 234 :107-116,2000 ; Kayyem 等人, Eur. J. Biochem. 208 :1-8,1992 ;Spiller 等人, J. Immunol. Methods 224 :51-60,1999)。

[0109] 含有特异性结合于靶蛋白质的特异性抗体之抗血清能够通过不同的标准方法进行制备;见,例如“Monoclonal Antibodies :A Practical Approach”Philip Shepherd, Christopher Dean ISBN 0-19-963722-9 ;“Antibodies :A Laboratory Manual”Harlow, David Lane 编, ISBN :0879693142 以及“Using Antibodies :A Laboratory Manual : Portable Protocol NO”Edward Harlow, David Lane, 编 Harlow ISBN0879695447。因此也可能产生识别天然形式的复合物膜蛋白质的姻亲 (affine) 和特异性抗体 (Azorsa 等人, J. Immunol. Methods 229 :35-48,1999 ;Anderson 等人, J. Immunol. 143 :1899-1904,1989 ; Gardsvoll, J. Immunol. Methods 234 :107-116,2000)。这尤其与用于治疗的抗体的制备相关,也与许多诊断应用相关。在这个方面,用全蛋白质,用细胞外部分序列以及用表达生理折叠形式的靶分子的细胞进行免疫是可能的。

[0110] 单克隆抗体传统上用杂交瘤技术进行制备。(技术细节见:“Monoclonal Antibodies:A Practical Approach”Philip Shepherd, Christopher Dean ISBN:0-19-963722-9;“Antibodies:A Laboratory Manual”Harlow, David Lane 编 ISBN:0879693142;“Using Antibodies:A Laboratory Manual:Portable Protocol NO”Edward Harlow, David Lane, 编 Harlow ISBN:0879695447)

[0111] 已知只有抗体分子的小部分,抗体决定簇,参与抗体与其抗原决定簇的结合(参看 Clark, W. R. (1986), The Experimental Foundations of Modern Immunology, Wiley&Sons, Inc., New York;Roitt, I. (1991), Essential Immunology, 第 7 版, Blackwell Scientific Publications, Oxford)。pFc' 和 Fc 区域是,例如,补体级联的效应物,但不参与抗原结合。其 pFc' 区域已经用酶去除或者生产的抗体没有 pFc' 区域的抗体,被提及为 F(ab') 片段,携带一个完全抗体的两个抗原结合位点。相似地,其 Fc 区域已经被酶去除或生产的抗体没有所述的 Fc 区域的抗体,被提及为 Fab 片段,携带一个完整抗体分子的一个抗原结合位点。此外, Fab 片段由抗体的共价结合轻链和所述抗体的部分重链组成,被提及为 Fd。Fd 片段是抗体特异性的主要决定因素(单个 Fd 片段能够与高达 10 个不同的轻链相联,不改变抗体的特异性)并且 Fd 片段当被分离时保留结合于抗原决定簇的能力。

[0112] 位于抗体的抗原结合部分内的是互补决定区域(CDR),所述互补决定区域与抗原的抗原决定簇和维持抗体决定簇三级结构的框架区域(FRs)直接相互作用。重链的 Fd 片段和 IgG 免疫球蛋白的轻链都含有四个框架区域(FR1 至 FR4),所述框架区域在每种情况下被三个补体决定区域(CDR1 至 CDR3)分开。CDR 并且尤其是 CDR3 区域,并且更加尤其是重链的 CDR3 区域很大程度上负责抗体的特异性。

[0113] 已知哺乳动物抗体的非 CDR 区域能够被有相同或不同特异性的抗体的相似区域替代,保留原先抗体对抗原的特异性。这就有可能开发“人源化”抗体,在所述抗体内,非人类 CDR 与人类 FR 和 / 或 Fc/pFc' 区域共价连接而产生功能性抗体。

[0114] 这被用在所谓的“SLAM”技术之中,其中来自全血的 B 细胞被分离并且细胞被单克隆。然后,对单个 B 细胞的上清液的抗体特异性进行分析。与杂交瘤技术相比,抗体基因的可变区域使用单细胞 PCR 进行扩增,并将其克隆进合适的载体中。这样,单克隆抗体的供应被加快了(de Wildt 等人, J. Immunol. Methods 207:61-67, 1997)。

[0115] 与另一个实施例一样,WO 92/04381 描述了人源化的鼠 RSV 抗体的生产和使用,所述抗体中至少鼠 FR 区域的部分被人源 FR 区域取代。此类抗体,包括有抗原结合能力的完整抗体的片段,被经常提及为“嵌合”抗体。

[0116] 本发明还提供抗体的 F(ab')₂、Fab、Fv 和 Fd 片段, Fc 和 / 或 FR 和 / 或 CDR1 和 / 或 CDR2 和 / 或轻链 - CDR3 区域被同源的人类或非人类序列取代的嵌合抗体, FR 和 / 或 CDR1 和 / 或 CDR2 和 / 或轻链 - CDR3 区域被同源的人类或非人类序列取代的嵌合 F(ab')₂ 片段抗体, FR 和 / 或 CDR1 和 / 或 CDR2 和 / 或轻链 - CDR3 区域被同源的人类或非人类序列取代的嵌合 Fab 片段抗体, 以及 FR 和 / 或 CDR1 和 / 或 CDR2 区域被同源的人类或非人类序列取代的嵌合 Fd 片段抗体。本发明也包含“单链”抗体。

[0117] 本发明还包括特异性结合肿瘤相关抗原的多肽。此类多肽结合物质,例如,可由简并性肽库(degenerate peptide libraries)来提供,所述简并性肽库简单地在溶液中被制备为固定的形态或被制备为噬菌体展示库(phage-display libraries)。也可能制备有一

个或多个氨基酸的肽之组合库。类肽 (Peptoid) 和非肽的合成残基的库也可被制备。

[0118] 噬菌体展示在鉴别本发明的结合肽尤其有效。就此而论,例如,呈现长度从 4 到大约 80 个氨基酸残基的插入物的噬菌体库被制备(使用,例如, M13、fd 或 lambda 噬菌体)。然后携带结合于肿瘤相关抗原的插入体的噬菌体被挑选。这个方法可通过两个或两个以上循环的结合于肿瘤相关抗原的噬菌体的再挑选而被重复。重复的回合导致携带指定序列的噬菌体的集中。为了鉴别表达多肽的序列可进行 DNA 序列的分析。结合于肿瘤相关抗原的序列的最小线性部分可被确定。酵母“双杂交系统”也用于鉴别结合于肿瘤相关抗原的多肽。为了鉴别和选择肿瘤相关抗原的肽结合伴侣,根据本发明描述的肿瘤相关抗原或其片段可用于筛选肽库,包括噬菌体陈列库。这样的分子,例如,可以用于筛选试验、纯化试验方案,用于肿瘤相关抗原功能的干扰以及用于其他技术工人已知的用途。

[0119] 上述抗体和其它结合分子可以用于,例如,鉴别表达肿瘤相关抗原的组织。为了陈列表达肿瘤相关抗原细胞和组织,抗体也可以偶联于特异性的诊断物质。它们也可以偶联于在治疗上有用的物质。诊断物质以非限制性的方式包含硫酸钡、碘西他酸、碘番酸、碘泊酸钙、泛影酸钠、泛影葡胺、甲泛影酰胺、酪洋酸钠以及放射性诊断试剂,包括正电子放射物质例如氟¹⁸、碳¹¹, γ -放射物质例如碘¹²³、钨^{99m}、碘¹³¹和钆¹¹¹,用于核磁共振的核素,例如氟和钆。根据本发明,术语“治疗上有用的物质”意思是任何治疗分子,如期望地,选择地被指引到表达一个或多个肿瘤相关抗原的细胞,包括抗癌剂、放射活性碘标记的化合物、毒素、细胞抑制或细胞溶解剂等等。所述抗癌剂包括,例如,氨鲁米特、硫唑嘌呤、博来霉素硫酸盐、白消安、卡氮芥、苯丁酸氮芥、顺铂、环磷酰胺、环孢霉素、cytarabidine、达卡巴嗪、放线菌素、柔红霉素、阿霉素、紫杉醇、依托泊甙、氟尿嘧啶、干扰素- α 、洛莫司汀、巯嘌呤、氨甲蝶呤、米托坦、甲苄肼盐酸盐、硫鸟嘌呤、长春碱硫酸盐和长春新碱硫酸盐。其它的抗癌剂在例如 Goodman 和 Gilman, “The Pharmacological Basis of Therapeutics”, 第 8 版, 1990, McGraw-Hill, Inc., 尤其是第 52 章 (Antineoplastic Agents (Paul Calabresi 和 Bruce A. Chabner)), 有描述。所述毒素可以是蛋白质,例如美洲商陆抗病毒蛋白质、霍乱毒素、百日咳毒素、蓖麻毒素、gelonin、相思豆毒素、白喉外毒素或假单胞菌外毒素。毒素残基也可以是高能量放射的放射性核素,例如钴⁶⁰。

[0120] 术语“患者”根据本发明意思是人类、非人灵长类或者其它的动物,尤其是哺乳动物例如母牛、马、猪、绵羊、山羊、狗、猫或啮齿动物例如小鼠和大鼠。在特别优选的实施方案中,患者是人类。

[0121] 根据本发明,术语“疾病”指的是表达或异常表达肿瘤相关抗原的任何病理状态。“异常表达”根据本发明意思是表达发生改变,优选地和健康个体的状态相比表达增加。表达增加指的是增加至少 10%,尤其是至少 20%、至少 50%或至少 100%。在一个实施方案中,肿瘤相关抗原仅表达在患病个体的组织中,然而在健康个体其表达被抑制。这样的疾病的一个实施例是癌症,其中术语“癌症”根据本发明包含白血病、精原细胞瘤、黑素瘤、畸胎瘤、神经胶质瘤、肾脏癌症、肾上腺癌症、甲状腺癌症、肠癌、肝癌、结肠癌、胃癌、胃肠道癌、淋巴结癌、食管癌、结直肠癌、胰腺癌、耳、鼻和喉 (ENT) 癌、乳腺癌、前列腺癌、子宫的癌症、卵巢癌和肺癌以及它们的转移灶。

[0122] 根据本发明,生物样品可以是组织样品和 / 或细胞样品,并且可用传统方式获得,例如通过组织活检,包括打孔组织活检,以及通过采集血样、支气管吸出物、唾液、尿液、粪

便或其它体液,以用于此处所述的各种方法。

[0123] 根据本发明,术语“免疫活性细胞”意思是在合适的刺激下能成熟分化成免疫细胞(例如 B 细胞、T 辅助细胞或溶细胞性 T 细胞)。免疫活性细胞包括 CD34⁺ 造血干细胞、不成熟和成熟 T 细胞以及不成熟和成熟 B 细胞。如果识别肿瘤相关抗原的溶细胞性 T 细胞或 T 辅助细胞的生产是期望的,那么在支持溶细胞性 T 细胞和 T 辅助细胞的生产、分化和 / 或选择的条件下将免疫活性细胞与表达肿瘤相关抗原的细胞接触。当暴露于抗原时, T 细胞的前体分化成溶细胞性 T 细胞与免疫系统的克隆选择相似。

[0124] 一些治疗方法以患者免疫系统的反应为基础,导致抗原递呈细胞例如递呈一个或多个肿瘤相关抗原的癌细胞的溶解。就此而论,例如特异于肿瘤相关抗原和 MHC 分子的复合物之自体同源的细胞毒性 T 淋巴细胞被给予具有细胞异常的患者。这样的细胞毒性 T 淋巴细胞体外的生产是已知的。使 T 细胞分化的方法的实施例能在 WO-A-9633265 中找到。一般地,含有细胞例如血细胞的样品从患者采集,并且将细胞与递呈复合物并导致细胞毒性 T 淋巴细胞(例如树突细胞)增殖(propagation)的细胞接触。靶细胞可以是被转染的细胞例如 COS 细胞。这些被转染的细胞在它们表面上递呈期望的复合物,并且当与细胞毒性 T 淋巴细胞接触时,可刺激淋巴细胞的增殖。然后克隆扩增的自体同源细胞毒性 T 淋巴细胞被给予患者。

[0125] 在另一个选择抗原特异性的细胞毒性 T 淋巴细胞的方法中, MHC I 型分子 / 肽复合物的荧光四聚体被用于检测细胞毒性 T 淋巴细胞的特异性克隆 (Altman 等人, Science 274 :94-96, 1996 ;Dunbar 等人, Curr. Biol. 8 :413-416, 1998)。可溶性 MHC I 型分子体外在 β_2 微球蛋白和结合于所述 I 型分子的肽抗原存在下是折叠的。MHC/ 肽的复合物被纯化并且用生物素对其进行标记。四聚体通过 4 : 1 的分子比例混合生物素化肽 -MHC 的复合物和标记的抗生物素蛋白(例如,藻红蛋白)而形成。然后四聚体与细胞毒性 T 淋巴细胞例如外周血或淋巴结接触。所述四聚体结合于识别肽抗原 /MHC I 型复合物的细胞毒性 T 淋巴细胞。结合于所述四聚体的细胞可以通过荧光控制的细胞分拣器进行分拣以分离反应的细胞毒性 T 淋巴细胞。被分离的细胞毒性 T 淋巴细胞然后可体外增殖。

[0126] 在被提及为过继性转移的治疗方法中 (Greenberg, J. Immunol. 136(5) :1917, 1986 ;Riddell 等人, Science 257 :238, 1992 ;Lynch 等人, Eur. J. Immunol. 21 :1403-1410, 1991 ;Kast 等人, Cell 59 :603-614, 1989), 递呈期望的复合物的细胞(例如树突细胞)联合于要接受治疗的患者的细胞毒性 T 淋巴细胞,导致特异性细胞毒性 T 淋巴细胞的增殖。增殖的细胞毒性 T 淋巴细胞然后被给予具有特征在于递呈特异性复合物的指定异常细胞的细胞异常之患者。然后细胞毒性 T 淋巴细胞溶解异常细胞,因此达到期望的治疗作用。

[0127] 经常,在患者的 T 细胞库中,只有对此类特异性复合物有低亲和力的 T 细胞能被增殖,因为那些有高亲和力的细胞由于耐受性的发展已经被消灭了。这里可选的是 T 细胞受体自身的转移。也为了这个,递呈期望复合物的细胞(例如树突细胞)联合于健康的人或另一个物种(例如鼠)的细胞毒性 T 淋巴细胞。如果 T 淋巴细胞源自以前未接触过特异性复合物的供体生物体,这就会导致有高亲和力的特异性细胞毒性 T 淋巴细胞的增殖。这些增殖的特异性 T 淋巴细胞的高亲和力 T 细胞受体被克隆。如果高亲和力 T 细胞受体从另一个物种克隆而来,它们能被不同程度的人源化。这样的 T 细胞受体然后通过基因转移,例如使用逆转录载体如期望地被转入患者的 T 细胞。然后过继性转移用这些基因改变的 T 淋巴细

胞进行 (Stanislawski 等人, *Nat Immunol.* 2 :962-70, 2001 ;Kessels 等人, *Nat Immunol.* 2 : 957-61, 2001)。

[0128] 上面的治疗方面从至少患者一些异常细胞递呈肿瘤相关抗原和 HLA 分子的复合物的事实出发。这样的细胞可用本身已知的方式进行鉴别。一旦递呈复合物的细胞被鉴别, 它们就可以联合来自患者的样品。如果细胞毒性 T 淋巴细胞溶解递呈复合物的细胞, 就能假设肿瘤相关抗原被递呈。

[0129] 过继性转移不是根据本发明能够应用的唯一的治疗形式。细胞毒性 T 淋巴细胞也可用本身已知的方式产生。一种方法使用表达复合物的非增殖细胞。这里使用的细胞将使那些通常表达复合物的细胞, 例如经照射的肿瘤细胞或被转染了一个或两个递呈复合物 (即抗原性肽和递呈 HLA 分子的复合物) 必须的基因的细胞。各种的细胞类型可被使用。此外, 可能使用携带一个或两个目的基因的载体。特别优选病毒或细菌载体。例如, 编码肿瘤相关抗原或其部分的核酸功能性地连接于控制所述肿瘤相关抗原或其片段在指定的组织或细胞类型中表达的启动子和增强子序列。核酸可被整合进表达载体。表达载体可以是外源核酸可被插入其中的非修饰的染色体外核酸、质粒或病毒基因组。编码肿瘤相关抗原的核酸也可被插入逆转录病毒基因组, 因此使得核酸能被整合进靶组织或靶细胞的基因组内。在这些系统中, 微生物例如疫苗病毒、痘病毒、单纯疱疹病毒、逆转录病毒或腺病毒携带目的基因并且实际地“感染”宿主细胞。另一个优选的形式是重组 RNA 形式的肿瘤相关抗原的引入, 所述重组 RNA 可例如通过脂质体转移或通过电穿孔被引入细胞。得到的细胞递呈目的复合物并且被后来增殖的自体同源细胞毒性 T 淋巴细胞识别。

[0130] 为了使体内整合进抗原递呈细胞成为可能, 相似的作用可通过将肿瘤相关抗原或其片段联合佐剂而获得。肿瘤相关抗原或其片段可被递呈为蛋白质、DNA (例如在载体内) 或 RNA。为了产生 HLA 分子的肽伴侣, 肿瘤相关抗原被加工, 但是其片段不需要进一步加工就可递呈。后者如果能结合于 HLA 分子尤其可行。优先选择完整抗原被树突细胞体内加工的给予形式, 因为这也可生产有效的免疫反应必需的 T 辅助细胞反应 (Ossendorp 等人, *Immunol Lett.* 74 :75-9, 2000 ;Ossendorp 等人, *J. Exp. Med.* 187 :693-702, 1998)。一般而言, 可能例如通过皮内注射将有效量的肿瘤相关抗原给予患者。然而, 也可进行淋巴内注射 (Maloy 等人, *Proc Natl Acad Sci USA* 98 :3299-303, 2001)。给药也可联合促进树突细胞摄入的试剂而进行。优选的肿瘤相关抗原包含与同种异体癌症抗血清或与许多癌症患者的 T 细胞反应的那些肿瘤相关抗原。尤其引人注目的, 然而, 是抗预先存在的非自发性免疫反应的那些。明显地, 有可能诱导抗能够溶解肿瘤的这些免疫反应 (Keogh 等人, *J. Immunol.* 167 :787-96, 2001 ;Appella 等人, *Biomed Pept Proteins Nucleic Acids* 1 : 177-84, 1995 ;Wentworth 等人, *Mol Immunol.* 32 :603-12, 1995)。

[0131] 根据本发明描述的药用组合物也可用于免疫的疫苗。根据本发明, 术语“免疫”或“疫苗接种”意思是对抗原的免疫反应的激活或增加。为了测试使用肿瘤相关抗原或编码肿瘤相关抗原的核酸的对癌症的免疫作用, 有可能使用动物模型。例如, 人类癌症细胞可被引入小鼠以产生肿瘤, 并且编码肿瘤相关抗原的一个或多个核酸可被给予。可测量对癌症细胞的作用 (例如肿瘤体积的减小) 作为测量用核酸免疫的有效性的手段。

[0132] 为了诱导免疫反应或增加免疫反应, 一个或多个肿瘤相关抗原或其刺激片段, 作为用于免疫的组合物的一部分, 可与一个或多个佐剂一起给予。佐剂是被整合进抗原或与

抗原一起给予并增强免疫反应的物质。佐剂通过提供抗原库（细胞之外地或在巨噬细胞内地）、激活巨噬细胞和 / 或刺激特定的巨噬细胞来增强免疫反应。佐剂是已知的,并且以非限制性的方法包含单磷酸脂 A(MPL, SmithKline Beecham)、皂甙例如 QS21 (SmithKline Beecham)、DQS21 (SmithKline Beecham; WO 96/33739)、QS7、QS17、QS18 和 QS-L1 (So 等人, Mol. Cells 7 :178-186, 1997)、弗氏不完全辅佐剂、弗氏完全辅佐剂、维生素 E、montanide、明矾、CpG 寡聚核苷酸（参见 Kreig 等人, Nature 374 :546-9, 1995）以及由生物可降解油例如角鲨烯和 / 或生育酚制备而来的各种油包水型乳剂。优选地,肽以与 DQS21/MPL 混合物的形式被给予。DQS21 与 MPL 的比例一般是大约 1 : 10 至 10 : 1, 优选的是大约 1 : 5 至 5 : 1, 并且尤其是 1 : 1。给予人类时,疫苗制剂一般含有大约 1 μ g 至大约 100 μ g 的 DQS21 和 MPL。

[0133] 刺激患者免疫反应的其它物质也可被给予。由于细胞因子对淋巴细胞的调节特性,有可能例如在疫苗接种中使用细胞因子。这样的细胞因子包含,例如显示能增加疫苗的保护作用的白细胞介素 -12 (IL-12) (参见 Science 268 :1432-1434, 1995), 粒细胞巨噬细胞 - 克隆刺激因子以及 IL-18。

[0134] 有很多增强免疫反应所以可以用于疫苗接种的化合物。所述的化合物包含以蛋白质或核酸形式提供的共刺激分子。这样的共刺激分子的实施例是表达于树突细胞 (DC) 上并与表达于 T 细胞上的 CD28 分子相互作用的 B7-1 和 B7-2 (分别是 CD80 和 CD86)。这种相互作用给抗原 /MHC/TCR 刺激 (信号 1) 的 T 细胞提供了共刺激 (信号 2), 因此增强所述 T 细胞的增殖和效应物功能。B7 也与 T 细胞上的 CTLA4 (CD152) 相互作用, 并且涉及 CTLA4 和 B7 配体的研究证明了 B7-CTLA4 的相互作用能增强抗肿瘤的免疫性和 CTL 的增殖 (Zheng, P. 等人, Proc. Natl. Acad. Sci. USA 95 (11) :6284-6289 (1998))。

[0135] B7 一般不在肿瘤细胞上表达, 所以这些对 T 细胞来说不是有效的抗原递呈细胞 (APC)。B7 表达的诱导使肿瘤细胞能刺激细胞毒性 T 淋巴细胞更加有效的增殖以及效应物的功能。通过 B7/IL-6/IL-12 联合的共刺激揭露 T 细胞群中 IFN- γ 和 Th1 细胞因子曲线 (profile) 的诱导, 导致 T 细胞活性的进一步增强 (Gajewski 等人, J. Immunol. 154 : 5637-5648 (1995))。

[0136] 细胞毒性 T 淋巴细胞的完全激活和完全的效应物功能需要 T 辅助细胞通过所述 T 辅助细胞上的 CD 配体和树突细胞表达的 CD40 分子的相互作用的参与 (Ridge 等人, Nature 393 :474 (1998), Bennett 等人, Nature 393 :478 (1998), **Schönberger** 等人, Nature 393 :480 (1998))。这个共刺激信号的机制很可能涉及由所述树突细胞 (抗原递呈细胞) 产生的 B7 以及产生的相关的 IL-6/IL-12 的增加。CD40-CD40L 相互作用因而补足了信号 1 (抗原 /MHC-TCR) 和信号 2 (B7-CD28)。

[0137] 抗 CD40 抗体用于刺激树突细胞被期望能直接增强对肿瘤抗原的反应, 通常所述肿瘤抗原在炎症反应的范围之外或由非专业的抗原递呈细胞 (肿瘤细胞) 递呈。在这些情况下, T 辅助细胞和 B7 共刺激信号未被提供。这个机制可与基于抗原刺激的 (antigen-pulsed) 树突细胞的治疗连接起来使用。

[0138] 本发明还提供了核酸、多肽或肽的给予方式。多肽和肽可用本身就已知的方式被给予。在一个实施方案中, 核酸可通过离体 (ex vivo) 方法被给予, 即通过将细胞从患者拿去, 为了整合肿瘤相关抗原对所述细胞的基因修饰, 并且将改变的细胞再引入患者。这一般

包含体外将基因的功能性拷贝引入患者的细胞内并且将基因改变的细胞重新引入患者。基因的功能性拷贝在调节元件的功能性控制之下,所述调节元件允许基因表达于基因改变的细胞中。转染和转导的方法是技术人员已知的。本发明还提供使用例如病毒和靶点控制的脂质体的载体体内给予核酸。

[0139] 在优选的实施方案中,用于给予编码肿瘤相关抗原的核酸的病毒载体选自腺病毒、腺伴随病毒、痘病毒包括痘苗病毒和减毒痘病毒、圣基利森林 (Semliki Forest) 病毒、逆转录病毒、辛德必斯 (Sindbis) 病毒和 Ty 病毒样颗粒。尤其优选腺病毒和逆转录病毒。逆转录病毒一般是复制缺陷的 (即它们不能产生有感染力的颗粒)。

[0140] 为了根据本发明将核酸体外或体内引入细胞内可使用各种方法。此类方法包含磷酸钙沉淀核酸的转染、与 DEAE 相关的核酸的转染、用上面携带目的核酸的病毒的转染或感染、脂质体介导的转染等等。在特定的实施方案中,优先选择将核酸指引至特定的细胞的方法。在这样的实施方案中,用于将核酸给予到细胞的载体 (例如逆转录病毒或脂质体) 具有结合靶点的控制分子。例如,分子例如特异于靶细胞上表面膜蛋白的抗体或者特异于靶细胞上受体的配体可以被整合进或附着于核酸载体。优选的抗体包含与肿瘤相关抗原选择性结合的抗体。如果通过脂质体给予核酸是期望的,结合于细胞内吞相关的表面膜蛋白的蛋白质可被整合进脂质体制剂中使得靶控制和 / 或摄取成为可能。这样的蛋白质包含特异于特定的细胞类型的衣壳蛋白或其片段,抗内化蛋白质的抗体,定位于细胞内位点的蛋白质等等。

[0141] 本发明的治疗组合物可以以药物相容性制剂的形式给予。这样的制剂通常含有药物相容浓度的盐、缓冲物质、防腐剂、载体、补充免疫增强物质例如辅佐剂、CpG 和细胞因子,以及在合适的情况下其它的治疗活性化合物。

[0142] 本发明的治疗活性化合物可以通过任何传统途径给予,包括注射或输入 (infusion)。给药方式可采用例如口服、静脉、腹腔、肌肉内、皮下或经皮地。优选地,抗体经由肺气雾剂被治疗性的给予。反义核酸优选地通过缓慢静脉给药方式被给予。

[0143] 本发明的组合物以有效的量被给予。“有效量”指的是自身或者与更多的剂量一起获得期望的反应或期望的作用的量。在治疗以表达一个或多个肿瘤相关抗原为特征的特定疾病或特定状态的情况下,期望的反应涉及疾病过程的抑制。这包含减慢疾病的发展,并且尤其是打断疾病的发展。在治疗疾病或不适中期望的反应也可以是所述疾病或所述不适的发作被延期或发作被阻止。

[0144] 本发明组合物的有效量将依赖于要治疗的不适、疾病的严重程度、患者的个人参数,包括年龄、生理状态、体积和体重,治疗的持续时间、伴随治疗的类型 (如果有的话),给药的特异性途径以及相似的因素。

[0145] 本发明药物组合物优选地是无菌的并且含有有效量的治疗活性物质以产生期望的反应或期望的作用。

[0146] 本发明组合物给予的剂量依赖于各种参数,例如给药形式、患者的状态、期望的给药时程等等。在初始剂量不足以引起患者的反应的情况下,可以使用更高的剂量 (或通过不同的更加局部化的给药途径以获得有效地更高的剂量)。

[0147] 一般,为了治疗或为了产生或增加免疫反应,从 1ng 至 1mg,优选地从 10ng 至 100 μ g 剂量的肿瘤相关抗原被制成制剂并被给予。如果编码肿瘤相关抗原的核酸 (DNA 和

RNA) 的给予是所期望的,从 1ng 至 0.1mg 的剂量被制成制剂并被给予。

[0148] 本发明的药物组合物一般以药物相容的量并且以药物相容的组合物被给予。术语“药物相容的”指的是不与药物组合物中活性成分的作用相互作用的非毒性物质。此类制剂通常含有盐、缓冲物质、防腐剂、载体,以及合适情况下的其它治疗活性物质。当用于医疗时,盐应当是药物相容的。然而,药物不相容的盐可以用于制备药物相容的盐并且包括在本发明中。此类药理和药物相容的盐以非限制性的方法包含那些从下列酸制备的盐:盐酸、氢溴酸、硫酸、硝酸、磷酸、马来酸、醋酸、水杨酸、柠檬酸、蚁酸、丙二酸、琥珀酸等等。药物相容盐还可以被制备成碱金属盐或碱土金属盐,例如钠盐、钾盐或钙盐。

[0149] 本发明的药物组合物包含药物相容性载体。根据本发明,术语“药物相容性载体”指的是适合给予人类的一个或多个相容的固体或液体填充剂、稀释剂或做成胶囊的物质。术语“载体”指的是天然或合成的有机或无机组分,其中活性组分被联合为应用提供方便。本发明药物组合物的组分通常是不会发生基本上削弱期望的药物效应相互作用的这些组分。

[0150] 本发明的药用组合物可以含有合适的缓冲物质例如醋酸盐、柠檬酸盐、硼酸盐和磷酸盐。

[0151] 药物组合物,在合适的情况下,也含有合适的防腐剂例如苯扎氯铵、三氯叔丁醇、paraben 和硫柳汞。

[0152] 药物组合物通常以统一的剂量形式提供,并且以本身已知的方式被制备。本发明的药物组合物是例如胶囊、片剂、锭剂、悬浮剂、糖浆剂、酏剂的形式或者是乳剂的形式。

[0153] 适合胃肠外给药的组合物通常包含活性化合物的无菌水或非水制剂,所述组合物优选地是与受者的血液等渗的。相容载体和溶剂的实施例是林格(Ringer)溶液和等渗的氯化钠溶液。另外,通常无菌的不易挥发的油类被用作溶液或悬浮液的介质。

[0154] 本发明通过下面的用于例证的目的并不是意在限定的插图和实例被详细地描述。由于这些叙述和实施例,同样包括在本发明中的进一步的实施方案很容易被本领域技术人员理解。

附图说明

[0155] 图 1. GPR35 在结肠癌活检组织中 mRNA 的表达

[0156] 对无 DNA 的 RNA 的 RT-PCR 调查显示,GPR35 在多数结肠癌活检组织中的表达。相反,在正常组织中没有检测到表达。(1- 乳腺,2- 肺,3- 淋巴结,4- 胸腺,5- 结肠,6-15 结肠癌,16- 阴性对照)

[0157] 图 2. 定量 PCR 分析 GUCY2C 在正常和肿瘤组织中 mRNA 的表达

[0158] 用 GUCY2C 特异性引物 (SEQ ID NO :22-23) 的实时 PCR 调查显示 GUCY2C 在正常回肠、结肠和所有结肠癌活检组织中选择性 mRNA 的表达。对结肠癌的肝脏转移灶中 GUCY2C 转录产物的不同的量也进行检测。

[0159] 图 3. 肿瘤特异性的 GUCY2C 剪接变异体的鉴别

[0160] 来自正常结肠组织和结肠癌的 PCR 产物被克隆,并且对两组的克隆通过限制性分析 (EcoR I) 进行检查并且进行测序。

[0161] 图 4. 正常肺和肺癌中选择性 SCGB3A 表达

[0162] 用基因特异性 SCGB3A2 引物 (SEQ ID NO :37,38) 的 RT-PCR 分析显示只在正常肺 (8 道,14-15 道) 以及肺癌 (16-24 道) 的活检组织中有 cDNA 扩增。(1- 肝 - 正常, 2- 外周血单核细胞 - 正常,3- 淋巴结 - 正常,4- 胃 - 正常,5- 睾丸 - 正常,6- 乳腺 - 正常, 7- 肾 - 正常,8- 肺 - 正常,9- 胸腺 - 正常,10- 卵巢 - 正常,11- 肾上腺 - 正常,12- 脾 - 正常,14-15- 肺 - 正常,16-24- 肺癌,25- 阴性对照)

[0163] 图 5. Claudin-18A2.1 在胃、食管、胃癌和胰腺癌中的表达

[0164] 用 claudin-18A2.1 特异性引物 (SEQ ID NO :39,40) 的 RT-PCR 分析显示根据本发明在 8/10 胃癌活检组织和 3/6 胰腺癌活检组织有 claudin-18A2.1 显著的表达。明显的表达还可在胃和正常食管组织中检测到。相反,表达没有在卵巢和卵巢癌中检测到。

[0165] 图 6. SLC13A1 在肾和肾细胞癌中的表达

[0166] 用 SCL13A1 特异性引物 (SEQ ID NO :49,50) 的 RT-PCR 分析显示在 7/8 肾细胞癌的样品中有表达。另外的,转录产物正常组织中只有在肾可检测到。(1-2 肾,3-10- 肾细胞癌,11- 乳腺,12- 肺,13- 肝,14- 结肠,15- 淋巴,16- 脾,17- 食管,18- 胸腺,19- 甲状腺, 20- 外周血单核细胞,21- 卵巢,22- 睾丸。)

[0167] 图 7. CLCA1 在结肠、结肠癌和胃癌中的表达

[0168] 用 CLCA1 特异性引物 (SEQ ID NO :67,68) 的 RT-PCR 调查证实在结肠中的选择性表达并且显示在 3/7 被调查的结肠癌和 1/3 被调查的胃癌样品中的高表达。其它的正常组织 (NT) 显示无表达或仅有很弱的表达。

[0169] 图 8. FLJ21477 在结肠和结肠癌中的表达

[0170] 用 FLJ21477 特异性引物 (SEQ ID NO :69,70) 的 RT-PCR 调查显示在结肠中的选择性表达,并且另外在 7/12 被调查的结肠癌样本中的不同水平的表达。其它的正常组织 (NT) 未显示有表达。

[0171] 图 9. FLJ20694 在结肠和结肠癌中的表达

[0172] 用 FLJ20694 特异性引物 (SEQ ID NO :71,72) 的 RT-PCR 调查显示在结肠中的选择性表达,并且另外在 (5/9) 被调查的结肠癌样本中的不同水平的表达。其它的正常组织 (NT) 未显示有表达。

[0173] 图 10. von Ebner 在胃、肺和肺癌中的表达

[0174] 用 von Ebner 特异性引物 (SEQ ID NO :73,74) 的 RT-PCR 调查显示在胃中、在肺中以及在 5/10 被调查的胃癌样本中的选择性表达。其它的正常组织 (NT) 未显示有表达。

[0175] 图 11. Plunc 在胸腺、肺和肺癌中的表达

[0176] 用 Plunc 特异性引物 (SEQ ID NO :75,76) 的 RT-PCR 调查显示在胸腺中、在肺中以及在 (6/10) 被调查的肺癌样品中的选择性表达。其它的正常组织未显示有表达。

[0177] 图 12. SLC26A9 在肺、肺癌和甲状腺中的表达

[0178] 用 SLC26A9 特异性引物 (SEQ ID NO :77,78) 的 RT-PCR 调查显示在肺中以及在所有 (13/13) 被调查的胃癌样本中的选择性表达。其它的正常组织 (NT) 除甲状腺外未显示有表达。

[0179] 图 13. THC1005163 在胃、卵巢、肺和肺癌中的表达

[0180] 用 THC1005163 特异性引物 (SEQ ID NO :79) 和非特异性寡聚 dT 标签引物的 RT-PCR 调查显示在胃、卵巢、肺中以及在 (5/9) 被调查的肺癌活检组织中的选择性表达。其

它的正常组织 (NT) 未显示有表达。

[0181] 图 14. LOC134288 在肾和肾细胞癌中的表达

[0182] LOC134288 特异性引物 (SEQ ID NO :80,81) 的 RT-PCR 调查显示在肾中以及在 (5/8) 被调查的肾细胞癌活检组织中的选择性表达。

[0183] 图 15. THC943866 在肾和肾细胞癌中的表达

[0184] 用 THC943866 特异性引物 (SEQ ID NO :82,83) 的 RT-PCR 调查显示在肾中以及在 (4/8) 被调查的肾细胞癌活检组织中的选择性表达。

[0185] 图 16. FLJ21458 在结肠和结肠癌中的表达

[0186] 用 FLJ21458 特异性引物 (SEQ ID NO :86,87) 的 RT-PCR 调查显示在结肠中以及在 (7/10) 被调查的结肠癌活检组织中的选择性表达。(1-2- 结肠, 3- 肝, 4- 外周血单核细胞, 5- 脾, 6- 前列腺, 7- 肾, 8- 卵巢, 9- 皮肤, 10- 回肠, 11- 肺, 12- 睾丸, 13-22 结肠癌, 23- 阴性对照。)

[0187] 图 17. GPR35 的细胞定位

[0188] 免疫荧光法用于检测在转染表达 GPR35-GFP 融合蛋白质的质粒之后的 GPR35 细胞定位。箭头鉴别了荧光 GFP 的膜相关荧光。

[0189] 图 18. GPR35 的定量表达

[0190] A. 用 GPR35 特异性引物 (SEQ ID NO :88,89) 的定量 RT-PCR 显示在肠道中、在结肠肿瘤样品中以及在来自肠道肿瘤的转移灶中的选择性表达。对下列正常组织进行分析：肝、肺、淋巴结、胃、脾、肾上腺、肾、食管、卵巢、睾丸、胸腺、皮肤、乳腺、胰腺、淋巴细胞、激活的淋巴细胞、前列腺、甲状腺、输卵管、子宫内膜、小脑、脑。

[0191] B. 结肠肿瘤及其转移灶中 GPR35 的流行。GPR35 在肿瘤和转移灶多于 90% 的案例都有表达。

[0192] 图 19. GUCY2C 的定量表达

[0193] 用 GUCY2C 特异性引物 (SEQ ID NO :98,99) 的定量 RT-PCR 显示在正常结肠和胃组织中的高水平并且选择性的表达 (A) 以及结肠和胃肿瘤样本中 GUCY2C 的特异性表达 (B)。GUCY2C 在 11/12 结肠癌中和 7/10 胃癌中可检测到。

[0194] 图 20. SCGB3A2 的定量表达

[0195] 用 SCGB3A2 特异性的引物 (SEQ ID NO :103,104) 的定量 RT-PCR 显示在肺样品和肺肿瘤样品中的选择性表达。19/20 肺肿瘤样品是 SCGB3A2 阳性的, 并且 SCGB3A2 在多于 50% 的样品中有至少 10 倍的过度表达。对下列正常组织进行分析：肝、肺、淋巴结、胃、脾、肾上腺、肾、食管、卵巢、睾丸、胸腺、皮肤、乳腺、胰腺、淋巴细胞、激活的淋巴细胞、前列腺、甲状腺、输卵管、子宫内膜、小脑、脑。

[0196] 图 21. 用 SCGB3A2 特异性抗体的免疫荧光法

[0197] COS7 细胞被转染了编码 SCGB3A2-GFP 融合蛋白质的质粒。A. 用 SCGB3A2 特异性的兔抗血清的转染的融合蛋白质之检测 (用 SEQ ID NO :105 免疫)。B. 通过 GFP 荧光的转染的融合蛋白质之检测。C. 从 A 和 B 得到的两个荧光的叠加。在两个荧光叠加的点产生黄色, 并且因而证明 SCGB3A2 抗血清的特异性。

[0198] 图 22. claudin-18 剪接变异体的图表叙述

[0199] 两个 claudin-18 的剪接变异体 A1 和 A2 在氨基端不同并且显示了不同的潜在糖

基化位点。

[0200] 图 23. claudin-18 变异体 A1 的定量表达

[0201] Claudin-A1 在大量的肿瘤组织中被高度激活。特别强的表达在胃肿瘤、肺肿瘤、胰腺癌和食管癌中可以找到。

[0202] 图 24. claudin-18 变异体 A2 的定量表达

[0203] 变异体 A2 (如同变异体 A1) 在许多肿瘤中被激活。

[0204] 图 25. claudin-18A2 特异性抗体 (细胞外结构域) 的使用

[0205] (上) 通过用肽 (SEQ ID NO :17) 的免疫产生的抗体对 claudin-18A2 阳性的胃癌细胞 (SNU-16) 的染色。膜染色似乎在细胞 / 细胞相互作用的区域特别强。A- 免疫前, MeOH ; B- 免疫血清 MeOH, 5 μ g/ml ; (下) 通过 claudin-18A2-GFP 转染的 293T 细胞的共定位分析对抗体特异性的证明。A-Claudin-18A2GFP ; B- 抗 claudin-A2 ; C- 叠加。

[0206] 图 26. claudin-18A2 特异性抗体 (细胞外结构域) 的使用

[0207] 通过用肽 (SEQ ID NO :113, 位于氨基端的细胞外结构域) 的免疫产生的抗体对 claudin-18A2 阳性的胃癌细胞 (SNU-16) 的膜染色。抗 E-cadherin 的单克隆抗体被用于复染。A- 抗体 ; B- 复染 ; C- 叠加。

[0208] 图 27. 抗 claudin-18 羧基端细胞外结构域的抗体的使用

[0209] (左, 上和下) 通过用肽 (SEQ ID NO :116, 位于羧基端的细胞外结构域) 的免疫产生的抗体对 claudin-18A2 阳性的胃癌细胞 (SNU-16) 的膜染色。抗 E-cadherin 的单克隆抗体被用于复染 (右上, 下)。

[0210] 图 28. claudin-18A1 特异性抗体的使用

[0211] (上) 用 claudin-18A1 特异性肽 (SEQ ID NO :115) 的免疫产生的抗体对胃癌细胞 (SNU-16 ; claudin18A2 阳性) 的弱或无染色。A- 抗 E-cadherin ; B- 抗 claudin-18A1 ; C- 叠加。

[0212] (下) 通过转染了 claudin-18A1-GFP 的 293T 细胞的共定位分析对抗体特异性的证明。A-GFP-Claudin-18A1 ; B- 抗 claudin-18A1 ; C- 叠加。

[0213] 图 29. Western 印迹中 claudin-18A2 的检测。

[0214] 用抗有 SEQ ID NO :17 之抗原决定簇的 claudin-18A2 特异性抗体对来自各种健康组织的溶解物进行 Western 印迹。1- 胃 ; 2- 睾丸 ; 3- 皮肤 ; 4- 乳腺 ; 5- 肝 ; 6- 结肠 ; 7- 肺 ; 8- 肾 ; 9- 淋巴结。

[0215] 图 30. 对来自胃和胃癌样品进行的 Claudin-18A2 的 Western 印迹

[0216] 使用抗具有 SEQ ID NO :17 的抗原的 claudin-18A2 特异性抗体对来自胃和胃肿瘤的溶解物进行 Western 印迹并且进行测试。胃肿瘤显示了较低糖基化形式的 claudin-18A2。对胃溶解物的肽 N- 糖苷酶 F (PNGase F) 处理导致了低糖基化形式的形成。

[0217] 左 : 1- 胃 No#A ; 2- 胃 Tu#A ; 3- 胃 No#B ; 4- 胃 Tu#B

[0218] 右 : 1- 胃 No#A ; 2- 胃 No#B ; 3- 胃 No#B+PNGase F ; 4- 胃 Tu#C ; 5- 胃 Tu#D ; 6- 胃 Tu#D+PNGase F

[0219] 图 31. 肺肿瘤中 claudin-18 的表达

[0220] 与图 30 一致, 低糖基化的 claudin-18A2 变异体在肺肿瘤中被检测到。1- 胃 No ; 2- 胃 Tu ; 3- 肺 Tu。

[0221] 图 32. 使用 claudin-18A2 特异性抗体对胃肿瘤组织中 claudin-18 的免疫组织化学分析

[0222] 图 33. 用 claudin-18 特异性多克隆抗血清对胃特异性 Snu16 细胞的间接免疫荧光

[0223] A. 用免疫前产生的免疫前血清的染色 ;B. 用 claudin-18 特异性血清的染色。

[0224] 图 34. SLC13A1 的定量表达

[0225] 用 SLC13A1 特异性引物 (SEQ ID NO :121,122) 的定量 RT-PCR 显示在正常肾组织中高水平并且选择性的表达 (A) 以及 SCL13A1 在肾细胞癌中的特异性表达 (B)。SLC13A1 的转录在 5/8 肾细胞癌中可检测到。

[0226] 图 35. SLC13A1 的细胞定位

[0227] 免疫荧光法证明在提供 SLC13A1-GFP 融合蛋白质的质粒的转染之后 SLC13A1 的细胞定位。SLC13A1 融合蛋白质的膜相关荧光清晰可见 (为环绕转染细胞的环)。

[0228] 图 36. CLCA1 的定量表达

[0229] 用 CLCA1 特异性引物 (SEQ ID NO :125,126) 的定量 RT-PCR 在显示正常结肠组织和胃组织中高水平并且选择性的表达 (A) 以及 CLCA1 在结肠和胃肿瘤样品中的特异性表达 (B)。CLCA1 在 6/12 结肠癌中和在 7/10 胃癌中可检测到。

[0230] 图 37. FLJ21477 的定量表达

[0231] 用 FLJ21477 特异性引物 (SEQ ID NO :127,128) 的定量 RT-PCR 显示在正常结肠和胃组织中高水平并且选择性的表达以及在胸腺、食管和脑中的微弱表达 (A) 以及 FLJ21477 在结肠肿瘤样品中的特异性表达 (B)。FLJ21477 在 11/12 结肠癌中可检测到。

[0232] 图 38. FLJ20694 的定量表达

[0233] 用 FLJ20694 特异性引物 (SEQ ID NO :129,130) 的定量 RT-PCR 显示在正常结肠和胃组织中高水平并且选择性的表达 (A) 以及 FLJ21477 在结肠和胃肿瘤样本中的特异性过度表达 (B)。FLJ20694 在 11/12 结肠癌中和在 7/10 胃癌中可检测到。

[0234] 图 39. FLJ21458 的定量表达

[0235] 用 FLJ21458 特异性引物 (SEQ ID NO :133,134) 的定量 RT-PCR 显示在睾丸、胃和肠道组织中的选择性表达。另外,FLJ21458 特异性转录产物在 20/20 结肠肿瘤中和在 7/11 结肠转移灶中可检测到。对下列正常组织进行分析 :肝、肺、淋巴结、脾、肾上腺、肾、食管、卵巢、睾丸、胸腺、皮肤、乳腺、胰腺、淋巴细胞、激活的淋巴细胞、前列腺、甲状腺、输卵管、子宫内膜、小脑、大脑。

[0236] 图 40. 用 FLJ21458 特异性抗体的免疫荧光法

[0237] (上) 293 细胞被转染了编码 FLJ21458-GFP 融合蛋白质的质粒。A :用 FLJ21458 特异性兔抗血清 (用 SEQ ID NO :136 免疫化) 对转染的融合蛋白质的检测。B :通过 GFP 荧光对转染的融合蛋白质的检测。C :来自 A 和 B 的两个荧光的叠加。在两个荧光叠加的地方产生黄色并且因此证明了 FLJ21458 抗血清的特异性。

[0238] (下) 对内源性合成 FLJ21458 的 Snu16 细胞的分析。A :使用 FLJ21458 特异性兔抗血清 (用 SEQ ID NO :136 免疫) 的蛋白质检测。B :膜蛋白质 E- 钙粘蛋白的检测。C :来自 A 和 B 的两个荧光的叠加。在两个荧光叠加的地方产生黄色,并且证明了 FLJ21458 的膜定位。

[0239] 图 41. 序列

[0240] 在此处显示了涉及的序列。

实施例：

[0241] 材料和方法

[0242] 术语“在芯片上”、“电子的”和“虚拟克隆”仅仅指的是基于数据库的方法的使用，并且还可用于刺激实验室试验方法。

[0243] 除非有不同表达性的定义，所有其它的术语和表达方法被使用以使得技术工人可以理解。提到的技术和方法以本身已知的方式进行并且在例如 Sambrook 等人, Molecular Cloning :A Laboratory Manual, 第 2 版 (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. 中有描述。所用的方法包括试剂盒和试剂的使用根据厂家的信息进行。

[0244] 用于测定新的肿瘤相关基因的基于数据采集 (datamining) 的策略

[0245] 两个“在芯片上”的策略，名为基因库 (GenBank) 关键词搜索和 cDNAxProfiler 被联合。利用 NCBI ENTREZ 搜索和重获系统 (Retrieval System) (<http://www.ncbi.nlm.nih.gov/Entrez>), 为被注解为在特异性的组织中特异性表达的候选基因可进行基因库搜索 (Wheeler 等人, Nucleic Acids Research 28 :10-14, 2000)

[0246] 用例如“结肠特异性基因”、“胃特异性基因”或“肾特异性基因”的关键词进行查询，候选基因 (GOI, 目的基因) 从数据库中被摘录出。通过对于有机体使用限定词“人类”，对于分子类型使用 mRNA, 使得搜索被限定为这些数据库全部信息的一部分。

[0247] 找到的 GOI 清单通过确定相同序列的不同名字以及消除这样的多余而被整理。

[0248] 对关键词搜索得到的所有候选基因通过“电子 Northern 杂交” (eNorthern) 法依次研究它们的组织分布。eNorthern 以将 GOI 的序列与 EST (表达序列标记) 数据库进行排列比较为基础 (Adams 等人, Science 252 :1651, 1991) (<http://www.ncbi.nlm.nih.gov/BLAST>)。每个发现与插入的 GOI 同源的 EST 之组织来源能够被确定并且这样所有 EST 的总和得到对 GOI 组织分布的初步评价。只对那些与来自非器官特异性正常组织的 EST 无同源性的 GOI 作进一步的研究。这个评价也考虑公共网域含有错误注解的 cDNA 库 (Scheurle 等人, Cancer Res. 60 :4037-4043, 2000) (www.fau.edu/cmabb/publications/cancergen6.htm)。

[0249] 利用的第二个数据采集方法是 NCBI 癌症基因组解剖计划的 cDNAxProfiler (Hillier 等人, Genome Research 6 :807-828, 1996 ;Pennisi, Science 276 :1023-1024, 1997)。这个允许转录物组群在通过逻辑操作系统互相有关联的数据库中存放。我们已经定义了一个群 A, 将例如从结肠的正常组织中制备的所有表达文库都分配于其中, 把混合的文库排除在外。从除结肠外的正常组织制备的所有 cDNA 文库被分配到群 B。一般, 所有的 cDNA 文库不依赖于潜在的制备方法而被利用, 但是只有大小 > 1000 的那些被承认。群 B 依靠 BUT NOT 操作系统从群 A 中被电子减除。对如此找到的一套 GOI 也进行 eNorthern 研究并通过文献研究进行验证。

[0250] 这种联合的数据采集包括所有在公共网域中大约 13,000 个全长基因, 并且预言了这些基因具有可能的器官特异性的表达。

[0251] 所有其它的基因首先依靠特异性 PT-PCR 在正常组织中被评价。所有已经被证实 在非器官特异性的正常组织中表达的 GOI 必须被认作假阳性,并且不对其作进一步的研究。在大量的各种各样的肿瘤组织中对剩下来的那些 GIO 进行研究。下面叙述的抗原这里 被证实 在肿瘤细胞中是被激活的。

[0252] RNA 的提取、用多聚 d(T) 作引物的 cDNA 的制备和传统的 RT-PCR 分析

[0253] 通过使用总 RNA 异硫氰酸胍作为离液试剂 (chaotropic agent) 从天然组织材料 中提取总 RNA (Chomczynski&Sacchi, Anal. Biochem. 162 :156-9, 1987)。用酸性苯酚提取和 用异丙醇沉淀后,将所述的 RNA 溶于 DEPC 处理的水。

[0254] 依靠 Superscript II (Invitrogen), 根据厂商的信息, 在 20 μ l 反应混合液中 从 2-4 μ g 总 RNA 进行第一链 cDNA 的合成。使用的引物是 dT(18) 寡聚核苷酸。cDNA 的 完整性和质量通过 30 个循环的 PCR 扩增 p53 (正义 CGTGAGCGCTTCGAGATGTTCCG, 反义 CCTAACCAGCTGCCCAACTGTAG, 杂交温度 67 $^{\circ}$ C) 来检查。

[0255] 第一链 cDNA 的档案从许多正常组织和肿瘤实体制备而来。为了表达研究, 使用 GOI 特异性的引物 (见下) 和 1U HotStarTaq DNA 多聚酶 (Qiagen), 对 0.5 μ l 的这些 cDNA 在 30 μ l 反应混合液中进行扩增。每个反应混合液含有 0.3mM 脱氧核糖核苷酸、0.3 μ M 的每 个引物和 3 μ l 的 10 $\times$ 反应缓冲液。对引物进行选择要使其定位在两个不同的外显子, 并且 由于污染基因组 DNA 的干扰是假阳性结果的原因, 通过测试作为模板的非逆转录的 DNA 来 证实干扰的消除。95 $^{\circ}$ C 激活 HotStarTaq DNA 多聚酶 15 分钟后, 进行 35 个循环的 PCR (94 $^{\circ}$ C 1 分钟, 特定的杂交温度 1 分钟, 72 $^{\circ}$ C 2 分钟, 以及 72 $^{\circ}$ C 终末延长 6 分钟)。

[0256] 20 μ l 的此反应物被分装并且在溴化乙啶染色的琼脂糖胶上进行分析。

[0257] 下列引物用于在指定的杂交温度下相应抗原的表达分析

[0258] GPR35 (65 $^{\circ}$ C)

[0259] 正义 :5' -AGGTACATGAGCATCAGCCTG-3'

[0260] 反义 :5' -GCAGCAGTTGGCATCTGAGAG-3'

[0261] GUCY2C (62 $^{\circ}$ C)

[0262] 正义 :5-GCAATAGACATTGCCAAGATG-3

[0263] 反义 :5-AACGCTGTTGATTCTCCACAG-3

[0264] SCGB3A2 (66 $^{\circ}$ C)

[0265] 正义 :5' -CAGCCTTTGTAGTTACTCTGC-3'

[0266] 反义 :5' -TGTCACACCAAGTGTGATAGC-3'

[0267] Claudin18A2 (68 $^{\circ}$ C)

[0268] 正义 1 :5-GGTTCGTGGTTTCACTGATTGGGATTGC-3

[0269] 反义 1 :5-CGGCTTTGTAGTTGGTTTCTTCTGGTG-3

[0270] 正义 2 :5-TGTTTTCAACTACCAGGGGC-3

[0271] 反义 2 :5-TGTTGGCTTTGGCAGAGTCC-3

[0272] Claudin18A1 (64 $^{\circ}$ C)

[0273] 正义 :5-GAGGCAGAGTTCAGGCTTCACCGA-3

[0274] 反义 :5-TGTTGGCTTTGGCAGAGTCC-3

[0275] SLC13A1 (64 $^{\circ}$ C)

- [0276] 正义 :5' CAGATGGTTGTGAGGAGTCTG-3
[0277] 反义 :5' -CCAGCTTTAACCATGTCAATG-3
[0278] CLCA1 (62°C)
[0279] 正义 :5' -ACACGAATGGTAGATACAGTG-3
[0280] 反义 :5' ATACTTGTGAGCTGTTCCATG-3
[0281] FLJ21477 (68°C)
[0282] 正义 :5' -ACTGTTACCTTGCATGGACTG-3
[0283] 反义 :5' -CAATGAGAACACATGGACATG-3
[0284] FLJ20694 (64°C)
[0285] 正义 :5' -CCATGAAAGCTCCATGTCTA-3
[0286] 反义 :5' -AGAGATGGCACATATTCTGTC
[0287] Ebner (70°C)
[0288] 正义 :5' -ATCGGCTGAAGTCAAGCATCG-3
[0289] 反义 :5' -TGGTCAGTGAGGACTCAGCTG-3
[0290] Plunc (55°C)
[0291] 正义 :5' -TTTCTCTGCTTGATGCACTTG-3
[0292] 反义 :5' -GTGAGCACTGGGAAGCAGCTC-3
[0293] SLC26A9 (67°C)
[0294] 正义 :5' -GGCAAATGCTAGAGACGTGA-3
[0295] 反义 :5' -AGGTGTCCTTCAGCTGCCAAG-3
[0296] THC1005163 (60°C)
[0297] 正义 :5' -GTTAAGTGCTCTCTGGATTTG-3
[0298] LOC134288 (64°C)
[0299] 正义 :5' -ATCCTGATTGCTGTGTGCAAG-3
[0300] 反义 :5' -CTCTTCTAGCTGGTCAACATC-3
[0301] THC943866 (59°C)
[0302] 正义 :5' -CCAGCAACAACCTTACGTGGTC-3
[0303] 反义 :5' -CCTTTATTCACCCAATCACTC-3
[0304] FLJ21458 (62°C)
[0305] 正义 :5' -ATTCATGGTTCCAGCAGGGAC-3'
[0306] 反义 :5' -GGGAGACAAAGTCACGTACTC-3'
[0307] 随机六聚体作引物的 cDNA 的制备和定量实时 PCR

[0308] 一些基因的表达用实时 PCR 定量。PCR 产物使用 SYBR 绿作为内嵌 (intercalating) 报告染料进行检测。SYBR 绿的报告荧光在溶液中被抑制,并且只有在结合于双链 DNA 片段后染料有活性。SYBR 绿荧光的上升作为用 GOI 特异性引物的特异性扩增的结果被用来进行定量分析。靶基因的表达通过绝对或相对于要调查的组织中恒量表达的对照基因的表达来定量分析。表达在使用 $\Delta\Delta-C_t$ 法 (PE Biosystems, USA) 将样品用所谓看家基因的 18sRNA 标准化后进行测量。反应进行双复样并且测定三复样。QuantiTect SYBR 绿 PCR 试剂盒 (Qiagen, Hilden) 依照厂商的说明书使用。使用高容量 cDNA 档案库试剂盒 (PE

Biosystems,USA) 并依照厂商的说明书使用六聚体引物合成 cDNA。每 5 μ l 的稀释的 cDNA 被用于总体积为 25 μ l 的 PCR:正义引物 300nM,反义引物 300nM;初始变性 95 $^{\circ}$ C 15 分钟;95 $^{\circ}$ C 30 秒;退火 30 秒;72 $^{\circ}$ C 30 秒;40 个循环。使用的引物的序列在各自的实施例中有指出。

[0309] 克隆和序列分析

[0310] 通过传统的方法进行全长基因和基因片段的克隆。为了确定序列,使用校正多聚酶 pfu(Stratagene) 来扩增相应的抗原。为了将片段依照厂家的说明书克隆进 TOPO-TA 载体,在完成 PCR 之后,腺苷依靠 HotStarTap DNA 多聚酶结于扩增子 (amplicon) 的末端。测序由商业服务商进行。使用传统预言程序和逻辑来分析序列。

[0311] Western 印迹

[0312] 来自细胞培养(编码靶蛋白质的表达载体转染后靶基因的内源性表达或者靶蛋白质的合成)或含有靶蛋白质的组织样品用 1% SDS 溶液裂解。SDS 使存在于裂解液中的蛋白质变性。试验混合物的裂解液通过在 8-15% 变性聚丙烯酰胺凝胶(含 1% SDS)上的电泳根据大小被分离(SDS-聚丙烯酰胺凝胶电泳,SDS-PAGE),胶的浓度依赖于期望的蛋白质的大小。然后蛋白质通过半干电转印法(Biorad)被转移到硝酸纤维素膜上(Schleicher&Schüll),在膜上面期望的蛋白质能被检测到。为了这个目的,膜先被封闭(例如,用奶粉),然后与稀释度为 1:20-1:200 的特异性抗体孵育(依赖于抗体的特异性)60 分钟。清洗步骤后,膜与偶联于识别第一抗体的标志物(例如酶例如过氧化物酶或碱性磷酸酶)的第二抗体孵育。在进一步的清洗步骤之后,随后靶蛋白质显现出颜色或依靠酶反应在膜上(例如 ECL, Amersham Bioscience)化学发光反应显现出来。结果用合适的相机照相做成文件。

