



(12)发明专利

(10)授权公告号 CN 102459646 B

(45)授权公告日 2016.08.10

(21)申请号 201080031553.3

(22)申请日 2010.05.14

(30)优先权数据

577012 2009.05.15 NZ

(85)PCT国际申请进入国家阶段日

2012.01.13

(86)PCT国际申请的申请数据

PCT/NZ2010/000089 2010.05.14

(87)PCT国际申请的公布数据

W02010/131984 EN 2010.11.18

(73)专利权人 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 黄革生

(51)Int.Cl.

C12Q 1/68(2006.01)

G01N 33/53(2006.01)

G01N 33/50(2006.01)

G01N 33/68(2006.01)

(56)对比文件

WO 98/44160 A1,1998.10.08,

CN 1867679 A,2006.11.22,

João Pinto-de-Sousa, 等.Mucins MUC1, MUC2, MUC5AC and MUC6 expression in the evaluation of differentiation and clinico-biological behaviour of gastric carcinoma.《Virchows Arch》.2002,第440卷第304-310页.

审查员 幸颖

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. 一种组合物在制备用于实施一种用于检测胃癌的方法的试剂盒中的用途,包括:
 - (i)提供血液或血清样品;和
 - (ii)检测所述样品中胃肿瘤标志物人酶原粒蛋白16(“ZG16”)的水平,其中所述组合物包括用于在血液或血清中检测的抗ZG16的抗体。
2. 权利要求1的用途,其中所述血液或血清样品中ZG16的过表达表明癌症。
3. 权利要求1或权利要求2的用途,包括检测一种或多种其他胃肿瘤标志物家族成员的水平。
4. 权利要求1至2任一项的用途,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”)的其他胃肿瘤标志物家族成员的水平。
5. 权利要求3的用途,其中所述一种或多种其他胃肿瘤标志物家族成员选自由MUC5AC; MUC17; 羧肽酶N多肽2的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”); β -诱导的转化生长因子(“TGFB1”); 含有EGF的fibulin-样胞外基质蛋白2(“EFEMP2”); 基膜聚糖(Iumican)(“LUM”); stannin(“SNN”); 分泌型磷蛋白1(“SPPI”); 硫酸软骨素蛋白聚糖2(“CSPG2”); N-乙酰基鞘氨醇酰胺水解酶(“ASAH1”); 丝氨酸蛋白酶11(“PRSS11”); 分泌型卷曲相关蛋白2(“SFRP2”); XIIB组磷脂酶A2(“PLA2G12B”); 胞外基质蛋白spondin 2(“SPON2”); 嗅介蛋白(oIfactomedin)1(“OLFM1”); 含有血小板反应蛋白1型重复序列的多肽(“TSRC1”); 血小板反应蛋白2(“THBS2”); adIican; 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CIade H成员1(“SERPINH1”); 丝氨酸或半胱氨酸蛋白酶抑制剂CIade 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5的用途,其中所测试的其它胃肿瘤标志物家族成员包括MUC5AC、MUC17、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的用途,其中通过检测胃肿瘤标志物蛋白的水平进行所述检测步骤。
8. 权利要求1的用途,其中通过检测胃肿瘤标志物肽的水平进行所述检测步骤。
9. 权利要求7或8的用途,其中使用针对所述胃肿瘤标志物的抗体进行所述检测步骤。
10. 权利要求7至8任一项的用途,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。
11. 权利要求9的用途,其中所述抗体是单克隆抗体。
12. 权利要求9的用途,其中所述抗体是多克隆抗血清。
13. 一种抗ZG16的抗体在生产用于检测胃肿瘤标志物的装置中的用途,所述装置包括: 基质,其上具有胃肿瘤标志物捕获试剂; 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胃肿瘤标志物,其中所述胃肿瘤标志物包括人酶原粒蛋白16(“ZG16”),其中所述胃肿瘤标志物捕获试

剂是用于在血液或血清中检测的抗ZG16的抗体。

14. 一种抗ZG16的抗体在生产用于检测胃癌的试剂盒中的用途,所述试剂盒包含:
基质,其上具有胃肿瘤标志物捕获试剂;
用于显示所述胃肿瘤标志物捕获试剂和胃肿瘤标志物复合物的试剂;和
使用说明书,其中所述胃肿瘤标志物包括人酶原粒蛋白16(“ZG16”),其中所述胃肿瘤标志物捕获试剂是用于在血液或血清中检测的抗ZG16的抗体。

15. 一种组合物在制备用于实施一种用于筛选胃癌的方法的试剂盒中的用途,包括步骤:

提供来自测试受试者的测试样品,所述测试样品是血液或血清样品;
测量所述测试样品中胃肿瘤标志物的存在;和
将所述测试样品中存在的胃肿瘤标志物的量与获自不患胃癌受试者的对照样品的值进行比较,其中所述胃肿瘤标志物包括人酶原粒蛋白16(“ZG16”),
其中所述组合物包括用于在血液或血清中检测的抗ZG16的抗体。
16. 权利要求15的用途,其中所述测量步骤使用ELISA测定法。

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。在胃癌的情况下,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(“TGFBI”)、含有EGF的fibulin-样胞外基质蛋白2(“EFEMP2”)、基膜聚糖(Iumican)(“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(oIfactomedin)1(“OLFM1”)、含有血小板反应蛋白重复1(“TSRC1”)、血小板反应蛋白2(“THBS2”)、adIican、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(“TGF β 1”)、丝氨酸或半胱氨酸蛋白酶抑制剂CIade H成员1(“SERPINH1”)、丝氨酸或半胱氨酸蛋白酶抑制剂CIade 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(如以下实施例中所述的那些)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,如肌昔)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%FicoII/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(50μg/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 & C.C.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-乙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”)、adIican、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CIade H成员1(“SERPINH1”)、丝氨酸或半胱氨酸蛋白酶抑制剂CIade 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HeId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HeIen 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°C,例如,约50-70°C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.NatI.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、IIIunima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(TeI-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(NoIte,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 PIex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] HaII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] Meile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,(J 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15uI。然后将Cy3和Cy5标记的cDNA与Ambion ULTRAhyb缓冲液混合,在100℃下变性2分钟,并在杂交室中,在42℃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 GTTTTTAGC),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_m1(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CACCAATGCCATTCAGGCCAGGT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℃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℃下在摇床上再生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Sonic 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0.45μm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Ni Sepharose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

℃。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℃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℃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phi 29在30℃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, NieIsen H, Rrunak S, von Heiine 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

```

atgtcgcctgc agttcaatgg gtttgtgctg tgtgatcaca cggtcctgcc cacagagctg      60
aatgggctat tggggtttcc ttggaggctg ctgagggaca gggcactctt ccccgccgtc      120
cacacaatga gtgttgccg gaggaagctg gccctgctct gggccctggc tctcgctctg      180
gcctgcaccc ggcacacagg ccatgccag gatggctcct ccgaatccag ctacaagcac      240
cacctgccc tctctcctat cgcccggggg cccagcgggg tcccgtccg tggggcgact      300
gtcttcccat ctctgaggac catccctgtg gtacgagcct ccaaccggc gcacaacggg      360
cgggtgtgca gcacctggg cagcttccac tacaagacct tcgacggcga cgtcttccgc      420
ttccccggcc tctgcaacta cgtgttctcc gagcactgct gtgccgcta cgaggatatt      480
aacatccagc tacgccgag ccaggagtca gcgccccca cgctgagcag ggtcctcatg      540
aagggtgatg gcgtggatc ccagctgacc aagggtccg tcctggtaa cggccaccgc      600
gtcctgctgc cttcagcca gtctggggtc ctcatcagc agagcagcag ctacaccaag      660
gtggaggcca ggctggcct tgcctcatg tggaaccacg atgacagcct gctgctggag      720
ctggacacca aatacgcaa caagacctgt gggctctgtg gggacttaa cgggatgcc      780
gtggtcagcg agctcctctc ccacaacacc aagctgacac ccatggaatt cgggaacctg      840
cagaagatgg acgacccac ggagcagtgt caggacctg tcctgaacc cccgaggaac      900
tgctccactg gctttggcat ctgtgaggag ctctgcacg gccagctgtt ctctggctgc      960
gtggccctgg tggacgtcgg cagctacctg gaggttgca ggcaagacct ctgcttctgt     1020
gaagacaccg acctgctcag ctgctctgc cacacccttg ccgagtactc ccggcagtgc     1080
accatgcag gggggttgcc ccaggactgg cggggccctg acttctgccc ccagaagtgc     1140
cccaacaaca tgcagtacca cgagtgccg tccccctgct cagacacctg ctccaaccag     1200
gagcactccc gggcctgtga ggaccactgt gtggccggct gcttctgccc tgaggggacg     1260
gtgcttgacg acatcgcca gaccggctgt gtccctgtgt caaagtgtgc ctgctctac     1320
aacggggctg cctatgccc aggggccacc tactccacag actgcaccaa ctgcacctgc     1380
tccggaggcc ggtggagctg ccaggagggt ccatgccgg gtacctgctc tgtgcttga     1440

