(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 103937877 B
(45)授权公告日 2016.09.28

(21)申请号 201410061691.6

(22)申请日 2011.02.04

(65)同一申请的已公布的文献号
申请公布号 CN 103937877 A

(43)申请公布日 2014.07.23

(30)优先权数据
61/337,465 2010.02.04 US

(62)分案原申请数据
201180005813.4 2011.02.04

(73)专利权人 自治医科大学
地址 日本东京

(72)发明人 马诺·真田广之 崔英李 苏达学

(74)专利代理机构 广州弘邦专利商标事务所有限公司 44236
代理人 张钇斌

(51)Int.Cl.
C12Q 1/68(2006.01)
G01N 33/68(2006.01)
A61K 45/00(2006.01)
A61P 35/00(2006.01)

(56)对比文件
CN 101466721 A,2009.06.24,
WO 01/04343 A2,2001.01.18,
Lihui Lu, et al.ALK mutants in the Kinase Domain Exhibit Altered Kinase Activity and Differential Sensitivity to Small Molecule ALK Inhibitors.《Biochemistry》.2009,第48卷(第16期),

审查员 吴希哲

权利要求书1页 说明书74页
序列表129页 附图5页

(54)发明名称

对ALK抑制剂具有先抗性或后抗性癌症的识别,判断和治疗

(57)摘要

本发明公开了那些能确定对象所患的癌症是否对ALK突变成阳性,和/或患有这种癌症的患者是否会有一个相对缓慢的疾病恶化过程的化合物,试剂盒和方法,所述的ALK突变可能在一种ALK抑制剂治疗后产生反应。本发明还公开了那些能预测患有这种癌症的患者其病情随时间发展的进程。

CN 103937877 B

1.一种用于确定一个癌症患者的化学敏感性，从而对其使用一种ALK抑制剂进行治疗的试剂盒，其特征在于包括：一种与一个或多个突变ALK多核苷酸分子或多肽形成特异性结合的试剂，所述ALK多核苷酸分子编码C1156Y突变和/或L1196M突变的ALK多肽，所述多肽包括C1156Y突变和/或L1196M突变；和使用说明书。

2.如权利要求1所述的试剂盒，其特征在于还包括一种ALK抑制剂。

3.如权利要求1所述的试剂盒，其特征在于所述的试剂包括一个或多个核酸探针，其中每一个探针包括一个多核苷酸序列，这个多核苷酸序列与一个核苷酸序列互补，这个互补的核苷酸序列编码包括C1156Y突变和/或L1196M突变的ALK多肽。

4.如权利要求3所述的试剂盒，其特征在于所述的探针包括长度为50-10^7个核苷酸的多核苷酸。

5.如权利要求3所述的试剂盒，其特征在于所述的探针包括寡核苷酸。

6.如权利要求3所述的试剂盒，其特征在于所述的探针包括cDNA分子。

7.如权利要求3所述的试剂盒，其特征在于所述的探针包括RNA分子。

8.如权利要求3所述的试剂盒，其特征在于所述的探针包括由碱基组成的合成基因探针。

9.如权利要求1所述的试剂盒，其特征在于所述的试剂包括一种抗体，一种抗体衍生物或一种抗体片段，其特异性结合包含C1156Y突变和/或L1196M突变的ALK多肽。

10.如权利要求1所述的试剂盒，其中所述突变ALK多肽是包括C1156Y突变的ALK多肽。

11.如权利要求1所述的试剂盒，其中所述突变ALK多肽是包括L1196M突变的ALK多肽。

12.一种确定一种检测化合物是否调节了一个或多个包含C1156Y突变和/或L1196M突变的ALK多肽的活性的方法，其特征在于包括：

(a)用检测化合物与哺乳动物细胞接触，其中所述哺乳动物细胞预先转染一个结构，使转染结构来编码一个或多个包含C1156Y突变和/或L1196M突变的ALK多肽；并

(b)测定该哺乳动物细胞来确定一个或多个突变ALK多肽的活性，其中相对于一个对照实验，检测化合物显著地调节了活性，因此它可以作为一个或多个突变ALK多肽的调节器。

13.如权利要求12所述的方法，其特征在于所述的一个或多个突变的ALK多肽是包含C1156Y突变的ALK多肽。

14.如权利要求12所述的方法，其特征在于所述的一个或多个突变的ALK多肽是包含L1196M突变的ALK多肽。

15.如权利要求12或13或14所述的方法，其特征在于所述的对照实验包括哺乳动物细胞表达一个野生型ALK多肽。

16.如权利要求15所述的方法，其特征在于所述的一个或多个突变ALK多肽的活性包括ATP结合，酪氨酸激酶活性，癌症细胞增殖，肿瘤生长，肿瘤数量，细胞凋亡和肿瘤转移。

17.如权利要求12或13或14所述的方法，其特征在于所述的对照实验包括哺乳动物细胞在所述的检测化合物中表达一个或多个突变ALK多肽。

18.如权利要求17所述的方法，其特征在于所述的一个或多个突变ALK多肽活性包括ATP结合，酪氨酸激酶活性，癌症细胞增殖，肿瘤生长，肿瘤数量，细胞凋亡和肿瘤转移。

对ALK抑制剂具有先抗性或后抗性癌症的识别,判断和治疗

[0001] 本分案申请是发明名称为“对ALK抑制剂具有先抗性或后抗性癌症的识别,判断和治疗”,申请号为2011800058134,申请日为2011年2月4日的原始发明专利申请的分案申请。

[0002] 相关申请

[0003] 本申请权利要求2010年02月04日申请的申请序号为61/337,465的美国临时申请的优先权。

[0004] 发明背景

[0005] 酪氨酸激酶是一类通过转移末端三磷酸腺苷来催化蛋白质底物上的酪氨酸残基磷酸化的酶。在许多情况下,酪氨酸激酶在细胞功能包括细胞增殖,癌变和细胞分化的信号转导方面具有关键作用。

[0006] EML4-ALK是一种融合型蛋白酪氨酸激酶,在5%的非小细胞肺癌(NSCLC)病例中出现,其可引起人类2号染色体短臂的微小倒位的结果(Soda,M.et al.(2007)Nature448:561-566;Mano,H.(2008)Cancer Sci.99:2349-2355)。EML4-ALK是通过每个单体EML4在卷曲螺旋结构域中相互作用而造成的二聚结构体,从而获得明显的致癌活性。用转基因老鼠来表达EML4-ALK,特别是肺上皮细胞,出生后不久就在双肺上生成了许许多多的腺癌小瘤,然后在给予口服特效的ALK酪氨酸激酶活性抑制剂后这些小瘤迅速地从肺上消除(Soda,M.et al.(2008)Proc.Natl.Acad.Sci.USA105:19893-19897)。这些观察表明,在NSCLC癌变中EML4-ALK具有生成这种融合激酶的重要作用,并进一步证明用ALK抑制剂来分子靶向治疗癌症的可行性。例如,临床验证,抑制剂PF-02341066,对ALK和MET都具有酪氨酸激酶活性,在对EML4-ALK的阳性NSCLC进行治疗,其间歇性结果较理想(Kwak,E.L.et al.(2009)J.Clin.Oncol.27(suppl):15s(abstract3509))。然而,EML4-ALK的阳性肿瘤对治疗失败的未知分子成份的抑制剂没有响应。

[0007] 除了PF-02341066,其他酪氨酸激酶抑制剂(TKIs)已被证明在对癌症患者治疗具有显的活性。甲磺酸伊马替尼,一种用于ABL1和KIT的TKI,例如,BCR-ABL1融合型激酶的阳性慢性骨髓性白血病或KIT的阳性胃肠道间质肿瘤的个案中具有明显的改善效果(Druker,B.J.et al.(2001)N.Engl.J.Med.344:1031-1037;Heinrich,M.C.et al.(2008)J.Clin.Oncol.26:5360-5367)。此外,吉非替尼和厄洛替尼,都是表皮生长因子受体(EGFR)的TKIs,在治疗EGFR活性的NSCLC中具有效果(Mok,T.S.et al.(2009)J.Clin.Oncol.27:5080-2087;Mok,T.S.et al.(2009)N.Engl.J.Med.361:947-957)。遗憾的是,肿瘤的靶向群要么从治疗开始就对相应的TKIs难以融合要么在开始响应后就产生耐药性。其次,在一些失败的案例中还发现肿瘤靶向激酶直接发生突变或变构影响ATP-结合组织的形状,从而导致影响TKI的结合(Deininger,M.et al.(2005)Blood105:2640-2653;Kobayashi,S.et al.(2005)N.Engl.J.Med.352:786-792;Pao,W.et al.(2005)PLoS Med.2:e73;Shah,N.P.et al.(2002)Cancer Cell2:117-125)。因此,迫切需要能识别出在酪氨酸激酶上的具有抗突变的位置,如EML4-ALK,以便更好的发展用于识别、判断,预防和治疗相关异常和/或活性表达的疾病的化合物,试剂盒和方法。

发明内容

[0008] 本发明至少提供了通过识别出对公知的ALK抑制剂产生抗性的新型的间变性淋巴瘤激酶(ALK)突变而识别、判断和治疗癌症的化合物，方法和试剂盒。同样，这些ALK突变能够在临床上鉴定药物组合物，这些药物组合物可以放入那些由新型ALK突变产生的异常ATP-结合区，来抑制ALK活性。

[0009] 一方面，本发明提供了一种用于识别一个患有癌症或具有罹患癌症风险的对象在使用ALK抑制剂治疗时是否会产生对治疗无反应的额外风险的方法，其包括从患者身上采集样本并分析样本来检查是否存在一个或多个突变的ALK多核苷酸分子，其中存在一个或多个突变的ALK多核苷酸分子表明对象已经具有对ALK抑制剂治疗无反应的额外风险。

[0010] 另一个方面，本发明提供了一种用于识别一个患有癌症或具有罹患癌症风险的对象在使用ALK抑制剂治疗时是否会产生对治疗无反应的额外风险的方法，其包括从患者身上采集样本并分析样本来检查一个或多个突变的ALK多肽的表达水平、结构和/或活性，其中存在一个或多个突变的ALK多肽表明对象已经具有对ALK抑制剂治疗无反应的额外风险。

[0011] 在本发明的一些实施例中，对象可以是之前没有使用一种ALK抑制剂来治疗，也可以是之前已经经用一种ALK抑制剂治疗，并至少部分地对ALK抑制剂(如PF-02341066，PDD，2-甲基-11-(2-甲基丙基)-4-氧-4,5,6,11,12,13-六氢-2H-吲唑[5,4-α]吡咯[3,4-c]咔唑-8-基[4-(二甲氨基)苄基]氨基甲酸甲酯，(1S,2S,3R,4R)-3-({5-氯-2-[(1-乙基-2,3,4,5-四氢-6-甲氧基-2-氧-1H-1-苯并氮杂卓-7-基)氨基]-4-嘧啶}氨基)双环[2.2.1]庚-5-烯-2-甲酰胺和NVP-TAE684)产生了抗性。在另一些实施例中，所述的癌症包括间变性大细胞淋巴瘤，神经母细胞瘤，乳腺癌，大肠癌，炎性肌纤维瘤，非小细胞肺癌。还有另一些实施例中，所述的样本包括痰液，支气管肺泡灌洗，胸腔积液，组织，全血，血清，血浆，口腔刮，唾液，脑脊液，尿液，粪便，循环肿瘤细胞，循环核酸和骨髓等。

[0012] 在还有另一些实施例中，所述的样本包括细胞或组织。在一些实施例中，组织是肿瘤或癌组织。在另一些实施例中，一个或多个突变的ALK多核苷酸分子或多肽包括如表1所列的突变ALK多核苷酸分子或多肽。还有一些实施例中，一个或多个突变的ALK通过核酸杂交法来判断。在另一些实施例中，一个或多个突变的ALK通过聚合酶链反应来判断。在另一些实施例中，一个或多个ALK多肽表达水平用特异性结合到一个或多个ALK多肽(如抗体，抗体衍生物，和抗体片段)的试剂来查明。在还有一些实施例中，一个或多个ALK多肽的数量，结构和/或活性与控制样本相比较。在还有一些实施例中，一个或多个突变的ALK通过启始点的时间和至少一个随后点的时间来判断。在另一些实施例中，样本包括生殖细胞或体细胞基因组DNA。

[0013] 另一个方面，本发明还提供了一种治疗癌症患者或有潜在的癌症患者的方法，包括从患者身上收集样本，分析样本来查明是否存在一个或多个如表1所列的突变ALK多核苷酸分子，然后给所述病人使用一个治疗有效的ALK抑制剂。在一些实施例中，ALK抑制剂包括

[0014] PF-02341066，PDD，2-甲基-11-(2-甲基丙基)-4-氧-4,5,6,11,12,13-六氢-2H-吲唑[5,4-α]吡咯[3,4-c]咔唑-8-基[4-(二甲氨基)苄基]氨基甲酸甲酯，

[0015] (1S,2S,3R,4R)-3-({5-氯-2-[(1-乙基-2,3,4,5-四氢-6-甲氧基-2-氧-1H-1-苯并氮杂卓-7-基)氨基]-4-嘧啶}氨基)双环[2.2.1]庚-5-烯-2-甲酰胺和NVP-TAE684。在另

一些实施例中，对象可以是之前没有使用一种ALK抑制剂来治疗，也可以是之前已经经用一种ALK抑制剂治疗，并至少部分地对ALK抑制剂产生了抗性。

[0016] 另一个方面，本发明还提供了一种用于确定一个癌症患者的化学敏感性，从而使用一种ALK抑制剂进行治疗的试剂盒，包括：一种与一个或多个突变的ALK多核苷酸分子或多肽产生特定结合的试剂；和使用说明书。在一些实施例中，该试剂盒还包括一种ALK抑制剂。在另一些实施例中，该试剂包括一个或多个多核苷酸探针，每个探针包括一个多核苷酸序列，这个多核苷酸序列与表1所列的其中一个核苷酸序列互补，或者与一个编码表1所列的其中一个多肽的核苷酸序列互补(例如寡核苷酸，cRNA分子，RNA分子和由碱基组成的合成基因探针)。还有一些实施例中，探针包括长度约为50至10^7个核苷酸的多核苷酸。有还有一些实施例中，试剂包括一种由表1所列的一个或多个核苷酸序列编码而成的一种抗体，抗体衍生物和抗体片段到一种多肽。

[0017] 另一个方面，本发明还提供了一种方法来确定一种检测化合物是否控制了一个或多个突变ALK多肽的活性，包括与哺乳动物细胞接触，使其转染一个结构，并用检测化合物来编码一个或多个突变的ALK多肽，并测定该哺乳动物细胞来确定一个或多个突变的ALK多肽的活性，其中由于被检测化合物能在一个控制表达中显著地调节活性，因此它可以作为一个或多个突变ALK多肽的调节器。在一些实施例中，所述的一个或多个突变的ALK多核苷酸分子或多肽包括表1所列的突变ALK多核苷酸分子或多肽。在另一些实施例中，所述的控制包括哺乳动物细胞表达一种野生型ALK多肽，多肽可以是表1所列的多肽。在另一些实施例中，一个或多个突变ALK多肽的活性可以是ATP结合，酪氨酸激酶活性，癌症细胞增殖，肿瘤生长，肿瘤数量，细胞凋亡和肿瘤转移。在另一些实施例中，所述的控制表达包括哺乳动物细胞在检测化合物的存在下，表达一个或多个突变ALK多肽，其中可以例如用一个或多个突变ALK多肽的活性(例如ATP结合，酪氨酸激酶活性，癌症细胞增殖，肿瘤生长，肿瘤数量，细胞凋亡和肿瘤转移)来确定检测化合物的存在。

附图说明

[0018] 图1所示的是本发明所述的能抗ALK酪氨酸激酶抑制剂的新型ALK突变。图1A是EML4-ALK蛋白的结构示意图。图中显示了激酶结构域中两个de novo突变的位置，并用实心或空心箭头分别图示了上述用于扩增激酶结构域或融合cDNA的那些PCR引物。图1B显示了ALK激酶结构域cDNA的深度测序结果。NSCLC细胞系H2228或编号为J-#1，J-#12，J-#113，J-#127或LK-#33的样本中获得的约1000bp的PCR产物是用GAII系统测序的。激酶结构域cDNA的每个位置上的总读数(总)和错配读数(错配)的数目分别用蓝色和红色菱形表示。并放大地显示了样本J-#1和J-#113中cDNA5'端处的插入(用绿色方块表示)。图1C是G4374和C4493周围ALK cDNA克隆的电泳图。PCR是用cDNA来完成的，所述的cDNA分别是从未治疗前采样的痰中(开始)和已经恶化后采样的胸腔积液细胞中(恶化)提取的。取代的A核苷酸用红色显示。

[0019] 图2图示了ALK cDNA中G4374和C4493对应位置处的基因序列。病患的胸腔积液细胞中分离出来的基因DNA可以用Platinum@DNA聚合酶(Invitrogen，Carlsbad，CA)，下述引物(5'-GGTAAGAAGTGGCTCACTCTTGAG-3'和5'-CACAACAACTGCAGCAAAGACTGG-3')进行PCR，94℃下15s，60℃下30s，68℃下2分钟，35个循环，获得的产物导入到质粒pT7Blue-2(Takara Bio)中。然后将插入后的质粒用3130xl基因分析仪测序，以鉴别含有G4374A(图左侧)or

C4493A(图右侧)取代的PCR克隆。取代的A核苷酸用红色显示。

[0020] 图3显示了用PF-02341066处理BA/F3细胞的结果。BA/F3表达了EML4-ALK(野生型),EML4-ALK(C1156Y),EML4-ALK(L1196M),或双突变的EML4-ALK(C1156Y/L1196M),它是在所示浓度的PF-02341066中培养48小时,然后用相差显微镜来观察其细胞形态。比例尺:20μm。图4显示了本发明所述的能抗ALK酪氨酸激酶抑制剂的新型ALK突变的性能。图4A显示了5×10^5个BA/F3细胞在PF-02341066的所示浓度下培养48h后表达EML4-ALK(野生型),EML4-ALK(C1156Y),EML4-ALK(L1196M)或双突变EML4-ALK(C1156Y/L1196M)的数量。图中显示了BA/F3表达野生型EML4-ALK相对的可用细胞的百分比。数据是指三个单独实验获得的±s.d.。图4B显示了在野生型的酪氨酸磷酸化作用中或者突变型EML4-ALK中PF-02341066的效果。BA/F3细胞表达了FLAG标记的野生型EML4-ALK或其突变,将其置于所示浓度的PF-02341066中15小时,然后EML4-ALK从细胞溶解产物中免疫沉淀出来,然后进行免疫印迹分析抗体对Tyr^{1604}磷酸化的ALK或FLAG表位(ALK).的特异性。将表达一种EML4-ALK(KM)不活跃突变的细胞作为阴性对照。图4C显示了一种体外激酶评估,它能用于评估从BA/F3细胞(未转入ALK抑制剂中)中免疫沉淀出来的FLAG标记的野生型EML4-ALK或其突变。所述的免疫沉淀是在[$\gamma-^{32}$P]ATP,一种合成肽和所述浓度的PF-02341066中进行的。肽底物免疫沉淀的磷酸化作用分别用于免疫印迹分析抗体抗FLAG的性能(图下侧)。

[0021] 图5是ALK激酶结构域的三维结构域模型图。ALK的氨基酸位置是重叠在胰岛素受体的晶体结构上,该胰岛素受体有一个结合的ATP同源物(日本蛋白信息银行中的编号为“1ir3”,可以在pdbj.org/index.html的万维网上查询)。右侧图显示了从左侧所示的模型左侧获得的蛋白结构。A螺旋和β折叠分别用红色和黄色来表示。图中还显示了螺旋αC,Cys^{1156},andLeu^{1196}的位置。

[0022] 发明详述

[0023] 本发明至少有部分内容是识别特定基因序列,包括如间变性淋巴瘤激酶(ALK)突变,然后判定ALK抑制剂是否能有效地治疗癌症。尤其是在本发明中,已经识别出的新型ALK基因突变(如EML4-ALK多肽编码的突变)会使多肽至少能部分地对ALK抑制剂疗法产生抗性。本发明还公开了那些使用本领域的公知技术来识别这些特定基因序列的方法,这些方法非限制性地包括寡核苷酸基因芯片(Brennan,et al.(2004)Cancer Res.64(14):4744-8;Lucito,et al.(2003)Genome Res.13:2291-2305;Bignell et al.(2004)Genome Res.14:287-295;Zhao,et al(2004)Cancer Research,64(9):3060-71)和本发明所述的其它方法,例如包括聚合酶链式反应(PCR)和直接测序法。本发明还公开了这些方法中使用的诊断试剂盒。

[0024] 本发明所述的其它方面将在下文中进一步详细说明。

[0025] I.定义

[0026] 本文中的“一个”和“一种”指一种或一种以上(如至少一种)的语法对象。例如,“一种成分”即一种成分或一种以上的成分。

[0027] 术语标记的“变化量”或标记的“变化水平”涉及标记或染色体区增加或减少的拷贝数,如ALK基因突变和/或基因产物(如表1所列的标记),和/或在一个癌症样本中,一个或一些特定标记基因与在一个控制样本中,与标记拷贝量的表达水平相比,增加或减少的表达量。术语标记的“可选量”也包括在一个如癌症样本等样本中,一个标记与正常控制样本

中的标记蛋白质水平相比，增加或减少的蛋白质水平。

[0028] 术语ALK基因突变和/或基因产物(如表1所列的标记)的“变化的表达水平”是指在一个检测样本，如从一个病症患者身上采集的样本中涉及一个标记的表达水平或拷贝数，它们大于或小于标准错误试验的表达水平或拷贝数，可以至少是一个控制样本(如从一个未患癌症的健康对象身上采集的样本)中ALK基因突变和/或基因产物(如表1所列的标记)的表达水平或拷贝数的两倍，两三倍，二四倍，二五倍或二十倍或更多倍，或者是在各种控制样本中，ALK基因突变和/或基因产物(如表1所列的标记)的平均表达水平或拷贝数。变化的表达水平大于或小于标准错误试验的表达水平或拷贝数，可以至少是一个控制样本(如从一个未患癌症的健康对象身上采集的样本)中ALK基因突变和/或基因产物(如表1所列的标记)的表达水平或拷贝数的两倍，三倍，四倍，五倍或十倍或更多倍，或者是各种控制样本中，ALK基因突变和/或基因产物(如表1所列的标记)的平均表达水平或拷贝数。

[0029] 术语标记的“变化活性”涉及一个标记的活性，在生病状态时，如一个癌症样本里会比正常控制样本中的标记活性更高或更低。标记的变化活性可以是例如由标记的表达变化，标记的蛋白质水平变化，标记的结构变化，如与其它在相同或不同的路径中作为标记的蛋白质发生一个变化的反应而引起的。

[0030] 术语标记的“结构变化”涉及突变的存在，或者是标记基因或标记蛋白质的突变，如较之正常或野生型基因或蛋白质，影响标记表达或活性的突变。例如，突变非限制性地包括插入和内部染色体重组，替换，删除和插入突变。突变可以在标记的编码区或未编码区发生。

[0031] 本发明中，“间变性淋巴瘤激酶”和“ALK”是同一个含意，均指任何来源(如啮齿类动物，人类和其他哺乳动物)的原生间变性淋巴瘤激酶和它的一些变化和突变。在一些实施例中，ALK蛋白质由参考序号识别号NP_004295的NCBI表达。除了识别之外，该术语还涉及人体蛋白。在本发明中，所述的基因编码ALK也可以称为“ALK”。在一些实施例中，ALK核苷酸序列可以由参考序号识别号NM_004304.3和的NCBI和GenBank登录号为29029631表达，本领域的普通技术人员能简单地识别本发明相关的序列(如编码，5’端非编码区，3’端非编码区，转录启动，翻译启动，转录停止，翻译停止等序列)。

[0032] 此外在本发明中，“间变性淋巴瘤激酶”和“ALK”还包括本领域普通技术人员公知的ALK融合蛋白激酶及其各种突变体。这些ALK融合蛋白激酶及其各种突变体包括ALK蛋白激酶活性，并将本发明所述的突变翻译成抗ALK抑制剂的ALK蛋白激酶活性。典型的实例包括EML4-ALK突变1(AB274722.1;BAF73611.1)，EML4-ALK突变2(AB275889.1;BAF73612.1)，EML4-ALK突变3a(AB374361.1;BAG55003.1)，EML4-ALK突变3b(AB374362.1;BAG55004.1)，EML4-ALK突变4(AB374363.1;BAG75147.1)，EML4-ALK突变5a(AB374364.1;BAG75148.1)，EML4-ALK突变5b(AB374365.1;BAG75149.1)，EML4-ALK突变6(AB462411.1;BAH57335.1)，EML4-ALK突变7(AB462412.1;BAH57336.1)，KIF5B-ALK(AB462413.1;BAH57337.1)，NPM-ALK，TPM3-ALK，TFGXL-ALK，TFGL-ALK，TFGS-ALK，ATIC-ALK，CLTC-ALK，MSN-ALK，TPM4-ALK，MYH9-ALK，RANBP2-ALK，ALO17-ALK和CARS-ALK(例如参见Pulford et al.，(2004)J.Cell.Physiol.199:330-358，本发明引用参考)。此外，本领域的普通技术人员明白一种ALK激酶与其融合体发生的特别融合能使ALK蛋白激酶突变增加(如EML4至少能与外显子2，6a，6b，13，14和/或15融合，例如如Horn and Pao，(2009)J.Clin.Oncol.27:4247-4253所

述,本发明引用参考)。例如,下面是本发明所述的典型ALK序列:

[0033] 表1

[0034] 野生型ALK cDNA序列(NM_004304.3;GI:29029631):

[0035] 1gggggcggca gcggtggtag cagctggtac ctcccgccgc ctctgttcgg agggtcgcgg

[0036] 61ggcaccgagg tgctttccgg ccgccctctg gtcggccacc caaagccgcg ggcgctgatg

[0037] 121atgggtgagg agggggcggc aagatttcgg gcgcccctgc cctgaacgcc ctcagctgct181gccgccgggg ccgctccagt gcctgcgaac tctgaggagc cgaggcgccg gtgagagcaa241ggacgctgca aacttgcgca gcgcgggggc tgggattcac gcccagaagt tcagcaggca301gacagtccga agccttcccg cagcggagag atagcttgag ggtgcgcaag acggcagcct361ccgccctcgg ttcccgccca gaccgggcag aagagcttgg aggagccaaa aggaacgcaa421aaggcggcca ggacagcgtg cagcagctgg gagccgccgt tctcagcctt aaaagttgca481gagattggag gctgccccga gaggggacag accccagctc cgactgcggg gggcaggaga541ggacggtacc caactgccac ctcccttcaa ccatagtagt tcctctgtac cgagcgcagc601gagctacaga cgggggcgcg gcactcggcg cggagagcgg gaggctcaag gtcccagcca661gtgagcccag tgtgcttgag tgtctctgga ctcgcccctg agcttccagg tctgtttcat721ttagactcct gctcgcctcc gtgcagttgg gggaaagcaa gagacttgcg cgcacgcaca781gtcctctgga gatcaggtgg aaggagccgc tgggtaccaa ggactgttca gagcctcttc841ccatctcggg gagagcgaag ggtgaggctg ggcccggaga gcagtgtaaa cggcctcctc901cggcgggatg ggagccatcg ggctcctgtg gctcctgccg ctgctgcttt ccacggcagc961tgtgggctcc gggatgggga ccggccagcg cgcgggctcc ccagctgcgg ggccgccgct1021gcagccccgg gagccactca gctactcgcg cctgcagagg aagagtctgg cagttgactt1081cgtggtgccc tcgctcttcc gtgtctacgc ccgggaccta ctgctgccac catcctcctc1141ggagctgaag gctggcaggc ccgaggcccg cggctcgcta gctctggact gcgccccgct1201gctcaggttg ctggggccgg cgccgggggt ctcctggacc gccggttcac cagccccggc1261agaggcccgg acgctgtcca gggtgctgaa gggcggctcc gtgcgcaagc tccggcgtgc1321caagcagttg gtgctggagc tgggcgagga ggcgatcttg gagggttgcg tcgggcccccc1381cggggaggcg gctgtggggc tgctccagtt caatctcagc gagctgttca gttggtggat1441tcgccaaggc gaagggcgac tgaggatccg cctgatgccc gagaagaagg cgtcggaagt1501gggcagagag ggaaggctgt ccgcggcaat tcgcgcctcc cagccccgcc ttctcttcca1561gatcttcggg actggtcata gctccttgga atcaccaaca aacatgcctt ctccttctcc1621tgattatttt acatggaatc tcacctggat aatgaaagac tccttccctt tcctgtctca1681tcgcagccga tatggtctgg agtgcagctt tgacttcccc tgtgagctgg agtattcccc1741tccactgcat gacctcagga accagagctg gtcctggcgc cgcatccccc ccgaggaggc1801ctcccagatg gacttgctgg atgggcctgg ggcagagcgt tctaaggaga tgcccagagg1861ctcctttctc cttctcaaca cctcagctga ctccaagcac accatcctga gtccgtggat1921gaggagcagc agtgagcact gcacactggc cgtctcggtg cacaggcacc tgcagccctc1981tggaaggtac attgcccagc tgctgcccca caacgaggct gcaagagaga tcctcctgat2041gcccactcca gggaagcatg gttggacagt gctccaggga agaatcgggc gtccagacaa2101cccatttcga gtggccctgg aatacatctc cagtggaaac cgcagcttgt

ctgcagtgga2161cttctttgcc ctgaagaact gcagtgaagg aacatcccca ggctccaaga
tggccctgca2221gagctccttc acttgttgga atgggacagt cctccagctt gggcaggcct
gtgacttcca2281ccaggactgt gcccagggag aagatgagag ccagatgtgc cggaaactgc
ctgtgggttt2341ttactgcaac tttgaagatg gcttctgtgg ctggacccaa ggcacactgt
caccccacac2401tcctcaatgg caggtcagga ccctaaagga tgcccggttc caggaccacc
aagaccatgc2461tctattgctc agtaccactg atgtccccgc ttctgaaagt gctacagtga
ccagtgctac2521gtttcctgca ccgatcaaga gctctccatg tgagctccga atgtcctggc
tcattcgtgg2581agtcttgagg ggaaacgtgt ccttggtgct agtggagaac aaaaccggga
aggagcaagg2641caggatggtc tggcatgtcg ccgcctatga aggcttgagc ctgtggcagt
ggatggtgtt2701gcctctcctc gatgtgtctg acaggttctg gctgcagatg gtcgcatggt
ggggacaagg2761atccagagcc atcgtggctt ttgacaatat ctccatcagc ctggactgct
acctcaccat2821tagcggagag gacaagatcc tgcagaatac agcacccaaa tcaagaaacc
tgtttgagag2881aaacccaaac aaggagctga aacccgggga aaattcacca agacagaccc
ccatctttga2941ccctacagtt cattggctgt tcaccacatg tggggccagc gggccccatg
gccccaccca3001ggcacagtgc aacaacgcct accagaactc caacctgagc gtggaggtgg
ggagcgaggg3061ccccctgaaa ggcatccaga tctggaaggt gccagccacc gacacctaca
gcatctcggg3121ctacggagct gctggcggga aaggcgggaa gaacaccatg atgcggtccc
acggcgtgtc3181tgtgctgggc atcttcaacc tggagaagga tgacatgctg tacatcctgg
ttgggcagca3241gggagaggac gcctgcccca gtacaaacca gttaatccag aaagtctgca
ttggagagaa3301caatgtgata gaagaagaaa tccgtgtgaa cagaagcgtg catgagtggg
caggaggcgg3361aggaggaggg ggtggagcca cctacgtatt taagatgaag gatggagtgc
cggtgcccct3421gatcattgca gccggaggtg gtggcagggc ctacggggcc aagacagaca
cgttccaccc3481agagagactg gagaataact cctcggttct agggctaaac ggcaattccg
gagccgcagg3541tggtggaggt ggctggaatg ataacacttc cttgctctgg gccggaaaat
ctttgcagga3601gggtgccacc ggaggacatt cctgccccca ggccatgaag aagtgggggt
gggagacaag3661agggggtttc ggagggggtg gagggggtg ctcctcaggt ggaggaggcg
gaggatatat3721aggcggcaat gcagcctcaa acaatgaccc cgaaatggat ggggaagatg
gggtttcctt3781catcagtcca ctgggcatcc tgtacacccc agctttaaaa gtgatggaag
gccacgggga3841agtgaatatt aagcattatc taaactgcag tcactgtgag gtagacgaat
gtcacatgga3901ccctgaaagc cacaaggtca tctgcttctg tgaccacggg acggtgctgg
ctgaggatgg3961cgtctcctgc attgtgtcac ccaccccgga gccacacctg ccactctcgc
tgatcctctc4021tgtggtgacc tctgccctcg tggccgccct ggtcctggct ttctccggca
tcatgattgt4081gtaccgccgg aagcaccagg agctgcaagc catgcagatg gagctgcaga
gccctgagta4141caagctgagc aagctccgca cctcgaccat catgaccgac tacaacccca
actactgctt4201tgctggcaag acctcctcca tcagtgacct gaaggaggtg ccgcggaaaa
acatcaccct4261cattcggggt ctgggccatg gcgcctttgg ggaggtgtat gaaggccagg
tgtccggaat4321gcccaacgac ccaagccccc tgcaagtggc tgtgaagacg ctgcctgaag
tgtgctctga4381acaggacgaa ctggatttcc tcatggaagc cctgatcatc agcaaattca
accaccagaa4441cattgttcgc tgcattgggg tgagcctgca atccctgccc cggttcatcc

tgctggagct4501catggcgggg ggagacctca agtccttcct ccgagagacc cgccctcgcc cgagccagcc4561ctcctccctg gccatgctgg accttctgca cgtggctcgg gacattgcct gtggctgtca4621gtatttggag gaaaaccact tcatccaccg agacattgct gccagaaact gcctcttgac4681ctgtccaggc cctggaagag tggccaagat tggagacttc gggatggccc gagacatcta4741cagggcgagc tactatagaa agggaggctg tgccatgctg ccagttaagt ggatgccccc4801agaggccttc atggaaggaa tattcacttc taaaacagac acatggtcct ttggagtgct4861gctatgggaa atcttttctc ttggatatat gccatacccc agcaaaagca accaggaagt4921tctggagttt gtcaccagtg gaggccggat ggacccaccc aagaactgcc ctgggcctgt4981ataccggata atgactcagt gctggcaaca tcagcctgaa gacaggccca actttgccat5041catttggag aggattgaat actgcaccca ggacccggat gtaatcaaca ccgctttgcc5101gatagaatat ggtccacttg tggaagagga agagaaagtg cctgtgaggc ccaaggaccc5161tgaggggttt cctcctctcc tggtctctca acaggcaaaa cgggaggagg agcgcagccc5221agctgcccca ccacctctgc ctaccacctc ctctggcaag gctgcaaaga aacccacagc5281tgcagagatc tctgttcgag tccctagagg gccggccgtg gaaggggac acgtgaatat5341ggcattctct cagtccaacc ctccttcgga gttgcacaag gtccacggat ccagaaacaa

[0038] 5401gcccaccagc ttgtggaacc caacgtacgg ctcctggttt acagagaaac ccaccaaaaa

[0039] 5461gaataatcct atagcaaaga aggagccaca cgacaggggt aacctggggc tggagggaag

[0040] 5521ctgtactgtc ccacctaacg ttgcaactgg gagacttccg ggggcctcac tgctcctaga

[0041] 5581gccctcttcg ctgactgcca atatgaagga ggtacctctg ttcaggctac gtcacttccc

[0042] 5641ttgtgggaat gtcaattacg gctaccagca acagggcttg cccttagaag ccgctactgc

[0043] 5701ccctggagct ggtcattacg aggataccat tctgaaaagc aagaatagca tgaaccagcc

[0044] 5761tgggccctga gctcggtcgc acactcactt ctcttccttg ggatccctaa gaccgtggag

[0045] 5821gagagagagg caatggctcc ttcacaaacc agagaccaaa tgtcacgttt tgttttgtgc

[0046] 5881caacctattt tgaagtacca ccaaaaaagc tgtattttga aaatgcttta gaaaggtttt

[0047] 5941gagcatgggt tcatcctatt ctttcgaaag aagaaaatat cataaaaatg agtgataaat

[0048] 6001acaaggccca gatgtggttg cataaggttt ttatgcatgt ttgttgtata cttccttatg

[0049] 6061cttctttcaa attgtgtgtg ctctgcttca atgtagtcag aattagctgc ttctatgttt

[0050] 6121catagttggg gtcatagatg tttccttgcc ttgttgatgt ggacatgagc catttgaggg

[0051] 6181gagagggaac ggaaataaag gagttatttg taatgactaa aa

[0052] 野生型cDNA序列TGC(4373至4375)编码除半胱氨酸之外的氨基酸的密码子突变或者与一种与之相同的对应突变。

[0053] 野生型cDNA序列CTG(4493至4495)编码除亮氨酸之外的氨基酸的密码子突变或者与一种与之相同的对应突变。

[0054] 野生型cDNA序列G4374A突变或一种与之相同的对应突变。

[0055] 野生型cDNA序列G4493A突变或一种与之相同的对应突变。

[0056] 野生型ALK蛋白质序列(NP_004295.2;GI:29029632):

[0057] 1mgaigllwll plllstaavg sgmgtgqrag spaagpplqp replsysrlq rkslavdfvv

[0058] 61pslfrvyard lllppsssel kagrpeargs laldcapllr llgpapgvsw tagspapaea

[0059] 121rtlsrvlkgg svrklrrakq lvlelgeeai legcvgppge aavgllqfnl selfswwirq

[0060] 181gegrlrirlm pekkasevgr egrlsaaira sqprllfqif gtghsslesp tnmpspspdy

[0061] 241ftwnltwimk dsfpflshrs ryglecsfdf pceleysppl hdlrnqswsw rripseeasq

[0062] 301mdlldgpgae rskemprgsf lllntsadsk htilspwmrs ssehctlavs vhrhlqpsgr

[0063] 361yiaqllphne aareillmpt pgkhgwtvlq grigrpdnpf rvaleyissg nrslsavdff

[0064] 421alkncsegts pgskmalqss ftcwngtvlq lgqacdfhqd caqgedesqm crklpvgfyc

[0065] 481nfedgfcgwt qgtlsphtpq wqvrtlkdar fqdhqdhall lsttdvpase satvtsatfp

[0066] 541apiksspcel rmswlirgvl rgnvslvlve nktgkeqgrm vwhvaayegl slwqwmvlpl

[0067] 601ldvsdrfwlq mvawwgqgsr aivafdnisi sldcyltisg edkilqntap ksrnlfernp

[0068] 661nkelkpgens prqtpifdpt vhwlfttcga sgphgptqaq cnnayqnsnl svevgsegpl

[0069] 721kgiqiwkvpa tdtysisgyg aaggkggknt mmrshgvsvl gifnlekddm lyilvgqqge

[0070] 781dacpstnqli qkvcigennv ieeeirvnrs vhewaggggg gggatyvfkm kdgvpvplii

[0071] 841aaggggrayg aktdtfhper lennssvlgl ngnsgaaggg ggwndntsll wagkslqega

[0072] 901tgghscpqam kkwgwetrgg fggggggcss ggggggyigg naasnndpem dgedgvsfis

[0073] 961plgilytpal kvmeghgevn ikhylncshc evdechmdpe shkvicfcdh gtvlaedgvs

[0074] 1021civsptpeph lplslilsvv tsalvaalvl afsgimivyr rkhqelqamq melqspeykl

[0075] 1081sklrtstimt dynpnycfag ktssisdlke vprknitlir glghgafgev yegqvsgmpn

[0076] 1141dpsplqvavk tlpevcseqd eldflmeali iskfnhqniv rcigvslqsl prfillelma

[0077] 1201ggdlksflre trprpsqpss lamldllhva rdiacgcqyl eenhfihrdi aarnclltcp

[0078] 1261gpgrvakigd fgmardiyra syyrkggcam lpvkwmppea fmegiftskt dtwsfgvllw

[0079] 1321eifslgympy psksnqevle fvtsggrmdp pkncpgpvyr imtqcwqhqp edrpnfaiil

[0080] 1381erieyctqdp dvintalpie ygplveeeek vpvrpkdpeg vppllvsqqa kreeerspaa

[0081] 1441ppplpttssg kaakkptaae isvrvprgpa vegghvnmaf sqsnppselh kvhgsrnkpt

[0082] 1501slwnptygsw ftekptkknn piakkephdr gnlglegsct vppnvatgrl pgaslleps

[0083] 1561sltanmkevp lfrlrhfpcg nvnygyqqqg lpleaatapg aghyedtilk sknsmnqpgp

[0084] 野生型蛋白质序列Cys1156Xaa突变，其中Xaa除一种半胱氨酸之外的氨基酸或一种与之相同的对应突变。

[0085] 野生型蛋白质序列Leu1196Xaa突变，其中Xaa除一种亮氨酸之外的氨基酸或一种与之相同的对应突变。

[0086] 野生型蛋白质序列Cys1156Tyr突变或一种与之相同的对应突变。

[0087] 野生型蛋白质序列Leu1196Met突变或一种与之相同的对应突变。

[0088] EML4-ALK突变1cDNA序列(AB274722.1;GI:152002652)

[0089] 1ggcggcgcgg cgcggcgctc gcggctgctg cctgggaggg aggccgggca ggcggctgag

[0090] 61cggcgcggct ctcaacgtga cggggaagtg gttcgggcgg ccgcggctta ctaccccagg

[0091] 121gcgaacggac ggacgacgga ggcgggagcc ggtagccgag ccgggcgacc tagagaacga181gcgggtcagg ctcagcgtcg gccactctgt cggtccgctg aatgaagtgc ccgcccctct241gagcccggag cccggcgctt tccccgcaag atggacggtt tcgccggcag tctcgatgat301agtatttctg ctgcaagtac ttctgatgtt caagatcgcc tgtcagctct tgagtcacga361gttcagcaac aagaagatga aatcactgtg ctaaaggcgg ctttggctga

tgttttgagg421cgtcttgcaa tctctgaaga tcatgtggcc tcagtgaaaa aatcagtctc
aagtaaaggc481caaccaagcc ctcgagcagt tattcccatg tcctgtataa ccaatggaag
tggtgcaaac541agaaaaccaa gtcataccag tgctgtctca attgcaggaa aagaaactct
ttcatctgct601gctaaaagtg gtacagaaaa aaagaaagaa aaaccacaag gacagagaga
aaaaaaagag661gaatctcatt ctaatgatca aagtccacaa attcgagcat caccttctcc
ccagccctct721tcacaacctc tccaaataca cagacaaact ccagaaagca agaatgctac
tcccaccaaa781agcataaaac gaccatcacc agctgaaaag tcacataatt cttgggaaaa
ttcagatgat841agccgtaata aattgtcgaa aataccttca acacccaaat taataccaaa
agttaccaaa901actgcagaca agcataaaga tgtcatcatc aaccaagaag gagaatatat
taaaatgttt961atgcgcggtc ggccaattac catgttcatt ccttccgatg ttgacaacta
tgatgacatc1021agaacggaac tgcctcctga gaagctcaaa ctggagtggg catatggtta
tcgaggaaag1081gactgtagag ctaatgttta ccttcttccg accggggaaa tagtttattt
cattgcatca1141gtagtagtac tatttaatta tgaggagaga actcagcgac actacctggg
ccatacagac1201tgtgtgaaat gccttgctat acatcctgac aaaattagga ttgcaactgg
acagatagct1261ggcgtggata aagatggaag gcctctacaa ccccacgtca gagtgtggga
ttctgttact1321ctatccacac tgcagattat tggacttggc acttttgagc gtggagtagg
atgcctggat1381ttttcaaaag cagattcagg tgttcattta tgtgttattg atgactccaa
tgagcatatg1441cttactgtat gggactggca gaagaaagca aaaggagcag aaataaagac
aacaaatgaa1501gttgttttgg ctgtggagtt tcacccaaca gatgcaaata ccataattac
atgcggtaaa1561tctcatattt tcttctggac ctggagcggc aattcactaa caagaaaaca
gggaatttttt1621gggaaatatg aaaagccaaa atttgtgcag tgtttagcat tcttggggaa
tggagatgtt1681cttactggag actcaggtgg agtcatgctt atatggagca aaactactgt
agagcccaca1741cctgggaaag gacctaaagt gtaccgccgg aagcaccagg agctgcaagc
catgcagatg1801gagctgcaga gccctgagta caagctgagc aagctccgca cctcgaccat
catgaccgac1861tacaacccca actactgctt tgctggcaag acctcctcca tcagtgacct
gaaggaggtg1921ccgcggaaaa acatcaccct cattcggggt ctgggccatg gagcctttgg
ggaggtgtat1981gaaggccagg tgtccggaat gcccaacgac ccaagccccc tgcaagtggc
tgtgaagacg2041ctgcctgaag tgtgctctga acaggacgaa ctggatttcc tcatggaagc
cctgatcatc2101agcaaattca accaccagaa cattgttcgc tgcattgggg tgagcctgca
atccctgccc2161cggttcatcc tgctggagct catggcgggg ggagacctca agtccttcct
ccgagagacc2221cgccctcgcc cgagccagcc ctcctccctg gccatgctgg accttctgca
cgtggctcgg2281gacattgcct gtggctgtca gtatttggag gaaaaccact tcatccaccg
agacattgct2341gccagaaact gcctcttgac ctgtccaggc cctggaagag tggccaagat
tggagacttc2401gggatggccc gagacatcta cagggcgagc tactatagaa agggaggctg
tgccatgctg2461ccagttaagt ggatgccccc agaggccttc atggaaggaa tattcacttc
taaaacagac2521acatggtcct ttggagtgct gctatgggaa atcttttctc ttggatatat
gccatacccc2581agcaaaagca accaggaagt tctggagttt gtcaccagtg gaggccggat
ggacccaccc2641aagaactgcc ctgggcctgt ataccggata atgactcagt gctggcaaca
tcagcctgaa2701gacaggccca actttgccat cattttggag aggattgaat actgcaccca

ggacccggat2761gtaatcaaca ccgctttgcc gatagaatat ggtccacttg tggaagagga agagaaagtg2821cctgtgaggc ccaaggaccc tgagggggtt cctcctctcc tggtctctca acaggcaaaa2881cgggaggagg agcgcagccc agctgcccca ccacctctgc ctaccacctc ctctggcaag2941gctgcaaaga aacccacagc tgcagaggtc tctgttcgag tccctagagg gccggccgtg3001gaagggggac acgtgaatat ggcattctct cagtccaacc ctccttcgga gttgcacagg3061gtccacggat ccagaaacaa gcccaccagc ttgtggaacc caacgtacgg ctcctggttt3121acagagaaac ccaccaaaaa gaataatcct atagcaaaga aggagccaca cgagaggggt3181aacctggggc tggagggaag ctgtactgtc ccacctaacg ttgcaactgg gagacttccg3241ggggcctcac tgctcctaga gccctcttcg ctgactgcca atatgaagga ggtacctctg3301ttcaggctac gtcacttccc ttgtgggaat gtcaattacg gctaccagca acagggcttg3361cccttagaag ccgctactgc ccctggagct ggtcattacg aggataccat tctgaaaagc3421aagaatagca tgaaccagcc tgggccctga gctcggtcac acactcactt ctcttccttg3481ggatccctaa gaccgtggag gagagagagg caatcaatgg ctccttcaca aaccagagac3541caaatgtcac gttttgtttt gtgccaacct attttgaagt accaccaaaa aagctgtatt3601ttgaaaatgc tttagaaagg ttttgagcat gggttcatcc tattctttcg aaagaagaaa3661atatcataaa aatgagtgat aaatacaagg cccagatgtg gttgcataag gtttttatgc3721atgtttgttg tatacttcct tatgcttctt ttaaattgtg tgtgctctgc ttcaatgtag3781tcagaattag ctgcttctat gtttcatagt tggggtcata gatgtttcct tgccttgttg3841atgtggacat gagccatttg aggggagagg gaacggaaat aaaggagtta tttgtaatga3901aaaaaaaaaa aaaaaaaaaa aaaaaa

[0092] EML4-ALK突变1蛋白质序列(BAF73611.1;GI:152002653)

[0093] 1mdgfagsldd sisaastsdv qdrlsalesr vqqqedeitv lkaaladvlr rlaisedhva

[0094] 61svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke

[0095] 121kpqgqrekke eshsndqspq iraspspqps sqplqihrqt pesknatptk sikrpspaek181shnswensdd srnklskips tpklipkvtk tadkhkdvii nqegeyikmf mrgrpitmfi241psdvdnyddi rtelppeklk lewaygyrgk dcranvyllp tgeivyfias vvvlfnyeer301tqrhylghtd cvkclaihpd kiriatgqia gvdkdgrplq phvrvwdsvt lstlqiiglg361tfergvgcld fskadsgvhl cviddsnehm ltvwdwqkka kgaeikttne vvlavefhpt421dantiitcgk shiffwtwsg nsltrkqgif gkyekpkfvq claflgngdv ltgdsggvml481iwskttvept pgkgpkvyrr khqelqamqm elqspeykls klrtstimtd ynpnycfagk541tssisdlkev prknitlirg lghgafgevy egqvsgmpnd psplqvavkt lpevcseqde6011dflmealii skfnhqnivr cigvslqslp rfillelmag gdlksflret rprpsqpssl661amldllhvar diacgcqyle enhfihrdia arnclltcpg pgrvakigdf gmardiyras721yyrkggcaml pvkwmppeaf megiftsktd twsfgvllwe ifslgympyp sksnqevlef781vtsggrmdpp kncpgpvyri mtqcwqhqpe drpnfaiile rieyctqdpd vintalpiey841gplveeeekv pvrpkdpegv ppllvsqqak reeerspaap pplpttssgk aakkptaaev901svrvprgpav egghvnmafs qsnppselhr vhgsrnkpts lwnptygswf tekptkknnp961iakkepherg nlglegsctv ppnvatgrlp gaslllepss ltanmkevpl frlrhfpcgn1021vnygyqqqgl pleaatapga ghyedtilks knsmnqpgp

