



(12) 发明专利申请

(10) 申请公布号 CN 102459646 A

(43) 申请公布日 2012.05.16

(21) 申请号 201080031553.3

G01N 33/53(2006.01)

(22) 申请日 2010.05.14

G01N 33/50(2006.01)

(30) 优先权数据

G01N 33/68(2006.01)

577012 2009.05.15 NZ

(85) PCT申请进入国家阶段日

2012.01.13

(86) PCT申请的申请数据

PCT/NZ2010/000089 2010.05.14

(87) PCT申请的公布数据

W02010/131984 EN 2010.11.18

(71) 申请人 环太平洋生物技术有限公司

地址 新西兰达尼丁

(72) 发明人 P·J·吉尔福德

(74) 专利代理机构 北京市中咨律师事务所

11247

代理人 柴云峰 黄草生

(51) Int. Cl.

C12Q 1/68(2006.01)

权利要求书 2 页 说明书 21 页

序列表 59 页 附图 3 页

(54) 发明名称

胃癌的检测标志物

(57) 摘要

肿瘤的早期检测是肿瘤(包括胃肿瘤)患者存活的主要决定因素。GTM 基因家族的成员可以在胃肿瘤组织中差异地表达,并且因此可以用作检测胃癌和其他类型癌症的标志物。本发明提供了用于检测肿瘤(包括胃肿瘤)的新 GTM,并且特别是人酶原粒蛋白 16(ZG16)。GTM 可以单独使用或与其他已知 GTM 一起使用,来提供用于肿瘤(包括胃肿瘤)检测的新标记。

1. 一种用于检测胃癌的方法,包括:
 - (i) 提供生物样品;和
 - (ii) 检测所述样品中人酶原粒蛋白 16(“ZG16”)的水平。
2. 权利要求 1 的方法,其中生物样品中 ZG16 的过表达表明癌症。
3. 权利要求 1 或权利要求 2 的方法,包括检测一种或多种其他 GTM 家族成员的水平。
4. 权利要求 1 至 3 任一项的方法,其中所述方法进一步包括检测选自黏蛋白 5AC(“MUC5AC”)和黏蛋白 17(“MUC17”)的其他 GTM 家族成员的水平。
5. 权利要求 1 至 4 任一项的方法,其中所述一种或多种其他 GTM 家族成员选自由 MUC5AC、MUC17、ZG16、羧肽酶 N、多肽 2,83kDa 链(“CPN2”)、基质金属蛋白酶 12(“MMP12”)、抑制素(“INHBA”)、胰岛素样生长因子 7(“IGFBP7”)、 γ -谷氨酰水解酶(“GGH”)、富含亮氨酸脯氨酸的蛋白聚糖(“LEPRE1”)、半胱氨酸蛋白酶抑制剂 S(“CST4”)、分泌型卷曲相关蛋白 4(“SFRP4”)、asporin(“ASPN”)、具有 EF 手型结构域的细胞生长调节剂 1(“CGREF1”)、激肽释放酶 10(KLK10)、金属蛋白酶的抑制剂 1(“TIMP1”)、分泌型酸性富含半胱氨酸蛋白(“SPARC”)、转化生长因子,13-诱导的(“TGFBI”)、含有 EGF 的 fibulin-样胞外基质蛋白 2(“EFEMP2”)、基膜聚糖(lumican)(“LUM”)、stannin(“SNN”)、分泌型磷蛋白 1(“SPP1”)、硫酸软骨素蛋白聚糖 2(“CSPG2”)、N-酰基鞘氨醇酰胺水解酶(“ASAHI”)、丝氨酸蛋白酶 11(“PRSS11”)、分泌型卷曲相关蛋白 2(“SFRP2”)、磷脂酶 A2, XIIB 组(“PLA2G12B”)、spondin 2,胞外基质蛋白(“SPON2”)、嗅介蛋白(olfactomedin)1(“OLFM1”)、含有血小板反应蛋白重复 1(“TSRC1”)、血小板反应蛋白 2(“THBS2”)、adlican、半胱氨酸蛋白酶抑制剂 SA(“CST2”)、半胱氨酸蛋白酶抑制剂 SN(“CST1”)、赖氨酰氧化酶样酶 2(“LOXL2”)、甲状腺球蛋白(“TG”)、转化生长因子 β 1(“TGFB1”)、丝氨酸或半胱氨酸蛋白酶抑制剂 Clade H 成员 1(“SERPINH1”)、丝氨酸或半胱氨酸蛋白酶抑制剂 Clade B 成员 5(“SERPINB5”)、基质金属蛋白酶 2(“MMP2”)、前蛋白转化酶枯草杆菌蛋白酶/kexin 5 型(“PCSK5”)、乙酰透明质酸糖蛋白连接蛋白 4(“HAPLN4”)、CA19-9、CA72-4、胃蛋白酶原和 CEA 组成的组。
6. 权利要求 1 至 5 任一项的方法,其中所测试的 GTM 标志物包括 MUC5AC、MUC17、ZG16、半胱氨酸蛋白酶抑制剂 SN、丝氨酸蛋白酶抑制蛋白 H1 和丝氨酸蛋白酶抑制蛋白 B5。
7. 权利要求 1 至 6 任一项的方法,其中通过检测 GTM mRNA 的水平进行所述检测步骤。
8. 权利要求 1 至 6 任一项的方法,其中通过检测 GTM cDNA 的水平进行所述检测步骤。
9. 权利要求 8 的方法,其中使用与所述 GTM cDNA 的至少一部分互补的寡核苷酸进行所述检测步骤。
10. 权利要求 8 的方法,其中使用正向引物和反向引物用 qRT-PCR 方法进行所述检测步骤。
11. 权利要求 1 至 6 任一项的方法,其中通过检测 GTM 蛋白的水平进行所述检测步骤。
12. 权利要求 1 至 6 任一项的方法,其中通过检测 GTM 肽的水平进行所述检测步骤。
13. 权利要求 11 或 12 的方法,其中使用针对所述 GTM 的抗体进行所述检测步骤。
14. 权利要求 11 至 13 任一项的方法,其中使用夹层型免疫测定方法,或使用抗体芯片进行所述检测步骤。
15. 权利要求 11 至 13 任一项的方法,其中所述抗体是单克隆抗体。
16. 权利要求 11 至 14 任一项的方法,其中所述抗体是多克隆抗血清。

17. 一种用于检测 GTM 的装置,其包括:
基质,其上具有 GTM 捕获试剂;和
与所述基质相连的检测仪,所述检测仪能够检测与所述捕获试剂结合的 GTM,其中所述 GTM 包括人酶原粒蛋白 16(“ZG16”)。
18. 权利要求 17 的装置,其中所述 GTM 捕获试剂是寡核苷酸。
19. 权利要求 17 的装置,其中所述 GTM 捕获试剂是对 GTM 寡核苷酸、GTM 蛋白或 GTM 肽特异的抗体。
20. 一种用于检测癌症的试剂盒,其包含:
基质,其上具有 GTM 捕获试剂;
用于显示所述 GTM 捕获试剂和 GTM 复合物的工具;
试剂;和
使用说明书,其中所述 GTM 包括人酶原粒蛋白 16(“ZG16”)。
21. 权利要求 20 的试剂盒,其中所述 GTM 捕获试剂是 GTM- 特异性寡核苷酸。
22. 权利要求 20 的试剂盒,其中所述 GTM 捕获试剂是对 GTM 寡核苷酸、GTM 蛋白或 GTM 肽选择性的 GTM 特异性抗体。
23. 一种检测胃癌的方法,包括步骤:
提供来自处于患胃癌风险中的患者的测试样品;测量所述测试样品中 GTM 蛋白的存在;和
将所述测试样品中存在的 GTM 的量与获自不患胃癌受试者的对照样品的值进行比较,其中所述 GTM 包含人酶原粒蛋白 16(“ZG16”)。
24. 一种用于筛选胃癌的方法,包括步骤:
提供来自测试受试者的测试样品;
测量所述测试样品中 GTM 的存在;和
将所述测试样品中存在的 GTM 的量与获自不患胃癌受试者的对照样品的值进行比较,其中所述 GTM 包括人酶原粒蛋白 16(“ZG16”)。
25. 权利要求 24 的方法,其中所述 GTM 是 GTM 蛋白或肽。
26. 权利要求 24 的方法,其中所述 GTM 是 GTM 特异性的寡核苷酸。
27. 权利要求 26 的方法,其中所述寡核苷酸是 DNA。
28. 权利要求 26 的方法,其中所述寡核苷酸是 RNA。
29. 权利要求 24 至 28 任一项的方法,其中所述测量步骤使用 ELISA 测定法。
30. 权利要求 24 至 29 任一项的方法,其中所述测试样品获自血浆。
31. 权利要求 24 至 30 任一项的方法,其中所述测试样品获自组织、尿液、胃液、血清和粪便。

胃癌的检测标志物

发明领域

[0001] 本发明涉及癌症的检测。具体地,本发明涉及遗传和 / 或蛋白质标志物用于检测癌症的用途,并且更特别地,涉及遗传和 / 或蛋白质标志物用于检测胃癌的用途。

[0002] 背景

[0003] 在早期检测到癌症并得到治疗时,癌症患者的存活得到很大程度的提高。在胃癌的情况中,与诊断为晚期疾病患者大约 10% 的 5- 年存活率相比,诊断为早期疾病的患者具有 90% 的 5- 年存活率。然而,大多数胃癌患者通常患有晚期疾病。因此,导致胃癌早期诊断的发展可以导致改善的患者预后。

[0004] 生物样品中特定癌症相关标志物的鉴定可以提供癌症早期诊断的有价值的方法,导致早期治疗和提高的预后,所述生物样品包括体液,例如,血液、尿液、腹膜冲洗液和粪便提取物。特定癌症标志物还可以提供监测疾病进展的手段,能够追踪外科手术治疗、放疗和化疗的疗效。然而,对于许多主要的癌症,可用的标志物存在灵敏度和特异性不足的问题。例如,对于胃癌最常用的标志物 ca19-9、ca72-4 和癌胚抗原 (CEA) 只能检测出约 15-50% 的任一阶段的胃肿瘤,对于早期疾病,降至大约 2-11%。因此,存在非常高频率的可导致患者和保健医生认为不存在疾病的假阴性测试,然而实际上,患者患有需要立即注意的严重癌症。此外,这些标志物在高达 1/3 受到良性胃病影响的个体中产生假阳性信号。

[0005] 发明概述

[0006] 本发明的几个方面提供了可以提供早期癌症检测并且降低假阳性和假阴性测试结果频率的方法、组合物和装置。

[0007] 在特定的实施方案中,分子分析可以用于鉴定在胃肿瘤组织中与非恶性胃组织相比高表达,但不必定是过表达的基因。这样的分析包括微阵列和定量聚合酶链式反应 (qPCR) 方法。癌症基因、RNA 以及由那些基因编码的蛋白质在此称为胃肿瘤标志物 (GTM)。将理解术语 GTM 不需要标志物只对胃肿瘤是特异性的。而是, GTM 的表达在其他类型的肿瘤中得到提高,所述肿瘤包括恶性或非恶性肿瘤,包括胃癌、膀胱癌、结直肠癌、胰腺癌、卵巢癌、皮肤癌 (例如,黑素瘤)、肝癌、食道癌、子宫内膜癌和脑癌。然而,应当理解,术语 GTM 不包括现有技术的标志物,如 CA19-9、CA72-4、胃蛋白酶原和 CEA,或任何其他之前已经鉴定为表示胃肿瘤的标志物。从肿瘤中分泌或逸出一些 GTM,其水平足以以高度可靠性来诊断胃癌,并且在其他情况中,两种或多种 GTM 的测量可以提供可靠的胃癌诊断。

[0008] 由细胞分泌或裂解出来的蛋白质,单独或互相结合,具有作为用于胃癌诊断的血清或体液标志物或作为用于监测已确定疾病进展标志物的效用。可以使用本领域已知的方法,并且包括使用单克隆抗体、多克隆抗血清等,来进行蛋白质标志物的检测。

[0009] 具体地,本发明提供了一种用于检测胃癌的方法,包括:

[0010] (i) 提供生物样品 ;和

[0011] (ii) 检测所述样品中人酶原粒蛋白 16 (“ZG16”) 的水平。

[0012] 在一个方面中,患者中 ZG16 的过表达表示患者患有胃癌。

[0013] 根据本发明的其他 GTM 家族成员可以选自黏蛋白 5AC (“MUC5AC”) 或黏蛋白

17(“MUC17”)。该方法可以涉及 ZG16 和 MUC5AC、ZG16 和 MUC17,或 ZG16 和 MUC5AC 和 MUC17 的检测。

[0014] 所述其他 GTM 家族成员还可以包含一个或多个更多的 GTM 家族成员,例如, MUC5AC、MUC17、ZG16、羧肽酶 N、多肽 2, 83kDa 链 (CPN2)、基质金属蛋白酶 12 (MMP12)、抑制素 (“INHBA”)、胰岛素样生长因子 7 (“IGFBP7”)、 γ -谷氨酰水解酶 (“GGH”)、富含亮氨酸脯氨酸的蛋白聚糖 (“LEPRE1”)、半胱氨酸蛋白酶抑制剂 (cystatin)S (“CST4”)、分泌型卷曲相关蛋白 4 (“SFRP4”)、asporin (“ASPN”)、具有 EF 手型结构域的细胞生长调节剂 1 (“CGREF1”)、激肽释放酶 10 (KLK10)、金属蛋白酶的抑制剂 1 (“TIMP1”)、分泌型酸性富含半胱氨酸蛋白 (“SPARC”)、转化生长因子, 13- 诱导的 (“TGFB1”)、含有 EGF 的 fibulin- 样胞外基质蛋白 2 (“EFEMP2”)、基膜聚糖 (lumican) (“LUM”)、stannin (“SNN”)、分泌型磷蛋白 1 (“SPP1”)、硫酸软骨素蛋白聚糖 2 (“CSPG2”)、N- 酰基鞘氨醇酰胺水解酶 (“ASAHI”)、丝氨酸蛋白酶 11 (“PRSS11”)、分泌型卷曲相关蛋白 2 (“SFRP2”)、磷脂酶 A2, XIIIB 组 (“PLA2G12B”)、spondin2, 胞外基质蛋白 (“SPON2”)、嗅介蛋白 (olfactomedin) 1 (“OLFM1”)、含有血小板反应蛋白重复 1 (“TSRC1”)、血小板反应蛋白 2 (“THBS2”)、adlcan、半胱氨酸蛋白酶抑制剂 SA (“CST2”)、半胱氨酸蛋白酶抑制剂 SN (“CST1”)、赖氨酰氧化酶样酶 2 (“LOXL2”)、甲状腺球蛋白 (“TG”)、转化生长因子 β 1 (“TGFB1”)、丝氨酸或半胱氨酸蛋白酶抑制剂 Clade H 成员 1 (“SERPINH1”)、丝氨酸或半胱氨酸蛋白酶抑制剂 Clade B 成员 5 (“SERPINB5”)、基质金属蛋白酶 2 (“MMP2”)、前蛋白 (proprotein) 转化酶枯草杆菌蛋白酶/kexin 5 型 (“PCSK5”)、乙酰透明质酸 (hyaluronan) 糖蛋白连接蛋白 4 (“HAPLN4”)、CA19-9、CA72-4、胃蛋白酶原、CEA、MUC5AC 和 MUC17。

[0015] 根据本发明的组合 GTM 标志物的一个实例是 MUC5AC、MUC17、ZG16、半胱氨酸蛋白酶抑制剂 SN、serpinH1 和 serpinB5。

[0016] 可以使用任何合适的用于检测 GTM 水平的方法,并且可以包括检测 GTM mRNA、GTM cDNA 的水平,使用与所述 GTM cDNA 的至少一部分互补的寡核苷酸,使用正向引物和反向引物的 qRT-PCR 方法,检测 GTM 蛋白的水平,检测 GTM 肽的水平,例如,使用针对所述 GTM 的抗体。可以使用任何合适的抗体,并且可以是单克隆抗体或多克隆抗血清。可以使用夹层型免疫试验方法或使用抗体芯片来进行该方法。

[0017] 本发明还提供了一种用于检测 GTM 的装置,其包括:其上具有 GTM 捕获试剂的基质;和与所述基质相连的检测仪,所述检测仪能够检测与所述捕获试剂相结合的 GTM。

[0018] GTM 捕获试剂可以是 GTM 寡核苷酸、GTM 蛋白或 GTM 肽特异性的寡核苷酸或抗体。

[0019] 本发明的再一个方面是用于检测癌症的试剂盒,其包括:

[0020] 其上具有 GTM 捕获试剂的基质;

[0021] 用于显现所述 GTM 捕获试剂和 GTM 的复合物的工具;和

[0022] 使用说明书,其中所述 GTM 包含人酶原蛋白 16 (“ZG16”)。

[0023] GTM 捕获试剂是对 GTM 寡核苷酸、GTM 蛋白或 GTM 肽有选择性的 GTM 特异性寡核苷酸或 GTM 特异性抗体。

[0024] 本发明还提供了用于检测胃癌的方法,其包括步骤:

[0025] 提供来自处于患有胃癌风险中的患者的测试样品;测量所述测试样品中 GTM 蛋白的存在;和

[0026] 将所述测试样品中存在的 GTM 的量与获自未患有胃癌受试者的对照样品的值进行比较,其中所述 GTM 包含人酶原粒蛋白 16(“ZG16”)。

[0027] 在进一步的方面中,本发明提供了一种用于筛选胃癌的方法,其包括步骤:

[0028] 提供来自测试受试者的测试样品;

[0029] 测量所述测试样品中 GTM 的存在;和

[0030] 将所述测试样品中存在的 GTM 的量与获自未患有胃癌受试者的对照样品的值进行比较,其中所述 GTM 包含人酶原粒蛋白 16(“ZG16”)。

[0031] GTM 可以是 GTM 蛋白或肽,或 GTM 特异性的寡核苷酸。寡核苷酸可以是 DNA 或 RNA。

[0032] 根据该方法,测量步骤可以使用 ELISA 试验。

[0033] 测试样品可以获自血浆、组织、尿液、胃液、血清和粪便。

[0034] 附图简述

[0035] 参照其特定实施方案和参照附图来描述本发明,其中:

[0036] 图 1 描绘了微阵列分析表,显示了在肿瘤组织中相对高表达的基因。将肿瘤组织和非恶性组织中每个基因的信号强度进行了分级。该表显示了比现有胃癌标志物 CEA(由基因 CEACAM5 编码)具有更高等级的 GTM。

[0037] 图 2 描绘了显示出抗体阵列分析中所用的血清样品特征的表。

[0038] 图 3 描绘了直方图,其显示了肿瘤和非恶性样品根据其 (a) ZG16 和 (b) MUC17 表达水平的分布。使用 RT-qPCR 获得了两个基因的表达水平。

[0039] 图 4 描绘了 boxplot,显示了使用抗体阵列和 RCA 检测的胃癌患者和对照的血清中 (a) MUC17 和 (b) ZG16 的检测。

[0040] 详述

[0041] 定义

[0042] 在详细描述本发明的实施方案之前,提供本文中术语的一些定义将是有益的。

[0043] 术语“GTM”或“胃肿瘤标志物”或“GTM 家族成员”意思是与胃癌相关或另外鉴定与胃癌相关分子的基因、基因片段、RNA、RNA 片段、蛋白质或蛋白质片段。作为本发明的一部分公开的 GTM 不包括现有技术中已知的与胃癌相关的分子,例如, CA19-9、CA72-4、胃蛋白酶原和 CEA。然而,本发明的标志物可以与之前公开的 GTM 以新的和创造性的组合来一起使用。

[0044] 术语“标志物”指与生物学现象的存在定量或定性相关的分子。“标志物”的实例包括多核苷酸,如基因或基因片段、RNA 或 RNA 片段;或基因产物,包括多肽,如肽、寡肽、蛋白质或蛋白质片段;或任何相关的代谢产物、副产物或任何其他鉴定分子,如抗体或抗体片段,不管是否与现象下的机理直接或间接相关。本发明的标志物包括如本文所公开的核苷酸序列(例如,GenBank 序列),特别是全长序列、其任何编码序列、任何片段或任何互补物,以及上述定义的其任何可测量标志物。

[0045] 如本文所用,“抗体”及类似术语指免疫球蛋白分子和免疫球蛋白(Ig)分子的免疫活性部分,即含有特异结合抗原(与之发生免疫反应)的抗原结合位点的分子。这些包括但不限于多克隆、单克隆、嵌合、单链、Fc、Fab、Fab' 和 Fab₂ 片段,以及 Fab 表达文库。抗体分子涉及 IgG、IgM、IgA、IgE 和 IgD 的任何一类,这些类别根据分子中存在的重链的性质而

彼此不同。这些还包括亚类,如 IgG1、IgG2 等。轻链可以是 κ 链或 λ 链。本文提及的抗体包括所有类、亚类和类型。还包括嵌合抗体,例如,对超过一种的来源(例如,小鼠或人序列)是特异性的单克隆抗体或其片段。还包括驼类抗体 (camelid antibody)、鲨鱼抗体或纳米抗体。

[0046] 术语“癌症”和“癌性的”是指或者描述了哺乳动物中通常以异常或失调的细胞生长为特征的生理状态。癌症和癌症病理可以与例如转移、干扰邻近细胞的正常机能、以异常水平释放细胞因子或其他分泌产物、抑制或加重炎性或免疫应答、瘤形成、癌前病变、恶性肿瘤、周围或远距离组织或器官如淋巴结的侵入等相关。特别包括的是黑素瘤。

[0047] 术语“肿瘤”指的是所有肿瘤细胞生长和增殖,不管恶性还是良性的,以及所有癌前和癌细胞和组织。

[0048] 术语“胃癌”指的是源自胃的肿瘤。这些肿瘤能够转移至任何器官。

[0049] 术语“差异地表达”“差异的表达”及相似短语是指相对于在对照受试者(例如,参照样品)中的表达,在患有病症(特别是癌症,如黑素瘤)的受试者(例如,测试样品)中的表达被激活至较高或较低水平的基因标志物。该术语还包括其表达在如下情况被激活至较高或较低的水平标志物:在相同疾病的不同阶段;在具有好或差的预后的疾病中;或在具有较高或较低增殖水平的细胞中。差异表达的标志物可以在多核苷酸水平或多肽水平被激活或抑制,或可以接受选择性剪接而导致不同的多肽产物。例如,这些差异可以通过 mRNA 水平、表面表达、多肽分泌或其他分配上的变化而得到证明。

[0050] 差异表达可以包括比较两个或多个标志物(例如,基因或其基因产物)间的表达;或比较两个或多个标志物(例如,基因或其基因产物)间的表达比率;或比较相同标志物的两种不同加工的产物(例如,转录产物或多肽),其在正常受试者和患病受试者之间不同;或在相同疾病的不同阶段之间不同;或在具有好或差的预后的疾病之间不同;或在具有较高和较低增殖水平的细胞之间不同;或在正常组织和患病组织(特别是癌症或黑素瘤)之间不同。差异表达包括基因或其表达产物在例如正常和患病细胞中、或在经历不同疾病事件或疾病阶段的细胞中、或在具有不同增殖水平的细胞中,在时间或细胞表达模式上的定量以及定性差异。

[0051] 术语“表达”包括多核苷酸和多肽的产生,特别地,从基因或基因的一部分产生 RNA(例如, mRNA),且包括 RNA 或基因或基因的一部分编码的多肽的产生,以及与表达相关的可检测物质的出现。例如,例如从多肽-多肽相互作用、多肽-核苷酸相互作用等形成复合物,也包括在术语“表达”的范围内。另一个实例是结合配体,如杂交探针或抗体,与基因或其他多核苷酸或寡核苷酸、多肽或蛋白质片段的结合以及结合配体的显影。因此,微阵列、杂交印迹(如 Northern 印迹),或免疫印迹(如 Western 印迹),或珠阵列上的,或通过 PCR 分析的斑点强度,包括在生物分子下的术语“表达”的范围内。

[0052] 术语“表达阈值”和“限定的表达阈值”可以互换使用,并且指的是正谈论的标志物的水平,在该水平外,多核苷酸或多肽作为患者存活的预测标志物。阈值将取决于所建立的预测模型,所述模型从临床研究(如以下实施例所述的那些)实验性地产生。根据所用的预测模型,设定表达阈值以获得最大的灵敏度,或最大的特异性,或最小的误差(最大的分类级别)。例如,设定较高的阈值,以获得最小误差,但这可能导致较低的灵敏度。因此,对于任何给定的预测模型,将使用临床研究来设定通常获得最高灵敏度同时具有最小误差

率的表达阈值。对于任何情况的表达阈值的确定在本领域技术人员知识范围内。

[0053] 术语“灵敏度”意思是（通过模型）测试的患病个体为阳性的比例。因此，提高的灵敏度意味着较低的假阴性测试结果。

[0054] 术语“特异性”意思是（通过模型）测试未患病个体为阴性的比例。因此，提高的特异性意味着较低的假阳性测试结果。

[0055] 术语“微阵列”指的是捕获剂（优选，多核苷酸（例如，探针）或多肽）在基质上的有序或无序排列。参见，例如，Microarray Analysis（微阵列分析），M. Schena, John Wiley & Sons, 2002；Microarray Biochip Technology（微阵列生物芯片技术），M. Schena 编辑，Eaton Publishing, 2000；Guide to Analysis of DNA Microarray Data（DNA 微阵列数据分析指导），S. Knudsen, John Wiley & Sons, 2004 和 Protein Microarray Technology, D. Kambhampati 编辑，John Wiley & Sons, 2004。

[0056] 术语“寡核苷酸”指的是多核苷酸，通常是探针或引物，包括，但不限于，单链脱氧核糖核苷酸、单链或双链核糖核苷酸，RNA:DNA 杂合体及双链 DNA。寡核苷酸，如单链 DNA 探针寡核苷酸，经常利用化学方法来合成，例如，使用可购得的自动寡核苷酸合成仪，或通过多种其他方法，包括体外表达系统、重组技术以及在细胞和生物中的表达。

[0057] 术语“过表达”或“过表达的”指的是患者中高于正常组织中所看到的基因或标志物的表达水平。如果是正常组织中表达的 1.2、1.3、1.4、1.5、1.6、1.7、1.8、1.9、2 倍或高于 2 倍，则认为表达是过表达的。

[0058] 术语“多核苷酸”，以单数或复数使用时，通常指的是任何多核糖核苷酸或多脱氧核糖核苷酸，其可以是未修饰的 RNA 或 DNA 或修饰的 RNA 或 DNA。这包括，但不限于，单链和双链 DNA、包括单链和双链区域的 DNA、单链和双链 RNA，以及包括单链和双链区域的 RNA，包含可以是单链或更一般地是双链或包括单链和双链区域的 DNA 和 RNA 的杂合分子。还包括包含 RNA 或 DNA 或既包含 RNA 又包含 DNA 的三链区域。特别包括 mRNA、cDNA 和基因组 DNA，及其任何片段。该术语包括包含一个或多个修饰碱基（如，氟化碱基或稀有碱基，如肌苷）的 DNA 和 RNA。本发明的多核苷酸可以包括编码或非编码序列，或正义或反义序列。将理解本文中每次涉及“多核苷酸”或类似术语，将包括全长序列及其任何片段、衍生物或变体。

[0059] 如本文所用，“多肽”指的是寡肽、肽或蛋白质序列，或其片段，以及天然产生的、重组的、合成的或半合成的分子。其中本文引用的“多肽”指的是天然产生的蛋白质分子的氨基酸序列，“多肽”及类似术语不意味着限定氨基酸序列为全长分子的完整天然氨基酸序列。将理解本文中每次涉及“多肽”或类似术语，将包括全长序列及其任何片段、衍生物或变体。

[0060] 术语“qPCR”或“QPCR”指的是例如 PCR Technique: Quantitative PCR (PCR 技术：定量 PCR)，J. W. Larrick 编辑，Eaton Publishing, 1997 和 A-Z of Quantitative PCR (定量 PCR 的 A-Z)，S. Bustin 编辑，IUL Press, 2004 中所述的定量聚合酶链式反应。

[0061] 术语 RCA 是滚环扩增的缩写。RCA 是一种涉及环形模版的重复拷贝而以线性方式来扩大信号的技术。

[0062] 杂交反应的“严谨性”是本领域普通技术人员容易测定的，并且通常取决于探针长度、洗涤温度和盐浓度根据经验计算。通常，较长的探针需要较高的温度，用于正确的退

火,而较短的探针需要较低的温度。杂交通常取决于在低于解链温度的环境中存在互补链时,变性 DNA 重新退火的能力。探针和可杂交序列之间所需的同源性程度越高,可以使用的相对温度越高。因此,断定较高的相对温度将倾向于使得反应条件更严谨,而较低的温度因此使反应条件不太严谨。杂交反应严谨性的其他详细内容和解释可以在例如 Ausuble 等, *Current Protocols in Molecular Biology* (分子生物学通用实验方案), Wiley Interscience Publishers, (1995) 中找到。

[0063] 如本文所定义,“严谨条件”或“高严谨条件”通常为:(1) 使用低离子强度和高温进行洗涤,例如 0.015M 氯化钠 /0.0015M 柠檬酸钠 /0.1% 十二烷基硫酸钠,50°C ;(2) 杂交期间使用变性剂,如甲酰胺,例如 50% (v/v) 甲酰胺,与 0.1% 牛血清白蛋白 /0.1% Ficoll/0.1% 聚乙烯吡咯烷酮 /50mM 磷酸钠缓冲液 pH 6.5,与 750mM 氯化钠,75mM 柠檬酸钠,42°C ;或(3) 使用 50% 甲酰胺,5×SSC (0.75M NaCl,0.075M 柠檬酸钠),50mM 磷酸钠 (pH 6.8),0.1% 焦磷酸钠,5×, Denhardt' s 溶液,超声波处理的鲑精 DNA (50 μ g/ml),0.1% SDS,和 10% 硫酸葡聚糖,42°C,并在 0.2×SSC (氯化钠 / 柠檬酸钠) 中于 42°C 和 50% 甲酰胺中于 55°C 洗涤,接着是包括于 55°C 在含 EDTA 的 0.1×SSC 中的高严谨洗涤。

[0064] “中等严谨条件”可如 Sambrook 等, *Molecular Cloning :A Laboratory Manual* (分子克隆 :实验室手册), New York, Cold Spring Harbor Press,1989 所述的来予以确认,并且包括使用不如上文所述那么严谨的洗涤溶液和杂交条件 (例如,温度、离子强度和 SDS%)。中等严谨条件的一个实例是于 37°C 在包含 20% 甲酰胺,5×SSC (150mM NaCl,15mM 柠檬酸三钠),50mM 磷酸钠 (pH 7.6),5×Denhardt' s 溶液,10% 硫酸葡聚糖和 20mg/ml 变性剪切鲑精 DNA 的溶液中过夜温育,接着在 1×SSC 中于约 37-50°C 下洗涤滤器。本领域技术人员将会明了如何按照需要对温度、离子强度等进行调整,以适应诸如探针长度等因素。

[0065] 术语“MUC5AC”意思是黏蛋白 5AC (Seq ID No. 1 和 4),并且包括标志物 MUC5AC,包括多核苷酸,如基因或基因片段、RNA 或 RNA 片段 ;或基因产物,包括多肽,如肽、寡肽、蛋白质,或蛋白质片段 ;或任何相关的代谢产物、副产物,或任何其他鉴定分子,如抗体或抗体片段。