[0313] 蛋白质修饰的分析通常用 Western 印迹进行。通常有几个 kDa 大小的糖基化修饰导致靶更大总质量的蛋白质,这可以用 SDS-PAGE 分离。为了检测特异性 O- 和 N- 糖苷联结,来自组织或细胞的蛋白质裂解物在用 SDS 变性之前与 O- 或 N- 糖苷酶(依照它们各自厂商的说明书,例如 PNgase、内糖苷酶 F、内糖苷酶 H,Roche Diagnostics)孵育。这之后就是如上所述的 Western 印迹。因而,如果与糖苷酶孵育后靶蛋白质的大小减少了,就有可能检测到特异性的糖基化,并且这样,还可能分析修饰的肿瘤特异性。被糖基化的氨基酸的确切位置能用逻辑和预报程序被预报。

[0314] 免疫荧光法

[0315] 使用已经建立细胞系的细胞,所述细胞内源性合成靶蛋白质(RT-PCR 中 RNA 的检测或者 Western 印迹中蛋白质的检测)或者在 IF 之前已经被转染了质粒 DNA。将 DNA 转染进细胞系的各种各样的方法(例如电穿孔、基于脂质体的转染、钙磷酸盐沉淀)已经完善建立(例如 Lemoine et al.Methods Mol.Biol.1997;75:441-7)。免疫荧光法中转染的质粒编码未修饰的蛋白质或者不同的氨基酸标志物被偶联于靶蛋白质。最重要的标志物是例如有各种差异荧光形式的发荧光的“绿荧光蛋白”(GFP) 以及其高亲和力和特异性抗体可得到的 6-12 个氨基酸的短肽序列。合成靶蛋白质的细胞用多聚甲醛、皂甙或甲醇固定。然后如果需要通过与去污剂(例如 0.2% Triton X-100)孵育通透细胞。固定/通透之后,细胞与抗靶蛋白或抗偶联标志物之一的第一抗体孵育。清洗步骤之后,混合物与偶联于与第一抗体结合的荧光标志物(例如荧光蛋白、Texas Red、Dako)的第二抗体孵育。这样标记的细

胞用一层甘油覆盖,并且根据厂商的说明书借助于荧光显微镜进行分析。在这种情况下特异性的荧光发射通过依赖于采用的物质的特异性的激发而获得。所述分析通常允许可靠的靶蛋白的定位、抗体的质量以及靶蛋白质在双染色法中被证实,所述双染色法除染色靶蛋白质外也染色其定位已经在文献中有描述的偶联氨基酸标志物或其它标志物蛋白质。GFP 及其衍生物代表了能被直接激活并且自身会发荧光的一个特殊情况,所以不需要抗体来检测。

[0316] 免疫组织化学法

[0317] IHC 特异性地用于 (1) 能够估计肿瘤和正常组织中靶蛋白质的量, (2) 分析肿瘤和健康组织中有多少细胞合成靶基因, 和 / 或 (3) 定义可检测到靶蛋白质的组织中的细胞类型 (肿瘤、健康的细胞)。必须依赖于各自的抗体而使用不同的试验方案 (例如 “Diagnostic Immunohistochemistry, David J., MD Dabbs ISBN:0443065667” 或 “Microscopy, Immunohistochemistry, and Antigen Retrieval Methods: For Light and Electron Microscopy, ISBN:0306467704”)。

[0318] 对特异性组织样品的免疫组织化学法 (IHC) 用于检测相应组织中的蛋白质。这个方法的目的是鉴别功能性完整的组织聚集体中蛋白质的定位。IHC 特异性的用于 (1) 能够估计肿瘤和正常组织中靶蛋白质的量, (2) 分析肿瘤和健康组织中有多少细胞合成靶基因, 和 (3) 定义可检测到靶蛋白质的组织中的细胞类型 (肿瘤、健康的细胞)。或者, 靶基因的蛋白质的量可用数码相机和合适的软件 (例如 Tillvision, Till-photronics, Germany) 通过组织免疫荧光法来进行定量分析。这个技术很频繁地被发表, 并且因此染色和显微镜使用的细节, 例如在 “Diagnostic Immunohistochemistry”, David J., MD Dabbs ISBN: 0443065667 或 “Microscopy, Immunohistochemistry, and Antigen Retrieval Methods: For Light and Electron Microscopy” ISBN:0306467704 之中都能找到。应该注意的是因为抗体的特性, 为了得到有效的结果必须使用不同的实验方案 (下面描述了一个实施例)。

[0319] 平常地, 组织学定义的肿瘤组织以及作为参照的可比较的健康组织被用于 IHC 中。还可能使用作为阳性或阴性对照的细胞系, 其中靶基因的存在通过 RT-PCR 分析是已知的。通常必须包括背景对照。

[0320] 将固定的组织 (例如用含乙醛的物质、甲醛、多聚甲醛或在乙醇溶液中固定) 或者 1-10 μm 厚的快速冷冻的组织碎片敷于玻璃支持物上。石蜡包埋的样品用例如二甲苯去除石蜡。样品用 TBS-T 清洗并且用血清封闭。这之后与第一抗体孵育 (稀释度: 1 : 2 至 1 : 2000) 1-18 个小时, 通常使用亲和纯化的抗体。清洗步骤之后与偶联于碱性磷酸酶 (或者, 例如过氧化物酶) 并且抗第一抗体的第二抗体孵育大约 30-60 分钟。这之后是使用可被结合的酶转化的颜色底物的颜色反应 (参见例如 Shi 等人, J. Histochem. Cytochem. 39 : 741-748, 1991 ; Shin 等人, Lab. Invest. 64 : 693-702, 1991)。为了证明抗体的特异性, 反应通过免疫原的预先加入能被阻断。

[0321] 免疫

[0322] (也 见 Monoclonal Antibodies : A Practical Approach, Philip Shepherd, Christopher Dean isbn 0-19-963722-9 ; Antibodies : A Laboratory Manual, Harlow, David Lane 编 ISBN:0879693142 ; Using Antibodies : A Laboratory Manual : Portable Protocol NO. Edward Harlow, David Lane, Harlow 编 ISBN:0879695447)。

[0323] 用于制备抗体的方法简述如下,并且细节能在引用的文献中找到。首先,动物(例如兔)用第一次注射期望的靶蛋白质进行免疫。动物对免疫原的免疫反应能通过在规定期间内(在前面的免疫之后大约 2-4 周)第二或第三次免疫被增强。此外在各种规定期间后(4 周后第一次采血,然后大约每两周共 5 次采血)从动物采集血样,并且从那里获得免疫血清。

[0324] 其它的方法也可以利用,但动物通常用四个已经完善建立的方法中的一个来免疫。还可能用特异于靶蛋白质的肽、用完整的蛋白质或用能通过试验或预报程序鉴别的蛋白质的细胞外部分序列来免疫。

[0325] (1) 在第一种情况下,结合于 KLH(匙孔血蓝蛋白)的肽(长度:8-12 个氨基酸)通过标准化的体外方法进行合成,并且这些肽用于免疫。通常,进行 3 次免疫,每次免疫肽的浓度为 5-1000 μ g。还能请服务供应商以服务的形式来进行免疫。

[0326] (2) 或者,能用重组蛋白质进行免疫。为了这个目的,靶基因的克隆的 DNA 被克隆进表达载体,并且靶蛋白质用与特定厂商提供的条件(例如 Roche Diagnostics, Invitrogen, Clontech, Qiagen)类似的条件,例如无细胞的体外、在细菌中(例如大肠杆菌)、在酵母中(例如裂变酵母)、在昆虫细胞中或在哺乳动物细胞中进行合成。在这些系统之一中合成后,靶蛋白质被纯化,在这种情况下下的纯化通常通过标准化的色谱方法来进行。也可能在这点上,为了免疫使用具有辅助纯化的分子锚(例如 Hia 标签, Qiagen; FLAG 标签, Roche Diagnostics; Gst 融合蛋白质)的蛋白质。大量的试验方案可以在例如“Current Protocols in Molecular Biology”, John Wiley&Sons Ltd., Wiley Interscience 中找到。

[0327] (3) 如果有内源性合成期望蛋白质的细胞可被利用,这个细胞系也能用于生产特异性的抗血清。在这种情况下,通过 1-3 次注射免疫,每次用大约 $1-5 \times 10^7$ 细胞进行免疫。

[0328] (4) 免疫还能通过 DNA 的注射进行(DNA 免疫)。为了这个目的,靶基因最初被克隆进表达载体,所以靶序列在强的真核启动子的控制之下(例如 CMV 启动子)。随后,5-100 μ g 的 DNA 作为免疫原用“基因枪”被转移到生物体(例如鼠、兔)内有强血流的毛细血管区域。转移的 DNA 被动物的细胞摄取,靶基因被表达,并且动物最终产生对靶基因的免疫反应(Jung 等人, Mol Cells 12:41-49, 2001; Kasinrerker 等人, Hybrid Hybridomics 21:287-293, 2002)。

[0329] 多克隆血清或抗体的质量控制

[0330] 基于细胞培养及其后 Western 印迹的试验最适合用来证明特异性(各种变异在例如“Current Protocols in Protein Chemistry”, John Wiley&Sons Ltd., Wiley Interscience 中有描述)。为了证明,细胞被转染了在强的真核启动子(例如巨细胞病毒启动子)控制下的编码靶蛋白的 cDNA。用于将 DNA 转染进细胞系的各种各样的方法(例如电穿孔、基于脂质体的转染、磷酸钙沉淀)都已经完善建立(例如 Lemoine 等人, Methods Mol. Biol. 75:441-7, 1997)。或者也可能使用内源性表达靶基因的细胞系(通过靶基因特异性的 RT-PCR 来证明)。为了能够证明在接下来的 Western blot 中分析的抗体的特异性,作为对照,理想的情况下同源基因在试验中也被转染。

[0331] 在随后的 Western 印迹中,来自细胞培养或组织样品的可能含有靶基因的细胞用 1% SDS 溶液裂解,并且因此使蛋白质变性。裂解液根据大小在 8-15% 变性聚丙烯酰胺凝胶(含 1% SDS)上通过电泳分离(SDS 变性聚丙烯酰胺凝胶电泳, SDS-PAGE)。蛋白质然后通

过多数的转印方法（例如半干电转印 ;Biorad）中的一个转移到特异性膜上（例如硝酸纤维素, Schleicher&Schüll）。期望的蛋白质能在这个膜上显示出来。为了这个目的,膜先与识别靶蛋白质的抗体（稀释度大约 1 : 20-1 : 200,依赖于抗体的特异性）孵育 60 分钟。清洗步骤之后,膜与偶联于标志物（例如酶例如过氧化物或碱性磷酸酶）并且识别第一抗体的第二抗体孵育。然后有可能通过颜色或化学发光反应使靶蛋白质在膜上显现出来（例如 ECL, Amersham Bioscience）。对靶蛋白质有高热特异性的抗体在理想情况下只识别期望的蛋白质本身。

[0332] 各种方法被用来证实“在芯片上”的方法鉴别的靶蛋白质的膜定位。使用上述抗体一个重要的并且完善建立的方法就是免疫荧光法 (IF)。为此,合成靶蛋白质 (RT-PCR 中 RNA 的检测或 Western 印迹中蛋白质的检测) 或者已经转染了质粒 DNA 的已经建立的细胞系的细胞被使用。用于将 DNA 转染进细胞的各种各样的方法（例如电穿孔、基因脂质体的转染、磷酸钙沉淀）已经完善建立（例如 Lemoine 等人, Methods Mol. Biol. 75 :441-7, 1997）。转染进细胞的质粒在免疫荧光法中编码非修饰的蛋白质或者将各种氨基酸标志物偶联于靶蛋白质。主要的标志物是例如有其各种差异荧光形式的有荧光的“绿荧光蛋白”(GFP), 其高亲和力和特异性抗体可得到的 6-12 个氨基酸的短肽序列, 或者能通过其半胱氨酸特异性荧光物质而结合的短氨基酸序列 Cys-Cys-X-X-Cys-Cys (Invitrogen)。合成靶蛋白质的细胞用例如多聚甲醛或甲醇固定。然后如果需要, 通过与去污剂（例如 0.2% TritonX-100）孵育通透细胞。细胞再与抗靶蛋白质的或抗偶联的标志物之一的第一抗体孵育。清洗步骤后, 混合物与偶联于荧光标志物（例如荧光素, Texas Red, Dako）并且结合于第一抗体的第二抗体孵育。这样标记的细胞再盖上一层甘油, 并且根据厂商的说明书借助于荧光显微镜进行分析。特异性的荧光发射在这种情况下通过依赖于采用的物质的特异性激发而获得。所述分析通常允许可靠的靶蛋白质的定位、抗体的质量以及靶蛋白质在双染色法中被证实, 所述双染色法除染色靶蛋白质外也染色其定位已经在文献中有描述的偶联氨基酸标志物或其它标志物蛋白质。GFP 及其衍生物代表了能被直接激活并且自身会发荧光的一个特殊情况。能过去去污剂的使用来控制的膜通透性允许在免疫荧光法中证明一个免疫原性抗原是位于细胞内还是细胞外。所选蛋白质的预报因而有试验的支持。另外一个可能性是依靠流式细胞仪检测细胞外结构域。为了这个目的, 细胞在非通透化的条件下被固定（例如, 用 PBS/ 叠氮钠 /0.2% 胎牛血清 /5mM EDTA）并且依照厂商的说明书用流式细胞仪进行分析。只有细胞外的抗原决定簇才能通过这个分析方法被抗体识别。免疫荧光的不同就是有可能使用例如碘化丙锭或台盼蓝区别死或活细胞, 并且因而避免假阳性结果。

[0333] 亲和纯化

[0334] 多克隆血清的纯化在肽抗体的情况下完全或者在抗重组蛋白质的情况下部分由合同公司提供服务来进行。为了这个目的, 在两种情况下, 合适的肽或重组蛋白质共价结合到基质上, 并且后者在偶联后用天然缓冲液 (native buffer) (PBS : 磷酸盐缓冲液) 平衡, 然后与粗制血清孵育。PBS 进一步清洗步骤后, 抗体用 100mM 甘氨酸 pH 2.7 洗脱, 并且立即用 2M TRIS, pH 8 中和洗脱液。这样纯化的抗体然后能用于通过 Western 印迹和免疫荧光法的靶蛋白质的特异性检测。

[0335] EGFP 转染体的制备

[0336] 为了免疫荧光显微镜检测异源性表达的肿瘤相关抗原, 抗原的完全 ORF 被克隆进

pEGFP-C1 和 pEGFP-N3 载体 (Clontech) 中。培养在玻片上的 CHO 和 NIH3T3 细胞用 Eugene 转染试剂 (Roche) 依照厂商的说明书被转染了合适的质粒构建体, 并且 12-24 小时后, 用免疫荧光显微镜进行分析。

[0337] 实施例 1 :GPR35 作为诊断和治疗癌症的靶点的鉴别

[0338] GPR35 (SEQ ID NO :1) 及其翻译产物 (SEQ ID NO :9) 被描述为推测的 G 蛋白偶联受体。其序列发表在基因库登记号 (accession No.) 为 AF089087 之下。该转录体编码一个 309 个氨基酸分子量为 34kDa 的蛋白质。预报 GPR35 属于有 7 个跨膜结构域的 G 蛋白质偶联受体超家族 (O'Dowd 等人, Genomics 47 :310-13, 1998)。为了证实预报的 GPR35 细胞内定位, 蛋白质与作为报告分子的 eGFP 相融合, 并且在合适的质粒转染后异源性地在 293 细胞中表达。然后荧光显微镜下分析其定位。根据本发明证实 GPR35 是整合跨膜分子 (图 17)。目前对人类 GPR35 的调查 (参见, 尤其是 Horikawa Y, Oda N, Cox NJ, Li X, Orho-Meland M, Hara M, Hinokio Y, Lindner TH, Mashima H, Schwarz PE, del Bosque-Plata L, Horikawa Y, Oda Y, Yoshiuchi I, Colilla S, Polonsky KS, Wei S, Concannon P, Iwasaki N, Schulze J, Baier LJ, Bogardus C, Groop L, Boerwinkle E, Hanis CL, Bell GI Nat Genet. 2000 Oct ;26 (2) :163-75) 提示 GPR35 在许多健康组织中是被激活的。基因的阅读框包含一个外显子。根据本发明, 为了分析以扩增结肠中和结肠癌 (13/26) 中的 cDNA, GPR35 的基因特异性引物对 (SEQ ID NO :20, 21) 被用于 RT-PCR。相反, 在其它的正常组织中未检测到显著表达。因为 GPR35 由一个外显子组成的特定的事实, 基因组 DNA 杂质不能用跨内含子的引物来检测。为了排除 RNA 样品中基因组的污染, 所以所有的 RNA 用 DNase 处理。使用无 DNA 的 RNA, GPR35 转录物根据本发明只有在结肠、在直肠、在睾丸和在结肠癌中检测到的。

[0339] 表 1. GPR35 在正常组织中的表达

[0340]

正常组织	表达
脑	-
小脑	-
心肌	-
骨骼肌	-
直肠	++
胃	-
结肠	++
胰腺	-

肾	-
睾丸	-
胸腺	-
乳腺	-
卵巢	-
子宫	未测定
皮肤	-
肺	-
甲状腺	-
淋巴结	-
脾	-
外周血单核细胞	-
肾上腺	-
食道	-
小肠	+
前列腺	-

[0341]

[0342] GPR35 转录物在正常结肠组织和在结肠癌活检组织中选择性和高水平的表达 (图 1) 以前并不为人所知, 并且根据本发明能用于分子诊断方法例如 RT-PCR, 以用来检测血清和骨髓中播散的肿瘤细胞以及检测其它组织中的转移灶。用特异性引物 (SEQ ID NO :88 和 89) 的定量 RT-PCR 也证实 GPR35 是高选择性的肠道特异性分化抗原, 在肠道肿瘤和肠道肿瘤转移灶中也含有 GPR35。与正常肠道相比, 在一些肠道肿瘤中, 它实际上过度表达一个对数数量级 (图 18)。抗体通过免疫兔生产, 用于检测 GPR35 蛋白质。下列肽被用于生产 (propagate) 这些抗体:

[0343] SEQ ID NO :90GSSDLTWPPAIKLG (氨基酸 9-23)

[0344] SEQ ID NO :91 :DRYVAVRHPLRARGLR (氨基酸 112-127)

[0345] SEQ ID NO :92 :VAPRAKAHKSQDSL(羧基端)

[0346] SEQ ID NO :93CFRSTRHNFNSMR(细胞外,结构域 2)

[0347] 例如在 Western 印迹中用这些抗体染色证实了肿瘤中的表达。GPR 35 的所有 4 个细胞外结构域(预报的细胞外结构域的位置在序列 SEQ ID NO :9 之中氨基酸 1-22(SEQ ID NO :94);氨基酸 81-94(SEQ ID NO :95);氨基酸 156-176(SEQ ID NO :96);氨基酸 280-309(SEQ ID NO :97))根据本发明可被用作单克隆抗体的靶结构。这些抗体与肿瘤细胞的细胞表面特异性结合,并且既可被用于诊断方法也可被用于治疗方法。GPR35 在肿瘤中的过度表达为这样的用途提供了额外的支持。另外,编码蛋白质的序列根据本发明能被用作疫苗(RNA、DNA、肽、蛋白质),用来诱导肿瘤特异性的免疫反应(T 细胞和 B 细胞介导的免疫反应)。另外,还很惊人的发现在一般已知的起始密码子 5'端前还有一个起始密码子,并且表达一个氨基端延长的蛋白质。

[0348] 因而根据本发明发现 GPR35,一个以前被描述为广泛表达的蛋白质,是肿瘤相关过度表达的,选择性地在胃肠道肿瘤中过度表达,尤其是结肠的肿瘤。GPR35 所以适合作为尤其是用于诊断和治疗这些肿瘤的分子靶结构。目前对人类 GPR35 的调查,参见例如 Horikawa Y, Oda N, Cox NJ, Li X, Orho-Melander M, Hara M, Hinokio Y, Lindner TH, Mashima H, Schwarz PE, del Bosque-Plata L, Horikawa Y, Oda Y, Yoshiuchi I, Colilla S, Polonsky KS, Wei S, Concannon P, Iwasaki N, Schulze J, Baier LJ, Bogardus C, Groop L, Boerwinkle E, Hanis CL, Bell GI Nat Genet. 2000 Oct ;26(2):163-75 提示 GPR35 在许多健康组织中是被激活的。相反,根据本发明的调查显示 GPR35 惊人地在多数正常组织中并不能明显地检测到,并且与此相反,在原发和转移性的结肠肿瘤被高度激活。另外,除了所述的 GPR35 序列,根据本发明发现了一个使用另外一个起始密码子的全新的翻译变体(SEQ ID NO :10)。

[0349] GPR35 是 G 偶联受体(GPCR)组的一员,GPCR 是一个非常大的蛋白质家族,其结构和功能已经被非常详尽地调查。GPCR 显著地适合作为靶结构来开发药物活性物质,因为为此必须的方法(例如受体表达、纯化、配体筛选、诱变处理、功能性抑制、激动剂配体和拮抗剂配体的选择、配体的放射性标记)都已经发展得非常好并且有详细的描述,参见例如,“G Protein-Coupled Receptors”, Tatsuya Haga, Gabriel Berstein 和 Gabriel Bernstein ISBN :0849333849 和“Identification and Expression of G-Protein Coupled Receptors Receptor Biochemistry and Methodology”, Kevin R. Lynch ASIN :0471183105。根据本发明对 GPR35 在多数健康组织中不能被检测到,但是在细胞表面有肿瘤相关表达的认识,使其能够被用作例如药物活性配体的肿瘤相关靶结构,尤其是与例如放射活性分子相结合被用作药物物质。在特定的实施方案中,有可能使用结合于 GPR35 的放射标记的配体用来检测肿瘤细胞或者用来体内治疗结肠肿瘤。

[0350] 实施例 2 :GUCY2C 在肝和卵巢肿瘤中的鉴别以及新型 GUCY2C 剪接变体作为诊断和治疗癌症的靶点的鉴别

[0351] 鸟苷酸环化酶 2C(SEQ ID NO :2;翻译产物:SEQ ID NO :11)-一种 I 型跨膜蛋白-属于利钠尿肽受体家族。序列发表在基因库注册号 NM_004963 之下。肽类鸟苷蛋白和尿鸟苷蛋白(uroguanylin)或者热稳定性肠毒素(STa)的结合增加细胞内 cGMP 的浓度,因而诱导细胞内的信号转导过程。

[0352] 最近的调查指出 GUCY2C 的表达也延伸到肠外区域,例如胃和食道的原发性和转移性腺癌 (Park 等人, *Cancer Epidemiol Biomarkers Prev.* 11 :739-44, 2002)。在正常和转化的肠道组织中都发现的 GUCYC 剪接变异体包含了外显子 1 中一个 142bp 的缺失,因而阻止了 GUCY2C 样产物的翻译 (Pearlman 等人, *Dig. Dis. Sci.* 45 :298-05, 2000)。至今描述的唯一剪接变异体导致没有翻译产物。

[0353] 根据本发明目的是鉴别能被用于诊断和被用于治疗 GUCY2C 的肿瘤相关剪接变异体。

[0354] 用 GUCY2C 特异性引物对 (SEQ ID NO :22, 23, 98, 99) 的 RT-PCR 调查显示 GUCY2C 转录物在正常结肠和胃的显著表达,以及在肝、睾丸、卵巢、胸腺、脾、脑和肺中的微弱表达 (表 2, 图 19)。在结肠和胃中的表达比所有其它正常组织高至少 50 倍。显著的 GUCY2C 转录物水平在结肠癌和胃癌中能检测到 (表 2)。这些结果通过定量 PCR 分析被阐述,并且显示在正常结肠、回肠和几乎所有被调查的结肠癌样品中有显著的 GUCY2C 表达 (图 2, 19B)。整块的过度表达在一些结肠癌样品中可检测到。另外,在 7/10 胃肿瘤中发现有表达。我们还惊讶地发现在很多其它以前未描述的肿瘤中,其中有卵巢、乳腺、肝和前列腺的肿瘤中,所述基因被激活 (图 19B, 表 2)。

[0355] 表 2 :GUC2C 在正常和肿瘤组织中的表达

[0356]

正常组织	表达
脑	+
小脑	
心肌	
骨骼肌	-
心肌	
胃	+++
结肠	+++
胰腺	-
肾	-
肝	+
睾丸	++
胸腺	+
乳腺	-
卵巢	+
子宫	+
皮肤	
肺	+
甲状腺	
淋巴结	-
脾	+
外周血单核细胞	-
前列腺	-

肿瘤类型	表达
结肠癌	+++
胰腺癌	-
食道癌	-
胃癌	+++
支气管癌	-
乳腺癌	-+
卵巢癌	+
子宫内膜癌	
耳鼻喉肿瘤	
肾细胞癌	
前列腺癌	+
肝癌	+

[0357] 下列的引物对被用于检测结肠组织和结肠癌组织中的剪接变异体：

[0358] GUCY2C-118s/GUCY2C-498 如 (SEQ ID NO :24, 29) ；

[0359] GUCY2C-621s/GUCY2C-1140 如 (SEQ ID NO :25,30) ;

[0360] GUCY2C-1450s/GUCY2C-1790 如 (SEQ ID NO :26,31) ;

[0361] GUCY2C-1993s/GUCY2C-2366 如 (SEQ ID NO :27,32) ;

[0362] GUCY2C-2717s/GUCY2C-3200 如 (SEQ ID NO :28,33) ;

[0363] GUCY2C-118s/GUCY2C-1140 如 (SEQ ID NO :24,30) ;

[0364] GUCY2C-621s/GUCY2C-1790 如 (SEQ ID NO :25,31) ;

[0365] GUCY2C-1450s/GUCY2C-2366 如 (SEQ ID NO :26,32) ;

[0366] GUCY2C-1993s/GUCY2C-3200 如 (SEQ ID NO :27,33) 。

[0367] 在调查结肠癌组织中剪接变异体的时候,三个以前未知的形式根据本发明被鉴别出来。

[0368] a) 外显子 3 (SEQ ID NO :3) 的缺失,导致 GUCY2C 的变异体只有 111 个氨基酸长,并且在其 111 位置的天冬氨酸被替换为脯氨酸。

[0369] b) 外显子 6 (SEQ ID NO :4) 的缺失,产生一个 258 个氨基酸长的表达产物。这将产生包含 13 个氨基酸的羧基端新抗原决定簇。

[0370] c) 一个变异体,其中位置在 1606-1614 的核苷酸以及相应的氨基酸 L(536)、L(537) 和 Q(538) 被缺失 (SEQ ID NO :5) 。

[0371] 根据本发明分别在外显子 3 和外显子 6 (SEQ ID NO :3,4) 缺失的剪接变异体尤其由于翻译产物没有跨膜结构域而有区别。在外显子 6 缺失的情况下,得到一个 13 个氨基酸的羧基端新抗原决定簇,它与任何以前已知的蛋白质无同源性。这个新抗原决定簇因而注定是免疫治疗的靶结构。在位置 1606-1614 有碱基缺失 (SEQ ID NO :5) 的本发明的剪接变异体及其翻译产物 (SEQ ID NO :14) 同样地包含新抗原决定簇。用于检测 GUCY2C 蛋白质的抗体通过免疫兔来生产。下列肽用于生产这些抗体 :

[0372] SEQ ID NO :100 :HNGSYEISVLMGNS (氨基酸 31-45)

[0373] SEQ ID NO :101 :NLPTPTVENQQRLA (氨基酸 1009-1023)

[0374] 这样的抗体主要能用于诊断和治疗的目的。

[0375] 尤其是 GUCY2C 细胞外结构域 (来自 SEQ ID NO :11 :氨基酸 454-1073 (SEQ ID NO :102) 的序列的预报的细胞外结构域之位置) 根据本发明能被用作单克隆抗体的靶结构。然而,结构预报有点不明确并且未被试验核实,所以其它的膜定向也是可以想象的。在这种情况下,氨基酸 1-431 将在细胞外并且适合作为单克隆抗体的出发点。这些抗体与肿瘤细胞的细胞表面特异性结合,并且能被用于诊断和用于治疗的方法。GUCY2C 尤其在结肠肿瘤中的过度表达,给这样的用途提供了额外的支持。编码蛋白质的序列还能根据本发明被用作疫苗 (RNA、DNA、肽、蛋白质),用来诱导肿瘤特异性免疫反应 (T 细胞和 B 细胞介导的免疫反应) 。

[0376] 还有可能依照 GUCY2C 分子的细胞功能根据本发明开发可调节肿瘤细胞上酶的功能的物质,尤其是小分子物质。酶反应的产物, cGMP, 是有各种各样功能的已知的细胞信号分子 (Tremblay 等人 .Mol Cell Biochem 230,31) 。

[0377] 实施例 3 :SCGB3A2 作为诊断和治疗癌症的靶点的鉴别

[0378] SCGB3A2 (SEQ ID NO :6) (翻译产物 :SEQ ID NO :15) 属于分泌球蛋白 (secretoglobin) 基因家族。这个序列发表在基因库注册号 NM_054023 之下。

SCGB3A2(UGPR1) 是同源二聚体的分泌蛋白质,大小为 17kDa,它只在肺中和气门中表达(Niimi 等人,Am J Hum Genet 70:718-25,2002)。用引物对(SEQ ID NO:37,38)的 RT-PCR 调查证实在正常肺组织中的选择性表达。例如编码表面活性蛋白质的肺和气管特异性基因,在恶性肿瘤去分化过程中高度下调并且在肺肿瘤中通常检测不到。惊人地发现 SCGB3A2 在原发性和转移性肺肿瘤中有活性。根据本发明的调查显示 SCGB3A2 在支气管癌中有很强地并且频繁地表达(图 4)。除了肺和支气管外,被检测的所有其它 23 个正常组织,未显示表达(参见图 20)。

[0379] 这另外在特异性定量 RT-PCR(SEQ ID NO:103,104)中被证实(图 20),它另外显示至少在超过 50%的支气管癌中过度表达至少一个数量级。

[0380] SCGB3A2 在正常肺组织中和肺癌活检组织中选择性并且高水平的表达根据本发明能被用于分子诊断方法,例如 RT-PCR,用来检测在血和骨髓、唾液、支气管吸出物或灌洗物中散布的肿瘤细胞以及用来检测其它组织例如局部淋巴结中的转移灶。在健康肺中,SCGB3A2 由专门的细胞只分泌到支气管中。相应地,不期望 SCGB3A2 蛋白质将在健康个体呼吸道外的体液中检测到。相反,尤其是转移的肿瘤细胞直接分泌它们的蛋白质产物到血流中。本发明的一个方面所以涉及通过特异性抗体试验检测患者的血清或血浆中 SCGB3A2 产物,作为肺肿瘤的诊断性发现。

[0381] 用于检测 SCGB3A2 蛋白质的抗体通过免疫兔来生产。下列肽被用于生产这些抗体:

[0382] SEQ ID NO:105:LINKVPLPVDKLAPL

[0383] SEQ ID NO:106:SEAVKKLLEALSHLV

[0384] SCGB3A2 特异性的反应在免疫荧光法中可检测到(图 21)。对于分泌的蛋白如期望的,SCGB3A2 在细胞的分布被指定在内质网和分泌颗粒(图 21A)。为了检查特异性,细胞被并列地转染了合成 SCGB3A2-GFP 融合蛋白质的质粒。在这个情况下,蛋白质的监测通过自发荧光的 GFP(绿荧光蛋白)来进行(图 21B)。两个荧光图的叠加清楚地显示免疫血清特异性地识别 SCGB3A2 蛋白质(图 21C)。

[0385] 这样的抗体根据本发明能被用于例如用于诊断和治疗目的的免疫试验中。

[0386] 实施例 4:claudin-18A1 和 claudin18A2 剪接变异体作为诊断和治疗癌症的靶点的鉴别

[0387] Claudin-18 基因编码具有 4 个跨膜结构与和细胞内氨基端和羧基端的表面膜分子。Niimi 和同事(Mol. Cell. Biol. 21:7380-90,2001)描述了鼠和人类 claudin-18 的两个剪接变异体,所述变异体在已经被描述为分别在肺组织中(claudin-18A1)和胃组织中(claudin-18A2)有选择性的表达。这些变异体在氨基端不同(图 22)。

[0388] 根据本发明调查了剪接变异体 claudin-18A2(SEQ ID NO:7)和 claudin-18A1(SEQ ID NO:117)以及他们各自的翻译产物(SEQ ID NO:16 and 118)被用作肿瘤的标志物或肿瘤治疗的靶结构到何种程度。能区别两个变异体的定量 PCR 通过选择 A1 特异性(SEQ ID NO:109&110)和 A2 特异性(SEQ ID NO:107&108)的引物对被建立。A2 剪接变异体还用第二对引物以传统 PCR(SEQ ID NO:39&40)进行测试。A1 变异体被描述为只在正常肺中有活性。然而,根据本发明惊人地发现 A1 变异体在胃粘膜中也有活性。胃和肺是唯一显示有显著激活的正常组织。所有其它的正常组织 claudin-A1 阴性。在调查肿瘤的时候,惊人地发现

claudin-A1 在大量的肿瘤组织中被高度激活。在胃肿瘤、肺肿瘤、胰腺癌、食道癌（图 23）、耳鼻喉肿瘤和前列腺癌中发现了特别强的表达。Claudin-A1 的表达水平在耳鼻喉、前列腺、胰腺和食道肿瘤中比相应的正常组织中的水平高 100-10000 倍。用于调查 claudin-A2 剪接变异体的寡聚核苷酸特异性地使这个转录物能被扩增（SEQ ID NO :39&40 和 107&108）。调查揭示 A2 变异剪接体在被调查的除胃粘膜外的 20 多个正常组织中无表达，并且还在睾丸组织中有少量表达。我们发现 A2 变异体还象 A1 变异体一样，在很多肿瘤中被激活（经由图 24 中的实施例有描绘）。这些包括胃肿瘤（8/10）、胰腺肿瘤（6/6）、食道癌（5/10）和肝癌。尽管 claudin-18A2 的激活在健康肺中未检测到，却惊人地发现一些肺肿瘤表达 A2.1 剪接变异体。

[0389] 表 3A. claudin-18A2 在正常和肿瘤组织中的表达

[0390]

正常组织	表达	肿瘤类型	表达
脑	-	结肠癌	-
小脑	-	胰腺癌	++
心肌	-	食道癌	++
骨骼肌	-	胃癌	+++
子宫内膜	-	支气管癌	++
胃	+++	乳腺癌	-
结肠	-	卵巢癌	-
胰腺	-	子宫内膜癌	n. i.
肾	-	耳鼻喉肿瘤	++
肝	-	肾细胞癌	-
睾丸	+	前列腺癌	-
胸腺	-		
乳腺	-		
卵巢	-		
子宫	-		
皮肤	-		
肺	-		
甲状腺	-		
淋巴结	-		
脾	-		
外周血单核细胞	-		
食管	-		

[0391] 表 3B. claudin-18A1 在正常和肿瘤组织中的表达

[0392]

正常组织	表达	肿瘤类型	表达
脑	-	结肠癌	-
小脑	-	胰腺癌	++
心肌	-	食道癌	++
骨骼肌	-	胃癌	+++
子宫内膜	-	支气管癌	++
胃	+++	乳腺癌	+
结肠	-	卵巢癌	n. i.
胰腺	-	子宫内膜癌	n. i.
肾	-	耳鼻喉肿瘤	++
肝	-	肾细胞癌	-
睾丸	+	前列腺癌	++
胸腺	-		
乳腺	-		
卵巢	-		
子宫	-		
皮肤	-		
肺	+++		
甲状腺	-		
淋巴结	-		
脾	-		
外周血单核细胞	-		
食管	-		

[0393] 作为独立的对照调查的传统 PCR 也证实了定量 PCR 的结果。用于传统 PCR 的寡聚核苷酸 (SEQ ID NO :39,40) 允许 A2 剪接变异体的特异性扩增。根据本发明显示 8/10 的胃癌以及一半的被测试的胰腺癌显示了这个剪接变异体的强表达 (图 5)。相反,表达用传统 PCR 在其它组织中未检测到。尤其是,在肺、肝、血、淋巴结、乳腺组织和肾组织没有表达 (表 3)。

[0394] 剪接变异体因而根据本发明代表了高度特异性于上胃肠道肿瘤以及肺肿瘤、耳鼻喉肿瘤、前列腺癌和它们的转移灶的分子标志物。这些分子标志物根据本发明能被用于检测肿瘤细胞。根据本发明用所述的寡聚核苷酸 (SEQ ID NO :39,40,107-110) 检测肿瘤是可能的。尤其合适的寡聚核苷酸是至少其中一个在严格条件下结合于转录物的片段的引物对,所述片段 180 个碱基对长并且特异于一个 (SEQ ID NO :8) 或另一个剪接变异体 (SEQ ID NO :119)。

[0395] 为了在蛋白质水平证实这些数据, claudin 特异性抗体和免疫血清通过免疫动

物来产生。Claudin-18 的浆膜定位和蛋白质的拓扑结构通过用生物信息工具 (TMHMM, TMPRED) 对跨膜结构域的分析以及表达附有增强的 GFP 标签的 claudin-18 融合蛋白质的细胞之免疫荧光调查而得到证实。Claudin-18 具有两个细胞外结构域。两个剪接变异体中氨基端细胞外结构的序列不同 (对于 A1SEQ ID NO:111 以及对于 A2SEQ ID NO:112)。两个变异体羧基端细胞外结构域一样 (SEQ ID NO:137)。至今,未有结合于 claudin-18 细胞外结构域的抗体的描述。根据本发明,定位于细胞外并特异于变异体 A1 或 A2 的或出现在两个变异体中的肽抗原决定簇被选择用于免疫。Claudin-18 的两个变异体没有传统的糖基化基序,所以并不期望有蛋白质的糖基化。然而,在抗原决定簇的选择中要考虑到包含天冬氨酸、丝氨酸、苏氨酸的抗原决定簇即便没有传统的糖基化位点在罕见的情况下有可能被糖基化。抗原决定簇的糖基化可以阻止特异于该抗原决定簇的抗体的结合。特别地,根据本发明选择抗原决定簇,所以由此产生的抗体允许抗原的糖基化状态可被区别出来。下列肽特别地被选择用于生产为了免疫的抗体:

[0396] SEQ ID NO:17:DQWSTQDLYN(氨基端细胞外结构域,A2 特异性的,结合不依赖于糖基化)

[0397] SEQ ID NO:18:NNPVTAVFNYQ(氨基端细胞外结构域,A2 特异性的,主要与非糖基化形式结合,N37)

[0398] SEQ ID NO:113:STQDLNNPVTAVF(氨基端细胞外结构域,A2 特异性的,只与非糖基化形式结合,N37)

[0399] SEQ ID NO:114:DMWSTQDLYDNP(氨基端细胞外结构域,A1 特异性的)

[0400] SEQ ID NO:115:CRPYFTILGLPA(氨基端细胞外结构域,主要特异于 A1 的)

[0401] SEQ ID NO:116:TNFWMSTANMYTG(羧基端细胞外结构域,识别 A1 和 A2)。

[0402] 经由实施例显示了用 SEQ ID NO:17 免疫产生的 A2 特异性抗体的数据。特异性抗体在不同的固定条件能被用于免疫荧光调查。用 RT-PCR 阳性和阴性细胞系的比较染色,很容易检测到的量的相应的蛋白质在胃癌细胞系中能被特异性地检测到,代表阳性 (图 25)。内源性蛋白质是定位于膜的并且在膜上形成相对大的局灶聚集体 (focal aggregates)。这个抗体又可用于 Western 印迹中蛋白质的检测。如期望地,蛋白质只在胃中而不在其它正常组织中可检测到,甚至在肺中也没有 (图 29)。来自患者的胃肿瘤和邻近的正常胃组织的比较染色惊人地揭示 claudin-18A2 在所有可检测到此种蛋白质的胃肿瘤中有较小的质量 (图 30 左)。根据本发明在一系列试验中发现当正常胃组织的裂解物用去糖基试剂 PNGase F 处理时,在该水平也出现一个条带 (图 30 右)。尽管只有 A2 变异体的糖基化形式在所有正常胃组织中可检测到,A2 在多于 60% 被调查的胃癌中可检测到,尤其是只有去糖基化形式的 A2。虽然 claudin-18 的 A2 变异体甚至在蛋白质水平也没有在正常肺中检测到,但是和以前定量 RT-PCR 发现的一样,它在支气管癌中被发现。仍然只有去糖基化的变异体存在 (图 31)。识别 claudin-18-A2 剪接变异体的细胞外结构域的抗体根据本发明本已经被生产。另外,选择形识别 claudin-18-A1 剪接变异体的氨基端结构域的抗体 (图 28) 以及结合于两个变异体羧基端细胞外结构域的区域 (图 27) 已经被生产。有可能根据本发明为了诊断的目的例如免疫组织学 (图 32),也为了上面解释的治疗目的使用这样的抗体。深一层重要的方面涉及 claudin-18 差异糖基化的结构域。只与非糖基化抗原决定簇结合的抗体根据本发明被生产。Claudin-18 本身是对胃组织 (A2) 和对肺和胃 (A1) 高选择性的

分化抗原。因为它明显地受肿瘤内糖基化机器 (machinery) 改变的影响,所以 A2 的一个特定的去糖基化的变异体在肿瘤内产生。这能被用于诊断和治疗。免疫血清例如这里描述的一个 (抗 SEQ ID NO:17 的肽) 能被用于诊断例如用在 Western 印迹中。如例如用 SEQ ID NO:113 的肽免疫得到的完全不能结合于糖基化的抗原决定簇的抗体 (图 26), 能在结合上区别肿瘤组织和正常组织。因为它们的高度选择性,尤其有可能治疗性地采用这样的抗体。产生的抗体还能直接用于产生嵌合的或人源化的重组抗体。这还能直接用得自于兔的抗体来进行 (关于这个,见 J Biol Chem. 2000 May 5;275(18):13668-76, Rader C, Ritter G, Nathan S, Elia M, Gout I, Jungbluth AA, Cohen LS, Welt S, Old LJ, Barbas CF 3rd. "The rabbit antibody repertoire as a novel source for the generation of therapeutic human antibodies")。为了这个目的,来自接受免疫的动物的淋巴细胞被保存。氨基酸 1-47 (SEQ ID NO:19 和 120) 也代表了用于免疫治疗方法例如疫苗和抗原特异性 T 淋巴细胞的过继性转移的抗原决定簇。

[0403] 实施例 5:SLC13A1 作为诊断和治疗癌症的靶点的鉴别

[0404] SLC13A1 属于钠硫酸盐共转运蛋白家族。与此基因的小鼠同源物相反,人类基因选择性地表达于肾 (Lee 等人, Genomics 70:354-63)。SLC13A1 编码一个 595 个氨基酸的蛋白质并且包含 13 个推测的跨膜结构域。选择性的剪接导致 4 个不同的转录物 (SEQ ID NO:41-44) 及其相应的翻译产物 (SEQ ID NO:45-48)。对 SLC13A1 是否能用作肾肿瘤的标志物进行了调查。能特异性扩增 SLC13A1 的寡聚核苷酸 (SEQ ID NO:49,50) 被用于此目的。

[0405] 表 4. SLC13A1 在正常和肿瘤组织中的表达

[0406]

正常组织	表达	肿瘤类型	表达
脑	-	结肠癌	未测定
小脑	未测定	胰腺癌	未测定
心肌	未测定	食道癌	未测定
骨骼肌	未测定	胃癌	未测定
子宫内膜	-	支气管癌	未测定
胃	-	乳腺癌	未测定
结肠	-	卵巢癌	未测定
胰腺	未测定	子宫内膜癌	未测定
肾	+++	耳鼻喉肿瘤	未测定
肝	-	肾细胞癌	+++
睾丸	+	前列腺癌	未测定
胸腺	-		
乳腺	-		
卵巢	-		
子宫	未测定		
皮肤	未测定		
肺	-		
甲状腺	-		
淋巴结	-		
脾	-		
外周血单核细胞	-		
乙状结肠	-		
食管	-		

[0407] 用 SLC13A1 特异性引物对的 RT-PCR 调查证实在肾中实际上选择性的表达,并且根据本发明显示了实际上在所有 (7/8) 被调查的肾细胞癌活检组织中的高表达 (表 4, 图 6)。用特异性引物 (SEQ ID NO:121,122) 的定量 PCR 也证实了这些数据 (图 34)。微弱的信号在下列正常组织中可检测到:结肠、胃、睾丸、乳腺、肝和脑。然而,在肾癌中的表达至少比在所有其它正常组织中高至少 100 倍。

[0408] 为了分析 SLC13A1 在细胞内的亚细胞定位,将蛋白质与作为报告分子的 eGFP 融合,并且在合适质粒的转染后异源地表达于 293 细胞中。其定位在荧光显微镜下进行分析。我们的数据令人印象深刻地证实了 SLC13A1 是一个整合跨膜分子 (图 35)。

[0409] 用于检测 SLC13A1 蛋白质的抗体通过免疫兔而产生。SEQ ID NO:123 和 124 的肽被用于生产这些抗体。这样的抗体主要被用于诊断和治疗目的。

[0410] SLC13A1 蛋白质具有 13 个跨膜结构域和 7 个细胞外区域。SLC13A1 的这些细胞外

结构域根据本发明尤其能被用作单克隆抗体的靶结构。SLC13A1 作为通道蛋白质参与离子的转运。SLC13A1 的细胞外结构域在健康肾中极性地被定向而朝着尿道的方向（腔膜方向地）。然而，用于治疗的高分子量的单克隆抗体不被分泌到尿道内，所以在健康肾中抗体不与 SCL13A1 结合。相反，肿瘤细胞内 SLC13A1 的极性被取消，通过血流直接就可得到为抗体靶点的蛋白质。SCL13A1 在肾细胞癌中的显著表达和高发生率使得该蛋白质根据本发明成为高度引人注目的诊断和治疗标志物。这包括根据本发明依靠 RT-PCR 在血清、骨髓、尿中散布的肿瘤细胞的检测以及其他器官中转移灶的检测。还有可能根据本发明依靠单克隆抗体将 SLC13A1 的细胞外结构域用作免疫诊断和治疗的靶结构。SLC13A1 根据本发明还能被用作疫苗（RNA、DNA、蛋白质、肽）以诱导肿瘤特异性的免疫反应（T 和 B 细胞介导的免疫反应）。这根据本发明还包括调节 SLC13A1 生物活性并能用于治疗肾肿瘤的所谓小分子化合物的开发。

[0411] 实施例 6 :CLCA1 作为诊断和治疗癌症的靶点的鉴别

[0412] CLCA1 (SEQ ID NO :51 ;翻译产物 :SEQ ID NO :60) 属于钙离子激活的氯通道家族。序列发表在基因库注册号 NM_001285 之下。CLCA1 只表达在肠腺上皮中和杯状细胞中 (Gruber 等人, Genomics 54 :200-14, 1998)。对 CLCA1 是否能被用作结肠和胃癌的标志物进行了调查。为了这个目的, 使用了能使 CLCA1 特异性扩增的寡聚核苷酸 (SEQ ID NO :67, 68)。用这个引物对的 RT-PCR 调查证实了在结肠中的选择性表达, 并且根据本发明显示在 (3/7) 被调查的结肠和 (1/3) 被调查的胃癌样品中的高度表达 (图 7)。其它的正常组织未显示或仅有很微弱的表达。这还用特异性定量 RT-PCR (SEQ ID NO :125, 126) 被证实, 其中在被分析的正常组织中未能检测到 CLCA1 的表达 (图 36)。在该试验中调查的肿瘤样品中, 6/12 结肠癌样品和 5/10 胃癌样品 CLCA1 阳性。总的说来, 该基因在肿瘤中的表达似乎是失控的。除了有非常强表达的样品外, CLCA1 在其它的样品中被显著下调。

[0413] 所述蛋白质被预报具有 4 个跨膜蛋白结构域, 共有 2 个细胞外区域。CLCA1 的这些细胞外结构域尤其根据本发明能被用作单克隆抗体的靶结构。

[0414] CLCA1 在胃癌和结肠癌中的显著表达和高发生率使得该蛋白质根据本发明成为引人注目的诊断和治疗标志物。这根据本发明包括依靠 RT-PCR 对血清、骨髓、尿中散布肿瘤细胞的检测以及其他器官中转移灶的检测。还有可能根据本发明依靠单克隆抗体将 CLCA1 的细胞外结构域用作免疫诊断和治疗的靶结构。CLCA1 还根据本发明能被用作疫苗（RNA、DNA、蛋白质、肽）以诱导肿瘤特异性的免疫反应（T 和 B 细胞介导的免疫反应）。这根据本发明还包括调节 CLCA1 生物活性并能被用于治疗胃肠道肿瘤的所谓小分子化合物的开发

[0415] 实施例 7 :FLJ21477 作为诊断和治疗癌症的靶点的鉴别

[0416] FLJ21477 (SEQ ID NO :52) 及其预报的翻译产物作为假设蛋白发表在基因库注册号 NM_025153 之下。它是一个整合膜蛋白质, 具有 ATPase 活性和 4 个跨膜结构域, 因此适合用与特异性抗体的治疗。用 FLJ21477 特异性引物 (SEQ ID NO :69, 70) 的 RT-PCR 调查显示在结肠中的选择性表达, 以及还在 (7/12) 被调查的结肠癌样品中不同水平的表达 (图 8)。其它正常组织未显示有表达。这还被特异性定量 RT-PCR (SEQ ID NO :127, 128) 所证实。FLJ21477 特异性的表达在结肠中 (图 37A) 和 11/12 的结肠癌中都能检测到。除了在结肠组织有表达外, 表达还可在胃组织中检测到。另外, 在定量 RT-PCR 的条件下, 脑、胸腺和食道中检测到的表达要显著弱于结肠和胃 (图 37A)。此外, 还有可能在下列的肿瘤样品

中检测 FLJ21477 的特异性表达 :胃、胰腺、食道和肝。所述蛋白质被预报具有 4 个跨膜蛋白结构域,共有 2 个细胞外区域。FLJ21477 的这些细胞外结构域尤其根据本发明能被用作单克隆抗体的靶结构。胃癌和结肠癌中 FLJ21477 的显著表达和高发生率使得该蛋白质根据本发明成为有价值的诊断和治疗标志物。这根据本发明包括依靠 RT-PCR 对血清、骨髓、尿中散布肿瘤细胞的检测以及其他器官中转移灶的检测。另外,CLCA1 的细胞外结构域根据本发明依靠单克隆抗体能用作免疫诊断和治疗的靶结构。另外,FLJ21477 的细胞外结构域根据本发明能被用作疫苗 (RNA、DNA、蛋白质、肽) 以诱导肿瘤特异性的免疫反应 (T 和 B 细胞介导的免疫反应)。

[0417] 实施例 8 :FLJ20694 作为诊断和治疗癌症的靶点的鉴别

[0418] FLJ20694 (SEQ ID NO :53) 及其翻译产物 (SEQ ID NO :62) 作为假设蛋白发表在基因库注册号 NM_017928 之下。该蛋白质是一个整合跨膜分子 (跨膜结构域氨基酸 33-54), 很可能具有硫氧化还原蛋白功能。用 FLJ20694 特异性引物 (SEQ ID NO :71,72) 的 RT-PCR 调查显示在结肠中的选择性表达,以及还在 (5/9) 被调查的结肠癌样品中不同水平的表达 (图 9)。其它正常组织未显示有表达。这还被特异性定量 RT-PCR (SEQ ID NO :129,130) 所证实 (图 38)。FLJ29694 的表达在除结肠和胃以外的任何其它正常组织中未检测到 (在第一个试验中未分析)。

[0419] 所述蛋白质被预报具有一个跨膜结构域,一个细胞外区域。FLJ20694 的这些细胞外结构域尤其根据本发明能被用作单克隆抗体的靶结构。

[0420] 另外,FLJ20694 根据本发明能被用作疫苗 (RNA、DNA、蛋白质、肽) 以诱导肿瘤特异性的免疫反应 (T 和 B 细胞介导的免疫反应)。这根据本发明还包括调节 FLJ20694 生物活性并能用于治疗胃肠道肿瘤的所谓小分子化合物的开发。

[0421] 实施例 9 :von Ebner 蛋白 (c20orf114) 作为诊断和治疗癌症的靶点的鉴别

[0422] Von Ebner 蛋白 (SEQ ID NO :54) 及其翻译产物 (SEQ ID NO :63) 被当作上气道的和鼻咽上皮的 Plunc 相关蛋白发表在基因库注册号 AF364078 之下。根据本发明对 von Ebner 蛋白是否能被用作肺癌的标志物进行调查。为此目的使用了能使 Ebner 蛋白特异性扩增的寡聚核苷酸 (SEQ ID NO :73,74)。用这个引物对的 RT-PCR 调查显示在肺和在 (5/10) 被调查的肺癌样本中的选择性表达 (图 10)。在正常组织这组中,胃中也有表达。其它正常组织未显示有表达。

[0423] 实施例 10 :Plunc 作为诊断和治疗癌症的靶点的鉴别

[0424] Plunc (SEQ ID NO :55) 及其翻译产物 (SEQ ID NO :64) 被发表在基因库注册号 NM_016583 之下。人类 Plunc 编码一个 256 个氨基酸的蛋白质,并且显示与鼠 Plunc 蛋白质有 72% 的同源性 (Bingle 和 Bingle, Biochem Biophys Acta 1493 :363-7,2000)。Plunc 的表达被限制在气管、上气道、鼻咽上皮和唾液腺。

[0425] 根据本发明对 Plunc 是否能被用作肺癌的标志物进行调查。为此目的使用了能使 Plunc 蛋白特异性扩增的寡聚核苷酸 (SEQ ID NO :76)。

[0426] 用这个引物对的 RT-PCR 调查显示在胸腺、在肺和在 (6/10) 被调查的肺癌样品中的选择性表达 (图 11)。其它正常组织未显示有表达。

[0427] 实施例 11 :SLC26A9 作为诊断和治疗癌症的靶点的鉴别

[0428] SLC6A9 (SEQ ID NO :56) 及其翻译产物 (SEQ ID NO :65) 被发表在基因库注册号

NM_134325 之下。SLC26A9 属于阴离子交换蛋白家族。SLC26A9 的表达被限制在肺的小支气管和肺泡上皮 (Lohi 等人, J Biol Chem 277:14246-54, 2002)。

[0429] 对 SLC26A9 是否能用作肺癌的标志物进行调查。为此目的使用能特异性扩增 SLC26A9 的寡聚核苷酸 (SEQ ID NO:77,78)。用 SLC26A9 特异性引物 (SEQ ID NO:77,78) 的 RT-PCR 调查显示在肺和在所有 (13/13) 被调查的肺癌样品中的选择性表达 (图 12)。除甲状腺外, 其它的正常组织未显示有表达。在定量 RT-PCR 试验中用 SEQ ID NO:131 和 132 首次有可能证实这些结果, 并且获得额外的信息。有可能在肺、结肠、胰腺和胃肿瘤的 4-5 个肿瘤组织的合并样品中检测到高表达水平的 SLC26A9 特异性 RNA。SLC26A9 是跨膜阴离子转运蛋白家族的一员。在健康肺中, 所述蛋白质朝着气道的方向被定向于腔膜面, 并且因而不能直接被来自血液中的 IgG 抗体利用。相反, 蛋白质的极性在肿瘤中被取消。所以根据本发明可能在阐述的肿瘤中, 其中包括肺肿瘤、胃癌、胰腺癌中, 将 SLC26A9 称为使用单克隆抗体的治疗靶点。肺、胃、胰腺和食道癌 SLC26A9 显著地高表达和高发生率使得这个蛋白质根据本发明成为极好的诊断和治疗标志物。这根据本发明包括依靠 RT-PCR 对血清、骨髓、尿中散布肿瘤细胞的检测以及其他器官中转移灶的检测。另外, FLJ20694 的这些细胞外结构域根据本发明依靠单克隆抗体能被用作免疫诊断和治疗的靶结构。SLC26A9 根据本发明还可能被用作疫苗 (RNA、DNA、蛋白质、肽) 以诱导肿瘤特异性的免疫反应 (T 和 B 细胞介导的免疫反应)。这根据本发明还包括调节 SLC26A9 生物活性并能用于治疗肺肿瘤和胃肠道肿瘤的所谓小分子化合物的开发。

[0430] 实施例 12:THC1005163 作为诊断和治疗癌症的靶点的鉴别

[0431] THC1005163 (SEQ ID NO:57) 是源自美国基因组研究所 (TIGR) 基因指标 (gene index) 的一个基因片段。该基因只有在 3' 端有阐述, 而 ORF 是缺乏的。用 THC1005163 特异性引物 (SEQ ID NO:79) 和寡聚 dT₁₈ 引物进行 RT-PCR 调查, 所述 dT₁₈ 引物的 5' 端具有一个 21 个特异性碱基的特异性标签。用数据搜索程序搜索来检查这个标签与已知序列的同源性。为了排除基因组 DNA 的污染, 这个特异性引物最初被用于 cDNA 合成。用这个引物对的 RT-PCR 调查显示在胃、卵巢、肺和 (5/9) 肺癌活检组织中的表达 (图 13)。其它的正常组织未显示有表达。

[0432] 实施例 13:LOC134288 作为诊断和治疗癌症的靶点的鉴别

[0433] LOC134288 (SEQ ID NO:58) 及其预报的翻译产物 (SEQ ID NO:66) 被发表在基因库注册号 XM_059703 之下。

[0434] 根据本发明对 LOC134288 是否能被用作肾细胞癌的标志物进行调查。为此目的使用了能特异性扩增 LOC134288 的寡聚核苷酸 (SEQ ID NO:80,81)。RT-PCR 调查显示在肾中和在 (5/8) 被调查的肾细胞癌活检组织中的选择性表达 (图 14)。

[0435] 实施例 14:THC943866 作为诊断和治疗癌症的靶点的鉴别

[0436] THC 943866 (SEQ ID NO:59) 是来源于 TIGR 基因指标的一个基因片段。对 THC943866 是否能被用作肾细胞癌的标志物进行调查。为此目的使用了能特异性扩增 THC943866 的寡聚核苷酸 (SEQ ID NO:82,83)。

[0437] 用 THC943866 特异性引物 (SEQ ID NO:82,83) 的 RT-PCR 调查显示在肾中和在 (4/8) 被调查的肾细胞癌活检组织中的选择性表达 (图 15)。

[0438] 实施例 15:FLJ21458 作为诊断和治疗癌症的靶点的鉴别

[0439] FLJ21458 (SEQ ID NO :84) 及其预报的翻译产物 (SEQ ID NO :85) 被发表在基因库注册号 NM_034850 之下。序列分析揭示所述蛋白质代表了 butyrophillin 家族的新成员。结构分析揭示它代表了有一个细胞外免疫球蛋白结构域的 I 型跨膜蛋白质。能特异性扩增 FLJ21458 的寡聚核苷酸 (SEQ ID NO :86,87) 被用于调查其表达。用 FLJ21458 特异性引物 (SEQ ID NO :86,87) 的 RT-PCR 调查显示在结肠和在 (7/10) 被调查的结肠癌活检组织中的选择性表达 (图 16, 表 5)。用特异性引物 (SEQ ID NO :133,134) 的定量 RT-PCR 证实这个选择性表达的轮廓 (图 39)。还有可能在试验中在结肠和在胃、在直肠和盲肠以及在睾丸中胃肠道特异性地检测 FLJ21458。7/11 结肠转移灶样品在定量 PCR 中也为阳性。FLJ21458 特异性的表达还延伸到其它肿瘤, 并且蛋白质特异性的表达在胃、胰腺和肝肿瘤中可检测到 (表 5)。用于检测 FLJ21458 蛋白质的抗体通过免疫兔产生。下列肽用于生产这些抗体:

[0440] SEQ ID NO :135 :QWQVFGPDKPVQAL

[0441] SEQ ID NO :136 :AKWKGPQGQDLSTDS

[0442] FLJ21458 特异性的反应在免疫荧光法中可检测到 (图 40)。为了检查抗体的特异性, 用编码 FLJ21458-GFP 融合蛋白质的质粒转染 293 细胞。特异性一方面通过使用 FLJ21458 特异性抗体的共定位调查, 另一方面通过自发荧光的 GFP 而被证明。两个荧光图的叠加清楚地显示免疫血清特异性地识别 FLJ21458 蛋白质 (图 40a)。由于蛋白质的过度表达, 得到的细胞染色是弥漫的, 并且不允许有不清楚的蛋白质定位。为了这个原因, 用内源性表达 FLJ21458 的胃肿瘤特异性的细胞系 Snu16 细胞进行进一步的免疫荧光试验 (图 41B)。细胞用 FLJ21458 特异性抗血清和用识别膜蛋白质 E- 钙粘蛋白的另一个抗体来染色。FLJ21458 特异性抗体至少微弱地染色细胞膜并且因而 FLJ21458 很明显的定位于细胞膜上。

[0443] 生物信息调查显示由 FLJ21458 编码的蛋白质代表了一个细胞表面分子并且具有一个免疫球蛋白超分子结构域。这个表面分子的选择性表达使其成为开发检测肿瘤细胞的诊断方法以及清除肿瘤细胞的治疗方法的一个好的靶点。

[0444] 胃癌和结肠癌 FLJ21458 的显著表达和高发生率根据本发明使该蛋白质成为一个高度引人注目的诊断和治疗标志物。这根据本发明包括依靠 RT-PCR 对血清、骨髓、尿中散布肿瘤细胞的检测以及其他器官中转移灶的检测。还有可能根据本发明将 FLJ21458 的细胞外结构域用作免疫诊断和治疗的靶结构。根据本发明还可能将 FLJ21458 用作疫苗 (RNA、DNA、蛋白质、肽) 以诱导肿瘤特异性的免疫反应 (T 和 B 细胞介导的免疫反应)。这根据本发明还包括调节 FLJ21458 生物活性并能用于治疗胃肠道肿瘤的所谓小分子化合物的开发。

[0445] 表 5 FLJ21458 在正常和肿瘤组中的表达

[0446]

正常组织	表达
脑	-
小脑	-
心肌	未测定
骨骼肌	-
子宫内膜	-
胃	++
结肠	+++
胰腺	-
肾	-
肝	-
睾丸	++
胸腺	未测定
乳腺	未测定
卵巢	-
子宫	-
皮肤	-
肺	-
甲状腺	未测定
淋巴结	-
脾	-
外周血单核细胞	-
肾上腺	未测定
食道	-
小肠	-
前列腺	-

肿瘤类型	表达
结肠癌	7/10
胰腺癌	5/6
食道癌	未测定
胃癌	8/10
支气管癌	未测定
乳腺癌	未测定
卵巢癌	未测定
子宫内膜癌	未测定
耳鼻喉肿瘤	未测定
肾细胞癌	未测定
前列腺癌	未测定
结肠的转移灶	7/11
肝癌	5/8

[0001]

序列表

<110> Ganymed Pharmaceuticals AG

Sahin Dr., Ugur

Tureci Dr., Ozlem

Koslowski Dr., Michael

<120> 肿瘤内差异表达的基因产物及其应用

<130> 342-4PCT

<150> DE 102 54 601.0

<151> 2002-11-22

<160> 137

<170> PatentIn version 3.1

<210> 1

<211> 1875

<212> DNA

<213> 人

<400> 1
caggeccagag tcccagctgt cctggactct gctgtgggga agggctgatg caggtgtgga 60
gtcaaatgtg ggtgectcct gcagccgggt gccaggagggt gtggaggggc caccctgggc 120
tttgtccggg agcctgggtct tcccgctcctt gggtgacag gtgctgctgc ctctgagccc 180
tcccctgtaa gagctgtgtg ctgggtaagg ctggtggccc ttgggctcc ctgtccagga 240
tttgtctct ggagggtagg gcttgcctgg ctggggactg gaggggaacg tggagctcct 300
tctgectcct ttctgcccc atgacagcag gcagatccca ggagagaaga gctcaggaga 360
tggaagagg atctgtccag gggtagacc tcaagggtga cttggagttc tttacggcac 420
ccatgcttcc tttagaggagt ttgtgtttg tgggtgtggg gtgggggctc acctcctccc 480
acatccctgc ccagaggtgg gcagagtggt ggcagtgctt tgcctccccc gctcgtcttc 540
tgctgacctc cggtccctg tgctgcccc gaacctgaa tggcacctac aacacctgtg 600
gctccagcga cctcacctgg ccccagcga tcaagctggg cttctacgcc tacttgggcg 660
tctgtctggt gctaggcctg ctgctcaaca gcttgccgt ctgggtgttc tgctgccga 720
tgacagcgtg gacggagacc cgcctctaca tgaccaacct ggcgggtggc gacctctgcc 780
tgctgtgcac ctggccctc gtgctgcact ccttgcgaga cacctcagac acgccgtgt 840
gccagctctc ccagggcacg tacctgacca acaggtacat gacatcagc ctggtcacgg 900
ccatgcctgt ggacctgtat gtggcctgct ggcacccgt gctgccccg gggtgcggt 960

[0002]

```

ccccaggca ggcgcggcc gtgtgcggc tcctctgggt gctggtcate ggctccctgg 1020
tggtctctg gctcctgggg attcaggagg gcggcttctg cttcaggagc acccggcaca 1080
atttcaactc catggcgttc ccgctgctgg gattctacct gcccctggcc gtggtggtct 1140
tctgctccct gaaggtggtg actgccctgg ccagaggccc acccaccgac gtggggcagg 1200
cagaggccac ccgcaaggct gcccgcattg tctgggccc cctcctgggt ttcgtggtct 1260
gcttctctgc cctgcacgtg gggctgacag tgcgcctcgc agtgggctgg aacgcctgtg 1320
ccctcctgga gacgatccgt cgcgcctgt acataaccag caagctctca gatgccaact 1380
gctgcctgga gcgcattctg tactactaca tggccaagga gttccaggag gcgtctgcac 1440
tgcccggtgc tcccagtgtc aagcccaca aaagccagga ctctctgtgc gtgacctcg 1500
cctaagaggc gtgctgtggg cgctgtgggc caggtctcgg ggctccggg aggtgtctgc 1560
tgccagggga agctggaacc agtagcaagg agcccgggat cagccctgaa ctcactgtgt 1620
attctcttgg agccttgggt gggcagggac ggcaccagta cctgctctct tgggaagaga 1680
gagggacagg gacaaggcca agaggactga ggcagagca aggccaatgt cagagacccc 1740
cgggatgggg cctcacactt gccaccccca gaaccagctc acctggccag agtgggttcc 1800
tgctggccag ggtgcagcct tgatgacacc tgcgcgtgcc cctcggggct ggaataaaac 1860
tccccacca gagtc 1875

```

<210> 2

<211> 3222

<212> DNA

<213> 人

<400> 2

```

atgaagacgt tgctgttga cttggctttg tggtcactgc tcttcagcc cgggtggctg 60
tcctttagtt cccagggtgag tcagaactgc cacaatggca gctatgaaat cagcgtcctg 120
atgatgggca actcagcctt tgcagagccc ctgaaaaact tggaaagatgc ggtgaatgag 180
gggctggaat tagtgagagg acgtctgcaa aatgctggcc taaatgtgac tglgaacgct 240
actttcatgt attcggatgg tctgattcat aactcaggcg actgccggag tagcaccgtt 300
gaaggcctcg acctactcag gaaaatttca aatgcacaac ggatgggctg tgtcctcata 360
gggcctctcat gtacatactc caccctccag atgtaccttg acacagaatt gagctacccc 420
atgatctcag ctggaagttt tggattgtca tgtactata aagaacacct aaccaggctg 480
atgtctccag ctgaaagtt gatgtacttc ttggttaact ttggaaaac caacgatctg 540
cccttcaaaa cttattcctg gaccactctg tatgtttaca agaattgtac agaaactgag 600
gactgtttct ggtaccttaa tgcctcggag gctagegttt cctatttctc ccacgaactc 660
ggctttaagg tgggtttaag acaagataag gagtttcagg atatcttaat ggaccacaac 720
agggaaaagca atgtgattat tatgtgtggt ggtccagagt tcctctacaa gctgaagggt 780
gaccgagcag ttgctgaaga cattgtcatt attctagtgg atcttttcaa tgaccagtac 840
ttggaggaca atgtcacagc cctgactat atgaaaaatg tcttgtttct gacgtgtctt 900
cctgggaatt cccttctaaa tagctctttc tccaggaatc taccaccaac aaaacgagac 960

```

[0003]

```

tttgcctcttg cctatttgaa tggaaacctg ctctttggac atatgctgaa gatatttctt 1020
gaaaatggag aaaatattac cacccecaaa tttgctcatg ctttcaggaa tctcactttt 1080
gaagggtatg acgggtccagt gaccttggat gactgggggg atgttgacag taccatgggtg 1140
cttctgtata cctctgtgga caccaagaaa tacaagggtc ttttgacctg tgatacccac 1200
gtaaataaga cctatcctgt ggatatgagc cccacattca ctggaagaa ctctaaactt 1260
cctaagtata ttacaggccg gggccctcag atcctgatga ttgcagtctt caccctcact 1320
ggagctgttg tgctgtcctt gctcgtcgt ctctgatgc tcagaaaata tagaaaagat 1380
tatgaacttc gtcagaaaaa atggteccac attcctcctg aaaatatctt tctctggag 1440
accaatgaga ccaatcatgt tagcctcaag atcgatgatg acaaaagacg agatacaatc 1500
cagagactac gacagtgcac atacgacaaa aagegagtga ttctcaaaga tctcaagcac 1560
aatgatggta atttactga aaacacagaag atagaattga acaagttgct tcagattgac 1620
tattacaacc tgaccaagtt ctacggcaca gtgaaacttg ataccatgat cttcgggggtg 1680
atagaatact gtgagagagg atccctccgg gaagttttaa atgacacaat ttctaccct 1740
gatggcacat tcatggattg ggagttaag atctctgtct tgtatgacat tgctaaggga 1800
atgtcatatc tgcactccag taagacagaa gtccatggtc gtctgaaatc taccactgc 1860
gtagtggaca gtagaatggt ggtagaagtc actgattttg gctgcaattc cattttacct 1920
ccaaaaaagg acctgtggac agctccagag caccctccgc aagccaacat ctctcagaaa 1980
ggagatgtgt acagctatgg gatcatcgca caggagatca ttctcggaa agaaacctc 2040
tacactttga gctgtcggga ccggaatgag aagattttca gagtggaaaa ttccaatgga 2100
atgaaacctc tccgccaga ttattcttg gaaacagcag aggaaaaaga gctagaagtg 2160
tacctacttg taaaaaactg ttgggaggaa gatccagaaa agagaccaga ttcaaaaaa 2220
attgagacta cacttgccaa gatatttga ctttttcctg accaaaaaa tgaaagctat 2280
atggatacct tgatccgacg tctacagta tattctcgaa acctggaaca tctggtagag 2340
gaaaggacac agctgtacaa ggcagagagg gacagggtg acagacttaa ctttatgttg 2400
cttccaagcg tagtggtaaa gtctctgaag gagaaaggt tttgtggagcc ggaactatat 2460
gaggaaglla caatctactl cagtgcatt glagglitca ctactatctg caaatacagc 2520
acccccatgg aagtgtgga catgctaat gacatctata agagttttga ccacattgtt 2580
gatcatcatg atgtctacaa ggtggaaacc atcgtgatg cgtacatggt ggctagtgg 2640
ttgcctaaga gaaatggcaa tcggcatgca atagacattg ccaagatggc ettggaatc 2700
ctcagcttca tggggacctt tgagctggag catcttctg gcttcccaat atggattcgc 2760
attggagttc actctgttc cgtgtctgt ggagltgtg gaatcaagat gccctgttat 2820
tgtctatttg gagatacgtt caacacagcc tctaggatgg aatccactgg cctcccttg 2880
agaattcacg tgagtggctc caccatagcc atcctgaaga gaactgagtg ccagttcctt 2940
tatgaagtga gaggagaaac atacttaaag ggaaggaggaa atgagactac ctactggctg 3000
actgggatga aggaccagaa attcaacctg ccaacccctc ctactgtgga gaatcaacag 3060
cgtttgcaag cagaattttc agacatgatt gccaaactct tacagaaaag acaggcagca 3120
gggataagaa gccaaaaacc cagacgggta gccagctata aaaaaggcac tctggaatac 3180
ttgcagctga ataccacaga caaggagagc acctattttt aa 3222

```

[0004]

<210>	3	
<211>	336	
<212>	DNA	
<213>	人	
<400>	3	
atgaagacgt	tgctgttgga	cttggtttg
tggtcactgc	tcttcagcc	cgggtggctg
		60
tcctttagtt	cccaggtgag	tcagaactgc
cacaatggca	gctatgaaat	cagcgtcctg
		120
atgatgggca	actcagcctt	tcagagccc
ctgaaaaact	tggaagatgc	ggtgaatgag
		180
gggcctggaaa	tagtgagagg	acgtctgcaa
aatgctggcc	taaatgtgac	tgtgaacgt
		240
actttcatgt	attcggatgg	tctgattcat
aactcaggcg	actgccggag	tagcacctgt
		300
gaaggcctcg	acctactcag	gaaaatttca
ccttga		336
<210>	4	
<211>	777	
<212>	DNA	
<213>	人	
<400>	4	
atgaagacgt	tgctgttgga	cttggtttg
tggtcactgc	tcttcagcc	cgggtggctg
		60
tcctttagtt	cccaggtgag	tcagaactgc
cacaatggca	gctatgaaat	cagcgtcctg
		120
atgatgggca	actcagcctt	tcagagccc
ctgaaaaact	tggaagatgc	ggtgaatgag
		180
gggctggaaa	tagtgagagg	acgtctgcaa
aatgctggcc	taaatgtgac	tgtgaacgt
		240
actttcatgt	attcggatgg	tctgattcat
aactcaggcg	actgccggag	tagcacctgt
		300
gaaggcctcg	acctactcag	gaaaatttca
aatgcacaac	ggatgggctg	tgctctcata
		360
gggcctcat	gtacatactc	cacctccag
atgtaccttg	acacagaatt	gagctacccc
		420
atgatctcag	ctggaagttt	tggattgtca
tgtgactata	aagaaacctt	aaccaggctg
		480
atgtctccag	ctagaaagtl	gatgtacttc
ttggllaacl	tttgaaaac	caacgatctg
		540
cccttcaaaa	cttattcctg	gagcacttcg
tatgtttaca	agaatggtac	agaaactgag
		600
gactgtttct	ggtaccttaa	tgtctggag
gctagcgttt	cctattttct	ccaacgaactc
		660
ggctttaagg	tggtgttaag	acaagataag
gagtttcagg	atatcttaat	ggaccacaac
		720
aggaaaagca	atgtgaccag	tacttggagg
acaatgtcac	agccctgac	tatatga
		777
<210>	5	
<211>	3213	
<212>	DNA	
<213>	人	
<400>	5	
atgaagacgt	tgctgttgga	cttggtttg
tggtcactgc	tcttcagcc	cgggtggctg
		60

[0005]

tecttttagtt cccaggtgag tcagaactgc cacaatggca gctatgaaat cagcgtcctg	120
atgatgggca actcagcctt tgcagagccc ctgaaaaact tggaagatgc ggtgaatgag	180
gggctggaaa tagtgagagg acgtctgcaa aatgctggcc taaatgtgac tgtgaacgct	240
actttcatgt attcggatgg tctgattcat aactcaggcg actgccggag tagcacctgt	300
gaaggcctcg acctactcag gaaaatttca aatgcacaac ggatgggctg tgtcctcata	360
gggccctcat gtacatactc caccctccag atgtaccttg acacagaatt gagctacccc	420
atgatctcag ctggaagttt tggattgtca tgtgactata aagaaacctt aaccaggctg	480
atgtctccag ctagaaagtt gatgtacttc ttggttaact ttggaaaaac caacgatctg	540
cccttcaaaa cttattcctg gagcacttcg tatgtttaca agaatggtag agaaactgag	600
gactgtttct ggtaccttaa tgcctcggag gctagcgttt cctattttct ccacgaactc	660
ggctttaagg tgggtttaag acaagataag gagtttcagg atatcttaat ggaccacaac	720
aggaaaagca atgtgattat tatgtgtggt ggtccagagt tcctctacaa gctgaagggt	780
gaccgagcag tggctgaaga cattgtcatt attctagtgg atcttttcaa tgaccagtac	840
ttggaggaca atgtcacagc ccttgacatc atgaaaaatg tcttlttctc gacgtgtctc	900
cctgggaatt ccttctaaa tagctcttc tccaggaatc tatccaac aaaacgagac	960
tttgccttg cctattttaa tggaaacctg ccttlttgac atatgtttaa gatatttctt	1020
gaaaaatggag aaaatattac cacccecaaa ttgtctcatg ctttcaggaa tctcactttt	1080
gaagggtatg acgggtccagt gaccttggat gactgggggg atgttgacag taccatggtg	1140
cttctgtata cctctgtgga caccaagaaa tacaaggttc ttttgacctg tgatacccac	1200
gtaataaaga cctatcctgt ggatatgagc cccacattca cttggaagaa ctctaaactt	1260
cctaatagata ttacaggccg gggccctcag atcctgatga ttgcagtctt caccctcact	1320
ggagctgtgg tgcgtctcct gctcgtcgtc ctcctgatgc tcagaaaata tagaaaagat	1380
tatgaacttc gtcagaaaaa atgggtccac attcctcctg aaaatatctt tcctctggag	1440
accaatgaga ccaatcatgt tagcctcaag atcgatgatg acaaaagacg agatacaatc	1500
cagagactac gacagtgcac atacgacaaa aagcgagtga ttctcaaaga tctcaagcac	1560
aatgatggta atttactga aaaacagaag atagaattga acaagattga ctattacaac	1620
ctgaccaagt tctacggcac agtgaactt gataccatga tcttcggggt gatagaatac	1680
tgtgagagag gatccctccg ggaagtttta aatgacacaa ttctctaccg tgatggcaca	1740
ttcatggatt gggagtttaa gatctctgtc ttgtatgaca ttgctaaggg aatgtcatat	1800
ctgcactcca gtaagacaga agtccatggt cgtctgaaat ctaccaactg cgtagtggac	1860
agtagaatgg tggtaagat cactgatttt ggctgcaatt ccattttacc tccaaaaaag	1920
gacctgtgga cagctccaga gcacctccg caagccaaca tctctcagaa aggagatgtg	1980
tacagctatg ggatcatgc acaggagatc attctgcgga aagaacctt ctacactttg	2040
agctgtcggg accggaatga gaagattttc agagtggaaa attccaatgg aatgaaacct	2100
ttccgccag atttattctt ggaaacagca gaggaaaaag agctagaagt gtacctactt	2160
gtaaaaaact gttgggagga agatccagaa aagagaccag atttcaaaaa aattgagact	2220
acacttgcca agatatttgg actttttcat gacaaaaaaa atgaaagcta tatggatacc	2280
ttgatecgac gtctacagct atattctega aacctggaac atctggtaga ggaaaggaca	2340

[0006]

```

cagctgtaca aggcagagag ggacagggt gacagactta acittatgtt gcttccaagg 2400
ctagtggtaa agtctctgaa ggagaaagge tttgtggagc cggaactata tgaggaagtt 2460
acaatctact tcagtgcacat ttaggttttc actactatct gcaatacag ccccccatg 2520
gaagtggtgg acatgcttaa tgacatcat aagagtttgg accacattgt tgatcatcat 2580
gatgtctaca aggtggaaac catcgggtgat gcgtacatgg tggctagtgg tttgcctaag 2640
agaaatggca atcggcatgc aatagacatt gccaaagtgg ccttggaat cctcagcttc 2700
atggggacct ttgagctgga gcattcttct ggctctccaa tatggattcg cattggagtt 2760
cactctgtgc cctgtctgc tggagtgtg ggaatcaaga tgcctcgtta ttgtctattt 2820
ggagatacgg tcaacacagc ctctaggatg gaatccactg gcctcccttt gagaattcac 2880
gtgagtggct ccacatagc catctgaag agaactgagt gccagtctct ttatgaagtg 2940
agaggagaaa catacttaaa gggaagagga aatgagacta cctactggct gactgggatg 3000
aaggaccaga aattcaacct gccaacacct cctactgtgg agaataaca gcgtttgcaa 3060
gcagaatddd cagacatgat tgccaactct ttacagaaaa gacaggcagc aggataaga 3120
agccaaaaac ccagacgggt agccagctat aaaaaaggca ctctggaata ctgcagctg 3180
aataccacag acaaggagag cactatddd taa 3213

```

<210> 6
 <211> 550
 <212> DNA
 <213> 人

```

<400> 6
ggggacactt tgatggcaa ggggaaccac tggctgggtg gatttggcta gatttctctg 60
attdtttaac tctgaaaaa tatcccagat aactgtcatg aagctggtaa ctatcttct 120
gtctgtgacc atcagccttt gtagttactc tgctactgcc ttctcatca acaagtgcc 180
ccttctgttt gacaagtgtg cacttttacc tctggacaac attcttccct ttatggatcc 240
attaaagctt cttctgaaaa ctctgggcat tctgttgag cacttgttg agggtctaag 300
gaagtgtgta aatgagctgg gaccagagge tctgaagct gtgaagaaac tgctggagge 360
getatcacac ttggtgtgac atcaagataa agagcggagg tggatgggga tggaagatga 420
tgctcttate ctctctgect gaaacctgtt ctaccaatta tagatcaaat gcctaaaaat 480
gtagtgacce gtgaaaagga caataaagc aatgaatact aaaaaaaaaa aaaaaaaaaa 540
aaaaaaaaaa 550

```

<210> 7
 <211> 786
 <212> DNA
 <213> 人

```

<400> 7
atggcctgta ctgctgtca gggcttgggg ttcgtgttt cactgattgg gattgaggge 60
atcattgtct ccacctgcat ggaccagtgg agcaccgaag acttgtacaa caaccccgta 120
acagctgttt tcaactacca ggggtgtggt cgtctctgtg tccagagag ctctggcttc 180

```

[0007]

```

accgagtgcc ggggetactt caccctgetg gggetgccag ccatgetgca ggcagtgcga    240
gccctgatga tcgttaggeat cgtectgggt gccattggcc tcctgggtatc catctttgcc    300
ctgaaatgca tcgcattgg cagcatggag gactctgcc aagccaacat gacactgacc    360
tcgggatca tgttcattgt ctcaggtett tgtgcaattg ctggagtgtc tgtgtttgcc    420
aacatgetgg tgactaactt ctggatgtcc acagetaaca tgtacaccgg catgggtggg    480
atggtgcaga ctgttcagac caggtacaca ttggtgcgg ctctgttcgt gggctgggtc    540
gtggagggcc tcacactaat tgggggtgtg atgatgtgca tcgctgcgc gggcctggca    600
ccagaagaaa ccaactacaa agcctttct tatcatgcct cagccacag tgttgcctac    660
aagcctggag gcttcaaggc cagcactggc ttgggtcca acacaaaaa caagaagata    720
tacgatggag gtgccegcac agaggacgag gtacaatett atccttccaa gcacgactat    780
gtgtaa                                           786

```

<210> 8

<211> 180

<212> DNA

<213> 人

<400> 8

```

tgcgccacca tggcctgac tgcctgtcag ggcttggggc tcgtggttc actgattggg    60
attgcgggca tcattgtgc cacctgatg gaccagtgga gcaccaaga cttgtacaac    120
aaccctgtaa cagctgtttt caactaccag gggctgtggc gctcctgtgt ccgagagagc    180

```

<210> 9

<211> 309

<212> PRT

<213> 人

<400> 9

```

Met Asn Gly Thr Tyr Asn Thr Cys Gly Ser Ser Asp Leu Thr Trp Pro
1           5           10           15

```

```

Pro Ala Ile Lys Leu Gly Phe Tyr Ala Tyr Leu Gly Val Leu Leu Val
          20           25           30

```

```

Leu Gly Leu Leu Leu Asn Ser Leu Ala Leu Trp Val Phe Cys Cys Arg
          35           40           45

```

```

Met Gln Gln Trp Thr Glu Thr Arg Ile Tyr Met Thr Asn Leu Ala Val
          50           55           60

```

```

Ala Asp Leu Cys Leu Leu Cys Thr Leu Pro Phe Val Leu His Ser Leu
65           70           75           80

```

```

Arg Asp Thr Ser Asp Thr Pro Leu Cys Gln Leu Ser Gln Gly Ile Tyr
          85           90           95

```

[0008]

Leu Thr Asn Arg Tyr Met Ser Ile Ser Leu Val Thr Ala Ile Ala Val
100 105 110

Asp Arg Tyr Val Ala Val Arg His Pro Leu Arg Ala Arg Gly Leu Arg
115 120 125

Ser Pro Arg Gln Ala Ala Ala Val Cys Ala Val Leu Trp Val Leu Val
130 135 140

Ile Gly Ser Leu Val Ala Arg Trp Leu Leu Gly Ile Gln Glu Gly Gly
145 150 155 160

Phe Cys Phe Arg Ser Thr Arg His Asn Phe Asn Ser Met Arg Phe Pro
165 170 175

Leu Leu Gly Phe Tyr Leu Pro Leu Ala Val Val Val Phe Cys Ser Leu
180 185 190

Lys Val Val Thr Ala Leu Ala Gln Arg Pro Pro Thr Asp Val Gly Gln
195 200 205

Ala Glu Ala Thr Arg Lys Ala Ala Arg Met Val Trp Ala Asn Leu Leu
210 215 220

Val Phe Val Val Cys Phe Leu Pro Leu His Val Gly Leu Thr Val Arg
225 230 235 240

Leu Ala Val Gly Trp Asn Ala Cys Ala Leu Leu Glu Thr Ile Arg Arg
245 250 255

Ala Leu Tyr Ile Thr Ser Lys Leu Ser Asp Ala Asn Cys Cys Leu Asp
260 265 270

Ala Ile Cys Tyr Tyr Tyr Met Ala Lys Glu Phe Gln Glu Ala Ser Ala
275 280 285

Leu Ala Val Ala Pro Arg Ala Lys Ala His Lys Ser Gln Asp Ser Leu
290 295 300

Cys Val Thr Leu Ala
305

<210> 10

<211> 394

<212> PRT

<213> 人

<400> 10

Met Thr Ala Gly Arg Ser Gln Glu Arg Arg Ala Gln Glu Met Gly Arg
1 5 10 15

Gly Ser Val Gln Gly Leu Asp Leu Lys Gly Asp Leu Glu Phe Phe Thr
20 25 30

[0009]

Ala Pro Met Leu Ser Leu Arg Ser Phe Val Phe Val Gly Val Gly Ser
 35 40 45
 Gly Leu Thr Ser Ser His Ile Pro Ala Gln Arg Trp Ala Glu Trp Gly
 50 55 60
 Gln Cys Leu Ala Pro Pro Ala Arg Ser Leu Leu Thr Ser Gly Ser Leu
 65 70 75 80
 Cys Cys Pro Arg Thr Met Asn Gly Thr Tyr Asn Thr Cys Gly Ser Ser
 85 90 95
 Asp Leu Thr Trp Pro Pro Ala Ile Lys Leu Gly Phe Tyr Ala Tyr Leu
 100 105 110
 Gly Val Leu Leu Val Leu Gly Leu Leu Leu Asn Ser Leu Ala Leu Trp
 115 120 125
 Val Phe Cys Cys Arg Met Gln Gln Trp Thr Glu Thr Arg Ile Tyr Met
 130 135 140
 Thr Asn Leu Ala Val Ala Asp Leu Cys Leu Leu Cys Thr Leu Pro Phe
 145 150 155 160
 Val Leu His Ser Leu Arg Asp Thr Ser Asp Thr Pro Leu Cys Gln Leu
 165 170 175
 Ser Gln Gly Ile Tyr Leu Thr Asn Arg Tyr Met Ser Ile Ser Leu Val
 180 185 190
 Thr Ala Ile Ala Val Asp Arg Tyr Val Ala Val Arg His Pro Leu Arg
 195 200 205
 Ala Arg Gly Leu Arg Ser Pro Arg Gln Ala Ala Ala Val Cys Ala Val
 210 215 220
 Leu Trp Val Leu Val Ile Gly Ser Leu Val Ala Arg Trp Leu Leu Gly
 225 230 235 240
 Ile Gln Glu Gly Gly Phe Cys Phe Arg Ser Thr Arg His Asn Phe Asn
 245 250 255
 Ser Met Ala Phe Pro Leu Leu Gly Phe Tyr Leu Pro Leu Ala Val Val
 260 265 270
 Val Phe Cys Ser Leu Lys Val Val Thr Ala Leu Ala Gln Arg Pro Pro
 275 280 285
 Thr Asp Val Gly Gln Ala Glu Ala Thr Arg Lys Ala Ala Arg Met Val
 290 295 300
 Trp Ala Asn Leu Leu Val Phe Val Val Cys Phe Leu Pro Leu His Val
 305 310 315 320
 Gly Leu Thr Val Arg Leu Ala Val Gly Trp Asn Ala Cys Ala Leu Leu
 325 330 335

[0010]

Glu Thr Ile Arg Arg Ala Leu Tyr Ile Thr Ser Lys Leu Ser Asp Ala
340 345 350

Asn Cys Cys Leu Asp Ala Ile Cys Tyr Tyr Tyr Met Ala Lys Glu Phe
355 360 365

Gln Glu Ala Ser Ala Leu Ala Val Ala Pro Ser Ala Lys Ala His Lys
370 375 380

Ser Gln Asp Ser Leu Cys Val Thr Leu Ala
385 390

<210> 11

<211> 1073

<212> PRT

<213> 人

<400> 11

Met Lys Thr Leu Leu Leu Asp Leu Ala Leu Trp Ser Leu Leu Phe Gln
1 5 10 15

Pro Gly Trp Leu Ser Phe Ser Ser Gln Val Ser Gln Asn Cys His Asn
20 25 30

Gly Ser Tyr Glu Ile Ser Val Leu Met Met Gly Asn Ser Ala Phe Ala
35 40 45

Glu Pro Leu Lys Asn Leu Glu Asp Ala Val Asn Glu Gly Leu Glu Ile
50 55 60

Val Arg Gly Arg Leu Gln Asn Ala Gly Leu Asn Val Thr Val Asn Ala
65 70 75 80

Thr Phe Met Tyr Ser Asp Gly Leu Ile His Asn Ser Gly Asp Cys Arg
85 90 95

Ser Ser Thr Cys Glu Gly Leu Asp Leu Leu Arg Lys Ile Ser Asn Ala
100 105 110

Gln Arg Met Gly Cys Val Leu Ile Gly Pro Ser Cys Thr Tyr Ser Thr
115 120 125

Phe Gln Met Tyr Leu Asp Thr Glu Leu Ser Tyr Pro Met Ile Ser Ala
130 135 140

Gly Ser Phe Gly Leu Ser Cys Asp Tyr Lys Glu Thr Leu Thr Arg Leu
145 150 155 160

Met Ser Pro Ala Arg Lys Leu Met Tyr Phe Leu Val Asn Phe Trp Lys
165 170 175

Thr Asn Asp Leu Pro Phe Lys Thr Tyr Ser Trp Ser Thr Ser Tyr Val
180 185 190

[0011]

Tyr Lys Asn Gly Thr Glu Thr Glu Asp Cys Phe Trp Tyr Leu Asn Ala
 195 200 205
 Leu Glu Ala Ser Val Ser Tyr Phe Ser His Glu Leu Gly Phe Lys Val
 210 215 220
 Val Leu Arg Gln Asp Lys Glu Phe Gln Asp Ile Leu Met Asp His Asn
 225 230 235 240
 Arg Lys Ser Asn Val Ile Ile Met Cys Gly Gly Pro Glu Phe Leu Tyr
 245 250 255
 Lys Leu Lys Gly Asp Arg Ala Val Ala Glu Asp Ile Val Ile Ile Leu
 260 265 270
 Val Asp Leu Phe Asn Asp Gln Tyr Leu Glu Asp Asn Val Thr Ala Pro
 275 280 285
 Asp Tyr Met Lys Asn Val Leu Val Leu Thr Leu Ser Pro Gly Asn Ser
 290 295 300
 Leu Leu Asn Ser Ser Phe Ser Arg Asn Leu Ser Pro Thr Lys Arg Asp
 305 310 315 320
 Phe Ala Leu Ala Tyr Leu Asn Gly Ile Leu Leu Phe Gly His Met Leu
 325 330 335
 Lys Ile Phe Leu Glu Asn Gly Glu Asn Ile Thr Thr Pro Lys Phe Ala
 340 345 350
 His Ala Phe Arg Asn Leu Thr Phe Glu Gly Tyr Asp Gly Pro Val Thr
 355 360 365
 Leu Asp Asp Trp Gly Asp Val Asp Ser Thr Met Val Leu Leu Tyr Thr
 370 375 380
 Ser Val Asp Thr Lys Lys Tyr Lys Val Leu Leu Thr Tyr Asp Thr His
 385 390 395 400
 Val Asn Lys Thr Tyr Pro Val Asp Met Ser Pro Thr Phe Thr Trp Lys
 405 410 415
 Asn Ser Lys Leu Pro Asn Asp Ile Thr Gly Arg Gly Pro Gln Ile Leu
 420 425 430
 Met Ile Ala Val Phe Thr Leu Thr Gly Ala Val Val Leu Leu Leu Leu
 435 440 445
 Val Ala Leu Leu Met Leu Arg Lys Tyr Arg Lys Asp Tyr Glu Leu Arg
 450 455 460
 Gln Lys Lys Trp Ser His Ile Pro Pro Glu Asn Ile Phe Pro Leu Glu
 465 470 475 480
 Thr Asn Glu Thr Asn His Val Ser Leu Lys Ile Asp Asp Asp Lys Arg
 485 490 495

[0012]

Arg Asp Thr Ile Gln Arg Leu Arg Gln Cys Lys Tyr Asp Lys Lys Arg
 500 505 510
 Val Ile Leu Lys Asp Leu Lys His Asn Asp Gly Asn Phe Thr Glu Lys
 515 520 525
 Gln Lys Ile Glu Leu Asn Lys Leu Leu Gln Ile Asp Tyr Tyr Asn Leu
 530 535 540
 Thr Lys Phe Tyr Gly Thr Val Lys Leu Asp Thr Met Ile Phe Gly Val
 545 550 555 560
 Ile Glu Tyr Cys Glu Arg Gly Ser Leu Arg Glu Val Leu Asn Asp Thr
 565 570 575
 Ile Ser Tyr Pro Asp Gly Thr Phe Met Asp Trp Glu Phe Lys Ile Ser
 580 585 590
 Val Leu Tyr Asp Ile Ala Lys Gly Met Ser Tyr Leu His Ser Ser Lys
 595 600 605
 Thr Glu Val His Gly Arg Leu Lys Ser Thr Asn Cys Val Val Asp Ser
 610 615 620
 Arg Met Val Val Lys Ile Thr Asp Phe Gly Cys Asn Ser Ile Leu Pro
 625 630 635 640
 Pro Lys Lys Asp Leu Trp Thr Ala Pro Glu His Leu Arg Gln Ala Asn
 645 650 655
 Ile Ser Gln Lys Gly Asp Val Tyr Ser Tyr Gly Ile Ile Ala Gln Glu
 660 665 670
 Ile Ile Leu Arg Lys Glu Thr Phe Tyr Thr Leu Ser Cys Arg Asp Arg
 675 680 685
 Asn Glu Lys Ile Phe Arg Val Glu Asn Ser Asn Gly Met Lys Pro Phe
 690 695 700
 Arg Pro Asp Leu Phe Leu Glu Thr Ala Glu Glu Lys Glu Leu Glu Val
 705 710 715 720
 Tyr Leu Leu Val Lys Asn Cys Trp Glu Glu Asp Pro Glu Lys Arg Pro
 725 730 735
 Asp Phe Lys Lys Ile Glu Thr Thr Leu Ala Lys Ile Phe Gly Leu Phe
 740 745 750
 His Asp Gln Lys Asn Glu Ser Tyr Met Asp Thr Leu Ile Arg Arg Leu
 755 760 765
 Gln Leu Tyr Ser Arg Asn Leu Glu His Leu Val Glu Glu Arg Thr Gln
 770 775 780
 Leu Tyr Lys Ala Glu Arg Asp Arg Ala Asp Arg Leu Asn Phe Met Leu
 785 790 795 800

[0013]

Leu Pro Arg Leu Val Val Lys Ser Leu Lys Glu Lys Gly Phe Val Glu
 805 810 815
 Pro Glu Leu Tyr Glu Glu Val Thr Ile Tyr Phe Ser Asp Ile Val Gly
 820 825 830
 Phe Thr Thr Ile Cys Lys Tyr Ser Thr Pro Met Glu Val Val Asp Met
 835 840 845
 Leu Asn Asp Ile Tyr Lys Ser Phe Asp His Ile Val Asp His His Asp
 850 855 860
 Val Tyr Lys Val Glu Thr Ile Gly Asp Ala Tyr Met Val Ala Ser Gly
 865 870 875 880
 Leu Pro Lys Arg Asn Gly Asn Arg His Ala Ile Asp Ile Ala Lys Met
 885 890 895
 Ala Leu Glu Ile Leu Ser Phe Met Gly Thr Phe Glu Leu Glu His Leu
 900 905 910
 Pro Gly Leu Pro Ile Trp Ile Arg Ile Gly Val His Ser Gly Pro Cys
 915 920 925
 Ala Ala Gly Val Val Gly Ile Lys Met Pro Arg Tyr Cys Leu Phe Gly
 930 935 940
 Asp Thr Val Asn Thr Ala Ser Arg Met Glu Ser Thr Gly Leu Pro Leu
 945 950 955 960
 Arg Ile His Val Ser Gly Ser Thr Ile Ala Ile Leu Lys Arg Thr Glu
 965 970 975
 Cys Gln Phe Leu Tyr Glu Val Arg Gly Glu Thr Tyr Leu Lys Gly Arg
 980 985 990
 Gly Asn Glu Thr Thr Tyr Trp Leu Thr Gly Met Lys Asp Gln Lys Phe
 995 1000 1005
 Asn Leu Pro Thr Pro Pro Thr Val Glu Asn Gln Gln Arg Leu Gln
 1010 1015 1020
 Ala Glu Phe Ser Asp Met Ile Ala Asn Ser Leu Gln Lys Arg Gln
 1025 1030 1035
 Ala Ala Gly Ile Arg Ser Gln Lys Pro Arg Arg Val Ala Ser Tyr
 1040 1045 1050
 Lys Lys Gly Thr Leu Glu Tyr Leu Gln Leu Asn Thr Thr Asp Lys
 1055 1060 1065
 Glu Ser Thr Tyr Phe
 1070

<210> 12

[0014]

<211> 111

<212> PRT

<213> 人

<400> 12

Met Lys Thr Leu Leu Leu Asp Leu Ala Leu Trp Ser Leu Leu Phe Gln
1 5 10 15

Pro Gly Trp Leu Ser Phe Ser Ser Gln Val Ser Gln Asn Cys His Asn
20 25 30

Gly Ser Tyr Glu Ile Ser Val Leu Met Met Gly Asn Ser Ala Phe Ala
35 40 45

Glu Pro Leu Lys Asn Leu Glu Asp Ala Val Asn Glu Gly Leu Glu Ile
50 55 60

Val Arg Gly Arg Leu Gln Asn Ala Gly Leu Asn Val Thr Val Asn Ala
65 70 75 80

Thr Phe Met Tyr Ser Asp Gly Leu Ile His Asn Ser Gly Asp Cys Arg
85 90 95

Ser Ser Thr Cys Glu Gly Leu Asp Leu Leu Arg Lys Ile Ser Pro
100 105 110

<210> 13

<211> 258

<212> PRT

<213> 人

<400> 13

Met Lys Thr Leu Leu Leu Asp Leu Ala Leu Trp Ser Leu Leu Phe Gln
1 5 10 15

Pro Gly Trp Leu Ser Phe Ser Ser Gln Val Ser Gln Asn Cys His Asn
20 25 30

Gly Ser Tyr Glu Ile Ser Val Leu Met Met Gly Asn Ser Ala Phe Ala
35 40 45

Glu Pro Leu Lys Asn Leu Glu Asp Ala Val Asn Glu Gly Leu Glu Ile
50 55 60

Val Arg Gly Arg Leu Gln Asn Ala Gly Leu Asn Val Thr Val Asn Ala
65 70 75 80

Thr Phe Met Tyr Ser Asp Gly Leu Ile His Asn Ser Gly Asp Cys Arg
85 90 95

Ser Ser Thr Cys Glu Gly Leu Asp Leu Leu Arg Lys Ile Ser Asn Ala
100 105 110

[0015]

Gln Arg Met Gly Cys Val Leu Ile Gly Pro Ser Cys Thr Tyr Ser Thr
115 120 125

Phe Gln Met Tyr Leu Asp Thr Glu Leu Ser Tyr Pro Met Ile Ser Ala
130 135 140

Gly Ser Phe Gly Leu Ser Cys Asp Tyr Lys Glu Thr Leu Thr Arg Leu
145 150 155 160

Met Ser Pro Ala Arg Lys Leu Met Tyr Phe Leu Val Asn Phe Trp Lys
165 170 175

Thr Asn Asp Leu Pro Phe Lys Thr Tyr Ser Trp Ser Thr Ser Tyr Val
180 185 190

Tyr Lys Asn Gly Thr Glu Thr Glu Asp Cys Phe Trp Tyr Leu Asn Ala
195 200 205

Leu Glu Ala Ser Val Ser Tyr Phe Ser His Glu Leu Gly Phe Lys Val
210 215 220

Val Leu Arg Gln Asp Lys Glu Phe Gln Asp Ile Leu Met Asp His Asn
225 230 235 240

Arg Lys Ser Asn Val Thr Ser Thr Trp Arg Thr Met Ser Gln Pro Leu
245 250 255

Thr Ile

<210> 14

<211> 1070

<212> PRT

<213> 人

<400> 14

Met Lys Thr Leu Leu Leu Asp Leu Ala Leu Trp Ser Leu Leu Phe Gln
1 5 10 15

Pro Gly Trp Leu Ser Phe Ser Ser Gln Val Ser Gln Asn Cys His Asn
20 25 30

Gly Ser Tyr Glu Ile Ser Val Leu Met Met Gly Asn Ser Ala Phe Ala
35 40 45

Glu Pro Leu Lys Asn Leu Glu Asp Ala Val Asn Glu Gly Leu Glu Ile
50 55 60

Val Arg Gly Arg Leu Gln Asn Ala Gly Leu Asn Val Thr Val Asn Ala
65 70 75 80

Thr Phe Met Tyr Ser Asp Gly Leu Ile His Asn Ser Gly Asp Cys Arg
85 90 95

[0016]

Ser Ser Thr Cys Glu Gly Leu Asp Leu Leu Arg Lys Ile Ser Asn Ala
 100 105 110
 Gln Arg Met Gly Cys Val Leu Ile Gly Pro Ser Cys Thr Tyr Ser Thr
 115 120 125
 Phe Gln Met Tyr Leu Asp Thr Glu Leu Ser Tyr Pro Met Ile Ser Ala
 130 135 140
 Gly Ser Phe Gly Leu Ser Cys Asp Tyr Lys Glu Thr Leu Thr Arg Leu
 145 150 155 160
 Met Ser Pro Ala Arg Lys Leu Met Tyr Phe Leu Val Asn Phe Trp Lys
 165 170 175
 Thr Asn Asp Leu Pro Phe Lys Thr Tyr Ser Trp Ser Thr Ser Tyr Val
 180 185 190
 Tyr Lys Asn Gly Thr Glu Thr Glu Asp Cys Phe Trp Tyr Leu Asn Ala
 195 200 205
 Leu Glu Ala Ser Val Ser Tyr Phe Ser His Glu Leu Gly Phe Lys Val
 210 215 220
 Val Leu Arg Gln Asp Lys Glu Phe Gln Asp Ile Leu Met Asp His Asn
 225 230 235 240
 Arg Lys Ser Asn Val Ile Ile Met Cys Gly Gly Pro Glu Phe Leu Tyr
 245 250 255
 Lys Leu Lys Gly Asp Arg Ala Val Ala Glu Asp Ile Val Ile Ile Leu
 260 265 270
 Val Asp Leu Phe Asn Asp Gln Tyr Leu Glu Asp Asn Val Thr Ala Pro
 275 280 285
 Asp Tyr Met Lys Asn Val Leu Val Leu Thr Leu Ser Pro Gly Asn Ser
 290 295 300
 Leu Leu Asn Ser Ser Phe Ser Arg Asn Leu Ser Pro Thr Lys Arg Asp
 305 310 315 320
 Phe Ala Leu Ala Tyr Leu Asn Gly Ile Leu Leu Phe Gly His Met Leu
 325 330 335
 Lys Ile Phe Leu Glu Asn Gly Glu Asn Ile Thr Thr Pro Lys Phe Ala
 340 345 350
 His Ala Phe Arg Asn Leu Thr Phe Glu Gly Tyr Asp Gly Pro Val Thr
 355 360 365
 Leu Asp Asp Trp Gly Asp Val Asp Ser Thr Met Val Leu Leu Tyr Thr
 370 375 380
 Ser Val Asp Thr Lys Lys Tyr Lys Val Leu Leu Thr Tyr Asp Thr His
 385 390 395 400

[0017]

Val Asn Lys Thr Tyr Pro Val Asp Met Ser Pro Thr Phe Thr Trp Lys
 405 410 415
 Asn Ser Lys Leu Pro Asn Asp Ile Thr Gly Arg Gly Pro Gln Ile Leu
 420 425 430
 Met Ile Ala Val Phe Thr Leu Thr Gly Ala Val Val Leu Leu Leu Leu
 435 440 445
 Val Ala Leu Leu Met Leu Arg Lys Tyr Arg Lys Asp Tyr Glu Leu Arg
 450 455 460
 Gln Lys Lys Trp Ser His Ile Pro Pro Glu Asn Ile Phe Pro Leu Glu
 465 470 475 480
 Thr Asn Glu Thr Asn His Val Ser Leu Lys Ile Asp Asp Asp Lys Arg
 485 490 495
 Arg Asp Thr Ile Gln Arg Leu Arg Gln Cys Lys Tyr Asp Lys Lys Arg
 500 505 510
 Val Ile Leu Lys Asp Leu Lys His Asn Asp Gly Asn Phe Thr Glu Lys
 515 520 525
 Gln Lys Ile Glu Leu Asn Lys Ile Asp Tyr Tyr Asn Leu Thr Lys Phe
 530 535 540
 Tyr Gly Thr Val Lys Leu Asp Thr Met Ile Phe Gly Val Ile Glu Tyr
 545 550 555 560
 Cys Glu Arg Gly Ser Leu Arg Glu Val Leu Asn Asp Thr Ile Ser Tyr
 565 570 575
 Pro Asp Gly Thr Phe Met Asp Trp Glu Phe Lys Ile Ser Val Leu Tyr
 580 585 590
 Asp Ile Ala Lys Gly Met Ser Tyr Leu His Ser Ser Lys Thr Glu Val
 595 600 605
 His Gly Arg Leu Lys Ser Thr Asn Cys Val Val Asp Ser Arg Met Val
 610 615 620
 Val Lys Ile Thr Asp Phe Gly Cys Asn Ser Ile Leu Pro Pro Lys Lys
 625 630 635 640
 Asp Leu Trp Thr Ala Pro Glu His Leu Arg Gln Ala Asn Ile Ser Gln
 645 650 655
 Lys Gly Asp Val Tyr Ser Tyr Gly Ile Ile Ala Gln Glu Ile Ile Leu
 660 665 670
 Arg Lys Glu Thr Phe Tyr Thr Leu Ser Cys Arg Asp Arg Asn Glu Lys
 675 680 685
 Ile Phe Arg Val Glu Asn Ser Asn Gly Met Lys Pro Phe Arg Pro Asp
 690 695 700

[0018]

```

Leu Phe Leu Glu Thr Ala Glu Glu Lys Glu Leu Glu Val Tyr Leu Leu
705                      710                      715                      720

Val Lys Asn Cys Trp Glu Glu Asp Pro Glu Lys Arg Pro Asp Phe Lys
725                      730                      735

Lys Ile Glu Thr Thr Leu Ala Lys Ile Phe Gly Leu Phe His Asp Gln
740                      745                      750

Lys Asn Glu Ser Tyr Met Asp Thr Leu Ile Arg Arg Leu Gln Leu Tyr
755                      760                      765

Ser Arg Asn Leu Glu His Leu Val Glu Glu Arg Thr Gln Leu Tyr Lys
770                      775                      780

Ala Glu Arg Asp Arg Ala Asp Arg Leu Asn Phe Met Leu Leu Pro Arg
785                      790                      795                      800

Leu Val Val Lys Ser Leu Lys Glu Lys Gly Phe Val Glu Pro Glu Leu
805                      810                      815

Tyr Glu Glu Val Thr Ile Tyr Phe Ser Asp Ile Val Gly Phe Thr Thr
820                      825                      830

Ile Cys Lys Tyr Ser Thr Pro Met Glu Val Val Asp Met Leu Asn Asp
835                      840                      845

Ile Tyr Lys Ser Phe Asp His Ile Val Asp His His Asp Val Tyr Lys
850                      855                      860

Val Glu Thr Ile Gly Asp Ala Tyr Met Val Ala Ser Gly Leu Pro Lys
865                      870                      875                      880

Arg Asn Gly Asn Arg His Ala Ile Asp Ile Ala Lys Met Ala Leu Glu
885                      890                      895

Ile Leu Ser Phe Met Gly Thr Phe Glu Leu Glu His Leu Pro Gly Leu
900                      905                      910

Pro Ile Trp Ile Arg Ile Gly Val His Ser Gly Pro Cys Ala Ala Gly
915                      920                      925

Val Val Gly Ile Lys Met Pro Arg Tyr Cys Leu Phe Gly Asp Thr Val
930                      935                      940

Asn Thr Ala Ser Arg Met Glu Ser Thr Gly Leu Pro Leu Arg Ile His
945                      950                      955                      960

Val Ser Gly Ser Thr Ile Ala Ile Leu Lys Arg Thr Glu Cys Gln Phe
965                      970                      975

Leu Tyr Glu Val Arg Gly Glu Thr Tyr Leu Lys Gly Arg Gly Asn Glu
980                      985                      990

Thr Thr Tyr Trp Leu Thr Gly Met Lys Asp Gln Lys Phe Asn Leu Pro
995                      1000                      1005

```

[0019]

Thr Pro Pro Thr Val Glu Asn Gln Gln Arg Leu Gln Ala Glu Phe
1010 1015 1020

Ser Asp Met Ile Ala Asn Ser Leu Gln Lys Arg Gln Ala Ala Gly
1025 1030 1035

Ile Arg Ser Gln Lys Pro Arg Arg Val Ala Ser Tyr Lys Lys Gly
1040 1045 1050

Thr Leu Glu Tyr Leu Gln Leu Asn Thr Thr Asp Lys Glu Ser Thr
1055 1060 1065

Tyr Phe
1070

<210> 15

<211> 93

<212> PRT

<213> 人

<400> 15

Met Lys Leu Val Thr Ile Phe Leu Leu Val Thr Ile Ser Leu Cys Ser
1 5 10 15

Tyr Ser Ala Thr Ala Lys Leu Ile Asn Lys Cys Pro Leu Pro Val Asp
20 25 30

Lys Leu Ala Pro Leu Pro Leu Asp Asn Ile Leu Pro Phe Met Asp Pro
35 40 45

Leu Lys Leu Leu Leu Lys Thr Leu Gly Ile Ser Val Glu His Leu Val
50 55 60

Glu Gly Leu Arg Lys Cys Val Asn Glu Leu Gly Pro Glu Ala Ser Glu
65 70 75 80

Ala Val Lys Lys Leu Leu Glu Ala Leu Ser His Leu Val
85 90

<210> 16

<211> 261

<212> PRT

<213> 人

<400> 16

Met Ala Val Thr Ala Cys Gln Gly Leu Gly Phe Val Val Ser Leu Ile
1 5 10 15

Gly Ile Ala Gly Ile Ile Ala Ala Thr Cys Met Asp Gln Trp Ser Thr
20 25 30

[0020]

Gln	Asp	Leu	Tyr	Asn	Asn	Pro	Val	Thr	Ala	Val	Phe	Asn	Tyr	Gln	Gly	
	35						40					45				
Leu	Trp	Arg	Ser	Cys	Val	Arg	Glu	Ser	Ser	Gly	Phe	Thr	Glu	Cys	Arg	
	50					55					60					
Gly	Tyr	Phe	Thr	Leu	Leu	Gly	Leu	Pro	Ala	Met	Leu	Gln	Ala	Val	Arg	
65					70					75					80	
Ala	Leu	Met	Ile	Val	Gly	Ile	Val	Leu	Gly	Ala	Ile	Gly	Leu	Leu	Val	
			85						90					95		
Ser	Ile	Phe	Ala	Leu	Lys	Cys	Ile	Arg	Ile	Gly	Ser	Met	Glu	Asp	Ser	
			100					105					110			
Ala	Lys	Ala	Asn	Met	Thr	Leu	Thr	Ser	Gly	Ile	Met	Phe	Ile	Val	Ser	
		115					120					125				
Gly	Leu	Cys	Ala	Ile	Ala	Gly	Val	Ser	Val	Phe	Ala	Asn	Met	Leu	Val	
	130					135					140					
Thr	Asn	Phe	Trp	Met	Ser	Thr	Ala	Asn	Met	Tyr	Thr	Gly	Met	Gly	Gly	
145					150					155					160	
Met	Val	Gln	Thr	Val	Gln	Thr	Arg	Tyr	Thr	Phe	Gly	Ala	Ala	Leu	Phe	
				165					170						175	
Val	Gly	Trp	Val	Ala	Gly	Gly	Leu	Thr	Leu	Ile	Gly	Gly	Val	Met	Met	
			180					185						190		
Cys	Ile	Ala	Cys	Arg	Gly	Leu	Ala	Pro	Glu	Glu	Thr	Asn	Tyr	Lys	Ala	
		195					200					205				
Val	Ser	Tyr	His	Ala	Ser	Gly	His	Ser	Val	Ala	Tyr	Lys	Pro	Gly	Gly	
		210				215					220					
Phe	Lys	Ala	Ser	Thr	Gly	Phe	Gly	Ser	Asn	Thr	Lys	Asn	Lys	Lys	Ile	
225					230					235					240	
Tyr	Asp	Gly	Gly	Ala	Arg	Thr	Glu	Asp	Glu	Val	Gln	Ser	Tyr	Pro	Ser	
				245					250					255		
Lys	His	Asp	Tyr	Val												
			260													
<210>	17															
<211>	10															
<212>	PRT															
<213>	人															
<400>	17															
Asp	Gln	Trp	Ser	Thr	Gln	Asp	Leu	Tyr	Asn							
1				5					10							

[0021]

<210> 18

<211> 11

<212> PRT

<213> 人

<400> 18

Asn Asn Pro Val Thr Ala Val Phe Asn Tyr Gln
1 5 10

<210> 19

<211> 47

<212> PRT

<213> 人

<400> 19

Met Ala Val Thr Ala Cys Gln Gly Leu Gly Phe Val Val Ser Leu Ile
1 5 10 15

Gly Ile Ala Gly Ile Ile Ala Ala Thr Cys Met Asp Gln Trp Ser Thr
20 25 30

Gln Asp Leu Tyr Asn Asn Pro Val Thr Ala Val Phe Asn Tyr Gln
35 40 45

<210> 20

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: 寡核苷酸

<400> 20

aggtacatga gcatcagcct g 21

<210> 21

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: 寡核苷酸

<400> 21

gcagcagttg gcatctgaga g 21

[0022]

<210>	22	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	22	
	gcaatagaca ttgccaagat g	21
<210>	23	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	23	
	aacgctgttg attctccaca g	21
<210>	24	
<211>	33	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	24	
	ggatcctcct ttagttecca ggtgagtcag aac	33
<210>	25	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	25	
	tgctctggag gctagcgttt c	21
<210>	26	

[0023]

<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	26	
	accaatcatg ttagcctcaa g	21
<210>	27	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	27	
	agctatggga tcategcaca g	21
<210>	28	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	28	
	cctttgagct ggagcatett c	21
<210>	29	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	29	
	ctttctagct ggagacatca g	21
<210>	30	
<211>	21	
<212>	DNA	

[0024]

<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	30	
	caccatggta ctgtcaacat c	21
<210>	31	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	31	
	atgtcataca agacagagat c	21
<210>	32	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	32	
	tctgccttgt acagctgtgt c	21
<210>	33	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	33	
	tctgtgttat tcagctgcaa g	21
<210>	34	
<211>	22	
<212>	DNA	
<213>	人工序列	
<220>		

[0025]

<223> 人工序列描述：寡核苷酸

<400> 34
tactcaggaa aatttcacct tg 22

<210> 35

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：寡核苷酸

<400> 35
gaccacaaca ggaaaagcaa tgtgacc 27

<210> 36

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：寡核苷酸

<400> 36
gatagaattg aacaagattg ac 22

<210> 37

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：寡核苷酸

<400> 37
cagcctttgt agttactctg c 21

<210> 38

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：寡核苷酸

[0026]

<400> 38 tgtcacacca agtgtgatag c	21
<210> 39	
<211> 28	
<212> DNA	
<213> 人工序列	
<220> <223> 人工序列描述: 寡核苷酸	
<400> 39 ggttcgtggt ttcactgatt gggattgc	28
<210> 40	
<211> 27	
<212> DNA	
<213> 人工序列	
<220> <223> 人工序列描述: 寡核苷酸	
<400> 40 cggttttgta gttggtttct tctggtg	27
<210> 41	
<211> 3814	
<212> DNA	
<213> 人	
<400> 41 ctattgaagc cacctgctca ggacaatgaa attcttcagt tacattctgg tttatcgccg atttctcttc gtggttttca ctgtgttggt tttactacct ctgcccatcg tctccacac caaggaagca gaatgtgcct acacactett tgtggtegcc acattttggc tcacagaagc attgcctctg tcggtaacag ctttgcctacc tagtttaatg ttaccatgt ttgggatcat gccttetaag aaggtggcat ctgcttattt caaggatttt cacttactgc taattggagt tatctgttta gcaacatcca tagaaaaatg gaatttgcac aagagaattg ctctgaaaat ggatgatgat gttggtgtaa atcttgcctg gctgacgctg gggttcatga gcagcactgc cttttctct atgtggctca gcaacacctc gacggctgcc atggatgac ccattgcgga ggctgtagtg cagcagatca tcaatgcaga agcagaggte gaggccactc agatgactta cttcaacgga tcaaccaacc acggactaga aattgatgaa agtggttaatg gacatgaaat aatgagagg aaagagaaaa caaaaccagt tccaggatac aataatgata cagggaaaaat ttcaagcaag gtggagtggg aaaagaactc aggcattgaga accaaatatc gaacaaagaa gggccacgtg acacgtaaac ttacgtgttt gtgcattgcc tactcttcta ccattgggtg	60 120 180 240 300 360 420 480 540 600 660 720 780