[0096] EML4-ALK突变2cDNA序列(AB275889.1;GI:152002654)

[0097] 1ggcggcgcgg cgcggcgctc gcggctgctg cctgggaggg aggccgggca ggcggctgag

[0098] 61cggcgcggct ctcaacgtga cggggaagtg gttcgggcgg ccgcggctta ctaccccagg

[0099] 121gcgaacggac ggacgacgga ggcgggagcc ggtagccgag ccgggcgacc tagagaacga181gcgggtcagg ctcagcgtcg gccactctgt cggtccgctg aatgaagtgc ccgcccctct241gagcccggag cccggcgctt tccccgcaag atggacggtt tcgccggcag tctcgatgat301agtatttctg ctgcaagtac ttctgatgtt caagatcgcc tgtcagctct tgagtcacga361gttcagcaac aagaagatga aatcactgtg ctaaaggcgg ctttggctga tgttttgagg421cgtcttgcaa tctctgaaga tcatgtggcc tcagtgaaaa aatcagtctc aagtaaaggc481caaccaagcc ctcgagcagt tattcccatg tcctgtataa ccaatggaag tggtgcaaac541agaaaaccaa gtcataccag tgctgtctca attgcaggaa aagaaactct ttcatctgct601gctaaaagtg gtacagaaaa aaagaaagaa aaaccacaag gacagagaga aaaaaaagag661gaatctcatt ctaatgatca aagtccacaa attcgagcat caccttctcc ccagccctct721tcacaacctc tccaaataca cagacaaact ccagaaagca agaatgctac tcccaccaaa781agcataaaac gaccatcacc agctgaaaag tcacataatt cttgggaaaa ttcagatgat841agccgtaata aattgtcgaa aataccttca acacccaaat taataccaaa agttaccaaa901actgcagaca agcataaaga tgtcatcatc aaccaagaag gagaatatat taaaatgttt961atgcgcggtc ggccaattac catgttcatt ccttccgatg ttgacaacta tgatgacatc1021agaacggaac tgcctcctga gaagctcaaa ctggagtggg catatggtta tcgaggaaag1081gactgtagag ctaatgttta ccttcttccg accggggaaa tagtttattt cattgcatca1141gtagtagtac tatttaatta tgaggagaga actcagcgac actacctggg ccatacagac1201tgtgtgaaat gccttgctat acatcctgac aaaattagga ttgcaactgg acagatagct1261ggcgtggata aagatggaag gcctctacaa ccccacgtca gagtgtggga ttctgttact1321ctatccacac tgcagattat tggacttggc acttttgagc gtggagtagg atgcctggat1381ttttcaaaag cagattcagg tgttcattta tgtgttattg atgactccaa tgagcatatg1441cttactgtat gggactggca gaagaaagca aaaggagcag aaataaagac aacaaatgaa1501gttgttttgg ctgtggagtt tcacccaaca gatgcaaata ccataattac atgcggtaaa1561tctcatattt tcttctggac ctggagcggc aattcactaa caagaaaaca gggaattttt1621gggaaatatg aaaagccaaa atttgtgcag tgtttagcat tcttggggaa tggagatgtt1681cttactggag actcaggtgg agtcatgctt atatggagca aaactactgt agagcccaca1741cctgggaaag gacctaaagg tgtatatcaa atcagcaaac aaatcaaagc tcatgatggc1801agtgtgttca cactttgtca gatgagaaat gggatgttat taactggagg agggaaagac1861agaaaaataa ttctgtggga tcatgatctg aatcctgaaa gagaaataga ggttcctgat1921cagtatggca caatcagagc tgtagcagaa ggaaaggcag atcaattttt agtaggcaca1981tcacgaaact ttattttacg aggaacattt aatgatggct tccaaataga agtacagggt2041catacagatg agctttgggg tcttgccaca catcccttca aagatttgct cttgacatgt2101gctcaggaca ggcaggtgtg cctgtggaac tcaatggaac acaggctgga atggaccagg2161ctggtagatg aaccaggaca ctgtgcagat tttcatccaa gtggcacagt ggtggccata2221ggaacgcact caggcaggtg gtttgttctg gatgcagaaa ccagagatct

agtttctatc2281cacacagacg ggaatgaaca gctctctgtg atgcgctact caatagatgg
taccttcctg2341gctgtaggat ctcatgacaa ctttatttac ctctatgtag tctctgaaaa
tggaagaaaa2401tatagcagat atggaaggtg cactggacat tccagctaca tcacacacct
tgactggtcc2461ccagacaaca agtatataat gtctaactcg ggagactatg aaatattgta
cttgtaccgc2521cggaagcacc aggagctgca agccatgcag atggagctgc agagccctga
gtacaagctg2581agcaagctcc gcacctcgac catcatgacc gactacaacc ccaactactg
ctttgctggc2641aagacctcct ccatcagtga cctgaaggag gtgccgcgga aaaacatcac
cctcattcgg2701ggtctgggcc atggagcctt tggggaggtg tatgaaggcc aggtgtccgg
aatgcccaac2761gacccaagcc ccctgcaagt ggctgtgaag acgctgcctg aagtgtgctc
tgaacaggac2821gaactggatt tcctcatgga agccctgatc atcagcaaat tcaaccacca
gaacattgtt2881cgctgcattg gggtgagcct gcaatccctg ccccggttca tcctgctgga
gctcatggcg2941gggggagacc tcaagtcctt cctccgagag acccgccctc gcccgagcca
gccctcctcc3001ctggccatgc tggaccttct gcacgtggct cgggacattg cctgtggctg
tcagtatttg3061gaggaaaacc acttcatcca ccgagacatt gctgccagaa actgcctctt
gacctgtcca3121ggccctggaa gagtggccaa gattggagac ttcgggatgg cccgagacat
ctacagggcg3181agctactata gaaagggagg ctgtgccatg ctgccagtta agtggatgcc
cccagaggcc3241ttcatggaag gaatattcac ttctaaaaca gacacatggt cctttggagt
gctgctatgg3301gaaatctttt ctcttggata tatgccatac cccagcaaaa gcaaccagga
agttctggag3361tttgtcacca gtggaggccg gatggaccca cccaagaact gccctgggcc
tgtataccgg3421ataatgactc agtgctggca acatcagcct gaagacaggc ccaactttgc
catcattttg3481gagaggattg aatactgcac ccaggacccg gatgtaatca acaccgcttt
gccgatagaa3541tatggtccac ttgtggaaga ggaagagaaa gtgcctgtga ggcccaagga
ccctgagggg3601gttcctcctc tcctggtctc tcaacaggca aaacgggagg aggagcgcag
cccagctgcc3661ccaccacctc tgcctaccac ctcctctggc aaggctgcaa agaaacccac
agctgcagag3721gtctctgttc gagtccctag agggccggcc gtggaagggg gacacgtgaa
tatggcattc3781tctcagtcca accctccttc ggagttgcac agggtccacg gatccagaaa
caagcccacc3841agcttgtgga acccaacgta cggctcctgg tttacagaga aacccaccaa
aaagaataat3901cctatagcaa agaaggagcc acacgagagg ggtaacctgg ggctggaggg
aagctgtact3961gtcccaccta acgttgcaac tgggagactt ccgggggcct cactgctcct
agagccctct4021tcgctgactg ccaatatgaa ggaggtacct ctgttcaggc tacgtcactt
cccttgtggg4081aatgtcaatt acggctacca gcaacagggc ttgcccttag aagccgctac
tgcccctgga4141gctggtcatt acgaggatac cattctgaaa agcaagaata gcatgaacca
gcctgggccc4201tgagctcggt cacacactca cttctcttcc ttgggatccc taagaccgtg
gaggagagag4261aggcaatcaa tggctccttc acaaaccaga gaccaaatgt cacgttttgt
tttgtgccaa4321cctattttga agtaccacca aaaaagctgt attttgaaaa tgctttagaa
aggttttgag4381catgggttca tcctattctt tcgaaagaag aaaatatcat aaaaatgagt
gataaataca4441aggcccagat gtggttgcat aaggttttta tgcatgtttg ttgtatactt
ccttatgctt4501cttttaaatt gtgtgtgctc tgcttcaatg tagtcagaat tagctgcttc
tatgtttcat4561agttggggtc atagatgttt ccttgccttg ttgatgtgga catgagccat

ttgaggggag4621agggaacgga aataaaggag ttatttgtaa tgaaaaaaaa aaaaaaaaaa aaaaaaaaa

[0100] EML4-ALK突变2蛋白质序列(BAF73612.1;GI:152002655)

[0101] 1mdgfagsldd sisaastsdv qdrlsalesr vqqqedeitv lkaaladvlr rlaisedhva

[0102] 61svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke

[0103] 121kpqgqrekke eshsndqspq iraspspqps sqplqihrqt pesknatptk sikrpspaek181shnswensdd srnklskips tpklipkvtk tadkhkdvii nqegeyikmf mrgrpitmfi241psdvdnyddi rtelppeklk lewaygyrgk dcranvyllp tgeivyfias vvvlfnyeer301tqrhylghtd cvkclaihpd kiriatgqia gvdkdgrplq phvrvwdsvt lstlqiiglg361tfergvgcld fskadsgvhl cviddsnehm ltvwdwqkka kgaeikttne vvlavefhpt421dantiitcgk shiffwtwsg nsltrkqgif gkyekpkfvq claflgngdv ltgdsggvml481iwskttvept pgkgpkgvyq iskqikahdg svftlcqmrn gmlltgggkd rkiilwdhdl541npereievpd qygtiravae gkadqflvgt srnfilrgtf ndgfqievqg htdelwglat601hpfkdllltc aqdrqvclwn smehrlewtr lvdepghcad fhpsgtvvai gthsgrwfvl661daetrdlvsi htdgneqlsv mrysidgtfl avgshdnfiy lyvvsengrk ysrygrctgh721ssyithldws pdnkyimsns gdyeilylyr rkhqelqamq melqspeykl sklrtstimt781dynpnycfag ktssisdlke vprknitlir glghgafgev yegqvsgmpn dpsplqvavk841tlpevcseqd eldflmeali iskfnhqniv rcigvslqsl prfillelma ggdlksflre901trprpsqpss lamldllhva rdiacgcqyl eenhfihrdi aarnclltcp gpgrvakigd961fgmardiyra syyrkggcam lpvkwmppea fmegiftskt dtwsfgvllw eifslgympy1021psksnqevle fvtsggrmdp pkncpgpvyr imtqcwqhqp edrpnfaiil erieyctqdp1081dvintalpie ygplveeeek vpvrpkdpeg vppllvsqqa kreeerspaa ppplpttssg1141kaakkptaae vsvrvprgpa vegghvnmaf sqsnppselh rvhgsrnkpt slwnptygsw1201ftekptkknn piakkepher gnlglegsct vppnvatgrl pgaslleps sltanmkevp1261lfrlrhfpcg nvnygyqqqg lpleaatapg aghyedtilk sknsmnqpgp

[0104] EML4-ALK突变3a核苷酸序列(AB374361.1;GI:194072592)

[0105] 1actctgtcgg tccgctgaat gaagtgcccg cccctctaag cccggagccc ggcgctttcc

[0106] 61ccgcaagatg gacggtttcg ccggcagtct cgatgatagt atttctgctg caagtacttc

[0107] 121tgatgttcaa gatcgcctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat181cactgtgcta aaggcggctt tggctgatgt tttgaggcgt cttgcaatct ctgaagatca241tgtggcctca gtgaaaaaat cagtctcaag taaaggccaa ccaagccctc gagcagttat301tcccatgtcc tgtataacca atggaagtgg tgcaaacaga aaaccaagtc ataccagtgc361tgtctcaatt gcaggaaaag aaactctttc atctgctgct aaaagtggta cagaaaaaaa421gaaagaaaaa ccacaaggac agagagaaaa aaaagaggaa tctcattcta atgatcaaag481tccacaaatt cgagcatcac cttctcccca gccctcttca caacctctcc aaatacacag541acaaactcca gaaagcaaga atgctactcc caccaaaagc ataaaacgac catcaccagc601tgaaaagtca cataattctt gggaaaattc agatgatagc cgtaataaat tgtcgaaaat661accttcaaca cccaaattaa taccaaaagt taccaaaact gcagacaagc ataaagatgt721catcatcaac caagtgtacc gccggaagca ccaggagctg caagccatgc

agatggagct781gcagagccct gagtacaagc tgagcaagct ccgcacctcg accatcatga ccgactacaa841ccccaactac tgctttgctg gcaagacctc ctccatcagt gacctgaagg aggtgccgcg901gaaaaacatc accctcattc ggggtctggg ccatggagcc tttggggagg tgtatgaagg961ccaggtgtcc ggaatgccca acgacccaag ccccctgcaa gtggctgtga agacgctgcc1021tgaagtgtgc tctgaacagg acgaactgga tttcctcatg gaagccctga tcatcagcaa1081attcaaccac cagaacattg ttcgctgcat tggggtgagc ctgcaatccc tgccccggtt1141catcctgctg gagctcatgg cgggggggaga cctcaagtcc ttcctccgag agacccgccc1201tcgcccgagc cagccctcct ccctggccat gctggacctt ctgcacgtgg ctcgggacat1261tgcctgtggc tgtcagtatt tggaggaaaa ccacttcatc caccgagaca ttgctgccag1321aaactgcctc ttgacctgtc caggccctgg aagagtggcc aagattggag acttcgggat1381ggcccgagac atctacaggg cgagctacta tagaaaggga ggctgtgcca tgctgccagt1441taagtggatg cccccagagg ccttcatgga aggaatattc acttctaaaa cagacacatg1501gtcctttgga gtgctgctat gggaaatctt ttctcttgga tatatgccat accccagcaa1561aagcaaccag gaagttctgg agtttgtcac cagtggaggc cggatggacc cacccaagaa1621ctgccctggg cctgtatacc ggataatgac tcagtgctgg caacatcagc ctgaagacag1681gcccaacttt gccatcattt tggagaggat tgaatactgc acccaggacc cggatgtaat1741caacaccgct ttgccgatag aatatggtcc acttgtggaa gaggaagaga aagtgcctgt1801gaggcccaag gaccctgagg gggttcctcc tctcctggtc tctcaacagg caaaacggga1861ggaggagcgc agcccagctg ccccaccacc tctgcctacc acctcctctg gcaaggctgc1921aaagaaaccc acagctgcag aggtctctgt tcgagtccct agagggccgg ccgtggaagg1981gggacacgtg aatatggcat tctctcagtc caaccctcct tcggagttgc acagggtcca2041cggatccaga aacaagccca ccagcttgtg gaacccaacg tacggctcct ggtttacaga2101gaaacccacc aaaaagaata atcctatagc aaagaaggag ccacacgaga ggggtaacct2161ggggctggag ggaagctgta ctgtcccacc taacgttgca actgggagac ttccgggggc2221ctcactgctc ctagagccct cttcgctgac tgccaatatg aaggaggtac ctctgttcag2281gctacgtcac ttcccttgtg ggaatgtcaa ttacggctac cagcaacagg gcttgcccct2341agaagccgct actgcccctg gagctggtca ttacgaggat accattctga aaagcaagaa2401tagcatgaac cagcctgggc cctgagctcg gtcgcacact cacttctctt ccttgggatc2461cctaagaccg tgg

[0108] EML4-ALK突变3a蛋白质序列(BAG55003.1;GI:194072593)

[0109] 1mdgfagsldd sisaastsdv qdrlsalesr vqqqedeitv lkaaladvlr rlaisedhva

[0110] 61svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke

[0111] 121kpqgqrekke eshsndqspq iraspspqps sqplqihrqt pesknatptk sikrpspaek181shnswensdd srnklskips tpklipkvtk tadkhkdvii nqvyrrkhqe lqamqmelqs241peyklsklrt stimtdynpn ycfagktssi sdlkevprkn itlirglghg afgevyegqv301sgmpndpspl qvavktlpev cseqdeldfl mealiiskfn hqnivrcigv slqslprfil361lelmaggdlk sflretrprp sqpsslamld llhvardiac gcqyleenhf ihrdiaarnc421lltcpgpgrv akigdfgmar diyrasyyrk ggcamlpvkw mppeafmegi ftsktdtwsf481gvllweifsl gympypsksn qevlefvtsg grmdppkncp gpvyrimtqc

wqhqpedrpn541faiileriey ctqdpdvint alpieygplv eeeekvpvrp kdpegvppll vsqqakreee601rspaappplp ttssgkaakk ptaaevsvrv prgpaveggh vnmafsqsnp pselhrvhgs661rnkptslwnp tygswftekp tkknnpiakk ephergnlgl egsctvppnv atgrlpgasl721llepssltan mkevplfrlr hfpcgnvnyg yqqqglplea atapgaghye dtilksknsm781nqpgp

[0112] EML4-ALK突变3b核苷酸序列(AB374362.1;GI:194072594)

[0113] 1actctgtcgg tccgctgaat gaagtgcccg cccctctaag cccggagccc ggcgctttcc

[0114] 61ccgcaagatg gacggtttcg ccggcagtct cgatgatagt atttctgctg caagtacttc

[0115] 121tgatgttcaa gatcgcctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat181cactgtgcta aaggcggctt tggctgatgt tttgaggcgt cttgcaatct ctgaagatca241tgtggcctca gtgaaaaaat cagtctcaag taaaggccaa ccaagccctc gagcagttat301tcccatgtcc tgtataacca atggaagtgg tgcaaacaga aaaccaagtc ataccagtgc361tgtctcaatt gcaggaaaag aaactctttc atctgctgct aaaagtggta cagaaaaaaa421gaaagaaaaa ccacaaggac agagagaaaa aaaagaggaa tctcattcta atgatcaaag481tccacaaatt cgagcatcac cttctcccca gccctcttca caacctctcc aaatacacag541acaaactcca gaaagcaaga atgctactcc caccaaaagc ataaaacgac catcaccagc601tgaaaagtca cataattctt gggaaaattc agatgatagc cgtaataaat tgtcgaaaat661accttcaaca cccaaattaa taccaaaagt taccaaaact gcagacaagc ataaagatgt721catcatcaac caagcaaaaa tgtcaactcg cgaaaaaaac agccaagtgt accgccggaa781gcaccaggag ctgcaagcca tgcagatgga gctgcagagc cctgagtaca agctgagcaa841gctccgcacc tcgaccatca tgaccgacta caaccccaac tactgctttg ctggcaagac901ctcctccatc agtgacctga aggaggtgcc gcggaaaaac atcaccctca ttcggggtct961gggccatgga gcctttgggg aggtgtatga aggccaggtg tccggaatgc ccaacgaccc1021aagccccctg caagtggctg tgaagacgct gcctgaagtg tgctctgaac aggacgaact1081ggatttcctc atggaagccc tgatcatcag caaattcaac caccagaaca ttgttcgctg1141cattggggtg agcctgcaat ccctgccccg gttcatcctg ctggagctca tggcgggggg1201agacctcaag tccttcctcc gagagacccg ccctcgcccg agccagccct cctccctggc1261catgctggac cttctgcacg tggctcggga cattgcctgt ggctgtcagt atttggagga1321aaaccacttc atccaccgag acattgctgc cagaaactgc ctcttgacct gtccaggccc1381tggaagagtg gccaagattg gagacttcgg gatggcccga gacatctaca gggcgagcta1441ctatagaaag ggaggctgtg ccatgctgcc agttaagtgg atgccccccag aggccttcat1501ggaaggaata ttcacttcta aaacagacac atggtccttt ggagtgctgc tatgggaaat1561cttttctctt ggatatatgc cataccccag caaaagcaac caggaagttc tggagtttgt1621caccagtgga ggccggatgg acccacccaa gaactgccct gggcctgtat accggataat1681gactcagtgc tggcaacatc agcctgaaga caggcccaac tttgccatca ttttggagag1741gattgaatac tgcacccagg acccggatgt aatcaacacc gctttgccga tagaatatgg1801tccacttgtg gaagaggaag agaaagtgcc tgtgaggccc aaggaccctg aggggttcc1861tcctctcctg gtctctcaac aggcaaaacg ggaggaggag cgcagcccag ctgccccacc1921acctctgcct accacctcct ctggcaaggc tgcaaagaaa cccacagctg

cagaggtctc1981tgttcgagtc cctagagggc cggccgtgga agggggacac gtgaatatgg cattctctca2041gtccaaccct ccttcggagt tgcacagggt ccacggatcc agaaacaagc ccaccagctt2101gtggaaccca acgtacggct cctggtttac agagaaaccc accaaaaaga ataatcctat2161agcaaagaag gagccacacg agaggggtaa cctggggctg gagggaagct gtactgtccc2221acctaacgtt gcaactggga gacttccggg ggcctcactg ctcctagagc cctcttcgct2281gactgccaat atgaaggagg tacctctgtt caggctacgt cacttccctt gtgggaatgt2341caattacggc taccagcaac agggcttgcc cttagaagcc gctactgccc ctggagctgg2401tcattacgag gataccattc tgaaaagcaa gaatagcatg aaccagcctg ggccctgagc2461tcggtcgcac actcacttct cttccttggg atccctaaga ccgtgg

[0116] EML4-ALK突变3b蛋白质序列(BAG55004.1;GI:194072595)

[0117] 1mdgfagsldd sisaastsdv qdrlsalesr vqqqedeitv lkaaladvlr rlaisedhva

[0118] 61svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke

[0119] 121kpqgqrekke eshsndqspq iraspspqps sqplqihrqt pesknatptk sikrpspaek181shnswensdd srnklskips tpklipkvtk tadkhkdvii nqakmstrek nsqvyrrkhq241elqamqmelq speyklsklr tstimtdynp nycfagktss isdlkevprk nitlirglgh301gafgevyegq vsgmpndpsp lqvavktlpe vcseqdeldf lmealiiskf nhqnivrcig361vslqslprfi llelmaggdl ksflretrpr psqpsslaml dllhvardia cgcqyleenh421fihrdiaarn clltcpgpgr vakigdfgma rdiyrasyyr kggcamlpvk wmppeafmeg481iftsktdtws fgvllweifs lgympypsks nqevlefvts ggrmdppknc pgpvyrimtq541cwqhqpedrp nfaiilerie yctqdpdvin talpieygpl veeeekvpvr pkdpegvppl601lvsqqakree erspaapppl pttssgkaak kptaaevsvr vprgpavegg hvnmafsqsn661ppselhrvhg srnkptslwn ptygswftek ptkknnpiak kephergnlg legsctvppn721vatgrlpgas lllepsslta nmkevplfrl rhfpcgnvny gyqqqglple aatapgaghy781edtilkskns mnqpgp

[0120] EML4-ALK突变4核苷酸序列(AB374363.1;GI:209837703)

[0121] 1actctgtcgg tccgctgaat gaagtgcccg cccctctaag cccggagccc ggcgctttcc

[0122] 61ccgcaagatg gacggtttcg ccggcagtct cgatgatagt atttctgctg caagtacttc

[0123] 121tgatgttcaa gatcgcctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat181cactgtgcta aaggcggctt tggctgatgt tttgaggcgt cttgcaatct ctgaagatca241tgtggcctca gtgaaaaaat cagtctcaag taaaggccaa ccaagccctc gagcagttat301tcccatgtcc tgtataacca atggaagtgg tgcaaacaga aaaccaagtc ataccagtgc361tgtctcaatt gcaggaaaag aaactctttc atctgctgct aaaagtggta cagaaaaaaa421gaaagaaaaa ccacaaggac agagagaaaa aaaagaggaa tctcattcta atgatcaaag481tccacaaatt cgagcatcac cttctcccca gccctcttca caacctctcc aaatacacag541acaaactcca gaaagcaaga atgctactcc caccaaaagc ataaaacgac catcaccagc601tgaaaagtca cataattctt gggaaaattc agatgatagc cgtaataaat tgtcgaaaat661accttcaaca cccaaattaa taccaaaagt taccaaaact gcagacaagc ataaagatgt721catcatcaac caagaaggag aatatattaa aatgtttatg cgcggtcggc caattaccat781gttcattcct tccgatgttg acaactatga tgacatcaga acggaactgc

ctcctgagaa841gctcaaactg gagtgggcat atggttatcg aggaaaggac tgtagagcta
atgtttacct901tcttccgacc ggggaaatag tttatttcat tgcatcagta gtagtactat
ttaattatga961ggagagaact cagcgacact acctgggcca tacagactgt gtgaaatgcc
ttgctataca1021tcctgacaaa attaggattg caactggaca gatagctggc gtggataaag
atggaaggcc1081tctacaaccc cacgtcagag tgtgggattc tgttactcta tccacactgc
agattattgg1141acttggcact tttgagcgtg gagtaggatg cctggatttt tcaaaagcag
attcaggtgt1201tcatttatgt gttattgatg actccaatga gcatatgctt actgtatggg
actggcagag1261gaaagcaaaa ggagcagaaa taaagacaac aaatgaagtt gttttggctg
tggagtttca1321cccaacagat gcaaatacca taattacatg cggtaaatct catattttct
tctggacctg1381gagcggcaat tcactaacaa gaaaacaggg aatttttggg aaatatgaaa
agccaaaatt1441tgtgcagtgt ttagcattct tggggaatgg agatgttctt actggagact
caggtggagt1501catgcttata tggagcaaaa ctactgtaga gcccacacct gggaaaggac
ctaaaggtgt1561atatcaaatc agcaaacaaa tcaaagctca tgatggcagt gtgttcacac
tttgtcagat1621gagaaatggg atgttattaa ctggaggagg gaaagacaga aaaataattc
tgtgggatca1681tgatctgaat cctgaaagag aaatagagat atgctggatg agccctgagt
acaagctgag1741caagctccgc acctcgacca tcatgaccga ctacaacccc aactactgct
ttgctggcaa1801gacctcctcc atcagtgacc tgaaggaggt gccgcggaaa aacatcaccc
tcattcgggg1861tctgggccat ggagcctttg gggaggtgta tgaaggccag gtgtccggaa
tgcccaacga1921cccaagcccc ctgcaagtgg ctgtgaagac gctgcctgaa gtgtgctctg
aacaggacga1981actggatttc ctcatggaag ccctgatcat cagcaaattc aaccaccaga
acattgttcg2041ctgcattggg gtgagcctgc aatccctgcc ccggttcatc ctgctggagc
tcatggcggg2101gggagacctc aagtccttcc tccgagagac ccgccctcgc ccgagccagc
cctcctccct2161ggccatgctg gaccttctgc acgtggctcg ggacattgcc tgtggctgtc
agtatttgga2221ggaaaaccac ttcatccacc gagacattgc tgccagaaac tgcctcttga
cctgtccagg2281ccctggaaga gtggccaaga ttggagactt cgggatggcc cgagacatct
acagggcgag2341ctactataga aagggaggct gtgccatgct gccagttaag tggatgcccc
cagaggcctt2401catggaagga atattcactt ctaaaacaga cacatggtcc tttggagtgc
tgctatggga2461aatcttttct cttggatata tgccataccc cagcaaaagc aaccaagaag
ttctggagtt2521tgtcaccagt ggaggccgga tggacccacc caagaactgc cctgggcctg
tataccggat2581aatgactcag tgctggcaac atcagcctga agacaggccc aactttgcca
tcattttgga2641gaggattgaa tactgcaccc aggacccgga tgtaatcaac accgctttgc
cgatagaata2701tggtccactt gtggaagagg aagagaaagt gcctgtgagg cccaaggacc
ctgagggggt2761tcctcctctc ctggtctctc aacaggcaaa acgggaggag gagcgcagcc
cagctgcccc2821accacctctg cctaccacct cctctggcaa ggctgcaaag aaacccacag
ctgcagaggt2881ctctgttcga gtccctagag ggccggccgt ggaaggggga cacgtgaata
tggcattctc2941tcagtccaac cctccttcgg agttgcacag ggtccacgga tccagaaaca
agcccaccag3001cttgtggaac ccaacgtacg gctcctggtt tacagagaaa cccaccaaaa
agaataatcc3061tatagcaaag aaggagccac acgagagggg taacctgggg ctggagggaa
gctgtactgt3121cccacctaac gttgcaactg ggagacttcc gggggcctca ctgctcctag

agccctcttc3181gctgactgcc aatatgaagg aggtacctct gttcaggcta cgtcacttcc cttgtgggaa3241tgtcaattac ggctaccagc aacagggctt gcccttagaa gccgctactg cccctggagc3301tggtcattac gaggatacca ttctgaaaag caagaatagc atgaaccagc ctgggcccctg3361agctcggtcg cacactcact tctcttcctt gggatcccta agaccgtgg

[0124] EML4-ALK突变4蛋白质序列(BAG75147.1;GI:209837704)

[0125] 1mdgfagsldd sisaastsdv qdrlsalesr vqqqedeitv lkaaladvlr rlaisedhva

[0126] 61svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke

[0127] 121kpqgqrekke eshsndqspq iraspspqps sqplqihrqt pesknatptk sikrpspaek181shnswensdd srnklskips tpklipkvtk tadkhkdvii nqegeyikmf mrgrpitmfi241psdvdnyddi rtelppeklk lewaygyrgk dcranvyllp tgeivyfias vvvlfnyeer301tqrhylghtd cvkclaihpd kiriatgqia gvdkdgrplq phvrvwdsvt lstlqiiglg361tfergvgcld fskadsgvhl cviddsnehm ltvwdwqrka kgaeikttne vvlavefhpt421dantiitcgk shiffwtwsg nsltrkqgif gkyekpkfvq claflgngdv ltgdsggvml481iwskttvept pgkgpkgvyq iskqikahdg svftlcqmrn gmlltgggkd rkiilwdhdl541npereieicw mspeyklskl rtstimtdyn pnycfagkts sisdlkevpr knitlirglg601hgafgevyeg qvsgmpndps plqvavktlp evcseqdeld flmealiisk fnhqnivrci661gvslqslprf illelmaggd lksflretrp rpsqpsslam ldllhvardi acgcqyleen721hfihrdiaar nclltcpgpg rvakigdfgm ardiyrasyy rkggcamlpv kwmppeafme781giftsktdtw sfgvllweif slgympypsk snqevlefvt sggrmdppkn cpgpvyrimt841qcwqhqpedr pnfaiileri eyctqdpdvi ntalpieygp lveeeekvpv rpkdpegvpp901llvsqqakre eerspaappp lpttssgkaa kkptaaevsv rvprgpaveg ghvnmafsqs961nppselhrvh gsrnkptslw nptygswfte kptkknnpia kkephergnl glegsctvpp1021nvatgrlpga slllepsslt anmkevplfr lrhfpcgnvn ygyqqqglpl eaatapgagh1081yedtilkskn smnqpgp

[0128] EML4-ALK突变5a核苷酸序列(AB374364.1;GI:209837705)

[0129] 1actctgtcgg tccgctgaat gaagtgcccg cccctctaag cccggagccc ggcgctttcc

[0130] 61ccgcaagatg gacggtttcg ccggcagtct cgatgatagt atttctgctg caagtacttc

[0131] 121tgatgttcaa gatcgcctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat181cactgtgcta aaggcggctt tggctgatgt tttgaggcgt cttgcaatct ctgaagatca241tgtggcctca gtgaaaaaat cagtctcaag taaagtgtac cgccggaagc accaggagct301gcaagccatg cagatggagc tgcagagccc tgagtacaag ctgagcaagc tccgcacctc361gaccatcatg accgactaca accccaacta ctgctttgct ggcaagacct cctccatcag421tgacctgaag gaggtgccgc ggaaaaacat caccctcatt cggggtctgg gccatggagc481ctttggggag gtgtatgaag gccaggtgtc cggaatgccc aacgacccaa gccccctgca541agtggctgtg aagacgctgc ctgaagtgtg ctctgaacag gacgaactgg atttcctcat601ggaagccctg atcatcagca aattcaacca ccagaacatt gttcgctgca ttggggtgag661cctgcaatcc ctgccccggt tcatcctgct ggagctcatg gcgggggggag acctcaagtc721cttcctccga gagacccgcc ctcgcccgag ccagccctcc tccctggcca tgctggacct781tctgcacgtg gctcgggaca ttgcctgtgg ctgtcagtat ttggaggaaa

accacttcat841ccaccgagac attgctgcca gaaactgcct cttgacctgt ccaggccctg gaagagtggc901caagattgga gacttcggga tggcccgaga catctacagg gcgagctact atagaaaggg961aggctgtgcc atgctgccag ttaagtggat gcccccagag gccttcatgg aaggaatatt1021cacttctaaa acagacacat ggtcctttgg agtgctgcta tgggaaatct tttctcttgg1081atatatgcca taccccagca aaagcaacca ggaagttctg gagtttgtca ccagtggagg1141ccggatggac ccacccaaga actgccctgg gcctgtatac cggataatga ctcagtgctg1201gcaacatcag cctgaagaca ggcccaactt tgccatcatt ttggagagga ttgaatactg1261cacccaggac ccggatgtaa tcaacaccgc tttgccgata gaatatggtc cacttgtgga1321agaggaagag aaagtgcctg tgaggcccaa ggaccctgag ggggttcctc ctctcctggt1381ctctcaacag gcaaaacggg aggaggagcg cagcccagct gccccaccac ctctgcctac1441cacctcctct ggcaaggctg caaagaaacc cacagctgca gaggtctctg ttcgagtccc1501tagagggccg gccgtggaag ggggacacgt gaatatggca ttctctcagt ccaaccctcc1561ttcggagttg cacagggtcc acggatccag aaacaagccc accagcttgt ggaacccaac1621gtacggctcc tggtttacag agaaacccac caaaaagaat aatcctatag caaagaagga1681gccacacgag aggggtaacc tggggctgga gggaagctgt actgtcccac ctaacgttgc1741aactgggaga cttccggggg cctcactgct cctagagccc tcttcgctga ctgccaatat1801gaaggaggta cctctgttca ggctacgtca cttcccttgt gggaatgtca attacggcta1861ccagcaacag ggcttgccct tagaagccgc tactgcccct ggagctggtc attacgagga1921taccattctg aaaagcaaga atagcatgaa ccagcctggg ccctgagctc ggtcgcacac1981tcacttctct tccttgggat ccctaagacc gtgg

[0132] EML4-ALK突变5a蛋白质序列(BAG75148.1;GI:209837706)

[0133] 1mdgfagsldd sisaastsdv qdrlsalesr vqqqedeitv lkaaladvlr rlaisedhva

[0134] 61svkksvsskv yrrkhqelqa mqmelqspey klsklrtsti mtdynpnycf agktssisdl

[0135] 121kevprknitl irglghgafg evyegqvsgm pndpsplqva vktlpevcse qdeldflmea181liiskfnhqn ivrcigvslq slprfillel maggdlksfl retrprpsqp sslamldllh241vardiacgcq yleenhfihr diaarncllt cpgpgrvaki gdfgmardiy rasyyrkggc301amlpvkwmpp eafmegifts ktdtwsfgvl lweifslgym pypsksnqev lefvtsggrm361dppkncpgpv yrimtqcwqh qpedrpnfai ilerieyctq dpdvintalp ieygplveee421ekvpvrpkdp egvppllvsq qakreeersp aappplptts sgkaakkpta aevsvrvprg481pavegghvnm afsqsnppse lhrvhgsrnk ptslwnptyg swftekptkk nnpiakkeph541ergnlglegs ctvppnvatg rlpgasllle pssltanmke vplfrlrhfp cgnvnygyqq601qglpleaata pgaghyedti lksknsmnqp gp

[0136] EML4-ALK突变5b蛋白质序列(AB374365.1;GI:209837707)

[0137] 1actctgtcgg tccgctgaat gaagtgcccg cccctctaag cccggagccc ggcgctttcc

[0138] 61ccgcaagatg gacggtttcg ccggcagtct cgatgatagt atttctgctg caagtacttc

[0139] 121tgatgttcaa gatcgcctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat181cactgtgcta aaggcggctt tggctgatgt tttgaggcgt cttgcaatct ctgaagatca241tgtggcctca gtgaaaaaat cagtctcaag taaaggttca gagctcaggg gaggatatgg301agatccaggg aggcttcctg taggaagtgg cctgtgtagt gcttcaaggg

ccaggctgcc361aggccatgtt gcagctgacc acccacctgc agtgtaccgc cggaagcacc aggagctgca421agccatgcag atggagctgc agagccctga gtacaagctg agcaagctcc gcacctcgac481catcatgacc gactacaacc ccaactactg ctttgctggc aagacctcct ccatcagtga541cctgaaggag gtgccgcgga aaaacatcac cctcattcgg ggtctgggcc atggagcctt601tggggaggtg tatgaaggcc aggtgtccgg aatgcccaac gacccaagcc ccctgcaagt661ggctgtgaag acgctgcctg aagtgtgctc tgaacaggac gaactggatt tcctcatgga721agccctgatc atcagcaaat tcaaccacca gaacattgtt cgctgcattg gggtgagcct781gcaatccctg ccccggttca tcctgctgga gctcatggcg gggggagacc tcaagtcctt841cctccgagag acccgccctc gcccgagcca gccctcctcc ctggccatgc tggaccttct901gcacgtggct cgggacattg cctgtggctg tcagtatttg gaggaaaacc acttcatcca961ccgagacatt gctgccagaa actgcctctt gacctgtcca ggccctggaa gagtggccaa1021gattggagac ttcgggatgg cccgagacat ctacagggcg agctactata gaaagggagg1081ctgtgccatg ctgccagtta agtggatgcc cccagaggcc ttcatggaag gaatattcac1141ttctaaaaca gacacatggt cctttggagt gctgctatgg gaaatctttt ctcttggata1201tatgccatac cccagcaaaa gcaaccagga agttctggag tttgtcacca gtggaggccg1261gatggaccca cccaagaact gccctgggcc tgtataccgg ataatgactc agtgctggca1321acatcagcct gaagacaggc ccaactttgc catcattttg gagaggattg aatactgcac1381ccaggacccg gatgtaatca acaccgcttt gccgatagaa tatggtccac ttgtggaaga1441ggaagagaaa gtgcctgtga ggcccaagga ccctgagggg gttcctcctc tcctggtctc1501tcaacaggca aaacgggagg aggagcgcag cccagctgcc ccaccacctc tgcctaccac1561ctcctctggc aaggctgcaa agaaacccac agctgcagag gtctctgttc gagtccctag1621agggccggcc gtggaagggg gacacgtgaa tatggcattc tctcagtcca accctccttc1681ggagttgcac agggtccacg gatccagaaa caagcccacc agcttgtgga acccaacgta1741cggctcctgg tttacagaga aacccaccaa aaagaataat cctatagcaa agaaggagcc1801acacgagagg ggtaacctgg ggctggaggg aagctgtact gtcccaccta acgttgcaac1861tgggagactt ccgggggcct cactgctcct agagccctct tcgctgactg ccaatatgaa1921ggaggtacct ctgttcaggc tacgtcactt cccttgtggg aatgtcaatt acggctacca1981gcaacagggc ttgcccttag aagccgctac tgcccctgga gctggtcatt acgaggatac2041cattctgaaa agcaagaata gcatgaacca gcctgggccc tgagctcggt cgcacactca2101cttctcttcc ttgggatccc taagaccgtg g

[0140] EML4-ALK突变5b蛋白质序列(BAG75149.1;GI:209837708)

[0141] 1mdgfagsldd sisaastsdv qdrlsalesr vqqqedeitv lkaaladvlr rlaisedhva

[0142] 61svkksvsskg selrggygdp grlpvgsglc sasrarlpgh vaadhppavy rrkhqelqam

[0143] 121qmelqspeyk lsklrtstim tdynpnycfa gktssisdlk evprknitli rglghgafge181vyegqvsgmp ndpsplqvav ktlpevcseq deldflmeal iiskfnhqni vrcigvslqs2411prfillelm aggdlksflr etrprpsqps slamldllhv ardiacgcqy leenhfihrd301iaarnclltc pgpgrvakig dfgmardiyr asyyrkggca mlpvkwmppe afmegiftsk361tdtwsfgvll weifslgymp ypsksnqevl efvtsggrmd ppkncpgpvy rimtqcwqhq421pedrpnfaii lerieyctqd pdvintalpi eygplveeee kvpvrpkdpe

gvppllvsqq481akreeerspa appplpttss gkaakkptaa evsvrvprgp avegghvnma fsqsnppsel541hrvhgsrnkp tslwnptygs wftekptkkn npiakkephe rgnlglegsc tvppnvatgr6011pgaslllep ssltanmkev plfrlrhfpc gnvnygyqqq glpleaatap gaghyedtil661ksknsmnqpg p

[0144] EML4-ALK突变6核苷酸序列(AB462411.1;GI:227452648)

[0145] 1tactctgtcg gtccgctgaa tgaagtgccc gcccctctaa gcccggagcc cggcgctttc

[0146] 61cccgcaagat ggacggtttc gccggcagtc tcgatgatag tatttctgct gcaagtactt

[0147] 121ctgatgttca agatcgcctg tcagctcttg agtcacgagt tcagcaacaa gaagatgaaa181tcactgtgct aaaggcggct ttggctgatg ttttgaggcg tcttgcaatc tctgaagatc241atgtggcctc agtgaaaaaa tcagtctcaa gtaaaggcca accaagccct cgagcagtta301ttcccatgtc ctgtataacc aatggaagtg gtgcaaacag aaaaccaagt cataccagtg361ctgtctcaat tgcaggaaaa gaaactcttt catctgctgc taaaagtggt acagaaaaaa421agaaagaaaa accacaagga cagagagaaa aaaaagagga atctcattct aatgatcaaa481gtccacaaat tcgagcatca ccttctcccc agccctcttc acaacctctc caaatacaca541gacaaactcc agaaagcaag aatgctactc ccaccaaaag cataaaacga ccatcaccag601ctgaaaagtc acataattct tgggaaaatt cagatgatag ccgtaataaa ttgtcgaaaa661taccttcaac acccaaatta ataccaaaag ttaccaaaac tgcagacaag cataaagatg721tcatcatcaa ccaagaagga gaatatatta aaatgtttat gcgcggtcgg ccaattacca781tgttcattcc ttccgatgtt gacaactatg atgacatcag aacggaactg cctcctgaga841agctcaaact ggagtgggca tatggttatc gaggaaagga ctgtagagct aatgtttacc901ttcttccgac cggggaaata gtttatttca ttgcatcagt agtagtacta tttaattatg961aggagagaac tcagcgacac tacctgggcc atacagactg tgtgaaatgc cttgctatac1021atcctgacaa aattaggatt gcaactggac agatagctgg cgtggataaa gatggaaggc1081ctctacaacc ccacgtcaga gtgtgggatt ctgttactct atccacactg cagattattg1141gacttggcac ttttgagcgt ggagtaggat gcctggattt ttcaaaagca gattcaggtg1201ttcatttatg tgttattgat gactccaatg agcatatgct tactgtatgg gactggcaga1261ggaaagcaaa aggagcagaa ataaagacaa caaatgaagt tgttttggct gtggagtttc1321acccaacaga tgcaaatacc ataattacat gcggtaaatc tcatattttc ttctggacct1381ggagcggcaa ttcactaaca agaaaacagg gaatttttgg gaaatatgaa aagccaaaat1441ttgtgcagtg tttagcattc ttggggaatg gagatgttct tactggagac tcaggtggag1501tcatgcttat atggagcaaa actactgtag agcccacacc tgggaaagga cctaaaggaa1561gtggcctgtg tagtgcttca agggccaggc tgccaggcca tgttgcagct gaccacccac1621ctgcagtgta ccgccggaag caccaggagc tgcaagccat gcagatggag ctgcagagcc1681ctgagtacaa gctgagcaag ctccgcacct cgaccatcat gaccgactac aaccccaact1741actgctttgc tggcaagacc tcctccatca gtgacctgaa ggaggtgccg cggaaaaaca1801tcaccctcat tcggggtctg ggccatggag cctttgggga ggtgtatgaa ggccaggtgt1861ccggaatgcc caacgaccca agccccctgc aagtggctgt gaagacgctg cctgaagtgt1921gctctgaaca ggacgaactg gatttcctca tggaagccct gatcatcagc aaattcaacc1981accagaacat tgttcgctgc attggggtga gcctgcaatc cctgccccgg

ttcatcctgc2041tggagctcat ggcgggggga gacctcaagt ccttcctccg agagacccgc cctcgcccga2101gccagccctc ctccctggcc atgctggacc ttctgcacgt ggctcgggac attgcctgtg2161gctgtcagta tttggaggaa aaccacttca tccaccgaga cattgctgcc agaaactgcc2221tcttgacctg tccaggccct ggaagagtgg ccaagattgg agacttcggg atggcccgag2281acatctacag ggcgagctac tatagaaagg gaggctgtgc catgctgcca gttaagtgga2341tgccccccaga ggccttcatg gaaggaatat tcacttctaa aacagacaca tggtcctttg2401gagtgctgct atgggaaatc ttttctcttg gatatatgcc ataccccagc aaaagcaacc2461aggaagttct ggagtttgtc accagtggag gccggatgga cccacccaag aactgccctg2521ggcctgtata ccggataatg actcagtgct ggcaacatca gcctgaagac aggcccaact2581ttgccatcat tttggagagg attgaatact gcacccagga cccggatgta atcaacaccg2641ctttgccgat agaatatggt ccacttgtgg aagaggaaga gaaagtgcct gtgaggccca2701aggaccctga gggggttcct cctctcctgg tctctcaaca ggcaaaacgg gaggaggagc2761gcagcccagc tgccccacca cctctgccta ccacctcctc tggcaaggct gcaaagaaac2821ccacagctgc agaggtctct gttcgagtcc ctagagggcc ggccgtggaa gggggacacg2881tgaatatggc attctctcag tccaaccctc cttcggagtt gcacagggtc cacggatcca2941gaaacaagcc caccagcttg tggaacccaa cgtacggctc ctggtttaca gagaaaccca3001ccaaaaagaa taatcctata gcaaagaagg agccacacga gaggggtaac ctggggctgg3061agggaagctg tactgtccca cctaacgttg caactgggag acttccgggg gcctcactgc3121tcctagagcc ctcttcgctg actgccaata tgaaggaggt acctctgttc aggctacgtc3181acttcccttg tgggaatgtc aattacggct accagcaaca gggcttgccc ttagaagccg3241ctactgcccc tggagctggt cattacgagg ataccattct gaaaagcaag aatagcatga3301accagcctgg gccctgagct cggtcgcaca ctcacttctc ttccttggga tccctaagac3361cgtgg

[0148] EML4-ALK突变6蛋白质序列(BAH57335.1;GI:227452649)

[0149] 1mdgfagsldd sisaastsdv qdrlsalesr vqqqedeitv lkaaladvlr rlaisedhva

[0150] 61svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke

[0151] 121kpqgqrekke eshsndqspq iraspspqps sqplqihrqt pesknatptk sikrpspaek181shnswensdd srnklskips tpklipkvtk tadkhkdvii nqegeyikmf mrgrpitmfi241psdvdnyddi rtelppeklk lewaygyrgk dcranvyllp tgeivyfias vvvlfnyeer301tqrhylghtd cvkclaihpd kiriatgqia gvdkdgrplq phvrvwdsvt lstlqiiglg361tfergvgcld fskadsgvhl cviddsnehm ltvwdwqrka kgaeikttne vvlavefhpt421dantiitcgk shiffwtwsg nsltrkqgif gkyekpkfvq claflgngdv ltgdsggvml481iwskttvept pgkgpkgsgl csasrarlpg hvaadhppav yrrkhqelqa mqmelqspey541klsklrtsti mtdynpnycf agktssisdl kevprknitl irglghgafg evyegqvsgm601pndpsplqva vktlpevcse qdeldflmea liiskfnhqn ivrcigvslq slprfillel661maggdlksfl retrprpsqp sslamldllh vardiacgcq yleenhfihr diaarncllt721cpgpgrvaki gdfgmardiy rasyyrkggc amlpvkwmpp eafmegifts ktdtwsfgvl781lweifslgym pypsksnqev lefvtsggrm dppkncpgpv yrimtqcwqh qpedrpnfai841ilerieyctq dpdvintalp ieygplveee ekvpvrpkdp egvppllvsq

qakreeersp901aappplptts sgkaakkpta aevsvrvprg pavegghvnm afsqsnppse lhrvhgsrnk961ptslwnptyg swftekptkk nnpiakkeph ergnlglegs ctvppnvatg rlpgaslple1021pssltanmke vplfrlrhfp cgnvnygyqq qglpleaata pgaghyedti lksknsmnqp1081gp

[0152] EML4-ALK突变7核苷酸序列(AB462412.1;GI:227452650)