[0066] 术语“MUC17”意思是细胞表面结合的人黏蛋白 17 (Seq ID No 2 和 5),并且包括标志物 MUC17,包括多核苷酸,如基因或基因片段、RNA 或 RNA 片段 ;或基因产物,包括多肽,如肽、寡肽、蛋白质,或蛋白质片段 ;或任何相关的代谢产物、副产物,或任何其他鉴定分子,如抗体或抗体片段。

[0067] 术语“ZG16”意思是人酶原粒蛋白 16 (Seq ID No 3 和 6),并且包括标志物 ZG16,包括多核苷酸,如基因或基因片段、RNA 或 RNA 片段 ;或基因产物,包括多肽,如肽、寡肽、蛋白质,或蛋白质片段 ;或任何相关的代谢产物、副产物,或任何其他鉴定分子,如抗体或抗体片段。

[0068] 除非另有说明,本发明的实施将采用分子生物学 (包括重组技术)、微生物学、细胞生物学和生物化学的常规技术,这落入本领域技术范围内。这样的技术在文献中有充分的说明,如 :*Molecular Cloning :A Laboratory Manual* (分子克隆 :实验室手册), 第 2 版, Sambrook 等,1989 ;*Oligonucleotide Synthesis* (寡核苷酸合成), MJ Gait 编辑,1984 ;*Animal Cell Culture*, R. I. Freshney 编辑,1987 ;*Methods in Enzymology* (酶学方法), Academic Press, Inc. ;*Handbook of Experimental Immunology* (实验免疫学

手册),第4版,D.M.Weir & CC.Blackwell 编辑,Blackwell Science Inc.,1987;Gene Transfer Vectors for Mammalian Cells(哺乳动物细胞的基因转移载体),J.M.Miller & M.P.Calos 编辑,1987;Current Protocols in Molecular Biology(分子生物学通用实验方案),F.M.Ausubel 等编辑,1987;以及 PCR:The Polymerase Chain Reaction(PCR:聚合酶链式反应),Mullis 等编辑,1994。

[0069] 将理解以上术语可以指蛋白质、DNA 序列和 / 或 RNA 序列。还将理解以上术语还指具有本文所示同源序列的非人蛋白质、DNA 和 / 或 RNA。

[0070] 发明详述

[0071] 通常,肿瘤标志物在肿瘤组织和相应的非恶性组织之间差异地表达。这提供了一种区分患有和未患有癌症的患者方法。然而,很可能即使那些标志物在肿瘤组织中不是过表达的,肿瘤组织的解剖学结构和生理学特征也将导致血清和其他生物液体中标志物累积的差异。特别地,预测与非恶性组织相比,异常的肿瘤细胞极性、渗漏的脉管系统和肿瘤组织的高间隙压力将促进特定的标志物流出肿瘤组织。因此,猜测在胃肿瘤组织中以非常高的水平表达,但与非恶性胃组织相比未必是过表达的分泌蛋白质将构成有用的胃癌标志物。

[0072] 使用微阵列分析和定量聚合酶链式反应(qPCR)的组合,已经鉴定出用于胃癌检测的新标志物。这种新的胃肿瘤标志物(GTM),在胃癌的早期检测中提供了更多工具。具体地,本发明包括新的GTM:MUC5AC(Seq ID No 1和4)、MUC17(Seq ID No 2和5)和ZG16(Seq ID No 3和6)。

[0073] 新的GTM可以单独使用,或备选地,可以组合在一起作为标记(包含两种或多种GTM)。根据本发明的标记包括MUC5AC、MUC17和ZG16中的至少一种,和至少一种另外的GTM,其可以是根据本发明的GTM,或任何其他GTM,包括已知的GTM。

[0074] 已知的适于结合本发明公开的GTM一起使用的GTM包括羧肽酶N、多肽2,83kDa链(CPN2)、基质金属蛋白酶12(MMP12)、抑制素(“INHBA”)、胰岛素样生长因子7(“IGFBP7”)、 γ -谷氨酰水解酶(“GGH”)、富含亮氨酸脯氨酸的蛋白聚糖(“LEPRE1”)、半胱氨酸蛋白酶抑制剂S(“CST4”)、分泌型卷曲相关蛋白4(“SFRP4”)、asporin(“ASPN”)、具有EF手型结构域的细胞生长调节剂1(“CGREF1”)、激肽释放酶10(KLK10)、金属蛋白酶的组织抑制剂1(“TIMP1”)、分泌型酸性富含半胱氨酸蛋白(“SPARC”)、转化生长因子,13-诱导的(“TGFB1”)、含有EGF的fibulin-样胞外基质蛋白2(“EFEMP2”)、基膜聚糖(“LUM”)、stannin(“SNN”)、分泌型磷蛋白1(“SPP1”)、硫酸软骨素蛋白聚糖2(“CSPG2”)、N-酰基鞘氨醇酰胺水解酶(“ASAHI”)、丝氨酸蛋白酶11(“PRSS11”)、分泌型卷曲相关蛋白2(“SFRP2”)、磷脂酶A2,XIIB组(“PLA2G12B”)、spondin2,胞外基质蛋白(“SPON2”)、嗅介蛋白1(“OLFM1”)、含有血小板反应蛋白重复1(“TSRC1”)、血小板反应蛋白2(“THBS2”)、adlcan、半胱氨酸蛋白酶抑制剂SA(“CST2”)、半胱氨酸蛋白酶抑制剂SN(“CST1”)、赖氨酰氧化酶样酶2(“LOXL2”)、甲状腺球蛋白(“TG”)、转化生长因子 β 1(“TGFB1”)、丝氨酸或半胱氨酸蛋白酶抑制剂Clade H成员1(“SERPINH1”)、丝氨酸或半胱氨酸蛋白酶抑制剂Clade B成员5(“SERPINB5”)、基质金属蛋白酶2(“MMP2”)、前蛋白转化酶枯草杆菌蛋白酶/kexin 5型(“PCSK5”)、乙酰透明质酸糖蛋白连接蛋白4(“HAPLN4”)、CA19-9、CA72-4、胃蛋白酶原和CEA,或之前已经鉴定为表示胃肿瘤的任何其他标志物。

[0075] 通过术语“可靠性”，我们包括假阳性和 / 或假阴性的低发生率。因此，使用较高可靠性的标志物，与使用该标志物形成的诊断相关的假阳性和 / 或假阴性较低。因此，在特定的实施方案中，提供了允许以高于现有技术标志物的可靠性约 50% 的可靠性来检测胃癌的标志物。在其他实施方案中，提供了可靠性高于约 70% 的标志物；在其他实施方案中，高于约 73%，在更多的其他实施方案中，高于约 80%，在更进一步的实施方案中，高于约 90%，在其他实施方案中，高于约 95%，在更多的实施方案中，高于约 98%，并且在特定的实施方案中，约 100% 的可靠性。

[0076] 癌症检测的一般方法

[0077] 以下列出了用于测定表达水平的一般方法，尽管将认识到用于测定表达水平的任何方法将是合适的。

[0078] 定量 PCR (qPCR)

[0079] 可以使用 GTM 特异性引物和探针，对肿瘤样品、血清和血浆进行定量 PCR (qPCR)。在受控的反应中，PCR 反应中形成的产物量 (Sambrook, J., E Fritsch, E. 和 T Maniatis, Molecular Cloning: A Laboratory Manual (分子克隆: 实验室手册), 第 3 版, Cold Spring Harbor Laboratory Press: Cold Spring Harbor) 与起始模板的量相关。可以通过 PCR 反应处于对数期时，在试剂变得限制之前，停止 PCR 反应来进行 PCR 产物的定量。然后将 PCR 产物在琼脂糖或聚丙烯酰胺凝胶中电泳，用溴化乙锭或相当的 DNA 染色剂进行染色，并且通过密度测定法来测量染色强度。或者，可以使用 PCR 仪器，如 Applied Biosystems' Prism 7000 或 Roche LightCycler 来测定 PCR 反应的进展，这些仪器实时地测量产物积累。实时 PCR 测量 DNA 嵌入染料如 Sybr 绿进入合成的 PCR 产物中的荧光，或报告分子从淬灭分子中裂解出来时释放的荧光；报告分子和淬灭分子结合至寡核苷酸探针中，所述寡核苷酸探针在 DNA 链从引物寡核苷酸延伸后与目标 DNA 分子杂交。寡核苷酸探针在下一个 PCR 循环中通过 Taq 聚合酶的酶作用得到置换和降解，将报告分子从淬灭分子中释放出来。在一种称为 Scorpion® 的变型中，探针与引物共价连接。

[0080] 反转录 PCR (RT-PCR)

[0081] RT-PCR 可以用于比较正常以及用或未用药物治疗的肿瘤组织中不同样品群中的 RNA 水平，以表征表达模式、分辨密切相关的 RNA 和分析 RNA 结构。

[0082] 对于 RT-PCR，第一步是从目标样品中分离出 RNA。起始材料通常是分别从人肿瘤或肿瘤细胞系以及相应的正常组织或细胞系分离的总 RNA。可以从多种样品分离 RNA，样品例如为来自乳房、肺、结肠（例如，大肠或小肠）、结直肠、胃、食管、肛门、直肠、前列腺、脑、肝脏、肾脏、胰腺、脾脏、胸腺、睾丸、卵巢、子宫、膀胱等组织的肿瘤样品，来自原发性肿瘤或肿瘤细胞系，以及来自健康供体收集的样品。如果 RNA 来源是肿瘤，例如，可以从冷冻或存档的石蜡包埋和固定的（例如，福尔马林固定的）组织样品中提取出 RNA。

[0083] 通过 RT-PCR 制作基因表达谱的第一步是将 RNA 模板反转录为 cDNA，接着在 PCR 反应中进行指数扩增。两种最常用的反转录酶是禽成髓细胞瘤病毒反转录酶 (AMV-RT) 和莫洛尼鼠白血病病毒反转录酶 (MMLV-RT)。反转录步骤通常利用特异性引物、随机六聚体或寡聚 dT 引物引发，这将视情况和制作表达谱的目标而定。例如，可以利用 GeneAmp RNA PCR 试剂盒 (Perkin Elmer, CA, USA)，按照生产商的说明，对提取的 RNA 进行反转录。然后将所得到的 cDNA 在随后的 PCR 反应中用作模板。

[0084] 虽然 PCR 步骤能够利用多种热稳定的 DNA 依赖性 DNA 聚合酶,但通常采用 Taq DNA 聚合酶,其具有 5'-3' 核酸酶活性,但缺乏 3'-5' 校对核酸内切酶活性。因此,TaqMan qPCR 通常利用 Taq 或 Tth 聚合酶的 5' 核酸酶活性来水解结合目标扩增子的杂交探针,但任何具有等同 5' 核酸酶活性的酶均可使用。

[0085] 可以利用两个寡核苷酸引物来产生 PCR 反应典型的扩增子。设计第三个寡核苷酸或探针来检测位于两个 PCR 引物之间的核苷酸序列。探针是不能通过 Taq DNA 聚合酶延伸的,并用报告荧光染料和淬灭荧光染料标记。当两种染料如其在探针上一样位置紧靠在一起时,任何激光诱导的报告染料的发射将被淬灭染料淬灭。在扩增反应期间,Taq DNA 聚合酶以模板依赖性的方式切割探针。所得到的探针片段在溶液中解离,并且来自所释放的报告染料的信号不受第二个荧光团的淬灭作用的影响。每合成一个新的分子,就释放一分子的报告染料,并且未淬灭的报告染料的检测提供了数据定量解释的基础。

[0086] TaqMan RT-PCR 可以使用可购得的设备来进行,如,例如 ABI PRISM 7700 序列检测系统 (Perkin-Elmer-Applied Biosystems, Foster City, CA, USA) 或 Lightcycler (Roche Molecular Biochemicals, Mannheim, 德国)。在优选的实施方案中,在实时定量 PCR 装置 (如 ABI PRISM 7700 序列检测) 上运行 5' 核酸酶程序。该系统由热循环仪、激光、电荷耦合器件 (CCD)、照相机和计算机组成。该系统在热循环仪上以 96 孔的形式扩增样品。在扩增过程中,通过光纤光缆实时采集所有 96 个孔的激光诱导的荧光信号,并在 CCD 处进行检测。该系统包括用于运行仪器和用于分析数据的软件。

[0087] 5' 核酸酶试验数据最初表达为 Ct 或阈值循环。如上文所讨论的那样,在每一循环中记录荧光值,其代表在扩增反应中到该点时所扩增的产物量。第一次记录到具有统计学显著性的荧光信号时的点为阈值循环。

[0088] 实时定量 PCR (qRT-PCR)

[0089] RT-PCR 技术较为新近的一种变型是实时定量 PCR,其通过双重标记的荧光生成探针 (即 TaqMan 探针) 测量 PCR 产物的积累。实时 PCR 与定量竞争 PCR 和定量比较 PCR 都兼容。前者利用每个目标序列的内部竞争物进行标准化,而后者利用样品内所含的归一化基因或持家基因进行 RT-PCR。更多细节由例如 Held 等,Genome Research 6 :986-994(1996) 来提供。

[0090] 可以使用固定的、石蜡包埋的组织作为 RNA 来源来测定表达水平。根据本发明的一个方面,设计 PCR 引物,使其在待扩增的基因中存在的内含子序列的两侧。在该实施方案中,引物 / 探针设计的第一步是描述出基因内的内含子序列。这可以通过公众可获得的软件来进行,如由 Kent, W. J., Genome Res. 12(4) :656-64(2002) 开发的 DNA BLAT 软件,或 BLAST 软件,包括其变型。后续步骤按照 PCR 引物和探针设计的成熟方法来进行。

[0091] 为了避免非特异性信号,在设计引物和探针时遮蔽内含子内的重复序列是有用的。这可以通过使用 Baylor College of Medicine 在线获得的 Repeat Masker 程序容易地实现,该程序相对于重复元件文库筛选 DNA 序列,并返回其中已经遮蔽了重复元件的查询序列。然后可以将已经遮蔽的序列用于设计引物和探针序列,利用任何可购得的或其他公众可获得的引物 / 探针设计包,如 Primer Express (Applied Biosystems) ;MGB 设计试验 (Applied Biosystems) ;Primer3 (Steve Rozen 和 Helen J. Skaletsky (2000) Primer3 on the VIMNV for general users and for biologist programmers (用于一般使用者和生

物学家程序员的 VIMNV 上的 Primer3), 见 :Krawetz S, Misener S(编辑)Bioinformatics Methods and Protocols :Methods in Molecular Biology(生物信息学方法和实验方案:分子生物学方法), Humana Press, Totowa, NJ, 365-386 页)。

[0092] PCR引物设计中考虑的最重要的因素包括引物长度、解链温度(T_m)和G/C含量、特异性、互补引物序列和3'端序列。通常,最佳PCR引物一般为17-30个碱基长,并含有约20-80%,如,例如约50-60%的G+C碱基。通常优选50-80℃,例如,约50-70℃的解链温度。有关PCR引物和探针设计的更多指南,参见,例如Dieffenbach,C.W.等,General Concepts for PCR Primer Design(PCR引物设计的一般概念),见:PCR Primer, A Laboratory Manual(PCR引物,实验室手册),Cold Spring Harbor Laboratory Press,New York,1995,133-155页;Innis和Gelfand,Optimization of PCR(PCR的优化),见:PCR Protocols, A Guide to Methods and Applications(PCR实验方案,方法和应用指导),CRC Press, London,1994,5-11页;和Plasterer,T.N.Primerselect:Primer and probe design(引物选择:引物和探针设计),Methods Mol.Biol.70:520-527(1997),在此特意将其全部公开内容明确引入作为参考。

[0093] 微阵列分析

[0094] 差异表达也可以使用微阵列技术来鉴定或验证。因此,可以使用微阵列技术在新鲜或者石蜡包埋的肿瘤组织中测量GTM的表达谱。在这种方法中,在微芯片基质上覆盖或排列目标多核苷酸序列(包括cDNA和寡核苷酸)。排列的序列(即,捕获探针)随之与来自目标细胞或组织(即,靶标)的特异性多核苷酸杂交。与RT-PCR方法一样,RNA来源通常为从人肿瘤或肿瘤细胞系以及相应的正常组织或细胞系中分离出来的总RNA。因此,RNA可以分离自多种原发性肿瘤或肿瘤细胞系。如果RNA的来源是原发性肿瘤,则可以例如从冷冻或存档的福尔马林固定的石蜡包埋的(FFPE)组织样品和固定的(例如,福尔马林固定的)组织样品中提取RNA,这些样品在日常的临床实践中都是常规制备和保存的。

[0095] 在微阵列技术的特定实施方案中,向基质上施加PCR扩增的cDNA克隆的插入片段。基质可以包括高达1、2、5、10、15、20、25、30、35、40、45、50或75个核苷酸序列。在其他方面中,基质可以包括至少10,000个核苷酸序列。固定在微芯片上的微阵列序列适合于在严谨条件下杂交。作为其他实施方案,用于微阵列的靶标可以为至少50、100、200、400、500、1000或2000个碱基长;或50-100、100-200、100-500、100-1000、100-2000或500-5000个碱基长。作为另外的实施方案,用于微阵列的捕获探针可以为至少10、15、20、25、50、75、80或100个碱基长;或10-15、10-20、10-25、10-50、10-75、10-80或20-80个碱基长。

[0096] 通过反转录从目标组织提取的RNA,通过荧光核苷酸的掺入可以生成荧光标记的cDNA探针。施加到芯片上的标记cDNA探针与阵列上的每个DNA斑点特异性地杂交。严谨洗涤以除去非特异性结合的探针后,通过共焦激光显微镜或通过另一种检测方法(如CCD照相机)来扫描芯片。对每个阵列化元素的杂交的定量允许评估相应mRNA的丰度。利用双色荧光,从两个RNA源生成分开标记的cDNA探针,其成对地与阵列杂交。由此,对应于每个特定基因的来自两个来源的转录物的相对丰度可以得以同时确定。

[0097] 这种小型化规模的杂交提供了对大量基因表达模式的方便快捷的评价。此类方法已经证明具有检测微量转录物和可再现地检测至少大约两倍的表达水平差异所需的灵敏度,其中所述微量转录物仅以每个细胞几个拷贝的水平表达(Schena等,Proc.Natl.Acad.

Sci. USA 93(2) :106-149(1996))。微阵列分析可以通过可购得的设备按照生产商的实验方案来进行,如使用 Affymetrix GenChip 技术、Illunima 微阵列技术或 Incyte 微阵列技术。对大规模分析基因表达的微阵列方法的研发,使得可以在多种肿瘤类型中系统地寻找癌症分类和结果预测的分子标志物。

[0098] RNA 分离、纯化和扩增

[0099] 用于 mRNA 提取的一般方法是本领域公知的,并公开在分子生物学的标准教科书中,包括 Ausubel 等,Current Protocols of Molecular Biology(分子生物学通用实验方案), John Wiley and Sons(1997)。例如, Rupp 和 Locker, Lab Invest. 56 :A67(1987),以及 De Sandres 等, BioTechniques 18 :42044(1995) 中公开了从石蜡包埋的组织中提取 RNA 的方法。特别地,可以使用来自商业生产商如 Qiagen 的纯化试剂盒、成套缓冲液和蛋白酶,按照生产商的说明,来进行 RNA 分离。例如,可以使用 Qiagen RNeasy 微型柱从培养细胞分离总 RNA。其他可购得的 RNA 分离试剂盒包括 MasterPure Complete DNA 和 RNA 纯化试剂盒 (EPICENTRE (D, Madison, WI) 和 Paraffin Block RNA 分离试剂盒 (Ambion, Inc.))。可以使用 RNA Stat-60 (Tel-Test 公司) 从组织样品分离总 RNA。例如,可以通过氯化铯密度梯度离心分离由肿瘤制备的 RNA。

[0100] 利用固定的石蜡包埋的组织作为 RNA 来源进行基因表达谱制作的代表性实验方案的步骤包括 mRNA 分离、纯化、引物延伸和扩增,在多篇发表的期刊文章中给出(例如: T. E. Godfrey 等, J. Molec. Diagnostics 2 :84-91(2000) ;K. Specht 等, Am. J. Pathol. 158 :419-29(2001))。简言之,一个代表性方法从切割约 10 微米厚的石蜡包埋肿瘤组织样品切片开始。然后提取 RNA,除去蛋白质和 DNA。分析 RNA 浓度后,如果需要,可以包括 RNA 修复和 / 或扩增步骤,并且利用基因特异性启动子反转录 RNA,接着进行 RT-PCR。最后,分析数据以基于所研究的肿瘤样品中鉴定到的特征性基因表达模式来确定患者可用的最佳治疗选择。

[0101] 免疫组织化学和蛋白质组学

[0102] 免疫组织化学法也适用于检测本发明的增殖标志物的表达水平。因此,利用每种标志物特异性的抗体或抗血清,优选多克隆抗血清,最优选单克隆抗体,来检测表达。可以通过用例如放射性标记、荧光标记、半抗原标记(如,生物素),或酶(如,辣根过氧化物酶或碱性磷酸酶)直接标记抗体本身来检测抗体。或者,可以将未标记的一级抗体与标记的二级抗体联合使用,所述二级抗体包括一级抗体特异性的抗血清、多克隆抗血清或单克隆抗体。免疫组织化学实验方案和试剂盒是本领域公知的并且可商购。

[0103] 蛋白质组学可以用来分析特定时间点样品(例如,组织、器官或细胞培养物)中存在的多肽。特别地,蛋白质组学技术可以用来评估样品中多肽表达的全局变化(也称为表达蛋白质组学)。蛋白质组学分析通常包括:(1) 通过 2-D 凝胶电泳 (2-D PAGE) 分离样品中单独的多肽;(2) 鉴定从凝胶中回收的单独的多肽,例如,通过质谱或 N-端测序,和 (3) 利用生物信息学分析数据。蛋白质组学方法对于其他基因表达谱研究方法是有价值的补充,并且可以单独使用或与其他方法结合使用,以检测本发明的增殖标志物的产物。使用标志物选择性的核酸探针的杂交方法

[0104] 这些方法涉及将核酸探针结合支持体,并且在合适的条件下与源自测试样品的 RNA 或 cDNA 杂交 (Sambrook, J., E Fritsch, E. 和 T Maniatis, Molecular Cloning :A

Laboratory Manual(分子克隆:实验室手册),第3版,Cold Spring Harbor Laboratory Press:Cold Spring Harbor(2001))。这些方法适用于源自肿瘤组织或液体样品的 GTM。通常用荧光或放射性分子来标记 RNA 或 cDNA,以能够检测和定量。在一些应用中,可以用分支的、荧光标记的结构将杂交 DNA 标记,以增强信号强度(Nolte, F. S. Branched DNA signal amplification for direct quantitation of nucleic acid sequences in clinical specimens(用于临床样品中核酸序列的直接定量的分支 DNA 信号扩增), Adv. Clin. Chem. 33, 201-35(1998))。在通过凝胶图像的荧光检测或密度测定定量杂交量之前,通过在低盐溶液(如, $0.1 \times \text{SSC}$, $0.5\% \text{ SDS}$) 中强烈洗涤来除去未杂交的标记。支持体可以是固体,如尼龙或硝基纤维素膜,或由在液体悬浮液中时杂交的微球体或珠子组成。为了可以洗涤和纯化,珠子可以是磁性的(Haukanes, B-1 和 Kvam, C., Application of magnetic beads in bioassays(磁性珠子在生物试验中的应用), Bio/Technology 11, 60-63(1993))或是荧光标记的,以能够进行流式细胞术(参见,例如: Spiro, A., Lowe, M. 和 Brown, D., A Bead-Based Method for Multiplexed Identification and Quantitation of DNA Sequences Using Flow Cytometry(使用流式细胞术的用于 DNA 序列的多路鉴定和定量的基于珠子的方法), Appl. Env. Micro. 66, 4258-4265(2000))。

[0105] 杂交技术的一种变型是 QuantiGene Plexe 测定(Genospectra, Fremont), 其结合了荧光珠子支持体和分支 DNA 信号扩增。杂交技术的再一种变型是 Quantikine[®] mRNA 测定(R&D Systems, Minneapolis)。方法按照生产商的说明书中所述的。简言之,测定使用与洋地黄毒苷缀合的寡核苷酸杂交探针。在比色试验中,使用偶联碱性磷酸酶的抗洋地黄毒苷抗体来检测杂交。

[0106] 其他的方法是本领域公知的,并且在本文中不需要更多描述。

[0107] 酶联免疫测定(ELISA)

[0108] 简言之,在夹层 ELISA 测定中,将对抗 GTM 的多克隆或单克隆抗体结合固相支持体(Crowther, J. R. The ELISA guidebook(ELISA 参考指南), Humana Press: New Jersey(2000); Harlow, E. 和 Lane, D., Using antibodies: a laboratory manual(使用抗体:实验室手册), Cold Spring Harbor Laboratory Press: Cold Spring Harbor(1999))或悬浮液珠子。其他方法是本领域公知的,并且在本文中不需要更多描述。单克隆抗体可以是杂交瘤衍生的,或选自噬菌体抗体文库(Hust M. 和 Dubel S., Phage display vectors for the in vitro generation of human antibody fragments(用于体外产生人抗体片段的噬菌体展示载体), Methods Mol Biol. 295: 71-96(2005))。用非目标蛋白制备物和去污剂将非特异性结合位点阻断。然后用来自含有 GTM 抗原的患者的样品或组织的制备物温育捕获抗体。在用检测目标 GTM 的二级抗体温育抗体/抗原复合物之前,洗涤混合物。二级抗体通常缀合荧光分子或可以在酶反应中检测或用缀合报告子的第三抗体检测的其他报告分子(Crowther, 同上)。或者,在直接 ELISA 中,含有 GTM 的制备物可以结合支持体或珠子,并且用抗体-报告子缀合物直接检测目标抗原(Crowther, 同上)。

[0109] 用于生产单克隆抗体和多克隆抗血清的方法是本领域公知的,并且在本文中不需要更多描述。

[0110] 免疫检测

[0111] 这些方法还可以用于免疫检测在外科手术取出肿瘤前后取自胃癌患者的血清或

血浆中的标志物家族成员、免疫检测患有其他癌症的患者中的标志物家族成员,所述其他癌症包括但不限于,结直肠癌、胰腺癌、卵巢癌、黑素瘤、肝癌、食道癌、胃癌、子宫内膜癌和脑癌,以及免疫检测来自胃癌患者的尿液和粪便中的标志物家族成员。

[0112] 还可以使用其他标准免疫检测技术,如免疫印迹或免疫沉淀,来检测组织或样品中的 GTM(Harlow,E. 和 Lane,D.,Using antibodies :a laboratory manual(使用抗体:实验室手册),Cold Spring Harbor Laboratory Press :Cold Spring Harbor(1999))。在免疫印迹中,将来自含有 GTM 的组织或液体的蛋白质制备物在变性或非变性条件下通过聚丙烯酰胺凝胶电泳。然后将蛋白质转移至膜支持体上,如尼龙。然后按照对免疫组织化学所述的,将 GTM 与单克隆或多克隆抗体直接或间接反应。或者,在一些制备物中,可以将蛋白质直接点在膜上,而没有之前的电泳分离。可以通过密度测定来定量信号。

[0113] 在免疫沉淀中,用抗 GTM 的单克隆或多克隆抗体温育含有 GTM 的可溶性制备物。然后用惰性珠子温育反应物,所述惰性珠子由具有共价连接的 A 蛋白或 G 蛋白的琼脂糖或聚丙烯酰胺制备。A 或 G 蛋白珠子与抗体特异性地反应,形成结合珠子的抗体-GTM-抗原的固定化复合物。洗涤后,可以通过免疫印迹或 ELISA 检测和定量结合的 GTM。

[0114] 阈值测定

[0115] 对于使用 GTM 的测试,将得到能够使样品对于胃癌称为阳性或阴性的阈值。这些阈值将通过正在研究胃癌存在的患者群的分析来测定。阈值对于不同的测试应用将不同;例如,使用大部分无泌尿科症状的患者群来测定用于群筛选测试中的阈值,并且这些阈值不同于用于监测胃癌复发患者的测试中所用的那些。可以选择阈值,以在所需的临床背景中提供实际水平的测试特异性;即,允许合理灵敏度但没有过量接受假阳性结果患者的特异性。该特异性可以在 80-90% 的范围内。获得测试阈值的备选方法是对于不同的测试阈值,将灵敏度相对于特异性作图(ROC 曲线),然后选择曲线的拐点。

[0116] 作为单个阈值的备选,测试可以使用测试间隔,其提供了不同程度的疾病存在的可能性,并且其具有与这些可能性相关的不同临床结果。例如,测试可以具有三个间隔;一个与高风险(例如,90%)的胃癌存在相关,第二个与低风险的胃癌相关,而第三个认为是怀疑患病的。“怀疑”的间隔将与限定的时间段内的重复测试的推荐相关。

[0117] 胃癌标志物的抗体

[0118] 在其他方面中,本发明包括制备针对 GTM 的抗体。使用本文中所述的方法,可以使用微阵列和/或 qRT-PCR 方法来鉴定新的 GTM。一旦鉴定出推定的标志物,可以足量生产以适合于引发免疫应答。在一些情况中,可以使用全长 GTM,而在其他情况中,GTM 的肽片段足够作为免疫原。可以将免疫原注入合适的宿主中(例如,小鼠、兔子等),并且如果需要,可以注入佐剂,如弗氏完全佐剂或弗氏不完全佐剂,以提高免疫应答。可以认识到制备抗体在免疫学领域中是常规的,并且在本文中不需要更多描述。因此,可以产生针对使用本文中所述方法鉴定的 GTM 的抗体,包括单克隆或噬菌体展示抗体。

[0119] 在进一步的实施方案中,可以制备针对本文中鉴定的肿瘤标志物的蛋白质或蛋白质核心的抗体或针对对于 GTM 独特的寡核苷酸序列的抗体。尽管特定的蛋白质可以糖基化,因此在特定的情况中,糖基化模式的变化可导致缺乏常规糖基化模式的 GTM 形式的错误检测。因此,在本发明的特定方面中,GTM 免疫原可以包括脱糖基化的 GTM 或脱糖基化的 GTM 片段。可以使用一种或多种本领域已知的糖苷酶来完成脱糖基化。或者,GTMcDNA 可以

在缺乏糖基化的细胞系中表达,如在原核细胞系中表达,包括大肠杆菌等。

[0120] 可以制备其中具有 GTM 编码寡核苷酸的载体。许多这样的载体可以基于本领域已知的标准载体。载体可以用于转染多种细胞系,以产生 GTM- 产生细胞系,其可以用于产生所需量的 GTM,用于研发特定的抗体或其他试剂,用于检测 GTM 或用于标准化为 GTM 研发的测定法。

[0121] 试剂盒

[0122] 基于本发明的发现,可以设想和产生几种类型的测试试剂盒。首先,可以制备具有预先装载了检测分子(或“捕获试剂”)的检测装置的试剂盒。在用于检测 GTM mRNA 的实施方案中,这样的装置可以包括基质(例如,玻璃,硅,石英,金属等),在其上结合了寡核苷酸,作为与待检测的 mRNA 杂交的捕获试剂。在一些实施方案中,可以通过 mRNA(用 cy3、cy5、放射性标记或其他标记来标记)与基质上的寡核苷酸杂交来完成 mRNA 的直接检测。在其他实施方案中,可以通过首先制备所需 mRNA 的互补 DNA(cDNA)来完成 mRNA 的检测。然后,可以将标记的 cDNA 与基质上的寡核苷酸杂交,并检测。