```

	ggtgccact tctcaacggt tgacgggaag caatacacgg tgcacggcga ctgcagctat	1500
	gtgctgacca agcctgtga cagcagtgcc ttcactgtac tggctgagct gcgcaggtgc	1560
	gggctgacgg acagcgagac ctgcctgaag agcgtgacac tgagcctgga tggggcgag	1620
	acgggtggtg tgatcaaggc cagtgggaa gtgttctga accagatcta caccagctg	1680
	cccatctctg cagccaacgt caccatcttc agaccctcaa ctttctcat catcgcccag	1740
	accagcctgg gcctgcagct gaacctgcag ctggtgcca ccatgcagct gttcatgcag	1800
	ctggcgccca agctccgtgg gcagacctgc ggtctctgtg ggaacttcaa cagcatccag	1860
	gccgatgact tccggaccct cagtgggtg gtggaggcca ccgctgcggc cttcttcaac	1920
	accttcaaga ccagggccgc ctgccccaac atcaggaaca gcttcgagga ccctgtctt	1980
	ctgagcgtgg agaatgcccc caggatgggt gcccggagcg cgtgccggac gctggtgtg	2040
	gctctggcgg ccatgctcgt ggtgcccgag gcagagacc agggccctgt ggagccgagc	2100
	tgggagaatg cagggcacac catggatggc ggtgccccga cgtcctcgcc caccggcgc	2160
	gtgagctttg ttccaccgt cactgtcttc ccagcctga gccgtaagca gatgtgccc	2220
	ctgccagccg ggaaggggt gtttgccagt ccaaagggtg ggggccaga tctaggggtg	2280
	cagctgccac cagccctgaa cccggcgcac aatggcggg tgtgcagcac ctgggtgac	2340
	ttccactaca agacctcga cggcgacgtc ttccgttcc ctggccttg caactacgtg	2400
	ttctctgagc actgccgcgc cgctacgag gacttcaacg tccagctacg ccgaggccta	2460
[0002]	gtgggtcca ggctgtggt caccgtgtt gtcatcaagg ccaggggtt ggtgtggag	2520
	gcgtccaacg gctccgtcct catcaatggg cagcgggagg agctgcctta cagccgact	2580
	ggcctcctgg tggagcagag cggggactac atcaaggta gcatccggct ggtgtgaca	2640
	ttctgtgga acggagagga cagtgcctg ctggagctgg atcccaaata cgccaaccag	2700
	acctgtggcc tgtgtggga cttcaacggc ctcccggcct tcaacgagtt ctatgcccac	2760
	aacgccaggc tgacccgct ccagtttggg aacctgcaga agttggatgg gccacggag	2820
	cagtgcccg acccgctgcc cttgccggcc ggcaactgca cggacgagga gggcatctgc	2880
	caccgaccc tgctggggcc ggcctttgcg gactgaccg cactggtgga cagcactgcg	2940
	tacctggccg cctgcgcca ggacctgtc cgctgcccc cctgcccgtg tgccacctt	3000
	gtggaatact cacgccagtg cgccacgcg gggggccagc cgcggaactg gaggtgccct	3060
	gagctctgcc cccggacctg cccctcaac atgcagcacc aggagtgtg ctaccctgc	3120
	acggacacct gtccaaccc ccagcgcg cagctctgcg aggaccactg tgtggacggc	3180
	tgcttctgcc cccaggcag gtcttgacg gtgctggatg acatcacgca ctctggctgc	3240
	ctgcccctg ggcagtgcc ctgcaccac ggcgccgca cctacagccc gggcacctcc	3300
	ttcaacacca cctgcagctc ctgcacctgc tccgggggc tatggcagtg ccaggacctg	3360
	ccgtgccctg gcacctgctc tgtgcagggc ggggccaca 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	cgtgggtgag	acaggcctcg	ggctgcagct	gctggtgcag	3720
ctggtgccac	tcatgcagg	gtttgtcagg	ctggacccc	cccaccagg	ccagatgtgc	3780
ggcctgtgtg	ggaacttcaa	ccagaaccag	gctgacgact	tcacggccct	cagcgggggtg	3840
gtggaggcca	cgggcgcagc	cttcgccaac	acctggaagg	cccaggctgc	ctgtgccaat	3900
gccaggaaca	gctttgagga	cccctgctcc	ctcagtgtgg	agaatgagaa	ctacgcccgg	3960
cactggtgct	cgcgcctgac	cgatcccaac	agtgccttct	cgcgctgcca	ctccatcatc	4020
aaccccaagc	ccttccactc	gaactgcatg	tttgacacct	gcaactgtga	gctggagcgag	4080
gactgcctgt	gcgccgcgct	gtcctcctac	gtgcacgcct	gtgccgcca	ggcgtacag	4140
ctcagcgact	ggagggacgg	cgtctgcacc	aagtacatgc	agaactgcc	caagtcccag	4200
cgctacgcct	acgtgggtga	tgcttgccag	cccacttgcc	gctggcctgag	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	ccccatgggtg	4500
tacctggact	gcagcaacag	ctcggcgggc	accctgggg	ccgagtgcct	ccggagctgc	4560
cacacgctgg	acgtgggctg	tttcagcaca	cactgcgtgt	ccggctgtgt	ctgtcccccg	4620
gggctgggtg	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
atcttgcccc	aggactactg	tggggacaac	accaccacg	ggaccttccg	catcgtcacc	4920
gagaacatcc	cctgtgggac	caccggcacc	acctgctcca	aggccatcaa	gctcttcgtg	4980
gaggtgagaa	cggccccagc	tgtgagcacc	cccgaccctg	cagccaacga	gccggcccc	5040
agggaaagctt	cggtcggctt	ccggcagcgt	ctgcctcccc	tgcagagcta	cgagctgatc	5100
ctccaagagg	ggacctttaa	ggcgggtggc	agagggccgg	gtggggacc	accctacaag	5160
atacgctaca	tggggatctt	cctggtcatc	gagaccacg	ggatggccgt	gtcctgggac	5220
cggaagacca	gcgtgttcat	ccgactgcac	caggactaca	agggcagggt	ctgcggcctg	5280
tgcggaact	tcgacgaaa	tgccatcaat	gactttgcca	cgcgtagccg	gtccgtgggtg	5340
ggggacgcac	tggagtgttg	gaacagctgg	aagctctccc	cctcctgcc	ggacgccctg	5400
gcaccaaggg	accctgcac	ggccaacccc	ttccgcaagt	cctgggcca	gaagcagtgc	5460
agcatcctcc	acggccccac	cttcgccgcc	tgccgctccc	aggttgactc	caccaagtac	5520
tacgaggcct	gcgtgaacga	cgcgtgtgcc	tgcgactcgg	gtggcgactg	cgagtgtttc	5580

	tgcacggctg tggctgccta cgcccaggcc tgccacgacg cgggcctgtg tgtgtcctgg	5640
	cggactccgg acacctgccc cttgttctgt gacttctaca acccacatgg gggctgtgag	5700
	tggcactacc agccctgcfg ggcaccctgc ctaaaaacct gccggaacce cagtgggcac	5760
	tgctgtgtgg acctgcctgg cctggaaggc tgctaccga agtgcccacc cagccagccc	5820
	ttcttcaatg aggaccagat gaagtgcgtg gcccagtggtg gctgctacga caaggacgga	5880
	aactactatg acgtcgggtg aagggtcccc acagcggaga actgccagag ctgtaactgc	5940
	acaccagtg gcatccagtg cgctcacagc cttgaggcct gcacctgcac ctatgaggac	6000
	aggacctaca gctaccagga cgctcatctac aacaccaccg atgggcttgg cgctgtctg	6060
	atcgccatct gcggaagcaa cggcaccatc atcaggaagg ctgtggcatg tcctggaact	6120
	ccagccaaa cgccattcac cttcaccacc gcctgggtcc cccactccac gacaagcccg	6180
	gccctcccgg tctccaccgt gtgtgtccgc gaggtctgcc gctgggtccag ctggtacaat	6240
	gggcaccgcc cagagcccgg cctgggaggc ggagactttg agacgtttga aaacctgagg	6300
	cagagagggg accaggtatg ccctgtgctg gctgacatcg agtgccgggc ggcgcagctt	6360
	cccacatgc cgctggagga gctgggcccag caggtggact gtgaccgcat gcgggggctg	6420
	atgtgcgcca acagccaaca gagtcccccg ctctgtcacg actacgagct gcgggttctc	6480
	tgctgcgaat acgtgccctg tggcccctcc ccggcccag gcaccagccc tcagccctcc	6540
	ctcagtgcc a gcacggagcc tgctgtgcct accccaacct agaccacagc aaccgaaaag	6600
[0004]	accaccctat gggtgacccc gagcatccgg tcgacggcgg ccctcacctc gcagactggg	6660
	tccagctcag gccccgtgac ggtcaccccc tcggcccag gtaccaccac ctgccagccc	6720
	cggtgtcagt ggacagagt gtttgatgag gactaccca agtctgaaca acttggaggg	6780
	gacgttgagt cctacgataa gatcagggcc gctggagggc acttatgcca gcagcctaag	6840
	gacatagagt gccaggccga gagcttcccc aactggaccc tggcacaggt ggggcagaag	6900
	gtgcaactgt acgtccactt cggcctggtg tgcaggaact gggagcagga gggcgtcttc	6960
	aagatgtgct acaactacag gatccgggtc ctctgtgca gtgacgacca ctgcagggga	7020
	cgtgccaaa ccccgccacc gaccacagag ctggagacgg ccaccaccac caccaccag	7080
	gccctgttct caacgccga gcctacgagt agcccggggc tgaccagggc tccccggcc	7140
	agcaccacag cagtccccac cctctcagaa ggactgacat cccccagata cacaagcacc	7200
	cttggtagag ccaccacggg aggccccacg acgcctgcag gctccacaga acccactgtc	7260
	ccaggggtgg ccacatccac cttccaaca cgctcagccc ttccagggac gacggggagc	7320
	ttgggacat ggcgcccctc acagccacc acgctggccc caacaacaat ggcaacctcc	7380
	agagctcgcc cgacaggcac agccagcacc gcttcaaag agccgctgac cacgagcctg	7440
	gcgccaacac tcacgagcga gctgtccacc tctcaggccg agaccagcac gccaggaca	7500
	gagacgaaa tgagcccctt gactaacacc accaccagcc agggcacgac ccgctgtcaa	7560
	ccgaagtgtg agtggacaga gtggtttgac gtggacttcc caacctcagg ggttgaggc	7620
	ggggacatgg aaacttttga aaacatcagg gctgctgggg gcaagatgtg ctgggcacca	7680

[0005]

aagagcatag	agtgccgggc	ggagaactac	cccgaggtaa	gcatcgacca	ggtcgggag	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	acgactgctg	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	gacccgcta	tcacagacca	ccacaccac	ggccacctg	8220
tccacagcca	caccctcctc	cactccagag	actgccaca	cctccacagt	gcttaccgcc	8280
acggccacca	caactggggc	caccggctct	gtggccacc	cctcctccac	cccaggaaca	8340
gctcacacta	ccaaagtgcc	aactaccaca	accacgggct	tcacagccac	ccccctctc	8400
agcccagga	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
tcctgcccc	ccagcagtc	ccacacggtg	cgcacagcct	ggacttcggc	cacctcgggc	8700
atcttgggca	ccaccacat	cacagagcct	tccacggtga	cttccacac	cctagcagca	8760
accaccggta	ccaccagca	ctcgactcca	gccctttcca	gccctcacc	tagcagcaga	8820
accaccgagt	caccctctc	tccagggagc	accaccccg	gccacaccac	ggccacctcc	8880
aggaccacag	ccacggccac	accagcaag	accgcacct	cgacctgct	gccagcagc	8940
cccacatcgg	ccccataac	cacggtggtg	accatgggct	gtgagccca	gtgtgcctgg	9000
tcagagtggc	tggactacag	ctacccatg	ccggggccct	ctggcgggga	ctttgacacc	9060
tactccaaca	tccgtgcggc	cggaggggcc	gtctgtgagc	agcccctggg	cctcagtgct	9120
cgtgcccagg	cccagcctgg	tgtccccctg	cgggagttgg	gccaggtcgt	ggaatgcagc	9180
ctggactttg	gcctggctg	caggaaccgt	gagcaggtgg	ggaagttaa	gatgtgcttc	9240
aactatgaaa	tccgtgtggt	ctgctgcaac	tacggccact	gccccagcac	cccggccacc	9300
agctctacgg	ccatgccctc	ctccactccg	gggacgacct	ggatcctcac	agagctgacc	9360
acaacagcca	ctacgactga	gtccactgga	tccacggcca	ccccgtcctc	caccacagg	9420
accacctgga	tcctcacaga	gccgagcact	acagccaccg	tgacggtgcc	caccggatcc	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