[0027]

actgacaaca atcactggta cctccaccaa cttgatcttt gcagagtatt tcaatacacg	840
ctatccigac tglcglgccc tcaactllgg atcatgggtt acgtlllccct lcccagcigc	900
ccttatcatt ctactcttat cctggatctg gcttcagtgg cttttcttag gattcaattt	960
taaggagatg ttcaaatgtg gcaaaaccaa aacagtccea caaaaagctt gtgctgaggt	1020
gattaagcaa gaataccaaa agcttgggcc aataaggatg caagaaattg tgacctgggt	1080
cctcttcatt ataattggctc tgctatgggt tagtcgagac cccggatttg ttcctgggtg	1140
gtctgcactt ttttcagagt accctgggtt tgctacagat tcaactgttg ctttacttat	1200
agggtcgeta ttctttctta tcccagctaa gacactgact aaaactacac ctacaggaga	1260
aattglgct lllgallact ctcactgat tactlggaaa gaattccagt catlcatgcc	1320
ctgggatata gccattcttg ttggtggagg gtttgcctg gcagatgggt gtgaggagtc	1380
tggattatct aagtggatag gaaataaatt atctcctctg gggttcattac cagcatggct	1440
aataattctg atatctctct tgatggtgac atctttaact gaggtagcca gcaatccagc	1500
taccattaca ctctttctcc caatattatc tccattggcc gaagccattc atgtgaaccc	1560
tctttatatt ctgatacctt ctactctgtg tacttcattt gcattcctcc taccagtagc	1620
aaatccaccc aatgctattg tcttttctta tggtcctctg aaagtcattg acatggttaa	1680
agctggactt ggtgtcaaca ttgttggtgt tgctgtggtt atgcttggca tatgtacttg	1740
gattgtaccc atgtttgacc tctacactta ccttctgtgg gctcctgcta tgagtaatga	1800
gaccatgcca taataagcac aaaatttctg actatcttgc ggtaatttct ggaagacatt	1860
aatgattgac tgtaaaatgt ggctctaaat aactaatgac acacatttaa atcagttatg	1920
gtgtagctgc tgcaattccc gtgaataccc gaaacctgct ggtataactc agagtccata	1980
tttgttattg cagtgcacct aaagagcatc tatgtgcctt catcaagaag cccatgtttt	2040
gagattttgc tcatgaacca tctgcaactt gettcatcat aagaataatt tataacttga	2100
ccttcaaaga gattagagca ttgttttcat cttacagttg gagttcaatg taacatttta	2160
aatgcaattt attatttcag aaatttccca tgaaactaaa aatagaaaat aagatataca	2220
agttaattcg gtacttggat aaatcatttc tgcattgttg ttccagagaa ttgtctgaga	2280
aatcaaagcc atgggtcatct ggtgatgaag agaaaaggtt aatctaaatg atatgtgcat	2340
ttctctattt aaaaaatcca attggattat tcttaataata tacatgtaat atgaaaattg	2400
agattgaagc actaattcca aaattatgac tgaatatact aaataacaga aaagttacag	2460
ataagaattt atttctactg aactctatag ttagtgtaat ataattcata tttttatgat	2520
attggcacac tgagaaatc attttgtaga gctatggata aggcttgcta tgatttgcac	2580
tattagtaca gtatagttag aaaggaaagc tgaacactat aaaactatta acatatttct	2640
gtatatgagt aacaactttg cttaatgtgt tatcttagtt cagaaataca taatgtcata	2700
tgttaaaaat aaagagatgt agaaatctaa atgaattatc actgtgtata cagacagaaa	2760
aatcacataa ctctgggtgtg ttaacattgc aatgaaaaaa tgaaaaaaag aaggaaaaaa	2820
gaataagaat gaaaactgct gacgtattac aaaacagaaa aataaatgat ttaaaatcaa	2880
atcaaaaaga aaaaaactaa acatttaaac aaaaatggga taagaatagt cttctagaag	2940
tgaggatgcg taaaagaatg agtttccaat taccctgatg tgacaattac acattgtaga	3000
caggtagcaa aatatcacat acaccccaaa aatatgtaca aatattatat atcaataaat	3060

[0028]

```

aaatttttaa agagtaagtg ctattggcat tccaaaattc agctaaagga aaaatgatca 3120
aaaacaaagt aaggtgcaca gttagcaaaa gatgcagatg ttatatcaca gcaattctca 3180
tgctaaaaat acaacaaaag acaaagcaaa aaataaacct ttgctttttt tttttttttt 3240
tttttttttt gagacggagc ctcgcctcgt cgcacaggct ggagtcagc gccgggaccc 3300
cggctcactg caagctccgc ctcccagggt cagcccatc tcctgcctca gccaaacctt 3360
tgctattttt aatcttcgtt ggcactttcc agctgttact gaccttgca tttttgttc 3420
aaataagatt atttacaac ttattcttga aactaaatat agtaaagagg gtttttaaaa 3480
taatatttaa catacgaatt attaattggc catgttcatt atttatctat gtttattaat 3540
gggccaatgc aaaaaatcat tttttcaaag aaaaatttgt ccatgtaaag cttaaattat 3600
aatattgctg ctttgtataa ctcttctatg tttattctat tcatttgctt ctttcctac 3660
catattttac acatgtatgt ataactgtga gtatttatta cattctgctt ttttctagt 3720
cattcaattt atcactgctg aattgcatca gatcatggat gcatttttat tatgaaaaaa 3780
taaaatgact ttcaaatta aaaaaaaaaa aaaa 3814

```

<210> 42

<211> 734

<212> DNA

<213> 人

<400> 42

```

caggacaatg aaattcttca gttacattct ggtttatcgc cgatttctct tcgtggtttt 60
cactgtgttg gttttactac ctctgcccat cgtcctccac accaaggaag cagaatgtgc 120
ctacacactc tttgtggtcg ccacattttg gtcacagaa gcattgectc tgcggtaac 180
agctttgcta cctagttaa tgttaccat gtttgggac atgccttcta agaagggtgc 240
atctgcttat ttcaaggatt ttcaattact gctaattgga gttatctgtt tagcaacatc 300
catagaaaaa tggaatttgc acaagagaat tgctctgaaa atggatgatg tggttggtgt 360
aaatcctgca tggctgacgc tgggttcat gacgagcact gcctttttgt ctatgtggct 420
cagcaacacc tcgacggctg ccatggtgat gccattgcg gaggtgtag tgcagcagat 480
catcaatgca gaagcagagg tcgagccac tcagatgact tacttcaacg gatcaaccaa 540
ccacggacta gaaattgatg aaagtgttaa tggacatgaa ataatgaga ggaaagagaa 600
aacaaaacca gttccaggat acaataatga tacagggaaa atttcaagca aggtggagt 660
ggaaaagact gttaactac tgaatgaag ctattctcct gactaaacat aactgaaaaa 720
ccattcatta aatg 734

```

<210> 43

<211> 539

<212> DNA

<213> 人

[0029]

<400> 43
 gccactcaga tgacttactt caacggatca accaaccacg gactagaaat tgatgaaagt 60
 gttaatggac atgaaataaa tgagaggaaa gagaaaacaa aaccagtcc aggatacaat 120
 aatgatacag ggaaaatttc aagcaagggtg gagttggaaa agcactggaa acttcagtt 180
 caagatggtt ccccatctcc ctctgtccat tctgtatcgc agctagctgc tcaaggaaag 240
 gagaaagtg aagcatatg tacttagaaa ttattctatt actttcctgg atttaagagt 300
 attcagattt tetatttcaa cateaacaat ttgcattttt aaaaagaaat ttatgtgttc 360
 catgtcaaat ttagtagtgt gtggtgtttt ataatatatt cttatatcta cttatatttc 420
 atagtattta tagttatatg tctttatttc taacattttt cttgtgcttt taaagattat 480
 ttaaagatta tttttaaata atctttattt catttaaata aaatatttta ttttaagtct 539

<210> 44

<211> 556

<212> DNA

<213> 人

<400> 44
 caccgactag aaattgatga aagtgttaat ggacatgaaa taaatgagag gaaagagaaa 60
 aaaaaaccag ttccaggata caataatgat acagggaaaa ttccaagcaa ggtggagtgg 120
 gaaaagaact caggcatgag aaccaaatat cgaacaaaga aggccacgt gacacgtaaa 180
 cttacgtgtt tgtgcattgc ctactcttct accattgggtg gactgacaac aatcactggt 240
 acctccacca acttgatctt tgcagagtat ttcaatacat tccatccaca cagaagagga 300
 gatcgtacaa ggcatgtaca ccaggaggca gaaatttgag gcatactctg gaactctgtc 360
 taccacatcc tgaacatcac acagtttcca ctcttggtgc cttcaatcct gagaatgcatt 420
 ccaggagcca tctgttttta tgtcaattac taattagatc atgtcagtt actaacttac 480
 tacgttccaa ttagtcctta ttgcatttgt aataaaatcc gcatacttcc ggactggcta 540
 caaggttata catgat 556

<210> 45

<211> 595

<212> PRT

<213> 人

<400> 45

Met Lys Phe Phe Ser Tyr Ile Leu Val Tyr Arg Arg Phe Leu Phe Val
 1 5 10 15

Val Phe Thr Val Leu Val Leu Leu Pro Leu Pro Ile Val Leu His Thr
 20 25 30

Lys Glu Ala Glu Cys Ala Tyr Thr Leu Phe Val Val Ala Thr Phe Trp
 35 40 45

[0030]

Leu Thr Glu Ala Leu Pro Leu Ser Val Thr Ala Leu Leu Pro Ser Leu
 50 55 60

Met Leu Pro Met Phe Gly Ile Met Pro Ser Lys Lys Val Ala Ser Ala
 65 70 75 80

Tyr Phe Lys Asp Phe His Leu Leu Leu Ile Gly Val Ile Cys Leu Ala
 85 90 95

Thr Ser Ile Glu Lys Trp Asn Leu His Lys Arg Ile Ala Leu Lys Met
 100 105 110

Val Met Met Val Gly Val Asn Pro Ala Trp Leu Thr Leu Gly Phe Met
 115 120 125

Ser Ser Thr Ala Phe Leu Ser Met Trp Leu Ser Asn Thr Ser Thr Ala
 130 135 140

Ala Met Val Met Pro Ile Ala Glu Ala Val Val Gln Gln Ile Ile Asn
 145 150 155 160

Ala Glu Ala Glu Val Glu Ala Thr Gln Met Thr Tyr Phe Asn Gly Ser
 165 170 175

Thr Asn His Gly Leu Glu Ile Asp Glu Ser Val Asn Gly His Glu Ile
 180 185 190

Asn Glu Arg Lys Glu Lys Thr Lys Pro Val Pro Gly Tyr Asn Asn Asp
 195 200 205

Thr Gly Lys Ile Ser Ser Lys Val Glu Leu Glu Lys Asn Ser Gly Met
 210 215 220

Arg Thr Lys Tyr Arg Thr Lys Lys Gly His Val Thr Arg Lys Leu Thr
 225 230 235 240

Cys Leu Cys Ile Ala Tyr Ser Ser Thr Ile Gly Gly Leu Thr Thr Ile
 245 250 255

Thr Gly Thr Ser Thr Asn Leu Ile Phe Ala Glu Tyr Phe Asn Thr Arg
 260 265 270

Tyr Pro Asp Cys Arg Cys Leu Asn Phe Gly Ser Trp Phe Thr Phe Ser
 275 280 285

Phe Pro Ala Ala Leu Ile Ile Leu Leu Leu Ser Trp Ile Trp Leu Gln
 290 295 300

Trp Leu Phe Leu Gly Phe Asn Phe Lys Glu Met Phe Lys Cys Gly Lys
 305 310 315 320

Thr Lys Thr Val Gln Gln Lys Ala Cys Ala Glu Val Ile Lys Gln Glu
 325 330 335

Tyr Gln Lys Leu Gly Pro Ile Arg Tyr Gln Glu Ile Val Thr Leu Val
 340 345 350

[0031]

Leu Phe Ile Ile Met Ala Leu Leu Trp Phe Ser Arg Asp Pro Gly Phe
355 360 365

Val Pro Gly Trp Ser Ala Leu Phe Ser Glu Tyr Pro Gly Phe Ala Thr
370 375 380

Asp Ser Thr Val Ala Leu Leu Ile Gly Leu Leu Phe Phe Leu Ile Pro
385 390 395 400

Ala Lys Thr Leu Thr Lys Thr Thr Pro Thr Gly Glu Ile Val Ala Phe
405 410 415

Asp Tyr Ser Pro Leu Ile Thr Trp Lys Glu Phe Gln Ser Phe Met Pro
420 425 430

Trp Asp Ile Ala Ile Leu Val Gly Gly Gly Phe Ala Leu Ala Asp Gly
435 440 445

Cys Glu Glu Ser Gly Leu Ser Lys Trp Ile Gly Asn Lys Leu Ser Pro
450 455 460

Leu Gly Ser Leu Pro Ala Trp Leu Ile Ile Leu Ile Ser Ser Leu Met
465 470 475 480

Val Thr Ser Leu Thr Glu Val Ala Ser Asn Pro Ala Thr Ile Thr Leu
485 490 495

Phe Leu Pro Ile Leu Ser Pro Leu Ala Glu Ala Ile His Val Asn Pro
500 505 510

Leu Tyr Ile Leu Ile Pro Ser Thr Leu Cys Thr Ser Phe Ala Phe Leu
515 520 525

Leu Pro Val Ala Asn Pro Pro Asn Ala Ile Val Phe Ser Tyr Gly His
530 535 540

Leu Lys Val Ile Asp Met Val Lys Ala Gly Leu Gly Val Asn Ile Val
545 550 555 560

Gly Val Ala Val Val Met Leu Gly Ile Cys Thr Trp Ile Val Pro Met
565 570 575

Phe Asp Leu Tyr Thr Tyr Pro Ser Trp Ala Pro Ala Met Ser Asn Glu
580 585 590

Thr Met Pro
595

<210> 46

<211> 224

<212> PRT

<213> 人

<400> 46

[0032]

Arg Thr Met Lys Phe Phe Ser Tyr Ile Leu Val Tyr Arg Arg Phe Leu
 1 5 10 15
 Phe Val Val Phe Thr Val Leu Val Leu Leu Pro Leu Pro Ile Val Leu
 20 25 30
 His Thr Lys Glu Ala Glu Cys Ala Tyr Thr Leu Phe Val Val Ala Thr
 35 40 45
 Phe Trp Leu Thr Glu Ala Leu Pro Leu Ser Val Thr Ala Leu Leu Pro
 50 55 60
 Ser Leu Met Leu Pro Met Phe Gly Ile Met Pro Ser Lys Lys Val Ala
 65 70 75 80
 Ser Ala Tyr Phe Lys Asp Phe His Leu Leu Leu Ile Gly Val Ile Cys
 85 90 95
 Leu Ala Thr Ser Ile Glu Lys Trp Asn Leu His Lys Arg Ile Ala Leu
 100 105 110
 Lys Met Val Met Met Val Gly Val Asn Pro Ala Trp Leu Thr Leu Gly
 115 120 125
 Phe Met Ser Ser Thr Ala Phe Leu Ser Met Trp Leu Ser Asn Thr Ser
 130 135 140
 Thr Ala Ala Met Val Met Pro Ile Ala Glu Ala Val Val Gln Gln Ile
 145 150 155 160
 Ile Asn Ala Glu Ala Glu Val Glu Ala Thr Gln Met Thr Tyr Phe Asn
 165 170 175
 Gly Ser Thr Asn His Gly Leu Glu Ile Asp Glu Ser Val Asn Gly His
 180 185 190
 Glu Ile Asn Glu Arg Lys Glu Lys Thr Lys Pro Val Pro Gly Tyr Asn
 195 200 205
 Asn Asp Thr Gly Lys Ile Ser Ser Lys Val Glu Leu Glu Lys Thr Val
 210 215 220
 <210> 47
 <211> 88
 <212> PRT
 <213> 人
 <400> 47
 Ala Thr Gln Met Thr Tyr Phe Asn Gly Ser Thr Asn His Gly Leu Glu
 1 5 10 15
 Ile Asp Glu Ser Val Asn Gly His Glu Ile Asn Glu Arg Lys Glu Lys
 20 25 30

[0033]

Thr Lys Pro Val Pro Gly Tyr Asn Asn Asp Thr Gly Lys Ile Ser Ser
35 40 45

Lys Val Glu Leu Glu Lys His Trp Lys Leu Ala Val Gln Asp Gly Ser
50 55 60

Pro Ser Pro Ser Val His Ser Val Ser Gln Leu Ala Ala Gln Gly Lys
65 70 75 80

Glu Lys Val Glu Gly Ile Cys Thr
85

<210> 48

<211> 112

<212> PRT

<213> 人

<400> 48

Ile Gly Leu Glu Ile Asp Glu Ser Val Asn Gly His Glu Ile Asn Glu
1 5 10 15

Arg Lys Glu Lys Thr Lys Pro Val Pro Gly Tyr Asn Asn Asp Thr Gly
20 25 30

Lys Ile Ser Ser Lys Val Glu Leu Glu Lys Asn Ser Gly Met Arg Thr
35 40 45

Lys Tyr Arg Thr Lys Lys Gly His Val Thr Arg Lys Leu Thr Cys Leu
50 55 60

Cys Ile Ala Tyr Ser Ser Thr Ile Gly Gly Leu Thr Thr Ile Thr Gly
65 70 75 80

Thr Ser Thr Asn Leu Ile Phe Ala Glu Tyr Phe Asn Thr Phe His Pro
85 90 95

His Arg Arg Gly Asp Arg Thr Arg His Val His Gln Glu Ala Glu Ile
100 105 110

<210> 49

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：寡核苷酸

<400> 49
ccagctttaa ccatgtcaat g

21

<210> 50

[0034]

<211> 21	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述: 寡核苷酸	
<400> 50	
cagatggttg tgaggagtct g	21
<210> 51	
<211> 3311	
<212> DNA	
<213> 人	
<400> 51	
tgctaattgct ttggtacaa atggatgtgg aatataattg aatattttct tgtttaagg	60
gagcatgaag aggtgttgag gttatgtcaa gcatctggca cagctgaagg cagatggaaa	120
tatttacaag tacgaattt gagactaaga tattgttate atttcctat tgaagacaag	180
agcaatagta aaacacatca ggtcaggggg ttaaagacct gtgataaacc acttcgata	240
agttggaaac gtgtgtctat attttcatat ctgtatata ataattggtaa agaaagacac	300
cttcgtaacc cgcattttcc aaagagagga atcacaggga gatgtacagc aatggggcca	360
tttaagagtt ctgtgttcat cttgattctt caccctctag aaggggccct gagtaattca	420
ctcattcagc tgaacaacaa tggctatgaa ggcattgtcg ttgcaatcga ccccaatgtg	480
ccagaagatg aaacactcat tcaacaata aaggacatgg tgaccagge atctctgtat	540
ctgtttgaag ctacaggaaa gcgattttat ttcaaaaatg ttgccatttt gattcctgaa	600
acatggaaga caaaggctga ctatgtgaga ccaaaacttg agacctacaa aatgctgat	660
gttctggttg ctgagttctac tctccaggt aatgatgaac cctacactga gcagatgggc	720
aactgtggag agaagggtga aaggatccac ctcactcctg atttcattgc agggaaaaag	780
ttagctgaat atggaccaca aggtaaggca tttgtccatg agtgggctca tctacgatgg	840
ggagtatttg acgagtacaa taatgatgag aaattctact tatecaatgg aagaatacaa	900
gcagtaagat gttcagcagg tattactggt acaaatgtag taaagaagtg tcagggagge	960
agctgttaca ccaaagatg cacattcaat aaagttacag gactctatga aaaaggatgt	1020
gagtttgttc tecaatcccg ccagacggag aaggettcta taatgtttgc acaacatgtt	1080
gattctatag ttgaattctg tacagaacaa aaccacaaca aagaagctcc aaacaagcaa	1140
aatcaaaaat geaatctccg aagcacatgg gaagtgatcc gtgattctga ggactttaag	1200
aaaaccactc ctatgacaac acagccacca aatcccact tctcattgct gcagattgga	1260
caaagaattg tgtgtttagt ccttgacaaa tctggaagca tggcgactgg taaccgctc	1320
aatcgactga atcaagcagg ccagcttttc ctgctgcaga cagttgagct ggggtcctgg	1380
gttgggatgg tgacatttga cagtgtgcc catgtacaaa gtgaactcat acagataaac	1440
agtggcagtg acagggacac actcgccaaa agattacctg cagcagcttc aggaggagc	1500

[0035]

```

tccatctgca gcgggcttcg atcggcattt actgtgatta ggaagaaata tccaactgat 1560
ggatctgaaa ttgtgtgct gacggatggg gaagacaaca ctataagtgg gtgctttaac 1620
gagggtcaaac aaagtgggtgc catcatccac acagtgcctt tggggccctc tgcagctcaa 1680
gaactagagg agctgtccaa aatgacagga ggtttacaga catatgcttc agatcaagtt 1740
cagaacaatg gcctcattga tgcctttggg gccctttcat caggaaatgg agctgtctct 1800
cagcgcacca tccagcttga gagtaaggga ttaacctcc agaacagcca gtcgatgaat 1860
ggcacagtga tcgtggacag caccgtggga aaggacactt tgtttcttat cacctggaca 1920
acgcagctc cccaaatect tctctgggat cccagtggac agaagcaagg tggctttgta 1980
gtggacaaaa acacaaaaat gccctacct caaatcccag gcattgctaa ggttggcact 2040
tggaataaca gtctgcaagc aagctcaca accctgacct tgactgtcac gtcctgtcgc 2100
tccaatgcta ccttgctcc aattacagt acttccaaa cgaacaagga caccagcaaa 2160
ttccccagcc ctctggtagt ttatgcaaat attcgcaag gagcctcccc aattctcagg 2220
gccagtgtca cagccctgat tgaatcagt aatggaaaa cagttacctt ggaactactg 2280
gataatggag caggtgtgta tctactaag gatgacggtg tctactcaag gtatttcaca 2340
acttatgaca cgaatggtag atacagtgt aaagtgcggg ctctgggagg agttaacga 2400
gccagacgga gagtgtacc ccagcagagt ggagcactgt acatacctgg ctggattgag 2460
aatgatgaaa tacaatggaa tccaccaaga cctgaaatta ataaggatga tgttcaacac 2520
aagcaagtgt gtttcagcag aacatcctc ggaggctcat ttgtggcttc tgatgtccca 2580
aatgtccca tacctgatct ctteccacct ggccaaatca ccgacctgaa ggccgaaatt 2640
cacgggggca gtctcattaa tctgacttgg acagctcctg gggatgatta tgacctgga 2700
acagctcaca agtatateat tcgaataagt acaagtattc ttgatctcag agacaagttc 2760
aatgaatctc ttcaagtga tactactgct ctcatccaa aggaagccaa ctctgaggaa 2820
gtctttttgt ttaaaccaga aaacattact ttgaaaatg gcacagatct ttctattgct 2880
attcaggctg ttgataaggt cgatctgaaa tcagaaatat ccaacattgc acgagtatct 2940
ttgtttattc ctccacagac tccgccagag acacctagtc ctgatgaac gtctgtcct 3000
tgtcctaata ttcataata cagcaccatt cctggcattc acattttaaa aattatgttg 3060
aagtggatag gagaactgca gctgtcaata gcctagggt gaatttttgt cagataaata 3120
aaataaatca ttcactcttt ttttgattat aaaattttct aaaatgtatt ttagacttcc 3180
tgtagggggc gatatactaa atgtatatag tacatttata ctaaatgtat tcctgtaggg 3240
ggcgatatac taaatgtatt ttagacttcc ttaggggggc gataaaataa aatgctaaac 3300
aactgggtaa a 3311

```

<210> 52

<211> 3067

<212> DNA

<213> 人

<400> 52

aattaaatta tgagaattaa aaagacaaca ttgagcagag atgaaaaagg aaggaggaa 60

[0036]

aaggtggaag	agaaaagaag	acaagaagcg	agtagtggtc	tctaacttgc	tctttgaagg	120
atggtctcac	aaagagaacc	ccaacagaca	tcacgtggg	aatcaaatca	agaccagcaa	180
gtacacgtg	tgtccttcg	tccccaaaa	catttttgag	cagctacace	ggtttgccaa	240
tctctatitt	gtgggcattg	cggttctgaa	ttttatccct	gtggccaatg	ctttccagcc	300
tgaggtgagc	atgataccaa	tctgtgttat	cctggcagtc	actgccaatc	aggacgcttg	360
ggaagacctc	cggaggtaca	aatcgataaa	agtcacatca	aaccgagagt	gcctcatcta	420
cagcagaaaa	gagcagacct	atgtgcagaa	gtgctggaag	gatgtgcgtg	tgggagacct	480
catccaaatg	aatgcaatg	agattgtccc	agcagacata	ctcctccttt	tttctcttga	540
ccccaatggg	atatgccatc	tggaaactgc	cagcttggtg	ggagagacaa	acctcaagca	600
aagacgtgtc	gtgaagggtc	tctcacagca	ggaggtacag	ttcgaaccag	agcttttcca	660
caataccatc	gtgtgtgaga	aacccaacaa	ccacctcaac	aaattttaag	gttatatgga	720
gcactctgac	cagaccagga	ctggcttttg	ctgtgagagt	cttctgcttc	gaggtgcac	780
catcagaaac	accgagatgg	ctgttgcat	tgtcatctat	gcaggccatg	agacgaaagc	840
catgtgaac	aacagtggcc	cccgtacaa	acgcagcaag	attgagcggc	gcattgaatat	900
agacatcttc	ttctgcatg	ggatctctat	cctcatgtgc	cttattggag	ctgtaggtca	960
cagcatclgg	aatgggacct	tlgaagaaca	cccctccttc	gatgtgccag	atgccaatgg	1020
cagcttcttc	cccagtgcct	ttgggggctt	ctacatgttc	ctcacaatga	tcactctgct	1080
ccaggtgctg	atcccatct	ctttgtatgt	ctccattgag	ctgggtgaagc	tegggcaagt	1140
gtctctcttg	agcaatgacc	tlgacctgta	lgatgaagag	accgalltat	ccattcaatg	1200
tcgagccctc	aacatgcag	aggacttggg	ccagatccag	tacatctctc	ccgataagac	1260
ggggaccttg	acagagaaca	agatgggtgt	cgcaggttgc	accatcatgg	gcagcgagta	1320
ttctacacaa	gaaaatggta	tagaagctcc	caagggtccc	atccctcttt	ctaaaaggaa	1380
atacctgct	ctcctaagaa	acgaggagat	aaaagacatt	ctcctggctc	tcttagaggc	1440
tgtgtggcat	ttccacaagt	tgettctgt	atccctgtgg	tcttctctgt	cacagatcag	1500
ggctgttcca	attacttgta	aactttcatt	tgtttacaaa	ggttagaagt	tatcccatat	1560
gtggttcccc	ttcagctgat	ctttgtctgg	tgcagacaa	agcactttat	gagacgagtt	1620
ttttatctgt	cagcaatgga	tggagacat	ttcccaattg	tgtgccagtc	acacaaccaa	1680
ggcttaggaa	tttctcaggc	caccttacct	gacatgtcag	ggcaggtctg	tgtctaggtg	1740
catggtcaga	tttaatacat	ccagaagatg	tctctatctc	taacagatct	cttagcttgt	1800
cactgaggca	aagttttgat	ttaggagata	gggtataaaa	atgcttgga	tgttaccttg	1860
catggactga	atatgactca	taaaactgat	ctgattcctt	cagccatcat	ctgcccact	1920
tggttccct	ccccacccc	ccacaacaca	cacacacact	ttctaagaaa	agaaaagaaa	1980
ttcttttttt	tcaatacttt	aagttctggg	atacatgtgc	agaatgtgca	ggtttgttac	2040
ataggtatac	atgtgtcatg	gtggtttgca	gcacccacca	acctatcctc	taccttaggt	2100
attctctcta	atgtatctcc	ttccctagcc	ccccaccccc	cgtgggctc	cagtgtgtga	2160
tgttccccct	catgtccatg	tgttctcatt	gttcaattcc	catttatgag	tgagaacatg	2220
cagtatttgg	ttttctgttc	ttgtgttagt	ttgtgatgg	tttctgttc	atcctgttcc	2280
ctgcaaagga	catgaactca	tcctttttta	tggctgcata	atattccatg	gtgtatatgt	2340

[0037]

gccacatttt ctttatecag tctatcgtg atgggcactg gggttgggtc caagtccttg	2400
ctattgtgaa cagtgcgtca ataaacttac atgtgcatgt gtcttttagta gaatgattta	2460
taatectttg ggtatatacc cagtaatggg attgctggtc aaatgggtatt tctggttcta	2520
gatecttgag gaatctttgt ctteccacaat ggttgaacta atttgtactc ccaccaacag	2580
tgtaaaagta ttectgtttc tctacatcct ctccagcacc tgttgtgtcc tgacatttta	2640
atgatcacta ttctcactgg cgtgagatgt tatctcattg tggttttgat ttgcatttct	2700
ctaattgacca gtaatgatga gcttttttct atatgtttgt tggctgcata aatgtcttct	2760
tttgagaagt gtctgttcat atccctcacc cattttttga agaaaacaaa ctcttaagag	2820
agcagtattc attcttttga gtgtgaggga tggagaaaga gaaagatgga gagagtatta	2880
taagcagctg tatccccctt gccatgggtga tagcagacca ttcacatggg agcttctggt	2940
ctctttgtaa taataataag agccacatta ccagtactta gagtatgcta gttattttta	3000
cacattgtat cattaaatct tcaaaacacc cctatgagtt agaaacctaa aaaaaaaaaa	3060
aaaaaaa	3067
<210> 53	
<211> 2778	
<212> DNA	
<213> 人	
<400> 53	
ctcattttga tgtctagaat caggggatcc aggatcatca ccaaggtcat ttccccaggt	60
atggaggggt ctttctgctt ctttcttgte atgcacagct gctgaggaag gggctgggag	120
taaagacagt gaaatgggga ggaggagtcc attcaaacg agaaacaaag tgtttggttt	180
ttcttaccce tgggttagaa gctaccaacc tttccaaga aagagggcct ggcccccttc	240
tgggtcttgg ctgggtgctt gctgtgcctc tctggcctcc ctctcgaagg gcaaccattc	300
ctcgggtgag tactaccggt ctgcaccgtc ttccagtggg gacagcctga gaagagagtc	360
tggggcctta ctccagtcac ttcttccact ggccaccacc tglgcaaatc atgccacacg	420
ctgcagcctc cttttcccta tctatataat aaaaatgacc ctgctctatc tcaactgggt	480
ggcaagaaca cactgttgtt gccttgcaga cagatgtgct gaggcgttag aaagtgcctt	540
ttatttgggt gggagcttgt gcataaatgc gagaggggct gcacalcga cggactagag	600
gtgactcatg gctgaaccgg aacaggacat cggggagaag ccagcagcca tgctgaactc	660
tccacagggc cctgtgaaaa gctcttcacc tctctgccc tctggatcta gtgaagccta	720
ttcatccttc agatgtcagc tcaataatac aaccttcacg gaggcctccc ttgaccctta	780
acatgctttc aaagtactgt gtatttcaca ttcatcatgc cccgacaaat gtgatttccc	840
atttattaat atctgtctct tctgtggccc tgcaaacctc aggagcacag agacatcttt	900
gggatttttg aacatgattt ccccagggtc tagcccagtg cctggtgcaa agcaggcttt	960
caacatgttc agtggatatt gtaagaaaga aagaataca caaaaggcct ggcatatgca	1020
aagcactcta aatattcact cttttccctt cctctgggt gagaaaattt ctcttataa	1080
agacaccctc ctaactgtat ctctgctaga gaactgaaga cataaagcac tctgtgccaa	1140

[0038]

```

aaatatttaa gtaaaaactt gagctaagca cagagattat aaatatttct tccccagatt 1200
acgcaccatt taaaaatact gtctcagctc cttttcatga tttgggtggg gattaaagaa 1260
aattactctt caagactgaa agtcattact gcccttttcc tgacttgccct tttcccttga 1320
gaaggggagg ataagctgca gggcaggaag tggaagtggg gcatecttgt cctttgtctg 1380
gcagacagcc aactggtcag gtactgtccc ttctcaactc ttctctgatt cccaggtgaa 1440
tataacaag aaggcacaaa tccacacttg ccaacaacgg acccaagtga taacaagaaa 1500
cccagtgaca cctgtctagg tgaagactca gccctatgtg gaccaggttg caaagccaaa 1560
ctgaccatct gctttccatt tggactttta gttcatactg tatcttctca ggacagttaa 1620
gttggaaatac aatgccactg tcctgaaaga tggtagaatt atcctatttc tggaggagtg 1680
ggggtgggtg gtaggaatct caagagcgat ttgtcctct gcacaatagc ttctttaagg 1740
acaccagggc ccccagggtc atacatttcc ctgaagcttt ccagataagc aacaagggtat 1800
gagcacctgc tatgtattgc ccaagggtga tgtgtttaaa tatccattgc atattttaaa 1860
tccttggctg gcttaaaget gcaagcttcc tgtcttcagt ggatataatg ggggcataca 1920
tcccagagct tgcccacac tccaagaaaa gaacctcag ctaatgcaaa gtgtgtatgt 1980
gcccataaaa gctccatgtc tacttaacat tcagttttta ggattattta tgctgtaata 2040
atagatatga aaatctctga caggatattt gtttcttta caaactgtat ttgaatttat 2100
gggtgattta gagcttgtgt ttaaagtcag aattcagaac cccaagaaa atgacttcat 2160
tgaaattgaa ctgaagagac aagaactgag ttaccaaaac ctactaaacg tgagttgctg 2220
tgaactgggg attaaaccag aacgagtgga gaagatcaga aagctacca acacactgct 2280
cagaaaggac aaagacattc gaagactgcg ggactttcag gaagtggaac tcattttaat 2340
gaaaaatgga agctccagat tgacagaata tgtgccatct ctgacagaaa ggccctgcta 2400
tgatagcaaa gctgcaaaaa tgacttatta aatactccca ggaatggccg cgcatgggtg 2460
ctcacccct gtaatcccag cactttggga agccaagggt ggccggtcac ctgaggtcag 2520
gagttctaga ccagctggc caacatatag tgaaaccag tctctactaa aaaaaataca 2580
aaaattagct aggtgtgggt gcgcacacct glagtagtcc cagctacatg ggaagctgag 2640
gcaggagaat cactgaacc caggaggcag aggttgctgt gagctgagat tgcgccactg 2700
cactccagcc tggcgacaga gcaagactct gtctctcaaa ataaataaat aaataaataa 2760
ataaataaat aaataatc
2778

```

<210> 54

<211> 1646

<212> DNA

<213> 人

<400> 54

```

gccccggaga ggagaggagc gggccgagga ctccagcgtg cccaggctctg gcatectgea 60
cttctgtccc tctgacacct gggaagatgg ccggcccggtg gaccttcacc cttctctgtg 120
gtttgtggc agccaccttg atccaagcca cctcagtc cactgcagtt ctcacctcg 180
gccccaaagt catcaagaa aagctgacac aggagctgaa ggaccacaac gccaccagca 240

```

[0039]

```

tccctgcagca gctgccgctg ctcagtgcc a tgcgggaaaa gccagccgga ggcatccctg 300
tgcctgggcag cctgggtgaac accgtcctga agcacatcat ctggctgaag gtcatcacag 360
ctaacatcct ccagctgcag gtgaagccct cggccaatga ccaggagctg ctagtcaaga 420
tccccctgga catggtggtt ggattcaaca cgcctctggt caagaccatc gtggagtctc 480
acatgacgac ttagggccaa gccaccatcc gcatggacac cagtgcaggt ggccccaccc 540
gcctgggtcct cagtgcctgt gccaccagcc atgggagcct gcgcacccaa ctgtgcata 600
agctctcctt cctgggtgaac gcccttagct agcaggctcat gaacctccta gtgccatccc 660
tgcccaatct agtgaaaaac cagctgtgtc ccgtgacga ggcttccttc aatggcatgt 720
atgcagacct cctgcagctg gtgaagggtg ccatttcctt cagcattgac cgtctggagt 780
ttgaccttct gtatcctgcc atcaagggtg acaccattca gctctacctg ggggccaagt 840
tgtttgactc acagggaag gtgaccaagt ggttcaataa ccttgcagct tccctgacaa 900
tgccaccctt ggacaacatc ccgttcagcc tcactgtgag tcaggacgtg gtgaaagctg 960
cagtggctgc tgtgtctctt ccagaagaat tcattgtcct gttggactct gtgttcctg 1020
agagtgccta tgggtgaag tcaagcatcg ggctgatcaa tgaaaaggct gcagataagc 1080
tgggatctac ccagatcgtg aagatcctaa ctcaggacac tcccgagttt tttatagacc 1140
aaggccatgc caagggtgcc caactgacg tcttggaagt gtttcctcc agtgaagccc 1200
tccgcccttt gttcacctg ggcatcgaag ccagctcgga agctcagttt tacaccaag 1260
gtgaccaact tatactcaac ttgaataaca tcagctctga tcggatccag ctgatgaact 1320
ctgggattgg ctggttccaa cctgatgttc tgaaaaacat cactactgag atcatccact 1380
ccatcctgct gccgaaccag aatggcaaat taagatctgg ggteccagtg tcatttgtga 1440
agcccttggg attcagagca gctgagtcct cactgaccaa ggatgccctt gtgcttactc 1500
cagcctcctt gtggaacccc agctctcctg tctcccagtg aagacttgga tggcagccat 1560
cagggaaggc tgggtcccag ctgggagtat ggggtgtgagc tctatagacc atccctctct 1620
gcaatcaata aacacttgcc tgtgat 1646

```

<210> 55

<211> 1049

<212> DNA

<213> 人

<400> 55

```

ggagtggggg agagagagga gaccaggaca gctgctgaga cctctaagaa gtccagatac 60
taagagcaaa gatgtttcaa actgggggccc tcattgtctt ctacgggctg ttagcccaga 120
ccatggccca gtttgagggc ctgccctgct ccctggacca gacctgccc ttgaatgtga 180
atccagccct gcccttgagt ccacaggtc ttgcaggaag cttgacaaat gccctcagca 240
atggcctgct gcttgggggc ctgttgggca ttctggaaaa ccttcgctc ctggacatcc 300
tgaagcctgg aggagtact tctgtgggcc tccttggggg actgcttgga aaagtgcgt 360
cagtatttcc tggcctgaac aacatcattg acataaaggt cactgacccc cagctgctgg 420
aacctggcct tglgcagagc cctgalggcc accgtctcta tglcaccalc cctctcgga 480

```

[0040]

taaagctcca agtgaatacg cccctggctg gtgcaagtct gttgaggctg gctgtgaage 540
 tggacatcac tgcagaaatc ttagctgtga gagataagca ggagaggatc cacctggctc 600
 ttggtgactg caccatttc cctggaagcc tgcaaatctc tctgttgat ggacttgcc 660
 cctccccat tcaaggtctt ctggacagcc tcacagggat cttgaataaa gtcttgctg 720
 agttggttca gggcaacgtg tgcctctgg tcaatgaggt tctcagaggc ttggacatca 780
 ccttggtgca tgacattgtt aacatgtga tccacggact acagtttctc atcaaggtct 840
 aagccttcca ggaaggggct ggcctctgct gactgcttc ccagtgtca cagatggctg 900
 gcccatgtgc tggaaatga cacagtgcc tctctccga ggaacctgcc cctctcctt 960
 tcccaccagg cgtgtgtaac atccatgtg cctcacctaa taaaatggct cttctctgc 1020
 aaaaaaaaa aaaaaaaaa aaaaaaaaa 1049

<210> 56

<211> 4815

<212> DNA

<213> 人

<400> 56

gagcagagcc ctttcacaca cctcaggaac accttcggc tgcctgctc ccagacacac 60
 ctgcagcct gccagccgg cttgtctac cactgcttg taaatgccc agatatgagc 120
 cagccaggc ccgctactg gtagacaga gcgcatact ccttacctt ctcgacgat 180
 gagtttgaga agaaggaccg gacataccca gtgggagaga aacttcgaa tgccttcaga 240
 tgttctcag ccaagatcaa agctgtggtg ttgggctgc tgcctgtgt ctcctggctc 300
 cccaagtaca agattaaaga ctacatcatt cctgacctgc tgggtggact cagcggggga 360
 tccatccagg tcccacaagg catggcattt gctctgttg ccaaccttc tgcagtcaat 420
 ggctctact cctcttctt cccctctctg acctactct tctgggggg tgttcaccag 480
 atggtgccag gtaccttgc cgttatcagc atcctggttg gtaacatct tctgcagctg 540
 gccccagagt cgaaattcca ggtcttcaac aatgccacca atgagagct tgtggacaca 600
 gcagccatgg aggtgagag gctgcacgtg tcagctacgc tagctgctt caccgccatc 660
 atccagatgg gtctgggctt catgcagttt gctttgttg ccatctacct ctcgagctc 720
 ttcacccggg gcttcatgac ggccgccggc ctgcagatcc tgatttcggt gctcaagtac 780
 atcttcggac tgaccatccc ctctacaca ggcccagggt ccatgtctt taccttcatt 840
 gacatttga aaaacctccc ccacaccaac atgcctcgc tcatcttcgc tctcatcagc 900
 ggtgccttc tgggtctggt gaaggagctc aatgctgct acatgcacaa gattcgcttc 960
 cccatcceta cagagatgat tgtggtggtg gtggcaacag ctatctccg ggctgtgaag 1020
 atgccccaaa agtatcacat gcagatcgtg ggagaaatcc aacgggggt cccaccccg 1080
 gtgtgcctg tggctcaca gtggaaggac atgataggca cagccttct cctagccatc 1140
 gtgagctacg tcatcaacct ggctatgggc cggacctgg ccaacaagca cgctacgac 1200
 gtggattcga accaggagat gatcgtctc ggctgcagca acttcttgg ctcctcttt 1260
 aaaattcatg tcatttgctg tgcgtttct gtcactctg ctgtggatgg agctggagga 1320

[0041]

aaateccagg	tggccagcct	gtgtgtgtct	ctgggtgtga	tgatcaccat	gctggtcctg	1380
gggatctatc	tgtatcctct	ccetaagtct	gtgctaggag	ccctgacgc	tgtaaatctc	1440
aagaactccc	tcaagcaact	caccgacccc	tactacctgt	ggaggaagag	caagctggac	1500
tgttgcatct	gggtagttag	cttctctctc	tccttcttcc	tcagcctgcc	ctatgggtgtg	1560
gcagtgggtg	tcgccttctc	cgtcctggtc	gtggtcttcc	agactcagtt	tcgaaatggc	1620
taigcactgg	cccagggtcat	ggacactgac	attatgtiga	atcccaagac	ctataatagg	1680
gcccaggata	tccaggggat	taaaatcctc	acgtactgct	cccctctcta	ctttgccaac	1740
tcagagatct	tcaggcaaaa	ggtcatcgcc	aagacaggca	tggaccccca	gaaagtatta	1800
ctagccaagc	aaaaatacct	caagaagcag	gagaagcgga	gaatgaggcc	cacacaacag	1860
aggaggctct	tattcatgaa	aaccaagact	gtctccctgc	aggagctgca	gcaggacttt	1920
gagaatgcgc	ccccaccca	ccccacaac	aaccagaccc	cggtacagg	caccagcgtg	1980
tcctatatca	ccttcagccc	tgacagctcc	tcacctgccc	agagtgcgc	accagcctcc	2040
gctgaggccc	ccggcgagcc	cagtgcactg	ctggccagcg	tcccaccctt	cgtcaccctc	2100
cacaccctca	tcctggacat	gagtggagtc	agcttcgttg	acttgatggg	catcaaggcc	2160
ctggccaagc	tgagctccac	ctatgggaag	atcgccgtga	aggtcttctt	ggtgaacatc	2220
catgcccagg	tgtacaatga	cattagccat	ggaggcgtct	ttgaggatgg	gagtctagaa	2280
tgcaagcagc	tctttcccag	catacatgac	gcagtcctct	ttgcccagge	aaatgctaga	2340
gaactgaccc	caggacacaa	cttccaaggg	gctccagggg	atgctgagct	ctctttgtac	2400
gactcagagg	aggacattcg	cagctactgg	gacttagagc	aggagatgtt	cgggagcatg	2460
tttcacgcag	agaccctgac	cgcctgtga	gggtcagccc	agtcctcatg	ctgcctacag	2520
agtgcctggc	acttgggact	tccataaagg	atgagcctgg	ggtcacaggg	ggtgtcgggc	2580
ggaggaaggt	gcacccccca	gagcttgggt	tcctctctcc	tctccccctc	tctctcctct	2640
tccttccctc	cccgcatctc	cagagagagc	ctctcagcag	caggggggtg	ctacccttac	2700
gggagtgaga	gtctgggtgag	cccactcttc	accgctcagg	ccctggccgc	aatggacaag	2760
cctcctgctc	actccacccc	accacatctt	gcccgttctt	tggcagctga	aggacacctt	2820
gacttccagc	ttttacgagt	gagccaaaaa	cagaaggaca	agtacaactg	tgctggcctg	2880
ctgtacaagc	ttcaaaaagt	gtccagagc	ccgcacggct	cgggtgtcaga	tgggtgcagg	2940
ctgtcacgga	catagggata	aacttggtta	ggactctggc	ttgccttccc	cagctgcctc	3000
aactctgtct	ctggcagctc	tgcacccagg	gaccatgtgc	tctccacacc	caggagtcta	3060
ggccttggta	actatgcgcc	ccccctccat	cacccccaa	gctgcccaca	ccaccactgc	3120
tgtcagcaag	cacatcagac	tctagcctgg	acagtggcca	ggaccgtcga	gaccaccaga	3180
gctacctccc	cggggacagc	ccactaaggt	tctgcctcag	cctcctgaaa	catcactgcc	3240
ctcagaggct	gctcccttcc	cctggaggct	ggctagaaac	cccaaagagg	gggatgggta	3300
gctggcagaa	tcactctggca	tcctagtaat	agataccagt	tattctgcac	aaaacttttg	3360
ggaattctct	tttgacacca	gagactcaga	ggggaagagg	gtgctagtag	caacacaggg	3420
aaaaeggatg	ggacctgggc	ccagacagtc	ccccctgacc	ccaggggcca	tcagggaagt	3480
gcctcccttt	ggtaaatctg	ccttaccctt	ctttacctgg	caaagagcca	atcatgttaa	3540
ctcttccctt	tcagcctgtg	gcccagagac	acaatggggg	ccttctgtag	gcaaagggtg	3600

[0042]

```

aagtectcca gggatceget acatecccta actgeatgca gatgtggaaa ggggctgate 3660
cagattgggt ctctctgcac aggaagactc tttaacaccc ttaggacctc aggccatctt 3720
ctcctatgaa gatgaaaata ggggttaagt ttccatatg tacaaggagg tattgagagg 3780
aaccctactg ttgacttgaa aataaatagg ttccatgtgt aagtgttttg taaaatttca 3840
gtggaaatgc acagaaaatc ttctggcctc tcactactgc ttttctcaag ctctctcagc 3900
ttaacaaccc ctctccctaac aggttgggct ggcccagcct aggaaaacat ccccatctct 3960
aacttcagcc agacctgcgt tgtgtgtctg tgtgttgagt gagctggtea gctaacaagt 4020
cttcttagag ttaaaggagg ggggtgctgc caagagccaa cacattcttg gcccaggagg 4080
attgcttttc tgtgaattca ttatgccatc tggttgccaa tggaactcaa aacttggaag 4140
ggaaggaca atgttatctg ggattcaccg tgcccagcac cgaagtgcc aaattccagg 4200
aggacaagag ccttagccaa tgacaactca ctctcccta ctccacctcc ttccaagtec 4260
agctcaggcc caggaggtgg gagaaggtea cagagcctca ggaatttcca agtcagagtc 4320
ccctttgaac caagtatcta gatccctga ggacttgatg aagtgatcct taaccccaa 4380
gtaatcatta acccccagac cagcctcaga actgaaggag attgttgacc cagtgcctg 4440
gagttgaggc tcaggagag atctgccaca tgtctgaggg ttgcagagcc cgtgtggag 4500
gtaagattgg aaacacatga ggcagaggga agacattgaa gaaaacatct ctgctggaat 4560
atttgaaaaa gaacactctt ctggacctgg ttgaagcagg aaagatggag gcaaagtagt 4620
gaaataatcc agaatttcaa tgettttgaa tgttcttagt gatactgacc tgtgataata 4680
taattcccag ggaggactgg gaaccttacc tcttgagata ttgcataat ttatttaatt 4740
taagcctcat tctcttttg ttcatcttgg taataaactg gatttgaatt gtgaacaaaa 4800
aaaaaaaaaaaa 4815

```

<210> 57

<211> 2572

<212> DNA

<213> 人

<400> 57

```

aatgctctaa gacctctcag caggggcgga agaaactccc ggagagctca cccaaaaaac 60
aaggagatcc catctagatt tcttcttgc tttgactcac agctggaagt tagaaaagcc 120
tcgatttcat ctttgagag gccaaatggt cttagcctca gtctctgtct ctaaatattc 180
caccataaaa cagctgagtt atttatgaat tagaggctat agctcacatt ttcaatctc 240
tatttctttt tttaaataa actttctact ctgatgagag aatgtggttt taatctctct 300
ctcacatttt gatgatttag acagactccc cctcttctc ctagtcaata aaccattga 360
tgatctatctt ccagcttat cccaagaaa acttttgaaa ggaagagta gacccaaaga 420
tgttattttc tgcgttttga attttgtct cccaccccca acttggetag taataaacac 480
ttactgaaga agaagcaata agagaaagat atttgtaac tctccagccc atgatctcg 540
ttttcttaca ctgtgatctt aaaagttacc aaaccaaagt cattttcagt ttgaggcaac 600
caaaccttct tactgtgttt gacatcttct tattacagca acaccattct aggagtctc 660

```

[0043]

```

tgagctctcc actggagtcc tctttctgtc gcgggtcaga aattgtccct agatgaatga 720
gaaaattatt ttttttaatt taagtcctaa atatagttaa aataaataat gttttagtaa 780
aatgatacac tatctctgtg aaatagcctc acccctacat gtggatagaa ggaaatgaaa 840
aaataattgc tttgacattg tctatatggg actttgtaaa gtcattgctta agtacaatt 900
ccatgaaaag ctcactgac ctaattcttt ccttttgagg tctctatggc tctgattgta 960
catgatagta agtgaagcc atgtaaaaag taaataatgt ctgggcacag tggtcacgc 1020
ctgtaatcct agcacttttg gagctgagg aggaaggatc acttgagccc agaagtcca 1080
gactagcctg ggcaacatgg agaagccctg tctctacaaa atacagagag aaaaaatcag 1140
ccagtcattg tggcatacac ctgtagtccc agcattccgg gagctgagg tgggaggatc 1200
acttgagccc agggaggttg gggctgcagt gagecatgat cacaccactg cactccagcc 1260
aggtgacata gcgagatcct gtctaaaaaa ataaaaata aataatggaa cacagcaagt 1320
cctaggaagt aggttaaaac taattcttta aaaaaaaaaa aaagttgagc ctgaattaaa 1380
tgtaatgttt ccaagtgaca ggtatccaca ttgcatggt tacaagccac tgccagttgg 1440
cagtagcact ttctggcac tgtggtcggg tttgtttgt tttgcttgt ttagagacgg 1500
ggtctcactt tccaggtgg cctcaaaact ctgcaactca gcaattcttc taccctggcc 1560
tcccaagtag ctggaattac aggtgtgcgc catcacaact agctggtggt cagttttgtt 1620
actctgagag ctgttcaact ctctgaatc accctagagt gttggaccat cagatgtttg 1680
ggcaaaactg aaagctcttt gcaaccacac accttccctg agcttacatc actgcccttt 1740
tgagcagaaa gtctaaatc cttccaagac agtagaattc catcccagta ccaaagccag 1800
atagggcccc taggaaactg aggtgaagac agtctctaaa aactaccac agcagcattg 1860
gtgcagggga acttgcccat taggttatta tttagagga aagtcctcac atcaatagta 1920
catatgaaa tgacctcaa ggggatttgt gaatactcat aaggatcttc aggtgaaca 1980
gactatgtct ggggaaagaa cggattatgc ccattaaat aacaagttgt gttcaagagt 2040
cagagcagtg agctcagagg ccttctcac tgagacagca acatttaaac caaaccagag 2100
gaagtatttg tggaactcac tgcctcagtt tgggtaaagg atgagcagac aagtcaacta 2160
aagaaaaaag aaaagcaagg aggagggttg agcaatctag agcatggagt ttgttaagtg 2220
ctctctggat ttgagttgaa gagcatccat ttgagttgaa ggccacaggg cacaatgagc 2280
tctccctct accaccagaa agtccctggt caggctcag gtagtgcggg tgggtcagc 2340
tgggttttta attagcgc atctctatcca acatttaatt gtttgaaagc ctccatag 2400
ttagattgtg ctttgaatt ttgttggtgt tgcctctatc tattgtatat gcattgagta 2460
ttaacctgaa tgttttgtta cttaaatatt aaaacactg ttatctaca aaaaaacct 2520
caaagctga aaataagaa ggaagatgga gacacctct ggggtcctc tc 2572

```

<210> 58

<211> 1324

<212> DNA

<213> 人

[0044]

<400> 58
 ctttgcagtg gatgcccttg gcagggtgag cccacaagga gcaatggagc agggcagcgg 60
 ccgcttggag gacttccctg tcaatgtgtt ctccgtcact ccttacacac ccagcaccgc 120
 tgacatccag gtgtccgatg atgacaaggc gggggccacc ttgtcttct caggeatctt 180
 tetgggactg gtggggatca cattcactgt catgggctgg atcaaatacc aaggtgtctc 240
 ccactttgaa tggacceagc tccittggcc cgtcctgctg tcagttgggg tgacattcat 300
 cctgattgct gtgtgcaagt tcaaaatgct ctctgccag ttgtgcaaag aaagtgagga 360
 aagggtcccg gactcggaac agacaccagg aggaccatca ttgtttttca ctggcatcaa 420
 ccaaceccatc acctccatg gggccactgt ggtgcagtac atccctcctc cttatggttc 480
 tccagagcct atggggataa ataccagcta cctgcagtct gtggtgagcc cctgeggcct 540
 cataacctct ggaggggcag cagccgccat gtcaagtctt cctcaatact acaccatcta 600
 cccctaagat aactctgcat ttgtggttga tgagggtgc ctttctttca cggacgggtg 660
 aaatcacagg cccaatcctg atgttgacca gctagaagag acacagctgg aagaggaggc 720
 ctgtgectgc ttctctcctc ccccttatga agaaatatac tctctccctc gctagagget 780
 attctgatat aataacacaa tgtcagctc agggagcaag tgtttccgtc attgttacct 840
 gacaaccgtg gtgttctatg ttgtaacctt cagaagttac agcagcgccc aggcagcctg 900
 acagagatca ttcaaggggg gaaaggggaa gtgggaggtg caatttctca gatttgtaaa 960
 aattagctg ggctggggaa attctctcc ggaacagttt caaatccct cgggtaagaa 1020
 atctctgta taaggttcag gacaggaat ttcactttt catccaccac cctccccctt 1080
 ctctgtagga aggcatttgg ggtcaattt taacccagc agccaatgga aaaatcacga 1140
 cttctgagac ttgggagtt tccacagagg tgagagtcgg gtgggaagga agcagggaag 1200
 agaaagcagg ccagctgga gatttccctg tggtctctct tggcccaaaa gcagactcac 1260
 taatcccaaa caactcagct gccatctggc ctctctgagg actctgggta ccttaaagac 1320
 tata 1324

<210> 59

<211> 683

<212> DNA

<213> 人

<400> 59
 caggaaagtt cgtctgcta ggcagaggaa ctgcagcttg ttggcaggtg aaggagacct 60
 gtlttagctg gtccagcaac aacllacgtg gtccgtcttg tglccaggt gaagcgtctg 120
 gccgcagc agaggaatca agacctgctc attcttctct cgggggatac atccagcaat 180
 gacatcatct catgtctcca caaggacccc aagtctgggc tgctggggac cagccacgct 240
 ccccactgct cattcttca tccagagac attctgactc tctccgact gcgtgtgca 300
 caggcgtgac aagctctttt acatctcagt ctgcacaact tcaggcactt agcagattga 360
 tatgcataca acaaatattg attgaatata tgctaaatac ccagtaatgt ttcattgagt 420
 attgggtgaa taaaggaatg ctggttccct ctggccatat taactcctgc acaatactaa 480
 gaaaaataaa ttgcactagc tgtggaataa tgtgaatccc aatgtcatct attgaaatat 540

[0045]

tacctgacta ttaagaggta tttatitttg tatcttttct agcaaagtaa ataaaattct 600
taatacagca tateccctta ttcacggggg gtatgttcca agacccccgg tggatgcctg 660
aaactatgga taataccaga tcc 683

<210> 60

<211> 914

<212> PRT

<213> 人

<400> 60

Met Gly Pro Phe Lys Ser Ser Val Phe Ile Leu Ile Leu His Leu Leu
1 5 10 15

Glu Gly Ala Leu Ser Asn Ser Leu Ile Gln Leu Asn Asn Asn Gly Tyr
20 25 30

Glu Gly Ile Val Val Ala Ile Asp Pro Asn Val Pro Glu Asp Glu Thr
35 40 45

Leu Ile Gln Gln Ile Lys Asp Met Val Thr Gln Ala Ser Leu Tyr Leu
50 55 60

Phe Glu Ala Thr Gly Lys Arg Phe Tyr Phe Lys Asn Val Ala Ile Leu
65 70 75 80

Ile Pro Glu Thr Trp Lys Thr Lys Ala Asp Tyr Val Arg Pro Lys Leu
85 90 95

Glu Thr Tyr Lys Asn Ala Asp Val Leu Val Ala Glu Ser Thr Pro Pro
100 105 110

Gly Asn Asp Glu Pro Tyr Thr Glu Gln Met Gly Asn Cys Gly Glu Lys
115 120 125

Gly Glu Arg Ile His Leu Thr Pro Asp Phe Ile Ala Gly Lys Lys Leu
130 135 140

Ala Glu Tyr Gly Pro Gln Gly Lys Ala Phe Val His Glu Trp Ala His
145 150 155 160

Leu Arg Trp Gly Val Phe Asp Glu Tyr Asn Asn Asp Glu Lys Phe Tyr
165 170 175

Leu Ser Asn Gly Arg Ile Gln Ala Val Arg Cys Ser Ala Gly Ile Thr
180 185 190

Gly Thr Asn Val Val Lys Lys Cys Gln Gly Gly Ser Cys Tyr Thr Lys
195 200 205

Arg Cys Thr Phe Asn Lys Val Thr Gly Leu Tyr Glu Lys Gly Cys Glu
210 215 220

[0046]