[0123] 抗体还可以作为捕获试剂用于试剂盒中。在一些实施方案中,基质(例如,多孔平板)可以具有与其连接的特异性 GTM 捕获试剂。在一些实施方案中,试剂盒可以包括阻断试剂。阻断试剂可以用于降低非特异性结合。例如,可以使用过量的来自不含有 GTM 寡核苷酸的任何常规来源的 DNA(如,鲑精 DNA)来降低非特异性的寡核苷酸结合。可以使用过量的阻断蛋白(如,血清白蛋白)来降低非特异性的抗体结合。可以认识到用于检测寡核苷酸和蛋白质的多种方法是本领域已知的,并且可以使用特异性地检测 GTM 相关分子的任何策略,并且认为在本发明的范围内。

[0124] 结合固相支持体时,例如,使用抗体芯片时,也可以使用抗体,抗体芯片允许使用单个芯片检测多种标志物。

[0125] 除了基质以外,测试试剂盒可以包含捕获试剂(如,探针)、洗涤溶液(例如,SSC、其他盐、缓冲剂、去污剂等),以及检测部分(例如,cy3、cy5、放射性标记等)。试剂盒还可以包括使用说明书和包装。

[0126] 可以使用任何合适的技术来检测样品中的癌症标志物,这些技术可以包括,但不限于,寡核苷酸探针、qPCR 或针对癌症标志物产生的抗体。

[0127] 将认识到待测试的样品不限于怀疑是肿瘤的组织样品。标志物可以分泌至血清或其他体液中。因此,样品可以包括任何身体样品,并且包括生检样品、血液、血清、腹膜冲洗物、脑脊髓液、尿液和粪便样品。

[0128] 还将认识到本发明不限于人类的癌症检测,而且还适于任何动物的癌症检测,包括,但不限于,狗、猫、马、牛、绵羊、鹿、猪和已知患癌症的任何其他动物。

[0129] 体液中胃癌标志物的测试

[0130] 在几个实施方案中,理想地可以对获自血液、血浆、血清、腹膜液(例如,获自腹膜冲洗物)或其他体液(如,尿液、淋巴、脑脊液、胃液)的样品或粪便样品进行 GTM 测定。

[0131] 通常,测试这些液体中寡核苷酸、蛋白质和肽的方法是本领域已知的。可以使用如 Northern 印迹、Southern 印迹或微阵列方法,或 qPCR 的杂交方法来进行寡核苷酸的检测。用于检测蛋白质的方法,包括如酶联免疫吸附测定(ELISA)、具有抗体的蛋白质芯片、悬浮液珠子放射免疫测定(RIA)、Western 印迹和凝集素结合。然而,为了说明的目的,可以使用

夹层型酶联免疫吸附测定 (ELISA) 来定量 GTM 的液体水平。对于血浆测定, 将 5uL 等份的适当稀释的样品或连续稀释的标准 GTM 和 75uL 的过氧化物酶缀合的抗-人 GTM 抗体加入微量滴定平板的孔中。在 30°C 下 30 分钟的温育期后, 用磷酸缓冲盐水 {PBS} 中的 0.05% 吐温 20 洗涤孔, 来除去未结合的抗体。然后在 30°C 下用含有邻-亚苯基二胺的 H₂O₂ 温育结合的 GTM 和抗-GTM 抗体复合物 15 分钟。通过加入 1M H₂SO₄ 来停止反应, 并且用微量滴定平板读出器来测量 492nm 处的吸光值。

[0132] 可以认识到抗-GTM 抗体可以是单克隆抗体或多克隆抗血清。还可以认识到可以合适地研究任何其他体液。

[0133] 在生理学意义上, 标志物不必是分泌的才是有用的。相反, 标志物蛋白质或基因进入血清的任何机制在产生可检测的、可定量水平的标志物中是有效的。因此, 可溶性蛋白从细胞的正常分泌、膜蛋白从质膜的脱落、选择性剪接形式的 mRNA 或从其表达的蛋白质的分泌、细胞死亡 (或凋亡) 可以产生足够水平的有用标志物。

[0134] 对于使用血清标志物作为诊断和 / 或评价多种癌症类型疗效的工具, 存在越来越多的支持。

[0135] Yoshikawa 等, (Cancer Letters, 151 :81-86 (2000)), 描述了胃癌患者血浆中的基质金属蛋白酶-1 的组织抑制剂。

[0136] Rudland 等, (Cancer Research 62 :3417-3427 (2002)) 描述了骨桥蛋白作为人乳腺癌中的转移相关蛋白。

[0137] Buckhaults 等, (Cancer Research 61 :6996-7001 (2002)) 描述了结直肠肿瘤中特定分泌的和表达的细胞表面基因。

[0138] Kim 等, (JAMA 287 (13) :1671-1679 (2002)) 描述了骨桥蛋白作为卵巢癌的潜在诊断生物标志物。

[0139] Hotte 等, (AJ. American Cancer Society 95 (3) :507-512 (2002)), 描述了血浆骨桥蛋白作为人体液中可检测的蛋白质, 并且与特定的恶性肿瘤相关。

[0140] Martin 等, (Prostate Cancer Prostatic Dis. 2004 年 3 月 9 日 (PMID :15007379) (摘要)), 描述了人激肽释放酶 2, 前列腺特异性抗原 (PSA) 和游离 PSA 作为检测前列腺癌的标志物的用途。

[0141] Hall 等, (Laryngoscope 113 (1) :77-81 (2003) (PMID :12679418) (摘要)), 描述了血清甲状腺球蛋白在甲状腺癌中的预测价值。

[0142] Mazzaferri 等, (J. Clin. Endocrinol. Metab. 88 (4) :1433-1441 (2003) (摘要)), 描述了甲状腺球蛋白作为甲状腺癌患者的潜在监测方法。

[0143] Whitley 等, (Dim Lab. Med. 24 (1) :29-47 (2004) (摘要)), 描述了甲状腺球蛋白作为甲状腺癌的血清标志物。

[0144] Kuo 等, (Clin. Chim. Acta. 294 (1-2) :157-168 (2000) (摘要)), 描述了 HCF- 和 HBV- 感染患者中的血清基质金属蛋白酶-2 和 -9。

[0145] Koopman 等, (Cancer Epidemiol. Biomarkers Prev 13 (3) :487-491 (2004) (摘要)), 描述了骨桥蛋白作为胰腺癌的生物标志物。

[0146] Pellegrini 等, (Cancer Immunol. Immunother. 49 (7) :388-394 (2000) (摘要)), 描述了可溶性癌胚抗原和 TIMP1 作为入侵前结直肠癌的标志物的测量。

[0147] Melle 等, (Clin. Chem. 53(4), 629-635(2007) (摘要)), 描述了 HSP27 作为胰腺癌的血清标志物。

[0148] Leman 等, (Urology, 69(4)714-20(2007) (摘要)), 描述了 EPCA-2 作为前列腺癌的血清标志物。

[0149] Tsigkou 等, (I Clin Endocrinol Metab, 92(7)2526-31(2007) (摘要)), 描述了总抑制素作为卵巢癌的潜在血清标志物。

[0150] Marchi 等, (Cancer 112, 1313-1324(2008) (摘要)), 描述了前载脂蛋白 AI 作为肺癌患者脑部转移的血清标志物。

[0151] 方法

[0152] 使用以下的一般方法来评价多种方法对胃肿瘤相关标志物的分子鉴定的适用性。

[0153] 肿瘤收集

[0154] 从首尔国立大学医院 (Seoul National University Hospital) 切除的外科手术样本收集胃肿瘤样品和非恶性胃组织。基于症状、生理发现和组织的组织学检查来进行胃癌的诊断。

[0155] RNA 提取

[0156] 在一些实施方案中, 通过测定取自肿瘤的样品中的 RNA 水平来分析胃肿瘤相关基因的表达。将冷冻外科手术样本包埋在 OCT 介质中。使用切片机从组织块切出 60 微米的切片, 在 TriReagent : 水 (3 : 1) 混合物中均质化, 然后氯仿萃取。然后使用 RNeasy™ 程序 (Qiagen) 从水相纯化总 RNA。总共从 58 个胃肿瘤和 58 个非恶性 (“正常”) 胃组织样品中提取了 RNA, 并用于以下所述的微阵列分析中。还从 16 个癌细胞系中提取出 RNA, 并合并作为参照 RNA。

[0157] 微阵列载玻片制备

[0158] 环氧涂布的玻璃载玻片获自 MWG Biotech AG, Ebersberg, 德国, 并且根据生产商的实验方案, 使用 Gene Machines 微阵列机器人用 -30,000 个 50mer 寡核苷酸印刷。

[0159] RNA 标记和杂交

[0160] 在含有 5-(3-氨基烯丙基)-2' 脱氧尿苷 -5' -三磷酸盐的反应物中, 使用 Superscript II 反转录酶 (Invitrogen) 从 10ug 总 RNA 转录 cDNA。然后将反应物在 Microcon 柱中去离子, 接着在室温下, 在碳酸氢盐缓冲液中用 Cy3 或 Cy5 温育。使用 Qiaquick 柱 (Qiagen) 除去未结合的染料, 并且将样品在 SpeedVac 中浓缩至 15ul。然后将 Cy3 和 Cy5 标记的 cDNA 与 Ambion ULTRAhyb 缓冲液混合, 在 100°C 下变性 2 分钟, 并在杂交室中, 在 42°C 下, 与微阵列载玻片杂交 16 小时。然后洗涤载玻片, 并在 Axon 4000A 扫描仪中, 在两个功率设定下扫描两次, 以产生基因表达的原始荧光数据。

[0161] 归一化程序

[0162] 为了测量肿瘤和非癌组织中癌基因的表达, 通过减去局部背景荧光强度来校正由 Genepix™ 软件检测的中值荧光强度。排除了背景校正强度低于零的点。为了促进归一化, 将强度比例和整体点强度进行了对数转化。使用 LOCFIT™ 包中实现的局部回归对染料和空间偏差校正了对数转化的强度比例。同时对于整体点强度和位置回归了对数转化的强度比例。局部回归的残差提供了校正的对数 - 倍数变化。对于质量控制, 相对于点强度和位置, 对每个归一化的微阵列的比例作图。随后目测检查图的可能剩余的假象。此外, 方差分析

(ANOVA) 模型适用于针尖 (pin-tip) 偏差的检测。将归一化的所有结果和参数插入用于统计学分析的 Postgres- 数据库中。

[0163] 标志物选择

[0164] 根据每个基因在肿瘤和非恶性组织中的相对信号强度,将 29,718 个基因的每个微阵列基因表达数据进行分级。进一步的分析限于 (i) 编码分泌蛋白的基因, (ii) 在肿瘤组织中具有高于对编码现有肿瘤标志物 CEA 的基因 (CEACAM5) 观察到的强度等级的基因和 (iii) 在血液或脉管组织中不具有显著表达的基因,如通过 Unigene 数据库中的 EST 计数所测定的 (Wheeler DL 等,2003)。通过鉴定预期含有 N- 端信号肽的转录物来预测分泌蛋白。丢弃具有不在头 20 个 N- 端氨基酸的预测的跨膜螺旋的蛋白 [Krogh A. 等,2001]。使用 TARGETP [Emanuelsson O 等,2000] 预测更多的亚细胞定位。

[0165] 相关寡核苷酸的参考号 (MWG 寡聚物 #),以及选定 GTM 的 NCBI mRNA 和蛋白质参考序列显示于图 1 中。图 1 还显示了选定的 GTM 在肿瘤和非恶性组织中的等级强度。下文中显示了本发明的 GTM 的全部 DNA 序列。

[0166] 定量实时 PCR

[0167] 在其他实施方案中,可以将实时或定量 PCR (qPCR) 用于 PCR 模板拷贝数的绝对或相对定量。使用 Primer Express V 2.0™ (Applied Biosystems) 设计 MUC17 的引物组 (正向 :GAGGTGGTCAGCAGCATTGAC ;反向 :CCTGGGAAGAGTG GTTTTTTAGC),并使用 SYBR 绿色标记来检测扩增的产物。通过 Assay-on-Demand™ 表达试验 Hs. 00380609_ml (Applied Biosystems) 来表示 ZG16。在 ABI Prism™ 7000 序列检测系统上在标准循环条件下进行了扩增。

[0168] 使用由单个 cDNA 表示的每个 RNA 样品在两个 96 孔平板上进行试验。分析了高达 45 个来自胃肿瘤和非恶性胃组织的 RNA 样品。每个平板含有参照 cDNA 标准曲线,在 625- 倍浓度范围内,以一式两份进行。分析由计算 ΔCT 组成 (目标基因 CT- 平均参照 cDNA CT)。 ΔCT 与负 \log_2 倍数变化成正比。然后计算相对于中值非恶性 \log_2 倍数变化的 \log_2 倍数变化 (\log_2 倍数变化 - 中值正常 \log_2 对数变化)。然后将这些倍数变化聚类成频率级并作图。

[0169] 蛋白质表达和抗体产生

[0170] 为了证实蛋白质水平的 ZG16,必须产生抗重组蛋白的新抗体。使用正向引物 CACCAATGCCATTTCAGGCCAGGT 和反向引物 TCAGCATCTGCTGCAGCTA 从人细胞系 cDNA PCR 扩增 ZG16 的编码区 17-167。将 PCR 产物凝胶纯化并且克隆至来自 Invitrogen 的“Gateway”进入载体“pENTR/dTOPO”中,然后测序以证实序列的正确插入。使用“Gateway”系统,然后从 pENTR/dTOPO 将 ZG16 克隆至含有 N 端 6×HIS 标记物的 Invitrogen 表达载体 pDEST17 中。在 BL21-AI 大肠杆菌细胞 (Invitrogen) 中进行 ZG16 的表达,将细胞在 37°C 下在摇床上生长,直至它们处于中间对数期 ($OD_{600} = 0.5$),由此在 0.2% 最终浓度的阿拉伯糖下诱导,并且在 37°C 下在摇床上再生长 3 小时。通过在 6000×g 下离心 15 分钟来收集细胞,并将上清液丢弃。将细胞重悬浮于 PBS (pH7.0) 中,并且使用 60% 功率的 Sonics Vibra cell 通过超声波来裂解。通过在 12000×g 下离心 10 分钟来澄清裂解的细胞,并且丢弃上清液。将细胞沉淀物在含有 0.5% TritonX-100 的 PBS 缓冲液 (pH7.0) 中洗涤三次,接着用 PBS (pH7.0) 洗涤一次。然后,使用 PBS 中的 8M 脲 (pH8.0) 将沉淀物再洗涤一次。通过在 12000×g 下

离心来澄清每个洗涤步骤,并且丢弃上清液。然后将沉淀物溶解于含有 10mM TRIS(pH8.0)、8M 脲、100mM NaCl 的增溶缓冲液中,在室温下过夜。将增溶缓冲液在 12000×g 下进一步离心,通过 0.45nm 膜过滤,并且装载于用含有 PBS(pH7.0)、8M 脲和 20mM 咪唑的洗涤缓冲液预先洗涤过的 NiSepharose 柱上。装载后,用 10 柱体积的洗涤缓冲液洗涤柱子,并且将溶解的蛋白质在补充 500mM 咪唑的洗涤缓冲液中洗脱。将洗脱的蛋白质脱盐至 PBS(pH7.0)和 8M 脲缓冲液中,并且在含有 50mM 乙酸钠 (pH4.5)、0.1M NDSB-201、10%甘油、1mM/0.1mM GSH/GSSH 的重折叠缓冲液中通过逐滴稀释来重折叠。通过在 12000×g 下离心来澄清重折叠缓冲液,并且使用具有 10KDa 标称分子截留的 Centriprep 滤器 (Millipore) 来浓缩重折叠蛋白。使用 G25 脱盐柱将重折叠蛋白缓冲液交换至补充了 10%甘油的含有 100mM 乙酸钠的缓冲液 (pH5.0) 中,并且将等份试样存储在 -80°C。考马斯染色的 10% SDS PAGE 凝胶和 Western 印迹分析总体表明了高达 95%纯度的 18KDa 的 His- 标记的蛋白质的存在。切除 18KDa 考马斯染色带并通过 MALDI-TOF/TOF MS/MS 鉴定含有 ZG16。

[0171] 通过用纯化的 ZG16 蛋白淘选噬菌体展示抗体文库来获得抗 ZG16 的抗体 (Antibodies by Design ;Morphosys AG 的一个分部,德国。 www.morphosys.com)。

[0172] 抗体阵列

[0173] 使用抗体阵列来证实候选标志物。血清样品获自胃癌、结直肠癌 (外科手术前后) 患者,和获自非恶性疾病的外科手术患者。通过 Dunedin Public Hospital,新西兰,和 Christchurch Cancer Society 组织库,Christchurch,新西兰,来制备样品。使用 GeneMachines OminGrid 100 阵列机器人将获自商业来源或选自噬菌体文库 (Morphosys) 的抗 ZG16 和 MUC17 的抗体印刷在玻璃载玻片上 (Schott Nexterion Slide H)。用疏水笔将每个阵列圈上范围。然后将载玻片在 3×PBS-0.5%吐温 20 (3×PBS-T) 中洗涤,接着用 50mM 硼酸钠缓冲液中的 50mM 乙醇胺, pH8.0 来封闭,接着用酪蛋白酸盐封闭缓冲液 (3×PBS-T, 1%酪蛋白酸钠) 封闭。然后将生物素标记的血清样品添加至载玻片上,接着在 4°C 下温育过夜。然后将载玻片在 3×PBS-T 中洗涤,接着空气干燥。然后使用滚环扩增 (RCA) 来检测结合的抗体,之前有大量描述 (Haab BB, Lizardi PM, RCA-enhanced protein detection arrays (RCA 增强的蛋白质检测阵列), Methods Mol Biol. 2006 ;328 :15-29)。简言之,用已经缀合寡核苷酸引物 (5'-CCT GGT GCT CAA ATT TCA GTT CTG C-3') 的抗生物素抗体温育载玻片。然后在潮湿的密封室中在 37°C 下将环状 DNA 模板与载玻片杂交 30min,接着将载玻片在递减浓度的 PBS-T (3×PBS-0.05%吐温 20, 1×PBS-0.05%吐温 20 和 0.1×PBS-0.05%吐温 20) 中洗涤并干燥。然后使用 phi29 在 30°C 下将模板延伸 3hr,接着洗涤载玻片,并通过离心干燥。然后使用同源荧光标记的探针来检测扩增的模板。用 Axon 4000A 扫描仪扫描载玻片,并用 GenePix Pro 6.1.0.4 软件测量信号。

[0174] 使用来自 R 的 limma (Smith, 2005) 包 (用于统计计算的 R 包 (R Development Core) 的 normalizeBetweenArrays 函数,使用分位数归一化来调节 Cy5 荧光强度。分位数归一化调节强度值,通过将来自不同区组的强度分位数设定至相同值,使得强度分布对于每个区组是相同的 (每个区组对应于单独的样品)。每个强度值的等级在该程序过程中没有改变,只有强度的相对量值发生了改变。设想描述抗原浓度范围的潜在概率分布函数对于所有样品是相同的。该程序提高了所有样品见的区组之间信号的平均相关性,并且还在考虑仅参照的区组时,其表明了数据质量的改进。在采取重复之间的中值来获得每种抗体

的归一化强度之前,除去 Genepix- 标记的点。

[0175] 因此,我们已经鉴定出对于研发试剂、检测和评价胃癌的装置和试剂盒有用的三种基因和 / 或蛋白质。可以使用一个或多个胃癌标志物,单独或结合使用,以提供用于胃癌的可靠分子测试。

实施例

[0176] 在此所述的实施例是为了说明本发明的实施方案的目的。其他实施方案、方法和分析类型在分子诊断领域的普通技术人员的能力范围内,并且在本文中不需要详细描述。认为本领域范围内的其他实施方案是本发明的一部分。

[0177] 实施例 1 :胃恶性肿瘤的标志物的鉴定

[0178] 使用获自胃肿瘤和非恶性样品的基因表达数据来选择标志物。将以下标准用于标志物选择:(i) 分泌蛋白特征性的信号序列的存在,(ii) 肿瘤组织中微阵列信号强度等级和(iii) 血液或脉管组织中相应 EST 的水平。这些标准的使用能够鉴定在肿瘤组织中大量表达但可能在血清、血液或血浆中具有低背景的分泌标志物。图 1 描绘了一张表,其显示了使用以上标准选择的胃恶性肿瘤的三个标志物,MUC5AC,MUC17 和 ZG16。图 1 包括基因符号(“符号”)、MWG 寡聚物号、NCBI mRNA 参照序列号、蛋白质参照序列号、使用肿瘤组织衍生的阵列上基因的等级强度和使用非恶性组织衍生的阵列上基因的等级强度。所有三个 GTM 具有比 CEACAM5 高的表达(强度)等级,CEACAM5 是编码现有胃癌标志物 CEA 的基因。最低的表达等级可能是 29,718。等级的检查还显示出这些 GTM 在肿瘤组织中的表达与非恶性组织相当,表明基因在癌形成过程中未得到强烈下调。血液和脉管组织中这三个 GTM 的 unigene EST 计数(Wheeler 等,2003)全部为零。

[0179] 实施例 2 :qRT-PCR 分析

[0180] 使用更灵敏和精确的基因表达定量技术 qPCR,在肿瘤组织中证实了 GTM ZG16 和 MUC17 的丰度和身份。使用方法部分中所述的引物和探针,通过 RT-qPCR 分析了来自相同患者的高达 45 个胃肿瘤样品和等量的非恶性胃组织样品。使用达到产物扩增阈值水平(Ct)所需的 PCR 循环数来定量这些基因的表达。

[0181] qPCR 分析证实了阵列数据:通过 qPCR 容易地在肿瘤组织中检测到两个标志物,并且不存在肿瘤组织与非恶性组织相比表达显著下降的证据。通过图 2a-b 中的直方图说明了这些 RNA 在肿瘤组织与非恶性组织相比的丰度。

[0182] 实施例 3 :血清中胃肿瘤标志物蛋白的检测

[0183] 在特定的实施方案中,可以使用针对完整蛋白质、蛋白质片段(肽)或蛋白质核心的抗体来完成 GTM 蛋白的检测。检测和定量蛋白质和肽表达的方法是本领域已知的,并且可以包括依赖于相对蛋白质或肽产生的特异性抗体的方法。可以使用本领域公知的方法来制备单克隆抗体和多克隆抗血清,并且在本文中不需要更多描述。

[0184] 为了检测血清中的 GTM,使用 Gene Machine OmniGrid™ 机器人技术将抗 GTM 的抗体印刷在玻璃载玻片上。将每个抗体在阵列上重复 8 次。然后在用抗体载玻片温育之前,用生物素标记来自 33 名胃癌患者和 41 名对照的血清样品。用抗-生物素抗体检测结合的蛋白,并且使用滚环扩增(RCA)和荧光标记来扩增信号。使用 Axon 4000a 扫描仪和 Genepix 6.1.0.4 软件来定量结合的蛋白质的含量。患者的特征显示于图 2 中。

[0185] 将来自阵列上每个抗体的荧光信号归一化,并且将 8 个重复的中值信号以任意荧光单位来表示。说明数据分布的盒形图显示于图 3 中。对于 MUC17 的中值信号,胃癌患者为 18,836AU,对照组为 16,130。这些中值是显著差异的 ($p = 0.007$)。对于获自 MorphoSys 的两个噬菌体展示 ZG16 抗体 (5902 和 5905),观察到中值之间的显著差异。胃癌患者样品中 ZG16_5902 的中值信号为 2139AU,相比,对照的为 1837AU;患者中的中值 ZG16_5905 信号为 3063AU,相比,对照的为 1675AU。患者和对照之间对于 ZG16_5902 和 ZG16_5905 的中值信号是显著不同的(各自为 $p = 0.05$ 和 $p = 0.005$)。

[0186] 该数据证明 MUC17 和 ZG16 以显著高于对照的水平存在于胃癌患者的血清中。患者和对照组之间的更多区别将通过免疫学测试程序的改进、对目标抗原具有更高特异性的抗体的鉴定和标志物的联合使用来实现。

[0187] 实施例 8:用含有 GTM 的载体转染的细胞

[0188] 在进一步的实施方案中,提供了可以表达 GTM、GTM 片段或肽标志物的细胞。因此可以使用原核细胞和真核细胞。例如,大肠杆菌(原核细胞)可以用来产生大量缺乏成熟糖基化的 GTM(如果特定 GTM 通常是糖基化的)。COS 细胞、293 细胞和多种其他真核细胞可以用于产生糖基化的 GTM,或具有正确的折叠,并且因此,具有天然形式的 GTM 蛋白的三维结构。转染这些细胞的方法是本领域已知的,并且在本文中不需要更多描述。

[0189] 实施例 9:试剂盒

[0190] 基于本发明的发现,可以产生几种类型的测试试剂盒。首先,可以制备具有预先装载了检测分子(或“捕获试剂”)的检测装置的试剂盒。在用于检测 GTM mRNA 的实施方案中,这样的装置可以包括基质(例如,玻璃,硅,石英,金属等),其上的寡核苷酸作为与待检测的 mRNA 杂交的捕获试剂。在一些实施方案中,可以通过 mRNA(用 cy3、cy5、放射性标记或其他标记来标记)与基质上的寡核苷酸杂交来完成 mRNA 的直接检测。在其他实施方案中,可以通过首先制备所需 mRNA 的互补 DNA(cDNA)来完成 mRNA 的检测。然后,可以将标记的 cDNA 与基质上的寡核苷酸杂交,并检测。

[0191] 与所用的检测方法无关,测试 GTM 表达与表达的标准测量的比较是所希望的。例如,可以将 RNA 表达对总细胞 DNA、组成型表达的 RNA(例如,核糖体 RNA)的表达或其他相对不变的标志物标准化。

[0192] 抗体也可以用于试剂盒中作为捕获试剂。在一些实施方案中,基质(例如,多孔平板)可以具有与其连接的特异性 GTM 捕获试剂。在一些实施方案中,试剂盒可以包括封闭试剂。封闭试剂可以用于降低非特异性结合。例如,可以使用来自不含有 GTM 寡核苷酸的任何常规来源的过量 DNA(如,鲑精 DNA)来降低非特异性的寡核苷酸结合。可以使用过量的封闭蛋白(如,血清白蛋白)来降低非特异性的抗体结合。可以认识到用于检测寡核苷酸和蛋白质的多种方法是本领域已知的,并且可以使用特异性地检测 GTM 相关分子的任何策略,并且认为在本发明的范围内。

[0193] 在依赖于抗体检测的实施方案中,可以基于每个细胞,或基于全部细胞、组织或液体蛋白、液体体积、组织质量(重量)来表示 GTM 蛋白或肽。此外,可以基于相对高丰度血清蛋白(如,白蛋白)来表示血清中的 GTM。

[0194] 除了基质,测试试剂盒可以包括捕获试剂(如,探针)、洗涤溶液(例如,SSC、其他盐、缓冲液、去污剂等)以及检测部分(例如,cy3、cy5、放射性标记等)。试剂盒还可以包

括使用说明书和包装。

[0195] 尽管参照其特定实施方案描述了本发明, 但将认识到可以使用涉及使用所公开标志物的其他实施方案, 而没有脱离本发明的范围。

[0196] 工业实用性

[0197] 检测 GTM 家族成员的方法包括使用微阵列和 / 或实时 PCR 方法来检测核酸, 以及检测蛋白质和肽。本发明的组合物和方法可用于诊断装置和试剂盒的制备、疾病的诊断、评价疗效和生产适用于测量生物样品中 GTM 家族成员表达的试剂。

[0198] 参考文献

[0199] Emanuelsson O, Nielsen H, Rrunak S, von Heine G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol. 2000 Jul 21 ;300(4) :1005-16.

[0200] Krogh A, Larsson B, von Heine G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model :application to complete genomes. J Mol Biol. 2001 Jan 19 ;305(3) :567-80.

[0201] Smyth, G. K. (2005). Limma :linear models for microarray data. In: ' Bioinformatics and Computational Biology Solutions using R and Bioconductor' . R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, W. Huber (eds), Springer, New York, pages 397--420.

[0202] R Development Core Team(2008). R :A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

[0203] Wheeler DL, et al. Database Resources of the National Center for Biotechnology. Nucl Acids Res 31 :28-33 ;2003.