	accatgtcca	cagccacacc	ctcctccact	ccagagactg	tccacacctc	cacagtgctt	9780
	accaccacgg	ccaccacaac	cggggccacc	ggctctgtgg	ccaccccctc	ctccacccca	9840
	ggaacagctc	acactaccaa	agtgtgact	accacaacca	cgggcttcac	agccaccccc	9900
	tcctccagcc	cagggacggc	acgcacgctt	ccagtgtgga	tcagcacaac	caccacacc	9960
	acaaccagag	gttccacggt	gaccccctcc	tccatcccgg	ggaccacca	cacccccaca	10020
	gtgctgacca	ccaccaccac	aactgtggcc	actggttcta	tggcaacacc	ctcctctagc	10080
	acacagacca	gtgggaccac	ccacacacc	ccagtgccga	acaccacggc	caccacacac	10140
	gggcgatccc	tgtccccag	cagtccccac	acggtgcgca	cagcctggac	ttcggccacc	10200
	tcaggcacct	tgggaccac	ccacatcaca	gagccttcca	cggggacttc	ccacacccca	10260
	gcagcaacca	ccggtaccac	ccagcactcg	actccagccc	tgtccagccc	tcaccctagc	10320
	agcaggacca	ccgagtcacc	cccttctcca	gggacgacca	ccccgggcca	caccagggcc	10380
	acctccagga	ccacggccac	ggccacacc	agcaagacc	gcacctcgac	cctgctgccc	10440
	agcagcccca	catcgcccc	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	aggtggggaa	gttcaagatg	10740
[0006]	tgcttcaact	atgaaatccg	tgtgttctgc	tgcaactacg	gccactgccc	cagcaccccg	10800
	gccaccagct	ctacggccac	gcccctctcc	actccaggga	cgacctggat	cctcacagag	10860
	cagaccacag	cagccactac	gaccgcaacc	actggatcca	cggccatccc	gtcctccacc	10920
	ccgggaacag	ctccccctcc	caaagtgtg	accagcacgg	ccaccacacc	cacagccacc	10980
	agttccaaag	ccacttcctc	ctccagtcca	aggactgcaa	ccacccttcc	agtgtgaca	11040
	agcacagcca	ccaaatccac	agctaccagc	tttacacca	tcccctcctt	cacccttggg	11100
	accaccggga	ccctcccaga	acagaccacc	acacccatgg	ccacctgtc	cacaatccac	11160
	ccctcttcca	ctccggagac	caccacacc	tccacagtgc	tgaccacgaa	ggccaccacg	11220
	acaagggcca	ccagttccat	gtccacccc	tcctccactc	cggggacgac	ctggatcctc	11280
	acagagctga	ccacagcagc	cactacaact	gcagccactg	gccccacggc	caccccgctc	11340
	tccaccccag	ggaccacctg	gatcctcaca	gagcccagca	ctacagccac	cgtgacggtg	11400
	cccaccggat	ccacggccac	cgctcctcc	accgggcaa	ctgctggcac	cctcaaagtg	11460
	ctgaccagca	cggccaccac	accacagtc	atcagctcca	gagccactcc	ctcctccagt	11520
	ccagggactg	caaccgcct	tccagcactg	agaagcacag	ccaccacacc	cacagctacc	11580
	agcgttacag	ccatcccctc	ttctcctg	ggcaccgct	ggaccgcct	atcacagacc	11640
	accacacca	cggccaccat	gtccacagcc	acaccctct	ctactccaga	gactgtccac	11700
	acctccacag	tgttaccac	cacgaccacc	acaaccaggg	ccaccggctc	tgtggccacc	11760
	ccctcctcca	ccccaggaac	agctcacact	accaaagtgc	cgactaccac	aaccacgggc	11820

[0007]

ttcacagcca	ccccctcctc	cagcccaggg	acggcactca	cgctccagt	gtggatcagc	11880
acaaccacca	caccacacaac	cagaggctcc	acggtgacct	cctcctccat	cccggggacc	11940
accacaccg	ccacagtgtc	gaccaccacc	accacaactg	tggccactgg	ttctatggca	12000
acaccctcct	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	tcacccctt	ctccaggac	gaccaccccg	12300
ggccacacca	ggggcacctc	caggaccaca	gccacagcca	caccagcaa	gacccgcacc	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	gccagcctg	gtgtccccct	gcgggagttg	12600
ggccaggtcg	tggaatgcag	cctggacttt	ggcctggtct	gcaggaaccg	tgagcaggtg	12660
gggaagtcca	agatgtgctt	caactatgaa	atccgtgtgt	tctgctgcaa	ctacggccac	12720
tgcccagca	ccccggccac	cagctctacg	gccacgcct	cctcaactcc	ggggacgacc	12780
tggatcctca	caaagctgac	cacaacagcc	actacgactg	agtccactgg	atccacggcc	12840
accccgctct	ccaccccagg	gaccacctgg	atcctcacag	agccgagcac	tacagccacc	12900
gtgacggtgc	ccaccggatc	cacggccacc	gcctcctcca	cccaggcaac	tgctggcacc	12960
ccacatgtga	gcaccacggc	cacgacacc	acagtcacca	gctccaaagc	cactcccttc	13020
tccagtccag	ggactgcaac	cgcccttcca	gactgagaa	gcacagccac	cacaccaca	13080
gctaccagct	ttacagccat	ccccctctcc	tcctgggca	ccacctggac	ccgcctatca	13140
cagaccacca	caccacggc	caccatgtcc	acagccacac	cctcctccac	tccagagact	13200
gccacacct	ccacagtgtc	taccaccag	gccaccaca	ccagggccac	cggtctgtg	13260
gccaccccct	cttccacccc	aggaacagct	cacactacca	aagtgccgac	taccacaacc	13320
acgggcttca	cagtcacccc	ctcctccagc	ccagggacgg	cacgcacgcc	tccagtgtgg	13380
atcagcacia	ccaccacacc	cacaaccagt	ggctccacgg	tgacccctc	ctccgtcccg	13440
gggaccaccc	acacccccac	agtgtctgacc	accaccacca	caactgtggc	cactggttct	13500
atggcaacac	cctcctctag	cacacagacc	agtgggacca	cccacacacc	cccagtgccg	13560
aacaccacgg	ccaccacaca	cgggcgatcc	ctgtccccca	gcagtccccca	cacggtgcgc	13620
acagcctgga	cttcggccac	ctcaggcacc	ttgggacca	cccacatcac	agagccttcc	13680
acggggactt	cccacacccc	agcagcaacc	accggtacca	cccagcactc	gactccagcc	13740
ctgtccagcc	ctcaccttag	cagcaggacc	accgagtcac	ccccttccc	agggacgacc	13800
acccggggcc	acaccacggc	cacctccagg	accacggcca	cggccacacc	cagcaagacc	13860

	cgcacctcga	ccctgctgcc	cagcagcccc	acatcggccc	ccataaccac	ggtggtgacc	13920
	acgggctgtg	agccccagtg	tgcctggtca	gagtggctgg	actacagcta	ccccatgccg	13980
	gggccctctg	gcggggactt	tgacacctac	tccaacatcc	gtgcggccgg	aggggccgtc	14040
	tgtgagcagc	ccctgggcct	cgagtgccgt	gccaggccc	agcctggtgt	ccccctgggg	14100
	gagttgggcc	aggtcgtgga	atgcagcctg	gactttggcc	tggtctgcag	gaaccgtgag	14160
	caggtgggga	agttcaagat	gtgcttcaac	tatgaaatcc	gtgtgttctg	ctgcaactac	14220
	ggccactgcc	ccagcacccc	ggccaccagc	tctacggcca	tgccctcctc	cactccgggg	14280
	acgacctgga	tcctcacaga	gctgaccaca	acagccacta	cgactgcatc	cactggatcc	14340
	acggccaccc	cgtcctccac	cccgggaaca	gctccccctc	caaagtgtct	gaccagcccc	14400
	gccaccacac	ccacagccac	cagttccaaa	gccacttctc	cctccagtcc	aaggactgca	14460
	accacccttc	cagtgtgac	aagcacagcc	accaaatacca	cagctaccag	cgttacacc	14520
	atcccctcct	ccacccttgg	gaccaccggg	accctcccag	aacagaccac	cacaccctgt	14580
	gccaccatgt	ccacaatcca	cccctcctcc	actccggaga	ccaccacac	ctccacagtg	14640
	ctgaccacga	aggccaccac	gacaagggcc	accagttcca	cgccacccc	ctcctccact	14700
	ccggggacga	cctggatcct	cacagagctg	accacagcag	ccactacaac	tgacagccact	14760
	ggccccacgg	ccaccccgtc	ctccaccca	gggaccacct	ggatcctcac	agagctgacc	14820
	acaacagcca	ctacgactgc	gtccactgga	tccacggcca	ccccgtcctc	caccccaggg	14880
[0008]	accacctgga	tcctcacaga	gccgagcact	acagccaccg	tgacgggtgcc	caccggatcc	14940
	acggccaccg	cctcctccac	ccaggcaact	gctggcacc	cacatgtgag	caccacggcc	15000
	acgacacca	cagtcaccag	ctccaaagcc	actcccctc	ccagtccagg	gactgcaact	15060
	gcccttccag	actgagaag	cacagccacc	acaccacag	ctaccagctt	tacagccatc	15120
	ccctcctcct	ccctgggcac	cacctggacc	cgctatcac	agaccaccac	accacgggcc	15180
	accatgtcca	cagccacacc	ctcctccact	ccagagactg	tccacacctc	cacagtgtct	15240
	accgccacgg	ccaccacaac	cggggccacc	ggctctgtgg	ccaccccctc	ctccaccca	15300
	ggaacagctc	acactacca	agtgccgact	accacaacca	cgggcttcc	agccacccc	15360
	tcctccagcc	cagggacggc	actcacgct	ccaaccacca	cacccatgtc	caccatgtcc	15420
	acaatccaca	cctcctctac	tccagagacc	accacacct	ccacagtgtc	gaccaccaca	15480
	gccaccatga	caagggccac	caattccag	gccacaccct	cctccactct	ggggacgacc	15540
	cggatcctca	ctgagctgac	cacaacagcc	actacaactg	cagccactgg	atccacggcc	15600
	accctgtcct	ccaccccagg	gaccacctgg	atcctcacag	agccgagcac	tatagccacc	15660
	gtgatggtgc	ccaccggttc	cacggccacc	gcctcctcca	ctctgggaac	agctcacacc	15720
	cccaaagtgg	tgaccaccat	ggccactatg	cccacagcca	ctgcctccac	ggttcccagc	15780
	tcgtccaccg	tggggaccac	ccgcaccct	gcagtgtctc	ccagcagcct	gccaaccttc	15840
	agcgtgtcca	ctgtgtcctc	ctcagtctc	accaccctga	gaccactgg	cttcccagc	15900
	tcccacttct	ctactccctg	cttctgcagg	gcatttggac	agtttttctc	gcccggggaa	15960