Phe Val Leu Gln Ser Arg Gln Thr Glu Lys Ala Ser Ile Met Phe Ala
 225 230 235 240
 Gln His Val Asp Ser Ile Val Glu Phe Cys Thr Glu Gln Asn His Asn
 245 250 255
 Lys Glu Ala Pro Asn Lys Gln Asn Gln Lys Cys Asn Leu Arg Ser Thr
 260 265 270
 Trp Glu Val Ile Arg Asp Ser Glu Asp Phe Lys Lys Thr Thr Pro Met
 275 280 285
 Thr Thr Gln Pro Pro Asn Pro Thr Phe Ser Leu Leu Gln Ile Gly Gln
 290 295 300
 Arg Ile Val Cys Leu Val Leu Asp Lys Ser Gly Ser Met Ala Thr Gly
 305 310 315 320
 Asn Arg Leu Asn Arg Leu Asn Gln Ala Gly Gln Leu Phe Leu Leu Gln
 325 330 335
 Thr Val Glu Leu Gly Ser Trp Val Gly Met Val Thr Phe Asp Ser Ala
 340 345 350
 Ala His Val Gln Ser Glu Leu Ile Gln Ile Asn Ser Gly Ser Asp Arg
 355 360 365
 Asp Thr Leu Ala Lys Arg Leu Pro Ala Ala Ala Ser Gly Gly Thr Ser
 370 375 380
 Ile Cys Ser Gly Leu Arg Ser Ala Phe Thr Val Ile Arg Lys Lys Tyr
 385 390 395 400
 Pro Thr Asp Gly Ser Glu Ile Val Leu Leu Thr Asp Gly Glu Asp Asn
 405 410 415
 Thr Ile Ser Gly Cys Phe Asn Glu Val Lys Gln Ser Gly Ala Ile Ile
 420 425 430
 His Thr Val Ala Leu Gly Pro Ser Ala Ala Gln Glu Leu Glu Glu Leu
 435 440 445
 Ser Lys Met Thr Gly Gly Leu Gln Thr Tyr Ala Ser Asp Gln Val Gln
 450 455 460
 Asn Asn Gly Leu Ile Asp Ala Phe Gly Ala Leu Ser Ser Gly Asn Gly
 465 470 475 480
 Ala Val Ser Gln Arg Ser Ile Gln Leu Glu Ser Lys Gly Leu Thr Leu
 485 490 495
 Gln Asn Ser Gln Trp Met Asn Gly Thr Val Ile Val Asp Ser Thr Val
 500 505 510
 Gly Lys Asp Thr Leu Phe Leu Ile Thr Trp Thr Thr Gln Pro Pro Gln
 515 520 525

[0047]

Ile Leu Leu Trp Asp Pro Ser Gly Gln Lys Gln Gly Gly Phe Val Val
 530 535 540

Asp Lys Asn Thr Lys Met Ala Tyr Leu Gln Ile Pro Gly Ile Ala Lys
 545 550 555 560

Val Gly Thr Trp Lys Tyr Ser Leu Gln Ala Ser Ser Gln Thr Leu Thr
 565 570 575

Leu Thr Val Thr Ser Arg Ala Ser Asn Ala Thr Leu Pro Pro Ile Thr
 580 585 590

Val Thr Ser Lys Thr Asn Lys Asp Thr Ser Lys Phe Pro Ser Pro Leu
 595 600 605

Val Val Tyr Ala Asn Ile Arg Gln Gly Ala Ser Pro Ile Leu Arg Ala
 610 615 620

Ser Val Thr Ala Leu Ile Glu Ser Val Asn Gly Lys Thr Val Thr Leu
 625 630 635 640

Glu Leu Leu Asp Asn Gly Ala Gly Ala Asp Ala Thr Lys Asp Asp Gly
 645 650 655

Val Tyr Ser Arg Tyr Phe Thr Thr Tyr Asp Thr Asn Gly Arg Tyr Ser
 660 665 670

Val Lys Val Arg Ala Leu Gly Gly Val Asn Ala Ala Arg Arg Arg Val
 675 680 685

Ile Pro Gln Gln Ser Gly Ala Leu Tyr Ile Pro Gly Trp Ile Glu Asn
 690 695 700

Asp Glu Ile Gln Trp Asn Pro Pro Arg Pro Glu Ile Asn Lys Asp Asp
 705 710 715 720

Val Gln His Lys Gln Val Cys Phe Ser Arg Thr Ser Ser Gly Gly Ser
 725 730 735

Phe Val Ala Ser Asp Val Pro Asn Ala Pro Ile Pro Asp Leu Phe Pro
 740 745 750

Pro Gly Gln Ile Thr Asp Leu Lys Ala Glu Ile His Gly Gly Ser Leu
 755 760 765

Ile Asn Leu Thr Trp Thr Ala Pro Gly Asp Asp Tyr Asp His Gly Thr
 770 775 780

Ala His Lys Tyr Ile Ile Arg Ile Ser Thr Ser Ile Leu Asp Leu Arg
 785 790 795 800

Asp Lys Phe Asn Glu Ser Leu Gln Val Asn Thr Thr Ala Leu Ile Pro
 805 810 815

Lys Glu Ala Asn Ser Glu Glu Val Phe Leu Phe Lys Pro Glu Asn Ile
 820 825 830

[0048]

Thr Phe Glu Asn Gly Thr Asp Leu Phe Ile Ala Ile Gln Ala Val Asp
835 840 845

Lys Val Asp Leu Lys Ser Glu Ile Ser Asn Ile Ala Arg Val Ser Leu
850 855 860

Phe Ile Pro Pro Gln Thr Pro Pro Glu Thr Pro Ser Pro Asp Glu Thr
865 870 875 880

Ser Ala Pro Cys Pro Asn Ile His Ile Asn Ser Thr Ile Pro Gly Ile
885 890 895

His Ile Leu Lys Ile Met Trp Lys Trp Ile Gly Glu Leu Gln Leu Ser
900 905 910

Ile Ala

<210> 61

<211> 501

<212> PRT

<213> 人

<400> 61

Met Lys Lys Glu Gly Arg Lys Arg Trp Lys Arg Lys Glu Asp Lys Lys
1 5 10 15

Arg Val Val Val Ser Asn Leu Leu Phe Glu Gly Trp Ser His Lys Glu
20 25 30

Asn Pro Asn Arg His His Arg Gly Asn Gln Ile Lys Thr Ser Lys Tyr
35 40 45

Thr Val Leu Ser Phe Val Pro Lys Asn Ile Phe Glu Gln Leu His Arg
50 55 60

Phe Ala Asn Leu Tyr Phe Val Gly Ile Ala Val Leu Asn Phe Ile Pro
65 70 75 80

Val Val Asn Ala Phe Gln Pro Glu Val Ser Met Ile Pro Ile Cys Val
85 90 95

Ile Leu Ala Val Thr Ala Ile Lys Asp Ala Trp Glu Asp Leu Arg Arg
100 105 110

Tyr Lys Ser Asp Lys Val Ile Asn Asn Arg Glu Cys Leu Ile Tyr Ser
115 120 125

Arg Lys Glu Gln Thr Tyr Val Gln Lys Cys Trp Lys Asp Val Arg Val
130 135 140

Gly Asp Phe Ile Gln Met Lys Cys Asn Glu Ile Val Pro Ala Asp Ile
145 150 155 160

[0049]

Leu Leu Leu Phe Ser Ser Asp Pro Asn Gly Ile Cys His Leu Glu Thr
 165 170 175
 Ala Ser Leu Asp Gly Glu Thr Asn Leu Lys Gln Arg Arg Val Val Lys
 180 185 190
 Gly Phe Ser Gln Gln Glu Val Gln Phe Glu Pro Glu Leu Phe His Asn
 195 200 205
 Thr Ile Val Cys Glu Lys Pro Asn Asn His Leu Asn Lys Phe Lys Gly
 210 215 220
 Tyr Met Glu His Pro Asp Gln Thr Arg Thr Gly Phe Gly Cys Glu Ser
 225 230 235 240
 Leu Leu Leu Arg Gly Cys Thr Ile Arg Asn Thr Glu Met Ala Val Gly
 245 250 255
 Ile Val Ile Tyr Ala Gly His Glu Thr Lys Ala Met Leu Asn Asn Ser
 260 265 270
 Gly Pro Arg Tyr Lys Arg Ser Lys Ile Glu Arg Arg Met Asn Ile Asp
 275 280 285
 Ile Phe Phe Cys Ile Gly Ile Leu Ile Leu Met Cys Leu Ile Gly Ala
 290 295 300
 Val Gly His Ser Ile Trp Asn Gly Thr Phe Glu Glu His Pro Pro Phe
 305 310 315 320
 Asp Val Pro Asp Ala Asn Gly Ser Phe Leu Pro Ser Ala Leu Gly Gly
 325 330 335
 Phe Tyr Met Phe Leu Thr Met Ile Ile Leu Leu Gln Val Leu Ile Pro
 340 345 350
 Ile Ser Leu Tyr Val Ser Ile Glu Leu Val Lys Leu Gly Gln Val Phe
 355 360 365
 Phe Leu Ser Asn Asp Leu Asp Leu Tyr Asp Glu Glu Thr Asp Leu Ser
 370 375 380
 Ile Gln Cys Arg Ala Leu Asn Ile Ala Glu Asp Leu Gly Gln Ile Gln
 385 390 395 400
 Tyr Ile Phe Ser Asp Lys Thr Gly Thr Leu Thr Glu Asn Lys Met Val
 405 410 415
 Phe Arg Arg Cys Thr Ile Met Gly Ser Glu Tyr Ser His Gln Glu Asn
 420 425 430
 Gly Ile Glu Ala Pro Lys Gly Ser Ile Pro Leu Ser Lys Arg Lys Tyr
 435 440 445
 Pro Ala Leu Leu Arg Asn Glu Glu Ile Lys Asp Ile Leu Leu Ala Leu
 450 455 460

[0050]

Leu Glu Ala Val Trp His Phe His Lys Leu Leu Pro Val Ser Leu Trp
465 470 475 480

Ser Ser Leu Ser Gln Ile Arg Ala Val Pro Ile Thr Cys Lys Leu Ser
485 490 495

Phe Val Tyr Lys Gly
500

<210> 62

<211> 154

<212> PRT

<213> 人

<400> 62

Met Gly Arg Arg Ser Pro Phe Lys Pro Arg Asn Lys Val Phe Gly Phe
1 5 10 15

Ser Tyr Pro Trp Cys Arg Ser Tyr Gln Pro Phe Pro Arg Lys Arg Ala
20 25 30

Trp Pro Pro Ser Arg Val Trp Leu Gly Ala Cys Cys Ala Ser Leu Ala
35 40 45

Ser Pro Pro Lys Gly Thr Ile Pro Ser Gly Glu Tyr Tyr Arg Pro Ala
50 55 60

Pro Ser Ser Ser Gly Asp Ser Leu Arg Arg Glu Ser Gly Ala Leu Leu
65 70 75 80

Gln Tyr Leu Pro Ser Leu Ala Ser Pro Cys Ala Asn His Ala Thr Arg
85 90 95

Cys Ser Leu Leu Phe Pro Ile Tyr Lys Ile Lys Met Thr Leu Leu Tyr
100 105 110

Leu Thr Gly Leu Ala Arg Thr His Cys Cys Cys Leu Ala Asp Arg Cys
115 120 125

Ala Glu Ala Val Glu Ser Ala Phe Tyr Leu Val Gly Ser Leu Cys Ile
130 135 140

Asn Ala Arg Gly Ala Ala His Leu Thr Asp
145 150

<210> 63

<211> 484

<212> PRT

<213> 人

<400> 63

[0051]

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

[0052]

Leu Leu Asp Ser Val Leu Pro Glu Ser Ala His Arg Leu Lys Ser Ser
305 310 315 320

Ile Gly Leu Ile Asn Glu Lys Ala Ala Asp Lys Leu Gly Ser Thr Gln
325 330 335

Ile Val Lys Ile Leu Thr Gln Asp Thr Pro Glu Phe Phe Ile Asp Gln
340 345 350

Gly His Ala Lys Val Ala Gln Leu Ile Val Leu Glu Val Phe Pro Ser
355 360 365

Ser Glu Ala Leu Arg Pro Leu Phe Thr Leu Gly Ile Glu Ala Ser Ser
370 375 380

Glu Ala Gln Phe Tyr Thr Lys Gly Asp Gln Leu Ile Leu Asn Leu Asn
385 390 395 400

Asn Ile Ser Ser Asp Arg Ile Gln Leu Met Asn Ser Gly Ile Gly Trp
405 410 415

Phe Gln Pro Asp Val Leu Lys Asn Ile Ile Thr Glu Ile Ile His Ser
420 425 430

Ile Leu Leu Pro Asn Gln Asn Gly Lys Leu Arg Ser Gly Val Pro Val
435 440 445

Ser Leu Val Lys Ala Leu Gly Phe Glu Ala Ala Glu Ser Ser Leu Thr
450 455 460

Lys Asp Ala Leu Val Leu Thr Pro Ala Ser Leu Trp Lys Pro Ser Ser
465 470 475 480

Pro Val Ser Gln

<210> 64

<211> 256

<212> PRT

<213> 人

<400> 64

Met Phe Gln Thr Gly Gly Leu Ile Val Phe Tyr Gly Leu Leu Ala Gln
1 5 10 15

Thr Met Ala Gln Phe Gly Gly Leu Pro Val Pro Leu Asp Gln Thr Leu
20 25 30

Pro Leu Asn Val Asn Pro Ala Leu Pro Leu Ser Pro Thr Gly Leu Ala
35 40 45

Gly Ser Leu Thr Asn Ala Leu Ser Asn Gly Leu Leu Ser Gly Gly Leu
50 55 60

[0053]

Leu Gly Ile Leu Glu Asn Leu Pro Leu Leu Asp Ile Leu Lys Pro Gly
65 70 75 80

Gly Gly Thr Ser Gly Gly Leu Leu Gly Gly Leu Leu Gly Lys Val Thr
85 90 95

Ser Val Ile Pro Gly Leu Asn Asn Ile Ile Asp Ile Lys Val Thr Asp
100 105 110

Pro Gln Leu Leu Glu Leu Gly Leu Val Gln Ser Pro Asp Gly His Arg
115 120 125

Leu Tyr Val Thr Ile Pro Leu Gly Ile Lys Leu Gln Val Asn Thr Pro
130 135 140

Leu Val Gly Ala Ser Leu Leu Arg Leu Ala Val Lys Leu Asp Ile Thr
145 150 155 160

Ala Glu Ile Leu Ala Val Arg Asp Lys Gln Glu Arg Ile His Leu Val
165 170 175

Leu Gly Asp Cys Thr His Ser Pro Gly Ser Leu Gln Ile Ser Leu Leu
180 185 190

Asp Gly Leu Gly Pro Leu Pro Ile Gln Gly Leu Leu Asp Ser Leu Thr
195 200 205

Gly Ile Leu Asn Lys Val Leu Pro Glu Leu Val Gln Gly Asn Val Cys
210 215 220

Pro Leu Val Asn Glu Val Leu Arg Gly Leu Asp Ile Thr Leu Val His
225 230 235 240

Asp Ile Val Asn Met Leu Ile His Gly Leu Gln Phe Val Ile Lys Val
245 250 255

<210> 65

<211> 791

<212> PRT

<213> 人

<400> 65

Met Ser Gln Pro Arg Pro Arg Tyr Val Val Asp Arg Ala Ala Tyr Ser
1 5 10 15

Leu Thr Leu Phe Asp Asp Glu Phe Glu Lys Lys Asp Arg Thr Tyr Pro
20 25 30

Val Gly Glu Lys Leu Arg Asn Ala Phe Arg Cys Ser Ser Ala Lys Ile
35 40 45

Lys Ala Val Val Phe Gly Leu Leu Pro Val Leu Ser Trp Leu Pro Lys
50 55 60

[0054]

Tyr Lys Ile Lys Asp Tyr Ile Ile Pro Asp Leu Leu Gly Gly Leu Ser
 65 70 75 80
 Gly Gly Ser Ile Gln Val Pro Gln Gly Met Ala Phe Ala Leu Leu Ala
 85 90 95
 Asn Leu Pro Ala Val Asn Gly Leu Tyr Ser Ser Phe Phe Pro Leu Leu
 100 105 110
 Thr Tyr Phe Phe Leu Gly Gly Val His Gln Met Val Pro Gly Thr Phe
 115 120 125
 Ala Val Ile Ser Ile Leu Val Gly Asn Ile Cys Leu Gln Leu Ala Pro
 130 135 140
 Glu Ser Lys Phe Gln Val Phe Asn Asn Ala Thr Asn Glu Ser Tyr Val
 145 150 155 160
 Asp Thr Ala Ala Met Glu Ala Glu Arg Leu His Val Ser Ala Thr Leu
 165 170 175
 Ala Cys Leu Thr Ala Ile Ile Gln Met Gly Leu Gly Phe Met Gln Phe
 180 185 190
 Gly Phe Val Ala Ile Tyr Leu Ser Glu Ser Phe Ile Arg Gly Phe Met
 195 200 205
 Thr Ala Ala Gly Leu Gln Ile Leu Ile Ser Val Leu Lys Tyr Ile Phe
 210 215 220
 Gly Leu Thr Ile Pro Ser Tyr Thr Gly Pro Gly Ser Ile Val Phe Thr
 225 230 235 240
 Phe Ile Asp Ile Cys Lys Asn Leu Pro His Thr Asn Ile Ala Ser Leu
 245 250 255
 Ile Phe Ala Leu Ile Ser Gly Ala Phe Leu Val Leu Val Lys Glu Leu
 260 265 270
 Asn Ala Arg Tyr Met His Lys Ile Arg Phe Pro Ile Pro Thr Glu Met
 275 280 285
 Ile Val Val Val Val Ala Thr Ala Ile Ser Gly Gly Cys Lys Met Pro
 290 295 300
 Lys Lys Tyr His Met Gln Ile Val Gly Glu Ile Gln Arg Gly Phe Pro
 305 310 315 320
 Thr Pro Val Ser Pro Val Val Ser Gln Trp Lys Asp Met Ile Gly Thr
 325 330 335
 Ala Phe Ser Leu Ala Ile Val Ser Tyr Val Ile Asn Leu Ala Met Gly
 340 345 350
 Arg Thr Leu Ala Asn Lys His Gly Tyr Asp Val Asp Ser Asn Gln Glu
 355 360 365

[0055]

Met Ile Ala Leu Gly Cys Ser Asn Phe Phe Gly Ser Phe Phe Lys Ile
 370 375 380
 His Val Ile Cys Cys Ala Leu Ser Val Thr Leu Ala Val Asp Gly Ala
 385 390 395 400
 Gly Gly Lys Ser Gln Val Ala Ser Leu Cys Val Ser Leu Val Val Met
 405 410 415
 Ile Thr Met Leu Val Leu Gly Ile Tyr Leu Tyr Pro Leu Pro Lys Ser
 420 425 430
 Val Leu Gly Ala Leu Ile Ala Val Asn Leu Lys Asn Ser Leu Lys Gln
 435 440 445
 Leu Thr Asp Pro Tyr Tyr Leu Trp Arg Lys Ser Lys Leu Asp Cys Cys
 450 455 460
 Ile Trp Val Val Ser Phe Leu Ser Ser Phe Phe Leu Ser Leu Pro Tyr
 465 470 475 480
 Gly Val Ala Val Gly Val Ala Phe Ser Val Leu Val Val Val Phe Gln
 485 490 495
 Thr Gln Phe Arg Asn Gly Tyr Ala Leu Ala Gln Val Met Asp Thr Asp
 500 505 510
 Ile Tyr Val Asn Pro Lys Thr Tyr Asn Arg Ala Gln Asp Ile Gln Gly
 515 520 525
 Ile Lys Ile Ile Thr Tyr Cys Ser Pro Leu Tyr Phe Ala Asn Ser Glu
 530 535 540
 Ile Phe Arg Gln Lys Val Ile Ala Lys Thr Gly Met Asp Pro Gln Lys
 545 550 555 560
 Val Leu Leu Ala Lys Gln Lys Tyr Leu Lys Lys Gln Glu Lys Arg Arg
 565 570 575
 Met Arg Pro Thr Gln Gln Arg Arg Ser Leu Phe Met Lys Thr Lys Thr
 580 585 590
 Val Ser Leu Gln Glu Leu Gln Gln Asp Phe Glu Asn Ala Pro Pro Thr
 595 600 605
 Asp Pro Asn Asn Asn Gln Thr Pro Ala Asn Gly Thr Ser Val Ser Tyr
 610 615 620
 Ile Thr Phe Ser Pro Asp Ser Ser Ser Pro Ala Gln Ser Glu Pro Pro
 625 630 635 640
 Ala Ser Ala Glu Ala Pro Gly Glu Pro Ser Asp Met Leu Ala Ser Val
 645 650 655
 Pro Pro Phe Val Thr Phe His Thr Leu Ile Leu Asp Met Ser Gly Val
 660 665 670

[0056]

Ser Phe Val Asp Leu Met Gly Ile Lys Ala Leu Ala Lys Leu Ser Ser
675 680 685

Thr Tyr Gly Lys Ile Gly Val Lys Val Phe Leu Val Asn Ile His Ala
690 695 700

Gln Val Tyr Asn Asp Ile Ser His Gly Gly Val Phe Glu Asp Gly Ser
705 710 715 720

Leu Glu Cys Lys His Val Phe Pro Ser Ile His Asp Ala Val Leu Phe
725 730 735

Ala Gln Ala Asn Ala Arg Asp Val Thr Pro Gly His Asn Phe Gln Gly
740 745 750

Ala Pro Gly Asp Ala Glu Leu Ser Leu Tyr Asp Ser Glu Glu Asp Ile
755 760 765

Arg Ser Tyr Trp Asp Leu Glu Gln Glu Met Phe Gly Ser Met Phe His
770 775 780

Ala Glu Thr Leu Thr Ala Leu
785 790

<210> 66

<211> 243

<212> PRT

<213> 人

<400> 66

Met Glu Gln Gly Ser Gly Arg Leu Glu Asp Phe Pro Val Asn Val Phe
1 5 10 15

Ser Val Thr Pro Tyr Thr Pro Ser Thr Ala Asp Ile Gln Val Ser Asp
20 25 30

Asp Asp Lys Ala Gly Ala Thr Leu Leu Phe Ser Gly Ile Phe Leu Gly
35 40 45

Leu Val Gly Ile Thr Phe Thr Val Met Gly Trp Ile Lys Tyr Gln Gly
50 55 60

Val Ser His Phe Glu Trp Thr Gln Leu Leu Gly Pro Val Leu Leu Ser
65 70 75 80

Val Gly Val Thr Phe Ile Leu Ile Ala Val Cys Lys Phe Lys Met Leu
85 90 95

Ser Cys Gln Leu Cys Lys Glu Ser Glu Glu Arg Val Pro Asp Ser Glu
100 105 110

Gln Thr Pro Gly Gly Pro Ser Phe Val Phe Thr Gly Ile Asn Gln Pro
115 120 125

[0057]

Ile Thr Phe His Gly Ala Thr Val Val Gln Tyr Ile Pro Pro Pro Tyr
130 135 140

Gly Ser Pro Glu Pro Met Gly Ile Asn Thr Ser Tyr Leu Gln Ser Val
145 150 155 160

Val Ser Pro Cys Gly Leu Ile Thr Ser Gly Gly Ala Ala Ala Ala Met
165 170 175

Ser Ser Pro Pro Gln Tyr Tyr Thr Ile Tyr Pro Gln Asp Asn Ser Ala
180 185 190

Phe Val Val Asp Glu Gly Cys Leu Ser Phe Thr Asp Gly Gly Asn His
195 200 205

Arg Pro Asn Pro Asp Val Asp Gln Leu Glu Glu Thr Gln Leu Glu Glu
210 215 220

Glu Ala Cys Ala Cys Phe Ser Pro Pro Pro Tyr Glu Glu Ile Tyr Ser
225 230 235 240

Leu Pro Arg

<210> 67
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：寡核苷酸

<400> 67
acacgaatgg tagatacagt g 21

<210> 68
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：寡核苷酸

<400> 68
atacttgtga gctgttccat g 21

<210> 69
<211> 21
<212> DNA
<213> 人工序列

[0058]

<220>
<223> 人工序列描述：寡核苷酸

<400> 69
actgttacct tgcattgact g 21

<210> 70

<211> 21

<212> DNA

<213> 人工序列

<220>
<223> 人工序列描述：寡核苷酸

<400> 70
caatgagaac acattggacat g 21

<210> 71

<211> 21

<212> DNA

<213> 人工序列

<220>
<223> 人工序列描述：寡核苷酸

<400> 71
ccattgaaagc tccattgtcta c 21

<210> 72

<211> 21

<212> DNA

<213> 人工序列

<220>
<223> 人工序列描述：寡核苷酸

<400> 72
agattatgca catattctgt c 21

<210> 73

<211> 21

<212> DNA

<213> 人工序列

<220>
<223> 人工序列描述：寡核苷酸

[0059]

<400> 73 atcggctgaa gtcaagcatc g	21
<210> 74 <211> 21 <212> DNA <213> 人工序列	
<220> <223> 人工序列描述: 寡核苷酸	
<400> 74 tggtcagtga ggactcagct g	21
<210> 75 <211> 21 <212> DNA <213> 人工序列	
<220> <223> 人工序列描述: 寡核苷酸	
<400> 75 tttctctgct tgatgcactt g	21
<210> 76 <211> 21 <212> DNA <213> 人工序列	
<220> <223> 人工序列描述: 寡核苷酸	
<400> 76 gtgagcactg ggaagcagct c	21
<210> 77 <211> 21 <212> DNA <213> 人工序列	
<220> <223> 人工序列描述: 寡核苷酸	
<400> 77 ggcaaattgct agagacgtga c	21

[0060]

<210>	78	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述：寡核苷酸	
<400>	78	
	agggtgtcctt cagctgccaa g	21
<210>	79	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述：寡核苷酸	
<400>	79	
	gttaagtgct ctctggattt g	21
<210>	80	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述：寡核苷酸	
<400>	80	
	atcctgattg ctgtgtgcaa g	21
<210>	81	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述：寡核苷酸	
<400>	81	
	ctcttctagc tggccaacat c	21
<210>	82	

[0061]

<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	82	
	ccagcaacaa cttacgtggt c	21
<210>	83	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述: 寡核苷酸	
<400>	83	
	cctttattca eccaatcact c	21
<210>	84	
<211>	2165	
<212>	DNA	
<213>	人	
<400>	84	
	agaacagegc agtttgccct ccgtcacgc agagcctctc cgtggcctcc gcaccttgag	60
	cattaggcca gttctectct tetctetaat ccatecgtea cctctcctgt catecgtttc	120
	catgecgtga ggtccattca cagaacacat ccatggetct catgetcagt ttggttctga	180
	gtctectcaa gtgggatca gggcagtggc aggtgtttgg gccagacaag cctgtccagg	240
	ccttggtggg ggaggacgca gcattctect gtttctgtc tcctaagacc aatgcagagg	300
	ccatggaagt gcggttcttc aggggccagt tctctagcgt ggtccacctc tacagggacg	360
	ggaaggacca gccatttatg cagatgccac agtatcaagg caggacaaaa ctggtgaagg	420
	attctattgc ggaggggcgc atctctctga ggetggaaaa cattactgtg ttggatgctg	480
	gcctctatgg gtgcaggatt agttcccagt cttaactacca gaaggccatc tgggagctac	540
	aggtgtcagc actgggctca gtctctctca ttccatcac gggatatgtt gatagagaca	600
	tccagctact ctgtcagtc tcgggctggt tcccccgcc cacagcgaag tggaaaggtc	660
	cacaaggaca ggatttgtcc acagactcca ggacaaacag agacatgcat ggctgtttg	720
	atgtggagat ctctctgacc gtccaagaga acgccgggag catatcctgt tccatcggc	780
	atgctcatct gagccgagag gtggaatcca gggtagagat aggagatacc tttttcgagc	840
	ctatatcgtg gcacctgget accaaagtac tgggaatact ctgctgtggc ctatTTTTTg	900
	gcattgttgg actgaagatt ttcttctcca aattccagtg taagcgagag agagaagcat	960

[0062]

```

gggcegggtgc cttattcatg gtccagcag ggacaggatc agagatgctc ccacatccag 1020
ctgtctctct tcttctagtc ctagcctcca ggggeccagg cccaaaaaag gaaaatccag 1080
gcggaactgg actggagaag aaagcacgga caggcagaat tgagagacgc ccggaaacac 1140
gcagtggagg tgactctgga tccagagacg gctcacccga agctctgcgt ttctgatctg 1200
aaaactgtaa cccatagaaa agctccccag gaggtgcctc actctgagaa gagatttaca 1260
aggaagagtg tgggtggcttc tcagagtttc caagcaggga aacattactg ggaggtggac 1320
ggaggacaca ataaaaggtg gcgcgtggga gtgtgccggg atgatgtgga caggaggaag 1380
gaglacgtga cttlglctcc cgatcatggg tactgggtcc tcagactgaa tggagaacat 1440
ttgtatttca cattaatatc ccgttttata agcgtcttcc ccaggacccc acctacaaaa 1500
ataggggtct tcttggaacta tgagtgtggg accatctcct tcttcaacat aaatgaccag 1560
tcccttattt atacctgac atgtcggttt gaaggettat tgagcccta cattgagtat 1620
ccgtcttata atgagcaaaa tggaactccc atagtcctct gcccagtcac ccaggaatca 1680
gagaaagagg cctcttggca aaggccctct gcaatcccag agacaagcaa cagtgagtcc 1740
tcttcacagg caaccacgcc ctctctcccc aggggtgaaa tgtaggatga atcacatccc 1800
acattcttct ttagggatat taaggtctct ctcccagatc caaagtcctg cagcagccgg 1860
ccaaggtggc ttccagatga agggggactg gcctgtccac atgggagtca ggtgtcatgg 1920
ctgccttgag ctgggaggga agaaggtga cattacattt agtttgcctc cactccatct 1980
ggctaagtga tcttgaaata ccacctctca ggtgaagaac cgtcaggaat tcccatctca 2040
caggctgtgg ttagatataa gtagacaagg aatgtgaata atgcttagat cttattgatg 2100
acagagtgtg tctaattggt ttgttcatta tattacattt tcagtaaaaa aaaaaaaaaa 2160
aaaaa 2165

```

<210> 85

<211> 347

<212> PRT

<213> 人

<400> 85

```

Met Ala Leu Met Leu Ser Leu Val Leu Ser Leu Leu Lys Leu Gly Ser
1          5          10          15

```

```

Gly Gln Trp Gln Val Phe Gly Pro Asp Lys Pro Val Gln Ala Leu Val
20          25          30

```

```

Gly Glu Asp Ala Ala Phe Ser Cys Phe Leu Ser Pro Lys Thr Asn Ala
35          40          45

```

```

Glu Ala Met Glu Val Arg Phe Phe Arg Gly Gln Phe Ser Ser Val Val
50          55          60

```

```

His Leu Tyr Arg Asp Gly Lys Asp Gln Pro Phe Met Gln Met Pro Gln
65          70          75          80

```

[0063]

Tyr Gln Gly Arg Thr Lys Leu Val Lys Asp Ser Ile Ala Glu Gly Arg
85 90 95

Ile Ser Leu Arg Leu Glu Asn Ile Thr Val Leu Asp Ala Gly Leu Tyr
100 105 110

Gly Cys Arg Ile Ser Ser Gln Ser Tyr Tyr Gln Lys Ala Ile Trp Glu
115 120 125

Leu Gln Val Ser Ala Leu Gly Ser Val Pro Leu Ile Ser Ile Thr Gly
130 135 140

Tyr Val Asp Arg Asp Ile Gln Leu Leu Cys Gln Ser Ser Gly Trp Phe
145 150 155 160

Pro Arg Pro Thr Ala Lys Trp Lys Gly Pro Gln Gly Gln Asp Leu Ser
165 170 175

Thr Asp Ser Arg Thr Asn Arg Asp Met His Gly Leu Phe Asp Val Glu
180 185 190

Ile Ser Leu Thr Val Gln Glu Asn Ala Gly Ser Ile Ser Cys Ser Met
195 200 205

Arg His Ala His Leu Ser Arg Glu Val Glu Ser Arg Val Gln Ile Gly
210 215 220

Asp Thr Phe Phe Glu Pro Ile Ser Trp His Leu Ala Thr Lys Val Leu
225 230 235 240

Gly Ile Leu Cys Cys Gly Leu Phe Phe Gly Ile Val Gly Leu Lys Ile
245 250 255

Phe Phe Ser Lys Phe Gln Cys Lys Arg Glu Arg Glu Ala Trp Ala Gly
260 265 270

Ala Leu Phe Met Val Pro Ala Gly Thr Gly Ser Glu Met Leu Pro His
275 280 285

Pro Ala Ala Ser Leu Leu Leu Val Leu Ala Ser Arg Gly Pro Gly Pro
290 295 300

Lys Lys Glu Asn Pro Gly Gly Thr Gly Leu Glu Lys Lys Ala Arg Thr
305 310 315 320

Gly Arg Ile Glu Arg Arg Pro Glu Thr Arg Ser Gly Gly Asp Ser Gly
325 330 335

Ser Arg Asp Gly Ser Pro Glu Ala Leu Arg Phe
340 345

<210> 86

<211> 21

<212> DNA

<213> 人工序列

[0064]

<220>		
<223>	人工序列描述：寡核苷酸	
<400>	86	
	attcatggtt ccagcaggga c	21
<210>	87	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述：寡核苷酸	
<400>	87	
	gggagacaaa gtcacgtact c	21
<210>	88	
<211>	22	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述：寡核苷酸	
<400>	88	
	tcctggtggtt cgtggtctgc tt	22
<210>	89	
<211>	22	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列描述：寡核苷酸	
<400>	89	
	gagagtctcg gcttttgtgg gc	22
<210>	90	
<211>	15	
<212>	PRT	
<213>	人	
<400>	90	
	Gly Ser Ser Asp Leu Thr Trp Pro Pro Ala Ile Lys Leu Gly Cys	
	1 5 10 15	
<210>	91	
<211>	16	
<212>	PRT	
<213>	人	
<400>	91	
	Asp Arg Tyr Val Ala Val Arg His Pro Leu Arg Ala Arg Gly Leu Arg	

[0065]

1	5	10	15
<210>	92		
<211>	15		
<212>	PRT		
<213>	人		
<400>	92		
Val Ala Pro Arg	Ala Lys Ala His Lys Ser Gln Asp Ser Leu Cys		
1	5	10	15
<210>	93		
<211>	13		
<212>	PRT		
<213>	人		
<400>	93		
Cys Phe Arg Ser Thr Arg His Asn Phe Asn Ser Met Arg			
1	5	10	
<210>	94		
<211>	22		
<212>	PRT		
<213>	人		
<400>	94		
Met Asn Gly Thr Tyr Asn Thr Cys Gly Ser Ser Asp Leu Thr Trp Pro			
1	5	10	15
Pro Ala Ile Lys Leu Gly			
	20		
<210>	95		
<211>	14		
<212>	PRT		
<213>	人		
<400>	95		
Arg Asp Thr Ser Asp Thr Pro Leu Cys Gln Leu Ser Gln Gly			
1	5	10	
<210>	96		
<211>	22		
<212>	PRT		
<213>	人		
<400>	96		
Gly Ile Gln Glu Gly Gly Phe Cys Phe Arg Ser Thr Arg His Asn Phe			
1	5	10	15
Asn Ser Met Arg Phe Pro			
	20		
<210>	97		
<211>	30		
<212>	PRT		
<213>	人		
<400>	97		
Ala Lys Glu Phe Gln Glu Ala Ser Ala Leu Ala Val Ala Pro Arg Ala			
1	5	10	15

[0066]

Lys Ala His Lys Ser Gln Asp Ser Leu Cys Val Thr Leu Ala
20 25 30

<210> 98
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述: 寡核苷酸

<400> 98
tcctgctcgt cgtctcctg at 22

<210> 99
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述: 寡核苷酸

<400> 99
tcgctttttg tcgtatttgc 20

<210> 100
<211> 15
<212> PRT
<213> 人

<400> 100

His Asn Gly Ser Tyr Glu Ile Ser Val Leu Met Met Gly Asn Ser
1 5 10 15

<210> 101
<211> 15
<212> PRT
<213> 人

<400> 101

Asn Leu Pro Thr Pro Thr Val Glu Asn Gln Gln Arg Leu Ala
1 5 10 15

<210> 102
<211> 619
<212> PRT
<213> 人

<400> 102

Arg Lys Tyr Arg Lys Asp Tyr Glu Leu Arg Gln Lys Lys Trp Ser His
1 5 10 15

Ile Pro Pro Glu Asn Ile Phe Pro Leu Glu Thr Asn Glu Thr Asn His
20 25 30

Val Ser Leu Lys Ile Asp Asp Asp Lys Arg Arg Asp Thr Ile Gln Arg
35 40 45

Leu Arg Gln Cys Lys Tyr Asp Lys Lys Arg Val Ile Leu Lys Asp Leu
50 55 60

[0067]

Lys 65	His	Asn	Asp	Gly	Asn 70	Phe	Thr	Glu	Lys	Gln 75	Lys	Ile	Glu	Leu	Asn 80
Lys	Leu	Leu	Gln	Ile	Asp	Tyr	Tyr	Asn	Leu 90	Thr	Lys	Phe	Tyr	Gly	Thr 95
Val	Lys	Leu	Asp	Thr	Met	Ile	Phe	Gly	Val 100	Ile	Glu	Tyr	Cys	Glu	Arg 110
Gly	Ser	Leu	Arg	Glu	Val	Leu	Asn	Asp	Thr	Ile	Ser	Tyr	Pro	Asp	Gly 115
Thr	Phe	Met	Asp	Trp	Glu	Phe	Lys	Ile	Ser	Val	Leu	Tyr	Asp	Ile	Ala 130
Lys	Gly	Met	Ser	Tyr	Leu	His	Ser	Ser	Lys	Thr	Glu	Val	His	Gly	Arg 145
Leu	Lys	Ser	Thr	Asn	Cys	Val	Val	Asp	Ser	Arg	Met	Val	Val	Lys	Ile 165
Thr	Asp	Phe	Gly	Cys	Asn	Ser	Ile	Leu	Pro	Pro	Lys	Lys	Asp	Leu	Trp 180
Thr	Ala	Pro	Glu	His	Leu	Arg	Gln	Ala	Asn	Ile	Ser	Gln	Lys	Gly	Asp 195
Val	Tyr	Ser	Tyr	Gly	Ile	Ile	Ala	Gln	Glu	Ile	Ile	Leu	Arg	Lys	Glu 210
Thr	Phe	Tyr	Thr	Leu	Ser	Cys	Arg	Asp	Arg	Asn	Glu	Lys	Ile	Phe	Arg 225
Val	Glu	Asn	Ser	Asn	Gly	Met	Lys	Pro	Phe	Arg	Pro	Asp	Leu	Phe	Leu 245
Glu	Thr	Ala	Glu	Glu	Lys	Glu	Leu	Glu	Val	Tyr	Leu	Leu	Val	Lys	Asn 260
Cys	Trp	Glu	Glu	Asp	Pro	Glu	Lys	Arg	Pro	Asp	Phe	Lys	Lys	Ile	Glu 275
Thr	Thr	Leu	Ala	Lys	Ile	Phe	Gly	Leu	Phe	His	Asp	Gln	Lys	Asn	Glu 290
Ser	Tyr	Met	Asp	Thr	Leu	Ile	Arg	Arg	Leu	Gln	Leu	Tyr	Ser	Arg	Asn 305
Leu	Glu	His	Leu	Val	Glu	Glu	Arg	Thr	Gln	Leu	Tyr	Lys	Ala	Glu	Arg 325
Asp	Arg	Ala	Asp	Arg	Leu	Asn	Phe	Met	Leu	Leu	Pro	Arg	Leu	Val	Val 340
Lys	Ser	Leu	Lys	Glu	Lys	Gly	Phe	Val	Glu	Pro	Glu	Leu	Tyr	Glu	Glu 355

[0068]

Val Thr Ile Tyr Phe Ser Asp Ile Val Gly Phe Thr Thr Ile Cys Lys
370 375 380

Tyr Ser Thr Pro Met Glu Val Val Asp Met Leu Asn Asp Ile Tyr Lys
385 390 395 400

Ser Phe Asp His Ile Val Asp His His Asp Val Tyr Lys Val Glu Thr
405 410 415

Ile Gly Asp Ala Tyr Met Val Ala Ser Gly Leu Pro Lys Arg Asn Gly
420 425 430

Asn Arg His Ala Ile Asp Ile Ala Lys Met Ala Leu Glu Ile Leu Ser
435 440 445

Phe Met Gly Thr Phe Glu Leu Glu His Leu Pro Gly Leu Pro Ile Trp
450 455 460

Ile Arg Ile Gly Val His Ser Gly Pro Cys Ala Ala Gly Val Val Gly
465 470 475 480

Ile Lys Met Pro Arg Tyr Cys Leu Phe Gly Asp Thr Val Asn Thr Ala
485 490 495

Ser Arg Met Glu Ser Thr Gly Leu Pro Leu Arg Ile His Val Ser Gly
500 505 510

Ser Thr Ile Ala Ile Leu Lys Arg Thr Glu Cys Gln Phe Leu Tyr Glu
515 520 525

Val Arg Gly Glu Thr Tyr Leu Lys Gly Arg Gly Asn Glu Thr Thr Tyr
530 535 540

Trp Leu Thr Gly Met Lys Asp Gln Lys Phe Asn Leu Pro Thr Pro Pro
545 550 555 560

Thr Val Glu Asn Gln Gln Arg Leu Gln Ala Glu Phe Ser Asp Met Ile
565 570 575

Ala Asn Ser Leu Gln Lys Arg Gln Ala Ala Gly Ile Arg Ser Gln Lys
580 585 590

Pro Arg Arg Val Ala Ser Tyr Lys Lys Gly Thr Leu Glu Tyr Leu Gln
595 600 605

Leu Asn Thr Thr Asp Lys Glu Ser Thr Tyr Phe
610 615

<210> 103
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述: 寡核苷酸

<400> 103
gctggttaact atcttcctgc

20

[0069]

$\langle 210 \rangle$ 110

126

<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述: 寡核苷酸

<400> 110
tggttgcttt ggcagagtcc 20

<210> 111
<211> 56
<212> PRT
<213> 人

<400> 111

Thr Gly Met Asp Met Trp Ser Thr Gln Asp Leu Tyr Asp Asn Pro Val
1 5 10 15

Thr Ser Val Phe Gln Tyr Glu Gly Leu Trp Arg Ser Cys Val Arg Gln
20 25 30

Ser Ser Gly Phe Thr Glu Cys Arg Pro Tyr Phe Thr Ile Leu Gly Leu
35 40 45

Pro Ala Met Leu Gln Ala Val Arg
50 55

<210> 112
<211> 53
<212> PRT
<213> 人

<400> 112

Asp Gln Trp Ser Thr Gln Asp Leu Tyr Asn Asn Pro Val Thr Ala Val
1 5 10 15

Phe Asn Tyr Gln Gly Leu Trp Arg Ser Cys Val Arg Glu Ser Ser Gly
20 25 30

Phe Thr Glu Cys Arg Gly Tyr Phe Thr Leu Leu Gly Leu Pro Ala Met
35 40 45

Leu Gln Ala Val Arg
50

<210> 113
<211> 14
<212> PRT
<213> 人

<400> 113

Ser Thr Gln Asp Leu Tyr Asn Asn Pro Val Thr Ala Val Phe
1 5 10

<210> 114
<211> 12
<212> PRT
<213> 人

<400> 114

[0071]

Asp Met Trp Ser Thr Gln Asp Leu Tyr Asp Asn Pro
1 5 10

<210> 115
<211> 12
<212> PRT
<213> 人

<400> 115

Cys Arg Pro Tyr Phe Thr Ile Leu Gly Leu Pro Ala
1 5 10

<210> 116
<211> 13
<212> PRT
<213> 人

<400> 116

Thr Asn Phe Trp Met Ser Thr Ala Asn Met Tyr Thr Gly
1 5 10

<210> 117
<211> 816
<212> DNA
<213> 人

<400> 117
gccaggatca tgtccaccac cacatgccaa gtggtggcgt tcctcctgtc ctccttgggg 60
ctggccggct gcacgcggc caccgggatg gacatgtgga gcaccagga cctgtacgac 120
aaceccgtca cctccgtgtt ccagtaagaa gggctctgga ggagctgcgt gaggcagagt 180
tcaggcttca ccgaatgcag gccctatttc accatcctgg gacttcacag catgctgcag 240
gcagtgcgag ccctgatgat cgtaggcacc gtcttgggtg ccattggcct cctgggtatcc 300
atctttgccc tgaaatgcat ccgcattggc agcatggagg actctgccaa agccaacatg 360
acactgacct ccgggatcat gttcattgtc tcagggtctt gtgcaattgc tggagtgtct 420
gtgtttgcc aacatgctgt gactaacttc tggatgtcca cagctaacat gtacaccggc 480
atgggtggga tgggtcagac tgttcagacc aggtacacat ttggtgcggc tctgttcgtg 540
ggctgggtcg ctggaggcct cacactaatt gggggtgtga tgatgtgcat cgcctgccgg 600
ggcctggcac cagaagaaac caactacaaa gccgtttctt atcatgcctc aggccacagt 660
gttgcttaca agcctggagg cttcaaggcc agcactggct ttgggtccaa caccaaaaac 720
aagaagatat acgatggagg tgccgcaca gaggacgagg tacaatttta tccttccaag 780
cacgactatg tgtaatgctc taagacctct cagcac 816

<210> 118
<211> 261
<212> PRT
<213> 人

<400> 118

Met Ser Thr Thr Thr Cys Gln Val Val Ala Phe Leu Leu Ser Ile Leu
1 5 10 15

Gly Leu Ala Gly Cys Ile Ala Ala Thr Gly Met Asp Met Trp Ser Thr
20 25 30

[0072]

Gln Asp Leu Tyr Asp Asn Pro Val Thr Ser Val Phe Gln Tyr Glu Gly	
35 40 45	
Leu Trp Arg Ser Cys Val Arg Gln Ser Ser Gly Phe Thr Glu Cys Arg	
50 55 60	
Pro Tyr Phe Thr Ile Leu Gly Leu Pro Ala Met Leu Gln Ala Val Arg	
65 70 75 80	
Ala Leu Met Ile Val Gly Ile Val Leu Gly Ala Ile Gly Leu Leu Val	
85 90 95	
Ser Ile Phe Ala Leu Lys Cys Ile Arg Ile Gly Ser Met Glu Asp Ser	
100 105 110	
Ala Lys Ala Asn Met Thr Leu Thr Ser Gly Ile Met Phe Ile Val Ser	
115 120 125	
Gly Leu Cys Ala Ile Ala Gly Val Ser Val Phe Ala Asn Met Leu Val	
130 135 140	
Thr Asn Phe Trp Met Ser Thr Ala Asn Met Tyr Thr Gly Met Gly Gly	
145 150 155 160	
Met Val Gln Thr Val Gln Thr Arg Tyr Thr Phe Gly Ala Ala Leu Phe	
165 170 175	
Val Gly Trp Val Ala Gly Gly Leu Thr Leu Ile Gly Gly Val Met Met	
180 185 190	
Cys Ile Ala Cys Arg Gly Leu Ala Pro Glu Glu Thr Asn Tyr Lys Ala	
195 200 205	
Val Ser Tyr His Ala Ser Gly His Ser Val Ala Tyr Lys Pro Gly Gly	
210 215 220	
Phe Lys Ala Ser Thr Gly Phe Gly Ser Asn Thr Lys Asn Lys Lys Ile	
225 230 235 240	
Tyr Asp Gly Gly Ala Arg Thr Glu Asp Glu Val Gln Ser Tyr Pro Ser	
245 250 255	
Lys His Asp Tyr Val	
260	
<210> 119	
<211> 227	
<212> DNA	
<213> 人	
<400> 119	
gccaggatca tgtccaccac cacatgccaa gtggtggcgt tcttcctgtc catcctgggg	60
ctggccggct gcatcgcggc caccgggatg gacatgtgga gcaccaggga cctgtacgac	120
aaccccgta cctccgtgtt ccagtagaa gggctctgga ggagctgcgt gaggcagagt	180
tcaggcttca ccgaatgcag gccctatttc accatcctgg gacttcc	227
<210> 120	

[0073]

<211> 69
<212> PRT
<213> 人

<400> 120
Met Ser Thr Thr Thr Cys Gln Val Val Ala Phe Leu Leu Ser Ile Leu
1 5 10 15

Gly Leu Ala Gly Cys Ile Ala Ala Thr Gly Met Asp Met Trp Ser Thr
20 25 30

Gln Asp Leu Tyr Asp Asn Pro Val Thr Ser Val Phe Gln Tyr Glu Gly
35 40 45

Leu Trp Arg Ser Cys Val Arg Gln Ser Ser Gly Phe Thr Glu Cys Arg
50 55 60

Pro Tyr Phe Thr Ile
65

<210> 121
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：寡核苷酸

<400> 121
aatgagagga aagagaaaac 20

<210> 122
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：寡核苷酸

<400> 122
atggtagaag agtaggcaat 20

<210> 123
<211> 15
<212> PRT
<213> 人

<400> 123
Glu Lys Trp Asn Leu His Lys Arg Ile Ala Leu Lys Met Val Cys
1 5 10 15

<210> 124
<211> 11
<212> PRT
<213> 人

<400> 124
Cys Leu Gly Phe Asn Phe Lys Glu Met Phe Lys
1 5 10

<210> 125
<211> 23

[0074]

<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述: 寡核苷酸	
<400> 125	
taatgatgaa ccctacactg agc	23
<210> 126	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述: 寡核苷酸	
<400> 126	
atggacaaat gccctacctt	20
<210> 127	
<211> 22	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述: 寡核苷酸	
<400> 127	
agtgtggaa ggatgtgcgt gt	22
<210> 128	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述: 寡核苷酸	
<400> 128	
ttgaggtggt tgttgggttt	20
<210> 129	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述: 寡核苷酸	
<400> 129	
agatgtgctg aggctgtaga	20
<210> 130	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述: 寡核苷酸	
<400> 130	
atgaaggttg attatttgag	20

[0075]

<210> 131
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述: 寡核苷酸

<400> 131
agccgcatac tcccttaccc tct 23

<210> 132
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述: 寡核苷酸

<400> 132
gcagcagccc aaacaccaca 20

<210> 133
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述: 寡核苷酸

<400> 133
ctgagccgag aggtggaatc 20

<210> 134
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述: 寡核苷酸

<400> 134
ctctctcgc t tacactggaa 20

<210> 135
<211> 14
<212> PRT
<213> 人

<400> 135
Gln Trp Gln Val Phe Gly Pro Asp Lys Pro Val Gln Ala Leu
1 5 10

<210> 136
<211> 15
<212> PRT
<213> 人

<400> 136
Ala Lys Trp Lys Gly Pro Gln Gly Gln Asp Leu Ser Thr Asp Ser
1 5 10 15

<210> 137
<211> 32

[0076]