[0001]

序列表

<110> 环太平洋生物技术有限公司

<120> 胃癌的检测标志物

<130> 231672PCT

<160> 6

<170> PatentIn 版本 3.3

<210> 1

<211> 18750

<212> DNA

<213> 人

<400> 1

```

atgtcgtgc agttcaatgg gtttgtgctg tgtgatcaca cggtcctgcc cacagagctg      60
aatgggctat tggggtttcc ttggaggctg ctgagggaca gggcactctt ccccgccgtc      120
cacacaatga gtgttgccg gaggaagctg gccctgctct gggccctggc tctcgtctg      180
gcctgcaccc ggcacacagg ccatgccag gatggctcct ccgaatccag ctacaagcac      240
cacctgccc tctctcctat cgcccggggg cccagcgggg tcccgtccg tggggcgact      300
gtttcccat ctctgaggac catccctgtg gtacgagcct ccaaccggc gcacaacggg      360
cgggtgtgca gcacctggg cagcttccac tacaagacct tcgacggcga cgtcttccgc      420
ttccccggc tctgcaacta cgtgttctcc gagcactgct gtgccgccta cgaggatitt      480
aacatccagc tacgccgag ccaggagtca gcggcccca cgctgagcag ggtcctcatg      540
aagggtgatg gcgtggtcat ccagctgacc aagggtccg tcctggtcaa cggccacccg      600
gtcctgctgc cttcagcca gtctggggtc ctattcagc agagcagcag ctacaccaag      660
gtggaggcca ggctgggct tgcctcatg tggaccacg atgacagcct gctgctggag      720
ctggacacca aatacgcca caagacctgt gggctctgtg gggacttcaa cgggatgccc      780
gtggtcagcg agctcctctc ccacaacacc aagctgacac ccatggaatt cgggaacctg      840
cagaagatgg acgacccac ggagcagtgt caggaccctg tcctgaacc cccgaggaac      900
tgctccactg gctttggcat ctgtgaggag ctctgcacg gccagctgtt ctctggctgc      960
gtggccctgg tggacgtcg cagctacctg gaggcttgca ggcaagacct ctgcttctgt     1020
gaagacaccg acctgctcag ctgctctgc cacaccctg ccgagtactc ccggcagctg     1080
accatgcag gggggttgcc ccaggactgg cggggcctg acttctgccc ccagaagtgc     1140
ccaacaaca tgacgtacca cgagtgccg tccccctgc cagacacctg ctccaaccag     1200
gagcactccc gggcctgtga ggacctgt gtggccggct gcttctgccc tgaggggacg     1260
gtgcttgacg acatcgcca gaccggctgt gtccctgtgt caaagtgtgc ctgctctac     1320
aacggggctg cctatgccc aggggccacc tactccacag actgcaccaa ctgcacctgc     1380
tccggaggcc ggtggagctg ccaggaggtt ccatgcccg gtacctgctc tgtgcttggg     1440

```

[0002]

ggtgcccact	tctcaacgtt	tgacgggaag	caatacacgg	tgacacggcga	ctgcagctat	1500
gtgctgacca	agccctgtga	cagcagtgcc	ttcactgtac	tggctgagct	gcgcaggtgc	1560
gggctgacgg	acagcgagac	ctgcctgaag	agcgtgacac	tgagcctgga	tggggcgag	1620
acgggtggtg	tgatcaaggc	cagtggggaa	gtgttcctga	accagatcta	caccagctg	1680
cccatctctg	cagccaacgt	caccatcttc	agaccctcaa	ccttcttcat	catcgcccag	1740
accagcctgg	gcctgcagct	gaacctgcag	ctggtgccca	ccatgcagct	gttcatgcag	1800
ctggcgccca	agctccgtgg	gcagacctgc	ggtctctgtg	ggaacttcaa	cagcatccag	1860
gccgatgact	tccggaccct	cagtgggggtg	gtggaggcca	ccgctgcggc	cttcttcaac	1920
accttcaaga	cccaggccgc	ctgccccaac	atcaggaaca	gcttcgagga	cccctgctct	1980
ctgagcgtgg	agaatgcccc	caggatgggt	gccccgagcg	cgtagccggac	gctggtgttg	2040
gctctggcgg	ccatgctcgt	ggtgccgcag	gcagagaccc	agggccctgt	ggagccgagc	2100
tgggagaatg	cagggcacac	catggatggc	ggtgccccga	cgctctcgcc	caccggcgc	2160
gtgagctttg	ttccaccctg	caactgtctt	cccagcctga	gccgtaagca	gatgctgcc	2220
ctgccagccg	ggaagggggt	gtttgccagt	cccaaagggt	ggggcccaga	tctaggggtg	2280
cagctgccac	cagccctgaa	cccggcgcac	aatgggcggg	tgtgcagcac	ctggggtgac	2340
ttccactaca	agaccttga	cggcgacgtc	ttccgcttcc	ctggcctttg	caactacgtg	2400
ttctctgagc	actgccgcgc	cgctacgag	gacttcaacg	tccagctacg	ccgaggccta	2460
gtgggctcca	ggcctgtggt	caccctgtgt	gtcatcaagg	cccaggggct	ggtgctggag	2520
gcgtccaacg	gctccgtcct	catcaatggg	cagcgggagg	agctgcctta	cagccgcact	2580
ggcctcctgg	tggagcagag	cggggactac	atcaaggctca	gcatccggct	ggtgctgaca	2640
ttcctgtgga	acggagagga	cagtgccttg	ctggagctgg	atcccaaata	cgccaaccag	2700
acctgtggcc	tgtgtgggga	cttcaacggc	ctcccggcct	tcaacgagtt	ctatgccac	2760
aacgccaggc	tgaccccgct	ccagtttggg	aacctgcaga	agttggatgg	gcccacggag	2820
cagtgcccg	acccgtgcc	cttgccggcc	ggcaactgca	cggacgagga	gggcatctgc	2880
caccgcacc	tgctggggcc	ggcctttgcg	gagtgccacg	caactggtga	cagcactgcg	2940
tacctggccg	cctgcgcca	ggacctgtgc	cgctgcccc	cctgcccgtg	tgccaccttt	3000
gtggaatact	cacgccagtg	cgccccagcg	gggggcccagc	cgcggaactg	gaggtgcct	3060
gagctctgcc	cccggacctg	ccccctcaac	atgcagcacc	aggagtgtgg	ctcaccctgc	3120
acggacacct	gctccaacc	ccagcgcgcg	cagctctgcg	aggaccactg	tgtggacggc	3180
tgcttctgcc	cccaggcag	gtcttgacg	gtgctggatg	acatcacgca	ctctggctgc	3240
ctgcccctcg	ggcagtgcc	ctgcacccac	ggcggccgca	cctacagccc	gggcacctcc	3300
ttcaacacca	cctgcagctc	ctgcacctgc	tccggggggc	tatggcagtg	ccaggacctg	3360
ccgtgccctg	gcacctgctc	tgtgcagggc	ggggcccaca	tctccaccta	tgatgagaaa	3420
ctctacgacc	tgcatggtga	ctgcagctac	gttctgtcca	agaaatgtgc	cgacagcagc	3480
ttcaccgtgc	tggctgagct	gcggaagtgc	ggcctgacgg	acaacgagaa	ctgcctgaaa	3540

[0003]

gcggtgacgc	tcagcctgga	cggcggggac	acggccatcc	gggtccaagc	ggacggcggc	3600
gtgttcctca	actccatcta	cacgcagctg	cccctgtcgg	cagccaacat	caccctgttc	3660
acaccctcga	gcttcttcat	cgtggtgcag	acaggcctcg	ggctgcagct	gctggtgcag	3720
ctggtgccac	tcatgcaggt	gtttgtcagg	ctggaccccc	cccaccaggg	ccagatgtgc	3780
ggcctgtgtg	ggaacttcaa	ccagaaccag	gctgacgact	tcacggcctc	cagcggggtg	3840
gtggaggcca	cgggcgcagc	cttcgccaac	acctggaagg	cccaggctgc	ctgtgccaat	3900
gccaggaaca	gctttgagga	cccctgctcc	ctcagtgtgg	agaatgagaa	ctacgcccgg	3960
cactggtgct	cgcgccctgac	cgatcccaac	agtgcccttct	cgcgctgcca	ctccatcatc	4020
aaccccaagc	ccttccactc	gaactgcatg	tttgacacct	gcaactgtga	gcggagcag	4080
gactgcctgt	gcgcccgcgt	gtcctcctac	gtgcacgcct	gtgccgcaa	ggcgtacag	4140
ctcagcgact	ggagggacgg	cgtctgcacc	aagtacatgc	agaactgcc	caagtcccag	4200
cgctacgcct	acgtggtgga	tgctgccag	cccacttgcc	gcgccctgag	tgaggccgac	4260
gtcacctgca	gcgtttcctt	cgtgcctgtg	gacggctgca	cctgccccgc	gggcaccttc	4320
ctcaatgacg	cgggcgcctg	tgtgcccgcc	caggagtgcc	cctgctacgc	tcacggcacc	4380
gtgctggctc	ctggagaggt	ggtgcacgac	gagggcgccg	tgtgttcatg	tacgggtggg	4440
aagctaagct	gcctgggagc	ctctctgcag	aaaagcacag	ggtgtgcagc	ccccatggtg	4500
tacctggact	gcagcaacag	ctcggcgggc	accctgggg	ccgagtgcct	ccggagctgc	4560
cacacgctgg	acgtgggctg	tttcagcaca	cactgcgtgt	ccggctgtgt	ctgtcccccg	4620
gggctggtgt	cggatgggag	tgggggctgc	attgccgagg	aggactgcc	ctgtgtgcac	4680
aacgaggcca	cctacaagcc	tggagagacc	atcagggtcg	actgcaacac	ctgcacctgc	4740
aggaaccgga	ggtgggagtg	cagccaccgg	ctctgcctgg	gcacctgcgt	ggcctacggg	4800
gatggccact	tcatcacctt	tgatggcgat	cgctacagct	ttgaaggcag	ctgcgagtac	4860
atcttggccc	aggactactg	tggggacaac	accaccacg	ggaccttccg	catcgtcacc	4920
gagaacatcc	cctgtgggac	caccggcacc	acctgctcca	aggccatcaa	gctcttcgtg	4980
gaggtgagaa	cggccccagc	tgtgagcacc	cccgaccctg	cagccaacga	gccggcccc	5040
agggaaagctt	cggtcggctt	ccggcagcgt	ctgcctcccc	tgcaagctta	cgagctgatc	5100
ctccaagagg	ggacctttaa	ggcgggtggc	agagggccgg	gtggggacc	accctacaag	5160
atacgtaca	tggggatctt	cctggctcct	gagaccacg	ggatggccgt	gtcctgggac	5220
cggaagacca	gcgtgttcat	ccgactgcac	caggactaca	aggcagggt	ctgcggcctg	5280
tgcggaact	tcgacgaaa	tgccatcaat	gactttgcca	cgcgtagccg	gtccgtggtg	5340
ggggacgcac	tggagtttg	gaacagctgg	aagctctccc	cctcctgcc	ggacgcctg	5400
gcaccaaggg	accctgcac	ggccaacccc	ttccgcaagt	cctgggcca	gaagcagtc	5460
agcatcctcc	acggccccac	cttcgccc	tgccgctccc	aggttgactc	caccaagtac	5520
tacgaggcct	gcgtgaacga	cgcggtgccc	tgcgactcgg	gtggcgactg	cgagtgttc	5580

[0004]

tgcacggctg	tggctgccta	cgcccaggcc	tgccacgacg	cgggcctgtg	tgtgtcctgg	5640
cggactccgg	acacctgccc	cttgttctgt	gacttctaca	accacatgg	gggctgtgag	5700
tggcactacc	agccctgcgg	ggcaccctgc	ctaaaaacct	gccggaaccc	cagtgggcac	5760
tgctgtgtg	acctgcctgg	cctggaaggc	tgctaccgga	agtgccacc	cagccagccc	5820
ttcttcaatg	aggaccagat	gaagtgcgtg	gcccagtgtg	gctgctacga	caaggacgga	5880
aactactatg	acgtcgggtg	aagggtcccc	acagcggaga	actgccagag	ctgtaactgc	5940
acaccagtg	gcatccagtg	cgctcacagc	cttgaggcct	gcacctgcac	ctatgaggac	6000
aggacctaca	gctaccagga	cgctcatctac	aacaccaccg	atgggcttgg	cgctgtcttg	6060
atcgccatct	gcggaagcaa	cggcaccatc	atcaggaagg	ctgtggcatg	tcctggaact	6120
ccagccacaa	cgccattcac	cttcaccacc	gcctgggtcc	cccactccac	gacaagcccc	6180
gccctcccgg	tctccaccgt	gtgtgtccgc	gaggtctgcc	gctggtccag	ctggtacaat	6240
gggcaccgcc	cagagcccgg	cctgggaggc	ggagactttg	agacgtttga	aaacctgagg	6300
cagagagggg	accaggatg	ccctgtgtg	gctgacatcg	agtgccgggc	ggcgcagctt	6360
cccgatgc	cgctggagga	gctgggccag	caggtggact	gtgaccgat	gcgggggctg	6420
atgtgcgcca	acagccaaca	gagtcccccg	ctctgtcacg	actacgagct	gcgggttctc	6480
tgctgcgaat	acgtgcctg	tgccccctcc	ccggccccag	gcaccagccc	tcagccctcc	6540
ctcagtgcc	gcacggagcc	tgctgtgct	accccaaccc	agaccacagc	aaccgaaaag	6600
accaccctat	gggtgacccc	gagcatccgg	tcgacggcgg	ccctcacctc	gcagactggg	6660
tccagctcag	gccccgtgac	ggtcaccccc	tcggccccag	gtaccaccac	ctgccagccc	6720
cggtgtcagt	ggacagagtg	gtttgatgag	gactacccca	agtctgaaca	acttgagggg	6780
gacgttgagt	cctacgataa	gatcagggcc	gctggagggc	acttatgcc	gcagcctaag	6840
gacatagagt	gccaggccga	gagcttcccc	aactggacc	tggcacaggt	ggggcagaag	6900
gtgactgtg	acgtccactt	cggcctggtg	tgccaggaact	gggagcagga	gggcgtcttc	6960
aagatgtgct	acaactacag	gatccgggtc	ctctgctgca	gtgacgacca	ctgcagggga	7020
cggtgccaaa	ccccgccacc	gaccacagag	ctggagacgg	ccaccaccac	caccaccag	7080
gccctgttct	caacgccgca	gcctacgagt	agcccggggc	tgaccagggc	tccccggcc	7140
agcaccacag	cagtccccac	cctctcagaa	ggactgacat	ccccagata	cacaagcacc	7200
cttggtacag	ccaccacggg	aggccccacg	acgcctgcag	gctccacaga	accactgtc	7260
ccaggggtgg	ccacatccac	ccttccaaca	cgctcagccc	ttccagggac	gacggggagc	7320
ttgggacat	ggcggccctc	acagccaccc	acgtggccc	caacaacaat	ggcaacctcc	7380
agagctcgcc	cgacaggcac	agccagcacc	gcttccaaag	agccgctgac	cacgagcctg	7440
gcgccaacac	tcacgagcga	gctgtccacc	tctcaggccg	agaccagcac	gcccaggaca	7500
gagacgacaa	tgagcccctt	gactaacacc	accaccagcc	agggcacgac	ccgctgtcaa	7560
ccgaagtgtg	agtggacaga	gtggtttgac	gtggacttcc	caacctcagg	ggttgaggcc	7620
ggggacatgg	aaacttttga	aaacatcagg	gctgctgggg	gcaagatgtg	ctgggcacca	7680

[0005]

aagagcatag	agtgccgggc	ggagaactac	cccgaggtaa	gcatcgacca	ggtcgggcag	7740
gtgctgacct	gcagcctgga	gacggggctg	acctgcaaga	acgaagacca	gacaggcagg	7800
ttcaacatgt	gcttcaacta	caacgtgctg	gtgctttgct	gtgacgacta	cagccactgc	7860
cccagtacc	cagccaccag	ctccacggcc	acgccctcct	caactccggg	gacgacctgg	7920
atcctcacia	agccgaccac	aacagccact	acgactgcgt	ccactggatc	cacggccacc	7980
ccgacctcca	ccctgagaac	agctccccct	cccaaagtgc	tgaccaccac	ggccaccaca	8040
cccacagtca	ccagctccaa	agccactccc	tcctccagtc	cagggactgc	aaccgccctt	8100
ccagcactga	gaagcacagc	caccacacc	acagctacca	gcgttacacc	catccccctt	8160
tcctccctgg	gcaccacctg	gacccgccta	tcacagacca	ccacaccac	ggccaccatg	8220
tccacagcca	caccctcctc	cactccagag	actgccca	cctccacagt	gcttaccgcc	8280
acggccacca	caactggggc	caccggctct	gtggccacc	cctcctccac	cccaggaaca	8340
gctcacacta	ccaaagtgcc	aactaccaca	accacgggt	tcacagccac	ccccctctcc	8400
agcccaggga	cggcactcac	gcctccagtg	tggatcagca	caaccaccac	accacaacc	8460
agaggctcca	cggtgacccc	ctcctccatc	ccggggacca	cccacaccgc	cacagtgctg	8520
accaccacca	ccacaactgt	ggccactggt	tctatggcaa	caccctcctc	tagcacacag	8580
accagtggga	ccaccacac	acccccagtg	ccgaacacca	tggccaccac	acacgggcga	8640
tccttgcccc	ccagcagtc	ccacacggtg	cgcacagcct	ggacttcggc	cacctcgggc	8700
atcttgggca	ccaccacat	cacagagcct	tccacggatga	cttccacac	cctagcagca	8760
accaccggta	ccaccagca	ctcgactcca	gccctttcca	gccctcacc	tagcagcaga	8820
accaccgagt	cacccccctc	tccagggagc	accaccccg	gccacaccac	ggccacctcc	8880
aggaccacag	ccacggccac	accagcaag	acccgcacct	cgaccctgct	gcccagcagc	8940
cccacatcgg	ccccataac	cacggtggtg	accatgggct	gtgagccca	gtgtgcctgg	9000
tcagagtggc	tggactacag	ctaccccatg	ccggggccct	ctggcgggga	ctttgacacc	9060
tactccaaca	tccgtgcggc	cggagggggc	gtctgtgagc	agcccctggg	cctcgagtgc	9120
cgtgcccagg	cccagcctgg	tgtccccctg	cgggagttgg	gccaggctctg	ggaatgcagc	9180
ctggactttg	gcctggtctg	caggaaccgt	gagcaggtgg	ggaagttaa	gatgtgcttc	9240
aactatgaaa	tccgtgtggt	ctgctgcaac	tacggccact	gccccagcac	cccggccacc	9300
agctctacgg	ccatgccctc	ctccactccg	gggacgacct	ggatcctcac	agagctgacc	9360
acaacagcca	ctacgactga	gtccactgga	tccacggcca	ccccgtcctc	caccccaggg	9420
accacctgga	tcctcacaga	gccgagcact	acagccaccg	tgacggtgcc	caccgatcc	9480
acggccaccg	cctcctccac	ccaggcaact	gctggcacc	cacatgtgag	caccacggcc	9540
acgacacca	cagtcaccag	ctccaaagcc	actcccttct	ccagtccagg	gactgcaacc	9600
gcccttccag	cactgagaag	cacagccacc	acaccacag	ctaccagctt	tacagccatc	9660
ccctcctcct	ccctgggcac	cacctggacc	cgctatcac	agaccaccac	accacggcc	9720

[0006]

accatgtcca	cagccacacc	ctcctccact	ccagagactg	tccacacctc	cacagtgttt	9780
accaccacgg	ccaccacaac	cggggccacc	ggctctgtgg	ccacccccctc	ctccacccca	9840
ggaacagctc	acactaccaa	agtgtgtact	accacaacca	cgggcttcac	agccaccccc	9900
tcctccagcc	cagggacggc	acgcacgctt	ccagtgtgga	tcagcacaac	caccacaccc	9960
acaaccagag	gttccacggt	gacccccctc	tccatccccg	ggaccaccca	cacccccaca	10020
gtgtgtacca	ccaccaccac	aactgtggcc	actggttcta	tggcaacacc	ctcctctagc	10080
acacagacca	gtgggaccac	ccacacaccc	ccagtgccga	acaccacggc	caccacacac	10140
gggcatccc	tgtccccag	cagtccccac	acgggtgcga	cagcctggac	ttcggccacc	10200
tcaggcacct	tgggcaccac	ccacatcaca	gagccttcca	cggggacttc	ccacacccca	10260
gcagcaacca	ccggtaccac	ccagcactcg	actccagccc	tgtccagccc	tcaccctagc	10320
agcaggacca	ccgagtcacc	cccttctcca	gggacgacca	ccccgggcca	caccagggcc	10380
acctccagga	ccacggccac	ggccacaccc	agcaagacc	gcacctgac	cctgtgtccc	10440
agcagcccca	catcggcccc	aataaccacg	gtggtgacca	tgggctgtga	gccccagtgt	10500
gcctggtcag	agtggctgga	ctacagctac	cccatgccgg	ggccctctgg	cggggacttt	10560
gacacctact	ccaacatccg	tgcggccgga	ggggccgtct	gtgagcagcc	cctgggcctc	10620
gagtgccgtg	cccaggccca	gcctggtgtc	cccctgcggg	agttgggcca	ggtcgtggaa	10680
tgcagcctgg	actttggcct	ggtctgcagg	aaccgtgagc	agtggggaa	gttcaagatg	10740
tgtttcaact	atgaaatccg	tgtgttctgc	tgcaactacg	gccactgccc	cagcaccctg	10800
gccaccagct	ctacggccac	gccctcctcc	actccagggg	cgacctggat	cctcacagag	10860
cagaccacag	cagccactac	gaccgcaacc	actggatcca	cggccatccc	gtcctccacc	10920
ccgggaacag	ctccccctcc	caaagtgtctg	accagcacgg	ccaccacacc	cacagccacc	10980
agttccaaag	ccacttctc	ctccagtcca	aggactgcaa	ccacccttcc	agtgtgtgaca	11040
agcacagcca	ccaaatccac	agctaccagc	tttacacca	tcccctcctt	cacccttggg	11100
accaccggga	ccctcccaga	acagaccacc	acacccatgg	ccaccatgtc	cacaatccac	11160
ccctcttcca	ctcgggagac	caccacaccc	tccacagtgc	tgaccacgaa	ggccaccacg	11220
acaagggcca	ccagttccat	gtccaccccc	tcttccactc	cggggacgac	ctggatcctc	11280
acagagctga	ccacagcagc	cactacaact	gcagccactg	gccccacggc	caccccgtcc	11340
tccaccccag	ggaccacctg	gatcctcaca	gagcccagca	ctacagccac	cgtgacggtg	11400
cccaccggat	ccacggccac	cgctcctcc	acccgggcaa	ctgctggcac	cctcaaagtg	11460
ctgaccagca	cggccaccac	accacagtc	atcagctcca	gagccactcc	ctcctccagt	11520
ccagggactg	caaccgccct	tccagcactg	agaagcacag	ccaccacacc	cacagctacc	11580
agcgttacag	ccatccccctc	ttctcctctg	ggcaccgcct	ggacccgcct	atcacagacc	11640
accacaccca	cggccaccat	gtccacagcc	acaccctcct	ctactccaga	gactgtccac	11700
acctccacag	tgcttaccac	cacgaccacc	acaaccaggg	ccaccggctc	tgtggccacc	11760
ccctcctcca	cccaggaac	agctcacact	accaaagtgc	cgactaccac	aaccacgggc	11820

[0007]

ttcacagcca	ccccctcctc	cagcccaggg	acggcactca	cgctccagt	gtggatcagc	11880
acaaccacca	caccacacaac	cagaggctcc	acggtgacct	cctcctccat	cccggggacc	11940
accacacccg	ccacagtgt	gaccaccacc	accacaactg	tggccactgg	ttctatggca	12000
acacctcct	ctagcacaca	gaccagtggg	accaccaca	cacccccagt	gccgaacacc	12060
acggccacca	cacacgggcg	gtccctgccc	cccagcagtc	cccacacggt	gcgcacagcc	12120
tggacttcgg	ccacctcggg	catcttgggc	accaccaca	tcacagagcc	ttccacggtg	12180
acttcccaca	ccccagcagc	aaccaccagt	accaccagc	actcgactcc	agccctgtcc	12240
agccctcacc	ctagcagcag	gaccaccgag	tcacccccct	ctccagggac	gaccaccccg	12300
ggccacacca	ggggcacctc	caggaccaca	gccacagcca	caccagcaa	gaccgcacc	12360
tcgaccctgc	tgcccagcag	ccccacatcg	gccccataa	ccacggtggt	gaccacgggc	12420
tgtgagcccc	agtgtgcctg	gtcagagtgg	ctggactaca	gctaccccat	gccggggccc	12480
tctggcgggg	actttgacac	ctactccaac	atccgtgcgg	ccggaggggc	agtctgtgag	12540
cagcccctgg	gcctcgagtg	ccgtgcccag	gcccagcctg	gtgtccccct	gcgggagttg	12600
ggccaggtcg	tggaatgcag	cctggacttt	ggcctggtct	gcaggaaccg	tgagcaggtg	12660
gggaagttca	agatgtgctt	caactatgaa	atccgtgtgt	tctgtgcaa	ctacggccac	12720
tgcccagca	ccccggccac	cagctctacg	gccacgcct	cctcaactcc	ggggacgacc	12780
tggatcctca	caaagctgac	cacaacagcc	actacgactg	agtccactgg	atccacggcc	12840
accccgctct	ccaccccag	gaccacctgg	atcctcacag	agccgagcac	tacagccacc	12900
gtgacggtgc	ccaccggtc	cacggccacc	gcctctcca	cccaggcaac	tgetggcacc	12960
ccacatgtga	gcaccacggc	cacgacacc	acagtcacca	gctccaaagc	cactcccttc	13020
tccagtccag	ggactgcaac	cgcccttcca	gactgagaa	gcacagccac	cacaccaca	13080
gctaccagct	ttacagccat	cccctcctcc	tccctgggca	ccacctggac	ccgcctatca	13140
cagaccacca	caccacggc	caccatgtcc	acagccacac	cctcctccac	tccagagact	13200
gccacacct	ccacagtgt	taccaccag	gccaccaca	ccagggccac	cggctctgtg	13260
gccaccccc	cttccacccc	aggaacagct	cacactacca	aagtgcgac	taccacaacc	13320
acgggcttca	cagtcacccc	ctcctccagc	ccagggacgg	cacgcacgcc	tccagtgtgg	13380
atcagcacia	ccaccacacc	cacaaccagt	ggctccacgg	tgacccccct	ctccgtcccg	13440
gggaccaccc	acacccccac	agtgtgtgacc	accaccacca	caactgtggc	cactggttct	13500
atggcaacac	cctcctctag	cacacagacc	agtgggacca	cccacacacc	cccagtgccg	13560
aacaccacgg	ccaccacaca	cgggcgatcc	ctgtccccca	gcagtccccca	cacggtgcgc	13620
acagcctgga	cttcggccac	ctcaggcacc	ttgggacca	cccacatcac	agagccttc	13680
acggggactt	cccacacccc	agcagcaacc	accggtacca	cccagcactc	gactccagcc	13740
ctgtccagcc	ctcaccttag	cagcaggacc	accgagtcac	ccccttcccc	agggacgacc	13800
accccgggcc	acaccacggc	cacctccagg	accacggcca	cggccacacc	cagcaagacc	13860

[0008]

cgcacctga ccctgctgcc cagcagcccc acatcggccc ccataaccac ggtggtgacc 13920
 acgggctgtg agccccagtg tgcctggtca gagtggctgg actacagcta ccccatgccc 13980
 gggccctctg gcggggactt tgacacctac tccaacatcc gtgcggccgg aggggcccgtc 14040
 tgtgagcagc ccctgggcct cgagtgccgt gccaggccc agcctggtgt cccctgggg 14100
 gagttgggccc aggtcgtgga atgcagcctg gactttggcc tggctgcag gaaccgtgag 14160
 cagggtgggga agttcaagat gtgcttcaac tatgaaatcc gtgtgttctg ctgcaactac 14220
 ggccactgcc ccagcaccac gccaccagc tctacggcca tgccctcctc cactccgggg 14280
 acgacctgga tcctcacaga gctgaccaca acagccacta cgactgcatc cactggatcc 14340
 acggccaccc cgtcctccac cccgggaaca gctccccctc ccaaagtgtc gaccagcccg 14400
 gccaccacac ccacagccac cagttccaaa gccacttctt cctccagtcc aaggactgca 14460
 accacccttc cagtgtgac aagcacagcc accaaatcca cagctaccag cgttacacc 14520
 atcccctcct ccacccttgg gaccaccggg accctcccag aacagaccac cacaccgtg 14580
 gccaccatgt ccacaatcca cccctcctc actccggaga ccaccacac ctccacagt 14640
 ctgaccacga aggccaccac gacaagggcc accagttcca cgtccacccc ctctccact 14700
 ccggggacga cctggatcct cacagagctg accacagcag ccaactacaac tgacgccact 14760
 ggccccacgg ccaccccgtc ctccaccca gggaccacct ggatcctcac agagctgacc 14820
 acaacagcca ctacgactgc gtccactgga tccacggcca cccgtcctc caccacagg 14880
 accacctgga tcctcacaga gccgagcact acagccaccg tgacgggtgcc caccggatcc 14940
 acggccaccg cctcctccac ccaggcaact gctggcacc cacatgtgag caccacggcc 15000
 acgacacca cagtcaccag ctccaaagcc actcctcct ccagtccagg gactgcaact 15060
 gcccttccag cactgagaag cacagccacc acaccacag ctaccagctt tacagccatc 15120
 ccctcctcct ccctgggcac cacctggacc cgcctatcac agaccaccac acccacggcc 15180
 accatgtcca cagccacacc ctctccact ccagagactg tccacacctc cacagtgtt 15240
 accgccacgg ccaccacaac cggggccacc ggctctgtgg ccacccctc ctccaccca 15300
 ggaacagctc aactaccaa agtgccgact accacaacca cgggcttcac agccacccc 15360
 tcctccagcc caggacggc actcacgcct ccaaccacca caccatgtc caccatgtc 15420
 acaatccaca cctcctctac tccagagacc acccacct ccacagtgtc gaccaccaca 15480
 gccaccatga caaggccac caattccacg gccacacct cctccactct ggggacgacc 15540
 cggatcctca ctgagctgac cacaacagcc actacaactg cagccactgg atccacggcc 15600
 accctgtcct ccacccagc gaccacctg atcctcacag agccgagcac tatagccacc 15660
 gtgatggtgc ccaccggtc cacggccacc gcctcctcca ctctgggaac agctcacacc 15720
 cccaaagtgg tgaccacat ggccactatg cccacagcca ctgcctccac ggttcccagc 15780
 tcgtccaccg tggggaccac ccgcaccct gcagtgtcc ccagcagcct gccaacctt 15840
 agcgtgtcca ctgtgtcctc ctcagtcctc accaccctga gaccactgg ctccccagc 15900
 tccacttct ctactccctg cttctgcagg gcatttgac agtttttctc gccggggaa 15960

[0009]

gtcatctaca ataagaccga cggagccggc tgccatttct acgcagtgtg caatcagcac 16020
tgtgacattg accgcttcca gggcgccctgt cccacctccc caccgccagt gtctctccgc 16080
ccgctgtcct cgccctcccc tgcccctggc tgtgacaatg ccatccctct ccggcaggtg 16140
aatgagacct ggaccctgga gaactgcacg gtggccaggt gcggtgggtga caaccgtgtc 16200
gtcctgctgg acccaaagcc tgtggccaac gtcacctgcg tgaacaagca cctgccccatc 16260
aaagtgtcgg acccgagcca gccctgtgac ttccactatg agtgcgagtg tgagtgcgtc 16320
ggtagcccg cggattacccc gggggcaggc atctgcagca tgtggggcgg ctcccactat 16380
tccacctttg acggcacctc ttacacctc cggggcaact gcacctatgt cctcatgaga 16440
gagatccatg cacgcttttg gaatctcagc ctctacctg acaaccacta ctgcacggcc 16500
tctgccactg ccgctgccgc tgccgcccgc tgccccgcg ccctcagcat cactacaag 16560
tccatggata tcgtctcac tgtcaccatg gtgcatggga aggaggagg cctgatcctg 16620
tttgacaaa ttccggtgag cagcggtttc agcaagaacg gcggtgctgt gtctgtgctg 16680
gggaccacca ccatgcgtgt ggacattcct gccctggcg tgagcgtcac cttcaatggc 16740
caagtcttcc aggcccggct gccctacagc ctcttcaca acaacaccga gggccagtgc 16800
ggcacctgca ccaacaacca gagggacgac tgtctccagc gggacggaac cactgccgc 16860
agttgcaagg acatggccaa gacgtggctg gtccccgaca gcagaaagga tggctgctgg 16920
gccccgactg gcacaccccc cactgccagc cccgcagccc cgggtgtctag cacaccacc 16980
cccccccat gccaccaca gccgctctgt gatctgatgc tgagccaggt cttgtctgag 17040
tgccacaacc ttgtgcccc gggcccattc ttcaacgct gcatcagcga cactgcagg 17100
ggccgccttg aggtgccctg ccagagcctg gaggcttacg cagagctctg ccgccccgg 17160
ggagtgtgca gtgactggcg aggtgcaacc ggtggcctgt gcgacctac ctgcccacc 17220
accaaagtgt acaagccatg cggcccata cagcctgcca cctgcaactc taggaaccag 17280
agccacagc tggaggggat ggcggagggc tgcttctgcc ctgaggacca gatcctctc 17340
aacgcacaca tgggcatctg cgtgcaggcc tgcccctgcg tgggaccga tgggtttcct 17400
aaatttccc gggagcgggtg ggtcagcaac tgccagtcct gcggtgtgta cgagggttca 17460
gtgtcgggtg agtgcaagcc cctgccctgt gacgccagc gtcagcccc gccgtgcaac 17520
cgtcccggct tcgtaaccgt gaccaggccc cgggccgaga acccctgctg ccccgagacg 17580
gtgtgcgtgt gcaacacaac cacctgcccc cagagcctgc ctgtgtgcc gccagggcag 17640
gagtccatct gcaccagga ggagggcgac tgctgtccca cttccgctg cagacctcag 17700
ctgtgttctg acaatggcac cttctacggg gttggtgcaa cttcccagg cgccttccc 17760
tgccacatgt gtacctgct ctctggggac acccaggacc caacggtgca atgtcaggag 17820
gatgcctgca acaatactac ctgtccccag ggctttgagt acaagagagt ggccgggag 17880
tgctgtgggg agtgcgtcca gaccgcctgc ctcagcccc atggccagcc agtccagctg 17940
aatgaaacct gggtaacag ccatgtggac aactgcaccg tgtacctctg tgaggctgag 18000