[0009]

gtcatctaca	ataagaccga	ccgagccggc	tgccatttct	acgcagtgtg	caatcagcac	16020
tgtgacattg	accgcttcca	ggg'gcctgt	cccacctccc	caccgccagt	gtcctccgcc	16080
ccgctgtcct	cgccctcccc	tgcccctggc	tgtgacaatg	ccatccctct	ccggcaggtg	16140
aatgagacct	ggaccctgga	gaactgcacg	gtggccaggt	gcg'tgggtga	caaccgtgtc	16200
gtcctgctgg	acccaaagcc	tgtggccaac	gtcacctgcg	tgaacaagca	cctgccccatc	16260
aaagtgtcgg	acccgagcca	gccctgtgac	ttccactatg	agtgcgagtg	tgagtgcgtc	16320
ggtggcccg	ggattacccc	gggggcaggc	atctgcagca	tgtggggcgg	ctcccactat	16380
tccacctttg	acggcacctc	ttacaccttc	cggggcaact	gcacctatgt	cctcatgaga	16440
gagatccatg	cacgctttgg	gaatctcagc	ctctacctgg	acaaccacta	ctgcacggcc	16500
tctgccactg	ccgctgccgc	tgccgcccgc	tgcccccgcg	ccctcagcat	ccactacaag	16560
tccatggata	tcgtcctcac	tgtcaccatg	gtgcatggga	aggaggagg	cctgatcctg	16620
tttgacaaa	ttccggtgag	cagcggtttc	agcaagaacg	gcg'tgcttgt	gtctgtgctg	16680
gggaccacca	ccatgcgtgt	ggacattcct	gccctggggc	tgagcgtcac	cttcaatggc	16740
caagtcttcc	aggcccggct	gccctacagc	ctcttccaca	acaacaccga	gggccagtg	16800
ggcacctgca	ccaacaacca	gagggacgac	tgtctccagc	gggacggaac	cactgccgcc	16860
agttgcaagg	acatggccaa	gacgtggctg	gtccccgaca	gcagaaagga	tggctgctgg	16920
gccccgactg	gcacaccccc	cactgccagc	cccgcagccc	cgg'tgtctag	cacaccacc	16980
cccaccccat	gcccaccaca	gccgctctgt	gatctgatgc	tgagccaggt	ctttgctgag	17040
tgccacaacc	ttgtgcccc	gggcccattc	ttcaacgcct	gcatcagcga	ccactgcagg	17100
ggccgccttg	aggtgccctg	ccagagcctg	gaggcttacg	cagagctctg	ccgcgcccgg	17160
ggagtgtgca	gtgactggcg	aggtgcaacc	ggtggcctgt	gcgacctcac	ctgcccacc	17220
accaaagtgt	acaagccatg	cggccccata	cagcctgcca	cctgcaactc	taggaaccag	17280
agcccacagc	tggaggggat	ggcggagggc	tgcttctgcc	ctgaggacca	gatcctcttc	17340
aacgcacaca	tgggcatctg	cgtgcaggcc	tgcccctgcg	tgggaccga	tgggtttcct	17400
aaatttcccg	gggagcgggtg	ggtcagcaac	tgccagtcct	gcg'tgtgtga	cgagggttca	17460
gtgtcgggtg	agtgcaagcc	cctgccctgt	gacgcccagg	gtcagcccc	gccgtgcaac	17520
cg'tcccggct	tcg'taacctg	gaccaggccc	cgggccgaga	accctgctg	ccccgagacg	17580
gtgtgcgtgt	gcaacacaac	cacctgcccc	cagagcctgc	ctgtgtgcc	gccagggcag	17640
gagtccatct	gcaccagga	ggagggcgac	tgctgtccca	ccttccgctg	cagacctcag	17700
ctgtgttctg	acaatggcac	cttctacggg	gttgg'tgcaa	ccttcccagg	cgcccttccc	17760
tgccacatgt	gtacctgcct	ctctggggac	accaggacc	caacggtgca	atgtcaggag	17820
gatgcctgca	acaatactac	ctgtccccag	ggctttgagt	acaagagagt	ggccgggcag	17880
tgctgtgggg	agtgcgtcca	gaccgcctgc	ctcacgcccg	atggccagcc	agtccagctg	17940
aatgaaacct	gggtcaacag	ccatgtggac	aactgcaccg	tgtacctctg	tgaggctgag	18000

ggtggagtcc atttgctgac cccacagcct gcatcctgcc cagatgtgtc cagctgcagg	18060
gggagcctca ggaaaaccgg ctgctgctac tcctgtgagg aggactcctg tcaagtccgc	18120
atcaacacga ccatacctgtg gcaccagggc tgcgagaccg aggtcaacat caccttctgc	18180
gagggctcct gccccggagc gtccaagtac tcagcagagg cccaggccat gcagcaccag	18240
tgcacctgct gccaggagag gcgggtccac gaggagacgg tgcccttgca 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	
[0010] tttcgccagc tcctctgggg gtgacaggca agtgagacgt gctcagagct ccgatgccaa	60
ggccagggac catggcgctg tgtctgctga ccttggtcct ctgctcttg cccccacaag	120
ctgctgcaga acaggacctc agtgtgaaca gggctgtgtg ggatggagga ggggtcatct	180
cccaagggga cgtcttgaac cgtcagtgcc agcagctgtc tcagcacggtt aggacaggtt	240
ctgcggaaca caccgccaca ggtacaacat ctacaaatgt cgtggagcca agaattgatt	300
tgagttgcag caccaaccct gagatgacct cgattgagtc cagtgtgact tcagacactc	360
ctggtgtctc cagtaccagg atgacaccaa cagaatccag aacaacttca gaatctacca	420
gtgacagcac cacacttttc cccagttcta ctgaagacac ttcatctcct acaactcctg	480
aaggcaccga cgtgccatg 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
ccagcacat gaaggtggcc agttcagagg ctatcacctt ttgacaact cctggtgaaa	780
tcagcacacc tgtgaccatt tctgctcaag ccagttcatc tcctacaact gctgaaggtc	840
ccagcctgtc aaactcagct cctagtggag gaagcactcc attaacaaga atgcctctca	900
gcgtgatgct ggtggtcagt tctgaggcta gcacccttc aacaactcct gctgccacca	960
acattcctgt gatcacttct actgaagcca gttcatctcc tacaacggct gaaggcacca	1020
gcataccaac ctcaacttat actgaaggaa gcactccatt aacaagtacg cctgccagca	1080

	ccatgccggt	tgccacttct	gaaatgagca	cactttcaat	aactcctggt	gacaccagca	1140
	cacttgtagc	cacttctact	gaaccagtt	cacttcctac	aactgctgaa	gctaccagca	1200
	tgtaacctc	aactcctagt	gaaggaagca	ctccattaac	aaatatgcct	gtcagcacca	1260
	tattggtggc	cagttctgag	gctagcacca	cttcaacaat	tctgttgtag	tccaaaactt	1320
	ttgtgaccac	tgctagtga	gccagctcat	ctcccacaac	tgctgaagat	accagcattg	1380
	caacctcaac	tcctagtga	ggaagcactc	cattaacaag	tatgcctgtc	agcaccactc	1440
	cagtgccag	ttctgaggct	agcaaccttt	caacaactcc	tgttgactcc	aaaactcagg	1500
	tgaccacttc	tactgaagcc	agttcatctc	ctccaactgc	tgaagttaac	agcatgccaa	1560
	cctcaactcc	tagtgaagga	agcactccat	taacaagtat	gtctgtcagc	accatgccgg	1620
	tgccagttc	tgaggctagc	accctttcaa	caactcctgt	tgacaccagc	acacctgtga	1680
	ccacttctag	tgaagccagt	tcattcttcta	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	tagcaccctt	tcaacaactc	ctggtgactc	caacactcct	gtgaccactt	2040
	caactgaagc	cacttcatct	tctacaactg	cggaaggtag	cagcatgcca	acctcaactt	2100
[0011]	atactgaagg	aagcactcca	ttaacaagta	tgctgtcaa	caccacactg	gtggccagtt	2160
	ctgaggctag	caccctttca	acaactcctg	ttgacaccag	cacacctgtg	accacttcaa	2220
	ctgaagccag	ttcctctcct	acaactgctg	atggtgccag	tatgccaaac	tcaactccta	2280
	gtgaaggaag	cactccatta	acaagtatgc	ctgtcagcaa	aacgctggtg	accagttctg	2340
	aggctagcac	cctttcaaca	actcctcttg	acacaagcac	acatatcacc	acttctactg	2400
	aagccagttg	ctctcctaca	accactgaag	gtaccagcat	gccaatctca	actcctagtg	2460
	aaggaagtcc	ttattaaca	agtatacctg	tcagcatcac	accggtgacc	agtcctgagg	2520
	ctagcaccct	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
	catctcctac	aactgctgaa	ggtaccagca	taccaacctc	gactcctagt	gaaggaacga	3000
	ctccattaac	aagcacacct	gtcagccaca	cgctggtggc	caattctgag	gctagcacc	3060
	tttcaacaac	tcctggtgac	tccaacactc	ctttgaccac	ttctactgaa	gccagttcac	3120
	ctcctccac	tgctgaagg	accagcatgc	caacctcaac	tcctagtga	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	actccgttag	3540
caagtatgcc	tgtcagcacc	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	catgccaaacc	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	cctgttgaca	4320
ccagcaccac	tgccgtagtc	agttctgagg	ctagcacctt	ttcagcaact	cctgttgaca	4380
tcagcatacc	aacctcaact	cctagtgaag	gaaagactcc	attaataaagt	atacctgtca	4440
gcaacacgcc	gggtggccaat	tctgaggcta	gcaccctttc	aacaactcct	gttgactcta	4500
acagtcctgt	ggtcacttct	acagcagtca	gttcatctcc	tacacctgct	gaaggtagca	4560
gcatagcaat	ctcaacgcct	agtgaaggaa	gactgcatt	aacaagtata	cctgtcagca	4620
ccacaacagt	ggccagttct	gaaatcaaca	gcctttcaac	aactcctgct	gtcaccagca	4680
cacctgtgac	cacttattct	caagccagtt	catctctac	aactgctgac	ggtaccagca	4740
tgcaaacctc	aacttatagt	gaaggaagca	ctccactaac	aagtttgctt	gtcagacca	4800
tgctggtggt	cagttctgag	gctaacacc	tttcaacaac	ccctattgac	tccaaaactc	4860
aggtgaccgc	ttctactgaa	gccagttcat	ctacaaccgc	tgaaggtagc	agcatgacaa	4920
tctcaactcc	tagtgaagga	agtcctctat	taacaagtat	acctgtcagc	accacgccgg	4980
tgccagctcc	tgaggctagc	accctttcaa	caactcctgt	tgactccaac	agtctgtgta	5040
tcacttctac	tgaagtcagt	tcattctccta	cacctgctga	aggtaccagc	atgccaacct	5100
caacttatac	tgaaggaaga	actcctttaa	caagtataac	tgtcagaaca	acaccggtgg	5160
ccagctctgc	aatcagcacc	ctttcaacaa	ctcccgttga	caacagcaca	cctgtgacca	5220