[0153] 1tactctgtcg gtccgctgaa tgaagtgccc gcccctctaa gcccggagcc cggcgctttc

[0154] 61cccgcaagat ggacggtttc gccggcagtc tcgatgatag tatttctgct gcaagtactt

[0155] 121ctgatgttca agatcgcctg tcagctcttg agtcacgagt tcagcaacaa gaagatgaaa181tcactgtgct aaaggcggct ttggctgatg ttttgaggcg tcttgcaatc tctgaagatc241atgtggcctc agtgaaaaaa tcagtctcaa gtaaaggcca accaagccct cgagcagtta301ttcccatgtc ctgtataacc aatggaagtg gtgcaaacag aaaaccaagt cataccagtg361ctgtctcaat tgcaggaaaa gaaactcttt catctgctgc taaaagtggt acagaaaaaa421agaaagaaaa accacaagga cagagagaaa aaaaagagga atctcattct aatgatcaaa481gtccacaaat tcgagcatca ccttctcccc agccctcttc acaacctctc caaatacaca541gacaaactcc agaaagcaag aatgctactc ccaccaaaag cataaaacga ccatcaccag601ctgaaaagtc acataattct tgggaaaatt cagatgatag ccgtaataaa ttgtcgaaaa661taccttcaac acccaaatta ataccaaaag ttaccaaaac tgcagacaag cataaagatg721tcatcatcaa ccaagaagga gaatatatta aaatgtttat gcgcggtcgg ccaattacca781tgttcattcc ttccgatgtt gacaactatg atgacatcag aacggaactg cctcctgaga841agctcaaact ggagtgggca tatggttatc gaggaaagga ctgtagagct aatgtttacc901ttcttccgac cggggaaata gtttatttca ttgcatcagt agtagtacta tttaattatg961aggagagaac tcagcgacac tacctgggcc atacagactg tgtgaaatgc cttgctatac1021atcctgacaa aattaggatt gcaactggac agatagctgg cgtggataaa gatggaaggc1081ctctacaacc ccacgtcaga gtgtgggatt ctgttactct atccacactg cagattattg1141gacttggcac ttttgagcgt ggagtaggat gcctggattt ttcaaaagca gattcaggtg1201ttcatttatg tgttattgat gactccaatg agcatatgct tactgtatgg gactggcaga1261ggaaagcaaa aggagcagaa ataaagacaa caaatgaagt tgttttggct gtggagtttc1321acccaacaga tgcaaatacc ataattacat gcggtaaatc tcatattttc ttctggacct1381ggagcggcaa ttcactaaca agaaaacagg gaatttttgg gaaatatgaa aagccaaaat1441ttgtgcagtg tttagcattc ttggggaatg gagatgttct tactggagac tcaggtggag1501tcatgcttat atggagcaaa actactgtag agcccacacc tgggaaagga cctaaaggtg1561tatatcaaat cagcaaacaa atcaaagctc atgatggcag tgtgttcaca ctttgtcaga1621tgagaaatgg gatgttatta actggaggag ggaaagacag aaaaataatt ctgtgggatc1681atgatctgaa tcctgaaaga gaaatagagc accaggagct gcaagccatg cagatggagc1741tgcagagccc tgagtacaag ctgagcaagc tccgcacctc gaccatcatg accgactaca1801accccaacta ctgctttgct ggcaagacct cctccatcag tgacctgaag gaggtgccgc1861ggaaaaacat caccctcatt cggggtctgg gccatggagc ctttggggag gtgtatgaag1921gccaggtgtc cggaatgccc aacgacccaa gccccctgca agtggctgtg aagacgctgc1981ctgaagtgtg ctctgaacag gacgaactgg atttcctcat ggaagccctg

atcatcagca2041aattcaacca ccagaacatt gttcgctgca ttggggtgag cctgcaatcc ctgccccggt2101tcatcctgct ggagctcatg gcgggggggag acctcaagtc cttcctccga gagacccgcc2161ctcgcccgag ccagccctcc tccctggcca tgctggacct tctgcacgtg gctcgggaca2221ttgcctgtgg ctgtcagtat ttggaggaaa accacttcat ccaccgagac attgctgcca2281gaaactgcct cttgacctgt ccaggccctg gaagagtggc caagattgga gacttcggga2341tggcccgaga catctacagg gcgagctact atagaaaggg aggctgtgcc atgctgccag2401ttaagtggat gcccccagag gccttcatgg aaggaatatt cacttctaaa acagacacat2461ggtcctttgg agtgctgcta tgggaaatct tttctcttgg atatatgcca taccccagca2521aaagcaacca ggaagttctg gagtttgtca ccagtggagg ccggatggac ccacccaaga2581actgccctgg gcctgtatac cggataatga ctcagtgctg gcaacatcag cctgaagaca2641ggcccaactt tgccatcatt ttggagagga ttgaatactg cacccaggac ccggatgtaa2701tcaacaccgc tttgccgata gaatatggtc cacttgtgga agaggaagag aaagtgcctg2761tgaggcccaa ggaccctgag ggggttcctc ctctcctggt ctctcaacag gcaaaacggg2821aggaggagcg cagcccagct gccccaccac ctctgcctac cacctcctct ggcaaggctg2881caaagaaacc cacagctgca gaggtctctg ttcgagtccc tagagggccg gccgtggaag2941ggggacacgt gaatatggca ttctctcagt ccaaccctcc ttcggagttg cacaaggtcc3001acggatccag aaacaagccc accagcttgt ggaacccaac gtacggctcc tggtttacag3061agaaacccac caaaaagaat aatcctatag caaagaagga gccacacgac aggggtaacc3121tggggctgga gggaagctgt actgtcccac ctaacgttgc aactgggaga cttccggggg3181cctcactgct cctagagccc tcttcgctga ctgccaatat gaaggaggta cctctgttca3241ggctacgtca cttcccttgt gggaatgtca attacggcta ccagcaacag ggcttgccct3301tagaagccgc tactgccccc ggagctggtc attacgagga taccattctg aaaagcaaga3361atagcatgaa ccagcctggg ccctgagctc ggtcgcacac tcacttctct tccttgggat3421ccctaagacc gtgga

[0156] EML4-ALK突变7蛋白质序列(BAH57336.1;GI:227452651)

[0157] 1mdgfagsldd sisaastsdv qdrlsalesr vqqqedeitv lkaaladvlr rlaisedhva

[0158] 61svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke

[0159] 121kpqgqrekke eshsndqspq iraspspqps sqplqihrqt pesknatptk sikrpspaek181shnswensdd srnklskips tpklipkvtk tadkhkdvii nqegeyikmf mrgrpitmfi241psdvdnyddi rtelppeklk lewaygyrgk dcranvyllp tgeivyfias vvvlfnyeer301tqrhylghtd cvkclaihpd kiriatgqia gvdkdgrplq phvrvwdsvt lstlqiiglg361tfergvgcld fskadsgvhl cviddsnehm ltvwdwqrka kgaeikttne vvlavefhpt421dantiitcgk shiffwtwsg nsltrkqgif gkyekpkfvq claflgngdv ltgdsggvml481iwskttvept pgkgpkgvyq iskqikahdg svftlcqmrn gmlltgggkd rkiilwdhdl541npereiehqe lqamqmelqs peyklsklrt stimtdynpn ycfagktssi sdlkevprkn601itlirglghg afgevyegqv sgmpndpspl qvavktlpev cseqdeldfl mealiiskfn661hqnivrcigv slqslprfil lelmaggdlk sflretrprp sqpsslamld llhvardiac721gcqyleenhf ihrdiaarnc lltcpgpgrv akigdfgmar diyrasyyrk ggcamlpvkw781mppeafmegi ftsktdtwsf gvllweifsl gympypsksn qevlefvtsg

grmdppkncp841gpvyrimtqc wqhqpedrpn faiileriey ctqdpdvint alpieygplv eeeekvpvrp901kdpegvppll vsqqakreee rspaappplp ttssgkaakk ptaaevsvrv prgpaveggh961vnmafsqsnp pselhkvhgs rnkptslwnp tygswftekp tkknnpiakk ephdrgnlgl1021egsctvppnv atgrlpgasl llepssltan mkevplfrlr hfpcgnvnyg yqqqglplea1081atapgaghye dtilksknsm nqpgp

[0160] KIF5B-ALK核苷酸序列(AB462413.1;GI:227452652)

[0161] 1tgcgagaaag atggcggacc tggccgagtg caacatcaaa gtgatgtgtc gcttcagacc

[0162] 61tctcaacgag tctgaagtga accgcggcga caagtacatc gccaagtttc agggagaaga

[0163] 121cacggtcgtg atcgcgtcca agccttatgc atttgatcgg gtgttccagt caagcacatc181tcaagagcaa gtgtataatg actgtgcaaa gaagattgtt aaagatgtac ttgaaggata241taatggaaca atatttgcat atggacaaac atcctctggg aagacacaca caatggaggg301taaacttcat gatccagaag gcatgggaat tattccaaga atagtgcaag atatttttaa361ttatatttac tccatggatg aaaatttgga atttcatatt aaggtttcat attttgaaat421atatttggat aagataaggg acctgttaga tgtttcaaag accaaccttt cagttcatga481agacaaaaac cgagttccct atgtaaaggg gtgcacagag cgttttgtat gtagtccaga541tgaagttatg gataccatag atgaaggaaa atccaacaga catgtagcag ttacaaatat601gaatgaacat agctctagga gtcacagtat atttcttatt aatgtcaaac aagagaacac661acaaacggaa caaaagctga gtggaaaact ttatctggtt gatttagctg gtagtgaaaa721ggttagtaaa actggagctg aaggtgctgt gctggatgaa gctaaaaaca tcaacaagtc781actttctgct cttggaaatg ttatttctgc tttggctgag ggtagtacat atgttccata841tcgagatagt aaaatgacaa gaatccttca agattcatta ggtggcaact gtagaaccac901tattgtaatt tgctgctctc catcatcata caatgagtct gaaacaaaat ctacactctt961atttggccaa agggccaaaa caattaagaa cacagtttgt gtcaatgtgg agttaactgc1021agaacagtgg aaaaagaagt atgaaaaaga aaaagaaaaa aataagatcc tgcggaacac1081tattcagtgg cttgaaaatg agctcaacag atggcgtaat ggggagacgg tgcctattga1141tgaacagttt gacaaagaga aagccaactt ggaagctttc acagtggata aagatattac1201tcttaccaat gataaaccag caaccgcaat tggagttata ggaaatttta ctgatgctga1261aagaagaaag tgtgaagaag aaattgctaa attatacaaa cagcttgatg acaaggatga1321agaaattaac cagcaaagtc aactggtaga gaaactgaag acgcaaatgt tggatcagga1381ggagcttttg gcatctacca gaagggatca agacaatatg caagctgagc tgaatcgcct1441tcaagcagaa aatgatgcct ctaaagaaga agtgaaagaa gttttacagg ccctagaaga1501acttgctgtc aattatgatc agaagtctca ggaagttgaa gacaaaacta aggaatatga1561attgcttagt gatgaattga atcagaaatc ggcaacttta gcgagtatag atgctgagct1621tcagaaactt aaggaaatga ccaaccacca gaaaaaacga gcagctgaga tgatggcatc1681tttactaaaa gaccttgcag aaataggaat tgctgtggga aataatgatg taaagcagcc1741tgagggaact ggcatgatag atgaagagtt cactgttgca agactctaca ttagcaaaat1801gaagtcagaa gtaaaaacca tggtgaaacg ttgcaagcag ttagaaagca cacaaactga1861gagcaacaaa aaaatggaag aaaatgaaaa ggagttagca gcatgtcagc ttcgtatctc1921tcaacatgaa gccaaaatca agtcattgac tgaatacctt caaaatgtgg

aacaaaagaa1981aagacagttg gaggaatctg tcgatgccct cagtgaagaa ctagtccagc
ttcgagcaca2041agagaaagtc catgaaatgg aaaaggagca cttaaataag gttcagactg
caaatgaagt2101taagcaagct gttgaacagc agatccagag ccatagagaa actcatcaaa
aacagatcag2161tagtttgaga gatgaagtag aagcaaaagc aaaacttatt actgatcttc
aagaccaaaa2221ccagaaaatg atgttagagc aggaacgtct aagagtagaa catgagaagt
tgaaagccac2281agatcaggaa aagagcagaa aactacatga acttacggtt atgcaagata
gacgagaaca2341agcaagacaa gacttgaagg gtttggaaga gacagtggca aaagaacttc
agactttaca2401caacctgcgc aaactctttg ttcaggacct ggctacaaga gttaaaaaga
gtgctgagat2461tgattctgat gacaccggag gcagcgctgc tcagaagcaa aaaatctcct
ttcttgaaaa2521taatcttgaa cagctcacta aagtgcacaa acagttggta cgtgataatg
cagatctccg2581ctgtgaactt cctaagttgg aaaagcgact tcgagctaca gctgagagag
tgaaagcttt2641ggaatcagca ctgaaagaag ctaaagaaaa tgcatctcgt gatcgcaaac
gctatcagca2701agaagtagat cgcataaagg aagcagtcag gtcaaagaat atggccagaa
gagggcattc2761tgcacagatt gtgtaccgcc ggaagcacca ggagctgcaa gccatgcaga
tggagctgca2821gagccctgag tacaagctga gcaagctccg cacctcgacc atcatgaccg
actacaaccc2881caactactgc tttgctggca agacctcctc catcagtgac ctgaaggagg
tgccgcggaa2941aaacatcacc ctcattcggg gtctgggcca tggcgccttt ggggaggtgt
atgaaggcca3001ggtgtccgga atgcccaacg acccaagccc cctgcaagtg gctgtgaaga
cgctgcctga3061agtgtgctct gaacaggacg aactggattt cctcatggaa gccctgatca
tcagcaaatt3121caaccaccag aacattgttc gctgcattgg ggtgagcctg caatccctgc
cccggttcat3181cctgctggag ctcatggcgg ggggagacct caagtccttc ctccgagaga
cccgccctcg3241cccgagccag ccctcctccc tggccatgct ggaccttctg cacgtggctc
gggacattgc3301ctgtggctgt cagtatttgg aggaaaacca cttcatccac cgagacattg
ctgccagaaa3361ctgcctcttg acctgtccag gccctggaag agtggccaag attggagact
tcgggatggc3421ccgagacatc tacagggcga gctactatag aaagggaggc tgtgccatgc
tgccagttaa3481gtggatgccc ccagaggcct tcatggaagg aatattcact tctaaaacag
acacatggtc3541ctttggagtg ctgctatggg aaatcttttc tcttggatat atgccatacc
ccagcaaaag3601caaccaggaa gttctggagt ttgtcaccag tggaggccgg atggacccac
ccaagaactg3661ccctgggcct gtataccgga taatgactca gtgctggcaa catcagcctg
aagacaggcc3721caactttgcc atcattttgg agaggattga atactgcacc caggacccgg
atgtaatcaa3781caccgctttg ccgatagaat atggtccact tgtggaagag gaagagaaag
tgcctgtgag3841gcccaaggac cctgaggggg ttcctcctct cctggtctct caacaggcaa
aacgggagga3901ggagcgcagc ccagctgccc caccacctct gcctaccacc tcctctggca
aggctgcaaa3961gaaacccaca gctgcagagg tctctgttcg agtccctaga gggccggccg
tggaaggggg4021acacgtgaat atggcattct ctcagtccaa ccctccttcg gagttgcaca
aggtccacgg4081atccagaaac aagcccacca gcttgtggaa cccaacgtac ggctcctggt
ttacagagaa4141acccaccaaa aagaataatc ctatagcaaa gaaggagcca cacgacaggg
gtaacctggg4201gctggaggga agctgtactg tcccacctaa cgttgcaact gggagacttc
cgggggcctc4261actgctccta gagccctctt cgctgactgc caatatgaag gaggtacctc

tgttcaggct4321acgtcacttc ccttgtggga atgtcaatta cggctaccag caacagggct tgcccttaga4381agccgctact gcccctggag ctggtcatta cgaggatacc attctgaaaa gcaagaatag4441catgaaccag cctgggccct gagctcggtc gcacactca

[0164] KIF5B-ALK蛋白质序列(BAH57337.1;GI:227452653)

[0165] 1madlaecnik vmcrfrplne sevnrgdkyi akfqgedtvv iaskpyafdr vfqsstsqeq

[0166] 61vyndcakkiv kdvlegyngt ifaygqtssg kthtmegklh dpegmgiipr ivqdifnyiy

[0167] 121smdenlefhi kvsyfeiyld kirdlldvsk tnlsvhedkn rvpyvkgcte rfvcspdevm181dtidegksnr hvavtnmneh ssrshsifli nvkqentqte qklsgklylv dlagsekvsk241tgaegavlde akninkslsa lgnvisalae gstyvpyrds kmtrilqdsl ggncrttivi301ccspssynes etkstllfgq raktikntvc vnveltaeqw kkkyekekek nkilrntiqw361lenelnrwrn getvpideqf dkekanleaf tvdkditltn dkpataigvi gnftdaerrk421ceeeiaklyk qlddkdeein qqsqlveklk tqmldqeell astrrdqdnm qaelnrlqae481ndaskeevke vlqaleelav nydqksqeve dktkeyells delnqksatl asidaelqkl541kemtnhqkkr aaemmasllk dlaeigiavg nndvkqpegt gmideeftva rlyiskmkse601vktmvkrckq lestqtesnk kmeenekela acqlrisqhe akikslteyl qnveqkkrql661eesvdalsee lvqlraqekv hemekehlnk vqtanevkqa veqqiqshre thqkqisslr721deveakakli tdlqdqnqkm mleqerlrve heklkatdqe ksrklheltv mqdrreqarq781dlkgleetva kelqtlhnlr klfvqdlatr vkksaeidsd dtggsaaqkq kisflennle841qltkvhkqlv rdnadlrcel pklekrlrat aervkalesa lkeakenasr drkryqqevd901rikeavrskn marrghsaqi vyrrkhqelq amqmelqspe yklsklrtst imtdynpnyc961fagktssisd lkevprknit lirglghgaf gevyegqvsg mpndpsplqv avktlpevcs1021eqdeldflme aliiskfnhq nivrcigvsl qslprfille lmaggdlksf lretrprpsq1081psslamldll hvardiacgc qyleenhfih rdiaarncll tcpgpgrvak igdfgmardi1141yrasyyrkgg camlpvkwmp peafmegift sktdtwsfgv llweifslgy mpypsksnqe1201vlefvtsggr mdppkncpgp vyrimtqcwq hqpedrpnfa iilerieyct qdpdvintal1261pieygplvee eekvpvrpkd pegvppllvs qqakreeers paappplptt ssgkaakkpt1321aaevsvrvpr gpavegghvn mafsqsnpps elhkvhgsrn kptslwnpty gswftekptk1381knnpiakkep hdrgnlgleg sctvppnvat grlpgaslll epssltanmk evplfrlrhf1441pcgnvnygyq qqglpleaat apgaghyedt ilksknsmnq pgp

[0168] NPM-ALK序列(t(2;5)(p23;q35染色体易位)*

[0169] TPM3-ALK序列(t(1;2)(p25;p23)染色体易位)*

[0170] TFGXL-ALK核苷酸序列(AF390893.1;GI:20269389)

[0171] 1atgaacggac agttggatct aagtgggaag ctaatcatca aagctcaact tggggaggat

[0172] 61attcggcgaa ttcctattca taatgaagat attacttatg atgaattagt gctaatgatg

[0173] 121caacgagttt tcagaggaaa acttctgagt aatgatgaag taacaataaa gtataaagat181gaagatggag atcttataac aatttttgat agttctgacc tttcctttgc aattcagtgc241agtaggatac tgaaactgac attatttgtt aatggccagc caagacccct tgaatcaagt301caggtgaaat atctccgtcg agaactgata gaacttcgaa ataaagtgaa tcgtttattg361gatagcttgg aaccacctgg agaaccagga ccttccacca atattcctga

aaatgatact421gtggatggta gggaagaaaa gtctgcttct gattcttctg gaaaacagtc
tactcaggtt481atggcagcaa gtatgtctgc ttttgatcct ttaaaaaacc aagatgaaat
caataaaaat541gttatgtcag cgtttggctt aacagatgat caggtttcag ggccacccag
tgctcctgca601gaagatcgtt caggaacacc cgacagcatt gcttcctcct cctcagcagc
tcacccacca661ggcgttcagc cacagcagcc accatataca ggagctcaga ctcaagcagg
tcagattgaa721gtgtaccgcc ggaagcacca ggagctgcaa gccatgcaga tggagctgca
gagccctgag781tacaagctga gcaagctccg cacctcgacc atcatgaccg actacaaccc
caactactgc841tttgctggca agacctcctc catcagtgac ctgaaggagg tgccgcggaa
aaacatcacc901ctcattcggg gtctgggcca tggcgccttt ggggaggtgt atgaaggcca
ggtgtccgga961atgcccaacg acccaagccc cctgcaagtg gctgtgaaga cgctgcctga
agtgtgctct1021gaacaggacg aactggattt cctcatggaa gccctgatca tcagcaaatt
caaccaccag1081aacattgttc gctgcattgg ggtgagcctg caatccctgc cccggttcat
cctgctggag1141ctcatggcgg ggggagacct caagtccttc ctccgagaga cccgccctcg
cccgagccag1201ccctcctccc tggccatgct ggaccttctg cacgtggctc gggacattgc
ctgtggctgt1261cagtatttgg aggaaaacca cttcatccac cgagacattg ctgccagaaa
ctgcctcttg1321acctgtccag gccctggaag agtggccaag attggagact tcgggatggc
ccgagacatc1381tacagggcga gctactatag aaagggaggc tgtgccatgc tgccagttaa
gtggatgccc1441ccagaggcct tcatggaagg aatattcact tctaaaacag acacatggtc
ctttggagtg1501ctgctatggg aaatcttttc tcttggatat atgccatacc ccagcaaaag
caaccaggaa1561gttctggagt ttgtcaccag tggaggccgg atggacccac ccaagaactg
ccctgggcct1621gtataccgga taatgactca gtgctggcaa catcagcctg aagacaggcc
caactttgcc1681atcattttgg agaggattga atactgcacc caggacccgg atgtaatcaa
caccgctttg1741ccgatagaat atggtccact tgtggaagag gaagagaaag tgcctgtgag
gcccaaggac1801cctgaggggg ttcctcctct cctggtctct caacaggcaa aacgggagga
ggagcgcagc1861ccagctgccc caccacctct gcctaccacc tcctctggca aggctgcaaa
gaaacccaca1921gctgcagagg tctctgttcg agtccctaga gggccggccg tggaaggggg
acacgtgaat1981atggcattct ctcagtccaa ccctccttcg gagttgcaca aggtccacgg
atccagaaac2041aagcccacca gcttgtggaa cccaacgtac ggctcctggt ttacagagaa
acccaccaaa2101aagaataatc ctatagcaaa gaaggagcca cacgacaggg gtaacctggg
gctggaggga2161agctgtactg tcccacctaa cgttgcaact gggagacttc cgggggcctc
actgctccta2221gagccctctt cgctgactgc caatatgaag gaggtacctc tgttcaggct
acgtcacttc2281ccttgtggga atgtcaatta cggctaccag caacagggct tgcccttaga
agccgctact2341gcccctggag ctggtcatta cgaggatacc attctgaaaa gcaagaatag
catgaaccag2401cctgggccct ga

[0174] TFGXL-ALK蛋白质序列(AAM17922.1;GI:20269390)*

[0175] 1mngqldlsgk liikaqlged irripihned itydelvlmm qrvfrgklls ndevtikykd

[0176] 61edgdlitifd ssdlsfaiqc srilkltlfv ngqprpless qvkylrreli elrnkvnrll

[0177] 121dsleppgepg pstnipendt vdgreeksas dssgkqstqv maasmsafdp lknqdeinkn181vmsafgltdd qvsgppsapa edrsgtpdsi assssaahpp gvqpqqppyt

gaqtqagqie241vyrrkhqelq amqmelqspe yklsklrtst imtdynpnyc fagktssisd
lkevprknit301lirglghgaf gevyegqvsg mpndpsplqv avktlpevcs eqdeldflme
aliiskfnhq361nivrcigvsl qslprfille lmaggdlksf lretrprpsq psslamldll
hvardiacgc421qyleenhfih rdiaarncll tcpgpgrvak igdfgmardi yrasyyrkgg
camlpvkwmp481peafmegift sktdtwsfgv llweifslgy mpypsksnqe vlefvtsggr
mdppkncpgp541vyrimtqcwq hqpedrpnfa iilerieyct qdpdvintal pieygplvee
eekvpvrpkd601pegvppllvs qqakreeers paappplptt ssgkaakkpt aaevsvrvpr
gpavegghvn661mafsqsnpps elhkvhgsrn kptslwnpty gswftekptk knnpiakkep
hdrgnlgleg721sctvppnvat grlpgaslll epssltanmk evplfrlrhf pcgnvnygyq
qqglpleaat781apgaghyedt ilksknsmnq pgp

[0178] TFGL-ALK核苷酸序列(AF143407.1;GI:6739534)

[0179] 1cctccgcaag ccgtctttct ctagagttgt atatatagaa catcctggag tccaccatga

[0180] 61acggacagtt ggatctaagt gggaagctaa tcatcaaagc tcaacttggg gaggatattc

[0181] 121ggcgaattcc tattcataat gaagatatta cttatgatga attagtgcta atgatgcaac181gagttttcag aggaaaactt ctgagtaatg atgaagtaac aataaagtat aaagatgaag241atggagatct tataacaatt tttgatagtt ctgacctttc ctttgcaatt cagtgcagta301ggatactgaa actgacatta tttgttaatg gccagccaag acccccttgaa tcaagtcagg361tgaaatatct ccgtcgagaa ctgatagaac ttcgaaataa agtgaatcgt ttattggata421gcttggaacc acctggagaa ccaggacctt ccaccaatat tcctgaaaat gatactgtgg481atggtaggga agaaaagtct gcttctgatt cttctggaaa acagtctact caggttatgg541cagcaagtat gtctgctttt gatcctttaa aaaaccaaga tgaaatcaat aaaaatgtta601tgtcagcgtt tggcttaaca gatgatcagg tttcagtgta ccgccggaag caccaggagc661tgcaagccat gcagatggag ctgcagagcc ctgagtacaa gctgagcaag ctccgcacct721cgaccatcat gaccgactac aaccccaact actgctttgc tggcaagacc tcctccatca781gtgacctgaa ggaggtgccg cggaaaaaca tcaccctcat tcggggtctg ggccatggcg841cctttgggga ggtgtatgaa ggccaggtgt ccggaatgcc caacgaccca agccccctgc901aagtggctgt gaagacgctg cctgaagtgt gctctgaaca ggacgaactg gatttcctca961tggaagccct gatcatcagc aaattcaacc accagaacat tgttcgctgc attggggtga1021gcctgcaatc cctgccccgg ttcatcctgc tggagctcat ggcgggggga gacctcaagt1081ccttcctccg agagacccgc cctcgcccga gccagccctc ctccctggcc atgctggacc1141ttctgcacgt ggctcgggac attgcctgtg gctgtcagta tttggaggaa aaccacttca1201tccaccgaga cattgctgcc agaaactgcc tcttgacctg tccaggccct ggaagagtgg1261ccaagattgg agacttcggg atggcccgag acatctacag ggcgagctac tatagaaagg1321gaggctgtgc catgctgcca gttaagtgga tgcccccaga ggccttcatg gaaggaatat1381tcacttctaa aacagacaca tggtcctttg gagtgctgct atgggaaatc ttttctcttg1441gatatatgcc ataccccagc aaaagcaacc aggaagttct ggagtttgtc accagtggag1501gccggatgga cccacccaag aactgccctg ggcctgtata ccggataatg actcagtgct1561ggcaacatca gcctgaagac aggcccaact ttgccatcat tttggagagg attgaatact1621gcacccagga cccggatgta atcaacaccg ctttgccgat agaatatggt

ccacttgtgg1681aagaggaaga gaaagtgcct gtgaggccca aggaccctga gggggttcct cctctcctgg1741tctctcaaca ggcaaaacgg gaggaggagc gcagcccagc tgccccacca cctctgccta1801ccacctcctc tggcaaggct gcaaagaaac ccacagctgc agaggtctct gttcgagtcc1861ctagagggcc ggccgtggaa gggggacacg tgaatatggc attctctcag tccaaccctc1921cttcggagtt gcacaaggtc cacggatcca gaaacaagcc caccagcttg tggaacccaa1981cgtacggctc ctggtttaca gagaaaccca ccaaaaagaa taatcctata gcaaagaagg2041agccacacga caggggtaac ctggggctgg agggaagctg tactgtccca cctaacgttg2101caactgggag acttccgggg gcctcactgc tcctagagcc ctcttcgctg actgccaata2161tgaaggaggt acctctgttc aggctacgtc acttcccttg tgggaatgtc aattacggct2221accagcaaca gggcttgccc ttagaagccg ctactgcccc tggagctggt cattacgagg2281ataccattct gaaaagcaag aatagcatga accagcctgg gccctgagct cggtcgcaca2341ctcacttctc ttccttggga tccctaagac cgtggaggag agagaggcaa tggctccttc2401acaaaccaga gaccaaatgt cacgttttgt tttgtgccaa cctattttga agtaccacca2461aaaaagctgt attttgaaaa tgctttagaa aggttttgag catgggttca tcctattctt2521tcgaaagaag aaaatatcat aaaaatgagt gataaataca aggcccagat gtggttgcat2581aaggttttta tgcatgtttg ttgtatactt ccttatgctt cttttaaatt gtgtgtgctc2641tgcttcaatg tagtcagaat tagctgcttc tatgtttcat agttggggtc atagatgttt2701ccttgccttg ttgatgtgga catgagccat ttgaggggag agggaacgga aataaaggag2761ttatttgtaa tgactaaaa

[0182] TFGL-ALK蛋白质序列(AAF27292.1;GI:6739535)*

[0183] 1mngqldlsgk liikaqlged irripihned itydelvlmm qrvfrgklls ndevtikykd

[0184] 61edgdlitifd ssdlsfaiqc srilkltlfv ngqprpless qvkylrreli elrnkvnrll121dsleppgepg pstnipendt vdgreeksas dssgkqstqv maasmsafdp lknqdeinkn181vmsafgltdd qvsvyrrkhq elqamqmelq speyklsklr tstimtdynp nycfagktss241isdlkevprk nitlirglgh gafgevyegq vsgmpndpsp lqvavktlpe vcseqdeldf3011mealiiskf nhqnivrcig vslqslprfi llelmaggdl ksflretrpr psqpsslaml361dllhvardia cgcqyleenh fihrdiaarn clltcpgpgr vakigdfgma rdiyrasyyr421kggcamlpvk wmppeafmeg iftsktdtws fgvllweifs lgympypsks nqevlefvts481ggrmdppknc pgpvyrimtq cwqhqpedrp nfaiilerie yctqdpdvin talpieygpl541veeeekvpvr pkdpegvppl lvsqqakree erspaapppl pttssgkaak kptaaevsvr601vprgpavegg hvnmafsqsn ppselhkvhg srnkptslwn ptygswftek ptkknnpiak661kephdrgnlg legsctvppn vatgrlpgas lllepsslta nmkevplfrl rhfpcgnvny721gyqqqglple aatapgaghy edtilkskns mnqpgp

[0185] TFGS-ALK核苷酸序列(AF125093.1;GI:7229260)

[0186] 1cctccgcaag ccgtctttct ctagagttgt atatatagaa catcctggag tccaccatga

[0187] 61acggacagtt ggatctaagt gggaagctaa tcatcaaagc tcaacttggg gaggatattc121ggcgaattcc tattcataat gaagatatta cttatgatga attagtgcta atgatgcaac181gagttttcag aggaaaactt ctgagtaatg atgaagtaac aataaagtat aaagatgaag241atggagatct tataacaatt tttgatagtt ctgacctttc ctttgcaatt

cagtgcagta301ggatactgaa actgacatta tttgttaatg gccagccaag acccccttgaa
tcaagtcagg361tgaaatatct ccgtcgagaa ctgatagaac ttcgaaataa agtgaatcgt
ttattggata421gcttggaacc acctggagaa ccaggacctt ccaccaatat tcctgaaaat
gtgtaccgcc481ggaagcacca ggagctgcaa gccatgcaga tggagctgca gagccctgag
tacaagctga541gcaagctccg cacctcgacc atcatgaccg actacaaccc caactactgc
tttgctggca601agacctcctc catcagtgac ctgaaggagg tgccgcggaa aaacatcacc
ctcattcggg661gtctgggcca tggcgccttt ggggaggtgt atgaaggcca ggtgtccgga
atgcccaacg721acccaagccc cctgcaagtg gctgtgaaga cgctgcctga agtgtgctct
gaacaggacg781aactggattt cctcatggaa gccctgatca tcagcaaatt caaccaccag
aacattgttc841gctgcattgg ggtgagcctg caatccctgc cccggttcat cctgctggag
ctcatggcgg901ggggagacct caagtccttc ctccgagaga cccgccctcg cccgagccag
ccctcctccc961tggccatgct ggaccttctg cacgtggctc gggacattgc ctgtggctgt
cagtatttgg1021aggaaaacca cttcatccac cgagacattg ctgccagaaa ctgcctcttg
acctgtccag1081gccctggaag agtggccaag attggagact tcgggatggc ccgagacatc
tacagggcga1141gctactatag aaagggaggc tgtgccatgc tgccagttaa gtggatgccc
ccagaggcct1201tcatggaagg aatattcact tctaaaacag acacatggtc ctttggagtg
ctgctatggg1261aaatcttttc tcttggatat atgccatacc ccagcaaaag caaccaggaa
gttctggagt1321ttgtcaccag tggaggccgg atggacccac ccaagaactg ccctgggcct
gtataccgga1381taatgactca gtgctggcaa catcagcctg aagacaggcc caactttgcc
atcattttgg1441agaggattga atactgcacc caggacccgg atgtaatcaa caccgctttg
ccgatagaat1501atggtccact tgtggaagag gaagagaaag tgcctgtgag gcccaaggac
cctgaggggg1561ttcctcctct cctggtctct caacaggcaa aacgggagga ggagcgcagc
ccagctgccc1621caccacctct gcctaccacc tcctctggca aggctgcaaa gaaacccaca
gctgcagagg1681tctctgttcg agtccctaga gggccggccg tggaaggggg acacgtgaat
atggcattct1741ctcagtccaa ccctccttcg gagttgcaca aggtccacgg atccagaaac
aagcccacca1801gcttgtggaa cccaacgtac ggctcctggt ttacagagaa acccaccaaa
aagaataatc1861ctatagcaaa gaaggagcca cacgacaggg gtaacctggg gctggaggga
agctgtactg1921tcccacctaa cgttgcaact gggagacttc cgggggcctc actgctccta
gagccctctt1981cgctgactgc caatatgaag gaggtacctc tgttcaggct acgtcacttc
ccttgtggga2041atgtcaatta cggctaccag caacagggct tgcccttaga agccgctact
gcccctggag2101ctggtcatta cgaggatacc attctgaaaa gcaagaatag catgaaccag
cctgggccct2161gagctcggtc gcacactcac ttctcttcct tgggatccct aagaccgtgg
aggagagaga2221ggcaatggct ccttcacaaa ccagagacca aatgtcacgt tttgttttgt
gccaacctat2281tttgaagtac caccaaaaaa gctgtatttt gaaaatgctt tagaaaggtt
ttgagcatgg2341gttcatccta ttctttcgaa agaagaaaat atcataaaaa tgagtgataa
atacaaggcc2401cagatgtggt tgcataaggt ttttatgcat gtttgttgta tacttcctta
tgcttctttt2461aaattgtgtg tgctctgctt caatgtagtc agaattagct gcttctatgt
ttcatagttg2521gggtcataga tgtttccttg ccttgttgat gtggacatga gccatttgag
gggagaggga2581acggaaataa aggagttatt tgtaatgact aaaa

[0188] TFGS-ALK蛋白质序列(AAF42734.1;GI:7229261)*

[0189] 1mngqldlsgk liikaqlged irripihned itydelvlmm qrvfrgklls ndevtikykd

[0190] 61edgdlitifd ssdlsfaiqc srilkltlfv ngqprpless qvkylrreli elrnkvnrll

[0191] 121dsleppgepg pstnipenvy rrkhqelqam qmelqspeyk lsklrtstim tdynpnycfa181gktssisdlk evprknitli rglghgafge vyegqvsgmp ndpsplqvav ktlpevcseq241deldflmeal iiskfnhqni vrcigvslqs lprfillelm aggdlksflr etrprpsqps301slamldllhv ardiacgcqy leenhfihrd iaarnclltc pgpgrvakig dfgmardiyr361asyyrkggca mlpvkwmppe afmegiftsk tdtwsfgvll weifslgymp ypsksnqevl421efvtsggrmd ppkncpgpvy rimtqcwqhq pedrpnfaii lerieyctqd pdvintalpi481eygplveeee kvpvrpkdpe gvppllvsqq akreeerspa appplpttss gkaakkptaa541evsvrvprgp avegghvnma fsqsnppsel hkvhgsrnkp tslwnptygs wftekptkkn601npiakkephd rgnlglegsc tvppnvatgr lpgaslllep ssltanmkev plfrlrhfpc661gnvnygyqqq glpleaatap gaghyedtil ksknsmnqpg p

[0192] ATIC-ALK序列(inv(2)(p23;q35)染色体易位)*

[0193] CLTC-ALK序列(t(2;17)(p23;q23)染色体易位)*

[0194] MSN-ALK核苷酸序列(AF295356.1;GI:14625823)

[0195] 1aactccgctg cctttgccgc caccatgccc aaaacgatca gtgtgcgtgt gaccaccatg

[0196] 61gatgcagagc tggagtttgc catccagccc aacaccaccg ggaagcagct atttgaccag

[0197] 121gtggtgaaaa ctattggctt gagggaagtt tggttctttg gtctgcagta ccaggacact181aaaggtttct ccacctggct gaaactcaat aagaaggtga ctgcccagga tgtgcggaag241gaaagccccc tgctctttaa gttccgtgcc aagttctacc ctgaggatgt gtccgaggaa301ttgattcagg acatcactca gcgcctgttc tttctgcaag tgaaagaggg cattctcaat361gatgatattt actgcccgcc tgagaccgct gtgctgctgg cctcgtatgc tgtccagtct421aagtatggcg acttcaataa ggaagtgcat aagtctggct acctggccgg agacaagttg481ctcccgcaga gagtcctgga acagcacaaa ctcaacaagg accagtggga ggagcggatc541caggtgtggc atgaggaaca ccgtggcatg ctcagggagg atgctgtcct ggaatatctg601aagattgctc aagatctgga gatgtatggt gtgaactact tcagcatcaa gaacaagaaa661ggctcagagc tgtggctggg ggtggatgcc ctgggtctca acatctatga gcagaatgac721agactaactc ccaagatagg cttcccctgg agtgaaatca ggaacatctc tttcaatgat781aagaaatttg tcatcaagcc cattgacaaa aaagccccgg acttcgtctt ctatgctccc841cggctgcgga ttaacaagcg gatcttggcc ttgtgcatgg ggaaccatga actatacatg901cgccgtcgca agcctgatac cattgaggtg cagcagatga aggcacaggc ccgggaggag961aagcaccaga agcagatgga gcgtgctatg ctggaaaatg agaagaagaa gcgtgaaatg1021gcagagaagg agaaagagaa gattgaacgg gagaaggagg agctgatgga gaggctgaag1081cagatcgagg aacagactaa gaaggctcag caagaactgg aagaacagac ccgtagggct1141ctggaacttg agcaggaacg gaagcgtgcc cagagcgagg ctgaaaagct ggccaaggag1201cgtcaagaag ctgaagaggc caaggaggcc ttgctgcagg cctcccggga ccagaaaaag1261actcaggaac agctggcctt ggaaatggca gagctgacag ctcgaatctc ccagctggag1321atggcccgac agaagaagga gagtgaggct gtggagtggc agcagaagca

ggagctgcaa1381gccatgcaga tggagctgca gagccctgag tacaagctga gcaagctccg cacctcgacc1441atcatgaccg actacaaccc caactactgc tttgctggca agacctcctc catcagtgac1501ctgaaggagg tgccgcggaa aaacatcacc ctcattcggg gtctgggcca tggcgccttt1561ggggaggtgt atgaaggcca ggtgtccgga atgcccaacg acccaag

[0198] MSN-ALK蛋白质序列(AAK71522.1;GI:14625824)*

[0199] 1mpktisvrvt tmdaelefai qpnttgkqlf dqvvktiglr evwffglqyq dtkgfstwlk

[0200] 61lnkkvtaqdv rkespllfkf rakfypedvs eeliqditqr lfflqvkegi lnddiycppe

[0201] 121tavllasyav qskygdfnke vhksgylagd kllpqrvleq hklnkdqwee riqvwheehr181gmlredavle ylkiaqdlem ygvnyfsikn kkgselwlgv dalglniyeq ndrltpkigf241pwseirnisf ndkkfvikpi dkkapdfvfy aprlrinkri lalcmgnhel ymrrrkpdti301evqqmkaqar eekhqkqmer amlenekkkr emaekekeki erekeelmer lkqieeqtkk361aqqeleeqtr raleleqerk raqseaekla kerqeaeeak eallqasrdq kktqeqlale421maeltarisq lemarqkkes eavewqqkqe lqamqmelqs peyklsklrt stimtdynpn481ycfagktssi sdlkevprkn itlirglghg afgevyegqv sgmpndp

[0202] TPM4-ALK小核苷酸突变序列(AF362887.1;GI:14010353)

[0203] 1cgagaagttg agggagaaag gcgggcccgg gaacaggctg aggctgaggt ggcctccttg

[0204] 61aaccgtagga tccagctggt tgaagaagag ctggaccgtg ctcaggagcg tgcggaggtg121tctgaactaa aatgtggtga cctggaagaa gaactcaaga atgttactaa caatctgaaa181tctctggagg ctgcatctga aaagtattct gaaaaggagg acaaatatga agaagaaatt241aaacttctgt ctgacaaact gaaagaggct gagacccgtg ctgaatttgc agagagaacg301gttgcaaaac tggaaaagac aattgatgac ctggaagtgt acctccggaa gcaccaagag361ctgcaagcca tgcagatgga gctgcagagc cctgagtaca agctgagcaa gctccgcacc421ctcgac

[0205] TPM4-ALK小蛋白质突变序列(AAK51964.1;GI:14010354)

[0206] 1revegerrar eqaeaevasl nrriqlveee ldraqeraev selkcgdlee elknvtnnlk

[0207] 61sleaasekys ekedkyeeei kllsdklkea etraefaert vaklektidd levylrkhqe121lqamqmelqs peyklsklrt ld

[0208] TPM4-ALK大核苷酸突变序列(AF362886.1;GI:14010351)

[0209] 1ctggcagagt cccgttgccg agagatggat gagcagatta gactgatgga ccagaacctg

[0210] 61aagtgtctga gtgctgctga agaaaagtac tctcaaaaag aagataaata tgaggaagaa121atcaagattc ttactgataa actcaaggag gcagagaccc gtgctgaatt tgcagagaga181acggttgcaa aactggaaaa gacaattgat gacctggaag tgtaccgccg gaagcaccag241gagctgcaag ccatgcagat ggagctgcag agccctgagt acaagctgag caagctccgc301acctcgac

[0211] TPM4-ALK大蛋白质突变序列(AAK51963.1;GI:14010352)

[0212] 1laesrcremd eqirlmdqnl kclsaaeeky sqkedkyeee ikiltdklke aetraefaer

[0213] 61tvaklektid dlevyrrkhq elqamqmelq speyklsklr tst

[0214] MYH9-ALK序列(t(2;22)(p23;q11.2)染色体易位)*

[0215] RANBP2-ALK序列(t(2;2)(p23;q13)或inv(2)(p23;q11-13)染色体易位)*

[0216] ALO17-ALK序列(t(2;17)(p23;q25)染色体易位)*

[0217] CARS-ALK序列(t(2;11;2)(p23;p15;q31)染色体易位)*

[0218] 除了MSN-ALK和MYH-9之外,所有融合蛋白质含有ALK的最后563个氨基酸。MSN-ALK和MYH9分别含有最后的567和566个氨基酸。

[0219] “ALK突变”通常涉及一个相关间变性淋巴瘤激酶序列中一种核苷酸和/或氨基酸序列的变化。但是在一些实施例中,“ALK突变”涉及特定的间变性淋巴瘤激酶在ALK抑制剂(如PF-02341066和/或PDD)处理时所产生的提前突变。例如,野生型ALK蛋白(NP_004295)中1156处半胱氨酸(C1156)和/或1196处亮氨酸(L1196)分别突变为一种不同的氨基酸,它们如本发明所述地对ALK抑制剂产生抗性。在一个实施例中,C1156位置包括一个酪氨酸和/或L1196位置包括一个蛋氨酸。本领域的普通技术人员也能看出与野生型ALK蛋白中C1156和L1196分别对应的氨基酸位置数量与其它相关序列(如ALK同族体,ALK融合蛋白等)不同,不会影响其响应ALK抑制剂(如PF-02341066和/或PDD)的提前突变值。本领域的普通技术人员还能看出一个特定蛋白的氨基酸序列和核苷酸序列之间公知的确定对应能根据遗传密码(图中未显示)确定地用于蛋白质编码。同样,一个特定核苷酸的核苷酸序列与氨基酸序列之间公知的确定对应也能根据遗传密码确定的核苷酸来编码。

[0220] 遗传密码

丙氨酸(Ala, A)	GCA, GCC, GCG, GCT
精氨酸(Arg, R)	AGA, ACG, CGA, CGC, CGG, CGT

[0221]

天冬酰胺(Asn, N)	AAC, AAT
天冬氨酸(Asp, D)	GAC, GAT
半胱氨酸(Cys, C)	TGC, TGT

[0222]

谷氨酸(Glu, E)	GAA, GAG
谷氨酰胺(Gln, Q)	CAA, CAG
甘氨酸(Gly, G)	GGA, GGC, GGG, GGT
组氨酸(His, H)	CAC, CAT
异亮氨酸(Ile, I)	ATA, ATC, ATT
亮氨酸(Leu, L)	CTA, CTC, CTG, CTT, TTA, TTG
赖氨酸(Lys, K)	AAA, AAG
甲硫氨酸(Met, M)	ATG
苯丙氨酸(Phe, F)	TTC, TTT
脯氨酸(Pro, P)	CCA, CCC, CCG, CCT
丝氨酸(Ser, S)	AGC, AGT, TCA, TCC, TCG, TCT
苏氨酸(Thr, T)	ACA, ACC, ACG, ACT
色氨酸(Trp, W)	TGG
酪氨酸(Tyr, Y)	TAC, TAT
缬氨酸(Val, V)	GTA, GTC, GTG, GTT
终止信号(end)	TAA, TAG, TGA

[0223] 遗传密码的一个重要而公知的特点是它过多，因此为了使用最多的氨基酸来制备蛋白质，就需要使用一个以上的编码核苷酸三联体(例如如图所示)。因此，一些不同的核苷酸序列可以用于编码一个特定的氨基酸序列。这些核苷酸序列被认为是功能相同，因为它们能在所有生物体中产生相同的氨基酸序列(虽然有些生物体能比它们更高效地翻译一些序列)。此外，有时会在一个特定的氨基酸序列中发现一个甲基化的嘌呤或嘧啶突变体。这种甲基化不会影响三核苷酸密码子和相应的氨基酸之间的编码关系。此外本领域的普通技术人员知道在一个特定的密码子中核苷酸如何突变，从而明确地改变相关密码子表上的一个编码氨基酸碱基。例如，Cys-1156的密码子是“TGC”，而Tyr可能是“TAT”或“TAC”。因此，将位置2处的密码子G替换为A就能编码酪氨酸，而不是半胱氨酸。本领域的普通技术人员可以进行类似的操作来设计其他突变。

[0224] “结合化合物”指一种结合化合物，如一种小分子，一种抗体，一种肽，一种肽或非肽配体，一种蛋白质，一种寡核苷酸，一种寡核苷酸类似物，如一种肽核苷酸，一种凝集素或任何其他的分子实体，它们能与一个目标蛋白或分子结合，或者能与一种分析物，如一种复杂蛋白结合形成稳定的复合物。

[0225] “结合基”是指任何分子标记可直接或间接地连接的分子，它可以与一种分析物特定地结合。结合基非限制性地包括分子量大约高达1000道尔顿的抗体，抗体结合化合物，肽，蛋白质，核酸和有机分子，它含有的原子包括氢，碳，氧，氮，硫和磷。

[0226] 一个“生物标记“或”标记“是可以变化的基因，mRNA或蛋白质，其中所述的变化与癌症有关。所述的变化可以是肿瘤组织或肿瘤细胞与一种正常或健康组织或细胞(如一个控制)及在患有癌症时的数量、结构和/或活性相比，其数量，结构和/或活性的变化。例如本发明所述的与癌症有关的一个标记或是提前应答抗癌疗法的一个标记可以在肿瘤组织或肿瘤细胞中有一个与在一个正常健康组织或细胞中不同的核苷酸序列，氨基酸序列，染色体易位，内染色体倒位，拷贝数，表达水平，蛋白质水平，蛋白质活性或甲基化状态。此外，一个“标记“包括一个结构改变的分子，例如突变(包含一个突变)，例如与野生型序列不同的核苷酸或者氨基酸水平，例如在患有癌症等疾病的组织或细胞中进行替换，删除或插入。