[0010]

ggtggagtcc atttgctgac cccacagcct gcacccctgcc cagatgtgtc cagctgcagg 18060
 gggagcctca ggaaaaccgg ctgctgctac tcctgtgagg aggactcctg tcaagtccgc 18120
 atcaacacga ccacccctgtg gcaccagggc tgcgagaccg aggtcaacat caccttctgc 18180
 gagggtcctt gccccggagc gtccaagtac tcagcagagg cccaggccat gcagcaccag 18240
 tgcacctgct gccaggagag gcgggtccac gaggagacgg tggccttgca ctgtcctaac 18300
 ggctcagcca tcctgcacac ctacaccac gctgtccagg tcctctgtgg gctgcttgcc 18360
 tggggcctgc aggctggagg acacatccgg ggggcagtcc aggaccccca gcagccactg 18420
 aaggaccagg aagcctctgg gaaggccagg caggggggtg gctacaggca gaccgtggcc 18480
 tggggagaca aaagcaatgc tcgtgcctgg ctgcagaagc ctgtggtgtg ggtgcagagc 18540
 ggggccttcc ccacaccggg gcctgcctct gccctgtgcc cttggaaaat gggcattcag 18600
 cctgaaacca ccaagcagct cagagatgct aacatcctga aggagagcaa acgctccatc 18660
 agtagagaaa gacaaaggca atgtgctcaa gctatcaggt ttaatagagg atttggagga 18720
 cagatctgga aatctcaacg ctttttctag 18750

<210> 2
 <211> 14360
 <212> DNA
 <213> 人

<400> 2
 tttcgccagc tcctctgggg gtgacaggca agtgagacgt gctcagagct cccgatgccaa 60
 ggccagggac catggcgctg tgtctgctga ccttggctct ctcgctcttg cccccacaag 120
 ctgctgcaga acaggacctc agtgtgaaca gggctgtgtg ggatggagga ggggtcatct 180
 cccaagggga cgtcttgaac cgtcagtgcc agcagctgtc tcagcacggtt aggacaggtt 240
 ctgcgccaaa caccgccaca ggtacaacat ctacaaatgt cgtggagcca agaattgatt 300
 tgagttgcag caccaaccct gagatgacct cgattgagtc cagtgtgact tcagacactc 360
 ctggtgtctc cagtaccagg atgacaccaa cagaatccag aacaacttca gaatctacca 420
 gtgacagcac cacacttttc cccagttcta ctgaagacac ttcattctctt acaactcctg 480
 aaggcaccga cgtgcccattg tcaacaccaa gtgaagaaag catttcatca acaatggctt 540
 ttgtcagcac tgcacctctt cccagttttg aggcctacac atctttaaca tataaggttg 600
 atatgagcac acctctgacc acttctactc aggcaagttc atctcctact actcctgaaa 660
 gcaccacat acccaaatca actaacagtg aaggaagcac tccattaaca agtatgcctg 720
 ccagcaccat gaagtgggcc agttcagagg ctatcacctt tttgacaact cctgttgaaa 780
 tcagcacacc tgtgaccatt tctgctcaag ccagttcatc tcctacaact gctgaaggtc 840
 ccagcctgtc aaactcagct cctagtggag gaagcactcc attaacaaga atgcctctca 900
 gcgtgatgct ggtggtcagt tctgaggcta gcacccttc acaactcct gctgccacca 960
 acattcctgt gatcacttct actgaagcca gttcatctcc tacaacggct gaaggacca 1020
 gcataccaac ctcaacttat actgaaggaa gcactccatt aacaagtacg cctgccagca 1080

[0011]

ccatgccggt	tgccacttct	gaaatgagca	cactttcaat	aactcctggt	gacaccagca	1140
cacttgtagc	cacttctact	gaacccagtt	cacttcctac	aactgctgaa	gctaccagca	1200
tgctaacctc	aactcctagt	gaaggaagca	ctccattaac	aaatatgcct	gtcagcacca	1260
tattggtggc	cagttctgag	gctagcacca	cttcaacaat	tctgttgac	tccaaaactt	1320
ttgtgaccac	tgctagttaa	gccagctcat	ctcccacaac	tgctgaagat	accagcattg	1380
caacctcaac	tcctagttaa	ggaagcactc	cattaacaag	tatgcctgtc	agcaccactc	1440
cagtggccag	ttctgaggct	agcaaccttt	caacaactcc	tgttgactcc	aaaactcagg	1500
tgaccacttc	tactgaagcc	agttcatctc	ctccaactgc	tgaagttaac	agcatgccaa	1560
cctcaactcc	tagtgaagga	agcactccat	taacaagtat	gtctgtcagc	accatgccgg	1620
tggccagttc	tgaggctagc	accctttcaa	caactcctgt	tgacaccagc	acacctgtga	1680
ccacttctag	tgaagccagt	tcatcttcta	caactcctga	aggtaccagc	ataccaacct	1740
caactcctag	tgaaggaagc	actccattaa	caaacatgcc	tgtcagcacc	aggctggtgg	1800
tcagttctga	ggctagcacc	acttcaacaa	ctcctgctga	ctccaacact	tttgtgacca	1860
cttctagtga	agctagttca	tcttctacaa	ctgctgaagg	taccagcatg	ccaacctcaa	1920
cttacagtga	aagaggcact	acaataacaa	gtatgtctgt	cagcaccaca	ctggtggcca	1980
gttctgaggc	tagcacctt	tcaacaactc	ctggtgactc	caacactcct	gtgaccactt	2040
caactgaagc	cacttcatct	tctacaactg	cggaaggtag	cagcatgcca	acctcaactt	2100
atactgaagg	aagcactcca	ttaacaagta	tgctgtcaaa	caccacactg	gtggccagtt	2160
ctgaggctag	caccctttca	acaactcctg	ttgacaccag	cacacctgtg	accacttcaa	2220
ctgaagccag	ttcctctcct	acaactgctg	atggtgccag	tatgcccaacc	tcaactccta	2280
gtgaaggaag	cactccatta	acaagtatgc	ctgtcagcaa	aacgctgttg	accagttctg	2340
aggctagcac	cctttcaaca	actcctcttg	acacaagcac	acatatcacc	acttctactg	2400
aagccagttg	ctctctaca	accactgaag	gtaccagcat	gccaatctca	actcctagt	2460
aaggaagtcc	tttattaaca	agtatacctg	tcagcatcac	accggtgacc	agtcctgagg	2520
ctagcacctt	ttcaacaact	cctggtgact	ccaacagtcc	tgtgaccact	tctactgaag	2580
tcagttcatc	tcctacacct	gctgaaggta	ccagcatgcc	aacctcaact	tatagtgaag	2640
gaagaactcc	tttaacaagt	atgcctgtca	gcaccacact	ggtggccact	tctgcaatca	2700
gcaccctttc	aacaactcct	gttgacacca	gcacacctgt	gaccaattct	actgaagccc	2760
gttcgtctcc	tacaacttct	gaaggtacca	gcatgccaac	ctcaactcct	ggggaaggaa	2820
gcactccatt	aacaagtatg	cctgacagca	ccacgccggt	agtcagttct	gaggctagaa	2880
cactttcagc	aactcctggt	gacaccagca	cacctgtgac	cacttctact	gaagccactt	2940
catctctac	aactgctgaa	ggtaccagca	taccaacctc	gactcctagt	gaaggaacga	3000
ctccattaac	aagcacacct	gtcagccaca	cgctggtggc	caattctgag	gctagcacc	3060
tttcaacaac	tcctgttgac	tccaacactc	ctttgaccac	ttctactgaa	gccagttcac	3120
ctcctccac	tgctgaaggt	accagcatgc	caacctcaac	tcctagttaa	ggaagcactc	3180

[0012]

cattaacacg	tatgcctgtc	agcaccacaa	tggtggccag	ttctgaaacg	agcacacttt	3240
caacaactcc	tgctgacacc	agcacacctg	tgaccactta	ttctcaagcc	agttcatctt	3300
ctacaactgc	tgacgggtacc	agcatgccaa	cctcaactta	tagtgaagga	agcactccac	3360
taacaagtgt	gcctgtcagc	accaggctgg	tggtcagttc	tgaggctagc	accctttcca	3420
caactcctgt	cgacaccagc	atacctgtca	ccacttctac	tgaagccagt	tcattctccta	3480
caactgctga	aggtaccagc	ataccaacct	cacctcccag	tgaaggaacc	actccgtag	3540
caagtatgcc	tgctcagcacc	acgctgggtg	tcagttctga	ggctaacacc	ctttcaacaa	3600
ctcctgtgga	ctccaaaact	caggtggcca	cttctactga	agccagttca	cctcctccaa	3660
ctgctgaagt	taccagcatg	ccaacctcaa	ctcctggaga	aagaagcact	ccattaacaa	3720
gtatgcctgt	cagacacacg	ccagtggcca	gttctgaggc	tagcacctt	tcaacatctc	3780
ccgttgacac	cagcacacct	gtgaccactt	ctgctgaaac	cagttcctct	cctacaaccg	3840
ctgaaggtag	cagcttgcca	acctcaacta	ctagtgaagg	aagtactcta	ttaacaagta	3900
tacctgtcag	caccacgctg	gtgaccagtc	ctgaggctag	caccctttta	acaactcctg	3960
ttgacactaa	aggtcctgtg	gtcacttcta	atgaagtcag	ttcatctcct	acacctgctg	4020
aaggtaccag	catgccaacc	tcaacttata	gtgaaggaag	aactccttta	acaagtatac	4080
ctgtcaacac	cacactgggtg	gccagttctg	caatcagcat	cctttcaaca	actcctgttg	4140
acaacagcac	acctgtgacc	acttctactg	aagcctgttc	atctcctaca	acttctgaag	4200
gtaccagcat	gccaaactca	aatcctagtg	aaggaaccac	tccgttaaca	agtatacctg	4260
tcagcaccac	gccggtagtc	agttctgagg	ctagcacctt	ttcagcaact	cctggtgaca	4320
ccagcaccac	tgggaccact	tctgctgaag	ccacttcatc	tcctacaact	gctgaaggta	4380
tcagcatacc	aacctcaact	cctagtgaag	gaaagactcc	attaaaaagt	atacctgtca	4440
gcaacacgcc	ggtggccaat	tctgaggcta	gcaccctttc	aacaactcct	gttgactcta	4500
acagtctctg	ggtcacttct	acagcagtca	gttcatctcc	tacacctgct	gaaggtacca	4560
gcatagcaat	ctcaacgcct	agtgaaggaa	gcactgcatt	aacaagtata	cctgtcagca	4620
ccacaacagt	ggccagttct	gaaatcaaca	gcctttcaac	aactcctgct	gtcaccagca	4680
cacctgtgac	cacttattct	caagccagtt	catctcctac	aactgctgac	ggtaccagca	4740
tgcaaacctc	aacttatagt	gaaggaagca	ctccactaac	aagtttgctt	gtcagcacca	4800
tgctgggtgt	cagttctgag	gctaacaccc	tttcaacaac	ccctattgac	tccaaaactc	4860
aggtgaccgc	ttctactgaa	gccagttcat	ctacaaccgc	tgaaggtagc	agcatgacaa	4920
tctcaactcc	tagtgaagga	agtcctctat	taacaagtat	acctgtcagc	accacgccgg	4980
tgccagctcc	tgaggctagc	accctttcaa	caactcctgt	tgactccaac	agtcctgtga	5040
tcacttctac	tgaagtcagt	tcattctccta	cacctgctga	aggtaccagc	atgccaacct	5100
caacttatac	tgaaggaaga	actcctttta	caagtataac	tgtcagaaca	acaccggtgg	5160
ccagctctgc	aatcagcacc	ctttcaacaa	ctcccgttga	caacagcaca	cctgtgacca	5220

[0013]

cttctactga agcccgttca tctcctacaa cttctgaagg taccagcatg ccaaactcaa	5280
ctcctagtga aggaaccact ccattaacaa gtatacctgt cagcaccacg ccggtactca	5340
gttctgaggc tagcacctt tcagcaactc ctattgacac cagcaccctt gtgaccactt	5400
ctactgaagc cacttcgtct cctacaactg ctgaaggtag cagcatacca acctcgactc	5460
ttagtgaagg aatgactcca ttaacaagca cacctgtcag ccacacgctg gtggccaatt	5520
ctgaggctag caccctttca acaactcctg ttgactctaa cagtctctgtg gtcacttcta	5580
cagcagtcag ttcatctcct acacctgctg aaggtagcag catagcaacc tcaacgccta	5640
gtgaaggaag cactgcatta acaagtatac ctgtcagcac cacaacagtg gccagttctg	5700
aaaccaacac cttttcaaca actcccgctg tcaccagcac acctgtgacc acttatgctc	5760
aagtcagttc atctctaca actgctgacg gtagcagcat gccaacctca actcctaggg	5820
aaggaaggcc tccattaaca agtatacctg tcagcaccac aacagtggcc agttctgaaa	5880
tcaacacctt ttcaacaact cttgctgaca ccaggacacc tgtgaccact tattctcaag	5940
ccagttcatc tcctacaact gctgatggta ccagcatgcc aaccccagct tatagtgaag	6000
gaagcactcc actaacaagt atgcctctca gcaccagct ggtggtcagt tctgaggcta	6060
gcactctttc cacaactcct gttgacacca gcactcctgc caccacttct actgaaggca	6120
gttcatctcc tacaactgca ggaggtacca gcatacaaac ctcaactcct agtgaacgga	6180
ccactccatt agcaggtatg cctgtcagca ctacgcttgt ggtcagttct gagggtaaca	6240
ccctttcaac aactcctggt gactccaaaa ctcaggtgac caattctact gaagccagtt	6300
catctgcaac cgctgaaggc agcagcatga caatctcagc tcctagtga ggaagtcctc	6360
tactaacaag tatacctctc agcaccacgc cgggtggccag tcctgaggct agcaccctt	6420
caacaactcc tgttgactcc aacagtctctg tgatcacttc tactgaagtc agttcatctc	6480
ctatacctac tgaaggtacc agcatgcaaa cctcaactta tagtgacaga agaactcctt	6540
taacaagtat gcctgtcagc accacagtggt tggccagttc tgcaatcagc accctttcaa	6600
caactcctgt tgacaccagc acacctgtga ccaattctac tgaagcccgt tcactctcta	6660
caacttctga aggtaccagc atgccaacct caactcctag tgaaggaagc actccattca	6720
caagtatgcc tgtcagcacc atgccggtag ttacttctga ggctagcacc ctttcagcaa	6780
ctcctgttga caccagcaca cctgtgacca cttctactga agccacttca tctcctacaa	6840
ctgctgaagg taccagcata ccaacttcaa ctcttagtga aggaacgact ccattaacaa	6900
gtatacctgt cagccacacg ctggtggcca attctgaggt tagcacctt tcaacaactc	6960
ctgttgactc caaactcct ttactactt ctactgaagc cagttcacct cctcccactg	7020
ctgaaggtag cagcatgcca acctcaactt ctagtgaagg aaacactcca ttaacacgta	7080
tgctgtcag caccacaatg gtggccagtt ttgaaacaag cacactttct acaactcctg	7140
ctgacaccag cacacctgtg actacttatt ctcaagccgg ttcatctcct acaactgctg	7200
acgatactag catgccaac tcaacttata gtgaaggaag cactccacta acaagtgtgc	7260
ctgtcagcac catgccggtg gtcagttctg aggctagcac ccattccaca actcctgttg	7320

[0014]

acaccagcac	acctgtcacc	acttctactg	aagccagttc	atctcctaca	actgctgaag	7380
gtaccagcat	accaacctca	cctcctagtg	aaggaaccac	tccgttagca	agtatgcctg	7440
tcagcaccac	gccggtggtc	agttctgagg	ctggcacctt	ttccacaact	cctggtgaca	7500
ccagcacacc	tatgaccact	tctactgaag	ccagttcatc	tcctacaact	gctgaagata	7560
tcgtcgtgcc	aatctcaact	gctagtgaag	gaagtactct	attaacaagt	atacctgtca	7620
gcaccacgcc	agtggccagt	cctgaggcta	gcaccctttc	aacaactcct	gttgactcca	7680
acagtccctg	ggtcacttct	actgaaatca	gttcatctgc	tacatccgct	gaaggtacca	7740
gcatgcctac	ctcaacttat	agtgaaggaa	gcactccatt	aagaagtatg	cctgtcagca	7800
ccaagccggt	ggccagttct	gaggctagca	ctctttcaac	aactcctggt	gacaccagca	7860
tacctgtcac	cacttctact	gaaaccagtt	catctcctac	aactgcaaaa	gataccagca	7920
tgccaatctc	aactcctagt	gaagtaagta	cttcattaac	aagtatactt	gtcagcacca	7980
tgccagtggc	cagttctgag	gctagcacc	tttcaacaac	tcctggtgac	accaggacac	8040
ttgtgaccac	ttccactgga	accagttcat	ctcctacaac	tgctgaaggt	agcagcatgc	8100
caacctcaac	tcctggtgaa	agaagcactc	cattaacaaa	tatacttgtc	agcaccacgc	8160
tgttggccaa	ttctgaggct	agcacccttt	caacaactcc	tgttgacacc	agcacacctg	8220
tcaccacttc	tgctgaagcc	agttcttctc	ctacaactgc	tgaaggtacc	agcatgcgaa	8280
tctcaactcc	tagtgatgga	agtactccat	taacaagtat	acttgtcagc	accctgccag	8340
tggccagttc	tgaggctagc	accgtttcaa	caactgctgt	tgacaccagc	atacctgtca	8400
ccacttctac	tgaagccagt	tcctctccta	caactgctga	agttaccagc	atgccaacct	8460
caactcctag	tgaacaaggt	actccattaa	ctagtatgcc	tgtcaaccac	acgccagtgg	8520
ccagttctga	ggctggcacc	ctttcaacaa	ctcctgttga	caccagcaca	cctgtgacca	8580
cttctactaa	agccagttca	tctcctacaa	ctgctgaagg	tatcgtcgtg	ccaatctcaa	8640
ctgctagtga	aggaagtact	ctattaacaa	gtatacctgt	cagcaccacg	ccggtggcca	8700
gttctgaggc	tagcaccctt	tcaacaactc	ctggtgatac	cagcatacct	gtcaccactt	8760
ctactgaagg	cagttcttct	cctacaactg	ctgaaggtac	cagcatgcca	atctcaactc	8820
ctagtgaagt	aagtactcca	ttaacaagta	tacttgtcag	caccgtgcca	gtggccggtt	8880
ctgaggctag	caccctttca	acaactcctg	ttgacaccag	gacacctgtc	accacttctg	8940
ctgaagctag	ttcttctcct	acaactgctg	aaggtaccag	catgccaatc	tcaactcctg	9000
gcgaaagaag	aactccatta	acaagtatgt	ctgtcagcac	catgccggtg	gccagttctg	9060
aggctagcac	cctttcaaga	actcctgctg	acaccagcac	acctgtgacc	acttctactg	9120
aagccagttc	ctctcctaca	actgctgaag	gtaccggcat	accaatctca	actcctagtg	9180
aaggaagtac	tccattaaca	agtatacctg	tcagcaccac	gccagtggcc	attcctgagg	9240
ctagcaccct	ttcaacaact	cctggtgact	ccaacagtcc	tgtggtcact	tctactgaag	9300
tcagttcatc	tcctacacct	gctgaaggta	ccagcatgcc	aatctcaact	tatagtgaag	9360

[0015]

gaagcactcc	attaacaggt	gtgcctgtca	gcaccacacc	ggtgaccagt	tctgcaatca	9420
gcaccctttc	aacaactcct	gttgacacca	gcacacctgt	gaccacttct	actgaagccc	9480
attcatctcc	tacaacttct	gaaggtacca	gcatgccaac	ctcaactcct	agtgaaggaa	9540
gtactccatt	aacatatatg	cctgtcagca	ccatgctggt	agtcagttct	gaggatagca	9600
ccctttcagc	aactcctggt	gacaccagca	cacctgtgac	cacttctact	gaagccactt	9660
catctacaac	tgctgaaggt	accagcattc	caacctcaac	tcctagttaa	ggaatgactc	9720
cattaactag	tgtacctgtc	agcaacacgc	cggtggccag	ttctgaggct	agcatccttt	9780
caacaactcc	tgttgactcc	aacactcctt	tgaccacttc	tactgaagcc	agttcatctc	9840
ctcccactgc	tgaaggtacc	agcatgccaa	cctcaactcc	tagtgaagga	agcactccat	9900
taacaagtat	gcctgtcagc	accacaacgg	tggccagttc	tgaaacgagc	accctttcaa	9960
caactcctgc	tgacaccagc	acacctgtga	ccacttatcc	tcaagccagt	tcctctctcc	10020
caattgctga	cggtagctagc	atgccaacct	caacttatag	tgaaggaagc	actccactaa	10080
caaatatgtc	tttcagcacc	acgccagtg	tcagttctga	ggctagcacc	ctttccacaa	10140
ctcctgttga	caccagcaca	cctgtcacca	cttctactga	agccagttta	tctcctacaa	10200
ctgctgaagg	taccagcata	ccaacctcaa	gtcctagtga	aggaaccact	ccattagcaa	10260
gtatgcctgt	cagcaccacg	ccggtgggtca	gttctgaggt	taacaccctt	tcaacaactc	10320
ctgtggactc	caacactctg	gtgaccactt	ctactgaagc	cagttcatct	cctacaatcg	10380
ctgaaggtac	cagcttgcca	acctcaacta	ctagtgaagg	aagcactcca	ttatcaatta	10440
tgctctcag	taccagccg	gtggccagtt	ctgaggctag	caccctttca	acaactcctg	10500
ttgacaccag	cacacctgtg	accacttctt	ctccaaccaa	ttcatctcct	acaactgctg	10560
aagttaccag	catgccaca	tcaactgctg	gtgaaggaag	cactccatta	acaaatagc	10620
ctgtcagcac	cacaccggtg	gccagttctg	aggctagcac	cctttcaaca	actcctgttg	10680
actccaacac	ttttgttacc	agttctagtc	aagccagttc	atctccagca	actcttcagg	10740
tcaccactat	gcgtatgtct	actccaagtg	aaggaagctc	ttcattaaca	actatgctcc	10800
tcagcagcac	atatgtgacc	agttctgagg	ctagcacacc	ttccactcct	tctgttgaca	10860
gaagcacacc	tgtgaccact	tctactcaga	gcaattctac	tcctacacct	cctgaagtta	10920
tcaccctgcc	aatgtcaact	cctagtgaag	taagcactcc	attaaccatt	atgcctgtca	10980
gcaccacatc	ggtgaccatt	tctgaggctg	gcacagcttc	aacacttctt	gttgacacca	11040
gcacacctgt	gatcacttct	acccaagtca	gttcatctcc	tgtgactcct	gaaggtacca	11100
ccatgccaat	ctggacgcct	agtgaaggaa	gcactccatt	aacaactatg	cctgtcagca	11160
ccacacgtgt	gaccagctct	gagggtagca	ccctttcaac	accttctggt	gtcaccagca	11220
cacctgtgac	cacttctact	gaagccattt	catcttctgc	aactcttgac	agcaccacca	11280
tgtctgtgtc	aatgcccattg	gaaataagca	cccttgggac	cactattctt	gtcagtagca	11340
cacctgttac	gaggtttcct	gagagtagca	ccccttccat	accatctggt	tacaccagca	11400
tgtctatgac	cactgcctct	gaaggcagtt	catctcctac	aactcttgaa	ggcaccacca	11460

[0016]

ccatgcctat gtcaactacg agtgaaagaa gcactttatt gacaactgtc ctcatcagcc 11520
ctatatctgt gatgagtcct tctgaggcca gcacactttc aacacctcct ggtgatacca 11580
gcacaccttt gctcacctct accaaagccg gttcattctc catacctgct gaagtcacta 11640
ccatacgtat ttcaattacc agtgaaagaa gcactccatt aacaactctc cttgtcagca 11700
ccacacttcc aactagcttt cctggggcca gcatagcttc gacacctcct cttgacacaa 11760
gcacaacttt tacccttct actgacactg cctcaactcc cacaattcct gtagccacca 11820
ccatatctgt atcagtgatc acagaaggaa gcacacctgg gacaaccatt tttattccca 11880
gcactcctgt caccagttct actgctgatg tctttcctgc aacaactggg gctgtatcta 11940
cccctgtgat aacttcact gaactaaaca caccatcaac ctccagtagt agtaccacca 12000
catctttttc aactactaag gaatttaca caccgcaat gactactgca gctcccctca 12060
catatgtgac catgtctact gccccagca caccagaac aaccagcaga ggctgacta 12120
cttctgcatc aacgctttct gcaaccagta cacctcacac ctctacttct gtcaccacc 12180
gtcctgtgac cccttcatca gaatccagca ggccgtcaac aattacttct cacaccatcc 12240
cacctacatt tcctcctgct cactccagta cacctccaac aacctctgcc tcctccacga 12300
ctgtgaacc tgaggctgtc accaccatga ccaccaggac aaaaccagc acacggacca 12360
cttccttccc cacggtgacc accaccgctg tccccagaa tactacaatt aagagcaacc 12420
ccacctcaac tcctactgtg ccaagaacca caacatgctt tggagatggg tgccagaata 12480
cggcctctcg ctgcaagaat ggaggcacct gggatgggct caagtgccag tgtccaacc 12540
tctattatgg ggagttgtgt gaggaggtgg tcagcagcat tgacataggg ccaccggaga 12600
ctatctctgc ccaaatggaa ctgactgtga cagtgaccag tgtgaagttc accgaagagc 12660
taaaaaacca ctcttcccag gaattccagg agttcaaaca gacattcacg gaacagatga 12720
atattgtgta ttccgggatc cctgagtatg tcggggtgaa catcacaag ctacgtcttg 12780
gcagtgtggg ggtggagcat gacgtcctcc taagaacca gtacacacca gaatacaaga 12840
cagtattgga caatgccacc gaagtagtga aagagaaaat cacaaaagt accacacagc 12900
aaataatgat taatgatatt tgctcagaca tgatgtgttt caacaccact ggcaccaag 12960
tgcaaacat tacggtgacc cagtacgacc ctgaagagga ctgccggaag atggccaagg 13020
aatatggaga ctacttcgta gtggagtacc gggaccagaa gccatactgc atcagcccct 13080
gtgagcctgg cttcagtgtc tccaagaact gtaacctcg caagtgccag atgtctctaa 13140
gtggacctca gtgcctctgc gtgaccacgg aaactcactg gtacagtggg gagacctgta 13200
accagggcac ccagaagagt ctggtgtacg gcctcgtggg ggcaggggtc gtgctgatgc 13260
tgatcatcct ggtagctctc ctgatgctcg ttttccgctc caagagagag gtgaaacggc 13320
aaaagtacag attgtctcag ttatacaagt ggcaagaaga ggacagtgga ccagctcctg 13380
ggaccttcca aaacattggc tttgacatct gccaatgta tgattccatc cacctggagt 13440
ccatctatag taatttccag ccctccttga gacacataga ccctgaaca agatccgaa 13500

[0017]

```

ttcagaggcc tcaggaatg acgacatcat ttttaaggcat ggagctgaga agtctgggag 13560
tgaggagatc ccagtcgccc taagcttggt ggagcatttt cccattgaga gccttccatg 13620
ggaactcaat gttcccattg taagtacagg aaacaagccc tgtacttacc aaggagaaag 13680
aggagagaca gcagtgctgg gagattctca aatagaaacc cgtggacgct ccaatgggct 13740
tgtcatgata tcaggctagg ctttctgct ctttttcaa agacgctcca gatttgaggg 13800
tactctgact gcaacatctt tcacccatt gatcgccagg attgatttgg ttgatctggc 13860
tgagcaggcg ggtgtccccg tcctccctca ctgccccata tgtgtccctc ctaaagctgc 13920
atgctcagtt gaagaggacg agaggacgac cttctctgat agaggaggac cacgcttcag 13980
tcaaaggcat acaagtatct atctggactt ccctgctagc acttccaaac aagctcagag 14040
atgttcctcc cctcatctgc ccgggttcag taccatggac agcgcctcg acccgctggt 14100
tacaaccatg accccttggg cactggactg catgcacttt acatatcaca aaatgctctc 14160
ataagaatta ttgcatacca tcttcatgaa aaacacctgt atttaaatat agagcattta 14220
ccttttggtg tataagattg tgggtatfff ttaagttctt attgttatga gttctgattt 14280
tttccttagt aaatattata atatatattt gtagtaacta aaaataataa agcaatttta 14340
ttacaatttt aaaaaaaaaa 14360

```

```

<210> 3
<211> 632
<212> DNA
<213> 人

```

```

<400> 3
cccagaatgt tgacagtcgc ttcctagcc cttctctgtg cctcagcctc tggcaatgcc 60
attcaggcca ggtcttctc ctatagtggg gagtatggaa gtgggtgggg aaagcgattc 120
tctcattctg gcaaccagtt ggacggcccc atcaccgccc tccgggtccg agtcaacaca 180
tactacatcg taggtcttca ggtgcgctat ggcaagggtg ggagcgacta tgtgggtggg 240
cgcaacggag acctggagga gatctttctg caccctgggg aatcagtgat ccaggtttct 300
gggaagtaca agtggtagct gaagaagctg gtattttgtg cagacaaggg ccgctatctg 360
tcttttgga aagacagtgg cacaagttc aatgccgtcc ccttgacacc caacaccgtg 420
ctccgcttca tcagtggccg gtctggttct ctcatcgatg ccattggcct gcactgggat 480
gtttaccca ctagctgcag cagatgctga gcctcctctc cttggcaggg gcactgtgat 540
gaggagtaag aactccctta tcaactaacc ccatcctaat ggctcaataa aaaaatatgg 600
ttaaggctaa aaaaaaaaaa aaaaaaaaaa aa 632

```

```

<210> 4
<211> 6207
<212> PRT
<213> 人

```

```

<400> 4
Met Ser Val Gly Arg Arg Lys Leu Ala Leu Leu Trp Ala Leu Ala Leu
1 5 10 15

```

[0018]

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

[0019]

Glu Ala Cys Arg Gln Asp Leu Cys Phe Cys Glu Asp Thr Asp Leu Leu
 290 295 300

Ser Cys Val Cys His Thr Leu Ala Glu Tyr Ser Arg Gln Cys Thr His
 305 310 315 320

Ala Gly Gly Leu Pro Gln Asp Trp Arg Gly Pro Asp Phe Cys Pro Gln
 325 330 335

Lys Cys Pro Asn Asn Met Gln Tyr His Glu Cys Arg Ser Pro Cys Ala
 340 345 350

Asp Thr Cys Ser Asn Gln Glu His Ser Arg Ala Cys Glu Asp His Cys
 355 360 365

Val Ala Gly Cys Phe Cys Pro Glu Gly Thr Val Leu Asp Asp Ile Gly
 370 375 380

Gln Thr Gly Cys Val Pro Val Ser Lys Cys Ala Cys Val Tyr Asn Gly
 385 390 395 400

Ala Ala Tyr Ala Pro Gly Ala Thr Tyr Ser Thr Asp Cys Thr Asn Cys
 405 410 415

Thr Cys Ser Gly Gly Arg Trp Ser Cys Gln Glu Val Pro Cys Pro Gly
 420 425 430 435

Thr Cys Ser Val Leu Gly Gly Ala His Phe Ser Thr Phe Asp Gly Lys
 435 440 445

Gln Tyr Thr Val His Gly Asp Cys Ser Tyr Val Leu Thr Lys Pro Cys
 450 455 460

Asp Ser Ser Ala Phe Thr Val Leu Ala Glu Leu Arg Arg Cys Gly Leu
 465 470 475 480

Thr Asp Ser Glu Thr Cys Leu Lys Ser Val Thr Leu Ser Leu Asp Gly
 485 490 495

Ala Gln Thr Val Val Val Ile Lys Ala Ser Gly Glu Val Phe Leu Asn
 500 505 510

Gln Ile Tyr Thr Gln Leu Pro Ile Ser Ala Ala Asn Val Thr Ile Phe
 515 520 525

Arg Pro Ser Thr Phe Phe Ile Ile Ala Gln Thr Ser Leu Gly Leu Gln
 530 535 540

Leu Asn Leu Gln Leu Val Pro Thr Met Gln Leu Phe Met Gln Leu Ala
 545 550 555 560

[0020]