<212> PRT

<213> 人

<400> 137

Asn Met Leu Val Thr Asn Phe Trp Met Ser Thr Ala Asn Met Tyr Thr
1 5 10 15

Gly Met Gly Gly Met Val Gln Thr Val Gln Thr Arg Tyr Thr Phe Gly
20 25 30

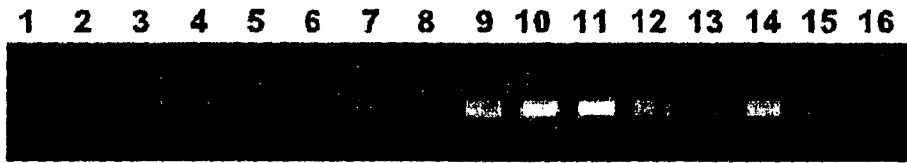


图 1

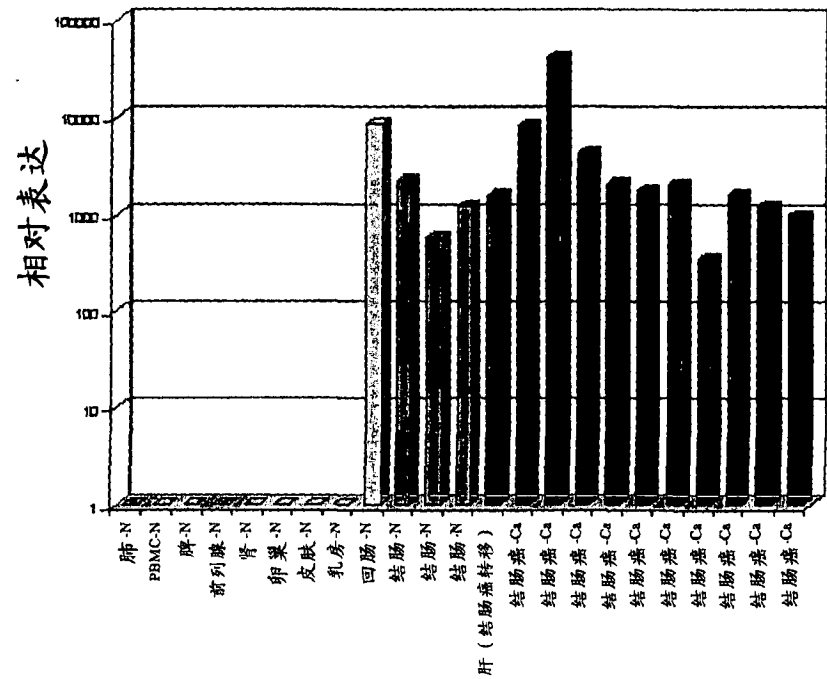


图 2

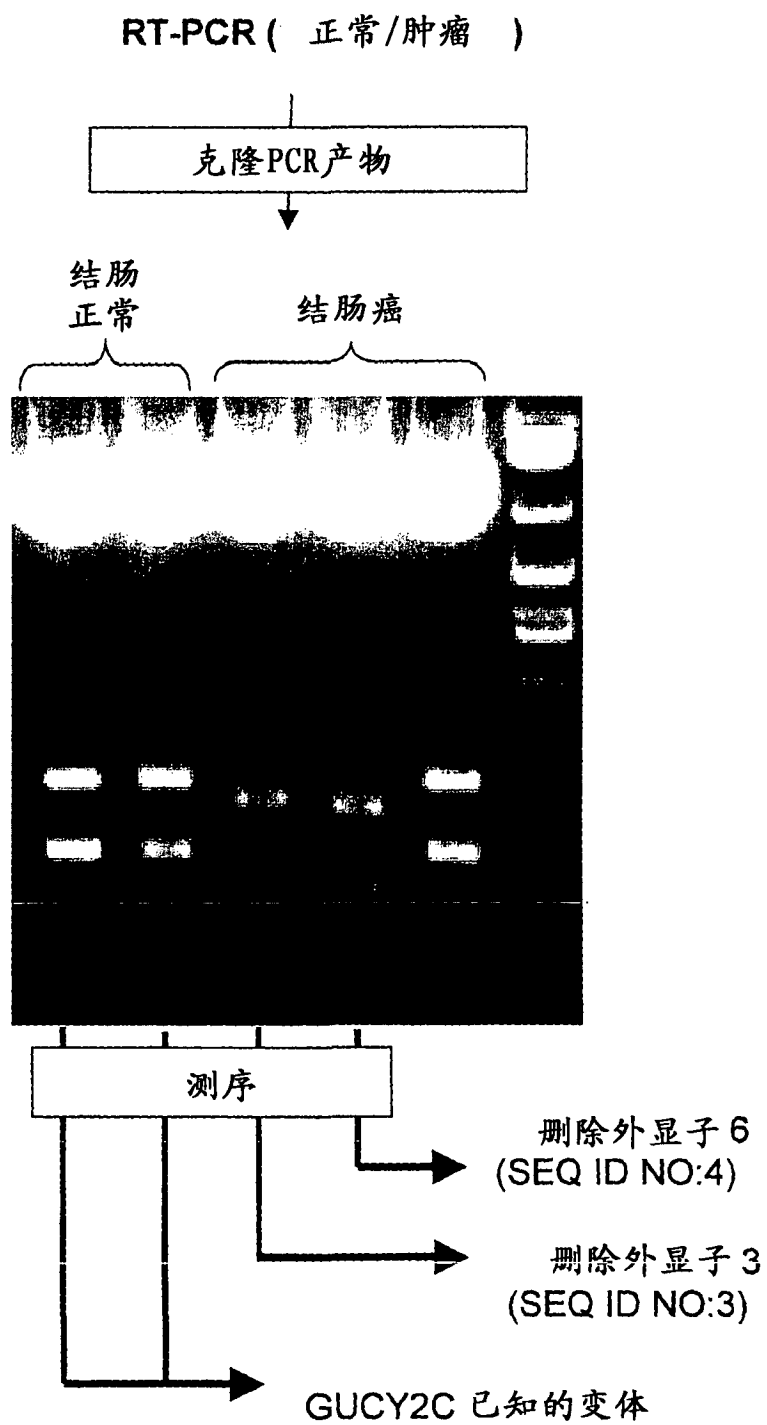


图 3

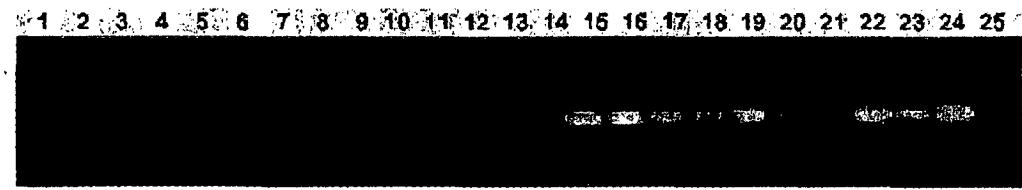


图 4



图 5

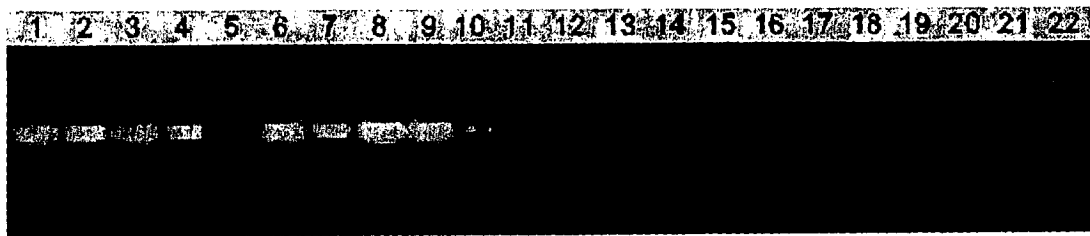


图 6

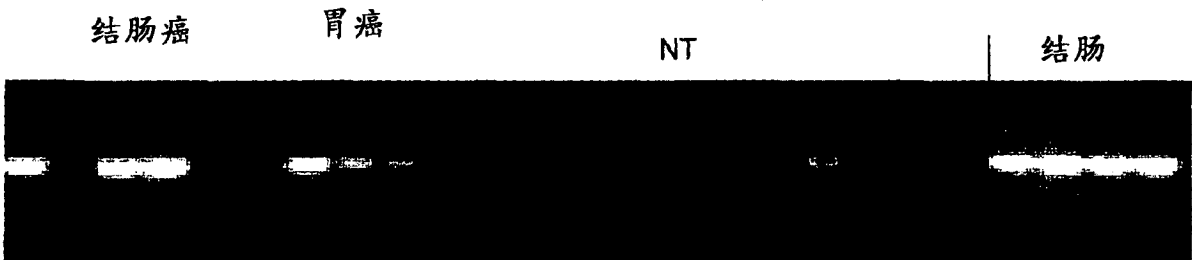


图 7



图 8

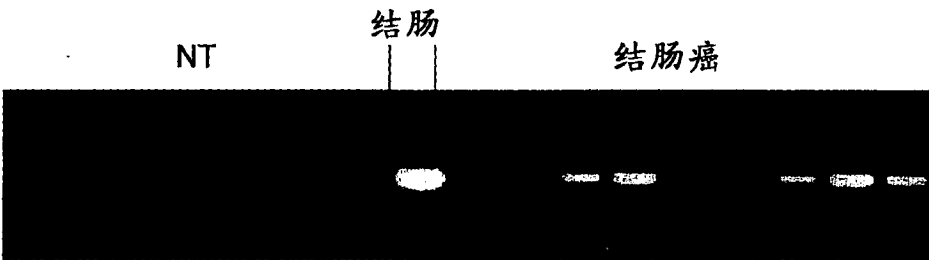


图 9

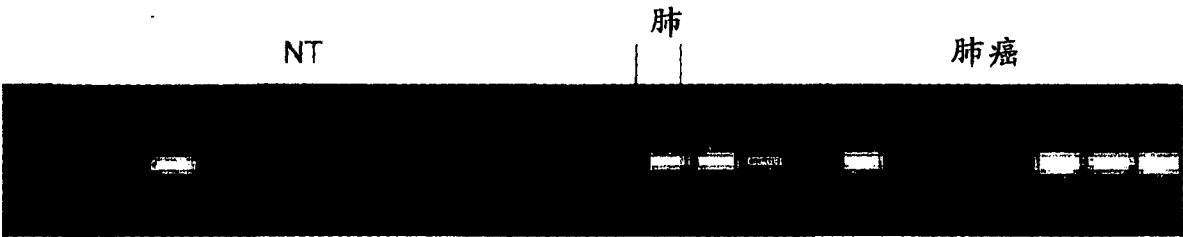


图 10

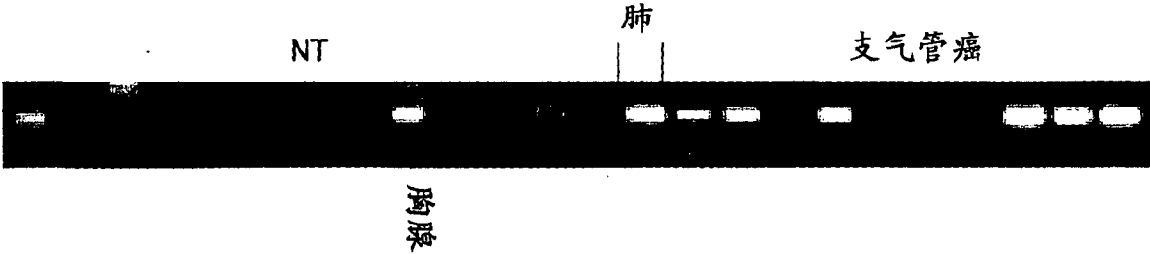


图 11

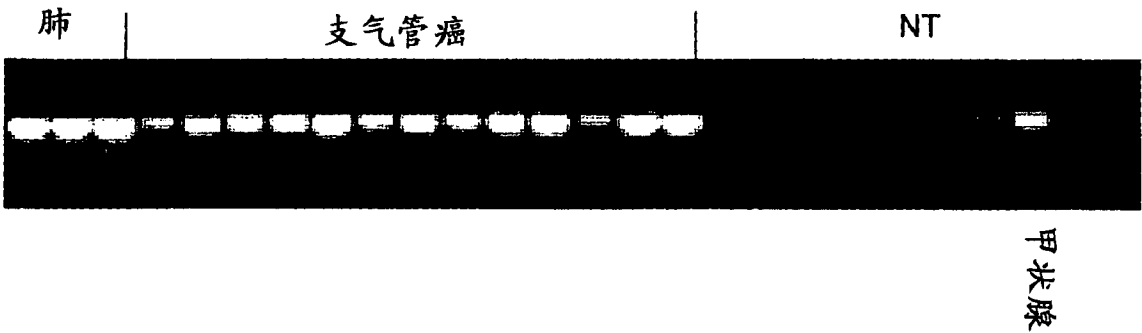


图 12

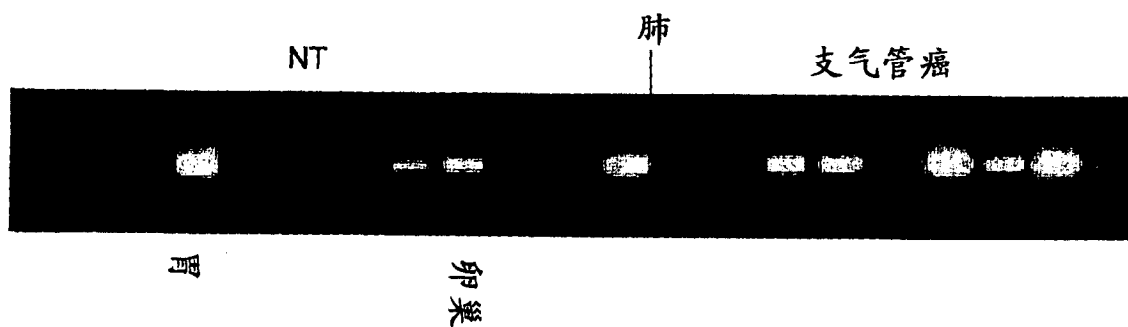


图 13

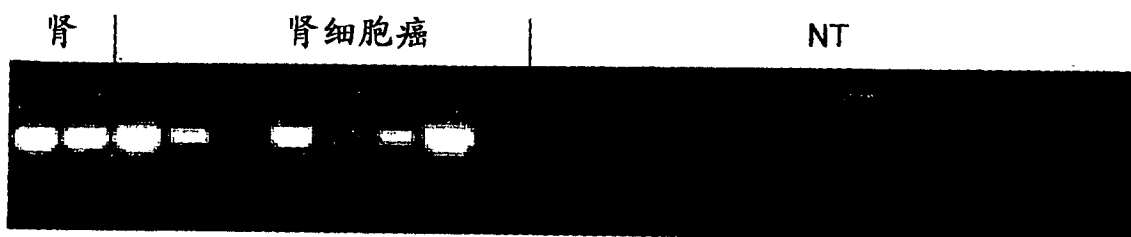


图 14

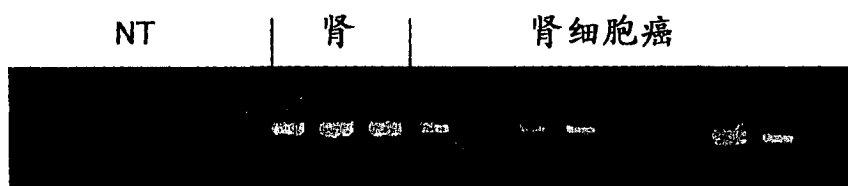


图 15



图 16

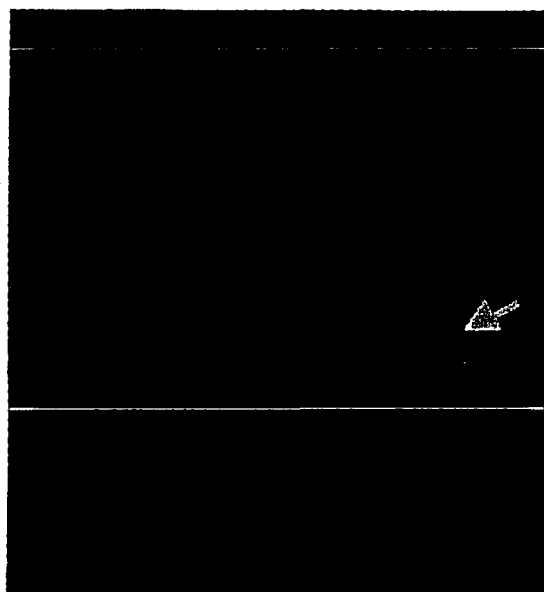


图 17

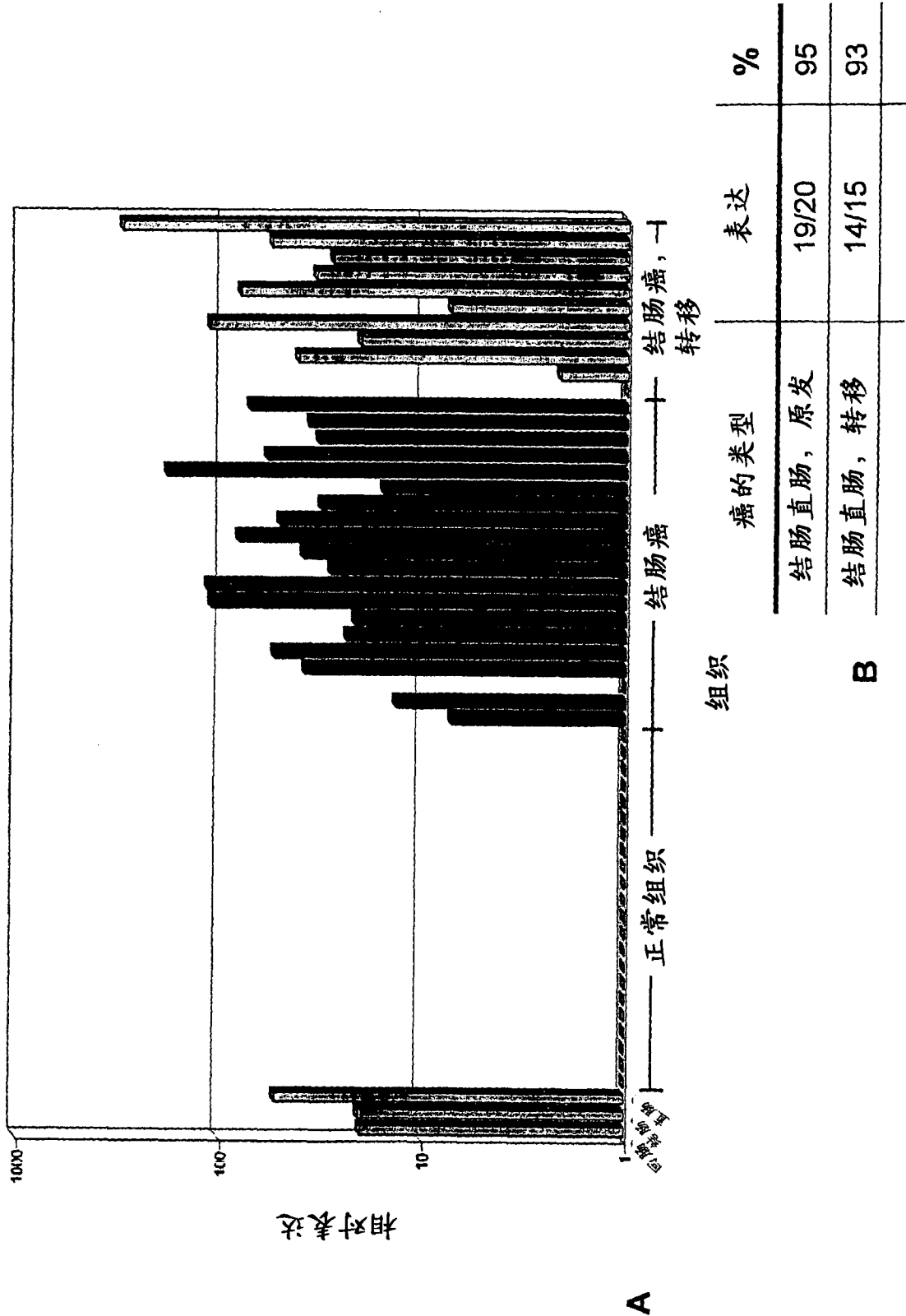


图 18

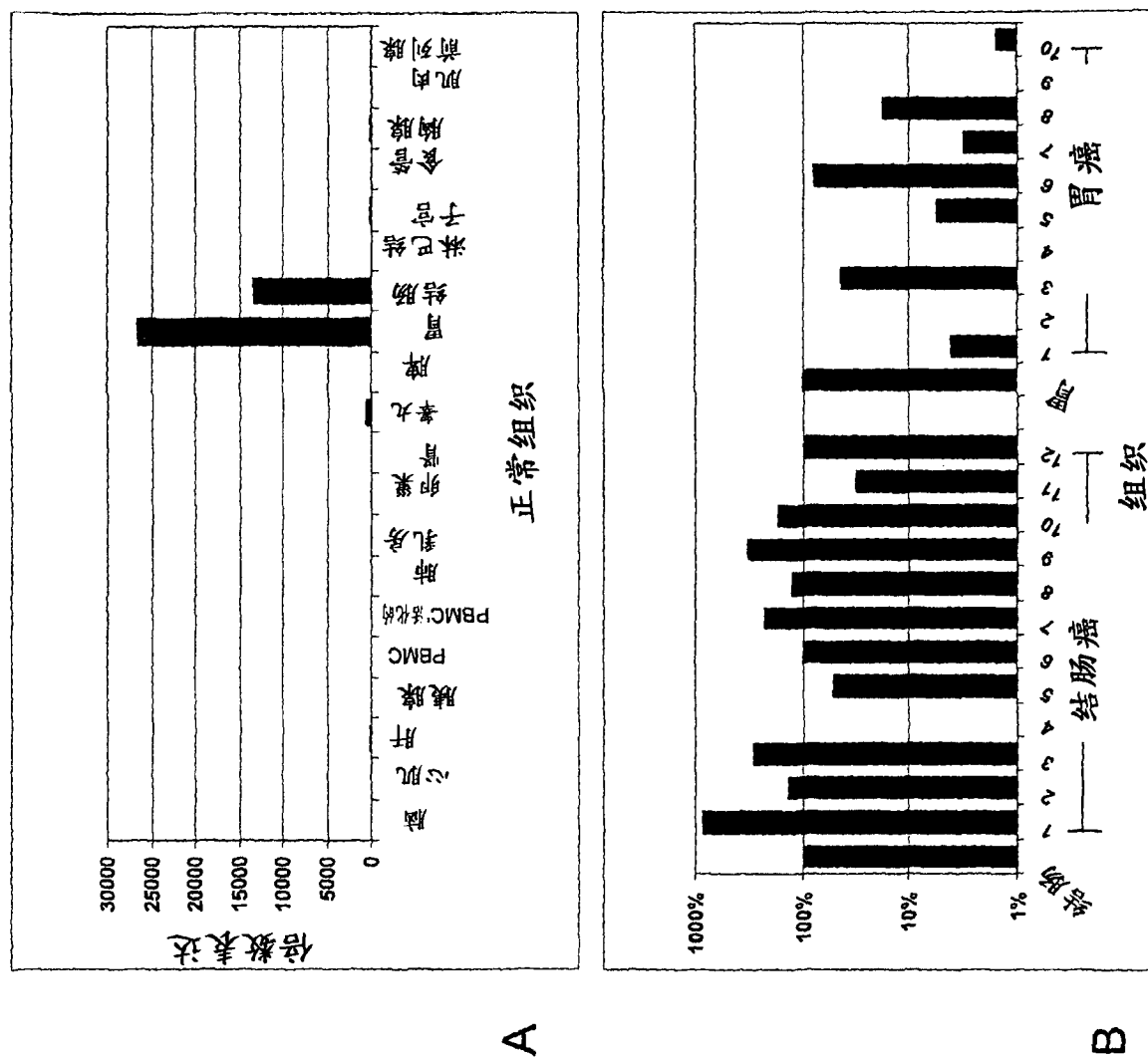


图 19

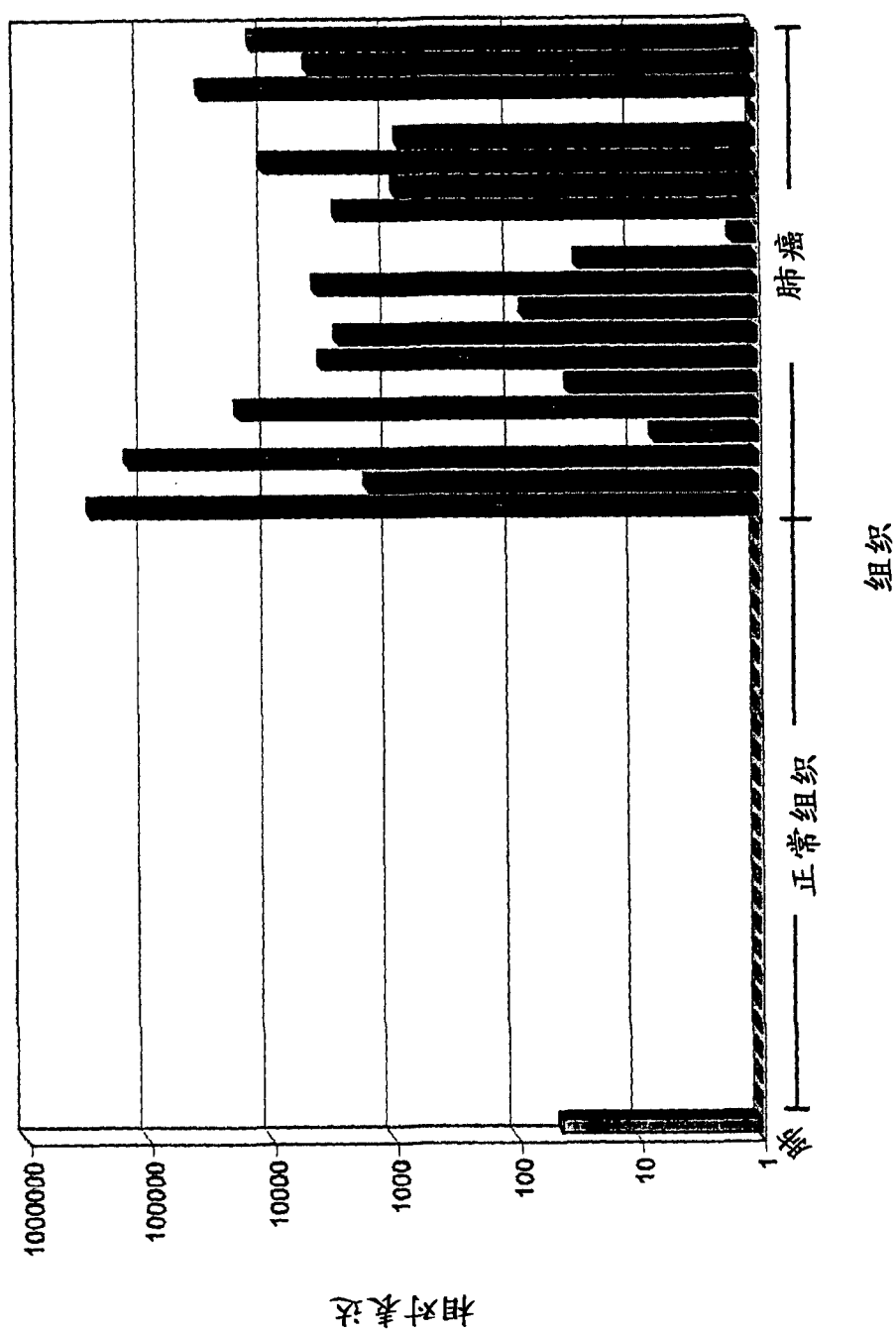


图 20

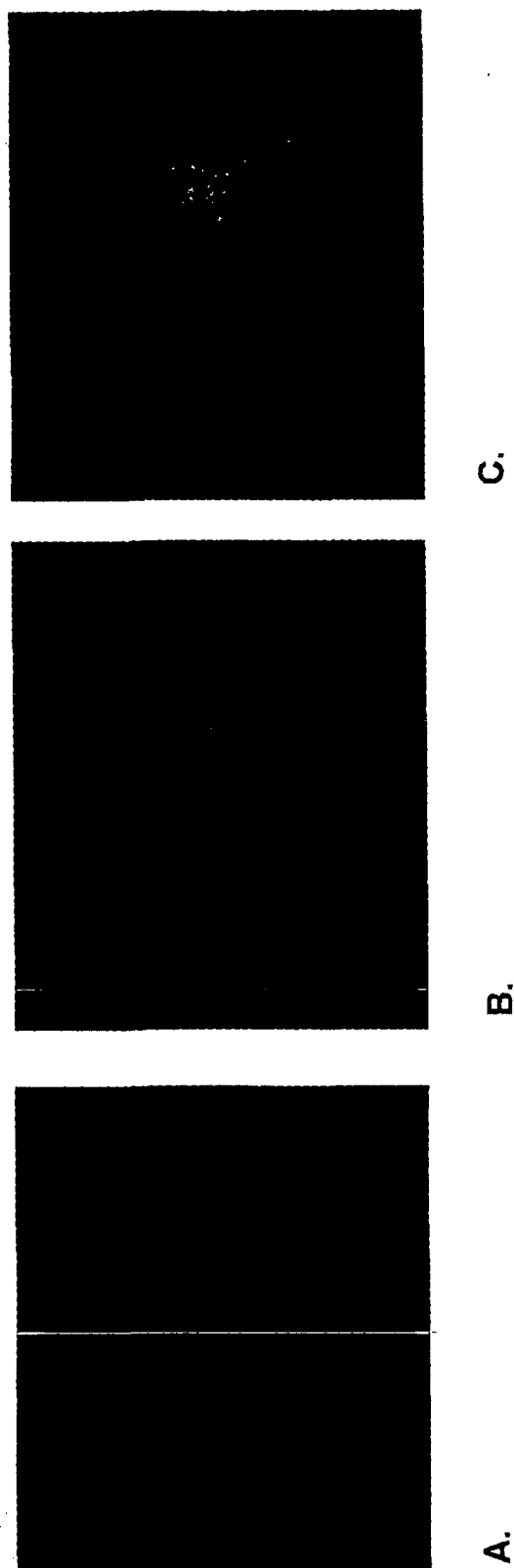


图 21

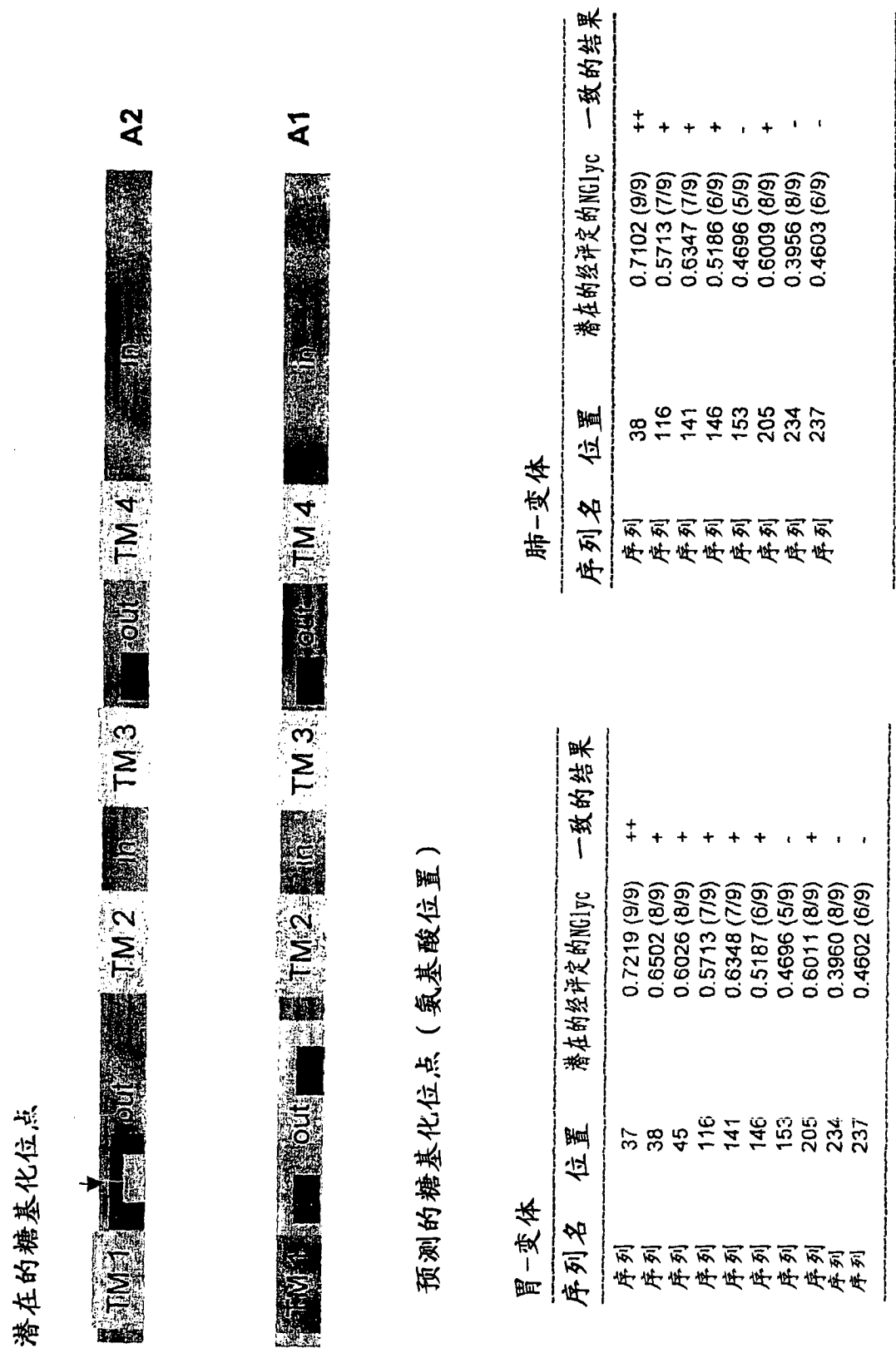


图 22

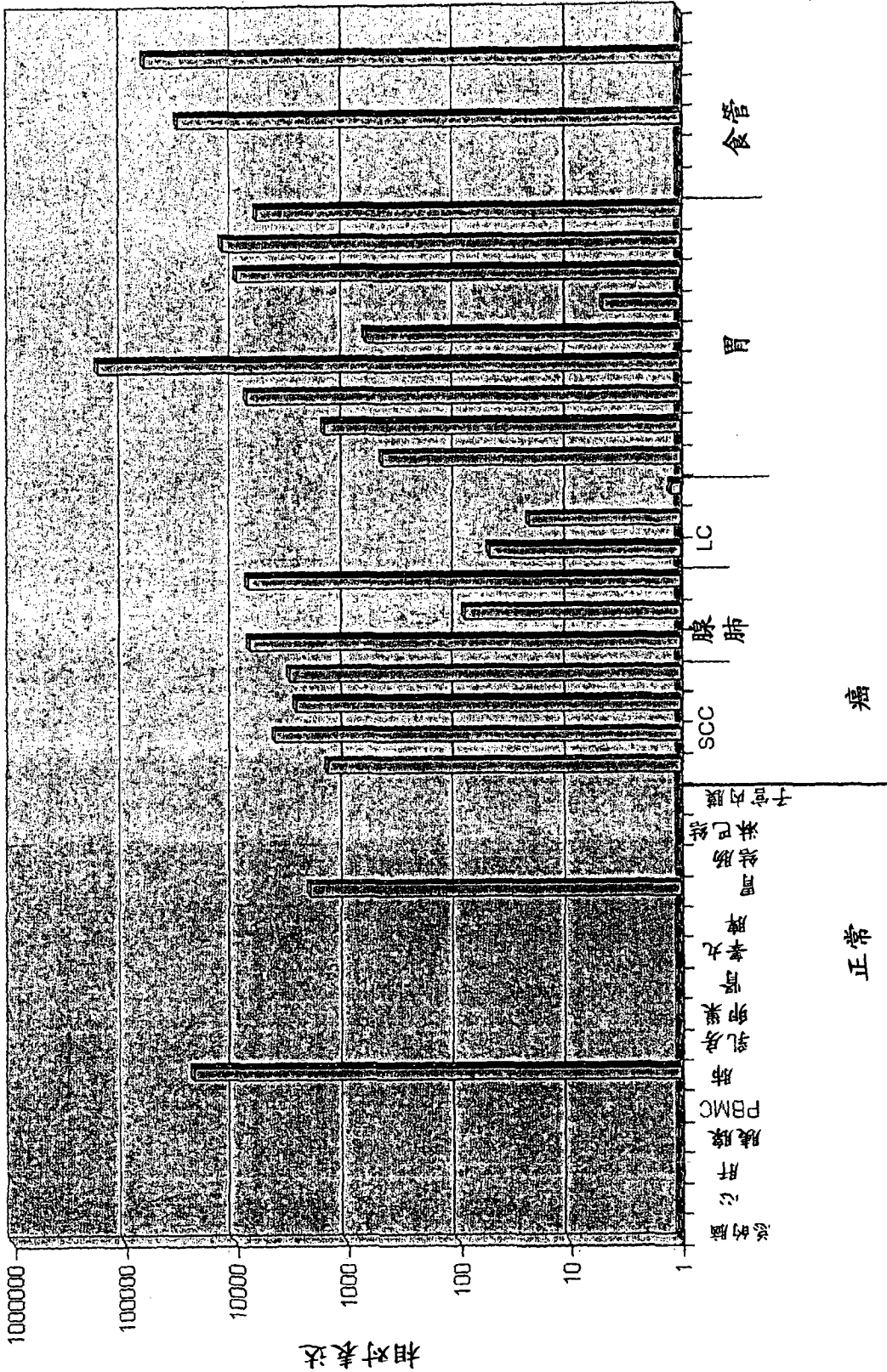


图 23

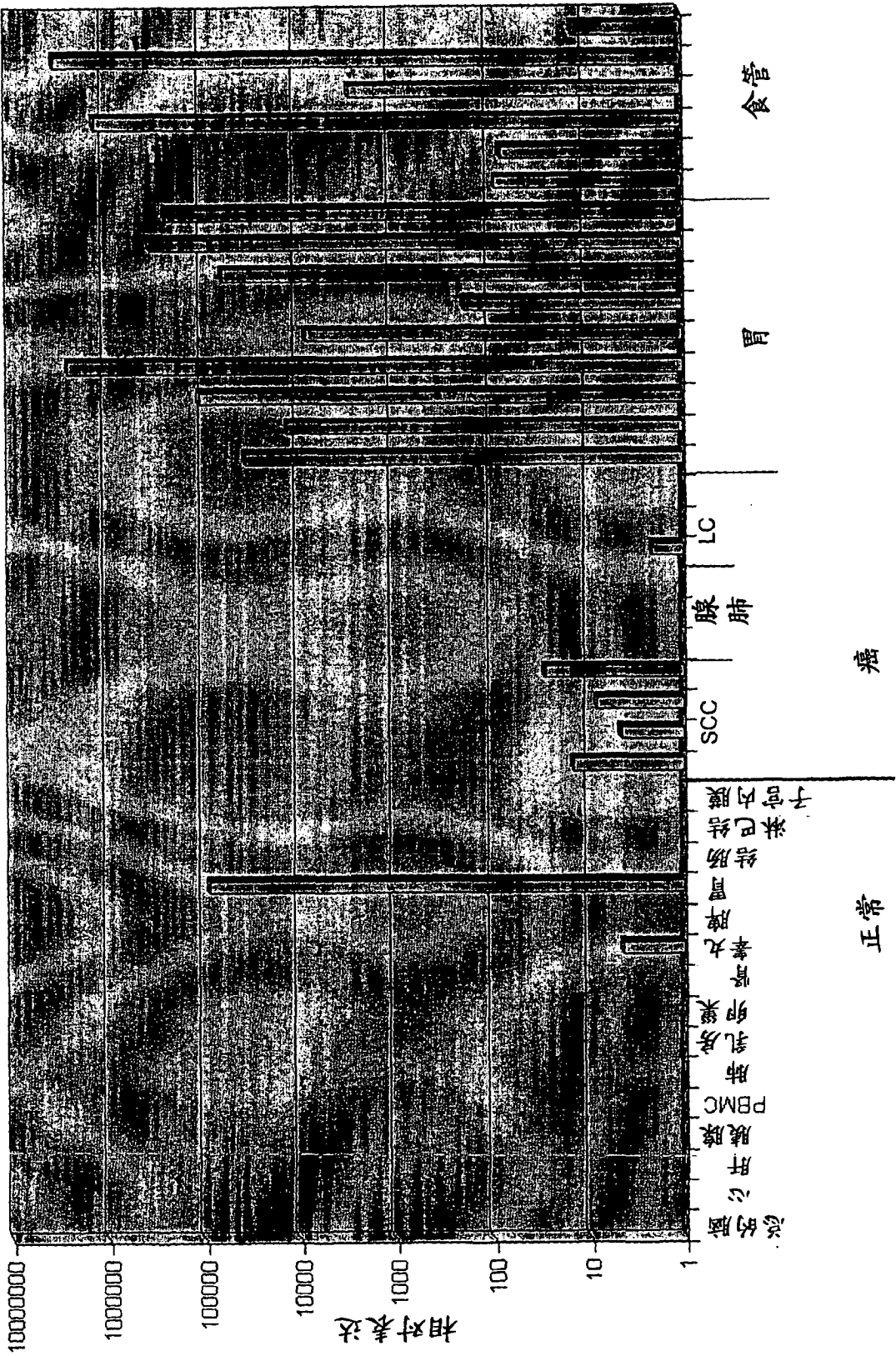


图 24

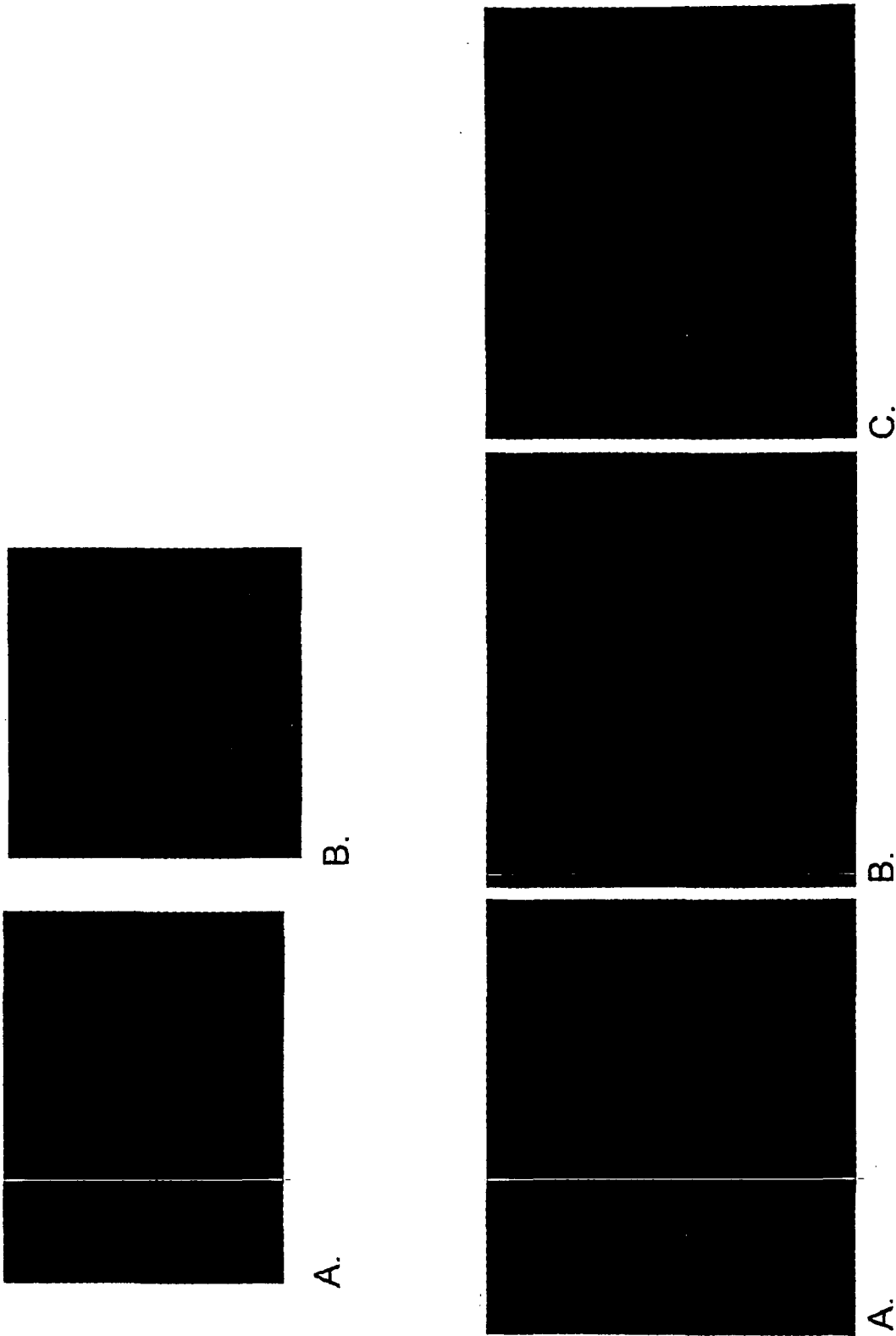


图 25

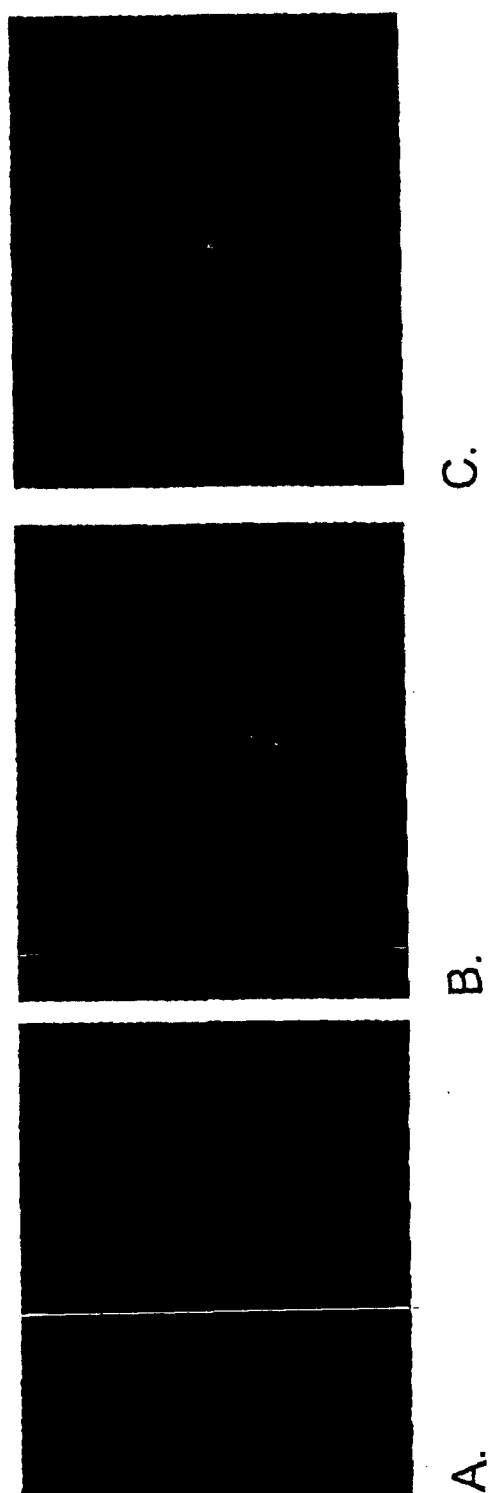


图 26

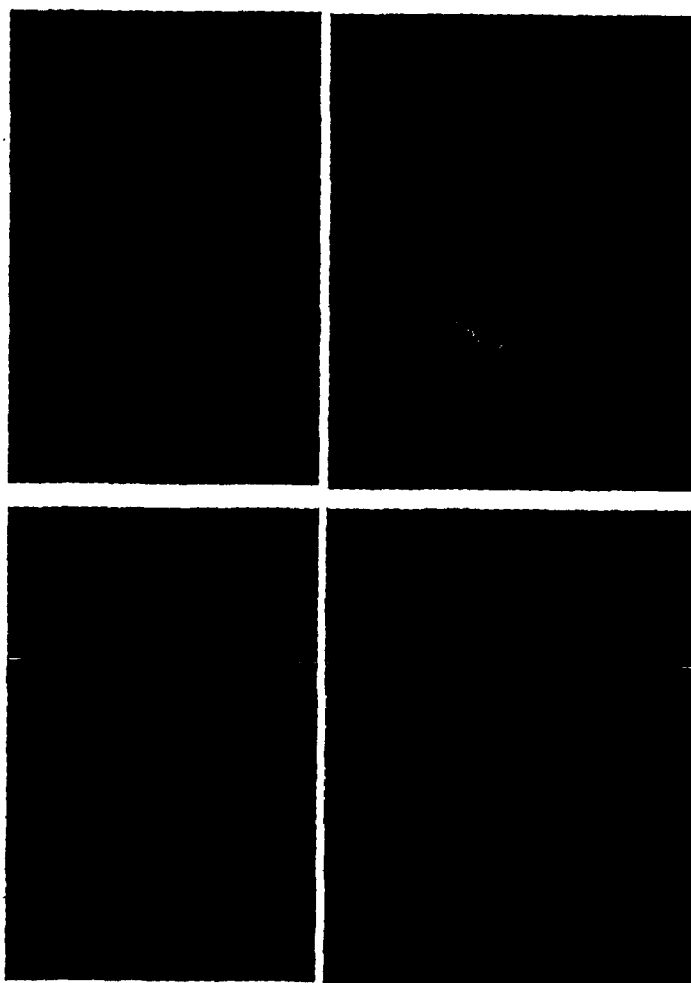


图 27

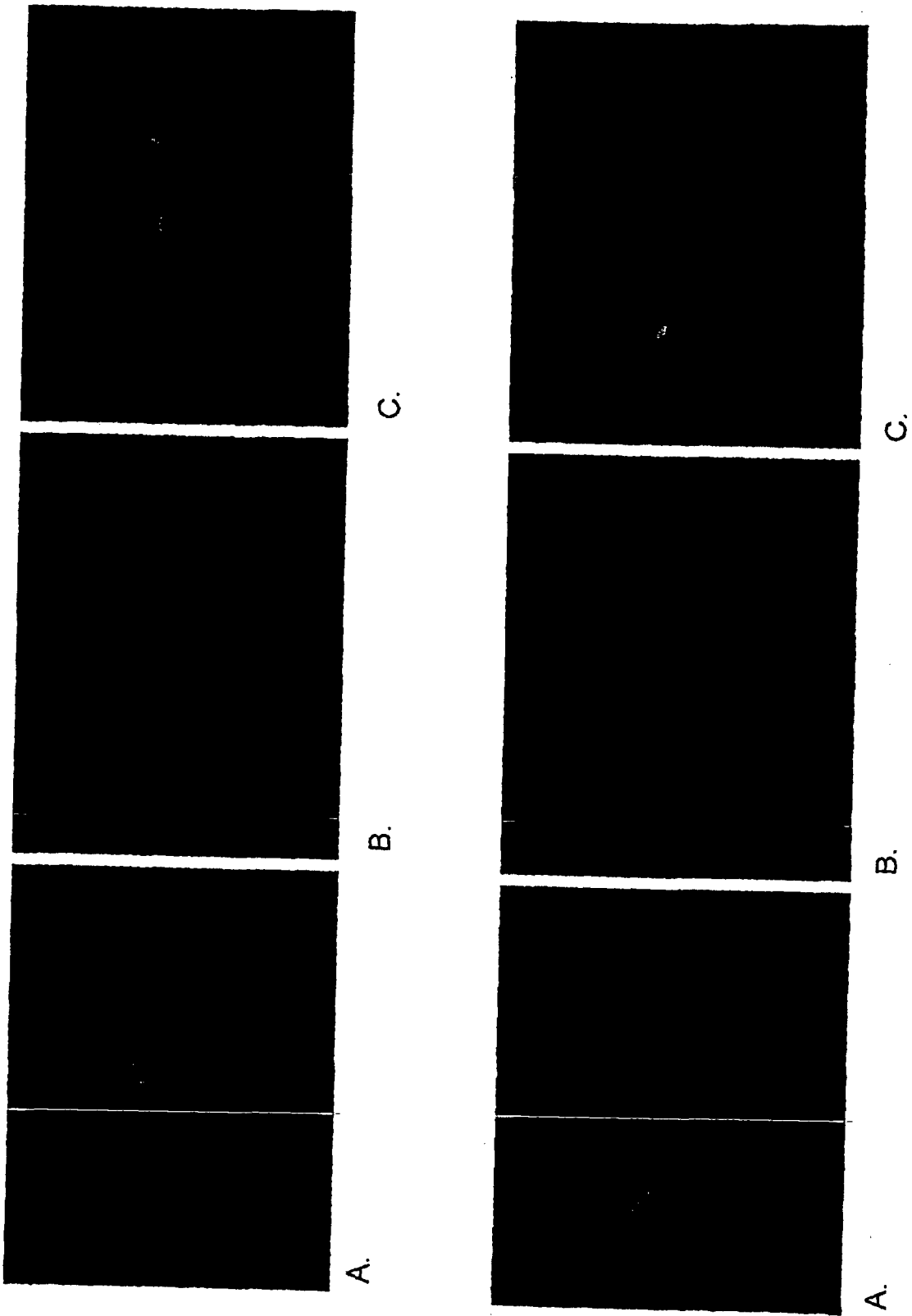


图 28



图 29

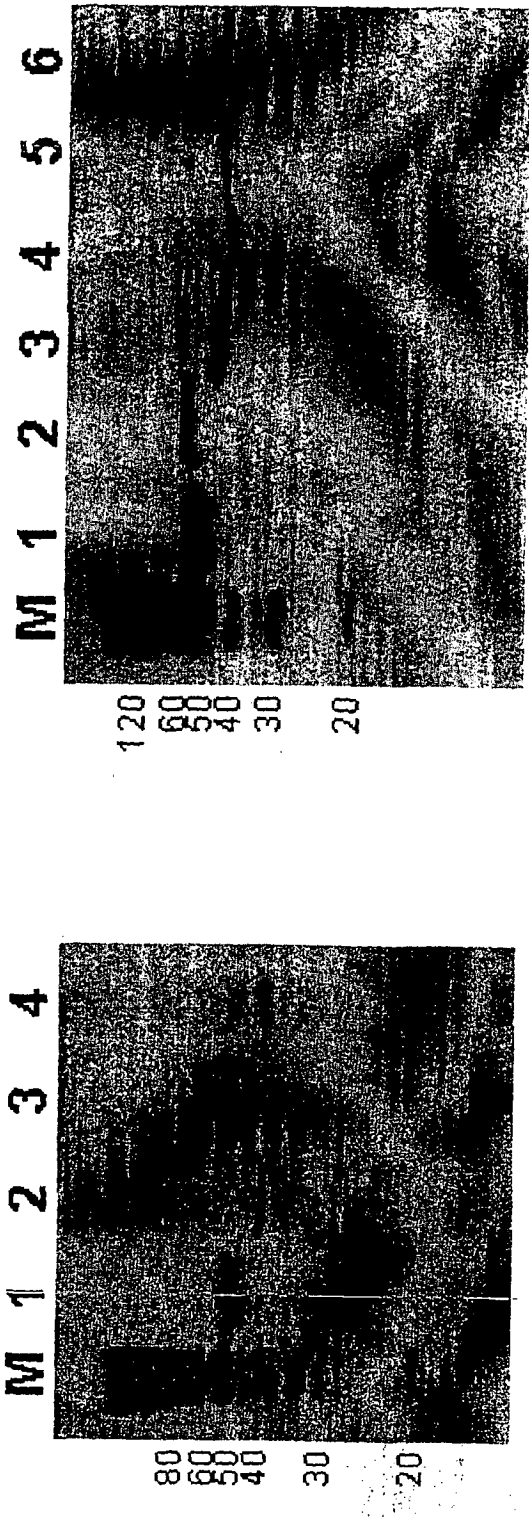


图 30

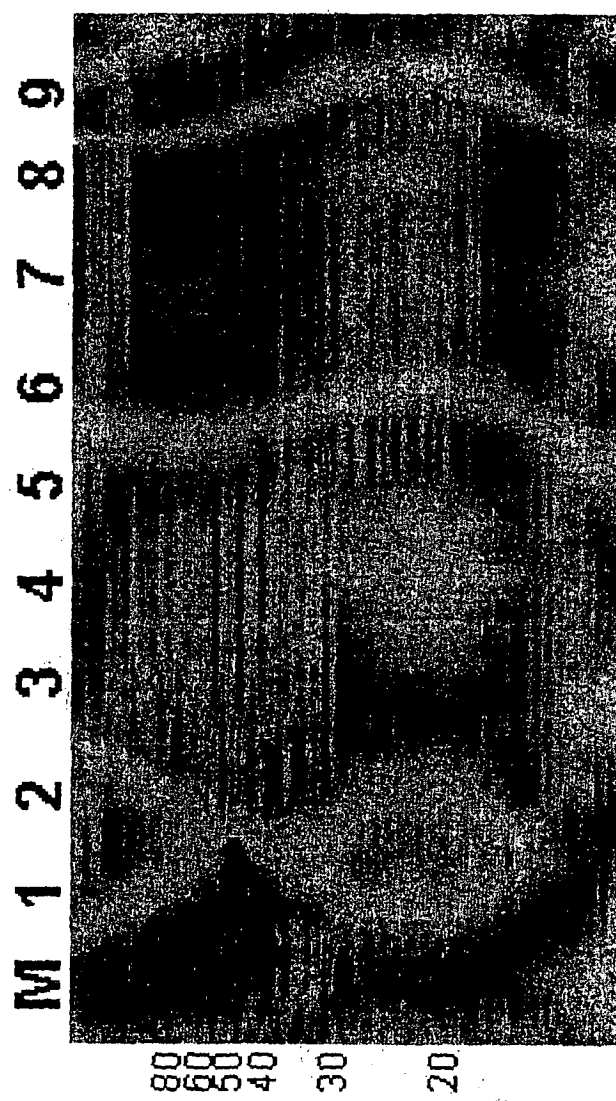


图 31



图 32

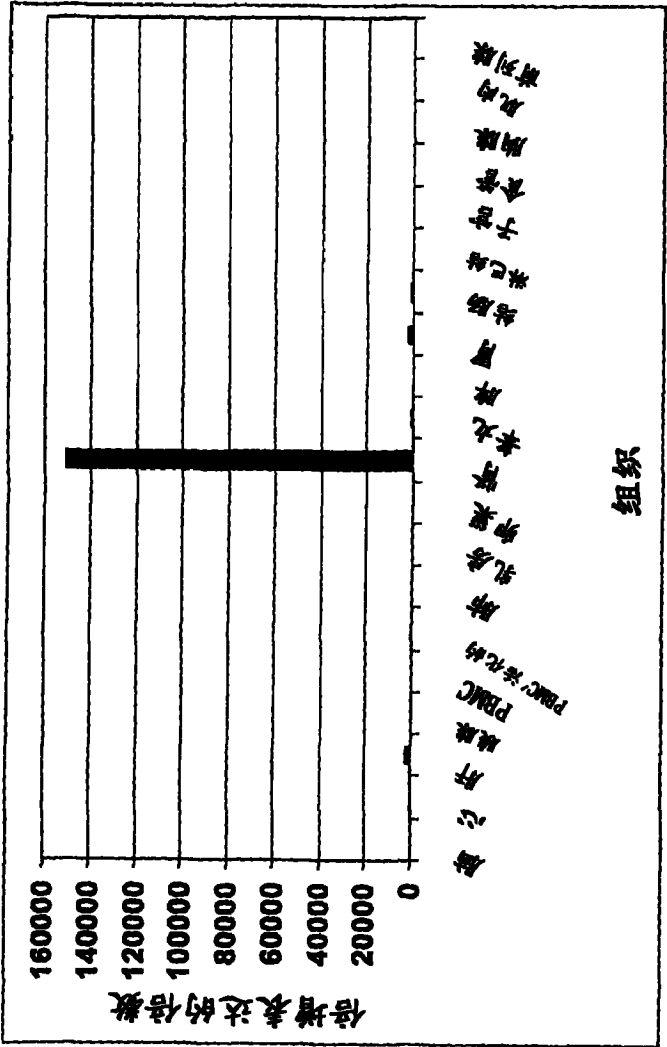


B.

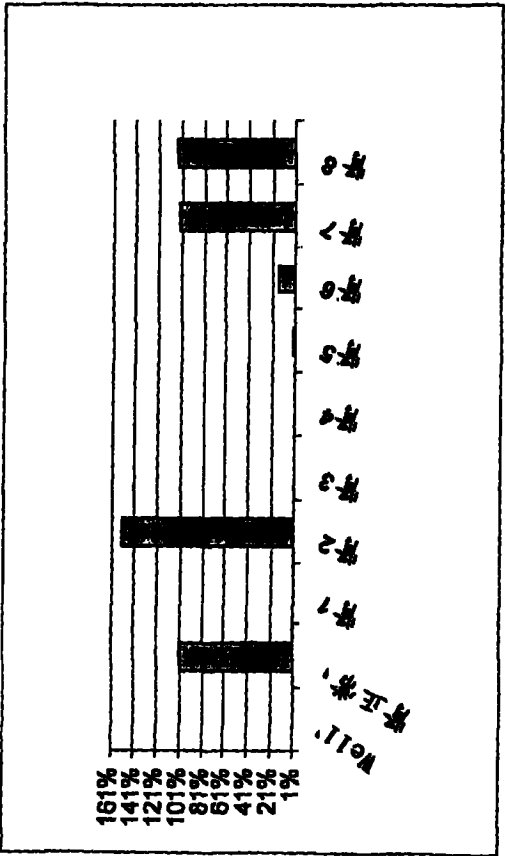


A.