[0227] 术语“癌症”或“肿瘤”是指具有典型致癌细胞的细胞，如不受控制其扩散，不死，转移潜能，快速生长和增殖率高，细胞形态特征的某些特征。癌细胞通常是一个肿瘤，但这种细胞可以单独地存在于一个动物中，也可以是一种非致瘤肿瘤细胞，如一种白血病细胞。本发明中术语“癌症”包括癌前病变及恶性肿瘤。癌症非限制性地包括B细胞肿瘤，如多发性骨髓瘤，瓦尔登斯特巨球蛋白血症，重链病，例如α链病，γ链病，mu链病，良性单克隆丙种球蛋白病，免疫细胞淀粉样变性，黑色素瘤，乳腺癌，肺癌(如非小细胞肺癌或NSCLC))，支气管癌，大肠癌，前列腺癌，胰腺癌，胃癌，卵巢癌，膀胱癌，脑或中枢神经系统肿瘤，周围神经系统肿瘤，食管癌，宫颈癌，子宫或子宫内膜癌，口腔或咽喉癌，肝癌，肾癌，睾丸癌，胆道癌，小肠或阑尾癌，唾液腺癌，甲状腺癌，肾上腺癌，骨肉瘤，软骨肉瘤，血液组织癌，腺癌，炎症肌纤维母细胞瘤，胃肠道间质瘤(GIST)，结肠直肠癌，多发性骨髓瘤(MM)，骨髓增生异常综合症(MDS)，骨髓增生性疾病(MPD)，急性淋巴细胞白血病(ALL)，急性髓系白血病(AML)，慢性粒细胞白血病(CML)，慢性淋巴细胞白血病(CLL)，真性红细胞增多症，霍奇金淋巴瘤，非霍奇金淋巴瘤(NHL)，软组织肉瘤，纤维肉瘤，粘液肉瘤，脂肪肉瘤，骨肉瘤，脊索瘤，血管肉瘤，内皮肉瘤，淋巴管肉瘤，淋巴管内皮肉瘤，滑膜瘤，间皮瘤，尤因氏瘤，平滑肌肉瘤，横纹肌肉瘤，鳞状细胞癌，基底细胞癌，腺癌，汗腺癌，皮脂腺癌，乳头状癌，乳头状腺癌，囊腺癌，髓样癌，支气管癌，肾细胞癌，肝癌，胆道癌，绒毛膜癌，精原细胞瘤，胚胎癌，肾母细胞瘤，膀胱癌，上皮癌，神经胶质瘤，星形细胞瘤，髓母细胞瘤，颅咽管瘤，室管膜瘤，松果体瘤，血管母细胞瘤，听神经瘤，小突神经瘤，脑膜瘤，神经母细胞瘤，视网膜母细胞瘤，滤泡性淋巴瘤，弥漫性大B细胞淋巴瘤，套细胞淋巴瘤，肝癌，甲状腺癌，胃癌，头颈部癌，小细胞癌，原发性血小板增多症，不明原因的骨髓组织异生症，慢性嗜酸粒细胞白血病，系统性肥大细胞增多症，类似嗜曙红细胞过多症，慢性嗜酸性粒细胞白血病，神经内分泌肿瘤，良性肿瘤等。

[0228] “化疗药物”是指一种化学物质，如一种细胞毒性或抑制细胞生长的试剂，它可以在一种情况如癌症时用于治疗。

[0229] “互补”是指两个核苷酸序列链之间或两个相同核苷酸序列链之间的序列互补的广泛概念。众所周知，一个第一核苷酸序列中的一个腺嘌呤残基可以和一个与之反平行的第二核苷酸序列中的一个残基形成特定的氢键(“碱基配对”)，其中的残基是胸腺嘧啶或尿嘧啶。同样众所周知的是，一个第一核苷酸链中的一个胞嘧啶残基可以和一个与之反平行的第二核苷酸序列中的一个残基形成特定的碱基配对，其中的残基是鸟嘌呤。一个核酸的一个第一序列和与其相同或相异的核苷酸的一个第二序列互补，如果这两个序列是反平行的，并且第一序列中至少有一个核苷酸残基可以与第二序列中的一个残基碱基配对。在一些实施例中，所述的第一序列包括一个第一部分，所述的第二区域包括一个第二部分，即当

第一和第二部分是反平行的，那么第一部分中至少有约50％，75％，90％或95的核苷酸残基可以与第二部分的核苷酸残基碱基配对。在其它实施例中，第一部分中的所有核苷酸残基都可以与第二部分中的核苷酸残基碱基配对。

[0230] 所述的“基因的拷贝数”或“标记的拷贝数”是指DNA序列在一个细胞中编码一个特定基因产品的数量。通常一个哺乳动物中的一个特定基因会有两个拷贝。但是所述的拷贝数可以通过基因扩增或复制来增加，或者通过删除来降低。

[0231] 一个标记是“固定”到底物上，如果它与底物是通过共价键或非共价键相连，那么底物就可以用一种液体漂洗（如标准柠檬酸盐，pH值7.4），而不需要将大量的标记从底物中分离出来。

[0232] 本发明所述的“危险系数”是指使用一种统计方法来获得一种对风险的评估。“危险系数”是一组预计的风险与另一组的比。例如将使用了一种ALK抑制剂进行治疗的病人与未使用一种ALK抑制剂的进行比较，确定ALK抑制剂是否能有效地增加疾病复发的间隔时间，尤其是对于ALK突变。例如如本发明所述，使用ALK抑制剂治疗那些肿瘤组织中已经产生ALK突变的对象，其效果会比未产生ALK突变的效果更差。

[0233] 本发明所述的“ALK抑制试剂”或“ALK抑制剂”是指一种能抑制ALK生物活性的化合物。生物活性也包括本申请所列的病人反应。典型的ALK抑制试剂非限制性地包括PF-02341066，PDD，2-甲基-11-(2-甲基丙基)-4-氧-4,5,6,11,12,13-六氢-2H-吲唑[5,4-α]吡咯[3,4-c]咔唑-8-基[4-(二甲氨基)苄基]氨基甲酸甲酯，(1S,2S,3R,4R)-3-({5-氯-2-[(1-乙基-2,3,4,5-四氢-6-甲氧基-2-氧-1H-1-苯并氮杂卓-7-基)氨基]-4-嘧啶}氨基)双环[2.2.1]庚-5-烯-2-甲酰胺和NVP-TAE684（参见例如PNAS104:270-275,2007;Choi,Y.L.et al.(2008)Cancer Res.68:4971-2976;and Biochemistry48:3600-3609,2009，本发明引用参考）。

[0234] 术语“同源”或“同源性”可以交换使用，它们是指两个多核苷酸序列或两个多肽序列之间的相似序列具有更严格的同源性比较。术语“同源性或同源百分率”和“同源性或同源％”是指两个或两个以上多核苷酸序列或两个或两个以上多肽序列序列中相似序列的百分率。“序列相似性”是指两个或两个以上的多核苷酸序列中碱基对序列的相似百分率（用任意合适的方法确定）。两个或两个以上的序列可以在任何地方具有0-100％的相似性或整体相似。同源性或相似性可以通过比较每个序列的相同位置来获得。当比较序列中的一个位置上是相同的核苷酸或氨基酸，那么这些分子在这个位置上就是同源的。多核苷酸序列的相似率或同源率是指这些多核苷酸序列共有位置上核苷酸同源或匹配的一个数量函数。多肽序列的一个同源率是指这些多肽序列共有位置上氨基酸同源的一个数量函数。多肽序列的一个同源率或相似率是指这些多肽序列共有位置上氨基酸的一个数量函数。本发明所述的术语“实质同源”是指至少50％，60％，65％，70％，75％，80％，85％，90％，95％或更多同源。

[0235] 癌症的“抑制”，如果至少有一个癌症症状缓解，终止，延缓或阻止。如本发明所使用的，如果能减少，减缓，减慢或预防癌症复发或转移，那么癌症也被“抑制”。

[0236] 一个“标记核酸”或“生物标记核酸”是一种由本发明所述的一个标记编码或与之对应的核酸（如DNA，mRNA，cDNA）。例如这种标记核酸分子包括DNA（例如，基因组DNA和cDNA），该DNA包括表1所列的任意核苷酸序列或这些序列中的一部分，或是这些序列的互补

序列或杂交序列。这种标记核酸分子也包括RNA，该RNA包括表1所列的任意核苷酸序列或这些序列中的一部分，或是这些序列的互补序列，其中所有的胸嘧啶残基被尿嘧啶残基替换。一种“标记蛋白质”是由本发明所述的一个标记编码的蛋白或对应的蛋白。一种标记蛋白质包括表1所列的任意序列或序列片段编码的整个蛋白序列或部分蛋白序列。术语“蛋白”和“多肽”在本发明中含意相同。

[0237] 一个标记的“正常”拷贝数或一个标记的“正常”表达水平是指在一个生物样本中，如从一个未患有癌症的目标对象，如人体中采集的含有痰液，支气管肺泡灌洗，胸腔积液，组织，全血，血清，血浆，口腔刮，唾液，脑脊液，尿液，粪便和骨髓的样本中，所述标记的表达水平和拷贝数。

[0238] ALK基因突变和/或基因产物(如表1所列的标记)的一种“过度表达”或“明显太高的表达水平，拷贝数和/或活性”是在一个检测样本中，一种表达水平，拷贝数和/或活性比标准的表达水平或拷贝数检测错误更大，它可以至少是一个控制样本(如从一个未患有癌症的健康对象中采集的样本)中ALK基因突变和/或基因产物(如表1所列的标记)表达水平或拷贝数的两，三，四，五或十或更高倍数，或者至少是数个控制样本中ALK基因突变和/或基因产物(如表1所列的标记)的平均表达水平或拷贝数的两，三，四，五或十或更高倍数。

[0239] 术语“探针”是指任何可以选择性地与一个特定目标分子结合的分子，例如本发明所述的一个标记。探针既可以由本发明的普通技术人员合成，也可以由合适的生物制剂制备。为了检测目标分子，可以如本发明所述的使探针特别地带有标记。可以用作探针的典型分子非限制性地包括核糖核酸，脱氧核糖核酸，蛋白质，抗体和有机单体。

[0240] “RECIST”是一个缩写，意为“固体肿瘤反应的评估标准”，它公布的标准是当癌症患者在治疗过程中病情好转(“反应”)，保持不变(“稳定”)或恶化(“进展“)。例如2000年2月2日国家癌症研究杂志92期3号已经公布了根据RECIST标准确定的反应，RECIST标准可以包括其它相似的定义和规则设置。本领域的普通技术人员明白本发明可以使用RECIST标准确定的定义，如"PR"、"CR"、"SD"和"PD"。

[0241] 本发明所述的“回应”治疗的“反应”及它的其它时态是指一个对象使用一种ALK抑制剂后产生的反应。在一个实施例中，一个对象使用一种ALK抑制剂后会产生反应，即该对象中的肿瘤会大约缓慢增长10％，20％，30％，40％，50％，60％，70％，80％，90％或更多。在另一个实施例中，一个对象使用一种ALK抑制剂后会产生反应，即使用合适的检测，如质量或体积等会发现该对象中的肿瘤大约减少5％，10％，20％，30％，40％，50％或更多。在另一个实施例中，一个对象使用一种ALK抑制剂后会产生反应，即其平均寿命会比未接受治疗的平均寿命大约增加5％，10％，20％，30％，40％，50％或更多。在另一个实施例中，一个对象使用一种ALK抑制剂后会产生反应，即这个对象有一个增加的无病存活率，整体存活率或增加恶化时间。可以使用几种方法来确定病人是否对治疗有反应，这些方法包括上述的RECIST标准。

[0242] “样本”，“组织样本”，“病人样本”，“病人细胞或组织样本”或“标本”均是指从一个对象或病人的一个组织中采集获得的相似细胞。该组织样本的来源可能是从一个新鲜的，冷冻和/或保存器官中获得的固体组织，组织样本，切片，或吸出物；血液或任何血液成分；体液如脑脊髓液，羊水，腹水或胞间液；或是在目标体中产生或生长任意时间的细胞。所述的组织样本可以含有与天然组织混合的非天然化合物，如防腐剂，抗凝血剂，缓冲剂，固定

剂，营养剂，抗生素等。

[0243] 如果通过一个比评估量的标准错误更多的量，或者至少是评估量的两，三，四，五，十或更多更好多倍的量使该标记的量分别比正常水平更高或更低，那么一个标记的量，例如ALK基因突变或基因产物（如表1所列的标记）的表达或拷贝数，在一个对象中是“明显”地高于或低于一个标记的正常量。或者在对象体中该标记的量可以被认为是“明显”高于或低于正常量，如果这个量至少是标记正常量的两，三，四，五倍，更高或更低倍。

[0244] 本发明中，“重要事件”是指病人患病时一个事件，本领域的普通技术人员认为其是重要的。典型的重要事件非限制性地包括初次诊断，死亡，复发，确定病人的病变发生转移，病人疾病复发或病人疾病从上述任何一个阶段发展到另一个阶段。一个重要事件可以是由本领域的普通技术人员使用OS，TTP和/或RECIST或其它反应标准评估确定的任何重要事件。

[0245] 本发明中，“对象”和“病人”可以互换使用。本发明中所使用的术语“对象”和“对象们”是指动物，如包括非灵长类动物（如牛，猪，马，驴，羊，骆驼，猫，狗，豚鼠，大鼠，小鼠，羊）和灵长类动物（如猴，如食蟹猴，大猩猩，黑猩猩和人类）在内的哺乳动物。

[0246] 本发明中所使用的“时间过程”是指初始事件和随后事件之间的时间量。例如一个病人的癌症，其时间过程可以涉及一个病人的疾病以及在该疾病过程中可以使用显著测量物来测量的，其中例如第一事件可以是被诊断，而后面的事件可以是转移。

[0247] “发展时间”或“TTP”是指从治疗开始到发展或一种癌症或检查的一个测量时间。检查可能来自一个研究结束或来自治疗的一个变化。发展时间也可以表示为一种概率，例如在一个评估事件发展点中，发展时间可以代表其在一个特定的时间内自由发展的可能性，时间是指治疗开始到发展或检查之间的时间。一个“多核苷酸转录”是一个多核苷酸（例如一个RNA，一个cDNA或者是一个RNA或cDNA的同源），它与一个成熟RNA的全部或一部分互补或同源，其中成熟RNA是由本发明的一个标记转录以及正常的转录后处理（如拼接），如果有的话，转录和转录的反转录获得的。

[0248] “治疗”，“处理”以及这个单词的其它形式都是指使用ALK抑制剂来限制一种癌症的生长，使一种癌症缩小重量或体积，延长对象的预计存活时间或者是肿瘤的发展时间等。

[0249] ALK基因突变和/或基因产物（如表1所列的标记）的一个“低表达”或“明显降低的表达水平，拷贝数和/或活性”是指一个检测样本的表达水平或拷贝数比用来评估表达或拷贝数的标准误差更大，至少比一个控制样本中ALK基因突变和/或基因产物（如表1所列的标记）的表达水平，拷贝数和/或活性低两倍，三倍，四倍，五倍，十倍或更多倍，或者是比几个控制样本中ALK基因突变和/或基因产物（如表1所列的标记）的平均表达水平，拷贝数和/或活性低两倍，三倍，四倍，五倍，十倍或更多倍。

[0250] II.发明实施方法

[0251] 本发明至少部分是基于对特定区域基因组的识别，包括如ALK突变体，及在治疗癌症中ALK抑制剂的效果判断。ALK基因表达序列的分析来鉴定对ALK多肽的新型突变（如表1所列的生物标记，包括ALK多肽），可以使至少部分的多肽具有对ALK抑制剂治疗的抵抗性。因此，在本发明的范围内这里所述不同的方法中存在和/或不存在一个或多个这样的生物标记。

[0252] 在一些实施方案，本发明的方法可以用来监测一个对象的癌症进程，其中，在癌症

进程中可识别出对象的样本存在一个或多个突变ALK(如EML4-ALK突变),如在最初点时间内和随后点时间内,然后癌症越来越少对ALK抑制剂介入治疗的响应,反之亦然。在另一些实施例中,在最初点时间和随后点时间之间,对象已接受如化疗,放疗,手术,或其他任何有用的抑制癌症的疗法等治疗,或已完成治疗,或正在缓解。

[0253] 在这里进一步说明,本发明的一个或多个生物标记(如突变ALK,包括突变EML4-ALK)可通过存在的基因组(如生殖和/或体细胞)序列与参考序列相比较来特定识别,如SEQ序号1。例如,这里所述的方法可以影响探测本发明所述的生物标记,通过对核酸靶向目标的DNA的延伸扩增反应包括表1所列的一个或多个突变,并分析存在的一个或多个突变核酸靶向目标来实现。

[0254] 在文献中各种用于核酸扩增技术是众所周知的,如:PCR(聚合酶链反应),公开于美国专利号4683195(引用参考),美国专利号4683202(引用参考)和美国专利号4800159(引用参考),和它的替代RT-PCR(逆转录聚合酶链反应),特别是它的一步式在专利EP-B-0.569.272中公开,LCR(连接酶链反应)如在专利申请EP-A-0.201.184公开,RCR(修复链反应)如在国际专利申请WO-A-90/01069公开(引用参考),3SR(自主序列复制系统)如在国际专利申请WO-A-90/06995公开(引用参考),NASBA(基于核酸序列扩增技术)如在专利EP-β-0.397.269和美国专利号5466586(引用参考)使用的DNA双链为模板,和TMA(转录介导扩增技术)如在美国专利号5399491公开(引用参考)。

[0255] 在文献中知道在不同的方式下可发现存在一个或多个突变的扩增产品,例如DNA测序方法如桑格测序和深度测序,限制性内切酶的应用,等位基因特异性扩增,肽核酸(PNA)-介导PCR,构象差异的发现,如链构象多态性检测(SSCP)和变性梯度凝胶电泳(DGGE),用标记的寡核苷酸探针来分析检测细胞膜(斑点杂交),分析检测微孔板,如反向杂交法,寡核苷酸链接检测法(OLA,MLPA),第一核苷酸变化(FNC)技术,交联技术,快速循环PCR和同步荧光分析(如5'核苷酸酶/荧光定量),和PCR后用微型测序用质谱测定法或毛细管电泳法。

[0256] III.核酸分子分离的实施

[0257] 本发明的一方面涉及分离核酸分子,相当于本发明的生物标记,包括相当于本发明的标记核酸编码的多肽或一部分这样的多肽。本发明的核酸分子包括那些核酸分子属于这里所识别的ALK或ALK-相关基因组(如生殖和/或体细胞)和/或编码ALK或ALK-相关(如EML4-ALK)多肽。在一些实施例中,本发明的核酸分子由表1所列的包括必须的或包括核序列或其片段组成。本发明的分离核酸分子也包括必须用于杂交探针来识别核酸分子的核酸分子,相当于本发明的标记,包括编码相当于标记的多肽的核酸分子,和这些核酸分子的片段,如那些适用作PCR扩增的引物或核酸分子突变。这里所使用的术语"核酸分子"是包括DNA分子(如cDNA或基因组DNA)和RNA分子(如mRNA)和用核苷酸生产的DNA或RNA类似物。核酸分子可以是单链或双链,在某些实施例中的核酸分子是双链DNA。

[0258] "分离"核酸分子是从其他核酸分子中分离出来的一种,其存在于核酸分子的天然来源中。在一些实施例中,"分离"核酸分子是自由序列(如蛋白质编码序列),其自然侧链核酸(如序列位于5'和3'末端的核酸)在核酸衍生而来的生物体基因组DNA中。例如,在各种实施例中,分离核酸分子包含少于5KB,少于4KB,少于3KB,少于2KB,少于1KB,少于0.5KB或少于0.1KB的核苷酸序列,其自然侧链核酸在细胞衍生而来的生物体基因组DNA中。此外,"分

离”核酸分子如cDNA分子，当通过重组技术生产时基本没有其他微孔材料或培养基，或当化学合成时基本没有化学前体或其他化学物质。

[0259] 语句“基本没有其他微孔材料或培养基”包括制备的核酸分子，是从分离或重组产生的细胞微孔结构中分离出来的。因此，核酸分子基本没有微孔材料包括制备的核酸分子具有少于30％，少于20％，少于10％，少于5％(干重)的其他微孔材料或培养基。

[0260] 本发明的核酸分子，如表1所列的突变ALK基因，用标准分子生物技术和这里所述的数据库记录中的序列信息来分离。使用全部或部分这样的核酸序列，本发明的核酸分子可使用标准的杂交分离和克隆技术来分离得到(即，在Sambrook et al.,ed.,Molecular Cloning:A Laboratory Manual,2nd ed.,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,NY,1989中所述)。

[0261] 本发明的核酸分子可以利用cDNA，mRNA，或基因组DNA(如生殖和/或体细胞基因组DNA)作为模版和依据标准的PCR扩增技术的寡核苷酸引物来扩增得到。通过DNA序列分析法可将扩增的核酸分子克隆到合适的载体中并表达。此外，寡核苷酸相当于全部或部分通过标准合成技术如使用自动DNA合成器分离得到本发明的核酸分子。

[0262] 在另一实施例中，本发明的分离核酸分子包括具有与本发明标记相一致的核苷酸序列互补的核酸序列的核酸分子或与本发明标记相一致的蛋白质编码的核酸序列的核酸分子。与给定的核苷酸序列互补的核酸分子是充分的与给定的核苷酸序列互补的，它可与给定的核苷酸序列进行杂交因此形成稳定的双链。

[0263] 此外，本发明的核酸分子可以仅仅包括核酸序列一部分，其中全长的核酸序列包括本发明标记或相当于本发明标记的编码多肽。这样的核酸分子可用作如探针或引物。探针/引物通常用作一个或多个基本纯化的寡核苷酸。寡核苷酸通常包括一段在严格条件下杂交的核苷酸序列区，至少约7，至少约8，至少约9，至少约10，至少约11，至少约12，至少约13，至少约14，至少约15，至少约16，至少约17，至少约18，至少约19，至少约20，至少约21，至少约22，至少约23，至少约24，至少约25，至少约26，至少约27，至少约28，至少约29，至少约30，至少约31，至少约32，至少约33，至少约34，至少约35，至少约36，至少约37，至少约38，至少约39，至少约40，至少约45，至少约50，至少约55，至少约60，至少约65，至少约70，至少约75，至少约80，至少约85，至少约90，至少约95，至少约100或更连续的本发明核酸的核苷酸。

[0264] 基于本发明核酸分子序列的探针可用于检测相当于一个或多个本发明标记的转录产品或基因序列。探针包括一组附属的标签如放射性同位素，荧光化合物，酶，或酶辅助因子。这些探针可以作用于识别不当表达蛋白蛋的细胞或组织的诊断试测试剂盒的一部分，如测量对象细胞样本的核酸分子编码蛋白质的水平，如检测mRNA水平或确定基因编码蛋白质是突变或删除。

[0265] 本发明另外包含的核酸分子与突变的ALK基因或基因产物(表1所列的标记)基本上一致，如此以致于它们至少60％，至少65％，至少有70％，至少75％，至少80％，至少85％，至少86％，至少87％，至少88％，至少89％，至少有90％，至少91％，至少92％，至少有93％，至少94％，至少95％，至少96％，至少97％，至少98％，至少99％，至少99.5％或更高是相同的。在另外一个实施例中，本发明另外包含的核酸分子与突变的ALK基因或基因产物(表1所列的标记)基本上一致，如此以致于它们之间只有仅仅一点或至少1，至少2，至少3，至少4，至少5，至少6个，至少7，至少8，至少9，至少10，至少11，至少12，至少13，至14，至少15，至少

16，至少17，至少18，至少19，至少20，至少21，至少22，至少23，至少24，至少25，至少26个，至少27，至少28，至少29，至少30，至少31，至少32，至少33，至少34，至少35，至少36，至少37，至少38个，至少39，至少40，至少45，至少50，至少55，至少60，至少65，至少70，至少75，至少80，至少85，至少90，至少95，至少100个核苷酸或任何范围的区别。

[0266] 术语“单核苷酸多态性”(SNP)是指单个核苷酸侵占多态性位置，即在等位基因序列中发生位置变化。该位置的前面和后面通常是高度保守的等位基因序列(如序列变化小于人口的1/100或1/1000)。SNPs也可以由在所涉及到的等位基因上删除一个核酸或插入一个核酸引起的。多态性位置通常被不同于所提到底物的底物侵占。例如，所涉及到的等位基因包括在多态性位置上的底物“T”(胸腺嘧啶)，变化等位基因包括在多态性位置上的“C”(胞嘧啶)，“G”(鸟嘌呤)，或“A”(腺嘌呤)。SNP可发生在蛋白质编码核酸序列上，在这种情况下，它们可能导致缺陷或其他突变蛋白质，或遗传疾病。这样SNP可改变基因的编码序列并因此得到其他氨基酸(错义SNP)或SNP可引入终止密码子(无意义SNP)。当SNP没有改变蛋白质的基氨酸序列，SNP被称为“沉默”。SNP也可发生在核酸序列的非编码区上，这可导致缺陷蛋白质表达，例如，导致具有额外刺激，或可能不会对蛋白质功能影响的结果。

[0267] 在另一实施例中，本发明的分离核酸分子是至少7，至少15，至少20，至少25，至少30，至少35，至少40，至少45，至少50，至少55个，至少60岁，至少65岁，至少70，至少75，至少80，至少85，至少90，至少95，至少100，至少125，至少150，至少175，至少200，至少250，至少300，至少350，至少400，至少450，至少550，至少650，至少700人，至少800，至少900，至少1000，至少1200，至少1400，至少1600，至少1800，至少2000，至少2200，至少有2400，至少2600，至少2800，至少3000，至少3500，至少4000，至少4500，或更长的核苷酸，并在严格条件下将核酸分子杂交成相当于本发明标记或将核酸分子的编码蛋白质成相应的本发明标记。这里所用的术语“在严格条件下杂交”是指用于杂交或洗涤的条件，其核苷酸序列通常至少60%，65%，70%，75%，80%，85%保持相同彼此互相杂交。这些严格的条件在文献John Wiley& Sons，N.Y.(1989)的分子生物学的普通实验6.3.1-6.3.6章节中知道并找到那些技术。另此，一个非限制性严格杂交条件的例子，在6X氯化钠/柠檬酸钠(SSC)45℃下，随后在0.2X SSC，0.1%SDS50-65℃下洗涤一次或多次。

[0268] 本发明还包括分子信标核酸分子，其至少具有一个与本发明的核酸分子互补的区域，这样的分子信标用来定量样本中存在本发明的核酸分子。“分子信标”核酸是包括一对互补区域和具有荧光和与此相关的荧光猝灭的核酸分子。荧光和猝灭在方向上与核酸不同的部分有关，当另一个互补区域退火时，荧光的荧光性通过猝灭得以猝灭。当核酸分子的另一个互补区域没有退火时，荧光的荧光性猝灭程度较轻。分子信标核酸分子被公开，例如，在美国专利5876930(引用参考)。

[0269] IV.分离蛋白质与抗体的实施

[0270] 一个方面，本发明涉及到相当于本发明单独标记的分离蛋白和其生物活性。在一实施例中，原生态多肽可从细胞或组织源中通过用标准蛋白质纯化技术的适当纯化方案来分离得到相当于本发明的标记。在另一实施例中，多肽通过重组DNA技术得到相当于本发明标记。重组表达，多肽相当于采用标准肽合成技术化学合成的本发明的标记。

[0271] “分离”或“净化”蛋白质或其生物活性部分基本不含蛋白质来源的细胞或组织源中的细胞物质或其他污染蛋白质，或当化学合成时基本不含化学前体或其他化学物。语句

“基本不含细胞物质”包括从分离或重组得到的细胞的细胞成分中分离得到的制备蛋白质。因此，基本不含细胞物质的蛋白质中包括制备的蛋白质具有至少约30%，至少约20%，至少约10%，至少约5%(干重)的异源蛋白(也就是这里所提到的“污染蛋白质”)。当蛋白质或其生物活性部分是通得重组得到，它可基本不含有培养基，如培养基至少约20%，至少约10%，至少约5%的制备蛋白质体积。当蛋白质是通过化学合成得到，它可基本不含有化学先体或其他化学物，如从合成蛋白质中所涉及到的化学先体或其他化学物中分离出来。因此，这样制备的蛋白质具有比添加物多肽少于约30%，少于约20%，少于约10%，少于约5%(干重)的化学先体或化合物。

[0272] 相当于本发明标记的多肽生物活性部分包括由本发明的氨基酸序列或采自与ALK基因突变和/或基因产品(表1所列标记)的蛋白质氨基酸序列组成的多肽。其包括比全长蛋白质更少的氨基酸，并呈现至少相当于全长蛋白质一倍的活力。通常情况下，生物活性部分包括一个具有至少相当于蛋白质一倍活力的区域或结构。本发明的蛋白质生物活性部分可以是多肽，如10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,30,35,40,45,50,55,60,65,70,75,80,85,90,95,100或更多氨基酸长度。此外，删除了蛋白质的其他区域的其他活性部分，可通过重组技术制备并识别一个或多个本发明原生态多肽的起作用的活性点。

[0273] 在某些实施例中，多肽具有表1所列核酸分子编码蛋白质的氨基酸序列。其他有用蛋白质与这些序列中的一个基本相同(例如，至少60，至少65，至少70，至少75，至少有80，至少85，至少86，至少87，至少88，至少89，至少90，至少91，至少92，至少93，至少94，至少95，至少96，至少97，至少98，至少99，至少99.5％或更大)，并保留不同氨基酸序列中全长蛋白质的有用的蛋白质活性(如对ALK抑制剂具有抵抗性或敏感性)。

[0274] 要确定两个氨基酸序列或核酸的百分率特性，对比序列以获得最佳比较目的(在第一个氨基酸或蛋白质核酸序列与第二个氨基酸或核酸序列连接处插入分歧点)。然后对氨基酸残基或核苷酸对应的氨基酸位置或核苷酸位置进行比较。当第一个序列被相当于第二个序列相同的氨基酸残体或核苷酸占用，则该位置的分子相同。两个序列中的百分率特性是序列完全共享位置的数量的函数(如%特性=相同位置数#/总位置数#(如重叠位置)×100)。在一个实施例中，这两个序列是相同长度。

[0275] 两个序列中的百分率特性可以通过一种数学算法完成。类似的，非限制性实施例，利用数学算法来比较两个序列可以是Karlin andAltschul(1990)Proc.Natl.Acad.Sci.USA87:2264-2268，改进Karlin and Altschul(1993)Proc.Natl.Acad.Sci.USA90:5873-5877中的算法。这种算法引用到Altschul,et al.(1990)J.Mol.Biol.215:403-410的NBLAST和XBLAST法中。BLAST核苷酸检索可通过NBLAST法来执行，分值=100，字长=12来获得与本发明核酸分子相同的核苷酸序列。BLAST蛋白质检索可通过XBLAST法来执行，分值=50，字长=3来获得与本发明蛋白质分子相同的氨基酸分子。为了获得比较目的的分歧线，可利用在Altschul et al.(1997)Nucleic Acids Res.25:3389-3402中所描述的分歧BLAST。另外，PSI-Blast可用作对出现分子间远亲关系重复检索。当使用BLAST，分歧BLAST和PSI-Blast 程序时，可用各个程序的默认参数(如XBLAST和NBLAST)(详见万维网ncbi.nlm.nih.gov 上的NCBI网页)。在另一非限制性实施例中，用于序列比较的数学算法是Myers and Miller,(1988)Comput Appl Biosci,4:11-7上的算法。这样一个算法引入到ALIGN程序(2.0版)中用于比较氨基酸序列，PAM120重残基表

中，可用分歧长度点为12，和分歧长度点为4。另一用于类似和线性局部序列的识别区域的有用算法是如在Pearson and Lipman(1988)Proc.Natl.Acad.Sci.USA85:2444-2448所描述的FASTA算法。当用FASTA算法来比较核苷酸或氨基酸序列时，例如，PAM120重残基表可用值为2的K-元数组。

[0276] 两个序列间的百分率特性可用类似以上所述的那些技术来确定具有或不具有分歧。在计算百分率特性中，仅有精确匹配的才计算。

[0277] 相当于本发明的标记的分离多肽或其碎片可用作使用多克隆和单克隆标准技术来生产制备抗体的免疫原。全长多肽或蛋白质可用作或本发明提供的抗原肽片段用作免疫原。本发明的抗原肽蛋白质含有至少8(或至少10，至少15，至少20，至少30或更多)本发明多肽的一个氨基酸序列的残体氨基酸，和包括蛋白质抗原表位，以致于抗体突出反向蛋白质结构来明确具有本发明标记的免疫综合体来与蛋白质相符。抗原肽包括的实施抗源表位位于蛋白质表面的区域，如亲水区。疏水性序列分析，亲水性序列分析，或类似的分析可以用来识别亲水区。

[0278] 免疫原通常是通过免疫接种(即免疫)合适的对象如兔子、山羊、老鼠或其他哺乳动物或脊椎动物来制备抗体。适当的免疫制剂包含如重组表达或化学合成多肽。制剂更进一步还包括辅助剂如弗氏完全或不完全辅助剂，或类似的免疫剂。

[0279] 因此，本发明的另一方面涉及到定向对抗本发明多肽的抗体。术语“抗体”和“抗体物质”在这里可相互替换，是指免疫球蛋白分子和免疫球蛋白分子的免疫活性部分，即分子含有抗原粘合位置，其特异性粘合抗原，如本发明的多肽。特异粘合到特定的本发明多肽上的分子是在同一个样品中如含有天然多肽的生物样品只粘合到多肽上但基本不粘合在其他分子上的分子。免疫球蛋白分子的免疫活性部分包括F(ab)和F(ab')2片段，其可通用采用酶如胃蛋白酶来处理抗体得到。本发明提供多克隆和单克隆抗体。在这里所用的术语“多克隆抗体”和“单克隆抗体成份”，涉及到抗体分子群，其仅仅含有一种具有特定抗原表位的免疫能力的抗原粘合位置。

[0280] 用本发明多肽作为免疫原免疫到合适的对象上来制备上述的多克隆抗体。免疫对象中的抗体浓度量可通过标准技术如用酶联免疫吸附测定法(ELISA)来随时间测定。如果需要，抗体分子可以从对象(如从对象的血液与血清中)中获得或分离得到并通过已知的技术如蛋白质层析纯化获得IgG片段。在免疫后的适当时间如特定的抗体浓度量最高，抗体生成细胞从对象上获得并通过标准技术如Kohler and Milstein(1975)Nature256:495-497上最早公开的杂种瘤技术，人类B细胞杂种瘤技术(见Kozbor et al.,1983,Immunol.Today4:72)，EBV-杂种瘤技术(见Cole et al.,pp.77-96In Monoclonal Antibodies and Cancer Therapy,Alan R.Liss,Inc.,1985)或三体杂交瘤技术来制备单克隆抗体。用于生产杂交瘤细胞的技术是众所周知的(见generally Current Protocols in Immunology,Coligan et al.ed.,John Wiley&Sons,New York,1994)。杂种瘤细胞生产本发明单克隆抗体通过筛选杂种瘤培养液的上浮液来得到抗体，其粘合在所目标的多肽上，即用标准ELISA试验法。

[0281] 制备单克隆体抗体-分泌杂种瘤的替代方法，定向对抗本发明的多肽可通过筛选具有所目标多肽的重组组合免疫球蛋白库(如噬菌体抗体陈列库)识别并分离得到。用于生产和筛选噬菌体抗体陈列库的试剂盒在市场上可得到(如法玛西亚重组噬菌体抗体系统，

产品编号27-9400-01;和Stratagene公司SurfZAP噬菌体试剂盒,产品编号240612)。此外,特别适合用在生产和筛选抗体陈列库的方法和试剂可在如美国专利号5223409(引用参考),PCT公布号WO92/18619(引用参考),PCT公布号WO91/17271(引用参考),PCT公布号WO92/20791(引用参考),PCT公布号WO92/15679(引用参考),PCT公布号WO93/01288(引用参考),PCT公布号WO92/01047(引用参考),PCT公布号WO92/09690(引用参考),PCT公布号WO90/02809(引用参考),福克斯等(1991)生物/技术9:1370-1372,Hay et al.(1992)Hum.Antibod.Hybridomas3:81-85,Huse et al.(1989)Science246:1275-1281,Griffiths et al.(1993)EMBO J.12:725-734等中找到。

[0282] 此外,重组抗体如嵌合的和人化的单克隆抗体,包括人类和非人类部分,在本发明范围内其可用标准重组DNA技术来制造。这种嵌合的和人化的单克隆抗体可通过在现有文献中所知的重组DNA技术来制备,如在PCT公开号WO87/02671(引用参考),欧洲专利申请号184187,欧洲专利申请号171496,欧洲专利申请号173494,PCT公开号WO86/01533(引用参考),美国专利号4816567(引用参考),欧洲专利申请125023,Better et al.(1988)Science240:1041-1043;Liu et al.(1987)Proc.Natl.Acad.Sci.USA84:3439-3443,Liu etal.(1987)J.Immunol.139:3521-3526,Sun et al.(1987)Proc.Natl.Acad.Sci.USA84:214-218,Nishimura et al.(1987)Cancer Res.47:999-1005,Wood et al.(1985)Nature314:446-449,和Shaw et al.(1988)J.Natl.Cancer Inst.80:1553-1559,Morrison(1985)Science229:1202-1207,Oi et al.(1986)Bio/Techniques4:214,美国专利5225539(引用参考),Jones et al.(1986)Nature321:552-525,Verhoeyan et al.(1988)Science239:1534,和Beidler et al.(1988)J.Immunol.141:4053-4060。

[0283] 用来治疗人类对象的完全人类抗体特别需要。这种抗体可用转基因老鼠来生产,其无能力表达内源性免疫球蛋白重链和轻链基因,但可表达人类重链和轻链基因。转基因老鼠在正常方式下有选择性的抗原免疫,如相当于本发明标记的全部或部分多肽。定向对抗单克隆抗体可用传统杂种瘤技术获得。人类免疫球蛋白转基因在转基因老鼠B细胞分化期间的重新排列生存,并随后进行类别转换和体细胞突变。因此,用这种技术,可生产用于治疗IgG,IgA和IgE的抗体。用于生产人类抗体的技术综述,详见Lonberg and Huszar(1995)Int.Rev.Immunol.13:65-93。用于生产人类抗体和人类单克隆抗体技术的详细论述和生产这种抗体的规约,详见美国专利5625126(引用参考),美国专利5633425(引用参考),美国专利5569825(引用参考),美国专利5661016(引用参考),和美国专利5545806(引用参考)。另外,公司如Abgenix,Inc.(Freemont,CA)可从事提供用类似上述的技术定向对抗选择性抗原的人类抗体。

[0284] 选择性识别抗原表位的完全人类抗体可使用被称为“导向选择”的技术来生产。这种方法用选择性非人类单克隆抗体如鼠科抗体来导向选择识别相同抗原表位的完全人类抗体(Jespers et al.,1994,Bio/technology12:899-903)。

[0285] 相当于本发明标记的定向对抗多肽的抗体(如单克隆抗体)可用于通过标准技术如亲和层析或免疫沉淀反应来分离多肽。而且,这种抗体可用于识别标记(如细胞溶解产物或细胞悬浮物)来判断标记的表达水平和模型。这抗体也可用于诊断监视在诊断测试程中组织或体内流体(如肿瘤细胞体内流体)的蛋白质水平,如确定一个给定治疗方案的效果。检测可促进抗体偶联到可见物质上。例如可见物质包括但不限于,各种酶,辅基,荧光材料,

发光材料，生物发光材料和放射性材料。例如适当的酶包括但不限于，辣根过氧化物酶，碱性磷酸酶，-半乳糖苷酶，或乙酰胆碱酯酶；例如适当的辅基化合物包括但不限于抗生蛋白链菌素/生物素，抗生素蛋白/生物素；例如适当的荧光材料包括但不限于伞形花内酯，荧光素，异硫氰酸荧光素，罗丹明，二氯三嗪胺荧光素，丹磺酰氯或藻红蛋白；例如发光材料包括但不限于鲁米诺，和例如适当的放射性材料包括但不限于125I，131I，35S或3H。

[0286] V.载体和宿主细胞重组表达的实施

[0287] 本发明的另一方面涉及到载体如表达载体，包括相当于本发明标记的编码多肽(或该多肽的部分)的氨基酸。这里所用的术语“载体”是指有能力运送到其已连接的另一核酸的核酸分子。载体的一种类型是“质粒”，是指可增加DNA片段循环连接到环形双链DNA上。载体的另一种类型是病毒载体，是指增加DNA片段可连接插入到病毒基因组中。一些载体有能力自主复制并引入宿主细胞中(如具有细菌复制启点的细菌载体和附加型哺乳动物载体)。其他载体(如非附加型哺乳动物载体)整合到宿主细胞基因上并引入宿主细胞，并随宿主基因一起复制。此外，一些载体，也叫表达载体，是有直接表达所合成连接基因的能力。一般来说，重组DNA技术的功能表达载体常常是从质粒(载体)中得到。然而，本发明计划包含其他形式的表达载体，如病毒载体(如复制缺陷逆转录病毒，腺病毒和腺相关病毒)，它们具有相同的功能。

[0288] 本发明的重组表达载体，包括适合在本发明宿主细胞中核酸表达的形式的核酸。这意味着重组表达基因包括一个或多个控制序列，基于宿主细胞有选择性的进行有用表达，并合成连接到核酸序列上来表达。在重组表达载体中，“合成连接”的意思是指所目标的核酸序列以一定的方式连接到允许核苷酸序列表达的控制序列上(如在体内转录/翻译系统中或在载体引进宿主细胞时的宿主细胞中)。术语“控制序列”是指含有启动子，增强子和其他表达控制元件(如多聚腺苷酸信号)。这种控制序列在如Goeddel，Methods in Enzymology:Gene Expression Technology vol.185，Academic Press，San Diego，CA(1991)中被描述。控制序列包括那些在多类型宿主细胞中直接结构表达的核苷酸序列和那些在一些宿主细胞中直接表达的核苷酸序列(如组织特异性控制序列)。通过在载体表达设计文献中的那些技能，如选择宿主细胞来改变些依赖因素，希望的蛋白质表达水平等等越来越被重视。本发明的表达载体可引入宿主细胞来生产蛋白质或肽，通过这里所述的核酸来编码具有功能性的蛋白质或肽。

[0289] 本发明的重组表达载体可用原核细胞(如大肠杆菌)或真核细胞(如昆虫细胞{使有了杆状病毒表达载体}，酵母细胞或哺乳动物细胞)来设计相当于本发明标记的多肽表达。适合的宿主细胞在Goeddel，supra中有更一步的论述。重组表达载体可以在如具有T7启动子控制序列和T7聚合酶的体外转录和翻译。

[0290] 原核生物的蛋白质表达最常用的是在含有构成或诱导启动子指导融合型或非融合性蛋白质表达的载体大肠杆菌中。其中在融合型载体通常加入一些氨基酸到编码蛋白质中作为重组蛋白质的氨基酸终止子。这种融合型载体通常具有三个作用：1)增加重组蛋白质的表达；2)增加重组蛋白质的溶解度；和3)通过在亲和纯化作用中的配合基有助于重组蛋白质的纯化。通常，在融合弄表达载体中，蛋白水解切割位点是在引入融合部分与重组蛋白质的联结点，其可从融合型蛋白质纯化后的融合部分中分离得到重组蛋白质。这种酶和他们同源识别序列，包括Xa因子，凝血酶和肠激酶。典型融合型表达载体包括pGEX

(Pharmacia Biotech Inc;Smith and Johnson,1988,Gene67:31-40),pMAL(New England Biolabs,Beverly,MA)和融合谷胱甘肽-S-转移酶(GST)的pRIT5(Pharmacia,Piscataway,NJ),麦芽糖E结合蛋白,或蛋白A,分别靶向重组蛋白。

[0291] 例如,合适的非融合型大肠杆菌表达载体包括pTrc(Amann et al.,1988,Gene 69:301-315)和pET11d(Studier et al.,p.60-89,In Gene Expression Technology: Methods in Enzymology vol.185,Academic Press,San Diego,CA,1991)。靶基因的表达载体pTrc依赖来自杂交trp-lac融合型启动子的宿主RNA聚合酶转录。靶基因的表达pET11d载体依赖来自杂交T7gn10-lac融合型启动子介导的共同表达病毒RNA聚合酶(T7gn1)转录。这些病毒聚合酶通过来自在lacUV5启动子的转录控制下位于前噬菌体上存在的T7gn1基因的宿主片段BL21(DE3)或HMS174(DE3)提供。

[0292] 在大肠杆菌中最大化重组蛋白质表达的实施之一是在具有蛋白水解切割重组蛋白质能力的受损宿主细菌中表达(Gottesman,p.119-128,In Gene Expression Technology:Methods in Enzymology vol.185,Academic Press,San Diego,CA,1990)。另一实施是核酸的核酸序列插入到表达载体中以致于在大肠杆菌中的那些独自的密码子都有各自的氨基酸(Wada et al.,1992,Nucleic Acids Res.20:2111-2118)。本发明氨基酸序列的这种改变可通过标准DNA合成技术实现。

[0293] 在另一个实施例中,表达载体是酵母表达载体。例如,用于表达的载体有酿酒酵母S包括pYepSec1(Baldari et al.,1987,EMBO J.6:229-234),pMFa(Kurjan and Herskowitz,1982,Cell30:933-943),pJRY88(Schultz et al.,1987,Gene54:113-123),pYES2(Invitrogen Corporation,San Diego,CA),和pPicZ(Invitrogen Corp,San Diego,CA)。

[0294] 另外,表达载体是杆状病毒表达载体。适用于蛋白质表达在培养昆虫细胞(如Sf9细胞)内的杆状病毒载体包括pAc系列(Smith et al.,1983,Mol.Cell Biol.3:2156-2165)和pVL系列(Lucklow and Summers,1989,Virology170:31-39)。

[0295] 在另一个实施例中,本发明的核酸是使用哺乳动物细胞中的哺乳动物表达载体来表达的。例如,哺乳动物表达载体包括pCDM8(Seed,1987,Nature329:840)和pMT2PC(Kaufman et al.,1987,EMBO J.6:187-195)。当用哺乳动物细胞时,表达载体的控制功能常常由病毒控制元件提供。例如,常用的启动子来自多瘤病毒,腺病毒2,巨细胞病毒和猿猴病毒40。对于其他的原核和真核细胞的合适表达系统详见Sambrook et al.,supra中的第16和17章节。

[0296] 在另一个实施例中,重组哺乳动物表达载体是在特定的细胞类型中指导核酸表达的能力(如组织特异性控制元件用于核酸表达)。组织特异性控制元件在文献中可知。非限制性例子,适合的组织特异性启动子包括白蛋白启动子(liver-specific;Pinkert et al.,1987,Genes Dev.1:268-277),淋巴特异性启动子(Calame and Eaton,1988,Adv.Immunol.43:235-275),如T细胞受体启动子(Winoto and Baltimore,1989,EMBO J.8:729-733)和免疫球蛋白(Banerji et al.,1983,Cell33:729-740;Queen and Baltimore,1983,Cell33:741-748),神经元特异性启动子(如神经丝启动子,Byrne and Ruddle,1989,Proc.Natl.Acad.Sci.USA86:5473-5477),胰腺特异性启动子(Edlund et al.,1985,Science230:912-916),和乳腺特异性启动子(如牛乳清启动子,美国专利号4873316(引用

参考)和欧洲专利公开号264166)。发展控制启动子也包括如小鼠同源盒蛋白启动子(Kessel and Gruss,1990,Science249:374-379)和-胎蛋白启动子(Camper and Tilghman,1989,Genes Dev.3:537-546)。

[0297] 本发明更进一步提供了重组表达载体,包括本发明反义方向克隆到表达载体里的DNA分子。也就是说,DNA分子以一定的方式合成连接到控制序列上,其允许本发明多肽的反义编码mRNA的RNA分子表达(通过DNA分子转录)。合成连接到反义方向克隆的核酸分子上的控制序列可有选择性的指导反义RNA分子在各种类弄的细胞中连续表达,比如,病毒启动子和/或增强子,或控制序列可有选择性的指导构成的,组织特异性的或细胞典型的反义RNA的特异性表达。反义表达载体可从重组质粒,噬菌体,或减弱病毒中得到,其反义核酸是在高效率控制区域的控制下生产,其活性可通过细胞类型引入载体中来确定。使用反义基因的基因表达控制论述详见Weintraub et al.,1986,Trends in Genetics,Vol.1(1)。

[0298] 本发明的另一个方面涉及到向宿主细胞引入本发明的重组表达载体。术语"宿主细胞"和"重组宿主细胞"在这里可互换。据了解,这些术语不仅涉及到特定的对象细胞还涉及到这些细胞的后代或潜在的后代。因为一些变化可在随后引起产生突变或环境影响,如实际上后代与母细胞不完全相同,但这些仍然包含在这里所用的术语范围内。