Pro Lys Leu Arg Gly Gln Thr Cys Gly Leu Cys Gly Asn Phe Asn Ser
 565 570 575

Ile Gln Ala Asp Asp Phe Arg Thr Leu Ser Gly Val Val Glu Ala Thr
 580 585 590

Ala Ala Ala Phe Phe Asn Thr Phe Lys Thr Gln Ala Ala Cys Pro Asn
 595 600 605

Ile Arg Asn Ser Phe Glu Asp Pro Cys Ser Leu Ser Val Glu Asn Ala
 610 615 620

Pro Arg Met Gly Ala Pro Ser Ala Cys Arg Thr Leu Val Leu Ala Leu
 625 630 635 640

Ala Ala Met Leu Val Val Pro Gln Ala Glu Thr Gln Gly Pro Val Glu
 645 650 655

Pro Ser Trp Glu Asn Ala Gly His Thr Met Asp Gly Gly Ala Pro Thr
 660 665 670

Ser Ser Pro Thr Arg Arg Val Ser Phe Val Pro Pro Val Thr Val Phe
 675 680 685

Pro Ser Leu Ser Arg Lys Gln Met Leu Pro Leu Pro Ala Gly Lys Gly
 690 695 700

Val Phe Ala Ser Pro Lys Gly Gly Gly Pro Asp Leu Gly Val Gln Leu
 705 710 715 720

Pro Pro Ala Leu Asn Pro Ala His Asn Gly Arg Val Cys Ser Thr Trp
 725 730 735

Gly Asp Phe His Tyr Lys Thr Phe Asp Gly Asp Val Phe Arg Phe Pro
 740 745 750

Gly Leu Cys Asn Tyr Val Phe Ser Glu His Cys Arg Ala Ala Tyr Glu
 755 760 765

Asp Phe Asn Val Gln Leu Arg Arg Gly Leu Val Gly Ser Arg Pro Val
 770 775 780

Val Thr Arg Val Val Ile Lys Ala Gln Gly Leu Val Leu Glu Ala Ser
 785 790 795 800

Asn Gly Ser Val Leu Ile Asn Gly Gln Arg Glu Glu Leu Pro Tyr Ser
 805 810 815

Arg Thr Gly Leu Leu Val Glu Gln Ser Gly Asp Tyr Ile Lys Val Ser
 820 825 830

Ile Arg Leu Val Leu Thr Phe Leu Trp Asn Gly Glu Asp Ser Ala Leu

[0021]

Cys Ala Asp Ser Ser Phe Thr Val Leu Ala Glu Leu Arg Lys Cys
 1115 1120 1125
 Gly Leu Thr Asp Asn Glu Asn Cys Leu Lys Ala Val Thr Leu Ser
 1130 1135 1140
 Leu Asp Gly Gly Asp Thr Ala Ile Arg Val Gln Ala Asp Gly Gly
 1145 1150 1155
 Val Phe Leu Asn Ser Ile Tyr Thr Gln Leu Pro Leu Ser Ala Ala
 1160 1165 1170
 Asn Ile Thr Leu Phe Thr Pro Ser Ser Phe Phe Ile Val Val Gln
 1175 1180 1185
 Thr Gly Leu Gly Leu Gln Leu Leu Val Gln Leu Val Pro Leu Met
 1190 1195 1200
 Gln Val Phe Val Arg Leu Asp Pro Ala His Gln Gly Gln Met Cys
 1205 1210 1215
 Gly Leu Cys Gly Asn Phe Asn Gln Asn Gln Ala Asp Asp Phe Thr
 1220 1225 1230
 Ala Leu Ser Gly Val Val Glu Ala Thr Gly Ala Ala Phe Ala Asn
 1235 1240 1245
 Thr Trp Lys Ala Gln Ala Ala Cys Ala Asn Ala Arg Asn Ser Phe
 1250 1255 1260
 Glu Asp Pro Cys Ser Leu Ser Val Glu Asn Glu Asn Tyr Ala Arg
 1265 1270 1275
 His Trp Cys Ser Arg Leu Thr Asp Pro Asn Ser Ala Phe Ser Arg
 1280 1285 1290
 Cys His Ser Ile Ile Asn Pro Lys Pro Phe His Ser Asn Cys Met
 1295 1300 1305
 Phe Asp Thr Cys Asn Cys Glu Arg Ser Glu Asp Cys Leu Cys Ala
 1310 1315 1320
 Ala Leu Ser Ser Tyr Val His Ala Cys Ala Ala Lys Gly Val Gln
 1325 1330 1335
 Leu Ser Asp Trp Arg Asp Gly Val Cys Thr Lys Tyr Met Gln Asn
 1340 1345 1350
 Cys Pro Lys Ser Gln Arg Tyr Ala Tyr Val Val Asp Ala Cys Gln
 1355 1360 1365

[0023]

Pro Thr Cys Arg Gly Leu Ser Glu Ala Asp Val Thr Cys Ser Val
 1370 1375 1380

Ser Phe Val Pro Val Asp Gly Cys Thr Cys Pro Ala Gly Thr Phe
 1385 1390 1395

Leu Asn Asp Ala Gly Ala Cys Val Pro Ala Gln Glu Cys Pro Cys
 1400 1405 1410

Tyr Ala His Gly Thr Val Leu Ala Pro Gly Glu Val Val His Asp
 1415 1420 1425

Glu Gly Ala Val Cys Ser Cys Thr Gly Gly Lys Leu Ser Cys Leu
 1430 1435 1440

Gly Ala Ser Leu Gln Lys Ser Thr Gly Cys Ala Ala Pro Met Val
 1445 1450 1455

Tyr Leu Asp Cys Ser Asn Ser Ser Ala Gly Thr Pro Gly Ala Glu
 1460 1465 1470

Cys Leu Arg Ser Cys His Thr Leu Asp Val Gly Cys Phe Ser Thr
 1475 1480 1485

His Cys Val Ser Gly Cys Val Cys Pro Pro Gly Leu Val Ser Asp
 1490 1495 1500

Gly Ser Gly Gly Cys Ile Ala Glu Glu Asp Cys Pro Cys Val His
 1505 1510 1515

Asn Glu Ala Thr Tyr Lys Pro Gly Glu Thr Ile Arg Val Asp Cys
 1520 1525 1530

Asn Thr Cys Thr Cys Arg Asn Arg Arg Trp Glu Cys Ser His Arg
 1535 1540 1545

Leu Cys Leu Gly Thr Cys Val Ala Tyr Gly Asp Gly His Phe Ile
 1550 1555 1560

Thr Phe Asp Gly Asp Arg Tyr Ser Phe Glu Gly Ser Cys Glu Tyr
 1565 1570 1575

Ile Leu Ala Gln Asp Tyr Cys Gly Asp Asn Thr Thr His Gly Thr
 1580 1585 1590

Phe Arg Ile Val Thr Glu Asn Ile Pro Cys Gly Thr Thr Gly Thr
 1595 1600 1605

Thr Cys ser Lys Ala Ile Lys Leu Phe Val Glu Val Arg Thr Ala
 1610 1615 1620

[0024]

Pro Ala Val Ser Thr Pro Asp Pro Ala Ala Asn Glu Pro Ala Pro
 1625 1630 1635
 Arg Glu Ala Ser Val Gly Phe Arg Gln Arg Leu Pro Pro Leu Gln
 1640 1645 1650
 Ser Tyr Glu Leu Ile Leu Gln Glu Gly Thr Phe Lys Ala Val Ala
 1655 1660 1665
 Arg Gly Pro Gly Gly Asp Pro Pro Tyr Lys Ile Arg Tyr Met Gly
 1670 1675 1680
 Ile Phe Leu Val Ile Glu Thr His Gly Met Ala Val Ser Trp Asp
 1685 1690 1695
 Arg Lys Thr Ser Val Phe Ile Arg Leu His Gln Asp Tyr Lys Gly
 1700 1710
 Arg Val Cys Gly Leu Cys Gly Asn Phe Asp Asp Asn Ala Ile Asn
 1715 1720 1725
 Asp Phe Ala Thr Arg Ser Arg Ser Val Val Gly Asp Ala Leu Glu
 1730 1735 1740
 Phe Gly Asn Ser Trp Lys Leu Ser Pro Ser Cys Pro Asp Ala Leu
 1745 1750 1755
 Ala Pro Lys Asp Pro Cys Thr Ala Asn Pro Phe Arg Lys Ser Trp
 1760 1765 1770
 Ala Gln Lys Gln Cys Ser Ile Leu His Gly Pro Thr Phe Ala Ala
 1775 1780 1785
 Cys Arg Ser Gln Val Asp Ser Thr Lys Tyr Tyr Glu Ala Cys Val
 1790 1795 1800
 Asn Asp Ala Cys Ala Cys Asp Ser Gly Gly Asp Cys Glu Cys Phe
 1805 1810 1815
 Cys Thr Ala Val Ala Ala Tyr Ala Gln Ala Cys His Asp Ala Gly
 1820 1825 1830
 Leu Cys Val Ser Trp Arg Thr Pro Asp Thr Cys Pro Leu Phe Cys
 1835 1840 1845
 Asp Phe Tyr Asn Pro His Gly Gly Cys Glu Trp His Tyr Gln Pro
 1850 1855 1860
 Cys Gly Ala Pro Cys Leu Lys Thr Cys Arg Asn Pro Ser Gly His
 1865 1870 1875
 Cys Leu Val Asp Leu Pro Gly Leu Glu Gly Cys Tyr Pro Lys Cys

[0025]

1880	1885	1890												
Pro 1895	Pro	Ser	Gln	Pro	Phe	Phe 1900	Asn	Glu	Asp	Gln	Met 1905	Lys	Cys	Val
Ala 1910	Gln	Cys	Gly	Cys	Tyr	Asp 1915	Lys	Asp	Gly	Asn	Tyr 1920	Tyr	Asp	Val
Gly 1925	Ala	Arg	Val	Pro	Thr	Ala 1930	Glu	Asn	Cys	Gln	Ser 1935	Cys	Asn	Cys
Thr 1940	Pro	Ser	Gly	Ile	Gln	Cys 1945	Ala	His	Ser	Leu	Glu 1950	Ala	Cys	Thr
Cys 1955	Thr	Tyr	Glu	Asp	Arg	Thr 1960	Tyr	Ser	Tyr	Gln	Asp 1965	Val	Ile	Tyr
Asn 1970	Thr	Thr	Asp	Gly	Leu	Gly 1975	Ala	Cys	Leu	Ile	Ala 1980	Ile	Cys	Gly
Ser 1985	Asn	Gly	Thr	Ile	Ile	Arg 1990	Lys	Ala	Val	Ala	Cys 1995	Pro	Gly	Thr
Pro 2000	Ala	Thr	Thr	Pro	Phe	Thr 2005	Phe	Thr	Thr	Ala	Trp 2010	Val	Pro	His
Ser 2015	Thr	Thr	Ser	Pro	Ala	Leu 2020	Pro	Val	Ser	Thr	Val 2025	Cys	Val	Arg
Glu 2030	Val	Cys	Arg	Trp	Ser	Ser 2035	Trp	Tyr	Asn	Gly	His 2040	Arg	Pro	Glu
Pro 2045	Gly	Leu	Gly	Gly	Gly	Asp 2050	Phe	Glu	Thr	Phe	Glu 2055	Asn	Leu	Arg
Gln 2060	Arg	Gly	Tyr	Gln	Val	Cys 2065	Pro	Val	Leu	Ala	Asp 2070	Ile	Glu	Cys
Arg 2075	Ala	Ala	Gln	Leu	Pro	Asp 2080	Met	Pro	Leu	Glu	Glu 2085	Leu	Gly	Gln
Gln 2090	Val	Asp	Cys	Asp	Arg	Met 2095	Arg	Gly	Leu	Met	Cys 2100	Ala	Asn	Ser
Gln 2105	Gln	Ser	Pro	Pro	Leu	Cys 2110	His	Asp	Tyr	Glu	Leu 2115	Arg	Val	Leu
Cys 2120	Cys	Glu	Tyr	Val	Pro	Cys 2125	Gly	Pro	Ser	Pro	Ala 2130	Pro	Gly	Thr
Ser 2135	Pro	Gln	Pro	Ser	Leu	Ser 2140	Ala	Ser	Thr	Glu	Pro 2145	Ala	Val	Pro

[0026]

Thr Pro Thr Gln Thr Thr Ala Thr Glu Lys Thr Thr Leu Trp Val
 2150 2155 2160
 Thr Pro Ser Ile Arg Ser Thr Ala Ala Leu Thr Ser Gln Thr Gly
 2165 2170 2175
 Ser Ser Ser Gly Pro Val Thr Val Thr Pro Ser Ala Pro Gly Thr
 2180 2185 2190
 Thr Thr Cys Gln Pro Arg Cys Gln Trp Thr Glu Trp Phe Asp Glu
 2195 2200 2205
 Asp Tyr Pro Lys Ser Glu Gln Leu Gly Gly Asp Val Glu Ser Tyr
 2210 2215 2220
 Asp Lys Ile Arg Ala Ala Gly Gly His Leu Cys Gln Gln Pro Lys
 2225 2230 2235
 Asp Ile Glu Cys Gln Ala Glu Ser Phe Pro Asn Trp Thr Leu Ala
 2240 2245 2250
 Gln Val Gly Gln Lys Val His Cys Asp Val His Phe Gly Leu Val
 2255 2260 2265
 Cys Arg Asn Trp Glu Gln Glu Gly Val Phe Lys Met Cys Tyr Asn
 2270 2275 2280
 Tyr Arg Ile Arg Val Leu Cys Cys Ser Asp Asp His Cys Arg Gly
 2285 2290 2295
 Arg Ala Thr Thr Pro Pro Pro Thr Thr Glu Leu Glu Thr Ala Thr
 2300 2305 2310
 Thr Thr Thr Thr Gln Ala Leu Phe Ser Thr Pro Gln Pro Thr Ser
 2315 2320 2325
 Ser Pro Gly Leu Thr Arg Ala Pro Pro Ala Ser Thr Thr Ala Val
 2330 2335 2340
 Pro Thr Leu Ser Glu Gly Leu Thr Ser Pro Arg Tyr Thr Ser Thr
 2345 2350 2355
 Leu Gly Thr Ala Thr Thr Gly Gly Pro Thr Thr Pro Ala Gly Ser
 2360 2365 2370
 Thr Glu Pro Thr Val Pro Gly Val Ala Thr Ser Thr Leu Pro Thr
 2375 2380 2385
 Arg Ser Ala Leu Pro Gly Thr Thr Gly Ser Leu Gly Thr Trp Arg
 2390 2395 2400

[0027]

Pro Ser Gln Pro Pro Thr Leu Ala Pro Thr Thr Met Ala Thr Ser
 2405 2410 2415
 Arg Ala Arg Pro Thr Gly Thr Ala Ser Thr Ala Ser Lys Glu Pro
 2420 2425 2430
 Leu Thr Thr Ser Leu Ala Pro Thr Leu Thr Ser Glu Leu Ser Thr
 2435 2440 2445
 Ser Gln Ala Glu Thr Ser Thr Pro Arg Thr Glu Thr Thr Met Ser
 2450 2455 2460
 Pro Leu Thr Asn Thr Thr Thr Ser Gln Gly Thr Thr Arg Cys Gln
 2465 2470 2475
 Pro Lys Cys Glu Trp Thr Glu Trp Phe Asp Val Asp Phe Pro Thr
 2480 2485 2490
 Ser Gly Val Ala Gly Gly Asp Met Glu Thr Phe Glu Asn Ile Arg
 2495 2500 2505
 Ala Ala Gly Gly Lys Met Cys Trp Ala Pro Lys Ser Ile Glu Cys
 2510 2515 2520
 Arg Ala Glu Asn Tyr Pro Glu Val Ser Ile Asp Gln Val Gly Gln
 2525 2530 2535
 Val Leu Thr Cys Ser Leu Glu Thr Gly Leu Thr Cys Lys Asn Glu
 2540 2545 2550
 Asp Gln Thr Gly Arg Phe Asn Met Cys Phe Asn Tyr Asn Val Arg
 2555 2560 2565
 Val Leu Cys Cys Asp Asp Tyr Ser His Cys Pro Ser Thr Pro Ala
 2570 2575 2580
 Thr Ser Ser Thr Ala Thr Pro Ser Ser Thr Pro Gly Thr Thr Trp
 2585 2590 2595
 Ile Leu Thr Lys Pro Thr Thr Thr Ala Thr Thr Thr Ala Ser Thr
 2600 2605 2610
 Gly Ser Thr Ala Thr Pro Thr Ser Thr Leu Arg Thr Ala Pro Pro
 2615 2620 2625
 Pro Lys Val Leu Thr Thr Thr Ala Thr Thr Pro Thr Val Thr Ser
 2630 2635 2640
 Ser Lys Ala Thr Pro Ser Ser Ser Pro Gly Thr Ala Thr Ala Leu
 2645 2650 2655

[0028]

Pro Ala Leu Arg Ser Thr Ala Thr Thr Pro Thr Ala Thr Ser Val
 2660 2665 2670
 Thr Pro Ile Pro Ser Ser Ser Leu Gly Thr Thr Trp Thr Arg Leu
 2675 2680
 Ser Gln Thr Thr Thr Pro Thr Ala Thr Met Ser Thr Ala Thr Pro
 2690 2695 2700
 Ser Ser Thr Pro Glu Thr Ala His Thr Ser Thr Val Leu Thr Ala
 2705 2710 2715
 Thr Ala Thr Thr Thr Gly Ala Thr Gly Ser Val Ala Thr Pro Ser
 2720 2725 2730
 Ser Thr Pro Gly Thr Ala His Thr Thr Lys Val Pro Thr Thr Thr
 2735 2740 2745
 Thr Thr Gly Phe Thr Ala Thr Pro Ser Ser Ser Pro Gly Thr Ala
 2750 2755 2760
 Leu Thr Pro Pro Val Trp Ile Ser Thr Thr Thr Thr Pro Thr Thr
 2765 2770 2775
 Arg Gly Ser Thr Val Thr Pro Ser Ser Ile Pro Gly Thr Thr His
 2780 2785 2790
 Thr Ala Thr Val Leu Thr Thr Thr Thr Thr Val Ala Thr Gly
 2795 2800 2805
 Ser Met Ala Thr Pro Ser Ser Ser Thr Gln Thr Ser Gly Thr Thr
 2810 2815 2820
 His Thr Pro Pro Val Pro Asn Thr Met Ala Thr Thr His Gly Arg
 2825 2830 2835
 Ser Leu Pro Pro Ser Ser Pro His Thr Val Arg Thr Ala Trp Thr
 2840 2845 2850
 Ser Ala Thr Ser Gly Ile Leu Gly Thr Thr His Ile Thr Glu Pro
 2855 2860 2865
 Ser Thr Val Thr Ser His Thr Leu Ala Ala Thr Thr Gly Thr Thr
 2870 2875 2880
 Gln His Ser Thr Pro Ala Leu Ser Ser Pro His Pro Ser Ser Arg
 2885 2890 2895
 Thr Thr Glu Ser Pro Pro Ser Pro Gly Thr Thr Thr Pro Gly His
 2900 2905 2910
 Thr Thr Ala Thr Ser Arg Thr Thr Ala Thr Ala Thr Pro Ser Lys

[0029]

2915	2920	2925
Thr Arg 2930	Thr Ser Thr Leu 2935	Leu Pro Ser Ser Pro Thr 2940
Ile Thr 2945	Thr Val Val Thr Met 2950	Gly Cys Glu Pro Gln 2955
Ser Glu 2960	Trp Leu Asp Tyr Ser 2965	Tyr Pro Met Pro Gly 2970
Gly Asp 2975	Phe Asp Thr Tyr Ser 2980	Asn Ile Arg Ala Ala 2985
Val Cys 2990	Glu Gln Pro Leu Gly 2995	Leu Glu Cys Arg Ala 3000
Pro Gly 3005	Val Pro Leu Arg Glu 3010	Leu Gly Gln Val Val 3015
Leu Asp 3020	Phe Gly Leu Val Cys 3025	Arg Asn Arg Glu Gln 3030
Phe Lys 3035	Met Cys Phe Asn Tyr 3040	Glu Ile Arg Val Phe 3045
Tyr Gly 3050	His Cys Pro Ser Thr 3055	Pro Ala Thr Ser Ser 3060
Pro Ser 3065	Ser Thr Pro Gly Thr 3070	Thr Trp Ile Leu Thr 3075
Thr Thr 3080	Ala Thr Thr Thr Glu 3085	Ser Thr Gly Ser Thr 3090
Ser Ser 3095	Thr Pro Gly Thr Thr 3100	Trp Ile Leu Thr Glu 3105
Thr Ala 3110	Thr Val Thr Val Pro 3115	Thr Gly Ser Thr Ala 3120
Ser Thr 3125	Gln Ala Thr Ala Gly 3130	Thr Pro His Val Ser 3135
Thr Thr 3140	Pro Thr Val Thr Ser 3145	Ser Lys Ala Thr Pro 3150
Pro Gly 3155	Thr Ala Thr Ala Leu 3160	Pro Ala Leu Arg Ser 3165
Thr Pro 3170	Thr Ala Thr Ser Phe 3175	Thr Ala Ile Pro Ser 3180

[0030]

Gly Thr Thr Trp Thr Arg Leu Ser Gln Thr Thr Thr Pro Thr Ala
 3185 3190 3195
 Thr Met Ser Thr Ala Thr Pro Ser Ser Thr Pro Glu Thr Val His
 3200 3205 3210
 Thr Ser Thr Val Leu Thr Thr Thr Ala Thr Thr Thr Gly Ala Thr
 3215 3220 3225
 Gly Ser Val Ala Thr Pro Ser Ser Thr Pro Gly Thr Ala His Thr
 3230 3235 3240
 Thr Lys Val Leu Thr Thr Thr Thr Thr Gly Phe Thr Ala Thr Pro
 3245 3250 3255
 Ser Ser Ser Pro Gly Thr Ala Arg Thr Leu Pro Val Trp Ile Ser
 3260 3265 3270
 Thr Thr Thr Thr Pro Thr Thr Arg Gly Ser Thr Val Thr Pro Ser
 3275 3280 3285
 Ser Ile Pro Gly Thr Thr His Thr Pro Thr Val Leu Thr Thr Thr
 3290 3295 3300
 Thr Thr Thr Val Ala Thr Gly Ser Met Ala Thr Pro Ser Ser Ser
 3305 3310 3315
 Thr Gln Thr Ser Gly Thr Thr His Thr Pro Pro Val Pro Asn Thr
 3320 3325 3330
 Thr Ala Thr Thr His Gly Arg Ser Leu Ser Pro Ser Ser Pro His
 3335 3340 3345
 Thr Val Arg Thr Ala Trp Thr Ser Ala Thr Ser Gly Thr Leu Gly
 3350 3355 3360
 Thr Thr His Ile Thr Glu Pro Ser Thr Gly Thr Ser His Thr Pro
 3365 3370 3375
 Ala Ala Thr Thr Gly Thr Thr Gln His Ser Thr Pro Ala Leu Ser
 3380 3385 3390
 Ser Pro His Pro Ser Ser Arg Thr Thr Glu Ser Pro Pro Ser Pro
 3395 3400 3405
 Gly Thr Thr Thr Pro Gly His Thr Arg Ala Thr Ser Arg Thr Thr
 3410 3415 3420
 Ala Thr Ala Thr Pro Ser Lys Thr Arg Thr Ser Thr Leu Leu Pro
 3425 3430 3435

[0031]

Ser	Ser	Pro	Thr	Ser	Ala	Pro	Ile	Thr	Thr	Val	Val	Thr	Met	Gly
	3440					3445					3450			
Cys	Glu	Pro	Gln	Cys	Ala	Trp	Ser	Glu	Trp	Leu	Asp	Tyr	Ser	Tyr
	3455					3460					3465			
Pro	Met	Pro	Gly	Pro	Ser	Gly	Gly	Asp	Phe	Asp	Thr	Tyr	Ser	Asn
	3470					3475					3480			
Ile	Arg	Ala	Ala	Gly	Gly	Ala	Val	Cys	Glu	Gln	Pro	Leu	Gly	Leu
	3485					3490					3495			
Glu	Cys	Arg	Ala	Gln	Ala	Gln	Pro	Gly	Val	Pro	Leu	Arg	Glu	Leu
	3500					3505					3510			
Gly	Gln	Val	Val	Glu	Cys	Ser	Leu	Asp	Phe	Gly	Leu	Val	Cys	Arg
	3515					3520					3525			
Asn	Arg	Glu	Gln	Val	Gly	Lys	Phe	Lys	Met	Cys	Phe	Asn	Tyr	Glu
	3530					3535					3540			
Ile	Arg	Val	Phe	Cys	Cys	Asn	Tyr	Gly	His	Cys	Pro	Ser	Thr	Pro
	3545					3550					3555			
Ala	Thr	Ser	Ser	Thr	Ala	Thr	Pro	Ser	Ser	Thr	Pro	Gly	Thr	Thr
	3560					3565					3570			
Trp	Ile	Leu	Thr	Glu	Gln	Thr	Thr	Ala	Ala	Thr	Thr	Thr	Ala	Thr
	3575					3580					3585			
Thr	Gly	Ser	Thr	Ala	Ile	Pro	Ser	Ser	Thr	Pro	Gly	Thr	Ala	Pro
	3590					3595					3600			
Pro	Pro	Lys	Val	Leu	Thr	Ser	Thr	Ala	Thr	Thr	Pro	Thr	Ala	Thr
	3605					3610					3615			
Ser	Ser	Lys	Ala	Thr	Ser	Ser	Ser	Ser	Pro	Arg	Thr	Ala	Thr	Thr
	3620					3625					3630			
Leu	Pro	Val	Leu	Thr	Ser	Thr	Ala	Thr	Lys	Ser	Thr	Ala	Thr	Ser
	3635					3640					3645			
Phe	Thr	Pro	Ile	Pro	Ser	Phe	Thr	Leu	Gly	Thr	Thr	Gly	Thr	Leu
	3650					3655					3660			
Pro	Glu	Gln	Thr	Thr	Thr	Pro	Met	Ala	Thr	Met	Ser	Thr	Ile	His
	3665					3670					3675			
Pro	Ser	Ser	Thr	Pro	Glu	Thr	Thr	His	Thr	Ser	Thr	Val	Leu	Thr
	3680					3685					3690			

[0032]

Thr Lys Ala Thr Thr Thr Arg Ala Thr Ser Ser Met Ser Thr Pro
 3695 3700 3705
 Ser Ser Thr Pro Gly Thr Thr Trp Ile Leu Thr Glu Leu Thr Thr
 3710 3715 3720
 Ala Ala Thr Thr Thr Ala Ala Thr Gly Pro Thr Ala Thr Pro Ser
 3725 3730 3735
 Ser Thr Pro Gly Thr Thr Trp Ile Leu Thr Glu Pro Ser Thr Thr
 3740 3745 3750
 Ala Thr Val Thr Val Pro Thr Gly Ser Thr Ala Thr Ala Ser Ser
 3755 3760 3765
 Thr Arg Ala Thr Ala Gly Thr Leu Lys Val Leu Thr Ser Thr Ala
 3770 3775 3780
 Thr Thr Pro Thr Val Ile Ser Ser Arg Ala Thr Pro Ser Ser Ser
 3785 3790 3795
 Pro Gly Thr Ala Thr Ala Leu Pro Ala Leu Arg Ser Thr Ala Thr
 3800 3805 3810
 Thr Pro Thr Ala Thr Ser Val Thr Ala Ile Pro Ser Ser Ser Leu
 3815 3820 3825
 Gly Thr Ala Trp Thr Arg Leu Ser Gln Thr Thr Thr Pro Thr Ala
 3830 3835 3840
 Thr Met Ser Thr Ala Thr Pro Ser Ser Thr Pro Glu Thr Val His
 3845 3850 3855
 Thr Ser Thr Val Leu Thr Thr Thr Thr Thr Thr Arg Ala Thr
 3860 3865 3870
 Gly Ser Val Ala Thr Pro Ser Ser Thr Pro Gly Thr Ala His Thr
 3875 3880 3885
 Thr Lys Val Pro Thr Thr Thr Thr Thr Gly Phe Thr Ala Thr Pro
 3890 3895 3900
 Ser Ser Ser Pro Gly Thr Ala Leu Thr Pro Pro Val Trp Ile Ser
 3905 3910 3915
 Thr Thr Thr Thr Pro Thr Thr Arg Gly Ser Thr Val Thr Pro Ser
 3920 3925 3930
 Ser Ile Pro Gly Thr Thr His Thr Ala Thr Val Leu Thr Thr Thr
 3935 3940 3945
 Thr Thr Thr Val Ala Thr Gly Ser Met Ala Thr Pro Ser Ser Ser

[0033]

Trp Ile Leu Thr Lys Leu Thr Thr Thr Ala Thr Thr Thr Glu Ser
 4220 4225 4230

Thr Gly Ser Thr Ala Thr Pro Ser Ser Thr Pro Gly Thr Thr Trp
 4235 4240 4245

Ile Leu Thr Glu Pro Ser Thr Thr Ala Thr Val Thr Val Pro Thr
 4250 4255 4260

Gly Ser Thr Ala Thr Ala Ser Ser Thr Gln Ala Thr Ala Gly Thr
 4265 4270 4275

Pro His Val Ser Thr Thr Ala Thr Thr Pro Thr Val Thr Ser Ser
 4280 4285 4290

Lys Ala Thr Pro Phe Ser Ser Pro Gly Thr Ala Thr Ala Leu Pro
 4295 4300 4305

Ala Leu Arg Ser Thr Ala Thr Thr Pro Thr Ala Thr Ser Phe Thr
 4310 4315 4320

Ala Ile Pro Ser Ser Ser Leu Gly Thr Thr Trp Thr Arg Leu Ser
 4325 4330 4335

Gln Thr Thr Thr Pro Thr Ala Thr Met Ser Thr Ala Thr Pro Ser
 4340 4345 4350

Ser Thr Pro Glu Thr Ala His Thr Ser Thr Val Leu Thr Thr Thr
 4355 4360 4365

Ala Thr Thr Thr Arg Ala Thr Gly Ser Val Ala Thr Pro Ser Ser
 4370 4375 4380

Thr Pro Gly Thr Ala His Thr Thr Lys Val Pro Thr Thr Thr Thr
 4385 4390 4395

Thr Gly Phe Thr Val Thr Pro Ser Ser Ser Pro Gly Thr Ala Arg
 4400 4405 4410

Thr Pro Pro Val Trp Ile Ser Thr Thr Thr Thr Pro Thr Thr Ser
 4415 4420 4425

Gly Ser Thr Val Thr Pro Ser Ser Val Pro Gly Thr Thr His Thr
 4430 4435 4440

Pro Thr Val Leu Thr Thr Thr Thr Thr Thr Val Ala Thr Gly Ser
 4445 4450 4455

Met Ala Thr Pro Ser Ser Ser Thr Gln Thr Ser Gly Thr Thr His
 4460 4465 4470

[0035]