	cttctactga agcccgttca tctcctacaa cttctgaagg taccagcatg ccaaactcaa	5280
	ctcctagtga aggaaccact ccattaacaa gtatacctgt cagcaccacg ccggtactca	5340
	gttctgaggc tagcaccctt 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 ttcattctct acacctgctg aaggtagcag catagcaacc tcaacgccta	5640
	gtgaaggaag cactgcatta acaagtatac ctgtcagcac cacaacagtg gccagttctg	5700
	aaaccaacac cttttcaaca actcccgtg tcaccagcac acctgtgacc acttatgctc	5760
	aagtcagttc atctcctaca 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 gcaccacgct ggtggtcagt tctgaggcta	6060
	gcactctttc cacaactcct gttgacacca gcactcctgc caccacttct actgaaggca	6120
	gttcatctcc tacaactgca ggaggtacca gcatacaaac ctcaactcct agtgaacgga	6180
	ccactccatt agcaggtatg cctgtcagca ctacgcttgt ggtagcttct gagggtaaca	6240
[0013]	ccctttcaac aactcctggt gactccaaaa ctcaggtgac caattctact gaagccagtt	6300
	catctgcaac cgctgaaggc agcagcatga caatctcagc tcctagtga ggaagtcctc	6360
	tactaacaag tatacctctc agcaccacgc cggtaggcccag tcctgaggct agcacccttt	6420
	caacaactcc tgttgactcc aacagtcctg tgatcacttc tactgaagtc agttcatctc	6480
	ctatacctac tgaaggtagc agcatgcaaa cctcaactta tagtgacaga agaactcctt	6540
	taacaagtat gcctgtcagc accacagtg gggccagttc tgcaatcagc accctttcaa	6600
	caactcctgt tgacaccagc acacctgtga ccaattctac tgaagcccgt tcattctccta	6660
	caacttctga aggtaccagc atgccaacct caactcctag tgaaggaagc actccattca	6720
	caagtatgcc tgtcagcacc atgccggtg ttacttctga ggctagcacc ctttcagcaa	6780
	ctcctgttga caccagcaca cctgtgacca ctctactga agccacttca tctcctacaa	6840
	ctgctgaagg taccagcata ccaacttcaa ctcttagtga aggaacgact ccattaacaa	6900
	gtatacctgt cagccacacg ctggtaggcca attctgaggc tagcaccctt tcaacaactc	6960
	ctggtgactc caactcctt ttcactactt ctactgaagc cagttcacct cctcccactg	7020
	ctgaaggtag cagcatgcca acctcaactt ctagtgaagg aaacttcca ttaacacgta	7080
	tgctgtcag caccacaatg gtggccagtt ttgaaacaag cacactttct acaactcctg	7140
	ctgacaccag cacacctgtg actacttatt ctcaagccgg ttcattctct acaactgctg	7200
	acgatactag catgccaaacc 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	ctggcaccct	ttccacaact	cctggtgaca	7500
ccagcacacc	tatgaccact	tctactgaag	ccagttcatc	tcctacaact	gctgaagata	7560
tcgtcgtgcc	aatctcaact	gctagtgaag	gaagtactct	attaacaagt	atacctgtca	7620
gcaccacgcc	agtggccagt	cctgaggcta	gcaccctttc	aacaactcct	gttgactcca	7680
acagtcctgt	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	tgaacaagt	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	ctgttgatac	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	actcctagt	9180
aaggaagtac	tccattaaca	agtatacctg	tcagcaccac	gccagtggcc	attcctgagg	9240
ctagcaccct	ttcaacaact	cctgttgact	ccaacagtcc	tgtggtcact	tctactgaag	9300
tcagttcatc	tcctacacct	gctgaaggta	ccagcatgcc	aatctcaact	tatagtgaag	9360

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	tgccagttc	tgaaacgagc	accctttcaa	9960	
caactcctgc	tgacaccagc	acacctgtga	ccacttattc	tcaagccagt	tcactctctc	10020	
caattgctga	cggtactagc	atgccaacct	caacttatag	tgaaggaagc	actccactaa	10080	
caaatatgtc	tttcagcacc	acgccagtgg	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	caactctctg	gtgaccactt	ctactgaagc	cagttcatct	cctacaatcg	10380	
[0015]	ctgaaggtac	cagcttgcca	acctcaacta	ctagtgaagg	aagcactcca	ttatcaatta	10440
	tgctctctag	taccacgccg	gtggccagtt	ctgaggctag	caccctttca	acaactcctg	10500
	ttgacaccag	cacacctgtg	accacttctt	ctccaaccaa	ttcatctcct	acaactgctg	10560
	aagttaccag	catgccaaaca	tcaactgctg	gtgaaggaag	cactccatta	acaaatatgc	10620
	ctgtcagcac	cacaccgggtg	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	tgtagcactt	tctactcaga	gcaattctac	tcctacacct	cctgaagtta	10920
	tcaccctgcc	aatgtcaact	cctagtgaag	taagcactcc	attaaccatt	atgcctgtca	10980
	gcaccacatc	ggtgaccatt	tctgaggctg	gcacagcttc	aacacttctt	gttgacacca	11040
	gcacacctgt	gactacttct	acccaagtca	gttcatctcc	tgtagactcct	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	gtcagtacca	11340
	cacctgttac	gaggtttcct	gagagtagca	cccttccat	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	tttattcca	11880
gcactcctgt	caccagttct	actgctgatg	tctttcctgc	aacaactggt	gctgtatcta	11940
cccctgtgat	aacttccact	gaactaaaca	caccatcaac	ctccagtagt	agtaccacca	12000
catctttttc	aactactaag	gaatttacia	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
ctgtgaaccc	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	catcaciaag	ctacgtcttg	12780
gcagtggtgt	ggtggagcat	gacgtcctcc	taagaaccaa	gtacacacca	gaatacaaga	12840
cagtattgga	caatgccacc	gaagtagtga	aagagaaaat	cacaaaagtg	accacacagc	12900
aaataatgat	taatgatatt	tgctcagaca	tgatgtgttt	caacaccact	ggcacccaag	12960
tgcaaaacat	tacggtgacc	cagtacgacc	ctgaagagga	ctgccggaag	atggccaagg	13020
aatatggaga	ctacttcgta	gtggagtacc	gggaccagaa	gccatactgc	atcagcccct	13080
gtgagcctgg	cttcagtgct	tccaagaact	gtaacctcgg	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	gccaaagtga	tgattccatc	cacctggagt	13440
ccatctatag	taatttccag	ccctccttga	gacacataga	ccctgaaaca	aagatccgaa	13500

```

ttcagaggcc tcaggtaatg acgacatcat tttaaggcat ggagctgaga agtctgggag 13560
tgaggagatc ccagtccggc taagcttggg ggagcatttt cccattgaga gccttccatg 13620
ggaactcaat gttcccattg taagtacagg aaacaagccc tgtacttacc aaggagaaag 13680
aggagagaca gcagtgtctg gagattctca aatagaaacc cgtggacgct ccaatgggct 13740
tgtcatgata tcaggctagg ctttctctgct ctttttcaa agacgctcca gatttgaggg 13800
tactctgact gcaacatctt tcacccatt gatcgccagg attgatttgg ttgatctggc 13860
tgagcaggcg ggtgtccccg tcctccctca ctgcccata tgtgtccctc ctaaagctgc 13920
atgctcagtt gaagaggacg agaggacgac cttctctgat agaggaggac cacgcttcag 13980
tcaaaggcat acaagtatct atctggactt ccctgctagc acttcaaac aagctcagag 14040
atgttctcc cctcatctgc cgggttcag taccatggac agcgcctcg acccgctgtt 14100
tacaaccatg accccttggg cactggactg catgcacttt acatatcaca aaatgctctc 14160
ataagaatta ttgcatacca tcttcatgaa aaacacctgt atttaaata agagcattta 14220
ccttttggta tataagattg tgggtathtt ttaagttctt attgttatga gttctgattt 14280
tttccttagt aaatattata atatatattt gtagtaacta aaaataataa agcaatttta 14340
ttacaatttt aaaaaaaaaa 14360

```

[0017]

```

<210> 3
<211> 632
<212> DNA
<213> 人

```

```

<400> 3
cccagaatgt tgacagtcgc tctcctagcc cttctctgtg cctcagcctc tggcaatgcc 60
attcaggcca ggtcttctc ctatagtggg gagtatggaa gtgggtgtgg aaagcgattc 120
tctcattctg gcaaccagtt ggacggcccc atcaccgccc tccgggtccg agtcaacaca 180
tactacatcg taggtcttca ggtgcgctat ggcaagggtg ggagcgacta tgtgggtggt 240
cgcaacggag acctggagga gatctttctg caccctgggg aatcagtgat ccaggtttct 300
gggaagtaca agtggtagct gaagaagctg gtatttgtga cagacaaggg ccgctatctg 360
tcttttggga aagacagtgg cacaagtttc aatgccgtcc ccttgcaccc caacaccgtg 420
ctccgcttca tcagtggccg gtctggttct ctcatcgatg ccattggcct gcaactggat 480
gtttacccca ctagtgcag cagatgctga gcctcctctc cttggcaggg gcaactgtgat 540
gaggagtaag aactccctta tcactaacc ccatccaaat 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

```

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

[0018] 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

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
 [0019] Thr Cys Ser Gly Gly Arg Trp Ser Cys Gln Glu Val Pro Cys Pro Gly
 420 425 430
 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

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

[0020]