[0299] 宿主细胞可以是任何原核细胞(如大肠杆菌)或真核细胞(例如,昆虫细胞,酵母或哺乳动物细胞)。

[0300] 通过传统的转录或转染技术载体DNA可引入原核或真核细胞中。这里所用的术语"转录"和"转染"的本意是指通过各种文献已知的技术将外源核酸引入宿主细胞,包括磷酸钙或氯化钙的共同沉淀,DEAE-葡聚糖介导转染,脂转染,或电穿孔。用于转录或转染到宿主细胞的适合方法可在Sambrook,et al.(supra)和其他实验手册上找到。

[0301] 众所周知,对了哺乳动物细胞的稳定转染依赖表达载体和所用的技术,只有一小部分细胞的基因组可与外源DNA整合。为了识别和选择这些整合体,基因编码有选择性标记(如对抗生素具有抗性),通常是把所目标的基因引入宿主细胞。典型实施,把选择性标记引入那些对药物如G418,潮霉素和甲氨蝶呤药物具有抗性的细胞中。具有插入核酸的细胞稳定转录可通过选择药物来识别(如具有选择性标记的基因的细胞则生存下来,而其他细胞则死亡)。

[0302] 本发明的宿主细胞,如在培养液中的原核或真核宿主细胞可用来生产相当于本发明标记的多肽。因此,本发明进一步提供用本发明的宿主细胞来生产相当于本发明标记的多肽的方法。在一个实施例中,生产标记的方法包括培养合适媒介的本发明宿主细胞(引入本发明多肽编码的重组表达载体)。在另一实施例中,方法更进一步包括从媒介或宿主细胞中分离得到标记多肽。

[0303] 本发明的宿主细胞也可用来生产非人类转基因动物,例如,在一个实施例中,本发明的宿主细胞是在受精卵母细胞或胎儿干细胞中引入相当于本发明标记的编码多肽序列。随后这种宿主细胞可在他们的基因组中引入编码本发明标记的外源序列用来产生非人类转基因动物,或改变相当于本发明标记的编码多肽的内源基因序列来产生重组动物。这种动物可用来研究相当于标记的多肽功能和/或活性,用来识别和/或判断多肽活动调节器,与治疗的临床前试验或诊断分子一样用来发现或识别标记,如治疗和诊断标记的发现或识别,或药物效力和特异性的替换。

[0304] 这里所用的“转基因动物”是非人类动物即哺乳动物，如啮齿目动物，即大老鼠或老鼠，其一个或多个动物细胞含有转基因。另一个例子，基因改造动物包括非人类灵长目动物，羊，狗，牛，羊，鸡，两栖类动物等。转基因是外源DNA整合到细胞的基因组上，其来自生长的转基因动物和成熟动物的基因组的残体，因此在一个或多个细胞类型或转基因动物的组织中指导编码基因的表达。这里所用的“同源重组动物”是非人类动物，例如哺乳动物即老鼠，其通过将内源基因和引入动物细胞的外源DNA分子间同源整合来改变，即动物的胎儿细胞，动物的前期到壮大。转基因动物也包括在Chan I.T.,et al.(2004)J Clin Invest.113(4):528-38and Chin L.et al(1999)Nature400(6743):468-72中所描述的那些转基因动物。

[0305] 本发明的转基因动物可通过引入相当于本发明标记的多肽编码核酸到受精卵母细胞的雄性原核中来生产，即通过显微注射，逆转录病毒感染并使卵母细胞在假孕的雌性动物中生长。也可包括转基因的内含子序列和多腺苷酸信号来提高转基因表达效率。组织特异性控制序列可整合连接到转基因上来指导本发明多肽表达的特异性细胞。用于生产转基因动物的方法是胚胎控制和显微注射，特别是动物如老鼠变为传统在文献和如在美国专利号4,736,866(引用参考)和4,870,009(引用参考)，美国专利号4,873,191(引用参考)和在Hogan,Manipulating the Mouse Embryo,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.,1986中公开。类似的方法用于其他转基因动物的生产。转基因创造的动物可基于在它的基因组和/或动物组织或细胞中编码转基因的mRNA表达中存在的转基因来识别。然后转基因创造的动物可用来繁殖携带额外转基因的动物。此外，携带转基因的动物可进一步繁殖其他携带其他转基因的转基因动物。

[0306] 为了创造同源性重组动物，制备的载体至少含有相当于本发明标记多肽编码的部分基因来删除，增加或替换已被引入的来改变基因，即打乱功能。在另一实施例中，载体设计成这样，同源基因结合，内生基因功能打乱(即不再编码功能蛋白质，也称为“淘汰”载体)。载体可设计成这样，同源基因结合，内生基因突变或其他改变但仍然编码功能蛋白(即向上游控制区域的改变可因此改变内生蛋白质的表达)。在同源结合载体中，通过增加基因的核酸来充许被携带到载体的外源基因和胎儿干细胞中的内生基因间发生同源基国再结合来改变部分基因的5’和3’端的侧链。增加侧链核酸序列足够的长度来来达到内生基因的同源再结合的目的。通常情况下，DNA侧链(均在5’和3’末端)的几个碱基包括在载体中(详见Thomas and Capecchi,1987,Cell51:503for a description of homologous recombination vectors)。载体引入到胎儿干细胞系中(如通过电穿孔)并且在细胞中引入基因与内生基因选择性的同源再结合(详见Li et al.,1992,Cell69:915)。然后选择性细胞注射到动物的胎泡中(如老鼠)来形成聚合嵌合体(详见Bradley,Teratocarcinomas and Embryonic Stem Cells:APractical Approach,Robertson,Ed.,IRL,Oxford,1987,pp.113-152)。然后嵌合体胚胎可注入到适合的假孕雌性孕育动物中并胚胎成长至分娩。在他们的胚胎细胞中怀有同源再结合DNA的后代可用来繁殖动物，通过转基因的种类转录使得在所有的动物细胞中含有同源再结合DNA。用于构造同源再结合载体和同源再结合动物的方法在Bradley(1991)Current Opinion in Bio/Technology2:823-829和PCT公开号WO90/11354(引用参考)，WO91/01140(引用参考)，WO92/0968(引用参考)，和WO93/04169(引用参考)中有更详细描述。

[0307] 在另一个实施，转基因非人类的动物可采用转基因的允许控制表达的选择性系统来生产。一个例子，这样的系统是噬菌体P1的cre/loxP重组酶系统。cre/loxP重组酶系统的详细描述详见Lakso et al.(1992)Proc.Natl.Acad.Sci.USA89:6232-6236。另一例子，重组系统是酿酒酵母的FLP重组酶系统(O'Gorman et al.,1991,Science251:1351-1355)。如果cre/loxP重组酶系统是用来控制含有转基因编码所需要CRE重组酶和选择性蛋白质的转基因，动物的表达。这种动物可通过建造"双"转基因动物来提供，即通过两种转基因动物杂交，一个含有编码选择性蛋白质的转基因和另一个含有编码重组酶的转基因。

[0308] 克隆所述的非人类转基因动物也可按照在Wilmut et al.(1997)Nature385:810-813和PCT公开号WO97/07668(引用参考)和WO97/07669(引用参考)中所公开的方法来生产。

[0309] V.试剂盒的实施

[0310] 试剂盒是任意产品(如包装或容器)包括至少一种用于本发明标记的特异性检测试剂如探针，产品可作为一个用于本发明的操作方法的单元被推销，输送或销售。当本发明的化合物，试剂盒子和方法用来实现本发明方法时本发明的ALK突变基因和/或基因产品(如表1所列的标记)是可选择性的，以致于对具有癌症的患者对象的时期，阶段，组织类型或良性/预恶性/恶性类型的阳性结果获得至少约20%，至少约40%，至少约60%，至少约80%，至少约90%，至少约95%，至少约99%，或100%。在一些实施例中，本发明的标记或标记板可选择性的，以致于总体人口的PPV(阳性预测值)获得大于约10%(即外加特异性分析大于99.5%)。

[0311] 本发明的多元化ALK突变基因和/或基因产品(如表1所列的标记)用在本发明的化合物，试剂盒和方法中时，每个标记的数量，结构和/或活性或表达水平或拷贝数可与正常的每个多元化标记的数量，结构和/或活性或表达水平或拷贝数进行比较，在相同类型的非癌症样本中，要么是单一反应混合物(即用反应剂，如为每个标记用不同的荧光探针)要么是相当于一个或多个ALK突变基因和/或基因产品(如表1所列的标记)独立反应混合物。如果多元化ALK突变基因和/或基因产品(如表1所列的标记)被使用，那么2,3,4,5,6,7,8,9,10,或更多的独立标记可用或识别。

[0312] 本发明包括用于对样本(即存档组织样本或从对象上获得的样本)分析癌症细胞的化合物，试剂盒和方法。这些化合物，试剂盒和方法与上述的那些基本上相同，不论那些，哪里需要，该化合物，试剂盒和方法都适合用于某些类型的样本。

[0313] 因此，本发明包括用于判断有或可能有对ALK抑制剂降低敏感性的癌症细胞的试剂盒(即在如对象样本的样本中)。试剂盒由一个或多个有能力识别出本发明的ALK突变基因和/或基因产品(即表1所列的标记)的试剂组成，如与相当于本发明ALK突变基因和/或基因产品(即表1所列的标记)的核酸或多肽特异性粘合的包括抗体，抗体衍生物，抗体片段等等。适合与核酸(即基因组DNA，mRNA，粘接mRNA，cDNA或类似的)粘合的试剂包括互补的核酸。例如，核酸试剂包括固定到底物的核苷酸(标记或非标记)，没有与底物粘合的标记核苷酸，PCR引物，分子信标探针等。在一些实施例中，试剂盒含有操作这里所描述方法的有用试剂，如至少含有一种能识别和杂交到得延长核酸的引物，周围至少一种延长的核酸，至少包括一种表1所列的突变体和用于检测存在所说突变体扩增靶向核酸的方法。

[0314] 本发明的试剂盒可选择性的增加对本发明方法操作有用的合化物。通过例如方式，试剂盒可含有适用于互补核酸退火或用于用蛋白质其特异性粘合抗体的液体(如SSC缓

冲液)，一个或多个间隔样本，一本描述本发明方法操作的教学材料，一个正常细胞样本，一个癌症细胞样本等等。

[0315] 本发明的试剂盒可包括用于检测标记蛋白质水平或蛋白质活性有用的试剂。在另一实施例中，本发明的试剂盒可包括一用于检测标记的甲基化程度的试剂，或包括用于检测标记结构变化的试剂，如存在的突变体。

[0316] VI.预测医学

[0317] 本发明还涉及到预测医学，包括诊断检测，药物基因组学，临床试验，它们可以提前预防地治疗一个人。因此，本发明的一个方面涉及确定与本发明的一个或多个标记相对应的多肽或核酸的数量，结构和/或活性，从而确定一个患有癌症或具有罹患癌症风险的人是否会有更大的可能来反应ALK抑制剂药物疗法。

[0318] 因此，本发明的一个方面目标了一种确定癌症患者是否对一种ALK抑制剂疗法有反应的方法。本发明另一个方面目标了一种预测疾病时间过程的方法。本发明还有一个方法是目标一种预测疾病的时间过程中可能发生的重要事件的方法。在一些实施例中，所述的方法包括检测一个生物标记或检测生物标记与本发明所述的一种ALK抑制剂治疗的反应物的结合，并检测所述的对象是否会对所述的ALK抑制剂治疗产生反应。

[0319] 在一些实施例中，所述的方法涉及对诊断有癌症或怀疑有癌症(如存在癌症症状)的患者身上采集生物组织样本，进行细胞遗传学筛查，检测ALK突变(如表1所列的那些)。

[0320] 所述的筛查方法及其说明的结果能预测患者对ALK抑制剂(如PF-02341066和/或PDD)治疗的反应。如本发明所述，一种ALK突变的存在表明与那些未产生一种ALK突变的患者相比，ALK抑制剂(如PF-02341066和/或PDD)疗法能更有效地治疗癌症细胞。

[0321] 在一个实施例中，本发明所述的方法包括与一个DNA样本接触，如一个生殖细胞系或体细胞基因组DNA，如一个染色体样本，它是从患者上采集的细胞分离为多核苷酸探针，这些探针能特异地用于在严格条件下与基因组DNA杂交，使病人细胞中的染色体序列发生细胞遗传学异常(如本发明所述的ALK突变)。这些分析的结果能预测患者对治疗试剂，尤其是ALK抑制剂(如PF-02341066和/或PDD)治疗可能产生的反应。

[0322] 在另一个实施例中，确定患者疾病过程中发生重要事件的时间间隔，从而计算一个时间过程，其中所述的计算可以预测一个患者是否有一个很长的时间过程。在另一个初稿例中，所述的重要事件是从初次诊断发展到死亡。在另一个实施例中，所述的重要事件是从初次诊断发展到疾病移植。在另一个实施中，所述的重要事件是从初次诊断发展到恶化。在另一个实施例中，所述的重要事件是从疾病移植发展到死亡。在另一个实施例中，所述的重要事件是从疾病移植发展到恶化。在另一个实施例中，所述的重要事件是从恶化发展到死亡。在一些实施例中，所述的时间过程是用整个存活率，发展时间和/或使用RECIST或其它反应标准来测量的。

[0323] 在一些实施例中，至少两个患者群中的单独患者样本会形成一个预测测量。在一些实施例中，群数是两个，因此患者样本被分成一个有一个ALK突变的患者群和一个没有一个ALK突变的患者群。在一些实施例中，对象中的ALK突变程度既和有ALK突变的相比，又和没有ALK突变的相比；如果所述的患者具有一种ALK突变，然后患者不太可能对一种ALK抑制剂(如PF-02341066和/或PDD)产生反应和/或患者不太可能有一个很长的时间过程。在一些实施例中，群数大于两个，非限制性地包括三个群，四个群，五个群和六个群，根据预测ALK

抑制剂效果的分层与特别ALK突变的关系。在一些实施例中，可能的反应是用整个存活率，发展时间和/或使用RECIST标准来测量的。在一些实施例中，ALK抑制剂是PF-02341066和/或PDD。

[0324] 本发明的另一个方面目标了一种确定一个患有ALK突变阳性癌症的对象是否有可能对一种ALK抑制剂(如PF-02341066和/或PDD)产生反应和/或疾病的时间过程是否很长的方法。本发明另一个方面目标了一种方法，它可以用于预测患有ALK突变阳性癌症的对象中一个重要事件的可能性。

[0325] 1.检测ALK突变的方法

[0326] 本领域的普通技术人员公知能检测ALK基因突变和/或基因产物(如表1所列的标记)的方法包括杂交底物检测。例如，用于检测一个样本中编码核苷酸的拷贝数的方法包括一种Southern印迹法。在一种Southern印迹法中，所述的基因组DNA(通常散布分离在一种电泳凝胶上)与一个特异目标序列的探针杂交。比较目标序列的探针发出的强烈杂交信号与正常mRNA(如相同或相关细胞，组织，器官等中的一个非增强部分)分析发出的控制探针信号能检测出目标核苷酸是否存在，并检测出其拷贝数。或者可以使用一种Northern印迹法来检测一个样本中的编码核苷酸拷贝数。在一种Northern印迹法中，mRNA与特异目标序列的探针杂交。比较目标序列的探针发出的强烈杂交信号与正常mRNA(如相同或相关细胞，组织，器官等中的一个非增强部分)分析发出的控制探针信号能检测出目标核苷酸是否存在，并检测出其拷贝数。

[0327] 确定拷贝数的一种选择性的方法是原位杂交(如Angerer(1987)Meth.Enzymol152:649)。原位杂交通常包括下述步骤：(1)固定组织或生物结构用于分析；(2)预处理生物结构以提高目标DNA的可接近性，并减少非特异结合；(3)将核苷酸混合物与所述的生物结构或组织中的核苷酸杂交；(4)杂交后漂洗掉那些没有结合在杂交中的核苷酸片段；(5)检测杂交核苷酸片段。在上述每一个步骤中使用的试剂和条件可根据实际应用进行改变。

[0328] 优选的杂交底物检测非限制性包括现有的“直接探针”法，如Southern印迹法或原位杂交法(如FISH和FISH加SKY)，以及“比较探针”法，如比较基因组杂交(CGH)，例如cDNA底物或寡核苷酸底物CGH。所述的方法可以以各种方式使用，其非限制性地包括底物(如薄膜或玻璃)结合法或矩阵底物法。

[0329] 一方面是使用FISH分析。细胞样本可以根据本领域普通技术人员公知的方法从患者身上获得，用本领域普通技术人员公知的一种合适的细胞生成检测法来检测，例如FISH法。在一个实施例中，FISH可以根据Vysis™系统(Abbott分子)操作，其制作商的使用说明可以成为本发明的参考。

[0330] 为使用的探针含有DNA片段，其与部分染色体中存在的DNA底物序列互补。Bittner等发明的Vysis公司的两份美国专利，其专利号5,491,224(引用参考)和6,277,569(引用参考)公开了本发明使用的探针实例，并将这些探针标记杂交到样本中。

[0331] 染色体探针的长度通常约为50-105。更长的探针通常包括长为100-500核苷酸的更小片段。与着丝粒DNA和位点特异性DNA杂交的探针是可以商用的，例如用于Vysis公司(唐纳斯格韦，伊利诺斯州)，分子探针公司(柳真，俄勒冈州)或者Cytocell(牛津郡，英国)。探针也可以选择性地使用染色体组或基因组DNA通过标准技术制成，而不是商业制备。例

如，可以使用的DNA来源包括基因组DNA，克隆的DNA序列，含有一个或一个中的一部分的体细胞杂交，染色体（如人体染色体）和宿主的正常染色体互补和流式细胞仪或显微纯化的染色体。所述的插入序列可以通过克隆分离，或者用聚合物链式反应（PCR）的位点特异扩增分离。例如可参见Nath and Johnson，Biotechnic Histochem.，1998，73(1):6-22，Wheeless et al.，Cytometry1994，17:319-326和美国专利5，491，224（引用参考）。

[0332] 与一个染色体特异序列杂交的探针能确定这个序列是否存在细胞遗传学异常。细胞遗传学异常之一是一个删除。虽然这些删除可以是一个或多个完整染色体，但是它们通常涉及一个或多个染色体的部分缺失。如果探针中所含的一个染色体的整个序列从一个细胞中删除，那么这个探针通常不会再与细胞中的DNA杂交，于是染色体上就不会产生信号。如果一个局部含有一个探针的染色体序列从一个细胞中删除，那么这个探针还是会和这个细胞中的DNA杂交，但是信号会降低。例如一个信号的损失较之与控制细胞DNA进行的探针杂交不含有探针预计检测的基因异常。在一些实施例中，细胞遗传学异常会产生至少1，5，10，20，30，40，50，60，70，80，90，100，110，120，130，140，150，160，170，180，190，200或更多的细胞。

[0333] 能检测的细胞遗传学异常非限制性地包括非相互易位，内染色体倒位，点突变，缺失，基因拷贝数变化，基因表达水平变化和生殖系突变。一种特别的细胞遗传学异常是一种复制。这些复制可以是整个染色体组，也可以是比一个完整染色体更小的序列。如果一个含有一个探针的染色体序列在一个细胞中被复制，那么探针与这个细胞DNA杂交与不存在含探针的染色体序列的控制细胞中产生的信号数相比，通常会至少产生一个额外的信号。虽然检测人体染色体组2p23或其直系同源或任何染色体序列的任意探针含有一个2p23或其直系同源的ALK基因易位，但是它也可以使用。现有技术中已经公开了合适的探针（如可以使用Vysis公司（唐纳斯格韦，伊利诺斯州）生产的）。

[0334] 染色体探针是可以标记的，所以染色体序列的杂交是可以检测的。探针通常可以用一个荧光团，一个有机分子直接标记，其中有机分子能在吸收低波/高能光后发荣光。所述的荧光团使探针不需要一个第二检测分子就能变成可见的。一个荧光团与一个核苷酸共价结合后，所述的核苷酸可以直接用标准技术，如缺失转录，随机引物和PCR标记等实现与探针结合。或者可以用一个连接使脱氧核苷酸与探针促转氨。所述的荧光团可以与促转氨后的脱氧核苷酸共价结合。参见美国专利5，491，224（引用参考）。

[0335] 美国专利5，491，224公开了标记为一些胞嘧啶残基的探针具有一个荧光团标记共价键。荧光团标记胞嘧啶底物的数量足够在一个时就能够产生一个可检测的荧光信号，因此标记的DNA片段基本上能保留与能检测的染色体组或染色体序列相关的特异互补键（杂交）特性。这些探针可以这样制备：吸引未标记的DNA探针序列，用一个连接基团在序列中促转氨一些脱氧核苷酸，将一个荧光标记与至少一部分促转氨的脱氧胞苷底物共价结合。

[0336] 探针也可以使用缺失转录，随机引物或PCR标记来标记。标记既可以使用荧光团（直接）标记核苷酸，也可以使用半抗原（间接）标记核苷酸来完成。标记的非限制性典型实例包括：AMCA-6-dUTP，级联蓝-4-dUTP，荣光素-12-dUTP，罗丹明-6-dUTP，德克萨斯红-6-dUTP，Cy3-6-dUTP，Cy5-dUTP，生物素（BIO）-11-dUTP，地高辛（DIG）-11-dUTP或二硝基苯（DNP）-11-dUTP。

[0337] 探针也可以用生物素或地高辛间接地标记，也可以用放射性同位素，如32p和3H等

标记，虽然然后还是需要第二检测分子或进一步的处理来识别探针。例如，一个生物素的探针标记可以利用抗生素蛋白与一个检测标记结合来检测。例如，抗生素蛋白可以与一个酶标记结合，如碱性磷酸酶或辣根过氧化物酶。酶标记可以使用酶的一个底物和/或催化剂在标准比色反应中检测。碱性磷酸酶的催化剂包括5-溴-4-氯-3-吲哚基磷酸盐和硝基四氮唑蓝。二氨基苯甲酸可以作为辣根过氧化物酶的催化剂使用。

[0338] 探针也可以制备为使杂交前或杂交中，一个荧光团或其它标记不会是一部分DNA，而在杂交后增加检测这个探针杂交为一个染色体组。例如，使用的探针可以具有与DNA结合的抗原分子。杂交后，这些抗原分子可以使用特异抗体与抗原分子反应来检测。这些抗体自身可以与一个荧光团结合，或者可以使用一个具有一个共价荧光团的第二抗体检测。

[0339] 但是处理后或修饰后，所述的探针DNA可以在杂交使用前进行普通的纯化，以去除未反应的剩余产物(如没有和DNA结合的荧光团分子)。

[0340] 在杂交前可以使用本领域普通技术人员公知的方法使染色体探针变性。杂交步骤通常包括：向标记探针混合物中加入一种过量的封阻DNA，在杂交条件下将封阻探针与用于检测的染色体序列接触，如在一个固相中，其中的DNA已经变性，漂洗未杂交的探针，检测与染色体或染色体序列结合的探针。

[0341] 探针在杂交条件下杂交或退火为所述的染色体DNA。“杂交条件”是有利于一个探针与目标染色体DNA间退火的条件。虽然不同探针的退火根据探针长度，底物浓度等情况的不同而不同，现有技术已经公开了有利于退火的各种探针浓度，杂交温度，盐浓度和其它因素。

[0342] 各种浓度，底物混合物，复合物和探针长度，以及盐浓度，温度和保温时间都能促进杂交。例如，原位杂交通常在杂交缓冲液中进行的，所述的杂交缓冲液含有1-2x SSC，50-65%甲酰胺和防止产生非特异杂交的封阻DNA。上述杂交条件通常包括约25℃到55℃的温度，约0.5到96小时的保温时间。

[0343] 可以使用各种漂洗来去除染色体探针与目标序列外DNA之间形成的非特异性结合。改变每次漂洗的温度和盐浓度来严格控制这些漂洗。例如在非常严格的条件下，漂洗的温度约为65-80℃，使用0.2x至约2x的SSC，以及约0.1%-1%的非离子洗涤剂，如诺乃洗涤剂P-40

[0344] (NP40)。

[0345] 增加漂洗时的漂洗温度或盐浓度可以降低漂洗时的严格条件。在一些应用中需要封阻重复序列的杂交量。因此在一些实施例中可以使用tRNA，人体基因组DNA或Cot-I DNA来封阻非特异性杂交。

[0346] 漂洗后，所述的载玻片可以用完并空气干燥，然后加入媒介，使用一种染色液如DAPI和一个盖玻片。随即观察载玻片并在复查前储存在-20℃。

[0347] 为了在荧光原位杂交(FISH)技术中使用荧光探针，可以为每个荧光设置一个带荧光过滤器的荧光显微镜设备，这样就可以看到荧光，或者使用双或三通带过滤设备来观察多荧光。例如可参见美国专利5,776,688(引用参考)。或者也可以使用如流动血细胞计数等技术来检查染色体探针的杂交方式。可以使用FISH来检测染色体拷贝数或染色体的重测序列。这些探针与互补DNA的杂交或结合，由于它们标记了荧光标签，因此研究者可以使用一个荧光显微镜来观察这些DNA序列的位置。与染色体研究中使用的大多数其它技术不同，

FISH不需要细胞是活性分隔的，它是在未分区的细胞中进行的，因此是一个高效多功能的方法。因此FISH可以使用分区细胞来实现，或者使用分区细胞或细胞分区循环来实现。Gray和Pinkel的美国专利5,447,841公开了FISH分析中的许多技术。

[0348] FISH结果可以参考控制细胞来说明设计的探针检测，其中控制细胞是公知的不含有特异细胞遗传学异常。将所述的探针与控制细胞中DNA进行FISH杂交的方法与同种探针与未测试或评估特异性细胞遗传学异常的细胞中的DNA杂交进行比较。当一个设计的探针用于检测一个染色体或染色体序列的缺失时，通常所述的探针与未测试或评估特异性细胞遗传学异常的细胞中的DNA杂交比与控制细胞中的DNA杂交更少。在这些测试细胞中通常没有一个探针信号，说明探针通常与之杂交的一个染色体序列缺失。当一个设计的探针检测到一种染色体重复或增加，那么通常所述的探针与已经测试的细胞中的DNA杂交比与控制细胞中的DNA杂交更多。测试细胞中通常会有一个额外的探针信号，说明有一个额外的染色体序列，它通常会和探针杂交。

[0349] 在CGH方法中，核苷酸的一个第一聚集区(如从一个样本，如一个可能的组织)标有一个第一标记，当核苷酸的一个第二聚集区(如一个控制，如从一个健康的细胞/组织)标有一个第二标记。这些核苷酸的杂交率是由矩阵中两种(第一和第二)标记与每个光纤的结合率决定的。染色体缺失或复制处检测的信号率与两种标记处不同，这能检测拷贝数。矩阵底物CGH也可以使用单独的颜色标记来实现(使用两种染料来不同地标记控制和可能的组织样本，并在杂交前混合它们，由于竞争性杂交会在矩阵上产生一个比率)。在单独的颜色CGH中，可以将所述的控制标记杂交为可读的一个矩阵和完整信号，将可能的组织样本标记杂交为可读的一个第一矩阵(有识别内容)和完整信号。两个矩阵中的完整信号计算出不同的拷贝数。

[0350] 如在Albertson(1984)EMBO J.3:1227-1234;Pinkel(1988)Proc.Natl.Acad.Sci.USA85:9138-9142;欧洲公开号430,402;Methods in Molecular Biology,Vol.33:In situ Hybridization Protocols,Choo,ed.,Humana Press,Totowa,N.J.(1994)等已经公开了适于使用本发明所述方法的杂交方案。在一个实施例中，使用了Pinkel,et al.(1998)Nature Genetics20:207-211或Kallioniemi(1992)Proc.Natl Acad Sci USA89:5321-5325(1992)公开的杂交法。美国专利6,455,258中公开了矩阵底板CGH，其中所有的内容可作为本发明的参考。

[0351] 在另一个实施例中，可以使用扩增底物评估来测量存在与否和拷贝数。在这个放大底物评估中，所述的核苷酸序列作为扩增反应(如聚合酶链式反应(PCR))中的一个模板。在一个定量扩增中，扩增产物的量与原始样本中的模板量成比例。比较合适的控制，如健康组织可以测量所述的拷贝数。

[0352] 本领域的普通技术人员均公知“定量”扩增法。例如，定量PCR涉及使用相同的引物同时共同扩增一个已知质量的控制序列。这就提供了一个可以用来校准PCR反应的内部标准。Innis,et al.(1990)PCR Protocols,A Guide to Methods and Applications,Academic Press,Inc.N.Y.已经公开了定量PCR的具体方案。Ginzonger,et al.(2000)Cancer Research60:5405-5409已经公开了使用定量PCR分析在微卫星中测量DNA的拷贝数。基因的已经核苷酸序列足够使一个本领域的普通技术人员能惯常地选择扩增基因任何部分的引物。也可以在本发明所述的方法使用荧光定量PCR。在荧光定量PCR中，定量是基于

荧光信号的数量，如TaqMan和sybr基因。

[0353] 其它适宜的扩增法非限制性地包括连接酶链式反应(LCR)(参见Wu and Wallace (1989)Genomics4:560,Landegren,et al.(1988)Science241:1077,and Barringer et al.(1990)Gene89:117)，转录扩增(Kwoh,et al.(1989)Proc.Natl.Acad.Sci.USA86:1173)，自动持续的基因复制(Guatelli,et al.(1990)Proc.Nat.Acad.Sci.USA87:1874)，点PCR和连接适配物PCR等。

[0354] 杂合性缺失(LOH)图(Wang,Z.C.,et al.(2004)Cancer Res64(1):64-71;Seymour,A.B.,et al.(1994)Cancer Res54,2761-4;Hahn,S.A.,et al.(1995)Cancer Res55,4670-5;Kimura,M.,et al.(1996)Genes Chromosomes Cancer17,88-93)也可以用于识别扩增序列或删除序列。

[0355] 2.评估基因表达的方法

[0356] 标记表达的水平也可以评估。本发明所述的一个标记的表达也可以用各种公知的检测一个转录分子或蛋白的方法来评估。这些方法的非限制性实例包括用于检测分泌，细胞表面，细胞质或核蛋白质的免疫学方法，蛋白质分离纯化法，蛋白质功能或活性检测，核酸杂交法，核酸反转录法和核酸扩增法。

[0357] 在一些实施例中，一个特异基因的活性可以通过测量基因转录(如mRNA)，通过测量转录蛋白的质量或通过测量基因产物活性来表征。标记表达可以用各种方法来监控，包括通过检测mRNA水平，蛋白质水平或蛋白质活性等可以使用标准技术来测量的。检测的内容包括基因表达水平量(如基因组DNA，cDNA，mRNA，蛋白质或酶活)或者可以是基因表达水平的一个评估量，尤其是与一个控制水平相比。这种被检测水平的类型可以从上下文中推出。

[0358] 本领域的普通技术人员(见Sambrook等supra)均公知那些使用核苷酸杂交技术来检测和/或对基因转录定量(这里的mRNA或cDNA制备)的方法。例如一种评估cDNA是否存在及其量的方法包括本发明所述的Southern印迹法。简而言之，所述的mRNA被分离(如使用一种酸胍-苯酚-氯仿，Sambrook等supra)且反转录为产物cDNA。然后所述的cDNA选择性地吸收并在缓冲液中跑凝胶，移植到薄膜。于是使用与目标cDNA特异的核苷酸探针能完成杂交。

[0359] 这些诊断和预测评估的一个基本原理包括制备一种样本或反应混合物，它们可以含有一个标记和探针来相互作用和结合，从而形成一个复合物，它能从反应混合物中移动和/或删除。这些评估可以用各种方法实施。

[0360] 例如，其中一种实现评估的方法是将标记或探针固定到一个固相支撑上，还是作为一个底物，并在反应结束时检测固定在固相上的目标标记/探针。在这种方法的一个实施例中，从对象上采集样本，评估其是否有标记及标记的浓度，并将其固定到一个载体或固相支撑上。在另一个实施例中，可以是相反的情况，其中探针固定在一个固相上，而对象上采集的一个样本可以作为评估中一个未固定成分反应。

[0361] 现有技术中公开了许多方法将评估成分固定到一个固相上。它们非限制性地包括将结合的生物素和抗生蛋白链菌素固定的标记或探针。这种生物素化的评估成分可以利用现有技术中公知的技术(如生物素化试剂盒，biotinylation kit，皮尔斯化工，罗克福德，伊利诺斯州)由生物素-NHS(N-羟基-琥珀酰亚胺)制备，并在抗生蛋白链菌素涂层96孔板的孔处固定(皮尔斯化工)。在一些实施例中，这些固定了评估成分的表面可以事先制备并保

存。

[0362] 这些评估中使用的其它合适载体或固相支撑包括任何可以与标记或探针所属级分子结合的材料。公知的支撑或载体非限制性地包括玻璃，聚苯乙烯，尼龙，聚丙烯，聚乙烯，葡聚糖，淀粉酶，天然和改性纤维素，聚丙烯酰胺，辉长岩和磁铁矿。

[0363] 为了使用上述方法实现评估，未固定的成分被加入到已经固定了第二成分的固相中。反应完成后，未复合的成分在条件下被去除(如漂洗)，使任何复合成分固定地保留在固相上。可以使用本发明所述的一些方法将固定在所述固相上的标记/探针复合物删除。

[0364] 在另一个实施例中，当所述的探针是未固化的评估成分时，它可以标有本发明所述的本领域普通技术人员公知的可检测标记，从而可以直接或间接地用于实现评估及读取评估结果。

[0365] 也可以直接检测标记/探针复合物成分，而不需要再操作或标记两种成分(标记或探针)，例如利用荧光能转移技术(例如参见Lakowicz等的美国专利5,631,169(引用参考))。选择第一“捐献”分子上的一个荧光标记，于是使用适当波长的光激发，所发射的荧光能被一个第二“接受”分子上的一个荧光标记吸收，由于所述的吸收能所以可以依次发荧光。或者所述的“捐献”蛋白分子可以简单地利用色氨酸残基的天然荧光能。所选的标记可以发出不同波长的光，使“接受”分子标记可以与“捐献”的不同。既然标记间转移的能效与分子分离距离有关，那么分子间的空间间隔也可以被评估。当分子间产生结合时，评估中“接受”分子发出的荧光持续的时间最长。一个FET结合情况甚至可以用公知的标准荧光检测法来进行常规检测(如使用一种荧光剂)。

[0366] 在另一个实施例中，确定一个探针识别一个标记的能力可以不需要标记或评估成分(探针或标记)来实现，而是利用一种技术，如实时生物分子相互作用分析(BIA)(参见如Sjolander,S.and Urbaniczky,C.,1991,Anal.Chem.63:2338-2345and Szabo et al.,1995,Curr.Opin.Struct.Biol.5:699-705)。如本发明所述，“BIA”或“表面等离子体共振”是一种实时研究生物特异性相互作用的技术，而没有标记任何相互作用(如BIA克隆)。结合表面(一个结合结构的相互作用)上的改变量会使表面附近的光折射率(表面等离子体共振的光学现象(SPR))发生变化，产生一个可检测的信号，它能作为一种生物分子反应的实时指示剂使用。

[0367] 或者在另一个实施例中，相似的诊断和预先的评估可以在一种液体相中使用标记和探针作为溶剂。在这样一种评估中，可以用任何一些标准技术，非限制性地包括差速离心，层析，电泳和免疫沉淀来分离所述的复合标记和探针。在差速离心中，标记/探针复合物可以使用一系列离心步骤从未复合的评估成分中分离出来，因为复合物的不同大小和密度使得它们的沉降平衡也不相同(例如参见Heegaard,N.H.,1998,J.Mol.Recognit.Winter11(1-6):141-8;Hage,D.S.,and Tweed,S.A.J Chromatogr B Biomed Sci Appl19970ct

[0368] 10;699(1-2):499-525)。标准层析技术也可以用于从未复合物中分离出复合分子。例如，凝胶过滤层析根据分子大小来分离分子，其在一个柱中使用一种合适的凝胶过滤树脂，例如较大的复合物可以从较小的未复合成分中分离出来。类似地，与未复合成分相比，具有不同电荷性能的标记/探针复合物可以被利用为不同于未复合成分的复合物，例如通过使用离子交换层板树脂。本领域普通技术人员公知这些树脂和层析技术(例如参见Heegaard,N.H.,1998,J.Mol.Recognit.Winter11(1-6):141-8;Hage,D.S.,和Tweed,S.A.J

Chromatogr BBiomed Sci Appl19970ct10;699(1-2):499-525)。凝胶电泳也可以将复合评估成分从未结合的成分中分离出来(例如参见Ausubel et al.,ed.,Current Protocols in Molecular Biology,John Wiley&Sons,New York,1987-1999)。例如在这个技术中是根据蛋白和核苷酸复合物的大小和电荷来分离它们的。为了使电泳过程中保持持续的结合作用,没有减少试剂的未变性凝胶基质材料和条件是典型的。本领域的普通技术人员公知特定评估的适当条件和成分。

[0369] 在一个优选的实施例中,mRNA相当于标记的水平可以使用本领域公知的方法在原位或在试管中确定。术语"生物样本"可以包括组织,细胞,生物液体及其分离物,它们从一个对象身上采集获得,和一个对象身上存在的组织,细胞和液体一样。在试管法中,任何RNA分离技术不会选择防止mRNA分离,它能用于从细胞中纯化RNA(例如参见Ausubel et al.,ed.,Current Protocols in Molecular Biology,John Wiley&Sons,New York1987-1999)。此外,可以准备大量的组织样本,并使用本领域普通技术人员公知的方法来处理它们,如彼得亚雷一步RNA分离法(1989,美国专利4,843,155(引用参考))。

[0370] 可以在杂交或扩增评估中使用的分离核苷酸非限制性地包括Southern或Northern分析,聚合物链式反应分析和探针矩阵。一种检测mRNA水平的诊断方法包括将分离的mRNA与一个核苷酸分子(探针)分离,然后与检测基因编码的mRNA杂交。所述的核苷酸探针可以是例如一种完整cDNA或其一部分,如长度至少为7,15,30,50,100,250或500核苷酸的一种寡核苷酸,它足够能在严格条件下特异杂交为一个编码本发明所述的一个标记的mRNA或基因组DNA。其它合适的探针也可以在本发明所述的诊断评估中使用。一个带探针的mRNA的杂交说明问题中的标记被表达。

[0371] 在一个实例中,将mRNA固定在一个固体表面并与一个探针接触,例如在一个琼脂糖凝胶上跑分离的mRNA,并将mRNA从凝胶转移到薄膜上,如硝化纤维。在另一个实例中,将探针固定在一个固体表面上,将mRNA与探针接触,例如在一个Affymetrix的基因芯片阵列。一个普通技术人员可以将公知的mRNA方法用于检测mRNA水平,其中所述的mRNA是由本发明所述的的标记编码的。

[0372] 所述的探针可以是全长核苷酸序列或比全长小的核苷酸序列编码的蛋白。更短的探针可以凭经验地检测特异性。例如优选的核苷酸探针有20个碱基或长度更长(例如参见Sambrook et al.for methods of selecting nucleic acid probe sequences for use in nucleic acidhybridization)。杂交蛋白可以定量地检测是否有cDNA。

[0373] 一种确定样本中与本发明所述的一个标记对应的转录水平的选择性方法包括核苷酸扩增法,如利用rtPCR(Mullis,1987的美国专利4,683,202所列的实验例(引用参考)),连接酶链反应(Barany,1991,Proc.Natl.Acad.Sci.USA,88:189-193),自动持续的序列复制(Guatelli et al.,1990,Proc.Natl.Acad.Sci.USA87:1874-1878),转录扩增系统(Kwoh et al.,1989,Proc.Natl.Acad.Sci.USA86:1173-1177),Q-β复制酶(Lizardi et al.,1988,Bio/Technology6:1197)滚动循环复制(Lizardi等,美国专利5,854,033(引用参考))或其它任何核苷酸扩增法,随后通过使用本领域普通技术人员公知的技术来检测扩增分子。荧光rtPCR也可以在本发明所述的方法中使用。在荣光rtPCR中,定量是根据荧光信号的量,如荧光定量和荧光染料。这些检测方案尤其可以用于检测数量非常小的核苷酸分子。如本文所述,扩增引物被确定为一对核苷酸分子,它们能退火到基因的5'或3'端(分别增加和

减少链，或反之亦然），并且之间含有一个短序列。扩增引物通常长度为10-30个核苷酸，其两侧有一个长约50-200核苷酸序列。在扩增条件和扩增试剂中，这些引物使一个核苷酸分子的扩增包括引物两侧的核苷酸序列。

[0374] 在原位法中mRNA不需要在检测前从细胞中分离出来。在这些方法中，可以使用公知的杂交方法来制备/获得一个细胞或组织样本。然后将所述的样本固定在一个支撑上，通常为一个玻璃固体，然后与一个探针接触，这个探针能与编码所述标记的mRNA杂交。

[0375] 由于需要根据所述标记的限制表达水平来进行选择性地预测，因此也可以根据所述标记的规范表达水平来确定。表达水平可以通常修正一个标记的规范表达水平来规范，将其表达水平与非标记的基因表达水平相比较，如一个管家基因等可以表达的。规范的适宜基因包括管家基因，如肌动蛋白基因或上皮细胞特异性基因。这种规范使得一个样本，如一个对象样本与另一个样本，如一个未患癌症样本中的表达水平可以被比较，或者是可以比较不同来源样本间的表达水平。

[0376] 或者表达水平也可以作为一个相对的表达水平。为了确定一个标记的相对表达水平，在确定问题样本的表达水平之前可以用10个或更多的规范样本与分离的癌症细胞样本进行比较获得所述标记的表达水平，或者甚至是使用50或更多样本大量样本中评估的每个基因的平均表达水平，并将其作为所述标记的一个基础表达水平。可以确定检测样本（限制表达水平）的所述标记表达水平于是被标记获得的平均表达值所区分。这产生了一个相对表达水平。

[0377] 在一些实施例中，基础确定中使用的样本可以取自同种组织类型的癌症细胞或规范细胞。细胞来源可以根据所使用的相对表达水平来选择。使用规范组织中发现的表达作为一种平均表达成分酸是确认评估的标记是否特异于所述的从细胞中分离出的组织（较之规范细胞）。此外当计算更具体信息时可以修改平均表达水平，并在计算后的信息基础上提供更好的相对表达值。规范细胞中获得的表达值可以形成一种方法来对癌症情况的严重程度分级。

[0378] 在另一个实施例中是通过从一个对象样本的细胞中制备基因组DNA或mRNA/cDNA（如转录多核苷酸），并将所述的基因组DNA或mRNA/cDNA与一个对照多核苷酸杂交来评估一个标记的表达，这个多核苷酸是否个完整的多核苷酸，其含有所述的标记及标记片段。在用对照多核苷酸扩增前，可以使用任意聚合酶链式反应来选择性地扩增cDNA。一个或多个标记的表达可以使用定量PCR（QPCR）来评估这个标记的表达水平。

[0379] 在一个相关的实施例中，将样本中获得的多核苷酸转录混合物与一个底物接触，该底物中固定了本发明所述一个标记中的一个完整的多核苷酸或至少一部分（如至少7，10，15，

[0380] 20，25，30，40，50，100，500或更多核苷酸残基）。如果本发明所述一个标记中的一个完整的多核苷酸或至少一部分可在底物上发现不同之处（如使用不同的染色体或荧光进行检测，或者固定到不同的选择性位置），那么若干个标记的表达水平可以使用一个单独的底物同时进行评估（如固定在选择性位置上的多核苷酸的“基因芯片”微矩阵）。如果使用一种能评估标记表达水平的方法，其包括将一个核苷酸与另一个杂交，那么所述的杂交可以在严格杂交条件下完成。

[0381] 在另一个实施例中，可以使用一个组合物的方法来评估一个标记的表达。

[0382] 由于本发明所述的组合物，试剂盒和方法是依赖于检测本发明所述的一个或多个标记的表达水平或拷贝数的差别，在一些实施例中，所述标记的表达水平或拷贝数会明显大于所述方法的最小检测底限，其中所述的方法是用于评估至少一个规范细胞和癌症细胞中的表达或拷贝数。

[0383] 3.评估表达蛋白的方法

[0384] 一个标记蛋白的活性或水平也可以通过检测或定量其表达的多肽来检测和/或定量。所述的多肽可以用任何一种本领域普通技术人员公知的方法来检测和定量。这些可以包括如电泳，毛细管电泳，高效液相色谱法(HPLC)，薄层色谱(TLC)，过扩散色谱法等分析生物学方法，或如液体或凝胶沉淀反应，免疫(单或双)，免疫电泳，放射免疫分析法(RIA)法，酶联免疫分析法(ELISA试剂盒)，免疫检测分析法，免疫印迹法，免疫组织化学法等各种免疫学方法。一个技术人员将公知的蛋白/抗体检测法用于检测细胞是否表达了本发明的一个标记。

[0385] 本发明所述的用于检测一种多肽的另一种试剂是一种抗体，其可以与一种多肽结合，该多肽与本发明的一个标记相对应，如一种带一个可检测标记的抗体。这些抗体可以多克隆或单克隆。一个完整抗体或它的一个片段(如Fab或F(ab')2)可以使用。与探针或抗体相关的术语“标记的”指通过与直接标记的另一个试剂反应，与直接标记探针或抗体相同地将一个可检测底物与探针或抗体连接(如物理连接)，使探针或抗体被直接标记。直接标记的实例包括使用一个荧光检测一个初级抗体，这个荧光标记了次级抗体标记和一个带生物素的DNA探针的末端标记，因此它可以用荧光标记的抗生蛋白链菌素来检测。

[0386] 在另一个实施例中，所述的抗体是标记了的，如一个无线电标记，染色体标记，荧光标记或酶标记的抗体。在另一个实施例中，一个抗体衍生物(如一个抗体与一个底物结合，或与一个蛋白配位体对中的蛋白或配位体结合{如生物素抗生蛋白链菌素})，或与一个抗体片段结合(如一个单链抗体，一个单独的抗体高突变区)，所述的片段能与一个标记对应的蛋白特异结合，如蛋白由与标记对应的开放读取片段编码的，或者这种蛋白有在全部或部分经过后转录修饰后被使用。

[0387] 免疫组织化学或IHC指利用抗原与生物组织中的抗原特异结合的原理，将抗原(如蛋白)定位在一个组织部分的细胞中的方法。免疫组织化学着色可以广泛地在异常细胞的诊断中使用，例如能在癌症肿瘤中找到。特异分子标记是特别多孔物的特征，如增殖或细胞死亡(凋亡)。IHC也可以广泛地用于研究理解生物标记和不同表达的蛋白在一个生物组织的不同部分中的分布与定位。设想一种抗体抗原相互作用可以用许多方法实现。在最常用的例子中，一种抗本与一种酶结合，如过氧化物酶，可以催化一种生成有色产物的反应。或者该抗体也可以用一个荧光团标记，如荧光素，罗丹明，DyLight Fluor或Alexa Fluor。

[0388] 细胞中的蛋白可以使用本领域普通技术人员公知的技术分离。该使用的蛋白分离法例如是那些Harlow和Lane所公开的(Harlow和Lane，1988，抗体：一种实验指南，冷泉港实验室出版社，冷泉港，纽约)。

[0389] 在一个实例中，这些抗体或抗本片段可以在各种方法中使用，如Western印迹法或免疫荧光技术来检测表达的蛋白。在这些使用中，可以将抗体或蛋白固定在一个固体支撑上。合适的固相支撑或载体包括任何可以与一个抗原或一个抗体结合的支撑。公知的支撑或载体包括玻璃，聚苯乙烯，尼龙，聚丙烯，聚乙烯，葡聚糖，淀粉酶，天然和改性纤维素，聚

丙烯酰胺，辉长岩和磁铁矿。

[0390] 本领域的普通技术人员公知许多其它的可以与抗体或抗原结合的合适载体，并可以针这些支撑在本发明中使用。例如，细胞中分离出来的蛋白可以跑一种聚丙烯酰胺凝胶，并固定在一个固相支撑上，如硝基纤维素。于是所述的支撑可以在用可检测的标记抗体处理后再用合适的缓冲液漂洗。于是可以用缓冲液漂洗所述的固相支撑两次以去除未结合的抗体。固相支撑上结合的标记量于是可以用现有的方法检测出来。使用电泳技术检测蛋白质的方法是本领域普通技术人员公知的(大致参见R.Scopes(1982)Protein Purification,

[0391] Springer-Verlag,N.Y.;Deutscher,(1990)Methods in Enzymology Vol.182:Guide to Protein Purification,Academic Press,Inc.,N.Y.)。

[0392] 在另一个实施例中，Western印迹(免疫印迹)分析可以用于检测样本中是否有一种多肽及其含量。这种技术通常包括利用根据分子质量进行的凝胶电泳使样本蛋白质分离出来，将分离出来的蛋白质转移到一个合适的固相支撑上(如一种醋酸纤维素膜，一种尼龙过滤器或一种衍生的尼龙过滤器)，将所述的抗体孵化到样本中，所述的抗体与一种多肽特异地结合。所述的抗多肽抗体与固相支撑上的多肽特异地结合。这些抗体可以使用这些标记的抗体(如标记羊抗人抗体)进行直接标记或者选择性地被整个检测，所述的标记抗体能与抗多肽特异地结合。