Thr	Pro	Pro	Val	Pro	Asn	Thr	Thr	Ala	Thr	Thr	His	Gly	Arg	Ser
	4475					4480					4485			
Leu	Ser	Pro	Ser	Ser	Pro	His	Thr	Val	Arg	Thr	Ala	Trp	Thr	Ser
	4490					4495					4500			
Ala	Thr	Ser	Gly	Thr	Leu	Gly	Thr	Thr	His	Ile	Thr	Glu	Pro	Ser
	4505					4510					4515			
Thr	Gly	Thr	Ser	His	Thr	Pro	Ala	Ala	Thr	Thr	Gly	Thr	Thr	Gln
	4520					4525					4530			
His	Ser	Thr	Pro	Ala	Leu	Ser	Ser	Pro	His	Pro	Ser	Ser	Arg	Thr
	4535					4540					4545			
Thr	Glu	Ser	Pro	Pro	Ser	Pro	Gly	Thr	Thr	Thr	Pro	Gly	His	Thr
	4550					4555					4560			
Thr	Ala	Thr	Ser	Arg	Thr	Thr	Ala	Thr	Ala	Thr	Pro	Ser	Lys	Thr
	4565					4570					4575			
Arg	Thr	Ser	Thr	Leu	Leu	Pro	Ser	Ser	Pro	Thr	Ser	Ala	Pro	Ile
	4580					4585					4590			
Thr	Thr	Val	Val	Thr	Thr	Gly	Cys	Glu	Pro	Gln	Cys	Ala	Trp	Ser
	4595					4600					4605			
Glu	Trp	Leu	Asp	Tyr	Ser	Tyr	Pro	Met	Pro	Gly	Pro	Ser	Gly	Gly
	4610					4615					4620			
Asp	Phe	Asp	Thr	Tyr	Ser	Asn	Ile	Arg	Ala	Ala	Gly	Gly	Ala	Val
	4625					4630					4635			
Cys	Glu	Gln	Pro	Leu	Gly	Leu	Glu	Cys	Arg	Ala	Gln	Ala	Gln	Pro
	4640					4645					4650			
Gly	Val	Pro	Leu	Gly	Glu	Leu	Gly	Gln	Val	Val	Glu	Cys	Ser	Leu
	4655					4660					4665			
Asp	Phe	Gly	Leu	Val	Cys	Arg	Asn	Arg	Glu	Gln	Val	Gly	Lys	Phe
	4670					4675					4680			
Lys	Met	Cys	Phe	Asn	Tyr	Glu	Ile	Arg	Val	Phe	Cys	Cys	Asn	Tyr
	4685					4690					4695			
Gly	His	Cys	Pro	Ser	Thr	Pro	Ala	Thr	Ser	Ser	Thr	Ala	Met	Pro
	4700					4705					4710			
ser	ser	Thr	Pro	Gly	Thr	Thr	Trp	Ile	Leu	Thr	Glu	Leu	Thr	Thr
	4715					4720					4725			

[0036]

Thr Ala Thr Thr Thr Ala Ser Thr Gly Ser Thr Ala Thr Pro Ser
 4730 4735 4740
 Ser Thr Pro Gly Thr Ala Pro Pro Pro Lys Val Leu Thr Ser Pro
 4745 4750
 Ala Thr Thr Pro Thr Ala Thr Ser Ser Lys Ala Thr Ser Ser Ser
 4760 4765 4770
 Ser Pro Arg Thr Ala Thr Thr Leu Pro Val Leu Thr Ser Thr Ala
 4775 4780 4785
 Thr Lys Ser Thr Ala Thr Ser Val Thr Pro Ile Pro Ser Ser Thr
 4790 4795 4800
 Leu Gly Thr Thr Gly Thr Leu Pro Glu Gln Thr Thr Thr Pro Val
 4805 4810
 Ala Thr Met Ser Thr Ile His Pro Ser Ser Thr Pro Glu Thr Thr
 4820 4825 4830
 His Thr Ser Thr Val Leu Thr Thr Lys Ala Thr Thr Thr Arg Ala
 4835 4840 4845
 Thr Ser Ser Thr Ser Thr Pro Ser Ser Thr Pro Gly Thr Thr Trp
 4850 4855 4860
 Ile Leu Thr Glu Leu Thr Thr Ala Ala Thr Thr Thr Ala Ala Thr
 4865 4870 4875
 Gly Pro Thr Ala Thr Pro Ser Ser Thr Pro Gly Thr Thr Trp Ile
 4880 4885 4890
 Leu Thr Glu Leu Thr Thr Thr Ala Thr Thr Thr Ala Ser Thr Gly
 4895 4900 4905
 Ser Thr Ala Thr Pro Ser Ser Thr Pro Gly Thr Thr Trp Ile Leu
 4910 4915 4920
 Thr Glu Pro Ser Thr Thr Ala Thr Val Thr Val Pro Thr Gly Ser
 4925 4930 4935
 Thr Ala Thr Ala Ser Ser Thr Gln Ala Thr Ala Gly Thr Pro His
 4940 4945 4950
 Val Ser Thr Thr Ala Thr Thr Pro Thr Val Thr Ser Ser Lys Ala
 4955 4960 4965
 Thr Pro Ser Ser Ser Pro Gly Thr Ala Thr Ala Leu Pro Ala Leu
 4970 4975 4980
 Arg Ser Thr Ala Thr Thr Pro Thr Ala Thr Ser Phe Thr Ala Ile

[0037]

4985	4990	4995
Pro Ser Ser Ser Leu Gly	Thr Thr Trp Thr Arg	Leu Ser Gln Thr
5000	5005	5010
Thr Thr Pro Thr Ala Thr	Met Ser Thr Ala Thr	Pro Ser Ser Thr
5015	5020	5025
Pro Glu Thr Val His Thr	Ser Thr Val Leu Thr	Ala Thr Ala Thr
5030	5035	5040
Thr Thr Gly Ala Thr Gly	Ser Val Ala Thr Pro	Ser Ser Thr Pro
5045	5050	5055
Gly Thr Ala His Thr Thr	Lys Val Pro Thr Thr	Thr Thr Thr Gly
5060	5065	5070
Phe Thr Ala Thr Pro Ser	Ser Ser Pro Gly Thr	Ala Leu Thr Pro
5075	5080	5085
Pro Thr Thr Thr Pro Met	Ser Thr Met Ser Thr	Ile His Thr Ser
5090	5095	5100
Ser Thr Pro Glu Thr Thr	His Thr Ser Thr Val	Leu Thr Thr Thr
5105	5110	5115
Ala Thr Met Thr Arg Ala	Thr Asn Ser Thr Ala	Thr Pro Ser Ser
5120	5125	5130
Thr Leu Gly Thr Thr Arg	Ile Leu Thr Glu Leu	Thr Thr Thr Ala
5135	5140	5145
Thr Thr Thr Ala Ala Thr	Gly Ser Thr Ala Thr	Leu Ser Ser Thr
5150	5155	5160
Pro Gly Thr Thr Trp Ile	Leu Thr Glu Pro Ser	Thr Ile Ala Thr
5165	5170	5175
Val Met Val Pro Thr Gly	Ser Thr Ala Thr Ala	Ser Ser Thr Leu
5180	5185	5190
Gly Thr Ala His Thr Pro	Lys Val Val Thr Thr	Met Ala Thr Met
5195	5200	5205
Pro Thr Ala Thr Ala Ser	Thr Val Pro Ser Ser	Ser Thr Val Gly
5210	5215	5220
Thr Thr Arg Thr Pro Ala	Val Leu Pro Ser Ser	Leu Pro Thr Phe
5225	5230	5235
Ser Val Ser Thr Val Ser	Ser Ser Val Leu Thr	Thr Leu Arg Pro
5240	5245	5250

[0038]

Thr Gly Phe Pro Ser Ser His Phe Ser Thr Pro Cys Phe Cys Arg
 5255 5260 5265
 Ala Phe Gly Gln Phe Phe Ser Pro Gly Glu Val Ile Tyr Asn Lys
 5270 5275 5280
 Thr Asp Arg Ala Gly Cys His Phe Tyr Ala Val Cys Asn Gln His
 5285 5290 5295
 Cys Asp Ile Asp Arg Phe Gln Gly Ala Cys Pro Thr Ser Pro Pro
 5300 5305 5310
 Pro Val Ser Ser Ala Pro Leu Ser Ser Pro Ser Pro Ala Pro Gly
 5315 5320 5325
 Cys Asp Asn Ala Ile Pro Leu Arg Gln Val Asn Glu Thr Trp Thr
 5330 5335 5340
 Leu Glu Asn Cys Thr Val Ala Arg Cys Val Gly Asp Asn Arg Val
 5345 5350 5355
 Val Leu Leu Asp Pro Lys Pro Val Ala Asn Val Thr Cys Val Asn
 5360 5365 5370
 Lys His Leu Pro Ile Lys Val Ser Asp Pro Ser Gln Pro Cys Asp
 5375 5380 5385
 Phe His Tyr Glu Cys Glu Cys Glu Cys Val Gly Gly Arg Gly Ile
 5390 5395 5400
 Thr Pro Gly Ala Gly Ile Cys Ser Met Trp Gly Gly Ser His Tyr
 5405 5410 5415
 Ser Thr Phe Asp Gly Thr Ser Tyr Thr Phe Arg Gly Asn Cys Thr
 5420 5425 5430
 Tyr Val Leu Met Arg Glu Ile His Ala Arg Phe Gly Asn Leu Ser
 5435 5440 5445
 Leu Tyr Leu Asp Asn His Tyr Cys Thr Ala Ser Ala Thr Ala Ala
 5450 5455 5460
 Ala Ala Ala Ala Arg Cys Pro Arg Ala Leu Ser Ile His Tyr Lys
 5465 5470 5475
 Ser Met Asp Ile Val Leu Thr Val Thr Met Val His Gly Lys Glu
 5480 5485 5490
 Glu Gly Leu Ile Leu Phe Asp Gln Ile Pro Val Ser Ser Gly Phe
 5495 5500 5505

[0039]

Ser	Lys	Asn	Gly	Val	Leu	Val	Ser	Val	Leu	Gly	Thr	Thr	Thr	Met
	5510					5515					5520			
Arg	Val	Asp	Ile	Pro	Ala	Leu	Gly	Val	Ser	Val	Thr	Phe	Asn	Gly
	5525					5530					5535			
Gln	Val	Phe	Gln	Ala	Arg	Leu	Pro	Tyr	Ser	Leu	Phe	His	Asn	Asn
	5540					5545					5550			
Thr	Glu	Gly	Gln	Cys	Gly	Thr	Cys	Thr	Asn	Asn	Gln	Arg	Asp	Asp
	5555					5560					5565			
Cys	Leu	Gln	Arg	Asp	Gly	Thr	Thr	Ala	Ala	Ser	Cys	Lys	Asp	Met
	5570					5575					5580			
Ala	Lys	Thr	Trp	Leu	Val	Pro	Asp	Ser	Arg	Lys	Asp	Gly	Cys	Trp
	5585					5590					5595			
Ala	Pro	Thr	Gly	Thr	Pro	Pro	Thr	Ala	Ser	Pro	Ala	Ala	Pro	Val
	5600					5605					5610			
Ser	Ser	Thr	Pro	Thr	Pro	Thr	Pro	Cys	Pro	Pro	Gln	Pro	Leu	Cys
	5615					5620					5625			
Asp	Leu	Met	Leu	Ser	Gln	Val	Phe	Ala	Glu	Cys	His	Asn	Leu	Val
	5630					5635					5640			
Pro	Pro	Gly	Pro	Phe	Phe	Asn	Ala	Cys	Ile	Ser	Asp	His	Cys	Arg
	5645					5650					5655			
Gly	Arg	Leu	Glu	Val	Pro	Cys	Gln	Ser	Leu	Glu	Ala	Tyr	Ala	Glu
	5660					5665					5670			
Leu	Cys	Arg	Ala	Arg	Gly	Val	Cys	Ser	Asp	Trp	Arg	Gly	Ala	Thr
	5675					5680					5685			
Gly	Gly	Leu	Cys	Asp	Leu	Thr	Cys	Pro	Pro	Thr	Lys	Val	Tyr	Lys
	5690					5695					5700			
Pro	Cys	Gly	Pro	Ile	Gln	Pro	Ala	Thr	Cys	Asn	Ser	Arg	Asn	Gln
	5705					5710					5715			
Ser	Pro	Gln	Leu	Glu	Gly	Met	Ala	Glu	Gly	Cys	Phe	Cys	Pro	Glu
	5720					5725					5730			
Asp	Gln	Ile	Leu	Phe	Asn	Ala	His	Met	Gly	Ile	Cys	Val	Gln	Ala
	5735					5740					5745			
Cys	Pro	Cys	Val	Gly	Pro	Asp	Gly	Phe	Pro	Lys	Phe	Pro	Gly	Glu
	5750					5755					5760			

[0040]

Arg Trp Val Ser Asn Cys Gln Ser Cys Val Cys Asp Glu Gly Ser
 5765 5770 5775
 Val Ser Val Gln Cys Lys Pro Leu Pro Cys Asp Ala Gln Gly Gln
 5780 5785 5790
 Pro Pro Pro Cys Asn Arg Pro Gly Phe Val Thr Val Thr Arg Pro
 5795 5800 5805
 Arg Ala Glu Asn Pro Cys Cys Pro Glu Thr Val Cys Val Cys Asn
 5810 5815 5820
 Thr Thr Thr Cys Pro Gln Ser Leu Pro Val Cys Pro Pro Gly Gln
 5825 5830 5835
 Glu Ser Ile Cys Thr Gln Glu Glu Gly Asp Cys Cys Pro Thr Phe
 5840 5845 5850
 Arg Cys Arg Pro Gln Leu Cys Ser Tyr Asn Gly Thr Phe Tyr Gly
 5855 5860 5865
 Val Gly Ala Thr Phe Pro Gly Ala Leu Pro Cys His Met Cys Thr
 5870 5875 5880
 Cys Leu Ser Gly Asp Thr Gln Asp Pro Thr Val Gln Cys Gln Glu
 5885 5890 5895
 Asp Ala Cys Asn Asn Thr Thr Cys Pro Gln Gly Phe Glu Tyr Lys
 5900 5905 5910
 Arg Val Ala Gly Gln Cys Cys Gly Glu Cys Val Gln Thr Ala Cys
 5915 5920 5925
 Leu Thr Pro Asp Gly Gln Pro Val Gln Leu Asn Glu Thr Trp Val
 5930 5935 5940
 Asn Ser His Val Asp Asn Cys Thr Val Tyr Leu Cys Glu Ala Glu
 5945 5950 5955
 Gly Gly Val His Leu Leu Thr Pro Gln Pro Ala Ser Cys Pro Asp
 5960 5965 5970
 Val Ser Ser Cys Arg Gly Ser Leu Arg Lys Thr Gly Cys Cys Tyr
 5975 5980 5985
 Ser Cys Glu Glu Asp Ser Cys Gln Val Arg Ile Asn Thr Thr Ile
 5990 5995 6000
 Leu Trp His Gln Gly Cys Glu Thr Glu Val Asn Ile Thr Phe Cys
 6005 6010 6015
 Glu Gly Ser Cys Pro Gly Ala Ser Lys Tyr Ser Ala Glu Ala Gln

[0041]

Asn Arg Gln Cys Gln Gln Leu Ser Gln His Val Arg Thr Gly Ser Ala
 50 55 60

Ala Asn Thr Ala Thr Gly Thr Thr Ser Thr Asn Val Val Glu Pro Arg
 65 70 75 80

Met Tyr Leu Ser Cys Ser Thr Asn Pro Glu Met Thr Ser Ile Glu Ser
 85 90 95

Ser Val Thr Ser Asp Thr Pro Gly Val Ser Ser Thr Arg Met Thr Pro
 100 105 110

Thr Glu Ser Arg Thr Thr Ser Glu Ser Thr Ser Asp Ser Thr Thr Leu
 115 120 125

Phe Pro Ser Ser Thr Glu Asp Thr Ser Ser Pro Thr Thr Pro Glu Gly
 130 135 140

Thr Asp Val Pro Met Ser Thr Pro Ser Glu Glu Ser Ile Ser Ser Thr
 145 150 155 160

Met Ala Phe Val Ser Thr Ala Pro Leu Pro Ser Phe Glu Ala Tyr Thr
 165 170 175

Ser Leu Thr Tyr Lys Val Asp Met Ser Thr Pro Leu Thr Thr Ser Thr
 180 185 190

Gln Ala Ser Ser Ser Pro Thr Thr Pro Glu Ser Thr Thr Ile Pro Lys
 195 200 205

Ser Thr Asn Ser Glu Gly Ser Thr Pro Leu Thr Ser Met Pro Ala Ser
 210 215 220

Thr Met Lys Val Ala Ser Ser Glu Ala Ile Thr Leu Leu Thr Thr Pro
 225 230 235 240

Val Glu Ile Ser Thr Pro Val Thr Ile Ser Ala Gln Ala Ser Ser Ser
 245 250 255

Pro Thr Thr Ala Glu Gly Pro Ser Leu Ser Asn Ser Ala Pro Ser Gly
 260 265 270

Gly Ser Thr Pro Leu Thr Arg Met Pro Leu Ser Val Met Leu Val Val
 275 280 285

Ser Ser Glu Ala Ser Thr Leu Ser Thr Thr Pro Ala Ala Thr Asn Ile
 290 295 300

Pro Val Ile Thr Ser Thr Glu Ala Ser Ser Ser Pro Thr Thr Ala Glu
 305 310 315 320

[0043]

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

[0044]

595			600			605									
Ser	Ser	Ser	Thr	Thr	Ala	Glu	Gly	Thr	Ser	Met	Pro	Thr	Ser	Thr	Tyr
	610					615					620				
Ser	Glu	Arg	Gly	Thr	Thr	Ile	Thr	Ser	Met	Ser	Val	Ser	Thr	Thr	Leu
625					630					635					640
Val	Ala	Ser	Ser	Glu	Ala	Ser	Thr	Leu	Ser	Thr	Thr	Pro	Val	Asp	Ser
				645					650					655	
Asn	Thr	Pro	Val	Thr	Thr	Ser	Thr	Glu	Ala	Thr	Ser	Ser	Ser	Thr	Thr
			660					665						670	
Ala	Glu	Gly	Thr	Ser	Met	Pro	Thr	Ser	Thr	Tyr	Thr	Glu	Gly	Ser	Thr
		675					680					685			
Pro	Leu	Thr	Ser	Met	Pro	Val	Asn	Thr	Thr	Leu	Val	Ala	Ser	Ser	Glu
	690					695					700				
Ala	Ser	Thr	Leu	Ser	Thr	Thr	Pro	Val	Asp	Thr	Ser	Thr	Pro	Val	Thr
	705				710					715					720
Thr	Ser	Thr	Glu	Ala	Ser	Ser	Ser	Pro	Thr	Thr	Ala	Asp	Gly	Ala	Ser
				725					730					735	
Met	Pro	Thr	Ser	Thr	Pro	Ser	Glu	Gly	Ser	Thr	Pro	Leu	Thr	Ser	Met
			740					745					750		
Pro	Val	Ser	Lys	Thr	Leu	Leu	Thr	Ser	Ser	Glu	Ala	Ser	Thr	Leu	Ser
		755					760					765			
Thr	Thr	Pro	Leu	Asp	Thr	Ser	Thr	His	Ile	Thr	Thr	Ser	Thr	Glu	Ala
	770					775									
Ser	Cys	Ser	Pro	Thr	Thr	Thr	Glu	Gly	Thr	Ser	Met	Pro	Ile	Ser	Thr
	785				790					795					800
Pro	Ser	Glu	Gly	Ser	Pro	Leu	Leu	Thr	Ser	Ile	Pro	Val	Ser	Ile	Thr
				805					810					815	
Pro	Val	Thr	Ser	Pro	Glu	Ala	Ser	Thr	Leu	Ser	Thr	Thr	Pro	Val	Asp
			820					825					830		
Ser	Asn	Ser	Pro	Val	Thr	Thr	Ser	Thr	Glu	Val	Ser	Ser	Ser	Pro	Thr
		835					840					845			
Pro	Ala	Glu	Gly	Thr	Ser	Met	Pro	Thr	Ser	Thr	Tyr	Ser	Glu	Gly	Arg
	850					855					860				
Thr	Pro	Leu	Thr	Ser	Met	Pro	Val	Ser	Thr	Thr	Leu	Val	Ala	Thr	Ser
	865				870					875					880

[0045]

Ala Ile Ser Thr Leu Ser Thr Thr Pro Val Asp Thr Ser Thr Pro Val
885 890 895

Thr Asn Ser Thr Glu Ala Arg Ser Ser Pro Thr Thr Ser Glu Gly Thr
900 905 910

Ser Met Pro Thr Ser Thr Pro Gly Glu Gly Ser Thr Pro Leu Thr Ser
915 920 925

Met Pro Asp Ser Thr Thr Pro Val Val Ser Ser Glu Ala Arg Thr Leu
930 935 940

Ser Ala Thr Pro Val Asp Thr Ser Thr Pro Val Thr Thr Ser Thr Glu
945 950 955 960

Ala Thr Ser Ser Pro Thr Thr Ala Glu Gly Thr Ser Ile Pro Thr Ser
965 970 975

Thr Pro Ser Glu Gly Thr Thr Pro Leu Thr Ser Thr Pro Val Ser His
980 985 990

Thr Leu Val Ala Asn Ser Glu Ala Ser Thr Leu Ser Thr Thr Pro Val
995 1000 1005

Asp Ser Asn Thr Pro Leu Thr Thr Ser Thr Glu Ala Ser Ser Pro
1010 1015 1020

Pro Pro Thr Ala Glu Gly Thr Ser Met Pro Thr Ser Thr Pro Ser
1025 1030 1035

Glu Gly Ser Thr Pro Leu Thr Arg Met Pro Val Ser Thr Thr Met
1040 1045 1050

Val Ala Ser Ser Glu Thr Ser Thr Leu Ser Thr Thr Pro Ala Asp
1055 1060 1065

Thr Ser Thr Pro Val Thr Thr Tyr Ser Gln Ala Ser Ser Ser Ser
1070 1075 1080

Thr Thr Ala Asp Gly Thr Ser Met Pro Thr Ser Thr Tyr Ser Glu
1085 1090 1095

Gly Ser Thr Pro Leu Thr Ser Val Pro Val Ser Thr Arg Leu Val
1100 1105 1110

Val Ser Ser Glu Ala Ser Thr Leu Ser Thr Thr Pro Val Asp Thr
1115 1120 1125

Ser Ile Pro Val Thr Thr Ser Thr Glu Ala Ser Ser Ser Pro Thr
1130 1135 1140

[0046]

Thr Ala Glu Gly Thr Ser Ile Pro Thr Ser Pro Pro Ser Glu Gly
 1145 1150 1155
 Thr Thr Pro Leu Ala Ser Met Pro Val Ser Thr Thr Leu Val Val
 1160 1165 1170
 Ser Ser Glu Ala Asn Thr Leu Ser Thr Thr Pro Val Asp Ser Lys
 1175 1180 1185
 Thr Gln Val Ala Thr Ser Thr Glu Ala Ser Ser Pro Pro Pro Thr
 1190 1195 1200
 Ala Glu Val Thr Ser Met Pro Thr Ser Thr Pro Gly Glu Arg Ser
 1205 1210 1215
 Thr Pro Leu Thr Ser Met Pro Val Arg His Thr Pro Val Ala Ser
 1220 1225 1230
 Ser Glu Ala Ser Thr Leu Ser Thr Ser Pro Val Asp Thr Ser Thr
 1235 1240 1245
 Pro Val Thr Thr Ser Ala Glu Thr Ser Ser Ser Pro Thr Thr Ala
 1250 1255 1260
 Glu Gly Thr Ser Leu Pro Thr Ser Thr Thr Ser Glu Gly Ser Thr
 1265 1270 1275
 Leu Leu Thr Ser Ile Pro Val Ser Thr Thr Leu Val Thr Ser Pro
 1280 1285 1290
 Glu Ala Ser Thr Leu Leu Thr Thr Pro Val Asp Thr Lys Gly Pro
 1295 1300 1305
 Val Val Thr Ser Asn Glu Val Ser Ser Ser Pro Thr Pro Ala Glu
 1310 1315 1320
 Gly Thr Ser Met Pro Thr Ser Thr Tyr Ser Glu Gly Arg Thr Pro
 1325 1330 1335
 Leu Thr Ser Ile Pro Val Asn Thr Thr Leu Val Ala Ser Ser Ala
 1340 1345 1350
 Ile Ser Ile Leu Ser Thr Thr Pro Val Asp Asn Ser Thr Pro Val
 1355 1360 1365
 Thr Thr Ser Thr Glu Ala Cys Ser Ser Pro Thr Thr Ser Glu Gly
 1370 1375 1380
 Thr Ser Met Pro Asn Ser Asn Pro Ser Glu Gly Thr Thr Pro Leu
 1385 1390 1395

[0047]

Thr Ser Ile Pro Val Ser Thr Thr Pro Val Val Ser Ser Glu Ala
 1400 1405 1410
 Ser Thr Leu Ser Ala Thr Pro Val Asp Thr Ser Thr Pro Gly Thr
 1415 1420 1425
 Thr Ser Ala Glu Ala Thr Ser Ser Pro Thr Thr Ala Glu Gly Ile
 1430 1435 1440
 Ser Ile Pro Thr Ser Thr Pro Ser Glu Gly Lys Thr Pro Leu Lys
 1445 1450 1455
 Ser Ile Pro Val Ser Asn Thr Pro Val Ala Asn Ser Glu Ala Ser
 1460 1465 1470
 Thr Leu Ser Thr Thr Pro Val Asp Ser Asn Ser Pro Val Val Thr
 1475 1480 1485
 Ser Thr Ala Val Ser Ser Ser Pro Thr Pro Ala Glu Gly Thr Ser
 1490 1495 1500
 Ile Ala Ile Ser Thr Pro Ser Glu Gly Ser Thr Ala Leu Thr Ser
 1505 1510 1515
 Ile Pro Val Ser Thr Thr Thr Val Ala Ser Ser Glu Ile Asn Ser
 1520 1525 1530
 Leu Ser Thr Thr Pro Ala Val Thr Ser Thr Pro Val Thr Thr Tyr
 1535 1540 1545
 Ser Gln Ala Ser Ser Ser Pro Thr Thr Ala Asp Gly Thr Ser Met
 1550 1555 1560
 Gln Thr Ser Thr Tyr Ser Glu Gly Ser Thr Pro Leu Thr Ser Leu
 1565 1570 1575
 Pro Val Ser Thr Met Leu Val Val Ser Ser Glu Ala Asn Thr Leu
 1580 1585 1590
 Ser Thr Thr Pro Ile Asp Ser Lys Thr Gln Val Thr Ala Ser Thr
 1595 1600 1605
 Glu Ala Ser Ser Ser Thr Thr Ala Glu Gly Ser Ser Met Thr Ile
 1610 1615 1620
 Ser Thr Pro Ser Glu Gly Ser Pro Leu Leu Thr Ser Ile Pro Val
 1625 1630 1635
 Ser Thr Thr Pro Val Ala Ser Pro Glu Ala Ser Thr Leu Ser Thr
 1640 1645 1650
 Thr Pro Val Asp Ser Asn Ser Pro Val Ile Thr Ser Thr Glu Val

[0048]

Gly Arg Pro Pro Leu Thr Ser Ile Pro Val Ser Thr Thr Thr Val
 1925 1930 1935
 Ala Ser Ser Glu Ile Asn Thr Leu Ser Thr Thr Leu Ala Asp Thr
 1940 1945 1950
 Arg Thr Pro Val Thr Thr Tyr Ser Gln Ala Ser Ser Ser Pro Thr
 1955 1960 1965
 Thr Ala Asp Gly Thr Ser Met Pro Thr Pro Ala Tyr Ser Glu Gly
 1970 1975 1980
 Ser Thr Pro Leu Thr Ser Met Pro Leu Ser Thr Thr Leu Val Val
 1985 1990 1995
 Ser Ser Glu Ala Ser Thr Leu Ser Thr Thr Pro Val Asp Thr Ser
 2000 2005 2010
 Thr Pro Ala Thr Thr Ser Thr Glu Gly Ser Ser Ser Pro Thr Thr
 2015 2020 2025
 Ala Gly Gly Thr Ser Ile Gln Thr Ser Thr Pro Ser Glu Arg Thr
 2030 2035 2040
 Thr Pro Leu Ala Gly Met Pro Val Ser Thr Thr Leu Val Val Ser
 2045 2050 2055
 Ser Glu Gly Asn Thr Leu Ser Thr Thr Pro Val Asp Ser Lys Thr
 2060 2065 2070
 Gln Val Thr Asn Ser Thr Glu Ala Ser Ser Ser Ala Thr Ala Glu
 2075 2080 2085
 Gly Ser Ser Met Thr Ile Ser Ala Pro Ser Glu Gly Ser Pro Leu
 2090 2095 2100
 Leu Thr Ser Ile Pro Leu Ser Thr Thr Pro Val Ala Ser Pro Glu
 2105 2110 2115
 Ala Ser Thr Leu Ser Thr Thr Pro Val Asp Ser Asn Ser Pro Val
 2120 2125 2130
 Ile Thr Ser Thr Glu Val Ser Ser Ser Pro Ile Pro Thr Glu Gly
 2135 2140 2145
 Thr Ser Met Gln Thr Ser Thr Tyr Ser Asp Arg Arg Thr Pro Leu
 2150 2155 2160
 Thr Ser Met Pro Val Ser Thr Thr Val Val Ala Ser Ser Ala Ile
 2165 2170 2175

[0050]

Ser Thr Leu Ser Thr Thr Pro Val Asp Thr Ser Thr Pro Val Thr
 2180 2185 2190
 Asn Ser Thr Glu Ala Arg Ser Ser Pro Thr Thr Ser Glu Gly Thr
 2195 2200 2205
 Ser Met Pro Thr Ser Thr Pro Ser Glu Gly Ser Thr Pro Phe Thr
 2210 2215 2220
 Ser Met Pro Val Ser Thr Met Pro Val Val Thr Ser Glu Ala Ser
 2225 2230 2235
 Thr Leu Ser Ala Thr Pro Val Asp Thr Ser Thr Pro Val Thr Thr
 2240 2245 2250
 Ser Thr Glu Ala Thr Ser Ser Pro Thr Thr Ala Glu Gly Thr Ser
 2255 2260 2265
 Ile Pro Thr Ser Thr Leu Ser Glu Gly Thr Thr Pro Leu Thr Ser
 2270 2275 2280
 Ile Pro Val Ser His Thr Leu Val Ala Asn Ser Glu Val Ser Thr
 2285 2290 2295
 Leu Ser Thr Thr Pro Val Asp Ser Asn Thr Pro Phe Thr Thr Ser
 2300 2305 2310
 Thr Glu Ala Ser Ser Pro Pro Pro Thr Ala Glu Gly Thr Ser Met
 2315 2320 2325
 Pro Thr Ser Thr Ser Ser Glu Gly Asn Thr Pro Leu Thr Arg Met
 2330 2335 2340
 Pro Val Ser Thr Thr Met Val Ala Ser Phe Glu Thr Ser Thr Leu
 2345 2350 2355
 Ser Thr Thr Pro Ala Asp Thr Ser Thr Pro Val Thr Thr Tyr Ser
 2360 2365 2370
 Gln Ala Gly Ser Ser Pro Thr Thr Ala Asp Asp Thr Ser Met Pro
 2375 2380 2385
 Thr Ser Thr Tyr Ser Glu Gly Ser Thr Pro Leu Thr Ser Val Pro
 2390 2395 2400
 Val Ser Thr Met Pro Val Val Ser Ser Glu Ala Ser Thr His Ser
 2405 2410 2415
 Thr Thr Pro Val Asp Thr Ser Thr Pro Val Thr Thr Ser Thr Glu
 2420 2425 2430