835	840	845																										
Leu	Glu	Leu	Asp	Pro	Lys	Tyr	Ala	Asn	Gln	Thr	Cys	Gly	Leu	Cys	Gly													
	850					855					860																	
Asp	Phe	Asn	Gly	Leu	Pro	Ala	Phe	Asn	Glu	Phe	Tyr	Ala	His	Asn	Ala													
865					870					875					880													
Arg	Leu	Thr	Pro	Leu	Gln	Phe	Gly	Asn	Leu	Gln	Lys	Leu	Asp	Gly	Pro													
				885					890					895														
Thr	Glu	Gln	Cys	Pro	Asp	Pro	Leu	Pro	Leu	Pro	Ala	Gly	Asn	Cys	Thr													
			900					905					910															
Asp	Glu	Glu	Gly	Ile	Cys	His	Arg	Thr	Leu	Leu	Gly	Pro	Ala	Phe	Ala													
		915					920					925																
Glu	Cys	His	Ala	Leu	Val	Asp	Ser	Thr	Ala	Tyr	Leu	Ala	Ala	Cys	Ala													
	930					935					940																	
Gln	Asp	Leu	Cys	Arg	Cys	Pro	Thr	Cys	Pro	Cys	Ala	Thr	Phe	Val	Glu													
945					950					955					960													
Tyr	Ser	Arg	Gln	Cys	Ala	His	Ala	Gly	Gly	Gln	Pro	Arg	Asn	Trp	Arg													
				965					970					975														
[0021]	Cys	Pro	Glu	Leu	Cys	Pro	Arg	Thr	Cys	Pro	Leu	Asn	Met	Gln	His	Gln												
			980						985					990														
Glu	Cys	Gly	Ser	Pro	Cys	Thr	Asp	Thr	Cys	Ser	Asn	Pro	Gln	Arg	Ala													
		995					1000					1005																
Gln	Leu	Cys	Glu	Asp	His	Cys	Val	Asp	Gly	Cys	Phe	Cys	Pro	Pro														
	1010					1015					1020																	
Gly	Arg	Ser	Cys	Thr	Val	Leu	Asp	Asp	Ile	Thr	His	Ser	Gly	Cys														
	1025					1030					1035																	
Leu	Pro	Leu	Gly	Gln	Cys	Pro	Cys	Thr	His	Gly	Gly	Arg	Thr	Tyr														
	1040					1045					1050																	
Ser	Pro	Gly	Thr	Ser	Phe	Asn	Thr	Thr	Cys	Ser	Ser	Cys	Thr	Cys														
	1055					1060					1065																	
Ser	Gly	Gly	Leu	Trp	Gln	Cys	Gln	Asp	Leu	Pro	Cys	Pro	Gly	Thr														
	1070					1075					1080																	
Cys	Ser	Val	Gln	Gly	Gly	Ala	His	Ile	Ser	Thr	Tyr	Asp	Glu	Lys														
	1085					1090					1095																	
Leu	Tyr	Asp	Leu	His	Gly	Asp	Cys	Ser	Tyr	Val	Leu	Ser	Lys	Lys														
	1100					1105					1110																	

	Cys 1115	Ala	Asp	Ser	Ser	Phe	Thr 1120	Val	Leu	Ala	Glu	Leu 1125	Arg	Lys	Cys
	Gly 1130	Leu	Thr	Asp	Asn	Glu	Asn 1135	Cys	Leu	Lys	Ala	Val 1140	Thr	Leu	Ser
	Leu 1145	Asp	Gly	Gly	Asp	Thr	Ala 1150	Ile	Arg	Val	Gln	Ala 1155	Asp	Gly	Gly
	Val 1160	Phe	Leu	Asn	Ser	Ile	Tyr 1165	Thr	Gln	Leu	Pro	Leu 1170	Ser	Ala	Ala
	Asn 1175	Ile	Thr	Leu	Phe	Thr	Pro 1180	Ser	Ser	Phe	Phe	Ile 1185	Val	Val	Gln
	Thr 1190	Gly	Leu	Gly	Leu	Gln	Leu 1195	Leu	Val	Gln	Leu	Val 1200	Pro	Leu	Met
	Gln 1205	Val	Phe	Val	Arg	Leu	Asp 1210	Pro	Ala	His	Gln	Gly 1215	Gln	Met	Cys
	Gly 1220	Leu	Cys	Gly	Asn	Phe	Asn 1225	Gln	Asn	Gln	Ala	Asp 1230	Asp	Phe	Thr
[0022]	Ala 1235	Leu	Ser	Gly	Val	Val	Glu 1240	Ala	Thr	Gly	Ala	Ala 1245	Phe	Ala	Asn
	Thr 1250	Trp	Lys	Ala	Gln	Ala	Ala 1255	Cys	Ala	Asn	Ala	Arg 1260	Asn	Ser	Phe
	Glu 1265	Asp	Pro	Cys	Ser	Leu	Ser 1270	Val	Glu	Asn	Glu	Asn 1275	Tyr	Ala	Arg
	His 1280	Trp	Cys	Ser	Arg	Leu	Thr 1285	Asp	Pro	Asn	Ser	Ala 1290	Phe	Ser	Arg
	Cys 1295	His	Ser	Ile	Ile	Asn	Pro 1300	Lys	Pro	Phe	His	Ser 1305	Asn	Cys	Met
	Phe 1310	Asp	Thr	Cys	Asn	Cys	Glu 1315	Arg	Ser	Glu	Asp	Cys 1320	Leu	Cys	Ala
	Ala 1325	Leu	Ser	Ser	Tyr	Val	His 1330	Ala	Cys	Ala	Ala	Lys 1335	Gly	Val	Gln
	Leu 1340	Ser	Asp	Trp	Arg	Asp	Gly 1345	Val	Cys	Thr	Lys	Tyr 1350	Met	Gln	Asn
	Cys 1355	Pro	Lys	Ser	Gln	Arg	Tyr 1360	Ala	Tyr	Val	Val	Asp 1365	Ala	Cys	Gln

	Pro Thr 1370	Cys Arg Gly Leu	Ser 1375	Glu Ala Asp Val	Thr 1380	Cys Ser Val
	Ser Phe 1385	Val Pro Val Asp	Gly 1390	Cys Thr Cys Pro	Ala 1395	Gly Thr Phe
	Leu Asn 1400	Asp Ala Gly Ala	Cys 1405	Val Pro Ala Gln	Glu 1410	Cys Pro Cys
	Tyr Ala 1415	His Gly Thr Val	Leu 1420	Ala Pro Gly Glu	Val 1425	Val His Asp
	Glu Gly 1430	Ala Val Cys Ser	Cys 1435	Thr Gly Gly Lys	Leu 1440	Ser Cys Leu
	Gly Ala 1445	Ser Leu Gln Lys	Ser 1450	Thr Gly Cys Ala	Ala 1455	Pro Met Val
	Tyr Leu 1460	Asp Cys Ser Asn	Ser 1465	Ser Ala Gly Thr	Pro 1470	Gly Ala Glu
	Cys Leu 1475	Arg Ser Cys His	Thr 1480	Leu Asp Val Gly	Cys 1485	Phe Ser Thr
[0023]	His Cys 1490	Val Ser Gly Cys	Val 1495	Cys Pro Pro Gly	Leu 1500	Val Ser Asp
	Gly Ser 1505	Gly Gly Cys Ile	Ala 1510	Glu Glu Asp Cys	Pro 1515	Cys Val His
	Asn Glu 1520	Ala Thr Tyr Lys	Pro 1525	Gly Glu Thr Ile	Arg 1530	Val Asp Cys
	Asn Thr 1535	Cys Thr Cys Arg	Asn 1540	Arg Arg Trp Glu	Cys 1545	Ser His Arg
	Leu Cys 1550	Leu Gly Thr Cys	Val 1555	Ala Tyr Gly Asp	Gly 1560	His Phe Ile
	Thr Phe 1565	Asp Gly Asp Arg	Tyr 1570	Ser Phe Glu Gly	Ser 1575	Cys Glu Tyr
	Ile Leu 1580	Ala Gln Asp Tyr	Cys 1585	Gly Asp Asn Thr	Thr 1590	His Gly Thr
	Phe Arg 1595	Ile Val Thr Glu	Asn 1600	Ile Pro Cys Gly	Thr 1605	Thr Gly Thr
	Thr Cys 1610	Ser Lys Ala Ile	Lys 1615	Leu Phe Val Glu	Val 1620	Arg Thr Ala

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 1705 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

[0024]

	1880					1885						1890			
	Pro	Pro	Ser	Gln	Pro	Phe	Phe	Asn	Glu	Asp	Gln	Met	Lys	Cys	Val
	1895					1900						1905			
	Ala	Gln	Cys	Gly	Cys	Tyr	Asp	Lys	Asp	Gly	Asn	Tyr	Tyr	Asp	Val
	1910						1915					1920			
	Gly	Ala	Arg	Val	Pro	Thr	Ala	Glu	Asn	Cys	Gln	Ser	Cys	Asn	Cys
	1925						1930					1935			
	Thr	Pro	Ser	Gly	Ile	Gln	Cys	Ala	His	Ser	Leu	Glu	Ala	Cys	Thr
	1940						1945					1950			
	Cys	Thr	Tyr	Glu	Asp	Arg	Thr	Tyr	Ser	Tyr	Gln	Asp	Val	Ile	Tyr
	1955						1960					1965			
	Asn	Thr	Thr	Asp	Gly	Leu	Gly	Ala	Cys	Leu	Ile	Ala	Ile	Cys	Gly
	1970						1975					1980			
	Ser	Asn	Gly	Thr	Ile	Ile	Arg	Lys	Ala	Val	Ala	Cys	Pro	Gly	Thr
	1985						1990					1995			
	Pro	Ala	Thr	Thr	Pro	Phe	Thr	Phe	Thr	Thr	Ala	Trp	Val	Pro	His
	2000						2005					2010			
[0025]	Ser	Thr	Thr	Ser	Pro	Ala	Leu	Pro	Val	Ser	Thr	Val	Cys	Val	Arg
	2015						2020					2025			
	Glu	Val	Cys	Arg	Trp	Ser	Ser	Trp	Tyr	Asn	Gly	His	Arg	Pro	Glu
	2030						2035					2040			
	Pro	Gly	Leu	Gly	Gly	Gly	Asp	Phe	Glu	Thr	Phe	Glu	Asn	Leu	Arg
	2045						2050					2055			
	Gln	Arg	Gly	Tyr	Gln	Val	Cys	Pro	Val	Leu	Ala	Asp	Ile	Glu	Cys
	2060						2065					2070			
	Arg	Ala	Ala	Gln	Leu	Pro	Asp	Met	Pro	Leu	Glu	Glu	Leu	Gly	Gln
	2075						2080					2085			
	Gln	Val	Asp	Cys	Asp	Arg	Met	Arg	Gly	Leu	Met	Cys	Ala	Asn	Ser
	2090						2095					2100			
	Gln	Gln	Ser	Pro	Pro	Leu	Cys	His	Asp	Tyr	Glu	Leu	Arg	Val	Leu
	2105						2110					2115			
	Cys	Cys	Glu	Tyr	Val	Pro	Cys	Gly	Pro	Ser	Pro	Ala	Pro	Gly	Thr
	2120						2125					2130			
	Ser	Pro	Gln	Pro	Ser	Leu	Ser	Ala	Ser	Thr	Glu	Pro	Ala	Val	Pro
	2135						2140					2145			