[0393] 在另一个实施例中，可以使用一种免疫测定来检测多肽。如本发明所述，一种免疫测定是一种评估，它利用一种抗体与被分析物产生特定结合。因此对一种多肽与一种抗体之间特定的结合进行检测，这种检测可以表征所述的免疫测定，这不同于分离，靶向或定量分析物的其它物理或化学特性。

[0394] 对所述多肽的检测和/或定量是使用任意一种良好识别的免疫结合评估(例如参见美国专利4,366,241(引用参考)；4,376,110(引用参考)；4,517,288(引用参考)；和4,837,168(引用参考)。了解通常的免疫可以参见Asai(1993)Methods in Cell Biology Volume37:Antibodies in Cell Biology,Academic Press,Inc.New York;Stites&Terr(1991)Basic and Clinical Immunology7th Edition。

[0395] 免疫结合评估(或免疫测定)通常使用一种“捕捉剂”来特异地结合分析物并常常固定分析物(多肽或序列)。所述的捕捉剂是一个能与分析物特异结合的成分。在另一个实施例中，所述的捕捉剂是一种抗原，它能特异地结合一种多肽。所述的抗体(多肽)可以由任意一种本领域普通技术人员公知的方法来制备。

[0396] 免疫测定通常也可以使用一种标记剂来特异地结合并标记所述的结合复合物，它是由捕捉剂和分析物形成的。所述的标记剂自身可以是一种包括抗体/抗原复合物的成分。因此，所述的标记剂可以是一种标记的多肽或一种标记的抗体。或者所述的标记剂可以是一个第三成分，如另一种抗体，它能与所述的抗体/多肽复合物特异地结合。

[0397] 在一个实施例中，所述的标记剂是一种带标记的第二人体抗体。或者所述的第二抗体可以缺失一个标记，但是它可以转而被一个标记的第三抗体结合，所述的第三抗体特异于第二抗体衍生特种的抗体。所述的第二抗体可以是用一个可检测的成分修饰的，如生物素，这样它就能被一种第三标记分子特异地结合，如酶标记的抗生蛋白链菌素。

[0398] 其它能与免疫球蛋白的恒定序列特异结合的蛋白，如蛋白A或蛋白G也可以作为标

记剂使用。这些蛋白是链状球菌细胞壁的普通成分。它们与各种特种的免疫球蛋白恒定序列产生一种强烈的免疫原反应(大致参见Kronval,et al.(1973)J.Immunol.,111:1401-1406,and Akerstrom(1985)J.Immunol.,135:2589-2542)。

[0399] 如上所述,检测和/或定量一种多肽的免疫测定可以使用本领域普通技术人员公知的各种方法。

[0400] 检测一种多肽的优选免疫测定可以是竞争性或非竞争性的。非竞争性免疫测定可以评估能直接检测的捕捉分析物的量。例如在一个"三明治"评估中,所述的捕捉剂(抗多肽抗体)可以与一个固相底物直接结合,并固定在底物上。于是这些固定的抗体可以捕捉测试样本中的多肽。因此所述的固定多肽于是与一个标记剂结合,如一种带标记的第二人体抗体。

[0401] 在竞争性评估中,样本中所含的分析物(多肽)量是通过检测一种添加(外源)分析物(多肽)的量来直接检测,该添加分析物可以用样本中所含的分析物从一种捕捉剂(抗肽抗体)中替换出来(或竞争)。在一个竞争性评估中,这时一种公知量的多肽可以加入到所述的样本中,使接本与一种捕捉剂接触。与抗体结合的多肽量与样本中所含的多肽浓度成反比。

[0402] 在另一个实施例中,所述的抗体是固定在一个固相底物中。测量一种多肽/抗体复合物中所含的多肽量或者选择性地测量残留的未复合多肽量可以确定与抗体结合的多肽量。多肽量可以通过设置一个标记的多肽来检测。

[0403] 本发明所述的评估可以使用本领域普通技术人员公知的标准方法来完成(作为多肽的阳性或阴性或定量)。完成的特定方法是由评估形式和选择标记来完成的。例如,酶标记形成的有色产物可以用于实现一种免疫印迹分析。当存在一种清晰可见的印迹或结合可以产生阴性结果时,正确分子量的一种清晰可见的有色结合或印迹可以产生阳性结果。强烈的结合或印迹可以定量地检测多肽。

[0404] 本发明所述的在各种免疫测定中使用的抗体可以如本发明所述的方法制备。

[0405] 在另一个实施例中,水平(活性)是通过测量基因产物的酶活而评估出来的。评估一种酶活的方法是本领域普通技术人员公知的。

[0406] 检测一个标记蛋白的原位技术包括向一个对象中引入一个抗蛋白检测的标记抗体。例如,所述的抗体可以标记上一个放射性标记,其存在定位于一个对象中,可以用标准成像技术来检测。

[0407] 由本发明所述的方法鉴定的一些标记可以是分泌蛋白。技术人员能简单地确定任何特定标记蛋白是否是一种分泌蛋白。为了实现这个确定,标记蛋白在一个哺乳动物细胞,如一个人细胞系中表达,收集胞外液体,评估胞外液体是否含有所述的蛋白(例如使用一个能与蛋白特异结合的标记抗体)。

[0408] 下文中所述的一种方法实例可以用于检测一种蛋白分泌。大约8x105293T细胞可以在生长培养基(杜尔贝科改进的伊格尔培养基{DMEM}其加有10%胎牛血清)培养,培养条件是37°,通入5%(v/v)CO2,95%空气进入约60-70%混合物中。然后使用一种标准转染混合物转染所述的细胞,其中每个孔的混合物中含有2微克的DNA和10微升的LipofectAMINE™(GIBCO/BRL目录号18342-012),其中所述的DNA具有一种编码蛋白的表达载体。所述的转染混合物放置约5小时,然后换为新鲜的生长培养基,并放置在空气中。用DMEM缓缓地冲洗每

个孔两次，其中DMEM中不含有蛋氨酸或半胱氨酸(DMEM-MC;ICN目录号16-424-54)。再向每个孔中加入约1毫升的DMEM-MC和约50微居里Trans-$^{35}S^{TM}$反应物(ICN目录号51006)。向这些孔通入5%CO_2气体，并在37℃下培养一个选择性的时间。然后在随后的培养中去除150微升的条件培养基，并离心去除浮动细胞和碎片。浮在表面上的蛋白即说明蛋白被分泌。

[0409] 应明白这些对象样本，如一个含痰，支气管肺泡灌洗，胸腔积液，组织，全血，血清，血浆，口腔刮，唾液，脑脊液，尿液，大便，骨髓的样本中可以含有细胞，尤其是当这些细胞可以是癌症的，尤其是当癌症是转移性的，因此可以在本发明所述的方法中使用。在评估所述样本中的标记表达水平之前，可以使用各种公知的后收集制备和储存技术(如核苷酸和/或蛋白提取，固定，储存，冷冻，超滤，浓缩，蒸发，离心等)来处理所述的细胞样本。因此本发明所述的组合物，试剂盒和方法可以用于检测与蛋白质对就的标记表达，这些蛋白质中至少有一种位于其表达的细胞表面。技术人员能简单地确定与任意特定标记对应的蛋白是否包括一个细胞表面蛋白。例如免疫学方法可以用于检测整个细胞中的这些蛋白，或者是公知的计算机支撑的序列分析法(如SIGNALP程序;Nielsen et al.,1997,Protein Engineering10:1-6)可以用于预测至少一种胞外域的存在(如包括两种分泌蛋白，这些蛋白至少有一个胞外域)。一个标记的表达可以不用裂解细胞就能检测，与这个标记对应的一个蛋白质至少有一部分是位于其表达的一个细胞的表面(例如使用一种标记的抗体，它与所述蛋白的一个胞外域特异地结合)。

[0410] 本发明还包括试剂盒，它们能检测一个生物样本，如一个含痰，支气管肺泡灌洗，胸腔积液，组织，全血，血清，血浆，口腔刮，唾液，脑脊液，尿液，大便，骨髓的样本中是否含有与本发明标记对应的一个多肽或核苷酸。这些试剂盒可以用于确定一个对象是否患有癌症或具有罹患癌症的风险。例如所述的试剂盒可以包括一个标记的化合物或试剂，它能检测一种多肽或mRNA，这种多肽或mRNA是编码与一个生物样本中本发明所述的标记对应的多肽，还包括确定样本所含的多肽或mRNA量的成分(如一种与多肽结合的抗体或一种与编码多肽的DNA或mRNA结合的寡核苷酸探针)。

[0411] 例如为了抗体底物的试剂盒，所述的试剂盒包括：(1)一个第一抗体(如附在一个固相支撑上)，它与多肽结合，所述的多肽与本发明所述的一种标记对应；并选择性的(2)一种第二不同抗体，它与多肽或第一抗体结合，并与一个可检测的标记共轭。

[0412] 例如为了抗体底物的试剂盒，所述的试剂盒包括：(1)一个寡核苷酸，如一个可检测标记的寡核苷酸，它与一个核苷酸序列杂交，这个核苷酸序列编码一个与本发明所述标记对应的多肽或(2)一对用于扩增一个核苷酸分子的引物，其中所述的核苷酸分子与本发明所述的一个标记对应。该试剂盒也可以包括，如一种缓冲剂，一种防腐剂或一种蛋白稳定剂。所述的试剂盒还包括那些检测可检测标记(如一种酶或底物)的必要成分。该试剂盒还含有一种控制样本或一系列控制样本，它们可以被评估并与测试样本对比。试剂盒中的每个成分可以用一个单独容器来封装，所有的容器和使用试剂盒所需的说明结果的说明书一起放入一个单独的包装中。

[0413] 4.评估结构改变的方法

[0414] 本发明还包括一种评估是否有一个结构改变，如突变的方法。

[0415] 另一种检测方法是探针与等位基因特异性杂交，它使用与多态位点重叠的探针，并在所述的多态区周围有约5，10，20，25或30个核苷酸。本发明的另一个实施例中，可以与

突变特异性杂交的各种探针是附在一个固相支撑上，如一个“芯片”。寡核苷酸可以用许多方法，包括平版印刷术来与一种固相支撑结合。例如一个芯片上可以有250,000条寡核苷酸(基因芯片，Affymetrix™)。例如Cronin et al.(1996)Human Mutation7:244公开了突变检测分析，它使用了包括寡核苷酸在内的这些芯片，也称之为“DNA探针矩阵。在一个实施例中，一个芯片包括至少一种基因多态区中的所有突变。然后所述的固相支撑与一个测试核苷酸接触，并与待检测的特异性探针杂交。因此，一个或多个基因上多个突变的密度可以用一个样本杂交实验来鉴定。例如可以用一个单独的杂交实验来鉴定5’上游控制因子中的核苷酸多态突变。

[0416] 在其它检测方法中，在鉴定突变之前需要首先扩增至少一部分标记。扩增可以使用本领域公知的方法，例如用PCR和/或LCR(见Wu and Wallace(1989)Genomics4:560)来实现。在一个实施例中，一个细胞的基因组可以裂解为两条PCR引物，并扩增许多循环直到其产物的量能达到扩增DNA需要的量。在一些实施例中，这些引物是在150和350碱基对之间。

[0417] 选择性扩增法包括：自维持序列复制(Guatelli,J.C.et al.,(1990)Proc.Natl.Acad.Sci.USA87:1874-1878)，Q-β复制酶(Lizardi,P.M.et al.,(1988)Bio/Technology6:1197)和自维持序列复制(Guatelli et al.,(1989)Proc.Nat.Acad.Sci.87:1874)和核苷酸碱基序列复制(NABSA)或其它任何核苷酸扩增法，随后通过使用本领域普通技术人员公知的技术来检测扩增分子。这些检测方案尤其可以用于检测那些低含量的核苷酸分子。

[0418] 在一个实施例中，本领域公知的任何测序反应都可以直接用于至少一部分标记的测序，

[0419] 并将样本序列与对应的相关(控制)序列对比，检测突变。优选的测序反应包括那些基于美加净和吉尔伯特改进的技术(Proc.Natl Acad Sci USA(1977)74:560)or Sanger(Sanger et al.(1977)Proc.Nat.Acad.Sci74:5463)。当进行对象评估时也可以使用任意排充程序(Biotechniques(1995)19:448)，包括用定量光谱测定法测序(例如参见美国专利5,547,835(引用参考))和H.Köster申请的公开号为WO94/16101，名称为通过核酸外切酶降解的定量光谱测定法进行DNA测序的国际专利申请(引用参考)，和美国专利5,605,798(引用参考)和H.Köster的申请号为PCT/US96/03651，名称为基于定量光谱测定法的DNA诊断的国际申请(引用参考)；Cohen et al.(1996)Adv Chromatogr36:127-162;and Griffin et al.(1993)Appl Biochem Biotechnol38:147-159)。在一些实施例中，本领域的普通技术人员能明白仅有一个，两个或三个核苷酸碱基需要在测序反应中确定。例如，A-轨迹等仅检测的一个核苷酸可以实现。

[0420] 名称为“使用一个混合的DNA聚合物链式探针的DNA测序方法”的美国专利5,580,732(引用参考)和名称为“检测原位DNA测序中错位的方法”的美国专利5,571,676(引用参考)。

[0421] 在一些实例中，限制酶分析物能显示出对象DNA中标记的特异等位基因。例如，一个特异性核苷酸多态可以在一个核苷酸序列中形成，其包括一个限制位点，并由于另一个核苷酸序列的突变而缺失。

[0422] 在一个更优选的实施例中，裂解剂(如一种核苷酸，羟胺或四氧化锇与哌啶)的保护可以用于检测RNA/RNA DNA/DNA或RNA/DNA杂交双链中产生的错配碱基(Myers,et al.

(1985)Science230:1242)。“错配裂解”技术通常是通过杂交一种控制核苷酸形成的杂交双链开始的,所述的控制核苷酸可以选择性地被标记,如被RNA或DNA标记,其包括一个标记突变的核苷酸序列,该序列有一个从一个组织样本中获得的样本核苷酸,如RNA或DNA。所述的双链可以用一种试剂处理,然后双链裂解成单链,如双链是控制和样本链之间的碱基对错配形成的。例如,RNA/DNA双链可以用RNase处理,S1核苷酸酶将DNA/DNA杂交物处理为分解错配区的酶。在另一些实施例中,DNA/DNA或RNA/DNA双链都可以用羟胺或四氧化锇处理,并用派啶处理来分解错配区。错配区被分解之后,产生的材料于是被涂上变性聚丙烯酰胺凝胶,确定控制和样本核苷酸是否有一个鉴定核苷酸序列或其中的核苷酸是否不同,于是被分离。例如参见Cotton et al(1988)Proc.Natl Acad Sci USA85:4397;Saleeba et al(1992)Methods Enzymol.217:286-295。在另一个实施例中,所述的控制或样本核苷酸是进行了用于检测的标记。

[0423] 在另一个实施例中,可以用高能液相色谱(DHPLC)来鉴定一种突变(Oefner and Underhill,(1995)Am.J.Human Gen.57:Suppl.A266)。DHPLC使用反相离子对色谱来检测杂交双链,这些杂交双链是在PCR片段扩增中,由带该片段的杂交单体上一个特定核苷酸位点产生(Oefner and Underhill(1995)Am.J.Human Gen.57:Suppl.A266)。PCR产物通常是使用PCR引物连接到所处理DNA两侧来产生的。DHPLC分析完成,然后分析所产生的色谱,根据特异性的色谱图鉴别出碱基对变化或删除(参见O'Donovan et al.(1998)Genomics52:44-49)。

[0424] 在其它实施例中,电泳迁移率的变化可用于鉴别标记突变的类型。例如单链构型多态性(SSCP)可以用于检测突变和各种核苷酸之间的电流迁移率差异(Orita et al.(1989)Proc Natl.Acad.Sci USA86:2766,see also Cotton(1993)Mutat Res285:125-144;and Hayashi(1992)Genet Anal Tech Appl9:73-79)。样本和控制核苷酸的单链DNA片段可以变性和复性。单链核苷酸的第二结构可以根据序列而变化,而电泳迁移率的变化甚至使一个单独碱基的检测发生变化。DNA片段可以标记或检测标记的。使用RNA(而不是DNA)可以提高评估敏感度,其中第二结构对序列中的变化更敏感。在另一个实施例中,所述的对象方法使用杂交分析并根据电泳迁移率的变化来分离双链杂交分子(Keen et al.(1991)Trends Genet7:5)。

[0425] 在另一个实施例中,分析一个核苷酸的移动可以鉴定出一个多态区的突变,包括聚丙烯酰胺凝胶中的多态区,其含有一个使用构象多态性凝胶电泳能评估的构象多态性(DGGE)(Myers et al.(1985)Nature313:495)。当使用DGGE作为分析方法时,DNA可以被修饰来保证它不会完全变性,例如利用PCR加入高溶解GC-浓缩DNA的一种约40bp GC。在一个进一步的实施例中,一个温度变化可以用于代替变性剂的变化,来鉴定修饰后的控制和样本DNA的差异(Rosenbaum and Reissner(1987)Biophys Chem265:1275)。

[0426] 检测两个核苷酸产至少一个核苷酸差异的技术实例非限制性的包括选择性寡核苷酸杂交,选择性扩增或选择性引物延伸。例如寡核苷酸探针制备时是集中地替换公知的多态性核苷酸(等位基因特异性探针),然后在仅发现一个最佳配对的杂交许可条件下与目标DNA杂交(Saiki et al.(1986)Nature324:163);Saiki et al(1989)Proc.Natl Acad.Sci USA86:6230;and Wallace et al.(1979)Nucl.Acids Res.6:3543)。这种等位基因特异性寡核苷酸杂交技术可以用于同时地检测在标记的不同多态区内发生的各种核苷

酸变化。例如如果寡核苷酸具有特异性突变的核苷酸序列，那么它附在一个杂交薄膜上，于是这个薄膜与标记的样本核苷酸杂交。然后分析杂交信号可以获得对样本核苷酸的核苷酸的检测。

[0427] 或者根据选择性PCR扩增的等位基因特异性扩增技术可以在本发明的实施例中使用。寡核苷酸可以作为引物在合适的条件下，在分子中间或者在一个引物的3'端，特异性地扩增需扩增的突变，这样能防止发生错配或减少聚合酶的延伸(因此扩增是根据不同的杂交进行的)(Gibbs et al(1989)Nucleic Acids Res.17:2437-2448)。这项技术也可以称为"PROBE"，它用于探针寡核苷酸碱基的延伸。此外它还在突变区引出了一个新的限制性位点，从而形成碱基裂解检测(Gasparini et al(1992)Mol.Cell Probes6:1)。

[0428] 在另一个实施例中，美国专利4,998,617和Landegren,U.et al.,(1988)Science241:1077-1080公开了使用一种寡核苷酸连接评估(OLA)来鉴定突变。所述的OLA方案使用两种寡核苷酸，它们可以与相邻的一个单独目标链序列杂交。一种寡核苷酸与一个单独的标记，如生物素化的标记连接，另一个被可检测地标记了。如果在一个目标分子中发现了精确互补的序列，那么这些寡核苷酸会杂交，因此它们的位置相邻，形成一个连接底物。然后连接可以使用抗生物素蛋白或其它生物素配位体来覆盖标记的寡核苷酸。Nickerson,D.A.等公开了一种核苷酸检测评估，它将PCR和OLA的特性结合起来(Nickerson,D.A.et al.,(1990)Proc.Natl.Acad.Sci.(U.S.A.)87:8923-8927)。在这个方法中，PCR用于完成目标DNA的指数扩增，而所述的目标DNA可以使用OLA来检测。

[0429] 本发明还提供了在一个标记中检测单独核苷酸多态性的方法。由于可以用同样序列区域在各种单独核苷酸多态性位点的两侧扩增，因此它们的分析不要求鉴定单独的核苷酸存在于各种位点上，不也需要为每个对象确定一个完整的基因序列。各种改良的方法可以便利于这些单独核苷酸多态性的分析。

[0430] 在一个实施例中，单独的碱基多态性可以使用一个特异性的抗核酸外切酶，例如如Mundy,C.R.所述的(美国专利4,656,127(引用参考))。根据所述的方法，一个引物与所述的定位基因序列紧接的3'互补，因此多态性位点可以与一个从特定动物或人体上获得的目标分子杂交。如果目标分子上的多态性位点含有一个核苷酸，它与所存在的特定抗核酸外切酶衍生物互补，然后衍生物可以与杂交引物的末端连接。这些连接使引物能抗外切酶，因此能实现对它的检测。虽然抗核酸外切酶衍生物的鉴定实例是公知的，但还是发现氘核这的引物可以变成抗外切酶反转的，目标分子的多态性位点上存在原核苷酸与反应中使用的核苷酸衍生物互补。这种方法的优点是它不需要确定大量外源序列的信息。

[0431] 在本发明的另一个实施例中，一种基于溶液的方法是用于鉴定一种多态性位点的核苷酸(Cohen,D.et al.French Patent2,650,840;PCT公布号:WO91/02087(引用参考))。如美国专利4,656,127(引用参考)所述的芒迪法所述，使用的引物与定位基因序列紧接的3'互补成为一个多态性位点。所述的方法能使用标记的双脱氧核苷酸衍生物来鉴定该位点的核苷酸，如果它与多态位点的核苷酸互补，那么就可以与引物的末端连接。

[0432] Goelet,P.等(PCT申请号92/15712(引用参考))公开了一种选择性方法，即基因印迹分析法或GBA。Goelet,P.等的这个方法使用标记的末端和一个引物的混合物，其中引物与序列3'的互补形成一种多态性位点。所述的标记末端被连接，因此利用所述的与之互补的核苷酸来检测，该核苷酸存在于所评估目标分子的多态位点中。与Cohen等的方法

(French Patent2,650,840;PCT公布号:WO91/02087(引用参考))相比,Goelet,P.等的方法是一种杂交相评估,其中所述的引物或目标分子是固定在一个固相上。

[0433] 用于评估DNA中多态位点的各种引导引物核苷酸连接程序已经公开(Komher,J.S.etal.,(1989)Nucl.Acids.Res.17:7779-7784;Sokolov,B.P.,(1990)Nucl.Acids Res.18:3671;Syvanen,A.-C.,et al.,(1990)Genomics8:684-692;Kuppuswamy,M.N.et al.,(1991)Proc.Natl.Acad.Sci.(U.S.A.)88:1143-1147;Prezant,T.R.et al.,(1992) Hum.Mutat.1:159-164;Ugozzoli,L.et al.,(1992)GATA9:107-112;Nyren,P.(1993)et al.,Anal.Biochem.208:171-175)。这些方法与GBA不同之处在于它们均是连接标记的脱氧核苷酸来区分一个多态性位点的多个碱基。这样虽然信号与相连的脱氧核苷酸数量成比例,但是跑相同核苷酸所出现的多态性可以产生与跑动长度成比例的信号(Syvanen,A.C.,etal.,(1993)Amer.J.Hum.Genet.52:46-59)。

[0434] 为了鉴别一个位于一个标记编码区的多态性序列的突变,还可以使用与上述方法不同的方法。例如,使用一种抗体可以鉴定一种编码一个突变标记的突变,所述的抗体能特异地识别突变的蛋白,如在免疫组织化学或免疫中。可以使用公知的方法来制备抗体到野生型标记或突变型标记。

[0435] 或者也可以测量一个标记的活性,如与一个标记配位体结合。结合评估是本领域公知的,包括如从一个对象体上获得细胞,用一个标记的配位体完成结合实验,确定与突变蛋白的结合是否不同于与野生型蛋白的结合。

[0436] VI.基于ALK抑制剂的优选筛选法

[0437] 本发明还提供了鉴定底物,抑制ALK多肽(如EML4-ALK多肽),从而抑制癌症细胞增殖,生长,突变,凋亡和/或转移的方法。这些方法包括将一个测试化合物与一个ALK多肽(如表1所列多肽)接触。在一些实施例中,所述的ALK多肽包括一个变体(如表1所列的多态),它会增加用一种或多种ALK抑制剂形成的抑制不完整或不响应的风险。一种化合物是一种肿瘤转移的抑制剂,根据一种测试化合物对ALK多肽变体的活性测量的效果就能鉴定这种化合物(例如包括配位体结合,如ATP结合和/或酪氨酸激酶的活性)。在一个特定的实施例中,一种测试化合物会抑制酪氨酸激酶的活性,将这个活性与没有测试化合物存在时的活性相比较就能鉴定出该测试化合物是否能作为一种酪氨酸激酶的抑制剂。如果一种ALK的化合物抑制剂活性改变,那么它就能进一步地评估其抑制肿瘤生长或转移的能力。

[0438] 优选的活化酪氨酸激酶突变包括本发明表1所列的新标记(如ALK突变),它能用于鉴定出用于治疗,改善或抗癌的化合物,例如掏或防止癌细胞增殖,生长,突变,凋亡和/或转移。可控制,如抑制,中和,激化或模拟分子的筛选化学库是本领域公知的。所述的化学库例如可以是肽库,拟肽类库,化学合成库,重组,如抗菌素显示库和原位底物转录库,其它非肽合成有机物库。

[0439] 筛选或形成,鉴定,选择本发明表1所列的新生物标记的合适高同源抑制剂(如ALK突变)可以用各种方法实现。概括地说它们非限制性地包括两类方案。一类是使用目标蛋白的结构知识来设计一个能精确作用的优选分子。一个实例是计算机辅助分子设计,尤其是基于图6所示的新的结构功能信息。另一类是使用组合的或其它分子库,因此一个大分子库可以被筛选用于亲和相关的目标酶,或者能够抑制目标酶的活性。在一个进一步的实施例中,可以筛选一板能抑制目标酶的抗体。

[0440] 本发明所述的一些实施例包括确定一个特定化合物能抑制本发明表1所列的新生物标记(如ALK突变)。可以比较测试化合物与公知用于治疗肿瘤性病变的化合物的活性,来评估测试化合物直接或间接治疗肿瘤性病变的能力。例如,测试化合物抑制抗本发明表1所列新生物标记的配位体结合,如ATP结合和/或酪氨酸激酶活性(如ALK突变)的能力可以与公知的ALK抑制剂,如PF-02341066和/或PDD相比。在一个实施例中,同样评估条件下,这些测试化合物对本发明表1中所列的新生物标记(如ALK突变)的抑制作用至少是公知ALK抑制剂的99.6%、99.5%、99.4%、99.3%、99.2%、99.1%、99%、98.5%、98%、97.5%、97%、96.5%、96%、95.5%、94%、93.5%、93%、92.5%、92%、91.5%、91%、90.5%、90%、89.5%、89%、88.5%、88%、87.5%、87%、86.5%、86%、85.5%、85%、84.5%、84%、83.5%、83%、82.5%、82%、81.5%、81%、80.5%、80%、79%、78%、77%、76%、75%、74%、73%、72%、71%、70%、69%、68%、67%、66%、65%、64%、63%、62%、61%、60%、59%、58%、57%、56%、55%、54%、53%、52%、51%、50%或在这些范围之内。在一些实施例中,细胞可以转染一个结构,该结构会编码本发明表1所列的新生物标记(如ALK突变),并与一个测试化合物接触,这个化合物标有一种可检测的标记,然后分析是否存在结合的测试化合物。在一些实施例中,所述的转染细胞会与测试化合物结合,将这些细胞与那些已经转染本发明表1所列的新生物标记(如ALK突变)的细胞相比,说明测试化合物已经与本发明表1所列的新生物标记(如ALK突变)结合,其中的生物标记是由这些细胞表达的。所述化合物的这种结合通常是由任意一种本领域公知的评估确定的,如ELISA,RIA和BIAcore评估。

[0441] 使用所述酶的一种重组表达突变或是从另一个物种分离得到的一个同族体或直系同源体都可以根据化合物对本发明表1中所列的新生物标记(如ALK突变)活性的抑制或其它效果来筛选这些化合物。或者这些新生物标记多肽的细胞表达可以用一种测试化合物处理,在一种特异性目标的磷酸化上的测试化合物效果可以被确定,例如使用一种本发明所述的技术。在一个实施例中,酪氨酸激酶活性被确定。确定酪氨酸激酶影响活性的方法(如抑制)是本领域普通技术人员公知的。在一些实施例中,酪氨酸激酶活性也可以通过评估一个标记的磷酸盐(如32P-标记的磷酸盐)与一个底物的结合来确定,其中所述的底物可以用本发明表1所列的新生物标记(如ALK突变)来磷酸化(如一部分或一个肽片段,尤其是那些下流信号成分)。在其它实施例中,酪氨酸激酶活性可以使用一种普通的酪氨酸激酶活性试剂盒来检测(例如普通的酪氨酸激酶检测试剂盒(Takara Bio,Inc.,Madison,Wis.);酪氨酸激酶检测试剂盒(Millipore,Billerica,Mass.))

[0442] 在另一个实施例中,所述的筛选方法还包括确定所述的化合物是否减缓了肿瘤细胞的生长,例如公知肿瘤细胞会表达一种活化的酪氨酸激酶突变,如本发明表1所列的一种新生物标记(如ALK突变)。可以使用各种细胞系,选择的依据其组织是本领域的普通技术人员公知能检测的(如BA/F3细胞)。例如各种细胞系被很好地表征,并且被例如美国国家癌症研究所(NCI)用在其研发的新型抗癌药物的筛选程序中。

[0443] 高效和显著高效的肿瘤细胞生长抑制剂,如在100μM,90μM,80μM,70μM,60μM,50μM,40μM,30μM,20μM,10μM,9μM,8μM,7μM,6μM,5μM,4.5μM,4μM,3.5μM,3μM,2.5μM,2μM,1.5μM,1μM,900nM,850nM,800nM,750nM,700nM,650nM,600nM,550nM,500nM,450nM,400nM,350nM,300nM,250nM,200nM,150nM,100nM,95nM,90nM,85nM,80nM,75nM,70nM,65nM,60nM,55nM,50nM,45nM,40nM,35nM,30nM,25nM,20nM,15nM,10nM,5nM,4nM,3nM,2nM,1nM或更低剂量时有效率大于50%,它还说明所述的化合物能有效地用于治疗肿瘤性病变。一种IC_{50}值可

以确定并用于对比。这个值是通过50%相对于控制所需要抑制肿瘤细胞生长的药物浓度。

[0444] 这些值还可以使用其它标准。例如在另一些实施例中，本发明所述的筛选法还包括确定测试化合物是否会引起培养中的肿瘤细胞发生凋亡。细胞死亡的两种不同形式可以用形态学和生物指示来描述：坏死和凋亡。增加等离子体薄膜的渗透率可以随之产生坏死，因此细胞膨胀，而等离子体薄膜在几秒钟内破裂。薄膜起泡，细胞质冷凝和内源性核酸内切酶的活性是凋亡的特征。

[0445] 正常组织更新和器官，四肢的胚芽发育时会自然地产生凋亡。各种刺激因素通过电离辐射和一些化疗药物也会诱导凋亡产生，包括细胞毒性T淋巴细胞和自然杀伤细胞。凋亡的异常控制是在包括癌症、AIDS或阿尔茨海默氏病等各种病理学情况下充当了一个重要的作用。

[0446] 在上述条件下培养肿瘤细胞，然后筛选测试化合物对细胞凋亡的诱导。这些筛选方法的一些实施例中，用测试化合物处理细胞包括在各种浓度的测试化合物中前或后融合培养基并处理一至七天。在培养基的附着或“漂浮”部分都可以测量凋亡的细胞。这两者都可以这样收集：去除表面漂浮物，胰酶消化附着细胞，在一个离心冲洗步骤(10分钟，2000rpm)之后组合两种制备。用一种测试化合物处理之后，培养基可以评估用于凋亡和坏死，例如在标记了吖啶橙和溴乙锭之后使用荧光显微镜。各种测量凋亡细胞的方法是本领域普通技术人员公知的，例如Duke&Cohen公开了一种测量凋亡细胞数量的方法(Curr.Prot.Immuno.,Coligan et al.,eds.,3.17.1-3.17.1,1992)。例如使用胰蛋白酶并用PBS冲洗三次，收集漂浮和附着的细胞。于是细胞的等分部分被离心。所述的颗粒再次悬浮在媒介中，一种含有吖啶橙和溴乙锭的染料混合物在PBS中制备而成，并被平缓地混合。然后所述的混合物可以固定在一个显微镜滑槽中，检查其凋亡的形态特征。测量细胞中DNA片段的增加量也可以定量地检测凋亡，其中的细胞已经用测试化合物处理了。可以使用商业光度酶免疫测定(EIA)来定量地检测胞质蛋白-相关的-DNA-片段(单体和寡核苷酸)(如Cell Death Detection ELISA,Boehringer Mannheim)。

[0447] 在另外的实施例中，本发明的筛选法还包括确定测试化合物是否降低了肿瘤的移植，如在一个移植的动物模型。评估肿瘤移植的方法是本领域普通技术人员公知的(例如参见Khanna and Hunter,Carcinogenesis26:513-523,2005)。一种移植模型包括人-鼠异种移植，其中人体癌细胞系或组织可以移植到免疫系统受损伤的小鼠中(如SCID小鼠或裸小鼠)。在相似的方法中，一个细胞系可以设计为表达本发明表1所列的一种新型生物标记(如ALK突变)，它可以移植到一种免疫系统受损伤的小鼠中。在一个实施例中，肿瘤细胞或细胞系可以直接注入到系统循环中。注入的位点大半地确定了移植生长在这些实验系统中的位点。实验转移模型中最常用的肿瘤细胞注入位点是在小鼠尾静脉两侧，它主要是在肺转移中产生。肿瘤细胞中的脾内或肝门静脉注射是在活体中培育移植的最常用位点，而细胞的心脏内注射会在转移中产生多个位点，包括骨。肿瘤细胞或其它细胞系注入到循环之后，在目标位点(如肺)上培育的转移可以在若干天或周之后监测到。

[0448] 评估肿瘤转移的另一种模型是使用常位移植，其中癌症细胞可以移植给常位或组织，于是一个肿瘤从这里衍生(例如直接注入或外科培植肿瘤片段)。常位肿瘤中产生的自然转移可以在几天或几周内评估。一种测试化合物降低或防止肿瘤转移的能力可以通过将一种测试化合物皮下注射，肌肉注射，血液循环或原位移植地注入到肿瘤细胞中。转移培育

的数量，大小或时间是可以评估的。一种抑制肿瘤转移的化合物会减少转移的数量，例如比控制样本至少减少10%、20%、30%、40%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%，oreven100%。一种能抑制肿瘤转移的化合物也能使转移的大小比一个控制样本小。相似地，一种抑制肿瘤转移的化合物也可以减缓转移培育的开始，例如至少一周，两周，一个月，6个月，一年或者甚至是无限期的。

[0449] VII.优选的ALK抑制剂

[0450] 本发明所述的方法包括鉴定一种对象是否能用本发明表1所列的新型生物标记(如ALK突变)来治疗，诱导肿瘤细胞死亡，阻碍肿瘤生长或降低肿瘤转移的风险。ALK多肽抑制剂是本领域公知的。例如PF-02341066，PDD，2-甲基-11-(2-甲基丙基)-4-氧-4,5,6,11,12,13-六氢-2H-吲唑[5,4-α]吡咯[3,4-c]咔唑-8-基[4-(二甲氨基)苄基]氨基甲酸甲酯，(1S,2S,3R,4R)-3-({5-氯-2-[(1-乙基-2,3,4,5-四氢-6-甲氧基-2-氧-1H-1-苯并氮杂卓-7-基)氨基]-4-嘧啶}氨基)双环[2.2.1]庚-5-烯-2-甲酰胺，和NVP-TAE684，例如参见PNAS104:270-275,2007;Choi,Y.L.et al.(2008)Cancer Res.68:4971-2976;以及Biochemistry48:3600-3609,2009，它们在此是引用参考的。

[0451] 引用参考

[0452] 每个单独出版物，专利或专利申请特别而单独地表示将纳入参考时，本发明所述的所有出版物，专利和专利申请全部被纳入作为参考。当它们相互矛盾时，本申请，包括本发明的所有定义将起决定性作用。

[0453] 另外在其全部纳入参考公共数据库中与其全文相关的评估数有关的任何多核苷酸和多肽序列，例如由基因组研究所(TIGR)在tigr.org处的万维网上保持的那些和/或国家生物技术信息中心(NCBI)在ncbi.nlm.nih.gov处的万维网上保持的那些。

[0454] 例子

[0455] 本发明通过以下例子作进一步地说明，不应理解为限制。所有涉及到的内容，数字，序列表，专利和引用的公开专利文献均属于本发明申请范围。

[0456] 例1-用于例2-4的原料与方法

[0457] a.DNA测序

[0458] 寡聚(dT)-引物cDNAs来自样品RNAs使用EZ1系统(Qiagen,Valencia,CA)萃取并用PrimeSTAR HS DNA聚合酶(Takara Bio Inc.,Shiga,Japan)和引物ALK-TK-F

[0459] (5'-TACAACCCCAACTACTGCTTTGCT-3')和ALK-TK-R1(5'-AGGCACTTTCTCTTCCTCTTCCAC-3')通过聚合酶链反应(PCR)30个循环(包括98℃时10秒钟和68℃时2分钟)得到。PCR产物与ALK激酶结构域相同，然后分散并用Illumina基因组分析系统II(GAII)进行测序，通过配对端测序系统(Illumina,San Diego,CA)得到两端为76基质。读取原始数据是以存在的PCR原始顺序和质量≥20为基础的质量-渗透的所有基质。渗透通过所读取的然后用鲍蒂算法(提供在万维网bowtie-bio.sourceforge.net/index.shtml上)来校准ALK cDNA顺序。

[0460] 用3130xl基因分析仪(应用生物系统公司，福斯特市，加利福尼亚州)来毛细管测序，PCR产物与所希望的cDNAs具有相同的引物结构或具有EA-F-g-S

[0461] (5'-CCACACCTGGGAAAGGACCTAAAG-3')和ALK-TK-R2(5'-CCTCCAAATACTGACAGCCACAGG-3')引物的序列。

[0462] b.突变EML4-ALK

[0463] 把cDNA编码FLAG抗原表位标记EML4-ALK变体1(Soda,M.et al.(2007)Nature448:561-566)插入到pMX-iresCD8逆转录病毒载体(Yamashita Y.et al.(2001)J.Biol.Chem.276:39012-39020)中得到同时表达的FLAG-标记EML4-ALK和鼠CD8。核苷酸变化与C1156Y和L1196M一致,ALK的突变引进独个的质粒或用于EML4-ALK(C1156Y),

[0464] EML4-ALK(L1196M),或EML4-ALK(C1156Y/L1196M)表达的组合。基于这些质粒的重组逆转录病毒采用包装细胞系,BOSC23(Pear,W.S.et al.(1993)Proc.Natl.Acad.Sci.USA90:8392-8396),和采用感染小鼠白细胞介素-3-依赖细胞系BA/F3(Palacious,R.et al.(1985)Cell41:727-734)来产生。其结果CD8-阳性细胞采用miniMACS细胞分离柱和磁性玻璃粉来纯化并用共轭到对CD8的抗体上(两者均来自Miltenyi Biotec,Gladbach,Germany)。PF-02341066购自塞莱克。

[0465] 为了检验EML4-ALK的酪氨酸磷酸化,BA/F3细胞表达融合蛋白暴露在ALK抑制剂下15小时,之后EML4-ALK从细胞裂解物中免疫沉淀出具有对FLAG(Sigma-Aldrich,St.Louis,MO)的抗体和用对Tyr1604-磷酸化ALK(Cell Signaling Technology,Danvers,MA)抗体的受免疫分析。体外激酶测定是在室温下用合成YFF肽(Operon Biotechnologies,Huntsville,AL)操作30分钟,如早先(Donella-Deana,A.et al.(2005)Biochemistry44:8533-8542)所述的。

[0466] 例2-新型突变ALK与抗ALK酪氨酸激酶抑制剂的联系

[0467] 患者是一名28岁的男人,没有吸烟史,在2008年4月临床阶段被诊断为肺腺癌T4N3M1。鉴于肿瘤没有任何突变EGFR生存,此患者接受常规化疗,其结果是疾病出现连续多次在大脑和骨骼中转移。在2008年11月,通过逆转录-PCR痰液分析以及通过组织样本的荧光原位杂交分析,发现mRNA的EML4-ALK变体1在肿瘤中存在。因此,该患者登记到PF-02341066试验中,并在他的体力状况明显改善(从4级减少为2)开始。虽然他在治疗过程中表现出“局部反应”,他的胸腔积液并非完全根除。然而,经过5个月的治疗,肿瘤突然开始再次增长,导致胸腔积液增大并在双肺上形成多个癌小瘤。患者在2009年5月放弃试验,并对胸腔积液进行分子学分析。

[0468] 尽管持续的给予ALK抑制剂但肿瘤还是恢复增长,这表明肿瘤产生了次生遗传变化具有对药物的抗药性。此外,由于抗TKIs往往是在激酶靶上产生突变导致的,并对EML4-ALK它自身经过了氨基酸的变化的可能进行了验证。

[0469] 分别对患者肿瘤治疗前和后的痰(IDJ-#1)及胸腔积液(IDJ-#113)标本进行了分子学分析可知,鉴于这两个标本肿瘤细胞的比例可能会有所不同,新一代测序仪用来对采自这样标本的EML4-ALK的cDNAs进行了深度的序列分析。两个标本的详情结果是cDNAs与ALK酪氨酸激酶结构域相同(图1A),分散并用GAII系统进行核苷酸顺序分析。作为比较,EML4-ALK的阳性非小细胞肺癌细胞系H2228,和其他三个也是融合蛋白质阳性的临床标本进行相同的分析。一个已知的单核苷酸多态性rs3795850,在四个标本中均发现有cDNAs(图1B)。此外,在J-#1cDNAs中低概率(8.9%)发现在人类野生型ALK cDNA(基因库登记号,NM_004304)的相当于核苷酸4230位置有一个T→C的变化。而且,两个新的变化,在J-#113cDNAs中发现在野生型ALK cDNA的相当于核苷酸4374和4493位置有G→A和C→A的变化,其概率分别是41.8%和14.0%。在采自其他标本的激酶结构域cDNAs中没有出现其他经常性的改变

(出现读取率在≥5%的)。

[0470] 这些核苷酸的变化随后使用桑格测序证实。要排除这种突变是发生在内源性野生型ALK内而不是在EML4-ALK内的可能性,用PCR来检测前端引物目标至EML4cDNA,以致于只将融合型cDNA扩增(图1A)。采自J-#1的几百个融合型cDNAs中T4230C没有发现变化,这表明在原始的PCR或GAII测序步骤中它是人工产物。

[0471] 然而,G4374A和C4493A的变化都可以通过桑格测序很容易的得到证实。为J-#113克隆测序73融合cDNA中,G4374A的34克隆(46.6%)为阳性,C4493A的11克隆(15.1%)为阳性,其余(38.4%)为野生型(图1C)。由于PCR分析在相同的产品中涵盖了两个核苷酸位点,没有产品包含两个突变,这表明每一个突变都是独立的。围绕G4374或C4493位置的基因组片段也通过PCR扩增并进行核苷酸测序,其结果是在肿瘤中的每个基因组各自变化得到证实(图2)。

[0472] G4374A和C4493A的替换结果是在相当于野生型人类ALK的氨基酸1156和1196位置上有半胱氨酸→酪氨酸和亮氨酸→蛋氨酸的变化。

[0473] 例3-新型突变ALK具有对ALK酪氨酸激酶抑制剂的抗药性

[0474] 下一步研究的是ALK抑制剂对EML4-ALK的氨基酸变化影响的敏感性。野生型EML4-ALK,单突变EML4-ALK(C1156Y)和EML4-ALK(L1196M),和双突变EML4-ALK(C1156Y/L1196M)分别在BA/F3细胞和与ALK抑制剂接触后的细胞上表达。PF-02341066在浓度依赖型生长的BA/F3细胞上表达野生型EML4-ALK的抑制作用(图4A)。与此相反,细胞突变表达要么C1156Y要么L1196M突变证实对药物敏感性显著减低,重复试验表明BA/F3细胞表达的EML4-ALK(L1196M)比表达的EML4-ALK(C1156Y)对PF-02341066的抗药性更大(图3)。存在的这两个突变并没有导致细胞对PF-02341066抗药性的叠加作用。因此,这些数据表明C1156Y和L1196M突变分别具有对药物的抗药性。

[0475] 通过使用在Tyr1604上ALK磷酸化特异性抗体的免疫分析来验证EML4-ALK酪氨酸磷酸化。虽然BA/F3细胞暴露于PF-02341066下可明显抑制野生型EML4-ALK酪氨酸磷酸化,它没有对EML4-ALK(C1156Y)或EML4-ALK(L1196M)的重大影响(图4B)。根据这些发现,在体外激酶试验中发现EML4-ALK的C1156Y和L1196M突变比野生型蛋白用PF-02341066对酶活性的抑制敏感性更小(图4C)。作为抑制细胞生长的案例(图4A),L1196M突变比C1156Y突变用PF-02341066对激酶活性的抑制更加困难(图4C)。

[0476] 例4-新型突变ALK与抗ALK酪氨酸激酶抑制剂之间的结构-功能关系

[0477] 图5显示Cys1156和Leu1196在基于激酶,胰岛素受体的晶体结构上ALK激酶结构域的三维立体结构模型中的位置,前者残体位于与预测螺旋αC相连的氨基端也就是接近ATP-结构组织帽子上端的位置。在其他酪氨酸激酶中还没有关于在这个位置上有活性突变的报道。ALK的Leu1196相当于ABL1的Thr315和EGFR的Thr790,在这些激酶中的这些位置每一个都非常容易突变成具有对TKIs的抗药性(Deininger,M.et al.(2005)Blood105:2640-2653;Linardou,H.et al.(2009)Nat.Rev.Clin.Oncol.6:352-366)。这些"看门人"位置位于ATP-结构组织的底面上(图5),并在这些位置具有大侧链的氨基酸存在是来阻止大量TKIs的粘合(Shah,N.P.et al.(2002)Cancer Cell2:117-125;Tsao,M.S.et al.(2005)N.Engl.J.Med.353:133-144)。

[0478] 因此,对在EML4-ALK的激酶结构域的两头具有对多重ALK抑制剂的抵抗性进行了

识别，鉴于没有观察到EML4-ALK cDNAs的两种突变，可认为每一个突变在肿瘤的不同亚克隆内独立发生。

[0479] 如果没有上述所述的反弹，从患者治疗前的痰液预先得到的cDNAs不含有与C1156Y或L1196M突变一样的变化核苷酸，很可能是在用PF-02341066治疗的肿瘤亚克隆过程中获得重新突变。然而，在治疗前无法对胸腔积液进行检查，患者入院初期没有完全排除的早已存在胸腔积液中可能生存有肿瘤细胞的C1156Y或L1196M突变。如果是这样的情况下，在5个月用PF-02341066治疗期内肿瘤可能已经获得其他目前未知的突变，并允许其随后的快速增长。然而，具有C1156Y或L1196M突变的肿瘤细胞亚克隆在治疗初期很难治疗，并在治疗过程中有扩增的可能。相反，在至少5个月内患者没有肿瘤扩增的信号，这表明C1156Y和L1196M突变发生在用PF-02341066治疗期间。这种观点通过EGFR的T790M突变实事进一步得到支持，在患者事先用TKIs治疗中经常发现对吉非替尼或埃罗替尼具有抵抗性，而很难找到没有进行治疗的个案(Pao,W.et al.(2005)PLoS Med.2:e73)。

[0480] 在用TKIs治疗肿瘤中发现一些酪氨酸激酶的看门人位置发生氨基酸替换(Kobayashi,S.et al.(2005)N.Engl.J.Med.352:786-792;Pao,W.et al.(2005)PLoS Med.2:e73;Shah,N.P.et al.(2002)Cancer Cell2:117-125;Cools,J.et al.(2003)N.Engl.J.Med.348:1201-1214;Tamborini,E.et al.(2004)Gastroenterology127:294-299)。但事先还没有用EML4-ALK或ALK在这个位置突变的报道，各种在NPM-ALK，其他融合型ALK的看门人位置的人为氨基酸替换的影响最近得到验证((Lu,L.et al.(2009)Biochemistry48:3600-3609)。与体内肿瘤细胞分析一致，发现在这个位置引进蛋氨酸使NPM-ALK对多重ALK抑制剂具有最大抵抗性。

[0481] 相比看门人替换，在其他酪氨酸激酶中还没有在αC螺旋接近氨基末端的(ALK中的Cys1156)这个位置有活性突变的报道。在NSCLC病例中有在EGFR的相应位置存在Thr→Ile变化的阐述，但没有对药物敏感性进行验证(Tsao,M.S.et al.(2005)N.Engl.J.Med.353:133-144)。用于丝氨酸-苏氨酸激酶的重要αC螺旋的酶活性变构规则已得到验证

[0482] (Hindie,V.et al.(2009)Nat.Chem.Biol.5:758-764)。ALK的Cys1156变化直接参与到了激酶结构域和TKIs之间的物理作用中。