图 33



A



B

图 34

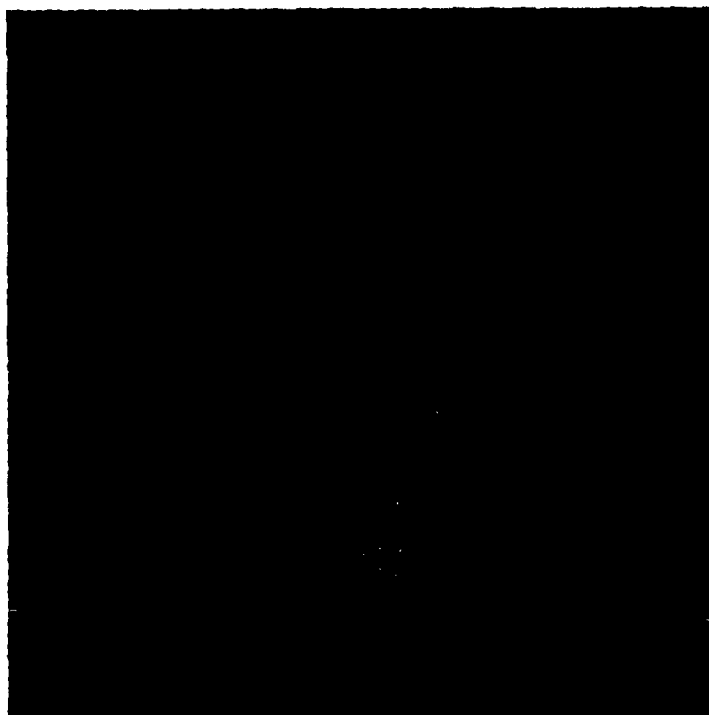


图 35

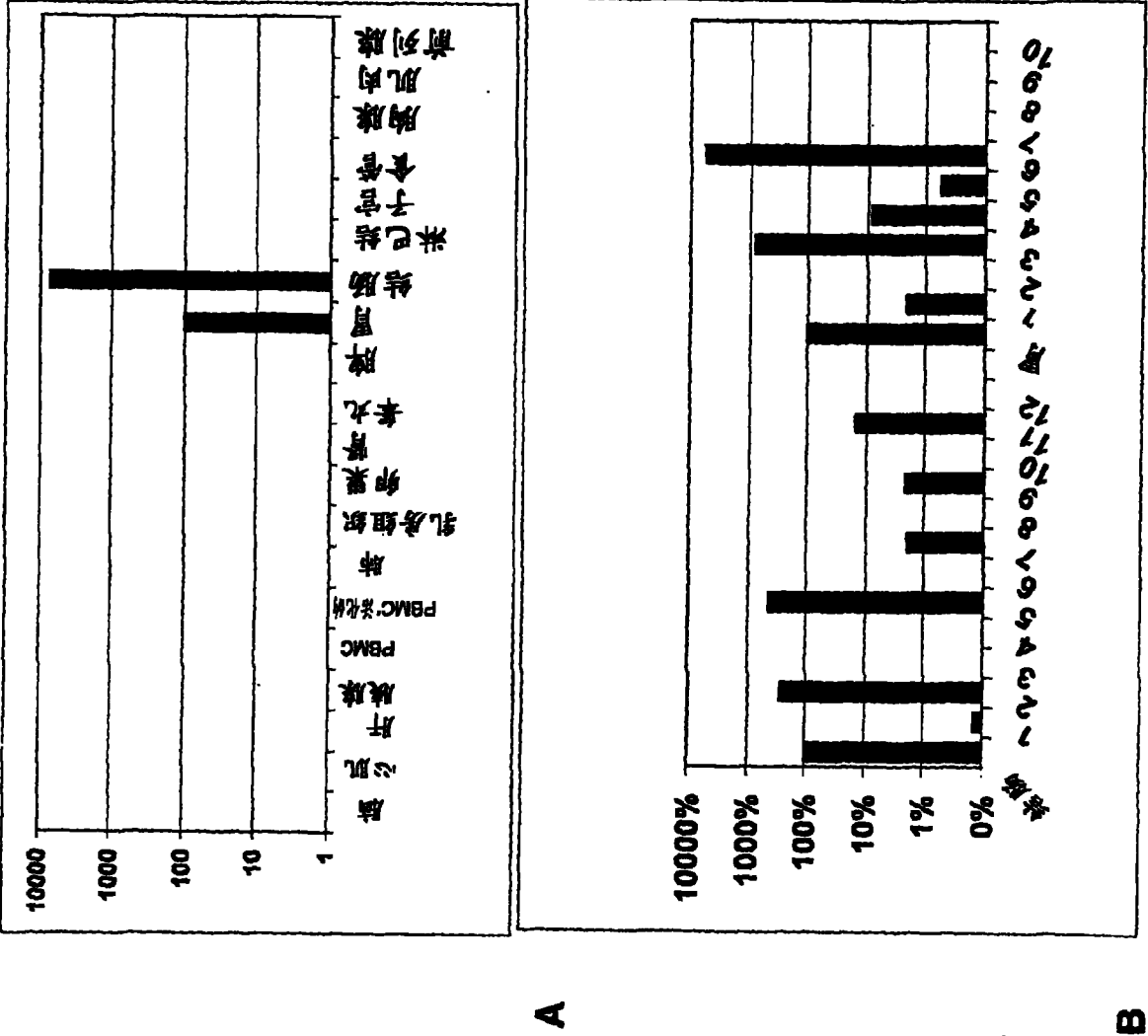


图 36

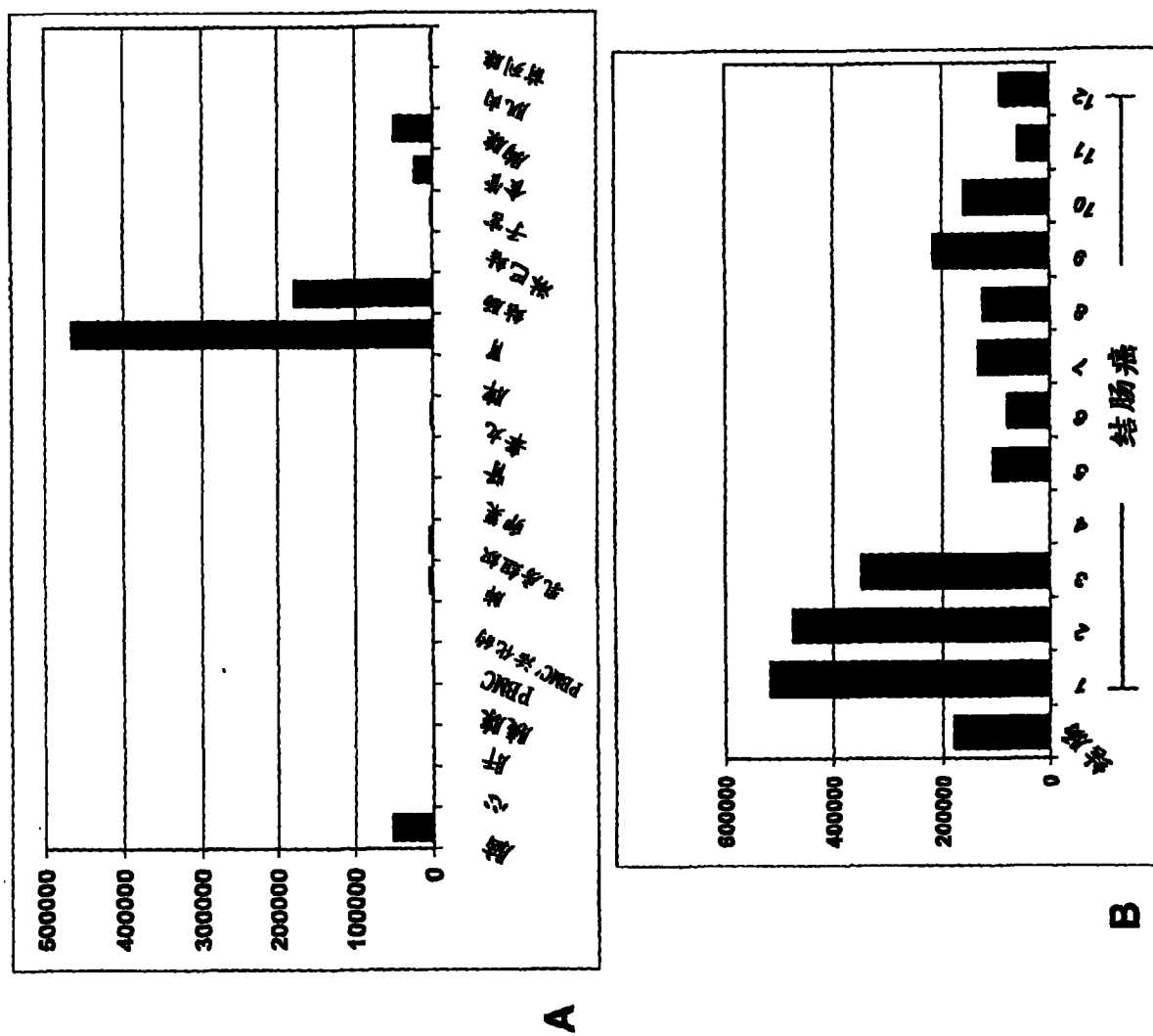


图 37

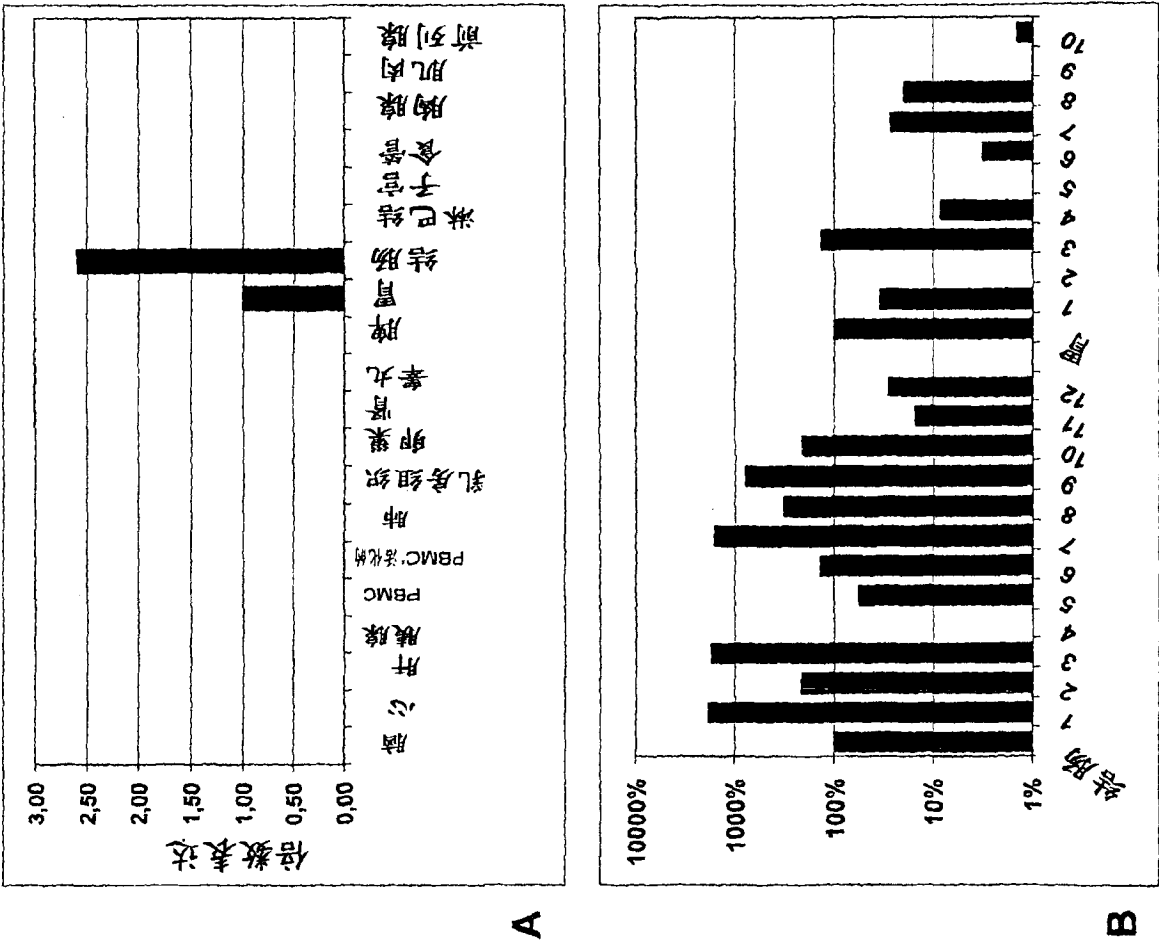


图 38

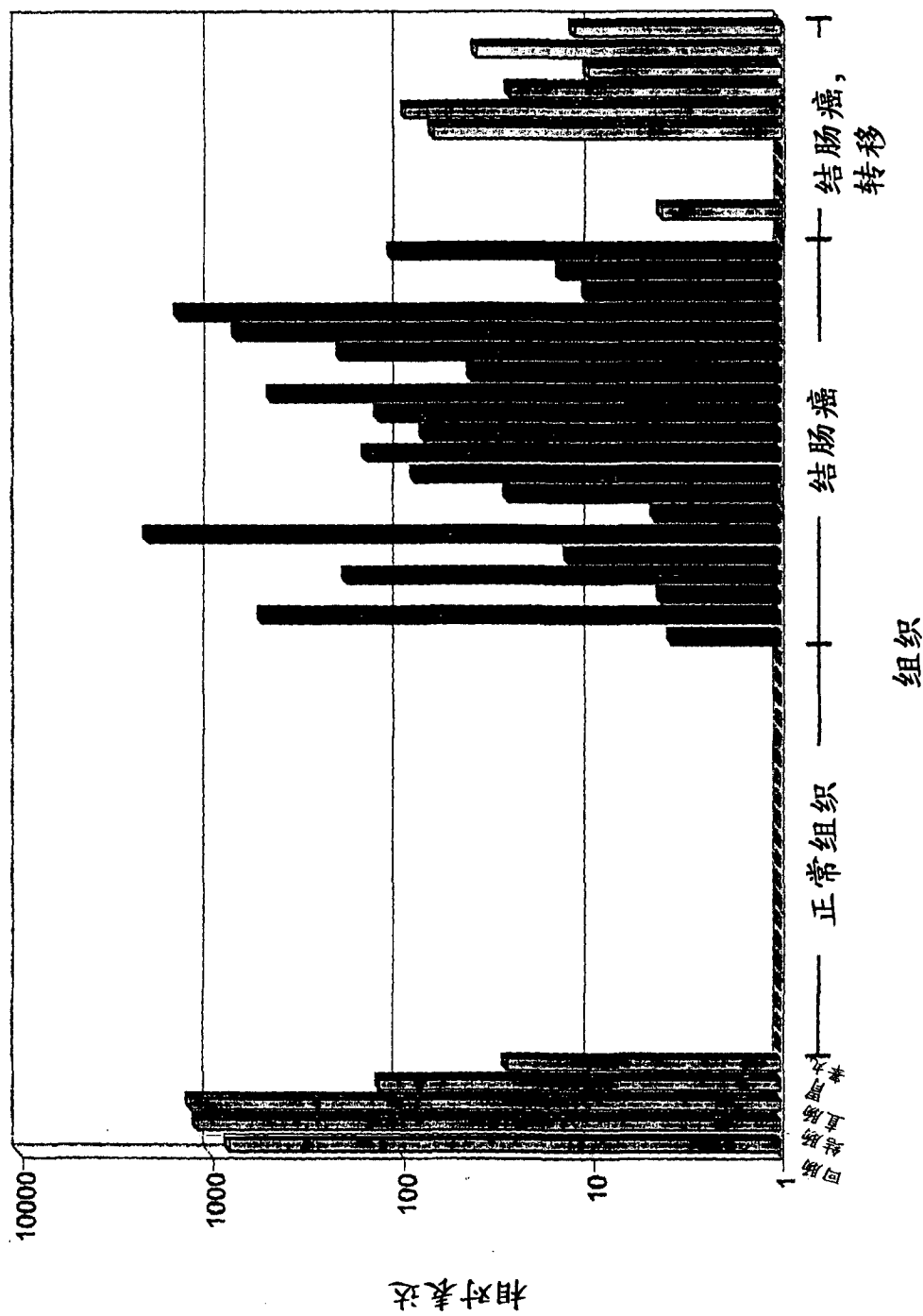


图 39

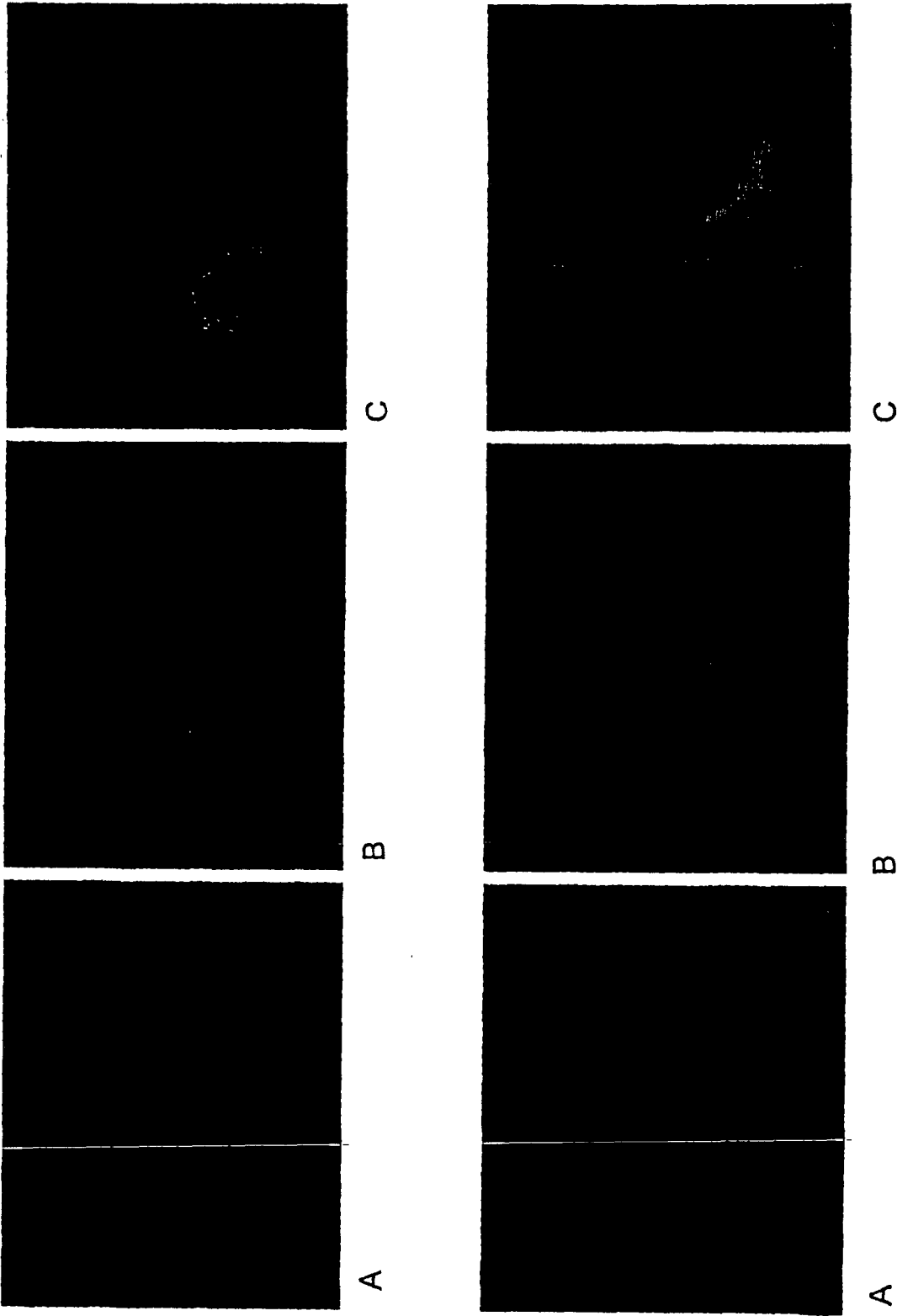


图 40

ID	序列
#1	caggccagagtcacagctgtcctggactctgctgtggggaagggtgatgcaggtgtgga gtcaaatgtgggtgcctcctgcagccgggtgccaggaggggtggagggccaccctgggc tttgtccgggagcctggtcttcccgctccttgggctgacaggtgctgctgcctctgagccc tcctgtctaagagctgtgtgctgggtaaggctggtggcccttgggctcctgtccagga tttgtgctctggagggtagggttgcctgggctggggaactggaggggaacgtggagctcct tctgcctccttctcctgccccatgacagcaggcagatcccaggagagaagagctcaggaga tgggaagaggatctgtccagggttagacctcaagggtgacttggagtctttacggcac ccatgctttctttaggagttttgtgtttgtgggtgtggggtcggggtcacctcctccc acatccctgcccagaggtgggcagagtggggcagtgcttgcctcccctgctcgtctc tgctgacctccggctcctgtgctgccccaggaccatgaatggcacctacaacacctgtg gctccagcgacctcacctggccccagcgatcaagctgggcttctacgctacttgggcg tcctgctggtgctaggcctgctgctcaacagcctggcgctctgggtgtctgctgccga tgcagcagtggaaggagaccgcctctacatgaccaacctggcggtggcgacctctgcc tgctgtgcaccttgccttctgctgctgcaactcctgagagacacctcagacacgcctgtg gccagctctcccagggtacatctacctgaccaacaggtacatgagcatcagcctggtcacgg ccatgcgctgggacgctatgtggcgtgcccgcacctgctgctgcccggggtgctgctg ccccaggcaggtgcccgtgtgctgctgctcctctgggtgctggtcatcggtcctgg tggtcgtgctgctcctggggttccagaggggctgctgcttccagagcaccgggcaca atttcaactccatggcgttccgctgctggtgatttctacctgcccctggcggtggtgct tctgctcctgaaggtggtgactgcctggcccagaggccaccacgcagctggggcagg cagagggcaccggcaaggctgcccgcctggtctgggccaacctcctggtgtctggtgct gcttctcctgcccctgcaagtggtgctgacagtgccctcgcagtggtggaacgcctgtg ccctcctggagacgatccgtcgcgcctgtacataaccagcaagctctcagatgccaa gctgctgggacgcctctgctactactacatggccaaggagtccaggaggcgtctgcac tgccgtggtcctccagtgctaaaggccacaaaaggcaggactctctgctgctgacctcg cctaagagcgctgctgtggcgctgtgggcaggctcctgggggtccgggaggtgctgcc tgccaggggagctggaaccagtagcaaggagccggggtcagccctgaactcactgtgt attctcttggagccttgggtggcgaggacggccaggtaacctgctctcttgggaagaga gaggggacagggaagggcaaggaggtgagggcagagcaaggccaatgtcagagacccc cgggatggggcctcacacttgcacccccagaaccagctcaactggccagagtgggttcc tgctggccagggtgcagccttgatgacacctgcccgtgcccctcggggtggaataaaac tccccaccagagtc
#2	ATGAAGACGTTGCTGTTGGACTTGGCTTTGTGGTCACTGCTCTCCAGCCCGGGTGGCTGTCCTTTAGTT CCCAGGTGAGTCAGAACTGCCACAATGGCAGCTATGAAATCAGCGTCTGATGATGGGCAACTCAGCCTT TGCAGAGCCCTGAAAACTTGAAGATGCGGTTGAATGAGGGGCTGGAAATAGTGAGAGGACGCTGCAAA AATGCTGGCCTAAATGTGACTGTGAACGCTACTTTTCATGTATTCCGATGGTCTGATTTCATAACTCAGGCG ACTGCCGGAGTAGCACCTGTGAAGGCTCGACCTACTCAGGAAAATTTCAAATGCACAAACGGATGGGCTG TGCTCTCATAGGGCCCTCATGTACATACCTCCACCTTCCAGATGTACCTTGACACAGAATTGAGCTACCCC ATGATCTCAGCTGGAAGTTTTGGATTGTATGTGACTATAAAGAAACCTTAACCAGGCTGATGTCTCCAG CTAGAAAGTTGATGTACTTCTTGGTTAACTTTTGGAAAACCAACGATCTGCCCTTCAAACTTATTCCTG GAGCACTTCGTATGTTTACAAGATGGTACAGAACTGAGGACTGTTTCTGGTACCTTAATGCTCTGGAG GCTAGCGTTTCTATTTCTCCACGAACTCGGCTTTAAGGTGGTGTAAAGACAAGATAAGGAGTTTCAGG ATATCTTAATGGACCACAACAGGAAAAGCAATGTGATTATTTATGTGTGGTGGTCCAGAGTTCTCTACAA GCTGAAGGGTGACCGAGCAGTGGCTGAAGACATTTGTCATTATTCTAGTGGATCTTTTCAATGACCAGTAC TTGGAGGACAATGTACAGCCCTGACTATATGAAAAATGTCTTGTCTGACGCTGTCTCCTGGGAATT CCCTTCTAAATAGCTCTTTCTCCAGGAATCTATCACCACAAAACGAGACTTTGCTCTTGCCTATTTGAA TGAATCCTGCTCTTTGGACATATGCTGAAGATATTTCTTGAAAAATGGAGAAAATATTAGCACCCCAAAA TTTGCTCATGCTTTTCAAGGAATCTCACTTTTGAAGGGTATGACGGTCCAGTGACCTTGGATGACTGGGGGG ATGTTGACAGTACCATGGTGTCTTCTGTATACCTCTGTGGACACCAAGAAATACAAGGTTCTTTGACCTA TGATACCCACGTAAATAAGACCTATCCTGTGGATATGAGCCCCACATTCACTTGAAGAATCTAACTT CCTAATGATATTACAGGCCGGGGCCCTCAGATCCTGATGATTGCACTTTCACCTCACTGGAGCTGTGG TGCTGCTCCTGCTCGTCTCTCTGATGCTCAGAAAATATAGAAAAGATTATGAACCTTCGTCAGAAAAA ATGGTCCCACATTTCTCTGAAAATATCTTTCTCTGGAGACCAATGAGACCAATCATGTTAGCCTCAAG ATCGATGATGACAAAAGACGAGATACATCCAGAGACTACGACAGTGCAAAATACGACAAAAGCGAGTGA TTCTCAAAGATCTCAAGCACAAATGATGGTAATTTCACTGAAAAACAGAGATAGAATTGAACAAGTTGCT TCAGATTGACTATTACAACCTGACCAAGTTCTACGGCACAGTGAAACTTGATACCATGATCTTCGGGGTG ATAGAATACTGTGAGAGAGGATCCCTCCGGGAAGTTTAAATGACACAATTTCTACCTGATGGCACAT TCATGGATTGGGAGTTAAGATCTCTGTCTTGTATGACATTTGCTAAGGGAATGTCATATCTGCACTCCAG TAAGACAGAAGTCCATGGTCTGAAATCTACCAACTGCGTAGTGGACAGTAGAATGGTGGTGAAGATC

	ACTGATTTTGGCTGCAATTCCATTTTACCTCCAAAAAGGACCTGTGGACAGCTCCAGAGCACCTCCGCC AAGCCAACATCTCTCAGAAAGGAGATGTGTACAGCTATGGGATCATCGCACAGGAGATCATTCTGCGGAA AGAAACCTTCTACACTTTGAGCTGTGCGGACCGGAATGAGAAGATTTTCAGAGTGGAAAATTCCATGGA ATGAAACCTTCCGCCAGATTATTCTTGGAAACAGCAGAGGAAAAAGAGCTAGAAGTGTACCTACTTG TAAAAACTGTTGGGAGGAAGATCCAGAAAAGAGACCAGATTTCAAAAAATTGAGACTACACTTGCCAA GATATTTGGACTTTTTCATGACCAAAAAAATGAAAGCTATATGGATACCTTGATCCGACGTCTACAGCTA TATTCTCGAAACCTGGAACATCTGGTAGAGGAAAGGACACAGCTGTACAAGGCAGAGAGGGACAGGGCTG ACAGACTTAACCTTATGTTGCTTCCAAGGCTAGTGGTAAAGTCTCTGAAGGAGAAAGGCTTTGTGGAGCC GGAATATATGAGGAAGTTACAATCTACTCAGTGACATGTAGGTTTCACTACTATCTGCAAAATACAGC ACCCCATGGAAGTGGTGGACATGCTTAATGACATCTATAAGAGTTTTGACCACATGTTGATCATCATG ATGCTACAAGGTGGAACCATCGGTGATGCGTACATGGTGGCTAGTGGTTGCCAAGAGAAATGGCAA TCGGCATGCCAATAGACATTGCCAAGATGGCTTGGAAATCCTCAGCTTCATGGGGACCTTTGAGCTGGAG CATCTTCTGGCCTCCCAATATGGATTGCGATTGGAGTTCAGTCTGGTCCCTGTGCTGCTGGAGTTGTGG GAATCAAGATGCCTCGTTATTGTCTATTTGGAGATACGGTCAACACAGCCTCTAGGATGGAATCCACTGG CCTCCCTTTGAGAATTCACGTGAGTGGCTCCACCATAGCCATCCTGAAGAGAATGAGTGCCAGTTCCCTT TATGAAGTGAGAGGAGAAACATACTTAAAGGGAAGAGGAAATGAGACTACCTACTGGCTGACTGGGATGA AGGACCAGAAATCAACCTGCCAACCCCTCTACTGTGGAGAATCAACAGCGTTTGAAGCAGAAATTTTC AGACATGATTGCCAACTCTTTACAGAAAAGACAGGCAGGAGGATAAGAAGCCAAAACCCAGACGGTA GCCAGCTATAAAAAAGGCACCTCTGGAATACCTGCAGCTGAATACCACAGACAAGGAGAGCACCTATTTTT AA
#3	ATGAAGACGTTGCTGTTGGACTTGGCTTTGTGGTCACTGCTCTTCCAGCCCGGGTGGCTGTCTTTAGTT CCCAGGTGAGTCAGAACTGCCACAATGGCAGCTATGAAATCAGCGTCTGATGATGGGCAACTCAGCCTT TGCAGAGCCCCTGAAAAACTTGAAGATGCGGTGAATGAGGGGCTGGAAATAGTGAGAGGACGTCTGCAA AATGCTGGCCTAAATGTGACTGTGAACGCTACTTTTCATGTATTGCGATGGTCTGATTCACTAAGTCAGGCG ACTGCCGGAGTAGCACCTGTGAAGGCCCTCGACCTACTCAGGAAAATTTCACTTGA
#4	ATGAAGACGTTGCTGTTGGACTTGGCTTTGTGGTCACTGCTCTTCCAGCCCGGGTGGCTGTCTTTAGTT CCCAGGTGAGTCAGAACTGCCACAATGGCAGCTATGAAATCAGCGTCTGATGATGGGCAACTCAGCCTT TGCAGAGCCCCTGAAAAACTTGAAGATGCGGTGAATGAGGGGCTGGAAATAGTGAGAGGACGTCTGCAA AATGCTGGCCTAAATGTGACTGTGAACGCTACTTTTCATGTATTGCGATGGTCTGATTCACTAAGTCAGGCG ACTGCCGGAGTAGCACCTGTGAAGGCCCTCGACCTACTCAGGAAAATTTCAATGCACAACGGATGGGCTG TGTCTCATAGGGCCCTCATGTACATACTCCACCTTCCAGATGTACCTTGACACAGAATTGAGCTACCCC ATGATCTCAGCTGGAAGTTTTGGATTGTGATGTGACTATAAAGAAACCTTAACCAGGCTGATGTCTCCAG CTAGAAAGTTGATGTACTTCTTGGTTAACTTTTGGAAAACCAACGATCTGCCCTTCAAACTTATTCTTG GAGCAGCTTCGTATGTTTACAAGAATGGTACAGAACTGAGGACTGTTTCTGGTACCTTAATGCTCTGGAG GCTAGCGTTTCTATTCTCCACGAACTCGGCTTTAAGGTGGTGTAAAGACAAGATAAGGAGTTTCAGG ATATCTTAATGACCACAACAGGAAAAGCAATGTGACCACTACTTGGAGGACAATGTCACAGCCCTGAC TATATGA
#5	ATGAAGACGTTGCTGTTGGACTTGGCTTTGTGGTCACTGCTCTTCCAGCCCGGGTGGCTGTCTTTAGTT CCCAGGTGAGTCAGAACTGCCACAATGGCAGCTATGAAATCAGCGTCTGATGATGGGCAACTCAGCCTT TGCAGAGCCCCTGAAAAACTTGAAGATGCGGTGAATGAGGGGCTGGAAATAGTGAGAGGACGTCTGCAA AATGCTGGCCTAAATGTGACTGTGAACGCTACTTTTCATGTATTGCGATGGTCTGATTCACTAAGTCAGGCG ACTGCCGGAGTAGCACCTGTGAAGGCCCTCGACCTACTCAGGAAAATTTCAATGCACAACGGATGGGCTG TGTCTCATAGGGCCCTCATGTACATACTCCACCTTCCAGATGTACCTTGACACAGAATTGAGCTACCCC ATGATCTCAGCTGGAAGTTTTGGATTGTGATGTGACTATAAAGAAACCTTAACCAGGCTGATGTCTCCAG CTAGAAAGTTGATGTACTTCTTGGTTAACTTTTGGAAAACCAACGATCTGCCCTTCAAACTTATTCTTG GAGCACTTCGTATGTTTACAAGAATGGTACAGAACTGAGGACTGTTTCTGGTACCTTAATGCTCTGGAG GCTAGCGTTTCTATTCTCCACGAACTCGGCTTTAAGGTGGTGTAAAGACAAGATAAGGAGTTTCAGG ATATCTTAATGACCACAACAGGAAAAGCAATGTGATTATATGTGTGGTGGTCCAGAGTTCTCTACAA GCTGAAGGGTGACCGAGCAGTGGCTGAAGACATGTCTATTATTCTAGTGGATCTTTTCAATGACCACTAC TTGGAGGACAATGTACAGCCCTGACTATATGAAAATGTCTTGTCTGACGCTGTCTCTGGGAATT CCCTTCTAAATAGCTCTTTCTCCAGGAATCTATCACCACAAAACGAGACTTTGCTCTTGCTATTGAA TGGAACTCTGCTCTTTGGACATATGCTGAAGATATTTCTTGAAAATGGAGAAAATATTACCACCCCAAA TTTGCTCATGCTTTTCAAGGAATCTCACTTTTGAAGGGTATGACGGTCCAGTGACCTTGGATGACTGGGGG ATGTTGACAGTACCATGGTGTCTGTATACCTCTGTGGACACCAAGAAATACAAGGTTCTTTTGACCTA TGATACCCACGTAAATAAGACCTATCCTGTGGATATGAGCCCCACATTCAGTTGGAAGAACTCTAACTT CCTAATGATATTACAGGCCGGGGCCCTCAGATCCTGATGATTGACGCTTCAACCTCAGTGAGCTGTGG TGCTGCTCCTGCTCGTCTCTCTGATGCTCAGAAAATATAGAAAAGATTATGAACCTCGTCAGAAAAA ATGGTCCCACATTCCTCTGAAAATATCTTCTCTGAGACCAATGAGACCAATCATGTTAGCCTCAAG ATCGATGATGACAAAAGACGAGATACAATCCAGAGACTACGACAGTGCAATACGACAAAAGCGAGTGA

	TTCTCAAAGATCTCAAGCACAATGATGGTAATTTCACTGAAAAACAGAAGATAGAATTGAACAAGATTGA CTATTACAACCTGACCAAGTTCTACGGCACAGTGAAACTTGATACCATGATCTTCGGGGTGATAGAATAC TGTGAGAGAGGATCCCTCCGGGAAGTTTAAATGACACAATTTCTACCCTGATGGCACATTCATGGATT GGGAGTTTAAGATCTCTGTCTTGTATGACATTGCTAAGGGAATGTCATATCTGCACTCCAGTAAGACAGA AGTCCATGGTCTCTGAAATCTACCAACTGCGTAGTGGACAGTAGAATGGTGGTGAAGATCACTGATTTT GGCTGCAATTCATTTTACCTCCAAAAAAGGACCTGTGGACAGCTCCAGAGCACCTCCGCCAAGCCAACA TCTCTCAGAAAGGAGATGTGTACAGCTATGGGATCATCGCACAGGAGATCATTCTGCGGAAAGAAACCTT CTACACTTTTGAGCTGTCTGGGACCGGAATGAGAAGATTTTCAGAGTGGAATTTCCAATGGAATGAACCC TTCCGCCCAGATTTATCTTGAAACAGCAGAGGAAAAAGAGCTAGAAGTGACCTACTTTGAAAAAACT GTTGGGAGGAAGATCCAGAAAAGAGACCAGATTTCAAAAAATTGAGACTACACTTGCCAGATATTTGG ACTTTTTCATGACCAAAAAATGAAAGCTATATGGATACCTTGATCCGACGTCTACAGCTATATTCTCGA AACCTGGAACATCTGGTAGAGGAAAGGACACAGCTGTACAAGGCAGAGAGGGACAGGGCTGACAGACTTA ACTTTATGTTGCTTCCAAGGCTAGTGGTAAAGTCTCTGAAGGAGAAAGGCTTTGTGGAGCCGAACTATA TGAGCAAGTTACAATCTACTTCAGTGACATTGTAGGTTTCACTACTATCTGCAAATACAGCACCCCCATG GAAGTGGTGGACATGCTTAATGACATCTATAAGAGTTTTCACCACATGTTGATCATCATGATGTCTACA AGGTGGAAACCATCGGTGATGCGTACATGGTGGCTAGTGGTTTGCTAAGAGAAATGGCAATCGGCATGC AATAGACATTGCCAAGATGGCCTTGGAATCCTCAGCTTCATGGGGACCTTTGAGCTGGAGCATCTTCCT GGCCTCCCAATATGGATTTCGCAATTGGAGTTCACTCTGGTCCCTGTGCTGCTGGAGTTGTGGGAATCAAGA TGCTCGTTATTGTCTATTGGAGATACGGTCAACACAGCCTCTAGGATGGAATCCACTGGCCTCCCTTT GAGAATTCACGTGAGTGGCTCCACCATAGCCATCCTGAAGAGAACTGAGTGCCAGTTTCCTTTATGAAGTG AGAGGAGAAACATACTTAAAGGGAAGAGGAAATGAGACTACCTACTGGCTGACTGGGATGAAGGACCAGA AATTCAACCTGCCAACCCCTCCTACTGTGGAGAATCAACAGCGTTTGCAAGCAGAATTTTCAGACATGAT TGCCAACTCTTTACAGAAAAGACAGGCAGCAGGATAAGAAGCCAAAAACCCAGACGGGTGCCAGCTAT AAAAAAGGCACTCTGGAATACTTGACAGCTGAATACCACAGACAAGGAGAGCACCTATTTTTTAA
#6	ggggacacttttgatggcaagtggaaaccactggcttggtg gattttgctagatttttctgatttttaaactcctgaaaaatatcccagataactgtcatgaagctggttaacta tcttctt gctggtgaccatcagcctttgtagttactctgctactgcc ttcctcatcaacaaagtggccttctctgttgacaagttggcacctttacctctggacaacattcttcccttta tggatcc attaaagcttcttctgaaaactctgggcatttctgttgag caccttgtggaggggctaaggaagtgtgtaaatgagctgggaccagaggcttctgaagctgtgaagaaactgc tggaggc gctatcacacttggtgtgacatcaagataaagagcggagg tggatggggatggaagatgatgctcctatcctccctgcctgaaacctgttctaccaattatagatcaaatgcc ctaaaatgtagtgaccctgtaaaaggacaaataaagcaatgaatactaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaa
#7	ATGGCCGTGACTGCCTGTGAGGGCTTGGGGTTCGTGGTTTCACTGATTGGGATTGCGGGCATCATTGCTG CCACCTGCATGGACAGTGGAGCACCCAAGACTTGTACAACAACCCCGTAACAGCTGTTTTCAACTACCA GGGGCTGTGGCCTCCTGTGTCCGAGAGAGCTCTGGCTTACCGAGTGCCGGGGCTACTTCACCTGCTG GGCTGCGACCATGCTGCAGGCAGTGGAGCCCTGATGATCGTAGGCATCGTCTGGGTGCCATTGGCC TCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAACAT GACACTGACCTCCGGGATCATGTTTATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTTTGCC AACATGCTGGTGAATACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGGGATGGTGCAGA CTGTTTCAAGCCAGGTACACATTGGTGCGGCTCTGTTCTGGGCTGGGTGCGTGGAGGCCCTCACATAAT TGGGGGTGTGATGATGTGCATCGCTGCCGGGGCTGGCACCAGAAGAAACCACTACAAAGCCGTTTCT TATCATGCCTCAGGCCACAGTGTTCCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCA ACACCAAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAA GCAGGACTATGTGTAA
#8	tgcgccaccatggccgtgactgcctgtcagggcttgggggttcgtggtttcactgattggg attgcgggcatcattgctgccacctgcattggaccagtggagcaccacaagacttgtacaac aaccctgaacagctgttttcaactaccaggggctgtggcgtcctgtgtccgagagagc
#9	MNGTYNTCGSSDLTWPPAIKLGFIYALGVLLVLGILLNSLALWVFCCRMQQWTETRIYMT NLAVADLCLLCTLPFVLHSLRDTSDTPLCQLSQGIYLTNRYSISLVTIAVDYVAVRH PLRARGLRSPRCAAACAVLWVLVIGSLVARWLLGIQEGGFCFRSTRHNFNSMRFFLLGF

	YLPLAVVVFCSLKVVTTALAQRPTDVGQAEATRKAARMVWANLLVFVVCFLPLHVGLTVR LAVGWNACALLETIRRALYITSKLSDANCCLDAICYYYMAKEFQEASALAVAPRAKAHKS QDSL CVTLA
#10	MTAGRSQERRAQEMGRGSVQGLDLKGDLEFFTAPMLSLRSFVTVGVGSGLTSSHIPAQKWAEWGQCLAPPARS LLTSGSLCCPRTMNGTYNTCGSSDLTWPPAIKLGFYAYLGVLLVLGL LLNSLALWVFCCRMQQWTETRIYMTNLAVADLCLLCTLPFVLSLRDTSOTFLCQLSQGI YLTNRYMSISLVTIAIVDRYVAVRHPLRARGLRSPQAAAVCAVLWVLVIGSLVARWLLG IQEGGF CFRSTRHNFNMSMAFPLLG FYLEPLAVVVFCSLKVVTTALAQRPTDVGQAEATRKA ARMVWANLLVFVVCFLPLHVGLTVRLAVGWNACALLETIRRALYITSKLSDANCCLDAIC YYMAKEFQEASALAVAPSAAHKSQDSL CVTLA
#11	MKTLLLDLALWSLLFQPGWLSFSSQVSNCHNGSYEISVLMGNSAFAPLKNLEDAVNEGLEIVRGRLO NAGLNVTVNATFMYSDGLIHNSGDCRSSTCEGLDLLRKISNAQRMGCVLIGPSCITYSTFQMYLDTELSYP MISAGSFGSLCDYKETLTRLMSPARKLMYFLVNFWKTNLDPFKTYSWSTSYVYKNGTETEDCFWYLNAL ASVSYFSHELGFKVVLQDKEFQDILMDHNRKSNVIMCGGPEFLYKLGDRVAEDIVILVDLFNDQY LEDNVTAPDYMKNVVLTLSPGNSLLNSSFSRNLSPTKRDFALAYLNGILLFGHMLKIFLENGENITPK FAHAFRNLTFEGYDGPVTLDDWGDVDSTMVLLYTSVDTKKYKVLTYDTHVNKTYPVDMSPTFTWKNSKL PNDITGRGPQILMIAVFTLTGAVVLLLVALLMLRKYRKDYELRQKKWSHIPPENIFPLETNETNHVSLK IDDDKRRDTIQR LRQCKYDKKRVILKDLKHN DGNFTEKQKIELNKLLQIDYNNLT KFYGT VKLDTMIFGV IEY CERGSLREV LNDTISYPDGTFMDWEFKISVLYDIAGMSYLHSSKTEVHGRLKSTNCVVD SRMVVKI TDFGCNSILPPKKDLWTAPEHLRQANISQKGDVSYGIIAQEII LRKETFYTLSCDRNEKIFRVENSNG MKPFRPDLFLETAEEKELEVYLLVKNCEEDPEKRPDFKKIETTLAKIFGLFHDQKNESYMDTLIRRLQY YSRNL EHLVEERTQLYKAERDRADRLNFMLLPRLVVKSLKEKGFVEPELYEVTIYFSDIVGFTTICKYS TPMEVVDMLNDIYKSFHDHVDVYKVTIGDAYMVASGLPKRNGNRHAIDIAKMALEILSFMGTFELE HLPGLPIWIRIGVHSGPCAAGVVGIMPRYCLFGDTVNTASRMESTGLPLRIHVSGSTIAILKRTECQFL YEV RGETY LKGRGNETTYWLTGMKDQKFNLPTPTVENQORLQAEFSDMIANSLOKQQAAGIRSQKPRRV ASYKKGTLEYLQLNTTDKESTYF*
#12	MKTLLLDLALWSLLFQPGWLSFSSQVSNCHNGSYEISVLMGNSAFAPLKNLEDAVNEGLEIVRGRLO NAGLNVTVNATFMYSDGLIHNSGDCRSSTCEGLDLLRKISP*
#13	MKTLLLDLALWSLLFQPGWLSFSSQVSNCHNGSYEISVLMGNSAFAPLKNLEDAVNEGLEIVRGRLO NAGLNVTVNATFMYSDGLIHNSGDCRSSTCEGLDLLRKISNAQRMGCVLIGPSCITYSTFQMYLDTELSYP MISAGSFGSLCDYKETLTRLMSPARKLMYFLVNFWKTNLDPFKTYSWSTSYVYKNGTETEDCFWYLNAL ASVSYFSHELGFKVVLQDKEFQDILMDHNRKSNVTSTWRTMSQPLTI*
#14	MKTLLLDLALWSLLFQPGWLSFSSQVSNCHNGSYEISVLMGNSAFAPLKNLEDAVNEGLEIVRGRLO NAGLNVTVNATFMYSDGLIHNSGDCRSSTCEGLDLLRKISNAQRMGCVLIGPSCITYSTFQMYLDTELSYP MISAGSFGSLCDYKETLTRLMSPARKLMYFLVNFWKTNLDPFKTYSWSTSYVYKNGTETEDCFWYLNAL ASVSYFSHELGFKVVLQDKEFQDILMDHNRKSNVIMCGGPEFLYKLGDRVAEDIVILVDLFNDQY LEDNVTAPDYMKNVVLTLSPGNSLLNSSFSRNLSPTKRDFALAYLNGILLFGHMLKIFLENGENITPK FAHAFRNLTFEGYDGPVTLDDWGDVDSTMVLLYTSVDTKKYKVLTYDTHVNKTYPVDMSPTFTWKNSKL PNDITGRGPQILMIAVFTLTGAVVLLLVALLMLRKYRKDYELRQKKWSHIPPENIFPLETNETNHVSLK IDDDKRRDTIQR LRQCKYDKKRVILKDLKHN DGNFTEKQKIELNKIDYNNLT KFYGT VKLDTMIFGVIEY CERGSLREV LNDTISYPDGTFMDWEFKISVLYDIAGMSYLHSSKTEVHGRLKSTNCVVD SRMVVKITDF GCNSILPPKKDLWTAPEHLRQANISQKGDVSYGIIAQEII LRKETFYTLSCDRNEKIFRVENSNGMKP FRPDLFLETAEEKELEVYLLVKNCEEDPEKRPDFKKIETTLAKIFGLFHDQKNESYMDTLIRRLQYLSR NLEHLVEERTQLYKAERDRADRLNFMLLPRLVVKSLKEKGFVEPELYEVTIYFSDIVGFTTICKYSTPM EVVDMLNDIYKSFHDHVDVYKVTIGDAYMVASGLPKRNGNRHAIDIAKMALEILSFMGTFELEHLP GLPIWIRIGVHSGPCAAGVVGIMPRYCLFGDTVNTASRMESTGLPLRIHVSGSTIAILKRTECQFLYEV RGETY LKGRGNETTYWLTGMKDQKFNLPTPTVENQORLQAEFSDMIANSLOKQQAAGIRSQKPRRVASY KKGTLEYLQLNTTDKESTYF*
#15	MKLVTIFLLVTISLCSYSATAKLINCKPLPVDKLA PLPLDNILPFMDPLK LLLKTLGISVEHLVEGLRKCVNELGPEASEAVKLLLEALSHLV
#16	MAVTACQLGFGVVS LIGIAGIIAATCMDQWSTQDLYNNPVTA VFN YQGLWRSCVRESSGFTECRGYFTLL

	GLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGLCAIAGVSVFA NMLVTNFWMTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIACRGLAPEETNYKAVS YHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDYV*
#17	DQWSTQDLYN
#18	NNPVTAVFNYQ
#19	MAVTACQGLGFVVSLLIGIAGTAAATCMDQWSTQDLYNPNPVTAVFNYQ
#20	AGGTACATGAGCATCAGCCTG
#21	GCAGCAGTTGGCATCTGAGAG
#22	GCAATAGACATTGCCAAGATG
#23	AACGCTGTTGATTCTCCACAG
#24	GGATCCTCCTTTAGTTCACAGGTGAGTCAGAAC
#25	TGCTCTGGAGGCTAGCGTTTC
#26	ACCAATCATGTTAGCCTCAAG
#27	AGCTATGGGATCATCGCACAG
#28	CCTTTGAGCTGGAGCATCTTC
#29	CTTTCTAGCTGGAGACATCAG
#30	CACCATGGTACTGTCAACATC
#31	ATGTCATACAAGACAGAGATC
#32	TCTGCCTTGTACAGCTGTGTC
#33	TCTGTGGTATTGAGCTGCAAG
#34	TACTCAGGAAAATTCACCTTG
#35	GACCACAACAGGAAAAGCAATGTGACC

#36	GATAGAATTGAACAAGATTGAC
#37	CAGCCTTTGTAGTTACTCTGC
#38	TGTCACACCAAGTGTGATAGC
#39	GGTTCGTGGTTTCTACTGATTGGGATTGC
#40	CGGCTTTGTAGTTGGTTTCTTCTGGTG
#41	ctattgaagccacctgctcaggacaatgaaattcttcagttacattctggttatcgccg atctctctctggtgttttactgtgtgtgtttactacctctgccatcgctccacac caaggaagcagaatgtgcctacacactctttgtggtcgccacattttggctcacagaagc attgcctctgtcggtaacagctttgtacctagttaaattgttacccatgtttgggatcat gccttctaagaaggtggcatctgcttatttcaaggattttcacttactgctaattggagt tatctgttttagcaacatccatagaaaaatggaatttgacaagagaattgctctgaaaat ggtgatgatggttgggtgaaatcctgcatggctgacgctgggttcatgagcagcactgc ctttttgtctatgtggctcagcaacacctcgacggtgcatggtgatgccattgcgga ggctgtagtgcagcagatcatcaatgcagaagcagaggtcgaggccactcagatgactta cttcaacggatcaaccaaccacggactagaaattgatgaaagtgttaattggacatgaaat aaatgagaggaagagaaaaacaaaccaggtccaggatataatgataggggaaaaat ttcaagcaaggtggagttggaaaaagaactcaggcatgagaaccaaatatcgaaacaaaga gggccagtgacacgttaaacttacgtgtttgtgcatgacctactcttctaccattggtg actgacaacaatcactggtaacctccaccaacttgatctttgcagagtatttcaatacacg ctatcctgactgtcgttgcctcaactttggatcatggtttacgttttcttcccagctgc ccttatcattctactcttctcctggatctggcttcagtggttttcttaggattcaattt taaggagatgttcaaatgtggcaaaacaaaacagtcacacaaaagcttgtgctgaggt gattaagcaagaataccaaaagcttgggccaataaggtatcaagaaattgtgacctggt cctcttcattataatggctctgctatggttttagtcgagaccccgatttgttctggttg gtctgcactttttcagagtaacctgggttttgctacagattcaactgttgcttacttat agggctgctattctttcttctccagctaagacactgactaaaactacacctacaggaga aattgttgcttttgattactctccactgattacttggaaagaattccagtcattcatgcc ctgggatatagccattcttgttgggtggagggttggccctggcagatgggtgtgaggagtc tggattatctaagtggataggaataaattatctcctctgggttcattaccagcatggct aataattctgatattcttcttgatggtgacatctttaaactgaggtagccagcaatccagc taccattacactcttctcccaatattatctccattggccgaagccattcatgtgaaccc tctttatattctgataccttctactctgtgtacttcatattgcattcctcctaccagtagc aaatccaccaatgctattgtcttttcatatgggtcatctgaaagtcattgacatgggttaa agctggacttgggtgcaacattgttgggtgtgctgtggttatgcttggcatatgtacttg gattgtacctatgttgacctctacacttaccttctgtgggtcctgctatgagtaatga gacctgccataataagcacaataattctgactatcttgcggttaatttctggaagacatt aatgattgactgtaaaaatgtggctctaaataactaatgacacacatttaaatcagttatg gtgtagctgctgcaattccgtgaataaccgaaacctgctggtataactcagagtcata tttgttattgcagtgcactaaagagcatctatgtgccttcatcaagaagcccatgtttt gagattttgtcatgaaccatctgcaacttgcttcatcataagaataatttataacttga ccttcaaagagattagagcatttgtttcatcttacagttggagttcaatgtaacatttta aatgcaattttattttcagaaatttcccatgaaactaaaaatagaaaaataagatataga agttaatccggtacttggataaatcatttctgcatgttgttccagagaatttgcagaga aatcaaagccatgggtcatctgggtgatgaagagaaaaaggttaattctaaatgatgtgcat ttcctcatttaaaaaatccaattggattattcttaatatatacatgtaatatgaaaattg agattgaagcactaattccaaaattatggctgaatataactaaataacagaaaagttacag ataagaattttatttctactgaactctatagttagtgtaatataattcatatttttatgat attggcacactgagaaattcattttgtagagctatggataaggcttgcctatgatttgac tattagtacagtatagttagaaggaaagctgaacactataaaactattaacatattttc gtatatagtagtaacaactttgcttaagtgtttatcttagttcagaaatacataatgtcata tgttaaaaaataaagagatgtagaatactaaatgaattatcactgtgtatacagacagaaa aatcacataactctggtgtgttaacattgcaatgaaaaaatgaaaaaagaaggaaaaaa gaataagaatgaaaactgctgacgtattacaaaacagaaaaataaatgatttaaatcaa

	atcaaaaaaagaaaaaactaaacattttaacaaaaatgggataagaatagtcttctagaag tgaggatgcgtaaaaagaatgagtttccaattaccctgatgtgacaattacacattgtaga caggtagcaaaatatcacatacaccacaaaatatgtacaaatattatatatcaataaat aaatttttaagagtaagtgtctattggcattccaaaattcagctaaaggaaaaatgatca aaaacaaagtaagggtgcacagtttagcaaaagatgcagatggttatatcacagcaattctca tgctaaaaatacaacaaaagacaaaagcaaaaaataaacctttgcttttttttttttttt tttttttttgagacggaggtctcgctctgtcgccaggctggagtgagtgggcggtct cggctcactgcaagctccgctccagggttcacgccaattctcctgctcagccaaacctt tgctatttttaactcttggtggcactttccagctgttactgacctgtcattttttgttc aaataagattatttacaacattattcttgaaactaaatatagtaaagggttttataaa taatatttaacatacgaattattaattggccatgttcatttttatctatgtttatataat gggccaatgcaaaaaatcattttttcaagaaaaatttgccatgtaaagcttaaatat aatattgtgctttgtataactcttctatgtttattctattcatttgttctcttccctac catattttacacatgtatttataatctgtagtattttattacatttctgctttttctagt cattcaatttatcactgtgaattgcatcagatcatggatgcattttttattatgaaaaa taaatgacttttcaaattaaaaaaaaaaaaaa
#42	caggacaatgaaattcttcagttacattctgggtttatcgccgatttctcttctggttttcaactgtgttggt ttactacctctgcccatogtctccacaccaagggaagcagaatgtgcotacacactctttgtggtcgccacat tttggtcacagaagcattgcctctgtcggttaacagctttgtctacctagttaatgttacctatgtttggat catgcttcttaagaaggtggcatctgcttatttcaaggattttcaactactgctaatggagttatctgttta gcaacatccatagaaaaatggaatttgacacaagagaattgtctgaaaatggtgatgatggttggtgtaaatc ctgcatggtgacgtgggttcatgagcagcactgctttttgtctatgtggctcagcaacacctcgacggc tgccatggtgatgcccattgaggaggtgtagtgcagcagatcatcaatgcagaagcagaggtcgaggccact cagatgacttacttcaacggatcaaccaaccaggactagaaattgatgaaagtgttaattggacatgaaataa atgagaggaagagaaaaacaaaaccagttccaggatacaataatgatcacagggaatttcaagcaaggtgga gttggaagagactgtttaactactgaaatgaagctattctctgactaaacataactgaaaaaccattcatta aatg
#43	gccactcagatgacttaactcaacggatcaaccaaccaaggactagaaattgatgaaagtgttaatggacatg aaataaatgagaggaaaagagaaaaacaaaccagttccaggatataatgatcacagggaatttcaagcaa ggtggagttggaaggaacttgcaacttgcaagatggctcccatctccctctgtccattctgtatcg cagctagctgctcaaggaaaggagaaagtgggaaggcatatgtacttagaaattattctattacttctctggat ttaagagtattcagattttctatttcaacatcaacaaattgcattttataaaagaaatttatgtgtccatgt caaattttagtagtgtgtgtgtttataatattttcttatctacttaatttctatagttattatagtata tgtctttattttcaacatttttctgtgttttaagattattttaagattatttttaataatctttatttc atttaataaaaatattttattttaagtct
#44	cacggactagaaattgatgaaagtgttaatggacatgaaataaatgagaggaaaagagaaaacaaaaccagttc caggatacaataatgatcacagggaatttcaagcaagggtggagttggaagaaactcaggcatgagaacca atatcgacaaaagaggccacgtgacacgtaaacttacgtgtttgtgcattgctactctctaccattggt ggactgacaacaatcactggtacctccaaccaacttgatottgcagagtatttcaatacattccatccacaca gaagaggagatcgtacaaggcatgtacaccaggaggcagaaatttgaggcatatcttgyaactctgtctacca catctgaacatcacacagtttccactottgttgcttcaatcctgagaatgcattcaggagccattctgttt tatgtcaattactaattagatcatgtcacgttactaacttactacgttccaattagtcttattgcatttgta ataaaatccgcatactttcggactggctacaagggttatcatgat
#45	MKFFSYILVYRRFLFVFTVLVLLPLPIVLHTKEAECAYTLFVV ATFWLTEALPLSVTALLPFLMLPMFGIMPSKKVASAYFKDFHLLLVICLATSIEKW NLHKRIALKMVMVGVNPAWLTGFMSSSTAFLSMWLSNTSTAAMVMPIAEAVVQOIIN AEAEVEATQMTYFNGSTNHGLEIDESVNGHEINERKEKTPVPGYNNDTGKISSKVEL EKNSGMRTRYRTKKGHVTRKLTCLCIAYSSTIGGLTTITGTSTNLIFAEYFNTRYPCD RCLNFGSWFTFSFPAALIILLWSIWLQWLFLGFNFKEMFKCGKTKTVQKACAEVIK QEYQKLGPIRYQEIIVTLVLFIMALLWFSRDPGFVPGWSALFSEYPGFATDSTVALLI GLLFLIPAKTLTKTPTGEIVAFDYSPLITWKEFQSFMPWDIAILVGGGFALADGCE ESGLSKWIGNKLSPLGSLPAWLIILISSLMVTSLTEVASNPATITLFLPILSPLAEAI HVNPLYILIPSTLCTSFALLPVANPPNAIVFSYGHKVIDMVKAGLVNIVGVAVVM LGICTWIVPMFDLYTYPWAPAMSNETME"

#46	RTMKFFSYILVYRRFLFVVFVTLVLLPLPIVLHTKEACAYTLFVVATFWLTEALPLSVTALLPSLMLPMFGL MPSKKVASAYFKDFHLLIGVICLATSIKWNLHKRIALKMVMVGVNPAWLTLGFMSSTAFLSMWLSNTSTA AMVMPIAEAVVQIINAEAEVEATQMTYFNGSTNHGLEIDESVNGHEINERKEKTKFPVPGYNNDTGKISSKVE LEKTV*
#47	ATQMTYFNGSTNHGLEIDESVNGHEINERKEKTKFPVPGYNNDTGKISSKVELEKHWKLAVQDGSPPSVHSVS QLAAQGEKVEGICT*
#48	HGLEIDESVNGHEINERKEKTKFPVPGYNNDTGKISSKVELEKNSGMRTKYRTKKGHVTRKLTCLCIAYSSSTIG GLTTITGTSTNLIFAEYFNTFHPHRRGDRTRHVHQAETI*
#49	CCAGCTTTAACCATGTCAATG
#50	CAGATGGTTGTGAGGAGTCTG
#51	TGCTAATGCTTTTGGTACAAATGGATGTGGAATATAATTGAATATTTCTTGTTTAAGGGAGCATGAAGAGG TGTTGAGGTTATGTCAAGCATCTGGCACAGCTGAAGGCAGATGGAATATTTACAGTACGCAATTTGAGACT AAGATATTGTTATCATCTCTCTATTGAAGACAAGAGCAATAGTAAACACATCAGGTGAGGGGTTAAAGACC TGTGATAAACCACCTTCGATAAGTTGGAACGCTGTGTCTATATTTTCATATCTGTATATATATATGTTAAAG AAAGACACCTTCGTAACCCGCATTTTCCAAAGAGAGGAATCACAGGGAGATGTACAGCAATGGGGCCATTTAA GAGTCTGTGTTTCTGTTGATCTTACCTTCTAGAGGGGCCCTGAGTAATTCATTCAGCTGAACAAC AATGGCTATGAAGGCATTTGTCGTTGCAATCGACCCCAATGTGCCAGAGATGAAACACTCATTCAACAAATAA AGGACATGGTGACCCAGGCATCTCTGTATCTGTTGAAGCTACAGGAAAGCGATTTTATTTCAAAAATGTTGC CATTTTGATTCTCTGAAACATGGAAGACAAAGGCTGACTATGTGAGACCAAACTTGAGACCTACAAAATGCT GATGTTCTGGTTGCTGAGTCTACTCCTCCAGGTAATGATGAACCTTACTGAGCAGATGGGCAACTGTGGAG AGAAGGGTGAAAGGATCCACCTCACTCCTGATTTGATGAGGAAAAAGTTAGCTGAATATGGACCACAAGG TAAGGCATTTGCTCCATGAGTGGGCTCATCTACGATGGGGGATATTTGACGAGTACAATAATGATGAGAAATTC TACTTATCCAATGGAAGAATACAAGCAGTAAGATGTTTACGAGGATTTACTGGTACAAATGTAGTAAAGAAGT GTCAGGGAGGCAGCTGTTACACCAAAAGATGCACATTCAATAAAGTTACAGGACTCTATGAAAAGGATGTGA GTTTGTCTTCCAATCCCGCCAGACGGAGAAGGCTTCTATAATGTTTGACACACATGTTGATTCTATAGTTGAA TTCTGTACAGAACAAAACCAACAAAGAAGCTCCAAACAGCAAAATCAAAAATGCAATCTCCGAGCACAT GGGAAGTGATCCGTGATTCTGAGGACTTTAAGAAAACCACTCCTATGACAAACAGCCACCAAAATCCCACCTT CTCATTGCTGCAGATTGGACAAAGAATTGCTGTTTGTAGTCTTGACAAATCTGGAAGCATGGCGACTGGTAAC CGCTCAATCGACTGAATCAAGCAGGCCAGCTTTTCTGCTGCAGACAGTTGAGCTGGGGTCTGGGTGGGA TGGTGACATTTGACAGTGCTGCCATGTACAAAGTGAACATACAGATAAAGAGTGGCAGTGACAGGGACAC ACTCGCCAAAAGATTACCTGCAGCAGCTTCAGGAGGGACGTCATCTGCAGCGGGCTTCGATCGGCATTTACT GTGATTAGGAAGAAATATCCAATGATGGATCTGAAATTGCTGCTGACCGATGGGGAAGACAACACTATAA GTGGGTGCTTTAACGAGGTCAAACAAAGTGGTGCCATCATCCACACAGTCCGCTTTGGGGCCCTCTGCAGTCA AGAAGTACAGGAGCTGTCCAAAATGACAGGAGGTTTACAGACATATGCTTCAGATCAAGTTCAGAACAATGGC CTCATTGATGCTTTTGGGGCCCTTTCTATCAGGAAATGGAGCTGTCTCTCAGCGCTCCATCCAGCTTGAGAGTA AGGGATTAAACCTCCGAAACAGCCAGTGGATGAATGGCACAGTGATCGTGGACAGCACCGTGGGAAAGGACAC TTTGTCTTCTATCACCTGGACACGCAGCCTCCCCAAATCCTTCTCTGGGATCCAGTGGACAGAAGCAAGGT GGCTTTGTAGTGGACAAAACACCAAAATGGCCTACCTCCAAATCCCAGGCATTGCTAAGGTTGGCACTTGGA AATACAGTCTGCAAGCAAGCTCACAAACCTTGACCCCTGACTGTACGTCCTGCGTCCAATGCTACCCCTGCC TCCAATTACAGTGACTTCCAAAACGAACAAGGACACCAGCAAAATCCCAGCCCTCTGGTAGTTTATGCAAT ATTCCGCAAGGAGCCTCCCCAATTCTCAGGGCCAGTGTCACAGCCCTGATTGAATCAGTGAATGAAAAACAG TTACCTTGGAACTACTGGATAATGGAGCAGGTGCTGATGCTACTAAGGATGACGGTGTCTACTCAAGGTATTT CACAACCTTATGACACGAATGGTAGATACAGTGTAAGTGGGGCTCTGGGAGGAGTTAAGCAGCCAGACGG AGAGTGATACCCAGCAGAGTGGAGCACTGTACATACCTGGCTGGATTGAGAATGATGAAATACAAATGGAATC CACCAGACCTGAAATTAATAAGGATGATGTTCAACACAAGCAAGTGTTTACAGAGAACATCTCGGGAGG CTCATTTGTGGCTTCTGATGTCCCAATGCTCCATACCTGATCTCTTCCACCTGGCCAAATCACCAGCCTG AAGGCGGAAATTCAGGGGGCAGTCTCATTAACTGACTTGGACAGCTCCTGGGGATGATTATGACCATGGAA CAGCTCACAGTATATCATTGCAATAAGTACAAGTATTCTTGATCTCAGAGACAAGTTCAATGAATCTCTTCA AGTGAATACTACTGCTCTCATCCCAAAGGAAGCCAACTCTGAGGAAGTCTTTTGTGTTAAACCAAGAAACATT ACTTTTGAAAAATGGCACAGATCTTTTCAATTGCTATTGAGGCTGTTGATAAGGTCGATCTGAAATCAGAAATAT CCAACATTGCAGAGTATCTTTGTTTATTCTCCACAGACTCCGCCAGAGACACCTAGTCTGATGAAACGTC TGCTCCTTGTCTTAATATTCATATCAACAGCACCATTCTGGCATTACATTTTAAAAATATGTGGAAGTGG ATAGGAGAACTGCAGCTGTCAATAGCCTAGGGCTGAATTTTGTGAGATAAATAAAATCAATTCATCCTT TTTTTGATTATAAAATTTCTAAATGTATTTTAGACTTCCTGAGGGGGCATATACTAAATGTATATAGTA