	Thr 2150	Pro	Thr	Gln	Thr	Thr	Ala 2155	Thr	Glu	Lys	Thr	Thr 2160	Leu	Trp	Val
	Thr 2165	Pro	Ser	Ile	Arg	Ser	Thr 2170	Ala	Ala	Leu	Thr	Ser 2175	Gln	Thr	Gly
	Ser 2180	Ser	Ser	Gly	Pro	Val	Thr 2185	Val	Thr	Pro	Ser	Ala 2190	Pro	Gly	Thr
	Thr 2195	Thr	Cys	Gln	Pro	Arg	Cys 2200	Gln	Trp	Thr	Glu	Trp 2205	Phe	Asp	Glu
	Asp 2210	Tyr	Pro	Lys	Ser	Glu	Gln 2215	Leu	Gly	Gly	Asp	Val 2220	Glu	Ser	Tyr
	Asp 2225	Lys	Ile	Arg	Ala	Ala	Gly 2230	Gly	His	Leu	Cys	Gln 2235	Gln	Pro	Lys
	Asp 2240	Ile	Glu	Cys	Gln	Ala	Glu 2245	Ser	Phe	Pro	Asn	Trp 2250	Thr	Leu	Ala
	Gln 2255	Val	Gly	Gln	Lys	Val	His 2260	Cys	Asp	Val	His	Phe 2265	Gly	Leu	Val
[0026]	Cys 2270	Arg	Asn	Trp	Glu	Gln	Glu 2275	Gly	Val	Phe	Lys	Met 2280	Cys	Tyr	Asn
	Tyr 2285	Arg	Ile	Arg	Val	Leu	Cys 2290	Cys	Ser	Asp	Asp	His 2295	Cys	Arg	Gly
	Arg 2300	Ala	Thr	Thr	Pro	Pro	Pro 2305	Thr	Thr	Glu	Leu	Glu 2310	Thr	Ala	Thr
	Thr 2315	Thr	Thr	Thr	Gln	Ala	Leu 2320	Phe	Ser	Thr	Pro	Gln 2325	Pro	Thr	Ser
	Ser 2330	Pro	Gly	Leu	Thr	Arg	Ala 2335	Pro	Pro	Ala	Ser	Thr 2340	Thr	Ala	Val
	Pro 2345	Thr	Leu	Ser	Glu	Gly	Leu 2350	Thr	Ser	Pro	Arg	Tyr 2355	Thr	Ser	Thr
	Leu 2360	Gly	Thr	Ala	Thr	Thr	Gly 2365	Gly	Pro	Thr	Thr	Pro 2370	Ala	Gly	Ser
	Thr 2375	Glu	Pro	Thr	Val	Pro	Gly 2380	Val	Ala	Thr	Ser	Thr 2385	Leu	Pro	Thr
	Arg 2390	Ser	Ala	Leu	Pro	Gly	Thr 2395	Thr	Gly	Ser	Leu	Gly 2400	Thr	Trp	Arg

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
 [0027] 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

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 2685
 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

[0028]

	2915					2920						2925						
	Thr	Arg	Thr	Ser	Thr	Leu	Leu	Pro	Ser	Ser	Pro	Thr	Ser	Ala	Pro			
		2930					2935					2940						
	Ile	Thr	Thr	Val	Val	Thr	Met	Gly	Cys	Glu	Pro	Gln	Cys	Ala	Trp			
		2945					2950					2955						
	Ser	Glu	Trp	Leu	Asp	Tyr	Ser	Tyr	Pro	Met	Pro	Gly	Pro	Ser	Gly			
		2960					2965					2970						
	Gly	Asp	Phe	Asp	Thr	Tyr	Ser	Asn	Ile	Arg	Ala	Ala	Gly	Gly	Ala			
		2975					2980					2985						
	Val	Cys	Glu	Gln	Pro	Leu	Gly	Leu	Glu	Cys	Arg	Ala	Gln	Ala	Gln			
		2990					2995					3000						
	Pro	Gly	Val	Pro	Leu	Arg	Glu	Leu	Gly	Gln	Val	Val	Glu	Cys	Ser			
		3005					3010					3015						
	Leu	Asp	Phe	Gly	Leu	Val	Cys	Arg	Asn	Arg	Glu	Gln	Val	Gly	Lys			
		3020					3025					3030						
	Phe	Lys	Met	Cys	Phe	Asn	Tyr	Glu	Ile	Arg	Val	Phe	Cys	Cys	Asn			
		3035					3040					3045						
[0029]	Tyr	Gly	His	Cys	Pro	Ser	Thr	Pro	Ala	Thr	Ser	Ser	Thr	Ala	Met			
		3050					3055					3060						
	Pro	Ser	Ser	Thr	Pro	Gly	Thr	Thr	Trp	Ile	Leu	Thr	Glu	Leu	Thr			
		3065					3070					3075						
	Thr	Thr	Ala	Thr	Thr	Thr	Glu	Ser	Thr	Gly	Ser	Thr	Ala	Thr	Pro			
		3080					3085					3090						
	Ser	Ser	Thr	Pro	Gly	Thr	Thr	Trp	Ile	Leu	Thr	Glu	Pro	Ser	Thr			
		3095					3100					3105						
	Thr	Ala	Thr	Val	Thr	Val	Pro	Thr	Gly	Ser	Thr	Ala	Thr	Ala	Ser			
		3110					3115					3120						
	Ser	Thr	Gln	Ala	Thr	Ala	Gly	Thr	Pro	His	Val	Ser	Thr	Thr	Ala			
		3125					3130					3135						
	Thr	Thr	Pro	Thr	Val	Thr	Ser	Ser	Lys	Ala	Thr	Pro	Phe	Ser	Ser			
		3140					3145					3150						
	Pro	Gly	Thr	Ala	Thr	Ala	Leu	Pro	Ala	Leu	Arg	Ser	Thr	Ala	Thr			
		3155					3160					3165						
	Thr	Pro	Thr	Ala	Thr	Ser	Phe	Thr	Ala	Ile	Pro	Ser	Ser	Ser	Leu			
		3170					3175					3180						

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

[0030] 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

	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			
[0031]	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			

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

[0032]

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

[0034] 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

	Thr 4475	Pro	Val	Pro	Asn	Thr 4480	Thr	Ala	Thr	Thr	His 4485	Gly	Arg	Ser	
	Leu 4490	Ser	Pro	Ser	Ser	Pro	His 4495	Thr	Val	Arg	Thr	Ala 4500	Trp	Thr	Ser
	Ala 4505	Thr	Ser	Gly	Thr	Leu	Gly 4510	Thr	Thr	His	Ile	Thr 4515	Glu	Pro	Ser
	Thr 4520	Gly	Thr	Ser	His	Thr	Pro 4525	Ala	Ala	Thr	Thr	Gly 4530	Thr	Thr	Gln
	His 4535	Ser	Thr	Pro	Ala	Leu	Ser 4540	Ser	Pro	His	Pro	Ser 4545	Ser	Arg	Thr
	Thr 4550	Glu	Ser	Pro	Pro	Ser	Pro 4555	Gly	Thr	Thr	Thr	Pro 4560	Gly	His	Thr
	Thr 4565	Ala	Thr	Ser	Arg	Thr	Thr 4570	Ala	Thr	Ala	Thr	Pro 4575	Ser	Lys	Thr
	Arg 4580	Thr	Ser	Thr	Leu	Leu	Pro 4585	Ser	Ser	Pro	Thr	Ser 4590	Ala	Pro	Ile
[0035]	Thr 4595	Thr	Val	Val	Thr	Thr	Gly 4600	Cys	Glu	Pro	Gln	Cys 4605	Ala	Trp	Ser
	Glu 4610	Trp	Leu	Asp	Tyr	Ser	Tyr 4615	Pro	Met	Pro	Gly	Pro 4620	Ser	Gly	Gly
	Asp 4625	Phe	Asp	Thr	Tyr	Ser	Asn 4630	Ile	Arg	Ala	Ala	Gly 4635	Gly	Ala	Val
	Cys 4640	Glu	Gln	Pro	Leu	Gly	Leu 4645	Glu	Cys	Arg	Ala	Gln 4650	Ala	Gln	Pro
	Gly 4655	Val	Pro	Leu	Gly	Glu	Leu 4660	Gly	Gln	Val	Val	Glu 4665	Cys	Ser	Leu
	Asp 4670	Phe	Gly	Leu	Val	Cys	Arg 4675	Asn	Arg	Glu	Gln	Val 4680	Gly	Lys	Phe
	Lys 4685	Met	Cys	Phe	Asn	Tyr	Glu 4690	Ile	Arg	Val	Phe	Cys 4695	Cys	Asn	Tyr
	Gly 4700	His	Cys	Pro	Ser	Thr	Pro 4705	Ala	Thr	Ser	Ser	Thr 4710	Ala	Met	Pro
	Ser 4715	Ser	Thr	Pro	Gly	Thr	Thr 4720	Trp	Ile	Leu	Thr	Glu 4725	Leu	Thr	Thr

[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

	4985		4990		4995
	Pro Ser 5000	Ser Ser	Leu Gly	Thr 5005	Thr Trp Thr Arg 5010
	Thr Thr 5015	Pro Thr	Ala Thr	Met 5020	Ser Thr Ala Thr 5025
	Pro Glu 5030	Thr Val	His Thr	Ser 5035	Thr Val Leu Thr 5040
	Thr Thr 5045	Gly Ala	Thr Gly	Ser 5050	Val Ala Thr Pro 5055
	Gly Thr 5060	Ala His	Thr Thr	Lys 5065	Val Pro Thr Thr 5070
	Phe Thr 5075	Ala Thr	Pro Ser	Ser 5080	Ser Pro Gly Thr 5085
	Pro Thr 5090	Thr Thr	Pro Met	Ser 5095	Thr Met Ser Thr 5100
	Ser Thr 5105	Pro Glu	Thr Thr	His 5110	Thr Ser Thr Val 5115
[0037]	Ala Thr 5120	Met Thr	Arg Ala	Thr 5125	Asn Ser Thr Ala 5130
	Thr Leu 5135	Gly Thr	Thr Arg	Ile 5140	Leu Thr Glu Leu 5145
	Thr Thr 5150	Thr Ala	Ala Thr	Gly 5155	Ser Thr Ala Thr 5160
	Pro Gly 5165	Thr Thr	Trp Ile	Leu 5170	Thr Glu Pro Ser 5175
	Val Met 5180	Val Pro	Thr Gly	Ser 5185	Thr Ala Thr Ala 5190
	Gly Thr 5195	Ala His	Thr Pro	Lys 5200	Val Val Thr Thr 5205
	Pro Thr 5210	Ala Thr	Ala Ser	Thr 5215	Val Pro Ser Ser 5220
	Thr Thr 5225	Arg Thr	Pro Ala	Val 5230	Leu Pro Ser Ser 5235
	Ser Val 5240	Ser Thr	Val Ser	Ser 5245	Ser Val Leu Thr 5250

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
 [0038] 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

	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			
[0039]	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			

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

[0040]