[0483] 等同

[0484] 那些在本领域技术人员将认识到或能够确定使用但不超过常规的试验，与本发明在这里所述的具体实施例有很多等同。这些等同包含在本发明的权利要求中。

<110> 自治医科大学

<120> 对 ALK 抑制剂具有先抗性或后抗性癌症的识别，判断和治疗

<140> 201410061691.6

<160> 40

<210> 1
<211> 6222
<212> DNA
<213> 人类

<400> 1
[0001]
ggggggcggca gcggtggtag cagctggtac ctcccgccgc ctctgttcgg agggtcgcgg 60

ggcaccgagg tgctttccgg ccgccctctg gtcggccacc caaagccgcg ggcgctgatg 120

atgggtgagg agggggcggc aagatttcgg gcgcccctgc cctgaacgcc ctcagctgct 180

gccgccgggg ccgctccagt gcctgcgaac tctgaggagc cgaggcgccg gtgagagcaa 240

ggacgctgca aacttgcgca gcgcgggggc tgggattcac gcccagaagt tcagcaggca 300

gacagtccga agccttcccg cagcggagag atagcttgag ggtgcgcaag acggcagcct 360

ccgccctcgg ttcccgccca gaccgggcag aagagcttgg aggagccaaa aggaacgcaa 420

aaggcggcca ggacagcgtg cagcagctgg gagccgccgt tctcagcctt aaaagttgca 480

gagattggag gctgcccccga gaggggacag accccagctc cgactgcggg gggcaggaga 540

ggacggtacc caactgccac ctcccttcaa ccatagtagt tcctctgtac cgagcgcagc 600

gagctacaga cgggggcgcg gcactcggcg cggagagcgg gaggctcaag gtcccagcca 660

gtgagcccag tgtgcttgag tgtctctgga ctcgcccctg agcttccagg tctgtttcat 720

ttagactcct gctcgcctcc gtgcagttgg gggaaagcaa gagacttgcg cgcacgcaca 780

gtcctctgga gatcaggtgg aaggagccgc tgggtaccaa ggactgttca gagcctcttc 840

ccatctcggg gagagcgaag ggtgaggctg ggcccggaga gcagtgtaaa cggcctcctc 900

cggcgggatg ggagccatcg ggctcctgtg gctcctgccg ctgctgcttt ccacggcagc 960

tgtgggctcc gggatgggga ccggccagcg cgcgggctcc ccagctgcgg ggccgccgct 1020

gcagccccgg gagccactca gctactcgcg cctgcagagg aagagtctgg cagttgactt 1080

cgtggtgccc tcgctcttcc gtgtctacgc ccgggaccta ctgctgccac catcctcctc 1140

ggagctgaag gctggcaggc ccgaggcccg cggctcgcta gctctggact gcgccccgct 1200

gctcaggttg ctggggccgg cgccgggggt ctcctggacc gccggttcac cagccccggc 1260

agaggcccgg acgctgtcca gggtgctgaa gggcggctcc gtgcgcaagc tccggcgtgc 1320

caagcagttg gtgctggagc tgggcgagga ggcgatcttg gagggttgcg tcgggccccc 1380

cggggaggcg gctgtggggc tgctccagtt caatctcagc gagctgttca gttggtggat 1440

tcgccaaggc gaagggcgac tgaggatccg cctgatgccc gagaagaagg cgtcggaagt 1500

gggcagagag ggaaggctgt ccgcggcaat tcgcgcctcc cagccccgcc ttctcttcca 1560

[0002]

gatcttcggg actggtcata gctccttgga atcaccaaca aacatgcctt ctccttctcc 1620

tgattatttt acatggaatc tcacctggat aatgaaagac tccttccctt tcctgtctca 1680

tcgcagccga tatggtctgg agtgcagctt tgacttcccc tgtgagctgg agtattcccc 1740

tccactgcat gacctcagga accagagctg gtcctggcgc cgcatcccct ccgaggaggc 1800

ctcccagatg gacttgctgg atgggcctgg ggcagagcgt tctaaggaga tgcccagagg 1860

ctcctttctc cttctcaaca cctcagctga ctccaagcac accatcctga gtccgtggat 1920

gaggagcagc agtgagcact gcacactggc cgtctcggtg cacaggcacc tgcagcctc 1980

tggaaggtac attgcccagc tgctgcccca caacgaggct gcaagagaga tcctcctgat 2040

gcccactcca gggaagcatg gttggacagt gctccaggga agaatcgggc gtccagacaa 2100

cccatttcga gtggccctgg aatacatctc cagtggaaac cgcagcttgt ctgcagtgga 2160

cttctttgcc ctgaagaact gcagtgaagg aacatcccca ggctccaaga tggccctgca 2220

gagctccttc acttgttgga atgggacagt cctccagctt gggcaggcct gtgacttcca 2280

ccaggactgt gcccagggag aagatgagag ccagatgtgc cggaaactgc ctgtgggttt 2340

ttactgcaac tttgaagatg gcttctgtgg ctggacccaa ggcacactgt caccccacac 2400

tcctcaatgg caggtcagga ccctaaagga tgcccggttc caggaccacc aagaccatgc 2460

tctattgctc agtaccactg atgtccccgc ttctgaaagt gctacagtga ccagtgctac 2520

gtttcctgca ccgatcaaga gctctccatg tgagctccga atgtcctggc tcattcgtgg 2580

agtcttgagg ggaaacgtgt ccttggtgct agtggagaac aaaaccggga aggagcaagg 2640

caggatggtc tggcatgtcg ccgcctatga aggcttgagc ctgtggcagt ggatggtgtt 2700

gcctctcctc gatgtgtctg acaggttctg gctgcagatg gtcgcatggt ggggacaagg 2760

atccagagcc atcgtggctt ttgacaatat ctccatcagc ctggactgct acctcaccat 2820

tagcggagag gacaagatcc tgcagaatac agcacccaaa tcaagaaacc tgtttgagag 2880

[0003]

aaacccaaac aaggagctga aacccgggga aaattcacca agacagaccc ccatctttga 2940

ccctacagtt cattggctgt tcaccacatg tggggccagc gggccccatg gccccaccca 3000

ggcacagtgc aacaacgcct accagaactc caacctgagc gtggaggtgg ggagcgaggg 3060

ccccctgaaa ggcatccaga tctggaaggt gccagccacc gacacctaca gcatctcggg 3120

ctacggagct gctggcggga aaggcgggaa gaacaccatg atgcggtccc acggcgtgtc 3180

tgtgctgggc atcttcaacc tggagaagga tgacatgctg tacatcctgg ttgggcagca 3240

gggagaggac gcctgcccca gtacaaacca gttaatccag aaagtctgca ttggagagaa 3300

caatgtgata gaagaagaaa tccgtgtgaa cagaagcgtg catgagtggg caggaggcgg 3360

aggaggaggg ggtggagcca cctacgtatt taagatgaag gatggagtgc cggtgccccc 3420

gatcattgca gccggaggtg gtggcagggc ctacggggcc aagacagaca cgttccaccc 3480

agagagactg gagaataact cctcggttct agggctaaac ggcaattccg gagccgcagg 3540

tggtggaggt ggctggaatg ataacacttc cttgctctgg gccggaaaat ctttgcagga 3600

gggtgccacc ggaggacatt cctgccccca ggccatgaag aagtgggggt gggagacaag 3660

agggggtttc ggaggggggtg gagggggggtg ctcctcaggt ggaggaggcg gaggatatat 3720

aggcggcaat gcagcctcaa acaatgaccc cgaaatggat ggggaagatg gggtttcctt 3780

catcagtcca ctgggcatcc tgtacacccc agctttaaaa gtgatggaag gccacgggga 3840

agtgaatatt aagcattatc taaactgcag tcactgtgag gtagacgaat gtcacatgga 3900

ccctgaaagc cacaaggtca tctgcttctg tgaccacggg acggtgctgg ctgaggatgg 3960

cgtctcctgc attgtgtcac ccaccccgga gccacacctg ccactctcgc tgatcctctc 4020

tgtggtgacc tctgccctcg tggccgccct ggtcctggct ttctccggca tcatgattgt 4080

gtaccgccgg aagcaccagg agctgcaagc catgcagatg gagctgcaga gccctgagta 4140

caagctgagc aagctccgca cctcgaccat catgaccgac tacaacccca actactgctt 4200

[0004]

tgctggcaag acctcctcca tcagtgacct gaaggaggtg ccgcggaaaa acatcaccct 4260

cattcggggt ctgggccatg gcgcctttgg ggaggtgtat gaaggccagg tgtccggaat 4320

gcccaacgac ccaagccccc tgcaagtggc tgtgaagacg ctgcctgaag tgtgctctga 4380

acaggacgaa ctggatttcc tcatggaagc cctgatcatc agcaaattca accaccagaa 4440

cattgttcgc tgcattgggg tgagcctgca atccctgccc cggttcatcc tgctggagct 4500

catggcgggg ggagacctca agtccttcct ccgagagacc cgccctcgcc cgagccagcc 4560

ctcctccctg gccatgctgg accttctgca cgtggctcgg gacattgcct gtggctgtca 4620

gtatttggag gaaaaccact tcatccaccg agacattgct gccagaaact gcctcttgac 4680

ctgtccaggc cctggaagag tggccaagat tggagacttc gggatggccc gagacatcta 4740

cagggcgagc tactatagaa agggaggctg tgccatgctg ccagttaagt ggatgccccc 4800

agaggccttc atggaaggaa tattcacttc taaaacagac acatggtcct ttggagtgct 4860

gctatgggaa atcttttctc ttggatatat gccataccc agcaaaagca accaggaagt 4920

tctggagttt gtcaccagtg gaggccggat ggacccaccc aagaactgcc ctgggcctgt 4980

ataccggata atgactcagt gctggcaaca tcagcctgaa gacaggccca actttgccat 5040

cattttggag aggattgaat actgcaccca ggacccggat gtaatcaaca ccgctttgcc 5100

gatagaatat ggtccacttg tggaagagga agagaaagtg cctgtgaggc ccaaggaccc 5160

tgagggggtt cctcctctcc tggtctctca acaggcaaaa cgggaggagg agcgcagccc 5220

agctgcccca ccacctctgc ctaccacctc ctctggcaag gctgcaaaga aacccacagc 5280

tgcagagatc tctgttcgag tccctagagg gccggccgtg gaaggggggac acgtgaatat 5340

ggcattctct cagtccaacc ctccttcgga gttgcacaag gtccacggat ccagaaacaa 5400

gcccaccagc ttgtggaacc caacgtacgg ctcctggttt acagagaaac ccaccaaaaa 5460

gaataatcct atagcaaaga aggagccaca cgacaggggt aacctggggc tggagggaag 5520

[0005]

ctgtactgtc ccacctaacg ttgcaactgg gagacttccg ggggcctcac tgctcctaga 5580

gccctcttcg ctgactgcca atatgaagga ggtacctctg ttcaggctac gtcacttccc 5640

ttgtgggaat gtcaattacg gctaccagca acagggcttg cccttagaag ccgctactgc 5700

ccctggagct ggtcattacg aggataccat tctgaaaagc aagaatagca tgaaccagcc 5760

tgggccctga gctcggtcgc acactcactt ctcttccttg ggatccctaa gaccgtggag 5820

gagagagagg caatggctcc ttcacaaacc agagaccaaa tgtcacgttt tgttttgtgc 5880

caacctattt tgaagtacca ccaaaaaagc tgtattttga aaatgcttta gaaaggtttt 5940

gagcatgggt tcatcctatt ctttcgaaag aagaaaatat cataaaaatg agtgataaat 6000

acaaggccca gatgtggttg cataaggttt ttatgcatgt ttgttgtata cttccttatg 6060

cttctttcaa attgtgtgtg ctctgcttca atgtagtcag aattagctgc ttctatgttt 6120

catagttggg gtcatagatg tttccttgcc ttgttgatgt ggacatgagc catttgaggg 6180

gagagggaac ggaaataaag gagttatttg taatgactaa aa 6222

<210> 2
<211> 24
<212> DNA
<213> 人工序列

<220>
<223>人工序列的描述：合成引物

<400> 2
ggtaagaagt ggctcactct tgag 24

<210> 3
<211> 24
<212> DNA
<213> 人工序列

<220>
[0006] <223> 人工序列描述：合成引物

<400> 3
cacaacaact gcagcaaaga ctgg 24

<210> 4
<211> 1620
<212> PRT
<213> 人类

<400> 4
Met Gly Ala Ile Gly Leu Leu Trp Leu Leu Pro Leu Leu Leu Ser Thr
1 5 10 15

Ala Ala Val Gly Ser Gly Met Gly Thr Gly Gln Arg Ala Gly Ser Pro
20 25 30

Ala Ala Gly Pro Pro Leu Gln Pro Arg Glu Pro Leu Ser Tyr Ser Arg
35 40 45

Leu Gln Arg Lys Ser Leu Ala Val Asp Phe Val Val Pro Ser Leu Phe
50 55 60

Arg Val Tyr Ala Arg Asp Leu Leu Leu Pro Pro Ser Ser Ser Glu Leu
65 70 75 80

Lys Ala Gly Arg Pro Glu Ala Arg Gly Ser Leu Ala Leu Asp Cys Ala
85 90 95

Pro Leu Leu Arg Leu Leu Gly Pro Ala Pro Gly Val Ser Trp Thr Ala
100 105 110

Gly Ser Pro Ala Pro Ala Glu Ala Arg Thr Leu Ser Arg Val Leu Lys
115 120 125

[0007] Gly Gly Ser Val Arg Lys Leu Arg Arg Ala Lys Gln Leu Val Leu Glu
130 135 140

Leu Gly Glu Glu Ala Ile Leu Glu Gly Cys Val Gly Pro Pro Gly Glu
145 150 155 160

Ala Ala Val Gly Leu Leu Gln Phe Asn Leu Ser Glu Leu Phe Ser Trp
165 170 175

Trp Ile Arg Gln Gly Glu Gly Arg Leu Arg Ile Arg Leu Met Pro Glu
180 185 190

Lys Lys Ala Ser Glu Val Gly Arg Glu Gly Arg Leu Ser Ala Ala Ile
195 200 205

Arg Ala Ser Gln Pro Arg Leu Leu Phe Gln Ile Phe Gly Thr Gly His
210 215 220

Ser Ser Leu Glu Ser Pro Thr Asn Met Pro Ser Pro Ser Pro Asp Tyr
225 230 235 240

Phe Thr Trp Asn Leu Thr Trp Ile Met Lys Asp Ser Phe Pro Phe Leu
245 250 255

Ser His Arg Ser Arg Tyr Gly Leu Glu Cys Ser Phe Asp Phe Pro Cys
260 265 270

Glu Leu Glu Tyr Ser Pro Pro Leu His Asp Leu Arg Asn Gln Ser Trp
275 280 285

Ser Trp Arg Arg Ile Pro Ser Glu Glu Ala Ser Gln Met Asp Leu Leu
290 295 300

[0008] Asp Gly Pro Gly Ala Glu Arg Ser Lys Glu Met Pro Arg Gly Ser Phe
305 310 315 320

Leu Leu Leu Asn Thr Ser Ala Asp Ser Lys His Thr Ile Leu Ser Pro
325 330 335

Trp Met Arg Ser Ser Ser Glu His Cys Thr Leu Ala Val Ser Val His
340 345 350

Arg His Leu Gln Pro Ser Gly Arg Tyr Ile Ala Gln Leu Leu Pro His
355 360 365

Asn Glu Ala Ala Arg Glu Ile Leu Leu Met Pro Thr Pro Gly Lys His
370 375 380

Gly Trp Thr Val Leu Gln Gly Arg Ile Gly Arg Pro Asp Asn Pro Phe
385 390 395 400

Arg Val Ala Leu Glu Tyr Ile Ser Ser Gly Asn Arg Ser Leu Ser Ala
405 410 415

Val Asp Phe Phe Ala Leu Lys Asn Cys Ser Glu Gly Thr Ser Pro Gly
420 425 430

Ser Lys Met Ala Leu Gln Ser Ser Phe Thr Cys Trp Asn Gly Thr Val
435 440 445

Leu Gln Leu Gly Gln Ala Cys Asp Phe His Gln Asp Cys Ala Gln Gly
450 455 460

Glu Asp Glu Ser Gln Met Cys Arg Lys Leu Pro Val Gly Phe Tyr Cys
465 470 475 480

[0009] Asn Phe Glu Asp Gly Phe Cys Gly Trp Thr Gln Gly Thr Leu Ser Pro
485 490 495

His Thr Pro Gln Trp Gln Val Arg Thr Leu Lys Asp Ala Arg Phe Gln
500 505 510

Asp His Gln Asp His Ala Leu Leu Leu Ser Thr Thr Asp Val Pro Ala
515 520 525

Ser Glu Ser Ala Thr Val Thr Ser Ala Thr Phe Pro Ala Pro Ile Lys
530 535 540

Ser Ser Pro Cys Glu Leu Arg Met Ser Trp Leu Ile Arg Gly Val Leu
545 550 555 560

Arg Gly Asn Val Ser Leu Val Leu Val Glu Asn Lys Thr Gly Lys Glu
565 570 575

Gln Gly Arg Met Val Trp His Val Ala Ala Tyr Glu Gly Leu Ser Leu
580 585 590

Trp Gln Trp Met Val Leu Pro Leu Leu Asp Val Ser Asp Arg Phe Trp
595 600 605

Leu Gln Met Val Ala Trp Trp Gly Gln Gly Ser Arg Ala Ile Val Ala
610 615 620

Phe Asp Asn Ile Ser Ile Ser Leu Asp Cys Tyr Leu Thr Ile Ser Gly
625 630 635 640

Glu Asp Lys Ile Leu Gln Asn Thr Ala Pro Lys Ser Arg Asn Leu Phe
645 650 655

[0010] Glu Arg Asn Pro Asn Lys Glu Leu Lys Pro Gly Glu Asn Ser Pro Arg
660 665 670

Gln Thr Pro Ile Phe Asp Pro Thr Val His Trp Leu Phe Thr Thr Cys
675 680 685

Gly Ala Ser Gly Pro His Gly Pro Thr Gln Ala Gln Cys Asn Asn Ala
690 695 700

Tyr Gln Asn Ser Asn Leu Ser Val Glu Val Gly Ser Glu Gly Pro Leu
705 710 715 720

Lys Gly Ile Gln Ile Trp Lys Val Pro Ala Thr Asp Thr Tyr Ser Ile
725 730 735

Ser Gly Tyr Gly Ala Ala Gly Gly Lys Gly Gly Lys Asn Thr Met Met
740 745 750

Arg Ser His Gly Val Ser Val Leu Gly Ile Phe Asn Leu Glu Lys Asp
755 760 765

Asp Met Leu Tyr Ile Leu Val Gly Gln Gln Gly Glu Asp Ala Cys Pro
770 775 780

Ser Thr Asn Gln Leu Ile Gln Lys Val Cys Ile Gly Glu Asn Asn Val
785 790 795 800

Ile Glu Glu Glu Ile Arg Val Asn Arg Ser Val His Glu Trp Ala Gly
805 810 815

Gly Gly Gly Gly Gly Gly Gly Ala Thr Tyr Val Phe Lys Met Lys Asp
820 825 830

[0011] Gly Val Pro Val Pro Leu Ile Ile Ala Ala Gly Gly Gly Gly Arg Ala
835 840 845

Tyr Gly Ala Lys Thr Asp Thr Phe His Pro Glu Arg Leu Glu Asn Asn
850 855 860

Ser Ser Val Leu Gly Leu Asn Gly Asn Ser Gly Ala Ala Gly Gly Gly
865 870 875 880

Gly Gly Trp Asn Asp Asn Thr Ser Leu Leu Trp Ala Gly Lys Ser Leu
885 890 895

Gln Glu Gly Ala Thr Gly Gly His Ser Cys Pro Gln Ala Met Lys Lys
900 905 910

Trp Gly Trp Glu Thr Arg Gly Gly Phe Gly Gly Gly Gly Gly Gly Cys
915 920 925

Ser Ser Gly Gly Gly Gly Gly Gly Tyr Ile Gly Gly Asn Ala Ala Ser
930 935 940

Asn Asn Asp Pro Glu Met Asp Gly Glu Asp Gly Val Ser Phe Ile Ser
945 950 955 960

Pro Leu Gly Ile Leu Tyr Thr Pro Ala Leu Lys Val Met Glu Gly His
965 970 975

Gly Glu Val Asn Ile Lys His Tyr Leu Asn Cys Ser His Cys Glu Val
980 985 990

Asp Glu Cys His Met Asp Pro Glu Ser His Lys Val Ile Cys Phe Cys
995 1000 1005

[0012]

Asp His Gly Thr Val Leu Ala Glu Asp Gly Val Ser Cys Ile Val
1010 1015 1020

Ser Pro Thr Pro Glu Pro His Leu Pro Leu Ser Leu Ile Leu Ser
1025 1030 1035

Val Val Thr Ser Ala Leu Val Ala Ala Leu Val Leu Ala Phe Ser
1040 1045 1050

Gly Ile Met Ile Val Tyr Arg Arg Lys His Gln Glu Leu Gln Ala
1055 1060 1065

Met Gln Met Glu Leu Gln Ser Pro Glu Tyr Lys Leu Ser Lys Leu
1070 1075 1080

Arg Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro Asn Tyr Cys Phe
1085 1090 1095

Ala Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu Val Pro Arg
1100 1105 1110

Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly His Gly Ala Phe Gly
1115 1120 1125

Glu Val Tyr Glu Gly Gln Val Ser Gly Met Pro Asn Asp Pro Ser
1130 1135 1140

Pro Leu Gln Val Ala Val Lys Thr Leu Pro Glu Val Cys Ser Glu
1145 1150 1155

Gln Asp Glu Leu Asp Phe Leu Met Glu Ala Leu Ile Ile Ser Lys
1160 1165 1170

[0013]

Phe Asn His Gln Asn Ile Val Arg Cys Ile Gly Val Ser Leu Gln
1175 1180 1185

Ser Leu Pro Arg Phe Ile Leu Leu Glu Leu Met Ala Gly Gly Asp
1190 1195 1200

Leu Lys Ser Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln Pro
1205 1210 1215

Ser Ser Leu Ala Met Leu Asp Leu Leu His Val Ala Arg Asp Ile
1220 1225 1230

Ala Cys Gly Cys Gln Tyr Leu Glu Glu Asn His Phe Ile His Arg
1235 1240 1245

Asp Ile Ala Ala Arg Asn Cys Leu Leu Thr Cys Pro Gly Pro Gly
1250 1255 1260

Arg Val Ala Lys Ile Gly Asp Phe Gly Met Ala Arg Asp Ile Tyr
1265 1270 1275

Arg Ala Ser Tyr Tyr Arg Lys Gly Gly Cys Ala Met Leu Pro Val
1280 1285 1290

Lys Trp Met Pro Pro Glu Ala Phe Met Glu Gly Ile Phe Thr Ser
1295 1300 1305

Lys Thr Asp Thr Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Phe
1310 1315 1320

Ser Leu Gly Tyr Met Pro Tyr Pro Ser Lys Ser Asn Gln Glu Val
1325 1330 1335

[0014]

Leu Glu Phe Val Thr Ser Gly Gly Arg Met Asp Pro Pro Lys Asn
1340 1345 1350

Cys Pro Gly Pro Val Tyr Arg Ile Met Thr Gln Cys Trp Gln His
1355 1360 1365

Gln Pro Glu Asp Arg Pro Asn Phe Ala Ile Ile Leu Glu Arg Ile
1370 1375 1380

Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile Asn Thr Ala Leu Pro
1385 1390 1395

Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu Glu Lys Val Pro Val
1400 1405 1410

Arg Pro Lys Asp Pro Glu Gly Val Pro Pro Leu Leu Val Ser Gln
1415 1420 1425

Gln Ala Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro Pro Pro
1430 1435 1440

Leu Pro Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr Ala
1445 1450 1455

Ala Glu Ile Ser Val Arg Val Pro Arg Gly Pro Ala Val Glu Gly
1460 1465 1470

Gly His Val Asn Met Ala Phe Ser Gln Ser Asn Pro Pro Ser Glu
1475 1480 1485

Leu His Lys Val His Gly Ser Arg Asn Lys Pro Thr Ser Leu Trp
1490 1495 1500

[0015]

Asn Pro Thr Tyr Gly Ser Trp Phe Thr Glu Lys Pro Thr Lys Lys
1505 1510 1515

Asn Asn Pro Ile Ala Lys Lys Glu Pro His Asp Arg Gly Asn Leu
1520 1525 1530

Gly Leu Glu Gly Ser Cys Thr Val Pro Pro Asn Val Ala Thr Gly
1535 1540 1545

Arg Leu Pro Gly Ala Ser Leu Leu Leu Glu Pro Ser Ser Leu Thr
1550 1555 1560

Ala Asn Met Lys Glu Val Pro Leu Phe Arg Leu Arg His Phe Pro
1565 1570 1575

Cys Gly Asn Val Asn Tyr Gly Tyr Gln Gln Gln Gly Leu Pro Leu
1580 1585 1590

Glu Ala Ala Thr Ala Pro Gly Ala Gly His Tyr Glu Asp Thr Ile
1595 1600 1605

Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly Pro
1610 1615 1620

<210> 5
<211> 3926
<212> DNA
<213> 人类

<400> 5
ggcggcgcgg cgcggcgctc gcggctgctg cctgggaggg aggccgggca ggcggctgag 60

cggcgcggct ctcaacgtga cggggaagtg gttcgggcgg ccgcggctta ctaccccagg 120

gcgaacggac ggacgacgga ggcgggagcc ggtagccgag ccgggcgacc tagagaacga 180

[0016] gcgggtcagg ctcagcgtcg gccactctgt cggtccgctg aatgaagtgc ccgccccctct 240

gagcccggag cccggcgctt tccccgcaag atggacggtt tcgccggcag tctcgatgat 300

agtatttctg ctgcaagtac ttctgatgtt caagatcgcc tgtcagctct tgagtcacga 360

gttcagcaac aagaagatga aatcactgtg ctaaaggcgg ctttggctga tgttttgagg 420

cgtcttgcaa tctctgaaga tcatgtggcc tcagtgaaaa aatcagtctc aagtaaaggc 480

caaccaagcc ctcgagcagt tattcccatg tcctgtataa ccaatggaag tggtgcaaac 540

agaaaaccaa gtcataccag tgctgtctca attgcaggaa aagaaactct ttcatctgct 600

gctaaaagtg gtacagaaaa aaagaaagaa aaaccacaag gacagagaga aaaaaaagag 660

gaatctcatt ctaatgatca aagtccacaa attcgagcat caccttctcc ccagccctct 720

tcacaacctc tccaaataca cagacaaact ccagaaagca agaatgctac tcccaccaaa 780

agcataaaac gaccatcacc agctgaaaag tcacataatt cttgggaaaa ttcagatgat 840

agccgtaata aattgtcgaa aataccttca acacccaaat taataccaaa agttaccaaa 900

actgcagaca agcataaaga tgtcatcatc aaccaagaag gagaatatat taaaatgttt 960

atgcgcggtc ggccaattac catgttcatt ccttccgatg ttgacaacta tgatgacatc 1020

agaacggaac tgcctcctga gaagctcaaa ctggagtggg catatggtta tcgaggaaag 1080

gactgtagag ctaatgttta ccttcttccg accggggaaa tagtttattt cattgcatca 1140

gtagtagtac tatttaatta tgaggagaga actcagcgac actacctggg ccatacagac 1200

tgtgtgaaat gccttgctat acatcctgac aaaattagga ttgcaactgg acagatagct 1260

ggcgtggata aagatggaag gcctctacaa ccccacgtca gagtgtggga ttctgttact 1320

ctatccacac tgcagattat tggacttggc acttttgagc gtggagtagg atgcctggat 1380

ttttcaaaag cagattcagg tgttcattta tgtgttattg atgactccaa tgagcatatg 1440

cttactgtat gggactggca gaagaaagca aaaggagcag aaataaagac aacaaatgaa 1500

[0017]

gttgttttgg ctgtggagtt tcaccccaaca gatgcaaata ccataattac atgcggtaaa 1560

tctcatattt tcttctggac ctggagcggc aattcactaa caagaaaaca gggaatttttt 1620

gggaaatatg aaaagccaaa atttgtgcag tgtttagcat tcttggggaa tggagatgtt 1680

cttactggag actcaggtgg agtcatgctt atatggagca aaactactgt agagcccaca 1740

cctgggaaag gacctaaagt gtaccgccgg aagcaccagg agctgcaagc catgcagatg 1800

gagctgcaga gccctgagta caagctgagc aagctccgca cctcgaccat catgaccgac 1860

tacaacccca actactgctt tgctggcaag acctcctcca tcagtgacct gaaggaggtg 1920

ccgcggaaaa acatcaccct cattcggggt ctgggccatg gagcctttgg ggaggtgtat 1980

gaaggccagg tgtccggaat gcccaacgac ccaagccccc tgcaagtggc tgtgaagacg 2040

ctgcctgaag tgtgctctga acaggacgaa ctggatttcc tcatggaagc cctgatcatc 2100

agcaaattca accaccagaa cattgttcgc tgcattgggg tgagcctgca atccctgccc 2160

cggttcatcc tgctggagct catggcgggg ggagacctca agtccttcct ccgagagacc 2220

cgccctcgcc cgagccagcc ctcctccctg gccatgctgg accttctgca cgtggctcgg 2280

gacattgcct gtggctgtca gtatttggag gaaaaccact tcatccaccg agacattgct 2340

gccagaaact gcctcttgac ctgtccaggc cctggaagag tggccaagat tggagacttc 2400

gggatggccc gagacatcta cagggcgagc tactatagaa agggaggctg tgccatgctg 2460

ccagttaagt ggatgccccc agaggccttc atggaaggaa tattcacttc taaaacagac 2520

acatggtcct ttggagtgct gctatgggaa atcttttctc ttggatatat gccatacccc 2580

agcaaaagca accaggaagt tctggagttt gtcaccagtg gaggccggat ggacccaccc 2640

aagaactgcc ctgggcctgt ataccggata atgactcagt gctggcaaca tcagcctgaa 2700

gacaggccca actttgccat cattttggag aggattgaat actgcaccca ggacccggat 2760

gtaatcaaca ccgctttgcc gatagaatat ggtccacttg tggaagagga agagaaagtg 2820

[0018]

cctgtgaggc ccaaggaccc tgagggggtt cctcctctcc tggtctctca acaggcaaaa 2880

cgggaggagg agcgcagccc agctgcccca ccacctctgc ctaccacctc ctctggcaag 2940

gctgcaaaga aacccacagc tgcagaggtc tctgttcgag tccctagagg gccggccgtg 3000

gaaggggggac acgtgaatat ggcattctct cagtccaacc ctccttcgga gttgcacagg 3060

gtccacggat ccagaaacaa gcccaccagc ttgtggaacc caacgtacgg ctcctggttt 3120

acagagaaac ccaccaaaaa gaataatcct atagcaaaga aggagccaca cgagaggggt 3180

aacctggggc tggagggaag ctgtactgtc ccacctaacg ttgcaactgg gagacttccg 3240

ggggcctcac tgctcctaga gccctcttcg ctgactgcca atatgaagga ggtacctctg 3300

ttcaggctac gtcacttccc ttgtgggaat gtcaattacg gctaccagca acagggcttg 3360

cccttagaag ccgctactgc ccctggagct ggtcattacg aggataccat tctgaaaagc 3420

aagaatagca tgaaccagcc tgggccctga gctcggtcac acactcactt ctcttccttg 3480

ggatccctaa gaccgtggag gagagagagg caatcaatgg ctccttcaca aaccagagac 3540

caaatgtcac gttttgtttt gtgccaacct attttgaagt accaccaaaa aagctgtatt 3600

ttgaaaatgc tttagaaagg ttttgagcat gggttcatcc tattctttcg aaagaagaaa 3660

atatcataaa aatgagtgat aaatacaagg cccagatgtg gttgcataag gtttttatgc 3720

atgtttgttg tatacttcct tatgcttctt ttaaattgtg tgtgctctgc ttcaatgtag 3780

tcagaattag ctgcttctat gtttcatagt tggggtcata gatgtttcct tgccttgttg 3840

atgtggacat gagccatttg aggggagagg gaacggaaat aaaggagtta tttgtaatga 3900

aaaaaaaaaa aaaaaaaaaa aaaaaa 3926

<210> 6
<211> 1059
<212> PRT
<213> 人类

[0019]

<400> 6
Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser
1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln
20 25 30

Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val
35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys
50 55 60

Ser Val Ser Ser Lys Gly Gln Pro Ser Pro Arg Ala Val Ile Pro Met
65 70 75 80

Ser Cys Ile Thr Asn Gly Ser Gly Ala Asn Arg Lys Pro Ser His Thr

85 90 95

Ser Ala Val Ser Ile Ala Gly Lys Glu Thr Leu Ser Ser Ala Ala Lys
100 105 110

Ser Gly Thr Glu Lys Lys Lys Glu Lys Pro Gln Gly Gln Arg Glu Lys
115 120 125

Lys Glu Glu Ser His Ser Asn Asp Gln Ser Pro Gln Ile Arg Ala Ser
130 135 140

Pro Ser Pro Gln Pro Ser Ser Gln Pro Leu Gln Ile His Arg Gln Thr
145 150 155 160

Pro Glu Ser Lys Asn Ala Thr Pro Thr Lys Ser Ile Lys Arg Pro Ser
165 170 175

[0020]

Pro Ala Glu Lys Ser His Asn Ser Trp Glu Asn Ser Asp Asp Ser Arg
180 185 190

Asn Lys Leu Ser Lys Ile Pro Ser Thr Pro Lys Leu Ile Pro Lys Val
195 200 205

Thr Lys Thr Ala Asp Lys His Lys Asp Val Ile Ile Asn Gln Glu Gly
210 215 220

Glu Tyr Ile Lys Met Phe Met Arg Gly Arg Pro Ile Thr Met Phe Ile
225 230 235 240

Pro Ser Asp Val Asp Asn Tyr Asp Asp Ile Arg Thr Glu Leu Pro Pro
245 250 255

Glu Lys Leu Lys Leu Glu Trp Ala Tyr Gly Tyr Arg Gly Lys Asp Cys

260 265 270

Arg Ala Asn Val Tyr Leu Leu Pro Thr Gly Glu Ile Val Tyr Phe Ile
275 280 285

Ala Ser Val Val Val Leu Phe Asn Tyr Glu Glu Arg Thr Gln Arg His
290 295 300

Tyr Leu Gly His Thr Asp Cys Val Lys Cys Leu Ala Ile His Pro Asp
305 310 315 320

Lys Ile Arg Ile Ala Thr Gly Gln Ile Ala Gly Val Asp Lys Asp Gly
325 330 335

Arg Pro Leu Gln Pro His Val Arg Val Trp Asp Ser Val Thr Leu Ser
340 345 350

[0021]

Thr Leu Gln Ile Ile Gly Leu Gly Thr Phe Glu Arg Gly Val Gly Cys
355 360 365

Leu Asp Phe Ser Lys Ala Asp Ser Gly Val His Leu Cys Val Ile Asp
370 375 380

Asp Ser Asn Glu His Met Leu Thr Val Trp Asp Trp Gln Lys Lys Ala
385 390 395 400

Lys Gly Ala Glu Ile Lys Thr Thr Asn Glu Val Val Leu Ala Val Glu
405 410 415

Phe His Pro Thr Asp Ala Asn Thr Ile Ile Thr Cys Gly Lys Ser His
420 425 430

Ile Phe Phe Trp Thr Trp Ser Gly Asn Ser Leu Thr Arg Lys Gln Gly

435 440 445

Ile Phe Gly Lys Tyr Glu Lys Pro Lys Phe Val Gln Cys Leu Ala Phe
450 455 460

Leu Gly Asn Gly Asp Val Leu Thr Gly Asp Ser Gly Gly Val Met Leu
465 470 475 480

Ile Trp Ser Lys Thr Thr Val Glu Pro Thr Pro Gly Lys Gly Pro Lys
485 490 495

Val Tyr Arg Arg Lys His Gln Glu Leu Gln Ala Met Gln Met Glu Leu
500 505 510

Gln Ser Pro Glu Tyr Lys Leu Ser Lys Leu Arg Thr Ser Thr Ile Met
515 520 525

[0022]

Thr Asp Tyr Asn Pro Asn Tyr Cys Phe Ala Gly Lys Thr Ser Ser Ile
530 535 540

Ser Asp Leu Lys Glu Val Pro Arg Lys Asn Ile Thr Leu Ile Arg Gly
545 550 555 560

Leu Gly His Gly Ala Phe Gly Glu Val Tyr Glu Gly Gln Val Ser Gly
565 570 575

Met Pro Asn Asp Pro Ser Pro Leu Gln Val Ala Val Lys Thr Leu Pro
580 585 590

Glu Val Cys Ser Glu Gln Asp Glu Leu Asp Phe Leu Met Glu Ala Leu
595 600 605

Ile Ile Ser Lys Phe Asn His Gln Asn Ile Val Arg Cys Ile Gly Val

610 615 620

Ser Leu Gln Ser Leu Pro Arg Phe Ile Leu Leu Glu Leu Met Ala Gly
625 630 635 640

Gly Asp Leu Lys Ser Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln
645 650 655

Pro Ser Ser Leu Ala Met Leu Asp Leu Leu His Val Ala Arg Asp Ile
660 665 670

Ala Cys Gly Cys Gln Tyr Leu Glu Glu Asn His Phe Ile His Arg Asp
675 680 685

Ile Ala Ala Arg Asn Cys Leu Leu Thr Cys Pro Gly Pro Gly Arg Val
690 695 700

[0023]

Ala Lys Ile Gly Asp Phe Gly Met Ala Arg Asp Ile Tyr Arg Ala Ser
705 710 715 720

Tyr Tyr Arg Lys Gly Gly Cys Ala Met Leu Pro Val Lys Trp Met Pro
725 730 735

Pro Glu Ala Phe Met Glu Gly Ile Phe Thr Ser Lys Thr Asp Thr Trp
740 745 750

Ser Phe Gly Val Leu Leu Trp Glu Ile Phe Ser Leu Gly Tyr Met Pro
755 760 765

Tyr Pro Ser Lys Ser Asn Gln Glu Val Leu Glu Phe Val Thr Ser Gly
770 775 780

Gly Arg Met Asp Pro Pro Lys Asn Cys Pro Gly Pro Val Tyr Arg Ile

785 790 795 800

Met Thr Gln Cys Trp Gln His Gln Pro Glu Asp Arg Pro Asn Phe Ala
805 810 815

Ile Ile Leu Glu Arg Ile Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile
820 825 830

Asn Thr Ala Leu Pro Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu Glu
835 840 845

Lys Val Pro Val Arg Pro Lys Asp Pro Glu Gly Val Pro Pro Leu Leu
850 855 860

Val Ser Gln Gln Ala Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro
865 870 875 880

[0024]

Pro Pro Leu Pro Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr
885 890 895

Ala Ala Glu Val Ser Val Arg Val Pro Arg Gly Pro Ala Val Glu Gly
900 905 910

Gly His Val Asn Met Ala Phe Ser Gln Ser Asn Pro Pro Ser Glu Leu
915 920 925

His Arg Val His Gly Ser Arg Asn Lys Pro Thr Ser Leu Trp Asn Pro
930 935 940

Thr Tyr Gly Ser Trp Phe Thr Glu Lys Pro Thr Lys Lys Asn Asn Pro
945 950 955 960

Ile Ala Lys Lys Glu Pro His Glu Arg Gly Asn Leu Gly Leu Glu Gly

965 970 975

Ser Cys Thr Val Pro Pro Asn Val Ala Thr Gly Arg Leu Pro Gly Ala
980 985 990

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

Pro Leu Phe Arg Leu Arg His Phe Pro Cys Gly Asn Val Asn Tyr
1010 1015 1020

Gly Tyr Gln Gln Gln Gly Leu Pro Leu Glu Ala Ala Thr Ala Pro
1025 1030 1035

Gly Ala Gly His Tyr Glu Asp Thr Ile Leu Lys Ser Lys Asn Ser
1040 1045 1050

[0025]

Met Asn Gln Pro Gly Pro
1055

<210> 7
<211> 4679
<212> DNA
<213> 人类

<400> 7
ggcggcgcgg cgcggcgctc gcggctgctg cctgggaggg aggccgggca ggcggctgag 60

cggcgcggct ctcaacgtga cggggaagtg gttcgggcgg ccgcggctta ctaccccagg 120

gcgaacggac ggacgacgga ggcgggagcc ggtagccgag ccgggcgacc tagagaacga 180

gcgggtcagg ctcagcgtcg gccactctgt cggtccgctg aatgaagtgc ccgccctctc 240

gagcccggag cccggcgctt tccccgcaag atggacggtt tcgccggcag tctcgatgat 300

agtatttctg ctgcaagtac ttctgatgtt caagatcgcc tgtcagctct tgagtcacga 360

gttcagcaac aagaagatga aatcactgtg ctaaaggcgg ctttggctga tgttttgagg 420

cgtcttgcaa tctctgaaga tcatgtggcc tcagtgaaaa aatcagtctc aagtaaaggc 480

caaccaagcc ctcgagcagt tattcccatg tcctgtataa ccaatggaag tggtgcaaac 540

agaaaaccaa gtcataccag tgctgtctca attgcaggaa aagaaactct ttcatctgct 600

gctaaaagtg gtacagaaaa aaagaaagaa aaaccacaag gacagagaga aaaaaaagag 660

gaatctcatt ctaatgatca aagtccacaa attcgagcat caccttctcc ccagccctct 720

tcacaacctc tccaaataca cagacaaact ccagaaagca agaatgctac tcccaccaaa 780

agcataaaac gaccatcacc agctgaaaag tcacataatt cttgggaaaa ttcagatgat 840

agccgtaata aattgtcgaa aataccttca acacccaaat taataccaaa agttaccaaa 900

actgcagaca agcataaaga tgtcatcatc aaccaagaag gagaatatat taaaatgttt 960

atgcgcggtc ggccaattac catgttcatt ccttccgatg ttgacaacta tgatgacatc 1020

[0026]

agaacggaac tgcctcctga gaagctcaaa ctggagtggg catatggtta tcgaggaaag 1080

gactgtagag ctaatgttta ccttcttccg accggggaaa tagtttattt cattgcatca 1140

gtagtagtac tatttaatta tgaggagaga actcagcgac actacctggg ccatacagac 1200

tgtgtgaaat gccttgctat acatcctgac aaaattagga ttgcaactgg acagatagct 1260

ggcgtggata aagatggaag gcctctacaa ccccacgtca gagtgtggga ttctgttact 1320

ctatccacac tgcagattat tggacttggc acttttgagc gtggagtagg atgcctggat 1380

ttttcaaaag cagattcagg tgttcattta tgtgttattg atgactccaa tgagcatatg 1440

cttactgtat gggactggca gaagaaagca aaaggagcag aaataaagac aacaaatgaa 1500

gttgttttgg ctgtggagtt tcacccaaca gatgcaaata ccataattac atgcggtaaa 1560

tctcatattt tcttctggac ctggagcggc aattcactaa caagaaaaca gggaatttttt 1620

gggaaatatg aaaagccaaa atttgtgcag tgtttagcat tcttggggaa tggagatgtt 1680

cttactggag actcaggtgg agtcatgctt atatggagca aaactactgt agagcccaca 1740

cctgggaaag gacctaaagg tgtatatcaa atcagcaaac aaatcaaagc tcatgatggc 1800

agtgtgttca cactttgtca gatgagaaat gggatgttat taactggagg agggaaagac 1860

agaaaaataa ttctgtggga tcatgatctg aatcctgaaa gagaaataga ggttcctgat 1920

cagtatggca caatcagagc tgtagcagaa ggaaaggcag atcaatttt agtaggcaca 1980

tcacgaaact ttatttttacg aggaacattt aatgatggct tccaaataga agtacagggt 2040

catacagatg agctttgggg tcttgccaca catcccttca aagatttgct cttgacatgt 2100

gctcaggaca ggcaggtgtg cctgtggaac tcaatggaac acaggctgga atggaccagg 2160

ctggtagatg aaccaggaca ctgtgcagat tttcatccaa gtggcacagt ggtggccata 2220

ggaacgcact caggcaggtg gtttgttctg gatgcagaaa ccagagatct agtttctatc 2280

cacacagacg ggaatgaaca gctctctgtg atgcgctact caatagatgg taccttcctg 2340

[0027]

gctgtaggat ctcatgacaa ctttatttac ctctatgtag tctctgaaaa tggaagaaaa 2400

tatagcagat atggaaggtg cactggacat tccagctaca tcacacacct tgactggtcc 2460

ccagacaaca agtatataat gtctaactcg ggagactatg aaatattgta cttgtaccgc 2520

cggaagcacc aggagctgca agccatgcag atggagctgc agagccctga gtacaagctg 2580

agcaagctcc gcacctcgac catcatgacc gactacaacc ccaactactg ctttgctggc 2640

aagacctcct ccatcagtga cctgaaggag gtgccgcgga aaaacatcac cctcattcgg 2700

ggtctgggcc atggagcctt tggggaggtg tatgaaggcc aggtgtccgg aatgcccaac 2760

gacccaagcc ccctgcaagt ggctgtgaag acgctgcctg aagtgtgctc tgaacaggac 2820

gaactggatt tcctcatgga agccctgatc atcagcaaat tcaaccacca gaacattgtt 2880

cgctgcattg gggtgagcct gcaatccctg ccccggttca tcctgctgga gctcatggcg 2940

gggggagacc tcaagtcctt cctccgagag acccgccctc gccccgagcca gccctcctcc 3000

ctggccatgc tggaccttct gcacgtggct cgggacattg cctgtggctg tcagtatttg 3060

gaggaaaacc acttcatcca ccgagacatt gctgccagaa actgcctctt gacctgtcca 3120

ggccctggaa gagtggccaa gattggagac ttcgggatgg cccgagacat ctacagggcg 3180

agctactata gaaagggagg ctgtgccatg ctgccagtta agtggatgcc cccagaggcc 3240

ttcatggaag gaatattcac ttctaaaaca gacacatggt cctttggagt gctgctatgg 3300

gaaatctttt ctcttggata tatgccatac cccagcaaaa gcaaccagga agttctggag 3360

tttgtcacca gtggaggccg gatggaccca cccaagaact gccctgggcc tgtataccgg 3420

ataatgactc agtgctggca acatcagcct gaagacaggc ccaactttgc catcattttg 3480

gagaggattg aatactgcac ccaggacccg gatgtaatca acaccgcttt gccgatagaa 3540

tatggtccac ttgtggaaga ggaagagaaa gtgcctgtga ggcccaagga ccctgagggg 3600

gttcctcctc tcctggtctc tcaacaggca aaacgggagg aggagcgcag cccagctgcc 3660

[0028]

ccaccacctc tgcctaccac ctcctctggc aaggctgcaa agaaacccac agctgcagag 3720

gtctctgttc gagtccctag agggccggcc gtggaagggg gacacgtgaa tatggcattc 3780

tctcagtcca accctccttc ggagttgcac agggtccacg gatccagaaa caagcccacc 3840

agcttgtgga acccaacgta cggctcctgg tttacagaga aacccaccaa aaagaataat 3900

cctatagcaa agaaggagcc acacgagagg ggtaacctgg ggctggaggg aagctgtact 3960

gtcccaccta acgttgcaac tgggagactt ccgggggcct cactgctcct agagccctct 4020

tcgctgactg ccaatatgaa ggaggtacct ctgttcaggc tacgtcactt cccttgtggg 4080

aatgtcaatt acggctacca gcaacagggc ttgcccttag aagccgctac tgccccctgga 4140

gctggtcatt acgaggatac cattctgaaa agcaagaata gcatgaacca gcctgggccc 4200

tgagctcggt cacacactca cttctcttcc ttgggatccc taagaccgtg gaggagagag 4260

aggcaatcaa tggctccttc acaaaccaga gaccaaatgt cacgttttgt tttgtgccaa 4320

cctattttga agtaccacca aaaaagctgt attttgaaaa tgctttagaa aggttttgag 4380

catgggttca tcctattctt tcgaaagaag aaaatatcat aaaaatgagt gataaataca 4440

aggcccagat gtggttgcat aaggttttta tgcatgtttg ttgtatactt ccttatgctt 4500

cttttaaatt gtgtgtgctc tgcttcaatg tagtcagaat tagctgcttc tatgtttcat 4560

agttggggtc atagatgttt ccttgccttg ttgatgtgga catgagccat ttgaggggag 4620

agggaacgga aataaaggag ttatttgtaa tgaaaaaaaa aaaaaaaaaa aaaaaaaaa 4679

<210> 8
<211> 1310
<212> PRT
<213> 人类

<400> 8
Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser
[0029] 1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln
20 25 30

Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val
35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys
50 55 60

Ser Val Ser Ser Lys Gly Gln Pro Ser Pro Arg Ala Val Ile Pro Met
65 70 75 80

Ser Cys Ile Thr Asn Gly Ser Gly Ala Asn Arg Lys Pro Ser His Thr
85 90 95

Ser Ala Val Ser Ile Ala Gly Lys Glu Thr Leu Ser Ser Ala Ala Lys
100 105 110

Ser Gly Thr Glu Lys Lys Lys Glu Lys Pro Gln Gly Gln Arg Glu Lys
115 120 125

Lys Glu Glu Ser His Ser Asn Asp Gln Ser Pro Gln Ile Arg Ala Ser
130 135 140

Pro Ser Pro Gln Pro Ser Ser Gln Pro Leu Gln Ile His Arg Gln Thr
145 150 155 160

Pro Glu Ser Lys Asn Ala Thr Pro Thr Lys Ser Ile Lys Arg Pro Ser
165 170 175

[0030] Pro Ala Glu Lys Ser His Asn Ser Trp Glu Asn Ser Asp Asp Ser Arg
180 185 190