	CATTTATACTAAATGTATTCTGTAGGGGGCGATATACTAAATGTATTTAGACTTCCTGTAGGGGGCGATAA AATAAAATGCTAAACAACCTGGGTAAA
#52	AATTAAATTATGAGAATTAAGACAAACATGAGCAGAGATGAAAAAGGAAGGGAGGAAAAGGTGGAAAAGA AAAGAAGACAAGAAGCGAGTAGTGGTCTCTAACTTGCTCTTTGAAGGATGGTCTCACAAAGAGAACCCCAACA GACATCATCGTGGGAATCAAATCAAGACCAGCAAGTACACCGTGTGTCTTCGTCCCCAAAAACATTTTGA GCAGCTACACCGGTTTGCCAATCTCTATTTTGTGGGCATTGCGGTTCTGAATTTATCCCTGTGGTCAATGCT TTCAGCCTGAGGTGAGCATGATACCAATCTGTGTATCTTGGCAGTCACTGCCATCAAGGACGCTTGGGAAG ACCTCCGGAGGTACAAATCGGATAAAGTCATCAATAACCGAGAGTGCCTCATCTACAGCAGAAAAGAGCAGAC CTATGTGCAGAAGTGCTGGAAGGATGTGCGTGTGGGAGACTTCATCCAAATGAAATGCAATGAGATTGTCCCA GCAGACATACTCCTCCTTTTTCTCTGACCCCAATGGGATATGCCATCTGGAACTGCCAGCTTGGATGGAG AGACAAACCTCAAGCAAAGACGTGTGCTGAAGGGCTTCTCACAGCAGGAGGTACAGTTCGAACCAGAGCTTTT CCACAATACCATCGTGTGTGAGAAACCAACAACACCTCAACAAATTTAAGGGTTATATGGAGCATCCTGAC CAGACCAGGACTGGCTTTGGCTGTGAGAGTCTTCTGCTTCGAGGCTGCACCATCAGAAACCCGAGATGGCTG TTGGCATTGTCTATGTCAGGCCATGAGACGAAAGCCATGCTGAACAACAGTGGCCCCGGTACAAACGAG CAAGATTGAGCGGCGCATGAATATAGACATCTTCTTCTGCATTGGGATCCTCATCCTCATGTGCCTTATTGGA GCTGTAGGTCACAGCATCTGGAATGGGACCTTTGAAGAACACCTCCCTTCGATGTGCCAGATGCCAATGGCA GCTTCCTTCCAGTGCCCTTGGGGGCTTCTACATGTTCTCACAATGATCATCCTGCTCCAGGTGCTGATCCC CATCTCTTTGTATGTCTCCATTGAGCTGGTGAAGCTCGGGCAAGTGTCTTCTTGGAGCAATGACCTTGACCTG TATGATGAAGAGACCGATTATCCATTCAATGTGAGCCCTCAACATCGCAGAGGACTTGGGCCAGATCCAGT ACATCTTCTCCGATAAGACGGGGACCTTGACAGAGAACAAGATGGTGTTCGACCTTGACCATCATGGCGAG CGAGTATTCTCACCAGAAAATGGTATAGAAGCTCCCAAGGGCTCCATCCCTCTTTCTAAAAGGAAATACCT GCTCTCCTAAGAAACGAGGAGATAAAAGACATCTCTCTGGCTCTCTTAGAGGCTGTGTGGCATTTCACAAGT TGCTTCCTGTATCCCTGTGGTCTTCTTGTACAGATCAGGGCTGTTCCAATTACTTTGAACTTTCAATTTGT TTACAAAGGTTAGAAGTTATCCCATATGTGGTTCCTTCAGCTGATCTTTGTCTGGTGCCAGACAAAGCACT TTATGAGACGAGTTTTTATCTGTGAGCAATGGATTGGAGACATTTCCCAATTGTGTGCCAGTCACACAACCA AGGCTTAGGAATTTCTCAGGCCACCTTACCTGACATGTCAGGGCAGGTCTGTGTCTAGGTGCATGGTCAGATT TAATACATCCAGAAGATGTCTTCTATTCTAACAGATCCTTAGCTTGTCACTGAGGCAAGTTTTGATTTAGG AGATAGGGCTATAAAATGCCCTGGACTGTTACCTTGATGGACTGAATATGACTCATAAACTGATCTGATTCC TTCAGCCATCATCTGCCCACTTGGTTCCCTCCCAACCCCCACACACACACACTTTCTAAGAAAA GAAAAGAAATCTTTTTTTTCAATACTTTAAGTTCTGGGATACATGTGCAGAATGTGCAGGTTTGTATACATAG GTATACATGTGTATGGTGGTTTGCAGCACCACCAACCCATCATCTACCTTAGGTATTCTCTTAATGCTAT CCCTCCCTTAGCCCCAACCCCCGATGGGCTCCAGTGTGTATGTTCCCTCCATGTCCATGTCTTCATT GTTCAATTTCCCACTTATGAGTGAGAACATGCAGTATTTGGTTTTCTGTTCTTGTGTTAGTTTGTCTGATGGTTT CCTGTTTCATCCGTGTCCCTGCAAAGGACATGAACCTCAICCTTTTTTATGGCTGCATAATATTCATGGTGTAT ATGTGCCACATTTTCTTTATCCAGTCTATCCCTGATGGGCACTGGGGTGGTTCCAAGTCTTTGCTATTGTGA ACAGTGCTGCAATAAACTTACATGTGCATGTGTCTTTAGTAGAATGATTATAATCCTTTGGGTATATACCCA GTAATGGGATTGCTGGTCAAATGGTATTTCTGGTCTAGATCCTTGAGGAATCTTTGTCTCCCAATGGTTG AACTAATTTGTACTCCCAACCAACAGTGTAAAGTATTCCCTGTTTCTCTACATCCTCTTCAGCATCTGTGTGT CCTGACATTTTAATGATCACTATTCTCACTGGCGTGAGATGTTATCTCATTGTGGTTTTGATTTGCATTTCTC TAATGACCAGTAATGATGAGCTTTTTTTCATATGTTTGTGGCTGCATAAATGTCTTCTTTTGAGAAGTGTCT GTTTCATATCCTTCACCCATTTTTTGAAGAAAACAACTCTTAAGAGAGCAGTATTCATTCTTTTGAGTGTGAG GGATGGAGAAAGAGAAAGATGGAGAGAGTATTATAAGCAGCTGTATCCCTTTGCCATGGTGATAGCAGACCA TTCACATGGGAGCTTCTGGTCTCTTTGTAATAATAAAGAGCCACATTACCAGTACTTAGAGTATGCTAGTT ATTTTAACACATTTGTATCATTAATCTTCAAACATCCCTATGAGTTAGAAACCTAAAAAATAAAAAA A
#53	CTCATTTTGTATGTCTAGAATCAGGGGATCCAGGATCATCACCAGGTCATTTTCCAGGTATGGAGGGGCTT TCTGCTTCTTTCTTGTGATGCACAGCTGCTGAGGAAGGGGCTGGGAGTAAAGACAGTGAATGGGGAGGAGGA GTCCATTCAAACCGAGAAACAAAGTGTGTTTCTTACCCCTGGTGTAGAAGCTACCAACCTTTTCCAAG AAAGAGGGCCTGGCCCCCTTCTCGGGTCTGGCTGGGTGCCTGCTGTGCCTCTCTGGCCTCCCCCGAAGGGC ACCATTCCCTCGGGTGAGTACTACCGGCTGCACCGTCTTCCAGTGGGGACAGCCTGAGAAGAGAGTCTGGG CCTTACTTCAGTACCTTCCTTCACTGGCCTCACCCTGTGCAAAATCATGCCACACGCTGCAGCCTCCTTTCCC TATCTATAAAATAAAATGACCCTGCTCTATCTCACTGGGCTGGCAAGAACACACTGTTGTTGCCTTGACAGAC AGATGTGTGAGGCTGTAGAAAGTGTCTTTTATTTGGTTGGGAGCTTGTGCATAAATGCAGAGGGGCTGCAC ATCTGACGGACTAGAGGTGACTCATGGCTGAACCGGAACAGGACATCGGGGAGAAGCCAGCAGCCATGCTGAA CTCTCCACAGGGCCCTGTGAAAAGCTCTTCACTCCTCTGCCCTCTGGATCTAGTGAAGCCTATTATCCCTTC AGATGTGAGCTCAAATAATCAACCTTCATGGAGGCTCCCTTGACCCCTAACATGCTTTCAAAGTACTGTGA TTTCACATTTCATCATGCCCCGACAACGTGATTTCCCATTTATTAATATCTGTCTCTTCTGCTGGCCTGCAAA CTCCAGGAGCACAGAGACATCTTGGGATTTTGAACATGATTTCCCAGGGCTTAGCCAGTGCCTGGTGCA AAGCAGGCTTTCAACATGTTCACTGGATATTGTAAGAAAGAAAGAAATACACAAAAGGCTGGCATATGCAAA GCACTCTAAATATTCACCTCCTTTCCCTTCCCTCTGGGTGAGAAAATTTCTCCTTATAAAGACACCTCCTAAC TGTATCTCTGTAGAGAACTGAAGACATAAGCACTCTGTGCCAAAATATTTAAGTAAAACTTGAGCTAAG CACAGAGATTATAAATATTTCTTCCCAGATTACGCACCATTTAAAAAATACTGTCTCAGCTCCTTTTCATGAT

	TTGGGTGGTGATTAAAGAAAATTACTCTTCAAGACTGAAAGTCATTACTGCCCTTTTCTGACTTGCCTTTTC CCTTGAGAAGGGGAGGATAAGCTGCAGGGCAGGAAGTGAAGTGGGGCATCCTTGTCTTGTCTGGCAGACA GCCAACTGGTCAGGTACTGCTCCTTCTCAACTCTTCTCTGATTCCCAGGTGAATATAACAAGAAGGCACAAA TCCACACTTGCCAAACAGGACCCAGTGATAACAAGAAACCCAGTGACACCTGTCTAGGTGAAGACTCAGCC CCTATGTGACCAGGTTGCAAAGCCAACTGACCATCTGCTTTCCATTTGGACTTTTAGTTTCATACTGTATCTT CTCAGGACAGTTAAGTTGGAATACAATGCCACTGTCTGAAAGATGGTAGAATTATCCTATTTCTGGAGGAGT GGGGTGGTGGGTAGGAATCTCAAGAGCGATTGGCTCTCTGCACAATAGCTTCTTTAAGGACACCAGGGCCC CCAGGGCTATACATTTCCCTGAAGCTTTCCAGATAAGCAACAAGGTATGAGCACCTGCTATGTATTGCCCAAG GGTGATGTGTTTTAAATATCCATTGCATATTTTAAATCCTTGGCTGGCTTAAAGCTGCAAGCTTTCTGTCTTCA GTGGATATAATGGGGGCATACATCCCAGAGCTTGCCCAACACTCCAAGAAAAGAACCCTCAGCTAATGCAAG TGTGTATGTGCCCATGAAAGCTCCATGTCTACTTAACATTCACTTTTAGGATTATTTATGCTGTAATAATAG ATATGAAAATCTCTGACAGGTATTTTGTTCCTTTACAAACTGTATTTGAATTTATGGGTGATTTAGAGCTTG TGTTTAAAGTCAGAATTCAGAACCCCAAAGAAAATGACTTATTGAAATGAAGTGAAGAGACAAGAACTGAG TTACCAAAAACCTACTAAACGTGAGTTGCTGTGAAGTGGGATTAAACCAGAACGAGTGAGAGAATCAGAAAG CTACCAAAACACTGCTCAGAAAGGACAAAGACATTCGAAGACTGCGGGACTTTTCAGGAAGTGAAGTCAATTT TAATGAAAATGGAAGCTCCAGATTGACAGAATATGTCCATCTCTGACAGAAAGGCCCTGCTATGATAGCAA AGCTGCAAAAATGACTTATTAATACTCCCAGGAATGGCCGCGCATGGTGGCTCACCCCTGTAATCCCAGCA CTTTGGGAAGCCAAGGTGGGCGGATCACCTGAGGTGAGGAGTCTAGACCAGCCTGGCCAACATATAGTGAAA CCAGTCTCTACTAAAAAAAATACAAAAATTAGCTAGGTGTGGTGGCGCACACCTGTAGTAGTCCCAGCTACA TGGGAAGCTGAGGCAGGAGAATCACCTGAACCCAGGAGGAGGTGTCAGTGAGCTGAGATTGCGCCACTGC ACTCCAGCCTGGCGACAGAGCAAGACTCTGTCTCTCAAATAAATAAATAAATAAATAAATAAATAAATAAAT AATC
#54	GCCCGGGAGAGGAGAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCT GACACCTGGGAAGATGGCCGGCCCGTGGACCTTACCCCTTCTCTGTGGTTTGGTGGCAGCCACCTTGATCCAA GCCACCCTCAGTCCCAGTGCAGTTCTCATCCTCGGCCCAAAAGTCATCAAAGAAAAGCTGACACAGGAGCTGA AGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGTGCTCAGTGCCATGCGGGAAAAGCCAGCCGAGG CATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCCTGAAGCACATCATCTGGCTGAAGGTCTACAGCTAAC ATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGG CTGGATTCAACACGCCCTTGGTCAAGACCATCGTGGAGTTCCACATGACGACTGAGGCCAAGCCACCATCCG CATGGACACCAGTGCAAGTGGCCCAACCCGCTGGTCCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGC ATCCAACCTGCTGCATAAGCTCTCCTTCTGGTGAACGCCCTAGCTAAGCAGGTGATGAACCTCCTAGTGCCAT CCCTGCCCAATCTAGTGAAAACAGCTGTGTCCCGTATCGAGGCTTCTTCAATGGCATGTATGCAGACCT CCTGCAGCTGGTGAAGGTGCCCATTTCCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTATCCTGCCATC AAGGGTGACACCATTCAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAAAGGTGACCAAGTGGTTCA ATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGACAACATCCCGTTCAGCCTCATCGTGAGTCAGGACGT GGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGATTATGGTCTGTGGACTCTGTGCTTCTGTGAG AGTGCCCATCGGCTGAAGTCAAGCATCGGGCTGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGA TCGTGAAGATCCTAACTCAGGACACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGAT CGTGCTGGAAGTGTTCCCTCCAGTGAAGCCCTCCGCCCTTTGTTACCCTGGGCATCGAAGCCAGCTCGGAA GCTAGCTTTTACACCAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTGA TGAAGTCTGGGATTGGCTGGTTCCAACTGATGTTCTGAAAAACATCATCACTGAGATCATCCACTCCATCCT GCTGCCGAACCAGAAATGGCAAATTAAGATCTGGGGTCCCAGTGTCAATTGGTGAAGGCCTTGGGATTGAGGCA GCTGAGTCTCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCCTTGTGGAACCCAGCTCTCCTGTCT CCCAGTCAAGACTTGGATGGCAGCCATCAGGGAAGGCTGGGTCCCAGCTGGGAGTATGGGTGTGAGCTCTATA GACCATCCCTCTCTGCAATCAATAAACACTTGCCTGTGAT
#55	GGAGTGGGGGAGAGAGAGGAGACCAGGACAGCTGCTGAGACCTCTAAGAAGTCCAGATACTAAGAGCAAAGAT GTTTCAAACCTGGGGGCTCATTGTCTTCTACGGGCTGTTAGCCAGACCATGGCCAGTTTGGAGGCCTGCCC GTGCCCTTGACCAGACCCTGCCCTGAATGTGAATCCAGCCCTGCCCTTGAGTCCCACAGGTCTTGCAGGAA GCTTGACAAATGCCCTCAGCAATGGCCTGCTGTCTGGGGGCTGTTGGGCATTCTGGAAAACCTTCCGCTCCT GGACATCCTGAAGCCTGGAGGAGTACTTCTGGTGGCCTCCTTGGGGGACTGCTTGGAAAAGTGACGTGAGT ATTCTTGGCCTGAACAACATCATTGACATAAAGTCACTGACCCCCAGCTGCTGGAATTTGGCCTTGTGCAGA GCCCTGATGGCCACCCTCTCTATGTACCATCCCTCTCGGCATAAAGCTCCAAGTGAATACGCCCTTGGTGG TGCAAGTCTGTTGAGGCTGGCTGTGAAGCTGGACATCACTGCAGAAATCTTAGCTGTGAGAGATAAGCAGGAG AGGATCCACCTGGTCTTGGTGACTGCACCCATTCCCTGGAAGCCTGCAAATTTCTCTGCTGTATGGACTTG GCCCTTCCCATTCAAGGTCTTCTGGACAGCCTCACAGGGATCTTGAATAAAGTCTGCTGAGTTGGTTCA GGGCAACGTGTGCCCTCTGGTCAATGAGGTCTCAGAGGCTTGGACATCACCCCTGGTGCATGACATTGTTAAC ATGCTGATCCACGGACTACAGTTTGTCTCAAGGTCTAAGCCTTCCAGGAAGGGGCTGGCCTCTGCTGAGCTG CTTCCAGTGCTCACAGATGGCTGGCCATGTGCTGGAAAGATGACACAGTTGCCTTCTCTCCGAGGAACCTGC CCCTCTCCTTTCCCAACAGGCGTGTGTAACATCCCATGTGCCTCACCTAATAAATGGCTCTTCTCTGCAA AAAAAAAAAAAAAAAAAAAAAAAAA

#56	GAGCAGAGCCCTTTCACACACCTCAGGAACACCTTTCGGCTGCCCGCTCCCCAGACACACCTGCAGCCCTGCC CAGCCGGCTTTGCTCACCCTGCTTGTAAATGCCCCAGATATGAGCCAGCCAGGCCCGCTACGTGGTAGA CAGAGCCGCATACTCCCTTACCCTCTTCGACGATGAGTTTGAGAAGAAGGACCGGACATACCCAGTGGGAGAG AAACTTCGCAATGCCTTCAGATGTTCCCTCAGCCAAGATCAAAGCTGTGGTGTTCGGCTGCTGCTGTGCTCT CCTGGCTCCCCAAGTACAAGATTAAAGACTACATCATTCCTGACCTGCTCGGTGGACTCAGCGGGGATCCAT CCAGGTCCCACAAGGCATGGCATTGCTCTGCTGGCCAACCTTCTGCAGTCAATGGCCTCTACTCCTCCTTC TTCCCCCTCCTGACCTACTTCTCCTGGGGGGTGTTCACCAGATGGTGCCAGGTACCTTTGCCGTATCAGCA TCCTGGTGGGTAACATCTGTCTGCAGCTGCCCCAGAGTCGAAATTCAGGTCTTCAACAATGCCACCAATGA GAGCTATGTGGACACAGCAGCCATGGAGGCTGAGAGGCTGCACGTGTGAGCTACGCTAGCCTGCCTCACCGCC ATCATCCAGATGGGTCTGGGCTTCATGCAGTTTGGCTTTGTGGCCATCTACCTCTCCGAGTCCTTCATCCGGG GCTTCATGACGGCCGCCGGCTGCAGATCCTGATTTGGTGCTCAAGTACATCTTCGGACTGACCATCCCCTC CTACACAGGCCCAGGGTCCATCGTCTTTACCTTCATTGACATTTGCAAAAACCTCCCCACACCAACATCGCC TCGCTCATCTTCGCTCTCATCAGCGGTGCCTTCTGGTGTGGTGAAGGAGCTCAATGCTCGTACATGCACA AGATTGCTTCCCCATCCCTACAGAGATGATTGTGGTGGTGGTGGCAACAGCTATCTCCGGGGGCTGTAAGAT GCCCAAAAGTATCACATGCAGATCGTGGGAGAAATCCAACGCGGTTCCCCACCCCGTGTGCCTGTGGTC TCACAGTGGGAAGGACATGATAGGCACAGCCTTCTCCCTAGCCATCGTGAAGTACGTCATCAACCTGGCTATGG GCCGGACCTTGCCCAACAAGCACGGCTACGACGTGGATTGAACAGGAGATGATCGCTCTCGGCTGCAGCAA CTTCTTTGGCTCCTTCTTTAAATTCATGTCATTTGCTGTGCGCTTCTGTCACTCTGGCTGTGGATGGAGCT GGAGGAAAATCCAGGTGGCCAGCCTGTGTGTCTCTGGTGGTGTGATGATCACCATGCTGGTCTGGGGATCT ATCTGTATCCTCTCCCTAAGTCTGTGCTAGGAGCCCTGATCGCTGTCAATCTCAAGAACTCCCTCAAGCACT CACCAGCCCTACTACCTGTGGAGGAAGCAAGCTGGATGTCATCTGGGTAGTGAGCTTCTCTCTCTCTCTC TTCTTCTCAGCCTGCCCTATGGTGTGGCAGTGGGTGTGCGCTTCTCCGTCCTGGTGTGGTCTTCCAGACTC AGTTTCGAAATGGCTATGCACTGGCCAGGTGATGGACACTGACATTTATGTGAATCCCAAGACCTATAATAG GGCCAGGATATCCAGGGGATTAAATCATCAGTACTGCTCCCCTCTCTACTTTGCCAAGTACAGAGATCTTC AGGCAAAAGGTGATCGCCAAGACAGGCATGGACCCCAAGAAATATTACTAGCCAAAGCAAAATACCTCAGA AGCAGGAGAAGCGGAGAATGAGGCCACACAACAGAGGAGTCTCTATTGATGAAAACCAAGACTGTCTCCCT GCAGGAGCTGCAGCAGGACTTTGAGAATGCGCCCCCAGCCAGCCCAACAACAACCAAGCCCGCTAACGGC ACCAAGCGTGTCTATATCACCTTCAGCCCTGACAGCTCCTCACCTGCCAGAGTGAGCCACCAGCCTCCGCTG AGGCCCCCGCGAGCCAGTGCATGCTGGCCAGCGTCCCACCCTTCTGTCACCTTCCACACCTCATCCTGGA CATGAGTGGAGTCAGCTTCGTGGACTTGATGGGCATCAAGGCCCTGGCCAAGCTGAGCTCCACCTATGGGAAG ATCGGCGTGAAGGTCTTCTTGGTGAACATCCATGCCAGGTGTACATGACATAGCCATGGAGGCGTCTTTG AGGATGGGAGTCTAGAATGCAAGCAGCTTTTCCAGCATACATGAGCAGTCCCTTTTGGCCAGGCAATGC TAGAGACGTGACCCAGGACACAACCTTCCAAGGGGCTCCAGGGGATGCTGAGCTCTCCTTTGACGATGAGCA GAGGACATTCGAGCTACTGGGACTTAGAGCAGGAGATGTTCCGGAGCATGTTTCAGCAGAGACCTGACCG CCCTGTGAGGGCTCAGCCAGTCTCATGCTGCCTACAGAGTGCCTGGCACTTGGGACTTCCATAAAGGATGAG CCTGGGGTTCAGGGGGGTGTCGGGCGGAGAAAGTGATCCCCCAGAGCTTGGGTCTCTCTCTCTCTCTCTCT TCT GGAGTGAGAGTCTGGTGAGCCACTTTCACCCGTGAGGCCCTGGCCGCAATGGACAAGCCTCTGCTCATC CACCACCCACATCTGCCCTGCTCTTGGCAGCTGAAGGACACCTTGACTTCCAGCTTTTACGATGAGCCAA AAACAGAAGGACAAGTACAACCTGTGCTGGCCTGCTGTACAAGCTTCAAAAAGTGTCCAGAGCCCGCACGGCT CGGTGTGAGATGGTGTGAGGCTGTACGGACATAGGGAATAAAGTTGGTTAGGACTCTGGCTTGCTTCCCCAG CTGCCTCAACTCTGTCTCTGTCAGCTCTGACCCAGGGACATGTGCTCTCACACCCAGGAGTCTAGGCCTT GGTAACATGCGCCCCCTCCATCATCCCCAAGGTGCCCAACCAACCACTGCTGTGACGAGCAGATCAGA CTTAGCCTGGACAGTGGCCAGGACCGTCGAGACCACCAAGAGTACCTCCCCGGGACAGCCCACTAAGGTT TGCTCAGCCTCCTGAAACATCACTGCCCTCAGAGGCTGCTCCCTTCCCCTGGAGGCTGGCTAGAAACCCCAA AGAGGGGATGGGTAGCTGGCAGAATCATCTGGCATCTAGTAATAGATACCACTTATTCTGCACAAACTTT TGGGAATTCCTCTTTGCACCAGAGACTCAGAGGGGAAGAGGGTGTAGTACCAACACAGGGAAACCGATGG GACCTGGGCCCAGACAGTCCCCCTTGACCCAGGGCCCATCAGGGAATGCCTCCCTTTGGTAAATCTGCCTT ATCCTTCTTTACCTGGCAAGAGCCAATCATGTTAACTCTTCTTATCAGCCTGTGGCCAGAGACACAATGG GGTCTTCTGTAGGCAAGGTGGAAGTCTCCAGGATCCCTACATCCCCTAAGTGCATGAGATGTGGAAA GGGGCTGATCCAGATTGGGTCTTCTGTCACAGGAAGACTCTTTAACACCTTAGGACCTCAGGCCATCTCTC CTATGAAGATGAAAATAGGGGTTAAGTTTTCATATGTACAAGGAGGTATTGAGAGGAACCTACTGTTGACT TGAAATATAATAGGTTCCATGTGTAAGTGTTTGTAAATTTCAAGTGGAAATGCACAGAAATCTTCTGGCCT CTCATCACTGCTTTTCTCAAGCTTCTTCACTTAACAACCCCTTCCCTAACAGGTTGGGCTGGCCAGCCTAG GAAACATCCCATTCTTAACCTCAGCCAGACCTGCGTTGTGTCTGTGTGTGTGAGTGAAGTGGTCAGCTAA CAAGTCTTCTTAGAGTTAAAGGAGGGGTGCTGGCCAAGAGCCAACACATTCTTGGCCAGGAGCATTGCTTT TCTGTGAATTCATTATGCCATCTGGCTGCCAATGGAACCTCAAACTTGAAGGCGAAGGACAATGTTATCTGG GATTCACCGTGCCAGCACCAGAGTGCCTAATTCAGGAGGACAAGAGCCTTAGCCAATGACAACCTCACTCT CCCCACTCCACCTCCTTCCAAGTCCAGCTCAGGCCAGGAGGTGGGAGAAGGTACAGAGCCTCAGGAATTT CCAAGTCAGAGTCCCCCTTGAACCAAGTATCTAGATCCCTGAGGACTGATGAAGTGATCCTTAACCCCAA GTAATCATTAACCCCAAGACAGCCTCAGAACTGAAGGAGATTGTTGACCCAGTGACCTGGAGTTGAGGCTCA
-----	---

	GGGAGAGATCTGCCACATGTCTGAGGGTTGCAGAGCCCGCTGTGGAGGTAAGATTGGAAACACATGAGGCAGA GGGAAGACATTGAAGAAAACATCTCTGCTGGAATATTTGGAAAAGAACACTCTTCTGGACCTGGTGAAGCAG GAAAGATGGAGGCAAAGTAGTGAAATAATCCAGAATTTCAATGCTTTTGAATGTTCTTAGTGATACGACCTG TGATAATAATAATCCCAGGAGGACTGGGAACCTTATCTCTTGAGATATTTGCATAATTTATTTAATTTAAGC CTCATTCTCCTTTTGTTCATTTTGGTAATAAACTGGATTGGAATTGTGAACAAAAAAAAAAAAAAAAAAAA
#57	AATGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCACCCAAAAACAAGGAGATCCCAT CTAGATTTCTTCTTGTCTTTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTTCATCTTTGGAGAGGCCAAA TGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTTCCACCATAAAACAGCTGAGTTATTTATGAATTAGAGGCTA TAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAACTTTCTACTCTGATGAGAGAATGTGTTTTTA ATCTCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTCTCCTCCTAGTCAATAAAACCCATTGATGATC TATTTCCCAGCTTATCCCCAAGAAAACCTTTTGAAAGGAAAGAGTAGACCCAAAGATGTTATTTTCTGCTGTTT GAATTTTGTCTCCCCACCCCCAAGCTTGGCTAGTAATAAACACTTACTGAAGAAGAAGCAATAAGAGAAAGATA TTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAGTTACCAAACCAAAGTCATT TTCAGTTTGGGCAACCAAACCTTTCTACTGCTGTGACATCTCTTATTACAGCAACACCATTCTAGGAGTT TCCTGAGCTCTCCACTGGAGTCTCTTTCTGTGCGGGGTGAGAAATTGTCCTAGATGAATGAGAAAATTATT TTTTTTAATTTAAGTCTTAAATATAGTTAAATAAATAATGTTTGTAGTAAATGATACACTATCTCTGTGAAA TAGCCTCACCCCTACATGTGGATAGAAGGAAATGAAAAATAATTGCTTTGACATTGCTATATGGTACTTTG TAAAGTCATGCTTAAGTACAAATTCATGAAAAGCTCACTGATCCTAATTTCTTCCCTTTGAGGTCTCTATGG CTCTGATTGTACATGATAGTAAGTGTAAGCCATGTAAAAAGTAAATAATGTCTGGGCAGTGGCTCAGCCT GTAATCTTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGTTCGAGACTAGCCTGGGCAA CATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAATCAGCCAGTCATGGTGGCATACACCTGTAGTC CCAGCATTCCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGTTGGGGCTGCACTGAGCCATGATC ACACCCTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAATAAAAAATAAATAATGGAACACA CGAAGTCCTAGGAAGTAGGTTAAACTAATTTCTTTAAAAAAGTTGAGCCTGAATTAATGTAATG TTTCCAAGTGACAGGTATCCACATTTGCATGGTTACAAGCCACTGCCAGTTGGCAGTAGCACTTTCTGGCAC TGTGGTGGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG CACTCAAGCAATTTCTTACCTGGCCTCCCAAGTAGCTGGAATTACAGGTGTGCGCCATCACAACCTAGCTGG TGGTCAGTTTTGTTACTCTGAGAGCTGTTCACTTCTCTGAATTCACCTAGAGTGGTTGGACCATCAGATGTTT GGGCAAACTGAAAGCTCTTTGCAACCACACACCTTCCCTGAGCTTACATCACTGCCCTTTTGAGCAGAAAGT CTAAATTTCTTCCAAGACAGTAGAATTTCCATCCCACTACCAAAGCCAGATAGGCCCCCTAGGAACTGAGGTA AGAGCAGTCTCTAAAAACTACCCACAGCAGCATTGGTGCAGGGGAAGTTGGCCATTAGGTTATTATTTGAGAG GAAAGTCTCACAATAGTACATATGAAAGTGACCTCCAAGGGGATTGGTGAATACATAGAGATCTTCA GGCTGAACAGACTATGTCTGGGGAAGAACGGATTATGCCCATTAATAACAAGTTGTGTTCAGAGTCAAG GCAGTGAGCTCAGAGGCCCTTCTCACTGAGACAGCAACATTTAAACCAAACAGAGGAAGTATTTGTGGAAGT CACTGCCTCAGTTTGGGTAAAGGATGAGCAGACAAGTCAACTAAAGAAAAAGAAAGGAGGAGGTTG AGCAATCTAGAGCATGGAGTTTGTAAAGTGCTCTCTGATTGAGTTGAGGTTGAAGAGCATCCATTTGAGTTGAAGGC CACAGGGCACATGAGCTCTCCCTTCTACCACCAGAAAGTCCCTGGTCAGGTCTCAGGTAGTGCGGTGTGGCT CAGCTGGGTTTTTAATTAGCGCATTCTCTATCCAACATTTAATTGTTTTGAAAGCCTCCATATAGTTAGATTGT GCTTTGTAATTTG TAAATATTAACCACTGTTATCCTACAAAAAACCTCAAAGGCTGAAATAAAGAAGGAAGATGGAGACAC CCTCTGGGGTCTCTC
#58	CTTTGCACTGGATGCCCTTGGCAGGGTGAGCCACAAGGAGCAATGGAGCAGGGCAGCGGCCCTTGGAGGAC TTCCCTGTCAATGTGTTCTCCGTCACTCCTTACACACCCAGCACCCTGACATCCAGGTGTCCGATGATGACA AGGCGGGGGCCACCTTGCTCTTCTCAGGCATCTTTCTGGGACTGGTGGGATCACATTCATGCTGATGGGCTG GATCAATAACCAAGGTGTCTCCACTTTGAATGGAGCCAGCTCCTTGGGCCCCCTGCTGCTGCTGCTGCTGCTG ACATTCATCCTGATTGCTGTGTGCAAGTTCAAATGCTCTCCTGCCAGTTGTGCAAGAAAGTGAGGAAAGGG TCCCGGACTCGGAACAGACACCAGGAGGACCATCATTGTTTTCCTGTCATCAACCAACCCATCACCTTCCA TGGGGCCACTGTGGTGAGTACATCCCTCCTCCTTATGGTTCTCCAGAGCCTATGGGGATAAATACCAGCTAC CTGCACTCTGTGGTGAGCCCCCTGCGGCCTCATAACCTCTGGAGGGGCAGCAGCCGCCATGTCAAGTCTCTCTC AATACTACACCATCTACCTCAAGATAACTCTGCATTTGTGGTTGATGAGGGCTGCCTTTCTTTCACGGACGG TGGAAATCACAGGCCCAATCCTGATGTTGACCAGCTAGAAGAGACACAGCTGGAAGAGGAGGCCCTGTGCTGC TTCTCTCCTCCCCCTTATGAAGAAATATACTCTCTCCTCGCTAGAGGCTATTCTGATATAATAACACAATGC TCAGCTCAGGGAGCAAGTGTTCCTGCTATTGTACCTGACAACCGTGGTGTCTATGTTGTAACCTTCAGAAG TTACAGCAGCGCCAGGCAGCCTGACAGAGATCATTCAAGGGGGAAAGGGGAAGTGGGAGGTGCAATTTCTC AGATTGGTAAAAATTAGGCTGGGCTGGGGAATTTCTCCTCCGGAACAGTTTCAAAATCCCTCGGGTAAGAAAT CTCCTGTATAAGGTTCAAGAGCAGGAATTTCACTTTTTCATCCACCACCTCCCCCTCTCTGTAGGAAGGCA TTGGTGGCTCAATTTTAACCCAGCAGCCAATGGAATAATCAGCACTTCTGAGACTTTGGGAGTTTCCACAGA GGTGAGAGTCCGGTGGGAAGGAAGCAGGGAAGAGAAAGCAGGCCAGCTGGAGATTCTCTGGTGGCTGTCTCT

	GGCCCCAAGCAGACTCACTAATCCCAACAACCTCAGCTGCCATCTGGCCTCTCTGAGGACTCTGGGTACCTTAAAGACTATA
#59	CAGGAAAGTTCTGCTGCTAGGCAGAGGAAGTGCAGCTTGTGGCAGGTGAAGGGAGCCTGTTAGCTGTGTCAGCAACAACCTTACGTGGTCCTGCTTGTGTCCAGGTGAAGCGTCTGGCCGCCGAGCAGAGGAATCAAGACCTGCTCATTCTTTCTCGGGGGATCCATCCAGCAATGACATCATCTCATGCTGCCACAAGGACCCCAAGTCTGGGCTGCTGGGGACAGCCAGCTCCCCACTGCTCATTCCCTTCATCCTAGAGACATTCTGACTCTCCTCCGACTGCGCTGTGGCAGGGCTGACAAGCTCTTTTACATCTCAGTCTGCACAACCTTCAGGCACTTAGCAGATTGATATGCATCCAACAAATATTGATTGAATATCTGCTAAATACCCAGTAATGTTTCATGAGTGATTGGGTGAATAAAGGAAATGCTGGTTCCTTCTGGCCATATTAACCTCCTGCACAATACTAAGAAAAATAAATGCAGCTAGCTGTGGAATAATGTGAATCCCAATGTCTATCTATTGAAATATTACCTGACTATTAAGAGGTATTTATTTTGTATCTTTTCTAGCAAGATAAATAAATCTTAATACAGCATATCCCCTTATTCACGGGGGTATGTTCAAGACCCCGGTGGATGCTGAAACTATGGATAATACCAGATCC
#60	MGPFFKSSVFILILHLEGALSNLIQLNNNGYEGIVVAIDPNVPEDETLIQQIKDMVTQASLYLFEATGKRFYFKNVAILIPETWTKADYVRPKLETYKNADVLVAESTPPGNDEPYEQMGNCGEKGERIHLTPDFIAGKKLAEYGPQGKAFVHEWAHLRWGVFDEYNNDKEFYLSNGRIQAVRCSAGITGTNVVKKCQGGSCYTKRCTFNKVTGLYEKGCEFLVQSRQTEKASIMFAQHVDSDIVEFCTEQNHNKEAPNKQNKCNLRSTWEVIRDSSEDFKKTTPMTTQPPNPTFSLQIGQIRIVCLVLDKSGSMATGNRLNRLNQAGQLFLQTVELGWSWGMVTFDSAHHVQSELIQINSGSDRDTLAKRLPAAASGGTSCSGLRSAFTVIRKKYPTDGEIVLLTDGEDNTISGCFNEVKQSGAIHTVALGPSAAQLEELSMTGGLQTYASDQVQNNGLIDAFGALSSGNGAVSQRSIQLESKGLTLQNSQWMNGTIVVDSTVGKDTLFLITWTTQPPQILLWDPGQKQGGFVVDKNTKMAYLQIPGIKVGWTKYSLQASSQTLTLTVTSRASNATLPPIITVTSKTNKDTSKFPSPLVYANIRQGASPILRASVTALIESVNGKTVTLELLDNGAGADATKDDGVYSRYFTTYDTNGRYSVKVRALGGVN AARRRVIPOQSGALYIPGWIENDEIQWNPPRPEINKDDVQHKQVCFSRSSSGGSFVASDVPNAPIPDLFPPGQITDLKAEIHGGSLINLTWTAPGDDYDHGTAHKYIIRISTSLDLRDKFNESLQVNTTALIPKEANSEEVFLFKPENITFENGTDLFIAIQAVDKVDLKSEISNIARVSLFIPPQTPPETPSPDETSAPCPNIHINSTIPGIHLKI MWKWIGELQLSIA
#61	MKKEGRKRWKRKEDKKRVVVSNNLFEGWSHKENPNRHRGNQIKTSKYTVLSFVPKNI FEQLHRFANLYFVGI AVLNFI PVVNAFQPEVSMIPICVILAVTAIKDAWEDLRYKSDKVINNRECLISYRKEQTYVQKCKWDRVVGDFIQMKCNEIVPADILLFSSDPNGICHLETASLDGETNLKQRRVVKGFSSQEVQFEPELFHNTIVCEKPNHLNKFKGMEHPDQTRTGFGCESLLRGCTIRNTEMAVGIVYAGHETKAMLNNSGPRYKRSKIERRMNIIDFFCIGILILMCLIGAVGHSIWNNGTFEEHPPFDVPDANGSFLPSALGGFYMFLTMILLQVLIPISLYVSTIELVKLGQVFFLSNDLDLYDEETDLSIQCRALNIAEDLQIQIYIFSDKTGTLTENKMVFRCTIMGSEYSHQENGIEAPKGSIPLSKRKYPALLRNEEIKDILLALLEAVWHFHKLLFVSLWSSLSQIRAVPITCKLSFVYKG
#62	MGRRSPEKPRNKVFGFSYPWCRSYQPFPRKRAWPPSRVWLACCASLASPPKGTIPSGEYYPAPSSSGDSLRESGALLQYLPSLASPCANHATRCSLFPYIKIMTLLYLTGLARTHCCCLADRCAEAVESAFYLVGSLCINARGAAHLTD
#63	MAGPWTFLLCGLLAATLIQATLSPTAVLILGPKVKEKLTQELKDHNATSILQQLPLLSAMREKPAGGIPVLGSLVNTVLKHIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIIVEFHMTEAQTIRMDTASGPTRLVLSDCATSHGSLRIQLLHKLSFLVNALAKQVMNLLVPSLPNLVKNQLCPVIEASFNMGYADLLQLVKVPSISLSIDRLEFDLLYPAIKGDTIQLYLGAALLDSQGVTKWFNNSAASLTMPITDNI PPSLIVSQDVVKAAVAAVLSPEEFMVLLDSVLPESAHLKSSIGLINEKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTLGIEASSEAQFYTKGDOLIILNLISSORIQLMNSGIGWFQPDVLKNIITETIHSILLPNQNGKLRSGVPVSLVKALGFEEAESSITKDALVLT PASLWKPSSPVSQ
#64	MFQTGGILIVFYGLLAQTMAQFGGLPVFLDQTLPLNVNPA LPSPTGLAGSLTNALSNGLLSGGLLGILENLPLLDILKPGGGTSGGLLGGLGKVTSVIPGLNNIIDIKVTD PQLELGLVQSPDCHRLYVTIPLGIKLQVNTPLVGASLLRLAVKLDITAEILAVRDKQERIHVLVDCTHSPGSLQISLLOGLGPLPIQGLLDSLTGILNKVLPVLVQGNVCPLVNEVLRGLDITLVHDIVNMLIHGLQFVIK
#65	MSQPRPRYVVDRAAYSILTFDDEFKDRTPVGEKLRNAFRCSAKIKAVVFGLLPVLSWLPKYKIKDYIIPDLGGSLSGSSIQVPOGMAFALLANLPAVNGLYSSFFPLTYFFLGGVHQMVPGTFAVISILVGNICLQLAPESKFQVFNNATNESYVDTAAMEERLHVSA TLACLTAIQMG LGFMQFGFVAIYLSSEFIRGFMTAAGLOILISV

	LKYIFGLTIPSYTGPFSIVFTFIDICKNLPHTNIASLIFALISGAFLVLVKELNARYMHKIRFPIPTMIVVV VATAISGGCKMPKKYHMQIVGEIQRGFPTVSPVVSQWKDMIGTAFSLAIVSYVINLAMGRTLANKHGYDVS NQEMIALGCSNFFGSSFFKIHVICCALSVTIADVGAGGKSQVASLCVSLVVMITMLVLGIYLYPLPKSVLGALI AVNLKNSLKQLTDPYYLWRKSKLDCCIWVVSFLSSFFLSLPYGVAVGVAFSVLVVVFQTFQFRNGYALAQVMDT DIYVNPKNYNRAQDIQGIKIITYCSPLYFANSEIFRQKVIKTGMDPQKVLAKQKYLKKQEKRRMRPTQQR SLFMKTKTVSLQELQDFENAPPTDPNNNQTPANGTSVSYITFSPDSSSPAQSEPPASAEAPGEPDMLASVP PFVTFHTLILDMSGVSFVDLMGIKA LAKLSSTYKGKIGVKVFLVNIHAQVYNDISHGGVFEDGSLECKHVFPSIHDAVLFAQANARDVTPGHNFQGAPG DAELSLYDSEEDIRSYWOLEQEMFGSMFHAETLTAL
#66	MEQSGRLEDFPVNVFSVTPYTPSTADIQVSDDDKAGATLLFSGIFLGLVGITFTVMGWIKYQGVSHFEWTQL LGPVLLSVGVTFILIAVCKFKMLSCQLCKESEERVDPDSEQTPGGPSFVFTGINQPIFHGATVVQYIPPPYGS PEPMGINTSYLQSVVSPCGLITSGGAAAAMSSPPQYYTIYPQDNSAFVVDGCLSFDTGGNHRPNPDVDQLEE TQLEEEACACFSPPPYEELYSPLR
#67	ACACGAATGGTAGATACAGTG
#68	ATACTTGTGAGCTGTTCCATG
#69	ACTGTTACCTTGCATGGACTG
#70	CAATGAGAACACATGGACATG
#71	CCATGAAAGCTCCATGTCTAC
#72	AGAGATGGCACATATTCTGTC
#73	ATCGGCTGAAGTCAAGCATCG
#74	TGGTCAGTGAGGACTCAGCTG
#75	TTTCTCTGCTTGATGCACTTG
#76	GTGAGCACTGGGAAGCAGCTC
#77	GGCAAAATGCTAGAGACGTGAC
#78	AGGTGTCCTTCAGCTGCCAAG
#79	GTTAAGTGCTCTCTGGATTG
#80	ATCCTGATTGCTGTGTGCAAG
#81	CTCTTCTAGCTGGTCAACATC
#82	CCAGCAACAACCTTACGTGGTC
#83	CCTTTATTACCCAATCACTC
#84	agaacagcgagctttgcctccgctcacgcagagcctctccgtggcctccgcaccttgag cattaggccagttctctctctctctctaatccatccgtcacctctcctgtcatccgttc catgccgtgaggtocattcacagaacacatccatggctctcatgctcagtttggtctga gtctcctcaagctgggatcagggcagtgccaggtgtttgggccagacaagcctgtccagg ccttggtggggaggacgcagcattctcctgtttcctgtctcctaagaccaatgcagagg ccatggaagtgcgggttcttcaggggacagttctctagcgtggtccacctctacagggacg

	ggaaggaccagccatttatgcagatgccacagtatcaaggcaggacaaaactggtgaagg attctattgaggagggcgcatctctctgaggtggaaaacattactgtgttgatgctg gcctctatgggtgcaggattagttcccagctcttactaccagaaggccatctgggagctac aggtgtcagcactgggctcagttcctctcatttccatcacgggatatgttgatagagaca tccagctactctgtcagtcctcggttccccgggcccacagcgaagtggaaaggtc cacaaggacaggatttgtccacagactccaggacaaacagagacatgcatggcctgtttg atgtggagatctctctgacgtccaagagaacggcgagcatatcctgttccatgcggc atgtcatctgagccgagaggtggaatccagggtacagataggagatacctttttcgagc ctatatcgtggcacctggtaccaaagtactgggaatactctgctgtggcctattttttg gcatgtgttgactgaagattttcttctccaaattccagtgtaagcgagagagagaagcat gggcccgtgccttattcatggttccagcagggacaggatcagagatgctcccatccag ctgcttctcttctctagtccctagcctccagggggccaggcccaaaaaaggaaaatccag gcggaactggactggagaagaaagcacggacaggcagaattgagagacgcccggaaacac gcagtgagggtgactctggatccagagacggctcaccgaagctctgcgtttctgatctg aaaactgtaacccatagaaaagctcccaggaggtgcctcactctgagaagagatttaca aggaagagtgtggtggcttctcagagttccaagcagggaacattactgggaggtggac ggaggacacaataaaaaggtggcgctgggagtggtccgggatgatgtggacaggaggag gagtacgtgactttgtctccgatcatgggtactgggtcctcagactgaatggagaacat ttgtatttcacattaaatcccgttttatcagcgtcttcccaggaccccacctacaaa ataggggtcttctggactatgagtgtgggaccatctccttctcaacataaatgaccag tccctattttataccctgacatgtcggtttgaaggcttattgaggccctacattgagtat ccgtctataatgagcaaaatggaactcccatagtcctgcccagtcacccaggaaatca gagaaagaggcctcttggcaaaggcctctgcaatcccagagacaagcaacagtgtgctc tctcacaggcaaccacgccttctcctcccagggtgaaatgtaggatgaatcacatccc acattcttctttaggatattaaggtctctctcccagatccaaagtccgcagcagccgg ccaagtggttccagatgaaggggactggcctgtccacatgggagtcagggtgtcatgg ctgcccgtgagctgggagggaagaaggtgacattacatttagtttgcctcactccatct ggctaagtgtcttgaataaccacctctcaggtgaagaaccgtcagggaattcccatctca caggctgtggtgtagattaagtagacaaggaatgtgaataatgcttagatcttatgatg acagagtgtatcctaaggtttgttcattatattacactttcagtaaaaaaaaaaaaaa aaaaa
#85	maimlslvlslklgsgqvwqvfpgdpkvqalvgedaafscflspktnaeamevrffrgqf ssvvhllyrdgkdqpfmqmpqyqgrtklvkdsiaegrislrlnitvldaglygcrissqs yyqkaiwelqvsalgsvglisityvdrdiqlcsgswfprptakwkpggqdlstdsr tnrdmhglfdveisltvqenagsiscsmrhahlsrevesrvqigdtffepiswhlatkvl gilccqlffgivglkiffskfqkrereawagalfmvpagtgsemplphpaaslllvlasr gpgpkkenpggtglekkartgrierrpetrsggdsgrdgspealrf
#86	ATTCATGGTTCAGCAGGGAC
#87	GGGAGACAAAGTCACGTACTC

#88	TCCTGGTGTTCTGGTCTGCTT
#89	GAGAGTCCTGGCTTTTGTGGGC
#90	GSSDLTWPPAIKLG
#91	DRYVAVRHPLRARGLR
#92	VAPRAKAHKSQDSL
#93	CFRSTRHNFSMR
#94	MNGTYNTCGSSDLTWPPAIKLG
#95	RDTSDTPLCQLSQG
#96	GIQEGGF CFRSTRHNFSMRFP
#97	AKEFQEASALAVAPRAKAHKSQDSLCLVTLA
#98	TCCTGCTCGTCGCTCTCTGAT
#99	TCGCTTTTGTCTGATTTGC
#100	HNGSYEISVLMGNS
#101	NLPTPPTVENQQLA

#102	RKYRKDYELRQKKWSHIPPENIFPLETNETNHVSLKIDDDKRRDTIQRLRQCKYDKKRVLKDLKHNDGN FTEKQKIELNKLQIDYYNLTKFYGTVKLDTMIFGVIEYCERGSRLREVLNDTISYPDGTGFMDFEKFISVL YDIAGKMSYLHSSKTEVHGRLKSTNCVVDSRMVVKITDFGCNSILPPKKDLWTAPHELRQANISQKGDVY SYGIIAQEII LRKETFYTLSCRDRNEKIFRVENSNGMKPFRPDLFLETAEEKELEVYLLVKNCWEDPEK RPDFKKIETTLAKIFGLFHDQKNESYMDTLIRRLQLYSRNLEHLVEERTQLYKAERDRADRLNFMLLPRL VVKSLKEKGFVEPELYEEVTIYFSDIVGFTTICKYSTPMEVVDMLNDIYKSFHDHVDHHDVYKVTIGDA YMVASGLPKRNGNRHAIDIAKMALEILSFMGTFELEHLPLPIWIRIGVHSGPCAAGVVGIKMPRYCLFG DTVNTASRMESTGLPLRIHVSGSTIALKRTECQFLYEVGRGETYLKGRGNETTYWLTGMKDQKFNLPPTP TVENQQRLOAEFSDMIANSLOKROAAGIRSQKPRRVASYKKGTLLEYLQLNTTDKESTYF
#103	GCTGGTAACATCTTCCTGC
#104	GAAGAATGTTGTCCAGAGGT
#105	LINKVPLPVDKLAPL
#106	SEAVKKLLEALSHLV
#107	TGTTTTCACTACCAGGGGC
#108	TGTTGGCTTTGGCAGAGTCC
#109	GAGGCAGAGTTCAGGCTTCACCGA
#110	TGTTGGCTTTGGCAGAGTCC
#111	TGMDMWSTQDLYDNPVTSVFQYEGLRWSCVRQSSGFTECRPYFTILGLPAMLQAVR
#112	DQWSTQDLYNNPVTAVFNYQGLWRSCVRESSGFTECRGYFTLL GLPAMLQAVR
#113	STQDLYNNPVTAVF
#114	DMWSTQDLYDNP
#115	CRPYFTILGLFA
#116	TNFWMSTANMYTG
#117	gccaggatca tgtccaccac cacatgccaa gtggtggcgt tcctcctgtc catcctgggg ctggccggct gcatcgccgc caccgggatg gacatgtgga gcaccagga cctgtacgac aaccctgca cctccgtgtt ccagtacgaa gggctctgga ggagctgcgt gaggcagagt tcaggcttca ccgaatgcag gccctatttc accatcctgg gacttccagc catgctgcag gcagtgcgag cctgatgat cgtaggcatc gtctgggtg ccattggcct cctggtatcc atctttgccc tgaaatgcat ccgcatggc agcatggagg actctgcaa agccaacatg acactgacct ccgggatcat gttcatgtc tcaggcttt gtgcaattgc tggagtgtct gtgtttgcca acatgctggt gactaacttc tggatgtcca cagctaacat gtacaccggc atgggtgga tgggtgcagac tgttcagacc aggtacacat ttggtgcggc tctgttcgtg ggctgggtcg ctggaggcct cacactaatt ggggtgtga tgatgtgcat cgctgccc ggcctggcac cagaagaaac caactacaaa gccgtttctt atcatgcctc aggccacagt gttgccata agcctggagg cttcaaggcc agcactggct ttgggtcaa caccaaaaac aagaagatat acgatggagg tgccgcaca gaggacgagg tacaatetta tccttccaag cacgactatg tgtaatgtc taagacctct cagcac
#118	MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVF QYEGLRWSCVRQSSGFTECRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALK CIRIGSMEDSAKANMTLTSGIMFIVSGLCAIAGVSFANMLVTNFWMSTANMYTGMGG MVQTVQTRYTGAALFVGWVAGGLTLIGGVMCIACRGLAPEETNYKAVSYHASGHSV AYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKRDYV
#119	gccaggatca tgtccaccac cacatgccaa gtggtggcgt tcctcctgtc catcctgggg ctggccggct gcatcgccgc caccgggatg gacatgtgga gcaccagga cctgtacgac aaccctgca cctccgtgtt ccagtacgaa gggctctgga ggagctgcgt gaggcagagt tcaggcttca ccgaatgcag gccctatttc accatcctgg gacttcc
#120	MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRWSCVRQSSGFTECRPYFTI
#121	AATGAGAGGAAAGAGAAAAC
#122	ATGGTAGAAGAGTAGGCAAT
#123	EKWNLRKRIALKMVC
#124	CLGFNFKEMFK
#125	TAATGATGAACCTACACTGAGC

#126	ATGGACAAATGCCCTACCTT
#127	AGTGCTGGAAGGATGTGCGTGT
#128	TTGAGGTGGTTGTTGGGTTT
#129	AGATGTGCTGAGGCTGTAGA
#130	ATGAAGGTTGATTATTTGAG
#131	AGCCGCATACTCCCTTACCCTCT
#132	GCAGCAGCCCAAACACCACA
#133	CTGAGCCGAGAGGTGGAATC
#134	CTCTCTCGCTTACACTGGAA
#135	QWQVFGPDKPVQAL
#136	AKWKGPQGQDLSTDS
#137	NMLVTNFWMSTANMYTGMGGMVQTVQTRYTFG

图 41

专利名称(译)	肿瘤内差异表达的基因产物及其应用		
公开(公告)号	CN101966337A	公开(公告)日	2011-02-09
申请号	CN201010279118.4	申请日	2003-11-21
[标]申请(专利权)人(译)	加尼梅德医药品有限公司		
申请(专利权)人(译)	加尼梅德医药品有限公司		
当前申请(专利权)人(译)	加尼梅德医药品有限公司		
[标]发明人	U萨海恩 O图雷茨 M柯斯勒司克		
发明人	U·萨海恩 O·图雷茨 M·柯斯勒司克		
IPC分类号	C12N15/12 A61P35/00 G01N33/574 G01N33/53 G01N33/535 C12N15/11 C12N5/10 C07K14/47 A61K45/00 G01N33/534 A61K38/17 A61K39/39 G01N33/577 A61K48/00 A61K39/395 C07K16/30 C12Q1/68 C12N15/63 A61K39/00		
CPC分类号	A61K47/48569 C07K14/47 C07K2317/34 C07K16/30 C07K16/3023 A61K2039/505 C07K16/3046 C07K16/18 G01N33/57484 A61K39/0011 A61K39/395 A61K48/005 C12N2503/02 C12Q1/6886 C12Q2600/136 C12Q2600/158 G01N33/5011 G01N33/5017 G01N33/57407 G01N2500/10 A61P1/04 A61P1/18 A61P11/00 A61P13/08 A61P13/12 A61P15/00 A61P17/00 A61K39/39558 A61P35/00 C07K2317/30 A61K47/549 C07K16/3076		
优先权	10254601 2002-11-22 DE		
外部链接	Espacenet SIPO		

摘要(译)

本发明涉及与肿瘤相关的基因产物的鉴别，涉及编码所述产物的核酸。本发明也涉及如下疾病的治疗和诊断，所述疾病中基因产物的异常表达与肿瘤相关，蛋白质、多肽和肽的表达与肿瘤相关。本发明还涉及编码所述多肽、肽和蛋白质的核酸。

正常组织	表达	肿瘤类型	表达
脑	+	结肠癌	+++
小脑		胰腺癌	-
心肌		食道癌	-
骨骼肌	-	胃癌	+++
心肌		支气管癌	-
胃	+++	乳腺癌	++
结肠	+++	卵巢癌	+
胰腺	-	子宫内膜癌	
肾	-	耳鼻喉肿瘤	
肝	+	肾细胞癌	
睾丸	++	前列腺癌	+
胸腺	+	肝癌	+
乳腺	-		
卵巢	+		
子宫	+		
皮肤			
脾	+		
甲状腺			
淋巴结	-		
脾	+		
外周血单核细胞	-		
前列腺	-		