[0051]

Ala Ser Ser Ser Pro Thr Thr Ala Glu Gly Thr Ser Ile Pro Thr
2435 2440 2445

Ser Pro Pro Ser Glu Gly Thr Thr Pro Leu Ala Ser Met Pro Val
2450 2455 2460

Ser Thr Thr Pro Val Val Ser Ser Glu Ala Gly Thr Leu Ser Thr
2465 2470 2475

Thr Pro Val Asp Thr Ser Thr Pro Met Thr Thr Ser Thr Glu Ala
2480 2485 2490

Ser Ser Ser Pro Thr Thr Ala Glu Asp Ile Val Val Pro Ile Ser
2495 2500 2505

Thr Ala Ser Glu Gly Ser Thr Leu Leu Thr Ser Ile Pro Val Ser
2510 2515 2520

Thr Thr Pro Val Ala Ser Pro Glu Ala Ser Thr Leu Ser Thr Thr
2525 2530 2535

Pro Val Asp Ser Asn Ser Pro Val Val Thr Ser Thr Glu Ile Ser
2540 2545 2550

Ser Ser Ala Thr Ser Ala Glu Gly Thr Ser Met Pro Thr Ser Thr
2555 2560 2565

Tyr Ser Glu Gly Ser Thr Pro Leu Arg Ser Met Pro Val Ser Thr
2570 2575 2580

Lys Pro Leu Ala Ser Ser Glu Ala Ser Thr Leu Ser Thr Thr Pro
2585 2590 2595

Val Asp Thr Ser Ile Pro Val Thr Thr Ser Thr Glu Thr Ser Ser
2600 2605 2610

Ser Pro Thr Thr Ala Lys Asp Thr Ser Met Pro Ile Ser Thr Pro
2615 2620 2625

Ser Glu Val Ser Thr Ser Leu Thr Ser Ile Leu Val Ser Thr Met
2630 2635 2640

Pro Val Ala Ser Ser Glu Ala Ser Thr Leu Ser Thr Thr Pro Val
2645 2650 2655

Asp Thr Arg Thr Leu Val Thr Thr Ser Thr Gly Thr Ser Ser Ser
2660 2665 2670

Pro Thr Thr Ala Glu Gly Ser Ser Met Pro Thr Ser Thr Pro Gly
2675 2680 2685

Glu Arg Ser Thr Pro Leu Thr Asn Ile Leu Val Ser Thr Thr Leu

[0052]

Val Thr Thr Ser Ala Glu Ala Ser Ser Ser Pro Thr Thr Ala Glu
 2960 2965 2970
 Gly Thr Ser Met Pro Ile Ser Thr Pro Gly Glu Arg Arg Thr Pro
 2975 2980 2985
 Leu Thr Ser Met Ser Val Ser Thr Met Pro Val Ala Ser Ser Glu
 2990 2995 3000
 Ala Ser Thr Leu Ser Arg Thr Pro Ala Asp Thr Ser Thr Pro Val
 3005 3010 3015
 Thr Thr Ser Thr Glu Ala Ser Ser Ser Pro Thr Thr Ala Glu Gly
 3020 3025 3030
 Thr Gly Ile Pro Ile Ser Thr Pro Ser Glu Gly Ser Thr Pro Leu
 3035 3040 3045
 Thr Ser Ile Pro Val Ser Thr Thr Pro Val Ala Ile Pro Glu Ala
 3050 3055 3060
 Ser Thr Leu Ser Thr Thr Pro Val Asp Ser Asn Ser Pro Val Val
 3065 3070 3075
 Thr Ser Thr Glu Val Ser Ser Ser Pro Thr Pro Ala Glu Gly Thr
 3080 3085 3090
 Ser Met Pro Ile Ser Thr Tyr Ser Glu Gly Ser Thr Pro Leu Thr
 3095 3100 3105
 Gly Val Pro Val Ser Thr Thr Pro Val Thr Ser Ser Ala Ile Ser
 3110 3115 3120
 Thr Leu Ser Thr Thr Pro Val Asp Thr Ser Thr Pro Val Thr Thr
 3125 3130 3135
 Ser Thr Glu Ala His Ser Ser Pro Thr Thr Ser Glu Gly Thr Ser
 3140 3145 3150
 Met Pro Thr Ser Thr Pro Ser Glu Gly Ser Thr Pro Leu Thr Tyr
 3155 3160 3165
 Met Pro Val Ser Thr Met Leu Val Val Ser Ser Glu Asp Ser Thr
 3170 3175 3180
 Leu Ser Ala Thr Pro Val Asp Thr Ser Thr Pro Val Thr Thr Ser
 3185 3190 3195
 Thr Glu Ala Thr ser ser Thr Thr Ala Glu Gly Thr ser Ile Pro
 3200 3205 3210

[0054]

Thr Ser Thr Pro Ser Glu Gly Met Thr Pro Leu Thr Ser Val Pro
 3215 3220 3225
 Val Ser Asn Thr Pro Val Ala Ser Ser Glu Ala Ser Ile Leu Ser
 3230 3235 3240
 Thr Thr Pro Val Asp Ser Asn Thr Pro Leu Thr Thr Ser Thr Glu
 3245 3250 3255
 Ala Ser Ser Ser Pro Pro Thr Ala Glu Gly Thr Ser Met Pro Thr
 3260 3265 3270
 Ser Thr Pro Ser Glu Gly Ser Thr Pro Leu Thr Ser Met Pro Val
 3275 3280 3285
 Ser Thr Thr Thr Val Ala Ser Ser Glu Thr Ser Thr Leu Ser Thr
 3290 3295 3300
 Thr Pro Ala Asp Thr Ser Thr Pro Val Thr Thr Tyr Ser Gln Ala
 3305 3310 3315
 Ser Ser Ser Pro Pro Ile Ala Asp Gly Thr Ser Met Pro Thr Ser
 3320 3325 3330
 Thr Tyr Ser Glu Gly Ser Thr Pro Leu Thr Asn Met Ser Phe Ser
 3335 3340 3345
 Thr Thr Pro Val Val Ser Ser Glu Ala Ser Thr Leu Ser Thr Thr
 3350 3355 3360
 Pro Val Asp Thr Ser Thr Pro Val Thr Thr Ser Thr Glu Ala Ser
 3365 3370 3375
 Leu Ser Pro Thr Thr Ala Glu Gly Thr Ser Ile Pro Thr Ser Ser
 3380 3385 3390
 Pro Ser Glu Gly Thr Thr Pro Leu Ala Ser Met Pro Val Ser Thr
 3395 3400 3405
 Thr Pro Val Val Ser Ser Glu Val Asn Thr Leu Ser Thr Thr Pro
 3410 3415 3420
 Val Asp Ser Asn Thr Leu Val Thr Thr Ser Thr Glu Ala Ser Ser
 3425 3430 3435
 Ser Pro Thr Ile Ala Glu Gly Thr Ser Leu Pro Thr Ser Thr Thr
 3440 3445 3450
 ser Glu Gly Ser Thr Pro Leu Ser Ile Met Pro Leu Ser Thr Thr
 3455 3460 3465

[0055]

Pro Val Ala Ser Ser Glu Ala Ser Thr Leu Ser Thr Thr Pro Val
 3470 3475 3480

Asp Thr Ser Thr Pro Val Thr Thr Ser Ser Pro Thr Asn Ser Ser
 3485 3490 3495

Pro Thr Thr Ala Glu Val Thr Ser Met Pro Thr Ser Thr Ala Gly
 3500 3505 3510

Glu Gly Ser Thr Pro Leu Thr Asn Met Pro Val Ser Thr Thr Pro
 3515 3520 3525

Val Ala Ser Ser Glu Ala Ser Thr Leu Ser Thr Thr Pro Val Asp
 3530 3535 3540

Ser Asn Thr Phe Val Thr Ser Ser Ser Gln Ala Ser Ser Ser Pro
 3545 3550 3555

Ala Thr Leu Gln Val Thr Thr Met Arg Met Ser Thr Pro Ser Glu
 3560 3565 3570

Gly Ser Ser Ser Leu Thr Thr Met Leu Leu Ser Ser Thr Tyr Val
 3575 3580 3585

Thr Ser Ser Glu Ala Ser Thr Pro Ser Thr Pro Ser Val Asp Arg
 3590 3595 3600

Ser Thr Pro Val Thr Thr Ser Thr Gln Ser Asn Ser Thr Pro Thr
 3605 3610 3615

Pro Pro Glu Val Ile Thr Leu Pro Met Ser Thr Pro Ser Glu Val
 3620 3625 3630

Ser Thr Pro Leu Thr Ile Met Pro Val Ser Thr Thr Ser Val Thr
 3635 3640 3645

Ile Ser Glu Ala Gly Thr Ala Ser Thr Leu Pro Val Asp Thr Ser
 3650 3655 3660

Thr Pro Val Ile Thr Ser Thr Gln Val Ser Ser Ser Pro Val Thr
 3665 3670 3675

Pro Glu Gly Thr Thr Met Pro Ile Trp Thr Pro Ser Glu Gly Ser
 3680 3685 3690

Thr Pro Leu Thr Thr Met Pro Val Ser Thr Thr Arg Val Thr Ser
 3695 3700 3705

Ser Glu Gly Ser Thr Leu Ser Thr Pro Ser Val Val Thr Ser Thr
 3710 3715 3720

Pro Val Thr Thr Ser Thr Glu Ala Ile Ser Ser Ser Ala Thr Leu

[0056]

3725	3730	3735												
Asp 3740	Ser 3740	Thr 3740	Thr 3740	Met 3740	Ser 3740	Val 3745	Ser 3745	Met 3745	Pro 3745	Met 3745	Glu 3750	Ile 3750	Ser 3750	Thr 3750
Leu 3755	Gly 3755	Thr 3755	Thr 3755	Ile 3755	Leu 3755	Val 3760	Ser 3760	Thr 3760	Thr 3760	Pro 3760	Val 3765	Thr 3765	Arg 3765	Phe 3765
Pro 3770	Glu 3770	Ser 3770	Ser 3770	Thr 3770	Pro 3770	Ser 3775	Ile 3775	Pro 3775	Ser 3775	Val 3775	Tyr 3780	Thr 3780	Ser 3780	Met 3780
Ser 3785	Met 3785	Thr 3785	Thr 3785	Ala 3785	Ser 3785	Glu 3790	Gly 3790	Ser 3790	Ser 3790	Ser 3790	Pro 3795	Thr 3795	Thr 3795	Leu 3795
Glu 3800	Gly 3800	Thr 3800	Thr 3800	Thr 3800	Met 3800	Pro 3805	Met 3805	Ser 3805	Thr 3805	Thr 3805	Ser 3810	Glu 3810	Arg 3810	Ser 3810
Thr 3815	Leu 3815	Leu 3815	Thr 3815	Thr 3815	Val 3815	Leu 3820	Ile 3820	Ser 3820	Pro 3820	Ile 3820	Ser 3825	Val 3825	Met 3825	Ser 3825
Pro 3830	Ser 3830	Glu 3830	Ala 3830	Ser 3830	Thr 3830	Leu 3835	Ser 3835	Thr 3835	Pro 3835	Pro 3835	Gly 3840	Asp 3840	Thr 3840	Ser 3840
Thr 3845	Pro 3845	Leu 3845	Leu 3845	Thr 3845	Ser 3845	Thr 3850	Lys 3850	Ala 3850	Gly 3850	Ser 3850	Phe 3855	Ser 3855	Ile 3855	Pro 3855
Ala 3860	Glu 3860	Val 3860	Thr 3860	Thr 3860	Ile 3860	Arg 3865	Ile 3865	Ser 3865	Ile 3865	Thr 3865	Ser 3870	Glu 3870	Arg 3870	Ser 3870
Thr 3875	Pro 3875	Leu 3875	Thr 3875	Thr 3875	Leu 3875	Leu 3880	Val 3880	Ser 3880	Thr 3880	Thr 3880	Leu 3885	Pro 3885	Thr 3885	Ser 3885
Phe 3890	Pro 3890	Gly 3890	Ala 3890	Ser 3890	Ile 3890	Ala 3895	Ser 3895	Thr 3895	Pro 3895	Pro 3895	Leu 3900	Asp 3900	Thr 3900	Ser 3900
Thr 3905	Thr 3905	Phe 3905	Thr 3905	Pro 3905	Ser 3905	Thr 3910	Asp 3910	Thr 3910	Ala 3910	Ser 3910	Thr 3915	Pro 3915	Thr 3915	Ile 3915
Pro 3920	Val 3920	Ala 3920	Thr 3920	Thr 3920	Ile 3920	Ser 3925	Val 3925	Ser 3925	Val 3925	Ile 3925	Thr 3930	Glu 3930	Gly 3930	Ser 3930
Thr 3935	Pro 3935	Gly 3935	Thr 3935	Thr 3935	Ile 3935	Phe 3940	Ile 3940	Pro 3940	Ser 3940	Thr 3940	Pro 3945	Val 3945	Thr 3945	Ser 3945
Ser 3950	Thr 3950	Ala 3950	Asp 3950	Val 3950	Phe 3950	Pro 3955	Ala 3955	Thr 3955	Thr 3955	Gly 3955	Ala 3960	Val 3960	Ser 3960	Thr 3960
Pro 3965	Val 3965	Ile 3965	Thr 3965	Ser 3965	Thr 3965	Glu 3970	Leu 3970	Asn 3970	Thr 3970	Pro 3970	Ser 3975	Thr 3975	Ser 3975	Ser 3975
Ser 3980	Ser 3980	Thr 3980	Thr 3980	Thr 3980	Ser 3980	Phe 3985	Ser 3985	Thr 3985	Thr 3985	Lys 3985	Glu 3990	Phe 3990	Thr 3990	Thr 3990

[0057]

Pro Ala Met Thr Thr Ala Ala Pro Leu Thr Tyr Val Thr Met Ser
 3995 4000 4005
 Thr Ala Pro Ser Thr Pro Arg Thr Thr Ser Arg Gly Cys Thr Thr
 4010 4015 4020
 Ser Ala Ser Thr Leu Ser Ala Thr Ser Thr Pro His Thr Ser Thr
 4025 4030 4035
 Ser Val Thr Thr Arg Pro Val Thr Pro Ser Ser Glu Ser Ser Arg
 4040 4045 4050
 Pro Ser Thr Ile Thr Ser His Thr Ile Pro Pro Thr Phe Pro Pro
 4055 4060 4065
 Ala His Ser Ser Thr Pro Pro Thr Thr Ser Ala Ser Ser Thr Thr
 4070 4075 4080
 Val Asn Pro Glu Ala Val Thr Thr Met Thr Thr Arg Thr Lys Pro
 4085 4090 4095
 Ser Thr Arg Thr Thr Ser Phe Pro Thr Val Thr Thr Thr Ala Val
 4100 4105 4110
 Pro Thr Asn Thr Thr Ile Lys Ser Asn Pro Thr Ser Thr Pro Thr
 4115 4120 4125
 Val Pro Arg Thr Thr Thr Cys Phe Gly Asp Gly Cys Gln Asn Thr
 4130 4135 4140
 Ala Ser Arg Cys Lys Asn Gly Gly Thr Trp Asp Gly Leu Lys Cys
 4145 4150 4155
 Gln Cys Pro Asn Leu Tyr Tyr Gly Glu Leu Cys Glu Glu Val Val
 4160 4165 4170
 Ser Ser Ile Asp Ile Gly Pro Pro Glu Thr Ile Ser Ala Gln Met
 4175 4180 4185
 Glu Leu Thr Val Thr Val Thr Ser Val Lys Phe Thr Glu Glu Leu
 4190 4195 4200
 Lys Asn His Ser Ser Gln Glu Phe Gln Glu Phe Lys Gln Thr Phe
 4205 4210 4215
 Thr Glu Gln Met Asn Ile Val Tyr Ser Gly Ile Pro Glu Tyr Val
 4220 4225 4230
 Gly Val Asn Ile Thr Lys Leu Arg Leu Gly Ser Val Val Val Glu
 4235 4240 4245

[0058]

His Asp Val Leu Leu Arg Thr Lys Tyr Thr Pro Glu Tyr Lys Thr
 4250 4255 4260
 Val Leu Asp Asn Ala Thr Glu Val Val Lys Glu Lys Ile Thr Lys
 4265 4270 4275
 Val Thr Thr Gln Gln Ile Met Ile Asn Asp Ile Cys Ser Asp Met
 4280 4285 4290
 Met Cys Phe Asn Thr Thr Gly Thr Gln Val Gln Asn Ile Thr Val
 4295 4300 4305
 Thr Gln Tyr Asp Pro Glu Glu Asp Cys Arg Lys Met Ala Lys Glu
 4310 4315 4320
 Tyr Gly Asp Tyr Phe Val Val Glu Tyr Arg Asp Gln Lys Pro Tyr
 4325 4330 4335
 Cys Ile Ser Pro Cys Glu Pro Gly Phe Ser Val Ser Lys Asn Cys
 4340 4345 4350
 Asn Leu Gly Lys Cys Gln Met Ser Leu Ser Gly Pro Gln Cys Leu
 4355 4360 4365
 Cys Val Thr Thr Glu Thr His Trp Tyr Ser Gly Glu Thr Cys Asn
 4370 4375 4380
 Gln Gly Thr Gln Lys Ser Leu Val Tyr Gly Leu Val Gly Ala Gly
 4385 4390 4395
 Val Val Leu Met Leu Ile Ile Leu Val Ala Leu Leu Met Leu Val
 4400 4410
 Phe Arg Ser Lys Arg Glu Val Lys Arg Gln Lys Tyr Arg Leu Ser
 4415 4420 4425
 Gln Leu Tyr Lys Trp Gln Glu Glu Asp Ser Gly Pro Ala Pro Gly
 4430 4435 4440
 Thr Phe Gln Asn Ile Gly Phe Asp Ile Cys Gln Asp Asp Asp Ser
 4445 4450 4455
 Ile His Leu Glu Ser Ile Tyr Ser Asn Phe Gln Pro Ser Leu Arg
 4460 4465 4470
 His Ile Asp Pro Glu Thr Lys Ile Arg Ile Gln Arg Pro Gln Val
 4475 4480 4485
 Met Thr Thr Ser Phe
 4490

[0059]

<210> 6
 <211> 167
 <212> PRT
 <213> 人

 <400> 6

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

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

 Gly Gly Gly Lys Arg Phe Ser His Ser Gly Asn Gln Leu Asp Gly Pro
 35 40 45

 Ile Thr Ala Leu Arg Val Arg Val Asn Thr Tyr Tyr Ile Val Gly Leu
 50 55 60

 Gln Val Arg Tyr Gly Lys Val Trp Ser Asp Tyr Val Gly Gly Arg Asn
 65 70 75 80

 Gly Asp Leu Glu Glu Ile Phe Leu His Pro Gly Glu Ser Val Ile Gln
 85 90 95

 Val Ser Gly Lys Tyr Lys Trp Tyr Leu Lys Lys Leu Val Phe Val Thr
 100 105 110

 Asp Lys Gly Arg Tyr Leu Ser Phe Gly Lys Asp Ser Gly Thr Ser Phe
 115 120 125

 Asn Ala Val Pro Leu His Pro Asn Thr Val Leu Arg Phe Ile Ser Gly
 130 135 140

 Arg Ser Gly Ser Leu Ile Asp Ala Ile Gly Leu His Trp Asp Val Tyr
 145 150 155 160

 Pro Thr Ser Cys Ser Arg Cys
 165

名称	符号	MWG 寡聚物	Unigene mRNA 参照序列	蛋白质 参照序列	等级强度 (肿瘤)	等级强度 (非恶性肿瘤)
黏蛋白5, 亚型A和C	MUC5AC	30K#B: 7561	NM_017511	NP_059981	404	328
黏蛋白17, 细胞表面结合的	MUC17	30K#C: 0346	NM_001040105	NP_001035 194	2043	1485
酶原粒蛋白16	ZG16	30K#C: 0346	NM_152338	NP_689551	3606	3249
癌胚抗原相关 细胞粘附分子 5	CEACAM5	30K#C: 0346	NM_004363	NP_004354	4873	7668

图 1

胃癌患者		对照	
阶段T1	2	结直肠癌阶段T2	1
阶段T2	15	结直肠癌阶段T3	9
阶段T3	13	结直肠癌阶段T4	1
阶段T4	1	结直肠癌外科手术后	7
胃肠基质肿瘤	2	非恶性疾病	10
		健康的	13
总计	33	总计	41

图 2

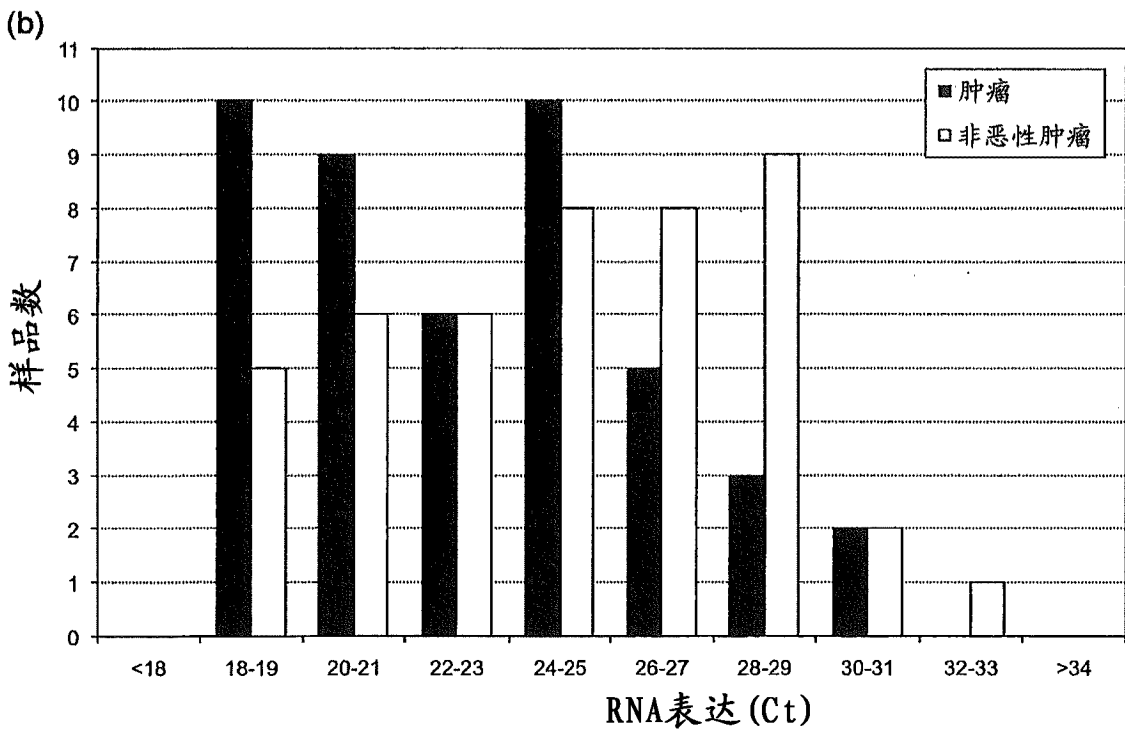
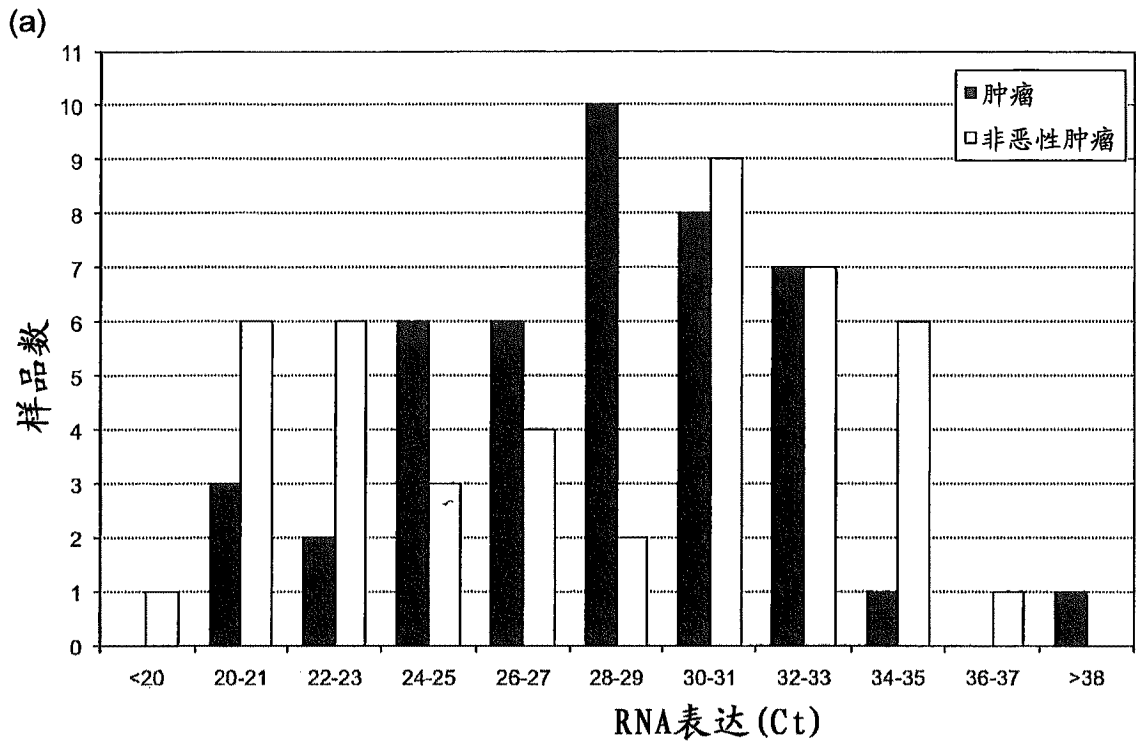


图 3

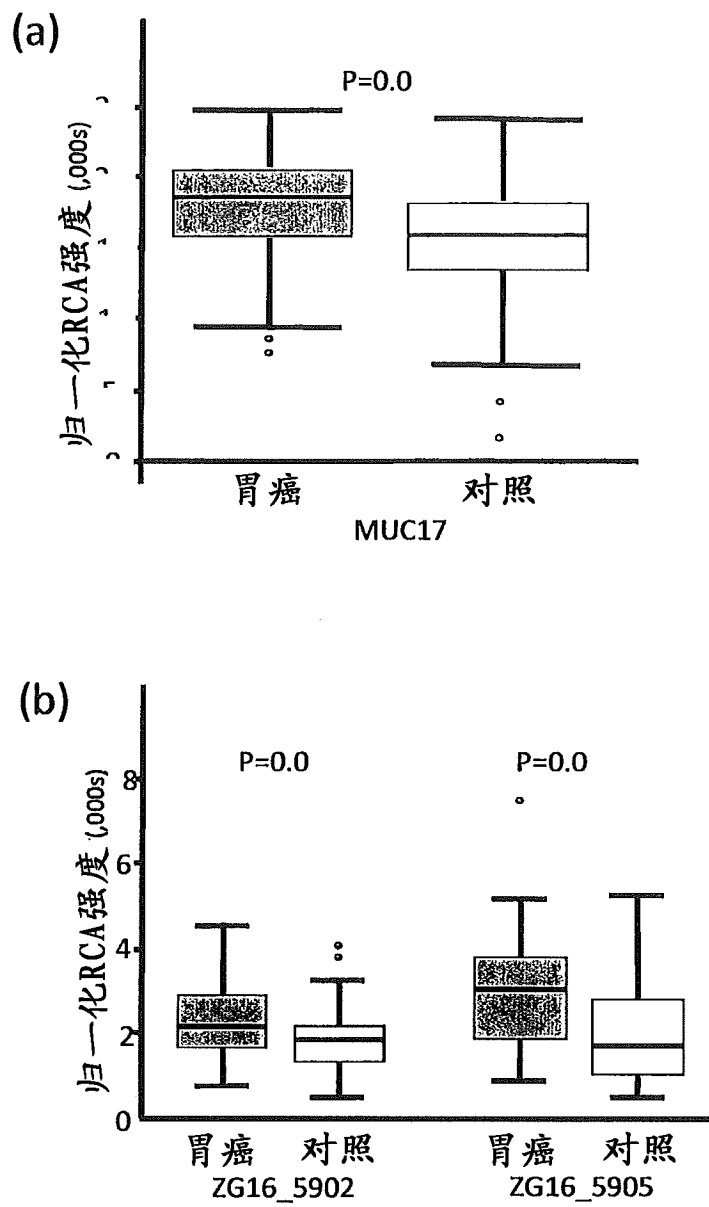


图 4

专利名称(译)	胃癌的检测标志物		
公开(公告)号	CN102459646A	公开(公告)日	2012-05-16
申请号	CN201080031553.3	申请日	2010-05-14
[标]申请(专利权)人(译)	环太平洋生物技术有限公司		
申请(专利权)人(译)	环太平洋生物技术有限公司		
当前申请(专利权)人(译)	环太平洋生物技术有限公司		
[标]发明人	PJ吉尔福德		
发明人	P· J· 吉尔福德		
IPC分类号	C12Q1/68 G01N33/53 G01N33/50 G01N33/68		
CPC分类号	G01N33/57446 C12Q2600/158 C12Q1/6886 C12Q2600/16 G01N2333/47 G01N2333/4725 G16B25/00		
代理人(译)	柴云峰		
优先权	577012 2009-05-15 NZ		
其他公开文献	CN102459646B		
外部链接	Espacenet SIPO		

摘要(译)

肿瘤的早期检测是肿瘤(包括胃肿瘤)患者存活的主要决定因素。GTM基因家族的成员可以在胃肿瘤组织中差异地表达，并且因此可以用作检测胃癌和其他类型癌症的标志物。本发明提供了用于检测肿瘤(包括胃肿瘤)的新GTM，并且特别是人酶原蛋白16(ZG16)。GTM可以单独使用或与其他已知GTM一起使用，来提供用于肿瘤(包括胃肿瘤)检测的新标记。

名称	符号	MWG 寡聚物	Unigene mRNA 参照序列	蛋白质 参照序列	等级强度 (肿瘤)	等级强度 (非恶性肿瘤)
黏蛋白5, 亚型A和C	MUC5AC	30K#B: 7561	NM_017511	NP_059981	404	328
黏蛋白17, 细胞表面结合的	MUC17	30K#C: 0346	NM_001040105	NP_001035194	2043	1485
酶原蛋白16	ZG16	30K#C: 0346	NM_152338	NP_689551	3606	3249
癌胚抗原相关细胞粘附分子 5	CEACAM5	30K#C: 0346	NM_004363	NP_004354	4873	7668