6020	6025	6030
Ala Met Gln His Gln Cys Thr	cys cys Gln Glu Arg	Arg Val His
6035	6040	6045
Glu Glu Thr Val Pro Leu His	Cys Pro Asn Gly Ser	Ala Ile Leu
6050	6055	6060
His Thr Tyr Thr His Ala Val	Gln Val Leu Cys Gly	Leu Leu Ala
6065	6070	6075
Trp Gly Leu Gln Ala Gly Gly	His Ile Arg Gly Ala	Val Gln Asp
6080	6085	6090
Pro Gln Gln Pro Leu Lys Asp	Gln Glu Ala Ser Gly	Lys Ala Arg
6095	6100	6105
Gln Gly Gly Gly Tyr Arg Gln	Thr Val Ala Trp Gly	Asp Lys Ser
6110	6115	6120
Asn Ala Arg Ala Trp Leu Gln	Lys Pro Val Val Trp	Val Gln Ser
6125	6130	6135
Gly Ala Phe Pro Thr Pro Gly	Pro Ala Ser Ala Leu	Cys Pro Trp
6140	6145	6150
Lys Met Gly Ile Gln Pro Glu	Thr Thr Lys Gln Leu	Arg Asp Ala
6155	6160	6165
Asn Ile Leu Lys Glu Ser Lys	Arg Ser Ile Ser Arg	Glu Arg Gln
6170	6175	6180
Arg Gln Cys Ala Gln Ala Ile	Arg Phe Asn Arg Gly	Phe Gly Gly
6185	6190	6195
Gln Ile Trp Lys Ser Gln Arg	Phe Phe	
6200	6205	

[0041]

<210> 5
 <211> 4493
 <212> PRT
 <213> 人
 <400> 5

Met Pro Arg Pro Gly Thr Met Ala Leu Cys Leu Leu Thr Leu Val Leu
1 5 10 15
Ser Leu Leu Pro Pro Gln Ala Ala Ala Glu Gln Asp Leu Ser Val Asn
20 25 30
Arg Ala Val Trp Asp Gly Gly Gly Cys Ile Ser Gln Gly Asp Val Leu
35 40 45

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

[0042]

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

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

[0043]

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

	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	
[0044]	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					780				
	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

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

[0045] 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

	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			
[0046]	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

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
 [0049] 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

	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			
[0050]	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			

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

[0051] 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

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
 [0053] 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

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
 [0054] 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

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

[0055]

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
 [0057] 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

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
 [0058] 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 4405 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

<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

 [0059] 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

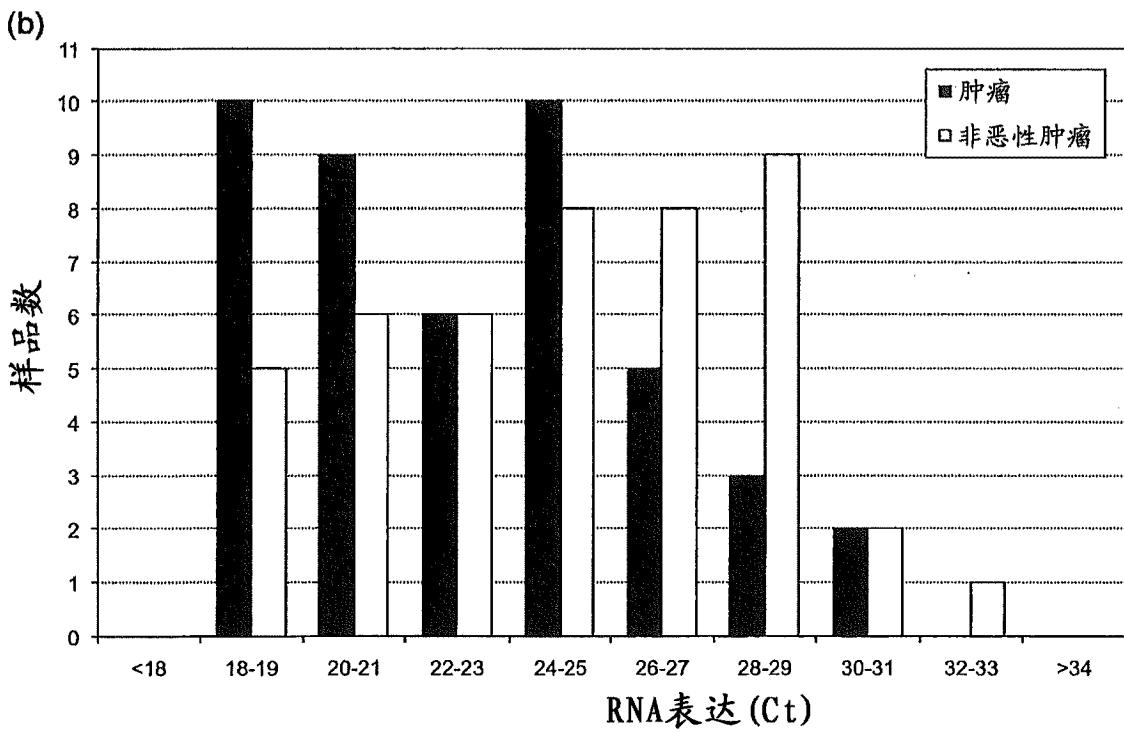
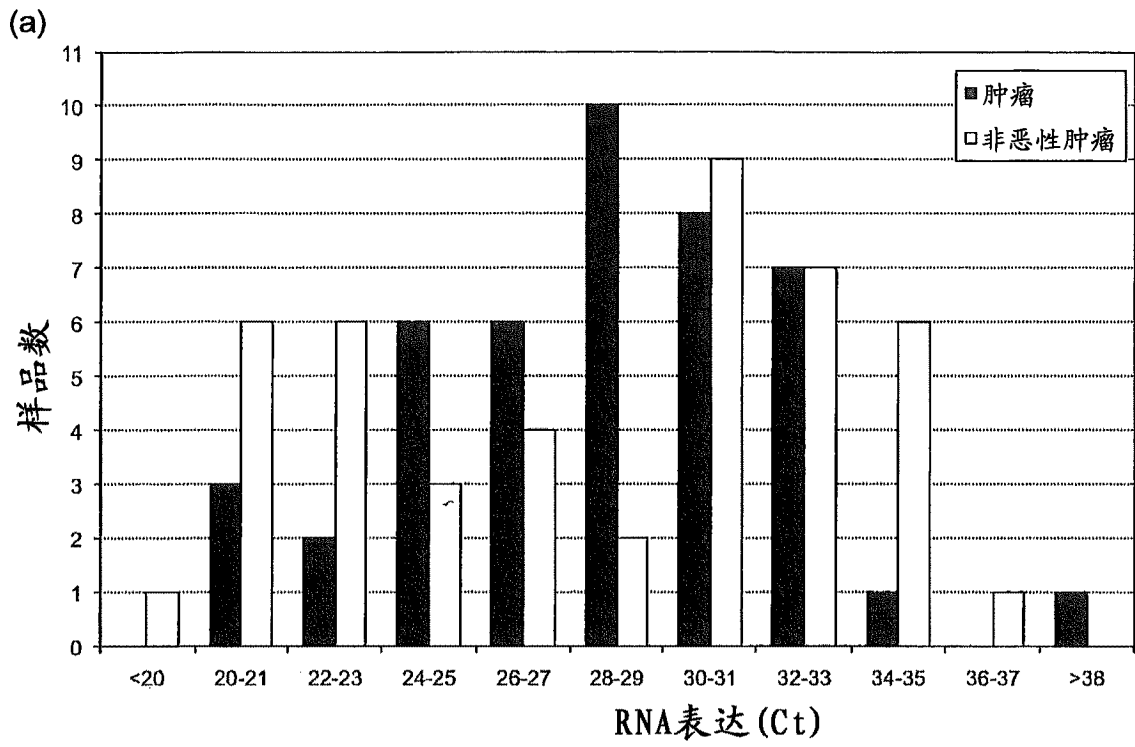


图3

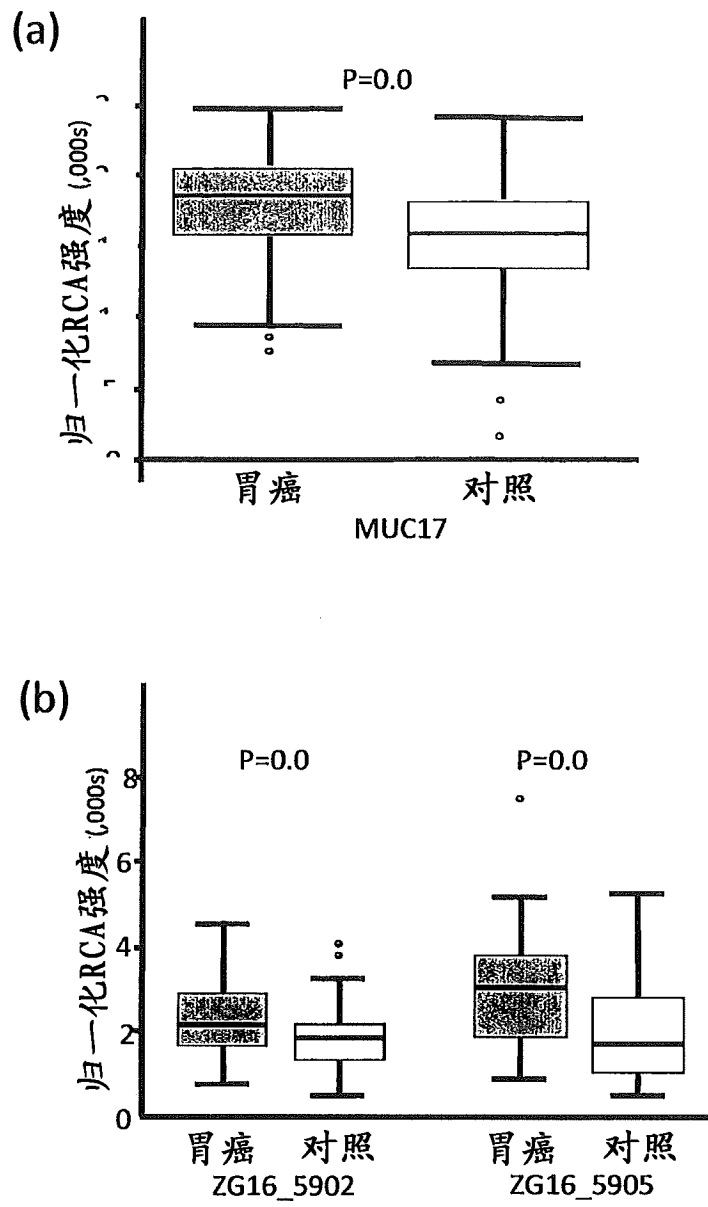


图4

专利名称(译)	胃癌的检测标志物		
公开(公告)号	CN102459646B	公开(公告)日	2016-08-10
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 C12Q1/6886 C12Q2600/158 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	CN102459646A		
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_001035 194	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