Asn Lys Leu Ser Lys Ile Pro Ser Thr Pro Lys Leu Ile Pro Lys Val
195 200 205

Thr Lys Thr Ala Asp Lys His Lys Asp Val Ile Ile Asn Gln Glu Gly
210 215 220

Glu Tyr Ile Lys Met Phe Met Arg Gly Arg Pro Ile Thr Met Phe Ile
225 230 235 240

Pro Ser Asp Val Asp Asn Tyr Asp Asp Ile Arg Thr Glu Leu Pro Pro
245 250 255

Glu Lys Leu Lys Leu Glu Trp Ala Tyr Gly Tyr Arg Gly Lys Asp Cys
260 265 270

Arg Ala Asn Val Tyr Leu Leu Pro Thr Gly Glu Ile Val Tyr Phe Ile
275 280 285

Ala Ser Val Val Val Leu Phe Asn Tyr Glu Glu Arg Thr Gln Arg His
290 295 300

Tyr Leu Gly His Thr Asp Cys Val Lys Cys Leu Ala Ile His Pro Asp
305 310 315 320

Lys Ile Arg Ile Ala Thr Gly Gln Ile Ala Gly Val Asp Lys Asp Gly
325 330 335

Arg Pro Leu Gln Pro His Val Arg Val Trp Asp Ser Val Thr Leu Ser
340 345 350

[0031] Thr Leu Gln Ile Ile Gly Leu Gly Thr Phe Glu Arg Gly Val Gly Cys
355 360 365

Leu Asp Phe Ser Lys Ala Asp Ser Gly Val His Leu Cys Val Ile Asp
370 375 380

Asp Ser Asn Glu His Met Leu Thr Val Trp Asp Trp Gln Lys Lys Ala
385 390 395 400

Lys Gly Ala Glu Ile Lys Thr Thr Asn Glu Val Val Leu Ala Val Glu
405 410 415

Phe His Pro Thr Asp Ala Asn Thr Ile Ile Thr Cys Gly Lys Ser His
420 425 430

Ile Phe Phe Trp Thr Trp Ser Gly Asn Ser Leu Thr Arg Lys Gln Gly
435 440 445

Ile Phe Gly Lys Tyr Glu Lys Pro Lys Phe Val Gln Cys Leu Ala Phe
450 455 460

Leu Gly Asn Gly Asp Val Leu Thr Gly Asp Ser Gly Gly Val Met Leu
465 470 475 480

Ile Trp Ser Lys Thr Thr Val Glu Pro Thr Pro Gly Lys Gly Pro Lys
485 490 495

Gly Val Tyr Gln Ile Ser Lys Gln Ile Lys Ala His Asp Gly Ser Val
500 505 510

Phe Thr Leu Cys Gln Met Arg Asn Gly Met Leu Leu Thr Gly Gly Gly
515 520 525

[0032] Lys Asp Arg Lys Ile Ile Leu Trp Asp His Asp Leu Asn Pro Glu Arg
530 535 540

Glu Ile Glu Val Pro Asp Gln Tyr Gly Thr Ile Arg Ala Val Ala Glu
545 550 555 560

Gly Lys Ala Asp Gln Phe Leu Val Gly Thr Ser Arg Asn Phe Ile Leu
565 570 575

Arg Gly Thr Phe Asn Asp Gly Phe Gln Ile Glu Val Gln Gly His Thr
580 585 590

Asp Glu Leu Trp Gly Leu Ala Thr His Pro Phe Lys Asp Leu Leu Leu
595 600 605

Thr Cys Ala Gln Asp Arg Gln Val Cys Leu Trp Asn Ser Met Glu His
610 615 620

Arg Leu Glu Trp Thr Arg Leu Val Asp Glu Pro Gly His Cys Ala Asp
625 630 635 640

Phe His Pro Ser Gly Thr Val Val Ala Ile Gly Thr His Ser Gly Arg
645 650 655

Trp Phe Val Leu Asp Ala Glu Thr Arg Asp Leu Val Ser Ile His Thr
660 665 670

Asp Gly Asn Glu Gln Leu Ser Val Met Arg Tyr Ser Ile Asp Gly Thr
675 680 685

Phe Leu Ala Val Gly Ser His Asp Asn Phe Ile Tyr Leu Tyr Val Val
690 695 700

[0033] Ser Glu Asn Gly Arg Lys Tyr Ser Arg Tyr Gly Arg Cys Thr Gly His
705 710 715 720

Ser Ser Tyr Ile Thr His Leu Asp Trp Ser Pro Asp Asn Lys Tyr Ile
725 730 735

Met Ser Asn Ser Gly Asp Tyr Glu Ile Leu Tyr Leu Tyr Arg Arg Lys
740 745 750

His Gln Glu Leu Gln Ala Met Gln Met Glu Leu Gln Ser Pro Glu Tyr
755 760 765

Lys Leu Ser Lys Leu Arg Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro
770 775 780

Asn Tyr Cys Phe Ala Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu
785 790 795 800

Val Pro Arg Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly His Gly Ala
805 810 815

Phe Gly Glu Val Tyr Glu Gly Gln Val Ser Gly Met Pro Asn Asp Pro
820 825 830

Ser Pro Leu Gln Val Ala Val Lys Thr Leu Pro Glu Val Cys Ser Glu
835 840 845

Gln Asp Glu Leu Asp Phe Leu Met Glu Ala Leu Ile Ile Ser Lys Phe
850 855 860

Asn His Gln Asn Ile Val Arg Cys Ile Gly Val Ser Leu Gln Ser Leu
865 870 875 880

[0034] Pro Arg Phe Ile Leu Leu Glu Leu Met Ala Gly Gly Asp Leu Lys Ser
885 890 895

Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln Pro Ser Ser Leu Ala
900 905 910

Met Leu Asp Leu Leu His Val Ala Arg Asp Ile Ala Cys Gly Cys Gln
915 920 925

Tyr Leu Glu Glu Asn His Phe Ile His Arg Asp Ile Ala Ala Arg Asn
930 935 940

Cys Leu Leu Thr Cys Pro Gly Pro Gly Arg Val Ala Lys Ile Gly Asp
945 950 955 960

Phe Gly Met Ala Arg Asp Ile Tyr Arg Ala Ser Tyr Tyr Arg Lys Gly
965 970 975

Gly Cys Ala Met Leu Pro Val Lys Trp Met Pro Pro Glu Ala Phe Met
980 985 990

Glu Gly Ile Phe Thr Ser Lys Thr Asp Thr Trp Ser Phe Gly Val Leu
995 1000 1005

Leu Trp Glu Ile Phe Ser Leu Gly Tyr Met Pro Tyr Pro Ser Lys
1010 1015 1020

Ser Asn Gln Glu Val Leu Glu Phe Val Thr Ser Gly Gly Arg Met
1025 1030 1035

Asp Pro Pro Lys Asn Cys Pro Gly Pro Val Tyr Arg Ile Met Thr
1040 1045 1050

[0035]

Gln Cys Trp Gln His Gln Pro Glu Asp Arg Pro Asn Phe Ala Ile
1055 1060 1065

Ile Leu Glu Arg Ile Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile
1070 1075 1080

Asn Thr Ala Leu Pro Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu
1085 1090 1095

Glu Lys Val Pro Val Arg Pro Lys Asp Pro Glu Gly Val Pro Pro
1100 1105 1110

Leu Leu Val Ser Gln Gln Ala Lys Arg Glu Glu Glu Arg Ser Pro
1115 1120 1125

Ala Ala Pro Pro Pro Leu Pro Thr Thr Ser Ser Gly Lys Ala Ala
1130 1135 1140

Lys Lys Pro Thr Ala Ala Glu Val Ser Val Arg Val Pro Arg Gly
1145 1150 1155

Pro Ala Val Glu Gly Gly His Val Asn Met Ala Phe Ser Gln Ser
1160 1165 1170

Asn Pro Pro Ser Glu Leu His Arg Val His Gly Ser Arg Asn Lys
1175 1180 1185

Pro Thr Ser Leu Trp Asn Pro Thr Tyr Gly Ser Trp Phe Thr Glu
1190 1195 1200

Lys Pro Thr Lys Lys Asn Asn Pro Ile Ala Lys Lys Glu Pro His
1205 1210 1215

[0036]

Glu Arg Gly Asn Leu Gly Leu Glu Gly Ser Cys Thr Val Pro Pro
1220 1225 1230

Asn Val Ala Thr Gly Arg Leu Pro Gly Ala Ser Leu Leu Leu Glu
1235 1240 1245

Pro Ser Ser Leu Thr Ala Asn Met Lys Glu Val Pro Leu Phe Arg
1250 1255 1260

Leu Arg His Phe Pro Cys Gly Asn Val Asn Tyr Gly Tyr Gln Gln
1265 1270 1275

Gln Gly Leu Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala Gly His
1280 1285 1290

Tyr Glu Asp Thr Ile Leu Lys Ser Lys Asn Ser Met Asn Gln Pro
1295 1300 1305

Gly Pro
1310

<210> 9
<211> 2473
<212> DNA
<213> 人类

<400> 9
actctgtcgg tccgctgaat gaagtgcccg cccctctaag cccggagccc ggcgctttcc 60

ccgcaagatg gacggtttcg ccggcagtct cgatgatagt atttctgctg caagtacttc 120

tgatgttcaa gatcgcctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat 180

cactgtgcta aaggcggctt tggctgatgt tttgaggcgt cttgcaatct ctgaagatca 240

tgtggcctca gtgaaaaaat cagtctcaag taaaggccaa ccaagccctc gagcagttat 300

tcccatgtcc tgtataacca atggaagtgg tgcaaacaga aaaccaagtc ataccagtgc 360

[0037]

tgtctcaatt gcaggaaaag aaactctttc atctgctgct aaaagtggta cagaaaaaaa 420

gaaagaaaaa ccacaaggac agagagaaaa aaaagaggaa tctcattcta atgatcaaag 480

tccacaaatt cgagcatcac cttctcccca gccctcttca caacctctcc aaatacacag 540

acaaactcca gaaagcaaga atgctactcc caccaaaagc ataaaacgac catcaccagc 600

tgaaaagtca cataattctt gggaaaattc agatgatagc cgtaataaat tgtcgaaaat 660

accttcaaca cccaaattaa taccaaaagt taccaaaact gcagacaagc ataaagatgt 720

catcatcaac caagtgtacc gccggaagca ccaggagctg caagccatgc agatggagct 780

gcagagccct gagtacaagc tgagcaagct ccgcacctcg accatcatga ccgactacaa 840

ccccaactac tgctttgctg gcaagacctc ctccatcagt gacctgaagg aggtgccgcg 900

gaaaaacatc accctcattc ggggtctggg ccatggagcc tttggggagg tgtatgaagg 960

ccaggtgtcc ggaatgccca acgacccaag ccccctgcaa gtggctgtga agacgctgcc 1020

tgaagtgtgc tctgaacagg acgaactgga tttcctcatg gaagccctga tcatcagcaa 1080

attcaaccac cagaacattg ttcgctgcat tggggtgagc ctgcaatccc tgccccggtt 1140

catcctgctg gagctcatgg cgggggggaga cctcaagtcc ttcctccgag agacccgccc 1200

tcgcccgagc cagccctcct ccctggccat gctggacctt ctgcacgtgg ctcgggacat 1260

tgcctgtggc tgtcagtatt tggaggaaaa ccacttcatc caccgagaca ttgctgccag 1320

aaactgcctc ttgacctgtc caggccctgg aagagtggcc aagattggag acttcgggat 1380

ggcccgagac atctacaggg cgagctacta tagaaaggga ggctgtgcca tgctgccagt 1440

taagtggatg ccccccagagg ccttcatgga aggaatattc acttctaaaa cagacacatg 1500

gtcctttgga gtgctgctat gggaaatctt ttctcttgga tatatgccat accccagcaa 1560

aagcaaccag gaagttctgg agtttgtcac cagtggaggc cggatggacc cacccaagaa 1620

ctgccctggg cctgtatacc ggataatgac tcagtgctgg caacatcagc ctgaagacag 1680

[0038]

gcccaacttt gccatcattt tggagaggat tgaatactgc acccaggacc cggatgtaat 1740

caacaccgct ttgccgatag aatatggtcc acttgtggaa gaggaagaga aagtgcctgt 1800

gaggcccaag gaccctgagg gggttcctcc tctcctggtc tctcaacagg caaaacggga 1860

ggaggagcgc agcccagctg ccccaccacc tctgcctacc acctcctctg gcaaggctgc 1920

aaagaaaccc acagctgcag aggtctctgt tcgagtccct agagggccgg ccgtggaagg 1980

gggacacgtg aatatggcat tctctcagtc caaccctcct tcggagttgc acagggtcca 2040

cggatccaga aacaagccca ccagcttgtg gaacccaacg tacggctcct ggtttacaga 2100

gaaacccacc aaaaagaata atcctatagc aaagaaggag ccacacgaga ggggtaacct 2160

ggggctggag ggaagctgta ctgtcccacc taacgttgca actgggagac ttccgggggc 2220

ctcactgctc ctagagccct cttcgctgac tgccaatatg aaggaggtac ctctgttcag 2280

gctacgtcac ttcccttgtg ggaatgtcaa ttacggctac cagcaacagg gcttgccctt 2340

agaagccgct actgcccctg gagctggtca ttacgaggat accattctga aaagcaagaa 2400

tagcatgaac cagcctgggc cctgagctcg gtcgcacact cacttctctt ccttgggatc 2460

cctaagaccg tgg 2473

<210> 10
<211> 785
<212> PRT
<213> 人类

<400> 10
Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser
1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln
20 25 30

[0039] Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val
35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys
50 55 60

Ser Val Ser Ser Lys Gly Gln Pro Ser Pro Arg Ala Val Ile Pro Met
65 70 75 80

Ser Cys Ile Thr Asn Gly Ser Gly Ala Asn Arg Lys Pro Ser His Thr
85 90 95

Ser Ala Val Ser Ile Ala Gly Lys Glu Thr Leu Ser Ser Ala Ala Lys
100 105 110

Ser Gly Thr Glu Lys Lys Lys Glu Lys Pro Gln Gly Gln Arg Glu Lys
115 120 125

Lys Glu Glu Ser His Ser Asn Asp Gln Ser Pro Gln Ile Arg Ala Ser
130 135 140

Pro Ser Pro Gln Pro Ser Ser Gln Pro Leu Gln Ile His Arg Gln Thr
145 150 155 160

Pro Glu Ser Lys Asn Ala Thr Pro Thr Lys Ser Ile Lys Arg Pro Ser
165 170 175

Pro Ala Glu Lys Ser His Asn Ser Trp Glu Asn Ser Asp Asp Ser Arg
180 185 190

Asn Lys Leu Ser Lys Ile Pro Ser Thr Pro Lys Leu Ile Pro Lys Val
195 200 205

[0040] Thr Lys Thr Ala Asp Lys His Lys Asp Val Ile Ile Asn Gln Val Tyr
210 215 220

Arg Arg Lys His Gln Glu Leu Gln Ala Met Gln Met Glu Leu Gln Ser
225 230 235 240

Pro Glu Tyr Lys Leu Ser Lys Leu Arg Thr Ser Thr Ile Met Thr Asp
245 250 255

Tyr Asn Pro Asn Tyr Cys Phe Ala Gly Lys Thr Ser Ser Ile Ser Asp
260 265 270

Leu Lys Glu Val Pro Arg Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly
275 280 285

His Gly Ala Phe Gly Glu Val Tyr Glu Gly Gln Val Ser Gly Met Pro
290 295 300

Asn Asp Pro Ser Pro Leu Gln Val Ala Val Lys Thr Leu Pro Glu Val
305 310 315 320

Cys Ser Glu Gln Asp Glu Leu Asp Phe Leu Met Glu Ala Leu Ile Ile
325 330 335

Ser Lys Phe Asn His Gln Asn Ile Val Arg Cys Ile Gly Val Ser Leu
340 345 350

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

Leu Lys Ser Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln Pro Ser
370 375 380

[0041] Ser Leu Ala Met Leu Asp Leu Leu His Val Ala Arg Asp Ile Ala Cys
385 390 395 400

Gly Cys Gln Tyr Leu Glu Glu Asn His Phe Ile His Arg Asp Ile Ala
405 410 415

Ala Arg Asn Cys Leu Leu Thr Cys Pro Gly Pro Gly Arg Val Ala Lys
420 425 430

Ile Gly Asp Phe Gly Met Ala Arg Asp Ile Tyr Arg Ala Ser Tyr Tyr
435 440 445

Arg Lys Gly Gly Cys Ala Met Leu Pro Val Lys Trp Met Pro Pro Glu
450 455 460

Ala Phe Met Glu Gly Ile Phe Thr Ser Lys Thr Asp Thr Trp Ser Phe
465 470 475 480

Gly Val Leu Leu Trp Glu Ile Phe Ser Leu Gly Tyr Met Pro Tyr Pro
485 490 495

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

Met Asp Pro Pro Lys Asn Cys Pro Gly Pro Val Tyr Arg Ile Met Thr
515 520 525

Gln Cys Trp Gln His Gln Pro Glu Asp Arg Pro Asn Phe Ala Ile Ile
530 535 540

Leu Glu Arg Ile Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile Asn Thr
545 550 555 560

[0042] Ala Leu Pro Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu Glu Lys Val
565 570 575

Pro Val Arg Pro Lys Asp Pro Glu Gly Val Pro Pro Leu Leu Val Ser
580 585 590

Gln Gln Ala Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro Pro Pro
595 600 605

Leu Pro Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr Ala Ala
610 615 620

Glu Val Ser Val Arg Val Pro Arg Gly Pro Ala Val Glu Gly Gly His
625 630 635 640

Val Asn Met Ala Phe Ser Gln Ser Asn Pro Pro Ser Glu Leu His Arg
645 650 655

Val His Gly Ser Arg Asn Lys Pro Thr Ser Leu Trp Asn Pro Thr Tyr
660 665 670

Gly Ser Trp Phe Thr Glu Lys Pro Thr Lys Lys Asn Asn Pro Ile Ala
675 680 685

Lys Lys Glu Pro His Glu Arg Gly Asn Leu Gly Leu Glu Gly Ser Cys
690 695 700

Thr Val Pro Pro Asn Val Ala Thr Gly Arg Leu Pro Gly Ala Ser Leu
705 710 715 720

Leu Leu Glu Pro Ser Ser Leu Thr Ala Asn Met Lys Glu Val Pro Leu
725 730 735

Phe Arg Leu Arg His Phe Pro Cys Gly Asn Val Asn Tyr Gly Tyr Gln
[0043] 740 745 750

Gln Gln Gly Leu Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala Gly His
755 760 765

Tyr Glu Asp Thr Ile Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly
770 775 780

Pro
785

<210> 11
<211> 2506
<212> DNA
<213> 人类

<400> 11
actctgtcgg tccgctgaat gaagtgcccg cccctctaag cccggagccc ggcgctttcc 60

ccgcaagatg gacggtttcg ccggcagtct cgatgatagt atttctgctg caagtacttc 120

tgatgttcaa gatcgcctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat 180

cactgtgcta aaggcggctt tggctgatgt tttgaggcgt cttgcaatct ctgaagatca 240

tgtggcctca gtgaaaaaat cagtctcaag taaaggccaa ccaagccctc gagcagttat 300

tcccatgtcc tgtataacca atggaagtgg tgcaaacaga aaaccaagtc ataccagtgc 360

tgtctcaatt gcaggaaaag aaactctttc atctgctgct aaaagtggta cagaaaaaaa 420

gaaagaaaaa ccacaaggac agagagaaaa aaaagaggaa tctcattcta atgatcaaag 480

tccacaaatt cgagcatcac cttctcccca gccctcttca caacctctcc aaatacacag 540

acaaactcca gaaagcaaga atgctactcc caccaaaagc ataaaacgac catcaccagc 600

tgaaaagtca cataattctt gggaaaattc agatgatagc cgtaataaat tgtcgaaaat 660

accttcaaca cccaaattaa taccaaaagt taccaaaact gcagacaagc ataaagatgt 720

[0044]

catcatcaac caagcaaaaa tgtcaactcg cgaaaaaaac agccaagtgt accgccggaa 780

gcaccaggag ctgcaagcca tgcagatgga gctgcagagc cctgagtaca agctgagcaa 840

gctccgcacc tcgaccatca tgaccgacta caaccccaac tactgctttg ctggcaagac 900

ctcctccatc agtgacctga aggaggtgcc gcggaaaaac atcaccctca ttcggggtct 960

gggccatgga gcctttgggg aggtgtatga aggccaggtg tccggaatgc ccaacgaccc 1020

aagccccctg caagtggctg tgaagacgct gcctgaagtg tgctctgaac aggacgaact 1080

ggatttcctc atggaagccc tgatcatcag caaattcaac caccagaaca ttgttcgctg 1140

cattggggtg agcctgcaat ccctgccccg gttcatcctg ctggagctca tggcgggggg 1200

agacctcaag tccttcctcc gagagacccg ccctcgcccg agccagccct cctccctggc 1260

catgctggac cttctgcacg tggctcggga cattgcctgt ggctgtcagt atttggagga 1320

aaaccacttc atccaccgag acattgctgc cagaaactgc ctcttgacct gtccaggccc 1380

tggaagagtg gccaagattg gagacttcgg gatggcccga gacatctaca gggcgagcta 1440

ctatagaaag ggaggctgtg ccatgctgcc agttaagtgg atgcccccag aggccttcat 1500

ggaaggaata ttcacttcta aaacagacac atggtccttt ggagtgctgc tatgggaaat 1560

cttttctctt ggatatatgc cataccccag caaaagcaac caggaagttc tggagtttgt 1620

caccagtgga ggccggatgg acccacccaa gaactgccct gggcctgtat accggataat 1680

gactcagtgc tggcaacatc agcctgaaga caggcccaac tttgccatca ttttggagag 1740

gattgaatac tgcacccagg acccggatgt aatcaacacc gctttgccga tagaatatgg 1800

tccacttgtg gaagaggaag agaaagtgcc tgtgaggccc aaggaccctg agggggttcc 1860

tcctctcctg gtctctcaac aggcaaaacg ggaggaggag cgcagcccag ctgccccacc 1920

acctctgcct accacctcct ctggcaaggc tgcaaagaaa cccacagctg cagaggtctc 1980

tgttcgagtc cctagagggc cggccgtgga agggggacac gtgaatatgg cattctctca 2040

[0045]

gtccaaccct ccttcggagt tgcacagggt ccacggatcc agaaacaagc ccaccagctt 2100

gtggaaccca acgtacggct cctggtttac agagaaaccc accaaaaaga ataatcctat 2160

agcaaagaag gagccacacg agaggggtaa cctggggctg gagggaagct gtactgtccc 2220

acctaacgtt gcaactggga gacttccggg ggcctcactg ctcctagagc cctcttcgct 2280

gactgccaat atgaaggagg tacctctgtt caggctacgt cacttccctt gtgggaatgt 2340

caattacggc taccagcaac agggcttgcc cttagaagcc gctactgccc ctggagctgg 2400

tcattacgag ataccattc tgaaaagcaa gaatagcatg aaccagcctg ggccctgagc 2460

tcggtcgcac actcacttct cttccttggg atccctaaga ccgtgg 2506

<210> 12
<211> 796
<212> PRT
<213> 人类

<400> 12

Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser
1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln
20 25 30

Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val
35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys
50 55 60

Ser Val Ser Ser Lys Gly Gln Pro Ser Pro Arg Ala Val Ile Pro Met
65 70 75 80

[0046] Ser Cys Ile Thr Asn Gly Ser Gly Ala Asn Arg Lys Pro Ser His Thr
85 90 95

Ser Ala Val Ser Ile Ala Gly Lys Glu Thr Leu Ser Ser Ala Ala Lys
100 105 110

Ser Gly Thr Glu Lys Lys Lys Glu Lys Pro Gln Gly Gln Arg Glu Lys
115 120 125

Lys Glu Glu Ser His Ser Asn Asp Gln Ser Pro Gln Ile Arg Ala Ser
130 135 140

Pro Ser Pro Gln Pro Ser Ser Gln Pro Leu Gln Ile His Arg Gln Thr
145 150 155 160

Pro Glu Ser Lys Asn Ala Thr Pro Thr Lys Ser Ile Lys Arg Pro Ser
165 170 175

Pro Ala Glu Lys Ser His Asn Ser Trp Glu Asn Ser Asp Asp Ser Arg
180 185 190

Asn Lys Leu Ser Lys Ile Pro Ser Thr Pro Lys Leu Ile Pro Lys Val
195 200 205

Thr Lys Thr Ala Asp Lys His Lys Asp Val Ile Ile Asn Gln Ala Lys
210 215 220

Met Ser Thr Arg Glu Lys Asn Ser Gln Val Tyr Arg Arg Lys His Gln
225 230 235 240

Glu Leu Gln Ala Met Gln Met Glu Leu Gln Ser Pro Glu Tyr Lys Leu
245 250 255

[0047] Ser Lys Leu Arg Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro Asn Tyr
260 265 270

Cys Phe Ala Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu Val Pro
275 280 285

Arg Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly His Gly Ala Phe Gly
290 295 300

Glu Val Tyr Glu Gly Gln Val Ser Gly Met Pro Asn Asp Pro Ser Pro
305 310 315 320

Leu Gln Val Ala Val Lys Thr Leu Pro Glu Val Cys Ser Glu Gln Asp
325 330 335

Glu Leu Asp Phe Leu Met Glu Ala Leu Ile Ile Ser Lys Phe Asn His
340 345 350

Gln Asn Ile Val Arg Cys Ile Gly Val Ser Leu Gln Ser Leu Pro Arg
355 360 365

Phe Ile Leu Leu Glu Leu Met Ala Gly Gly Asp Leu Lys Ser Phe Leu
370 375 380

Arg Glu Thr Arg Pro Arg Pro Ser Gln Pro Ser Ser Leu Ala Met Leu
385 390 395 400

Asp Leu Leu His Val Ala Arg Asp Ile Ala Cys Gly Cys Gln Tyr Leu
405 410 415

Glu Glu Asn His Phe Ile His Arg Asp Ile Ala Ala Arg Asn Cys Leu
420 425 430

[0048] Leu Thr Cys Pro Gly Pro Gly Arg Val Ala Lys Ile Gly Asp Phe Gly
435 440 445

Met Ala Arg Asp Ile Tyr Arg Ala Ser Tyr Tyr Arg Lys Gly Gly Cys
450 455 460

Ala Met Leu Pro Val Lys Trp Met Pro Pro Glu Ala Phe Met Glu Gly
465 470 475 480

Ile Phe Thr Ser Lys Thr Asp Thr Trp Ser Phe Gly Val Leu Leu Trp
485 490 495

Glu Ile Phe Ser Leu Gly Tyr Met Pro Tyr Pro Ser Lys Ser Asn Gln
500 505 510

Glu Val Leu Glu Phe Val Thr Ser Gly Gly Arg Met Asp Pro Pro Lys
515 520 525

Asn Cys Pro Gly Pro Val Tyr Arg Ile Met Thr Gln Cys Trp Gln His
530 535 540

Gln Pro Glu Asp Arg Pro Asn Phe Ala Ile Ile Leu Glu Arg Ile Glu
545 550 555 560

Tyr Cys Thr Gln Asp Pro Asp Val Ile Asn Thr Ala Leu Pro Ile Glu
565 570 575

Tyr Gly Pro Leu Val Glu Glu Glu Glu Lys Val Pro Val Arg Pro Lys
580 585 590

Asp Pro Glu Gly Val Pro Pro Leu Leu Val Ser Gln Gln Ala Lys Arg
595 600 605

[0049] Glu Glu Glu Arg Ser Pro Ala Ala Pro Pro Pro Leu Pro Thr Thr Ser
610 615 620

Ser Gly Lys Ala Ala Lys Lys Pro Thr Ala Ala Glu Val Ser Val Arg
625 630 635 640

Val Pro Arg Gly Pro Ala Val Glu Gly Gly His Val Asn Met Ala Phe
645 650 655

Ser Gln Ser Asn Pro Pro Ser Glu Leu His Arg Val His Gly Ser Arg
660 665 670

Asn Lys Pro Thr Ser Leu Trp Asn Pro Thr Tyr Gly Ser Trp Phe Thr
675 680 685

Glu Lys Pro Thr Lys Lys Asn Asn Pro Ile Ala Lys Lys Glu Pro His
690 695 700

Glu Arg Gly Asn Leu Gly Leu Glu Gly Ser Cys Thr Val Pro Pro Asn
705 710 715 720

Val Ala Thr Gly Arg Leu Pro Gly Ala Ser Leu Leu Leu Glu Pro Ser
725 730 735

Ser Leu Thr Ala Asn Met Lys Glu Val Pro Leu Phe Arg Leu Arg His
740 745 750

Phe Pro Cys Gly Asn Val Asn Tyr Gly Tyr Gln Gln Gln Gly Leu Pro
755 760 765

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

[0050] Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly Pro
785 790 795

<210> 13
<211> 3409
<212> DNA
<213> 人类

<400> 13
actctgtcgg tccgctgaat gaagtgcccg cccctctaag cccggagccc ggcgctttcc 60

ccgcaagatg gacggtttcg ccggcagtct cgatgatagt atttctgctg caagtacttc 120

tgatgttcaa gatcgcctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat 180

cactgtgcta aaggcggctt tggctgatgt tttgaggcgt cttgcaatct ctgaagatca 240

tgtggcctca gtgaaaaaat cagtctcaag taaaggccaa ccaagccctc gagcagttat 300

tcccatgtcc tgtataacca atggaagtgg tgcaaacaga aaaccaagtc ataccagtgc 360

tgtctcaatt gcaggaaaag aaactctttc atctgctgct aaaagtggta cagaaaaaaa 420

gaaagaaaaa ccacaaggac agagagaaaa aaaagaggaa tctcattcta atgatcaaag 480

tccacaaatt cgagcatcac cttctcccca gccctcttca caacctctcc aaatacacag 540

acaaactcca gaaagcaaga atgctactcc caccaaaagc ataaaacgac catcaccagc 600

tgaaaagtca cataattctt gggaaaattc agatgatagc cgtaataaat tgtcgaaaat 660

accttcaaca cccaaattaa taccaaaagt taccaaaact gcagacaagc ataaagatgt 720

catcatcaac caagaaggag aatatattaa aatgtttatg cgcggtcggc caattaccat 780

gttcattcct tccgatgttg acaactatga tgacatcaga acggaactgc ctcctgagaa 840

gctcaaactg gagtgggcat atggttatcg aggaaaggac tgtagagcta atgtttacct 900

tcttccgacc ggggaaatag tttatttcat tgcatcagta gtagtactat ttaattatga 960

ggagagaact cagcgacact acctgggcca tacagactgt gtgaaatgcc ttgctataca 1020

[0051]

tcctgacaaa attaggattg caactggaca gatagctggc gtggataaag atggaaggcc 1080

tctacaaccc cacgtcagag tgtgggattc tgttactcta tccacactgc agattattgg 1140

acttggcact tttgagcgtg gagtaggatg cctggatttt tcaaaagcag attcaggtgt 1200

tcatttatgt gttattgatg actccaatga gcatatgctt actgtatggg actggcagag 1260

gaaagcaaaa ggagcagaaa taaagacaac aaatgaagtt gttttggctg tggagtttca 1320

cccaacagat gcaaatacca taattacatg cggtaaatct catattttct tctggacctg 1380

gagcggcaat tcactaacaa gaaaacaggg aatttttggg aaatatgaaa agccaaaatt 1440

tgtgcagtgt ttagcattct tggggaatgg agatgttctt actggagact caggtggagt 1500

catgcttata tggagcaaaa ctactgtaga gcccacacct gggaaaggac ctaaaggtgt 1560

atatcaaatc agcaaacaaa tcaaagctca tgatggcagt gtgttcacac tttgtcagat 1620

gagaaatggg atgttattaa ctggaggagg gaaagacaga aaaataattc tgtgggatca 1680

tgatctgaat cctgaaagag aaatagagat atgctggatg agccctgagt acaagctgag 1740

caagctccgc acctcgacca tcatgaccga ctacaacccc aactactgct ttgctggcaa 1800

gacctcctcc atcagtgacc tgaaggaggt gccgcggaaa aacatcaccc tcattcgggg 1860

tctgggccat ggagcctttg gggaggtgta tgaaggccag gtgtccggaa tgcccaacga 1920

cccaagcccc ctgcaagtgg ctgtgaagac gctgcctgaa gtgtgctctg aacaggacga 1980

actggatttc ctcatggaag ccctgatcat cagcaaattc aaccaccaga acattgttcg 2040

ctgcattggg gtgagcctgc aatccctgcc ccggttcatc ctgctggagc tcatggcggg 2100

gggagacctc aagtccttcc tccgagagac ccgccctcgc ccgagccagc cctcctccct 2160

ggccatgctg gaccttctgc acgtggctcg ggacattgcc tgtggctgtc agtatttgga 2220

ggaaaaccac ttcatccacc gagacattgc tgccagaaac tgcctcttga cctgtccagg 2280

ccctggaaga gtggccaaga ttggagactt cgggatggcc cgagacatct acagggcgag 2340

[0052]

ctactataga aagggaggct gtgccatgct gccagttaag tggatgcccc cagaggcctt 2400

catggaagga atattcactt ctaaaacaga cacatggtcc tttggagtgc tgctatggga 2460

aatcttttct cttggatata tgccataccc cagcaaaagc aaccaagaag ttctggagtt 2520

tgtcaccagt ggaggccgga tggacccacc caagaactgc cctgggcctg tataccggat 2580

aatgactcag tgctggcaac atcagcctga agacaggccc aactttgcca tcattttgga 2640

gaggattgaa tactgcaccc aggacccgga tgtaatcaac accgctttgc cgatagaata 2700

tggtccactt gtggaagagg aagagaaagt gcctgtgagg cccaaggacc ctgagggggt 2760

tcctcctctc ctggtctctc aacaggcaaa acgggaggag gagcgcagcc cagctgcccc 2820

accacctctg cctaccacct cctctggcaa ggctgcaaag aaacccacag ctgcagaggt 2880

ctctgttcga gtccctagag ggccggccgt ggaaggggga cacgtgaata tggcattctc 2940

tcagtccaac cctccttcgg agttgcacag ggtccacgga tccagaaaca agcccaccag 3000

cttgtggaac ccaacgtacg gctcctggtt tacagagaaa cccaccaaaa agaataatcc 3060

tatagcaaag aaggagccac acgagagggg taacctgggg ctggagggaa gctgtactgt 3120

cccacctaac gttgcaactg ggagacttcc gggggcctca ctgctcctag agccctcttc 3180

gctgactgcc aatatgaagg aggtacctct gttcaggcta cgtcacttcc cttgtgggaa 3240

tgtcaattac ggctaccagc aacagggctt gcccttagaa gccgctactg ccctggagc 3300

tggtcattac gaggatacca ttctgaaaag caagaatagc atgaaccagc ctgggccctg 3360

agctcggtcg cacactcact tctcttcctt gggatcccta agaccgtgg 3409

<210> 14
<211> 1097
<212> PRT
<213> 人类

<400> 14

[0053] Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser
1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln
20 25 30

Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val
35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys
50 55 60

Ser Val Ser Ser Lys Gly Gln Pro Ser Pro Arg Ala Val Ile Pro Met
65 70 75 80

Ser Cys Ile Thr Asn Gly Ser Gly Ala Asn Arg Lys Pro Ser His Thr
85 90 95

Ser Ala Val Ser Ile Ala Gly Lys Glu Thr Leu Ser Ser Ala Ala Lys
100 105 110

Ser Gly Thr Glu Lys Lys Lys Glu Lys Pro Gln Gly Gln Arg Glu Lys
115 120 125

Lys Glu Glu Ser His Ser Asn Asp Gln Ser Pro Gln Ile Arg Ala Ser
130 135 140

Pro Ser Pro Gln Pro Ser Ser Gln Pro Leu Gln Ile His Arg Gln Thr
145 150 155 160

Pro Glu Ser Lys Asn Ala Thr Pro Thr Lys Ser Ile Lys Arg Pro Ser
165 170 175

[0054] Pro Ala Glu Lys Ser His Asn Ser Trp Glu Asn Ser Asp Asp Ser Arg
180 185 190

Asn Lys Leu Ser Lys Ile Pro Ser Thr Pro Lys Leu Ile Pro Lys Val
195 200 205

Thr Lys Thr Ala Asp Lys His Lys Asp Val Ile Ile Asn Gln Glu Gly
210 215 220

Glu Tyr Ile Lys Met Phe Met Arg Gly Arg Pro Ile Thr Met Phe Ile
225 230 235 240

Pro Ser Asp Val Asp Asn Tyr Asp Asp Ile Arg Thr Glu Leu Pro Pro
245 250 255

Glu Lys Leu Lys Leu Glu Trp Ala Tyr Gly Tyr Arg Gly Lys Asp Cys
260 265 270

Arg Ala Asn Val Tyr Leu Leu Pro Thr Gly Glu Ile Val Tyr Phe Ile
275 280 285

Ala Ser Val Val Val Leu Phe Asn Tyr Glu Glu Arg Thr Gln Arg His
290 295 300

Tyr Leu Gly His Thr Asp Cys Val Lys Cys Leu Ala Ile His Pro Asp
305 310 315 320

Lys Ile Arg Ile Ala Thr Gly Gln Ile Ala Gly Val Asp Lys Asp Gly
325 330 335

Arg Pro Leu Gln Pro His Val Arg Val Trp Asp Ser Val Thr Leu Ser
340 345 350

[0055] Thr Leu Gln Ile Ile Gly Leu Gly Thr Phe Glu Arg Gly Val Gly Cys
355 360 365

Leu Asp Phe Ser Lys Ala Asp Ser Gly Val His Leu Cys Val Ile Asp
370 375 380

Asp Ser Asn Glu His Met Leu Thr Val Trp Asp Trp Gln Arg Lys Ala
385 390 395 400

Lys Gly Ala Glu Ile Lys Thr Thr Asn Glu Val Val Leu Ala Val Glu
405 410 415

Phe His Pro Thr Asp Ala Asn Thr Ile Ile Thr Cys Gly Lys Ser His
420 425 430

Ile Phe Phe Trp Thr Trp Ser Gly Asn Ser Leu Thr Arg Lys Gln Gly
435 440 445

Ile Phe Gly Lys Tyr Glu Lys Pro Lys Phe Val Gln Cys Leu Ala Phe
450 455 460

Leu Gly Asn Gly Asp Val Leu Thr Gly Asp Ser Gly Gly Val Met Leu
465 470 475 480

Ile Trp Ser Lys Thr Thr Val Glu Pro Thr Pro Gly Lys Gly Pro Lys
485 490 495

Gly Val Tyr Gln Ile Ser Lys Gln Ile Lys Ala His Asp Gly Ser Val
500 505 510

Phe Thr Leu Cys Gln Met Arg Asn Gly Met Leu Leu Thr Gly Gly Gly
515 520 525

[0056] Lys Asp Arg Lys Ile Ile Leu Trp Asp His Asp Leu Asn Pro Glu Arg
530 535 540

Glu Ile Glu Ile Cys Trp Met Ser Pro Glu Tyr Lys Leu Ser Lys Leu
545 550 555 560

Arg Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro Asn Tyr Cys Phe Ala
565 570 575

Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu Val Pro Arg Lys Asn
580 585 590

Ile Thr Leu Ile Arg Gly Leu Gly His Gly Ala Phe Gly Glu Val Tyr
595 600 605

Glu Gly Gln Val Ser Gly Met Pro Asn Asp Pro Ser Pro Leu Gln Val
610 615 620

Ala Val Lys Thr Leu Pro Glu Val Cys Ser Glu Gln Asp Glu Leu Asp
625 630 635 640

Phe Leu Met Glu Ala Leu Ile Ile Ser Lys Phe Asn His Gln Asn Ile
645 650 655

Val Arg Cys Ile Gly Val Ser Leu Gln Ser Leu Pro Arg Phe Ile Leu
660 665 670

Leu Glu Leu Met Ala Gly Gly Asp Leu Lys Ser Phe Leu Arg Glu Thr
675 680 685

Arg Pro Arg Pro Ser Gln Pro Ser Ser Leu Ala Met Leu Asp Leu Leu
690 695 700

[0057] His Val Ala Arg Asp Ile Ala Cys Gly Cys Gln Tyr Leu Glu Glu Asn
705 710 715 720

His Phe Ile His Arg Asp Ile Ala Ala Arg Asn Cys Leu Leu Thr Cys
725 730 735

Pro Gly Pro Gly Arg Val Ala Lys Ile Gly Asp Phe Gly Met Ala Arg
740 745 750

Asp Ile Tyr Arg Ala Ser Tyr Tyr Arg Lys Gly Gly Cys Ala Met Leu
755 760 765

Pro Val Lys Trp Met Pro Pro Glu Ala Phe Met Glu Gly Ile Phe Thr
770 775 780

Ser Lys Thr Asp Thr Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Phe
785 790 795 800

Ser Leu Gly Tyr Met Pro Tyr Pro Ser Lys Ser Asn Gln Glu Val Leu
805 810 815

Glu Phe Val Thr Ser Gly Gly Arg Met Asp Pro Pro Lys Asn Cys Pro
820 825 830

Gly Pro Val Tyr Arg Ile Met Thr Gln Cys Trp Gln His Gln Pro Glu
835 840 845

Asp Arg Pro Asn Phe Ala Ile Ile Leu Glu Arg Ile Glu Tyr Cys Thr
850 855 860

Gln Asp Pro Asp Val Ile Asn Thr Ala Leu Pro Ile Glu Tyr Gly Pro
865 870 875 880

[0058] Leu Val Glu Glu Glu Glu Lys Val Pro Val Arg Pro Lys Asp Pro Glu
885 890 895

Gly Val Pro Pro Leu Leu Val Ser Gln Gln Ala Lys Arg Glu Glu Glu
900 905 910

Arg Ser Pro Ala Ala Pro Pro Pro Leu Pro Thr Thr Ser Ser Gly Lys
915 920 925

Ala Ala Lys Lys Pro Thr Ala Ala Glu Val Ser Val Arg Val Pro Arg
930 935 940

Gly Pro Ala Val Glu Gly Gly His Val Asn Met Ala Phe Ser Gln Ser
945 950 955 960

Asn Pro Pro Ser Glu Leu His Arg Val His Gly Ser Arg Asn Lys Pro
965 970 975

Thr Ser Leu Trp Asn Pro Thr Tyr Gly Ser Trp Phe Thr Glu Lys Pro
980 985 990

Thr Lys Lys Asn Asn Pro Ile Ala Lys Lys Glu Pro His Glu Arg Gly
995 1000 1005

Asn Leu Gly Leu Glu Gly Ser Cys Thr Val Pro Pro Asn Val Ala
1010 1015 1020

Thr Gly Arg Leu Pro Gly Ala Ser Leu Leu Leu Glu Pro Ser Ser
1025 1030 1035

Leu Thr Ala Asn Met Lys Glu Val Pro Leu Phe Arg Leu Arg His
1040 1045 1050

Phe Pro Cys Gly Asn Val Asn Tyr Gly Tyr Gln Gln Gln Gly Leu
[0059] 1055 1060 1065

Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala Gly His Tyr Glu Asp
1070 1075 1080

Thr Ile Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly Pro
1085 1090 1095

<210> 15
<211> 2014
<212> DNA
<213> 人类

<400> 15
actctgtcgg tccgctgaat gaagtgcccg cccctctaag cccggagccc ggcgctttcc 60

ccgcaagatg gacggtttcg ccggcagtct cgatgatagt atttctgctg caagtacttc 120

tgatgttcaa gatcgcctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat 180

cactgtgcta aaggcggctt tggctgatgt tttgaggcgt cttgcaatct ctgaagatca 240

tgtggcctca gtgaaaaaat cagtctcaag taaagtgtac cgccggaagc accaggagct 300

gcaagccatg cagatggagc tgcagagccc tgagtacaag ctgagcaagc tccgcacctc 360

gaccatcatg accgactaca accccaacta ctgctttgct ggcaagacct cctccatcag 420

tgacctgaag gaggtgccgc ggaaaaacat caccctcatt cggggtctgg gccatggagc 480

ctttggggag gtgtatgaag gccaggtgtc cggaatgccc aacgacccaa gcccctgca 540

agtggctgtg aagacgctgc ctgaagtgtg ctctgaacag gacgaactgg atttcctcat 600

ggaagccctg atcatcagca aattcaacca ccagaacatt gttcgctgca ttggggtgag 660

cctgcaatcc ctgccccggt tcatcctgct ggagctcatg gcgggggggag acctcaagtc 720

cttcctccga gagacccgcc ctcgcccgag ccagccctcc tccctggcca tgctggacct 780

tctgcacgtg gctcgggaca ttgcctgtgg ctgtcagtat ttggaggaaa accacttcat 840

[0060]

ccaccgagac attgctgcca gaaactgcct cttgacctgt ccaggccctg gaagagtggc 900

caagattgga gacttcggga tggcccgaga catctacagg gcgagctact atagaaaggg 960

aggctgtgcc atgctgccag ttaagtggat gcccccagag gccttcatgg aaggaatatt 1020

cacttctaaa acagacacat ggtcctttgg agtgctgcta tgggaaatct tttctcttgg 1080

atatatgcca taccccagca aaagcaacca ggaagttctg gagtttgtca ccagtggagg 1140

ccggatggac ccacccaaga actgccctgg gcctgtatac cggataatga ctcagtgctg 1200

gcaacatcag cctgaagaca ggcccaactt tgccatcatt ttggagagga ttgaatactg 1260

cacccaggac ccggatgtaa tcaacaccgc tttgccgata gaatatggtc cacttgtgga 1320

agaggaagag aaagtgcctg tgaggcccaa ggaccctgag ggggttcctc ctctcctggt 1380

ctctcaacag gcaaaacggg aggaggagcg cagcccagct gccccaccac ctctgcctac 1440

cacctcctct ggcaaggctg caaagaaacc cacagctgca gaggtctctg ttcgagtccc 1500

tagagggccg gccgtggaag ggggacacgt gaatatggca ttctctcagt ccaaccctcc 1560

ttcggagttg cacagggtcc acggatccag aaacaagccc accagcttgt ggaacccaac 1620

gtacggctcc tggtttacag agaaacccac caaaaagaat aatcctatag caaagaagga 1680

gccacacgag aggggtaacc tggggctgga gggaagctgt actgtcccac ctaacgttgc 1740

aactgggaga cttccggggg cctcactgct cctagagccc tcttcgctga ctgccaatat 1800

gaaggaggta cctctgttca ggctacgtca cttcccttgt gggaatgtca attacggcta 1860

ccagcaacag ggcttgccct tagaagccgc tactgcccct ggagctggtc attacgagga 1920

taccattctg aaaagcaaga atagcatgaa ccagcctggg ccctgagctc ggtcgcacac 1980

tcacttctct tccttgggat ccctaagacc gtgg 2014

<210> 16
<211> 632
[0061] <212> PRT
<213> 人类

<400> 16
Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser
1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln
20 25 30

Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val
35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys
50 55 60

Ser Val Ser Ser Lys Val Tyr Arg Arg Lys His Gln Glu Leu Gln Ala
65 70 75 80

Met Gln Met Glu Leu Gln Ser Pro Glu Tyr Lys Leu Ser Lys Leu Arg
85 90 95

Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro Asn Tyr Cys Phe Ala Gly
100 105 110

Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu Val Pro Arg Lys Asn Ile
115 120 125

Thr Leu Ile Arg Gly Leu Gly His Gly Ala Phe Gly Glu Val Tyr Glu
130 135 140

Gly Gln Val Ser Gly Met Pro Asn Asp Pro Ser Pro Leu Gln Val Ala
145 150 155 160

[0062] Val Lys Thr Leu Pro Glu Val Cys Ser Glu Gln Asp Glu Leu Asp Phe
165 170 175

Leu Met Glu Ala Leu Ile Ile Ser Lys Phe Asn His Gln Asn Ile Val
180 185 190

Arg Cys Ile Gly Val Ser Leu Gln Ser Leu Pro Arg Phe Ile Leu Leu
195 200 205

Glu Leu Met Ala Gly Gly Asp Leu Lys Ser Phe Leu Arg Glu Thr Arg
210 215 220

Pro Arg Pro Ser Gln Pro Ser Ser Leu Ala Met Leu Asp Leu Leu His
225 230 235 240

Val Ala Arg Asp Ile Ala Cys Gly Cys Gln Tyr Leu Glu Glu Asn His
245 250 255

Phe Ile His Arg Asp Ile Ala Ala Arg Asn Cys Leu Leu Thr Cys Pro
260 265 270

Gly Pro Gly Arg Val Ala Lys Ile Gly Asp Phe Gly Met Ala Arg Asp
275 280 285

Ile Tyr Arg Ala Ser Tyr Tyr Arg Lys Gly Gly Cys Ala Met Leu Pro
290 295 300

Val Lys Trp Met Pro Pro Glu Ala Phe Met Glu Gly Ile Phe Thr Ser
305 310 315 320

Lys Thr Asp Thr Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Phe Ser
325 330 335

[0063] Leu Gly Tyr Met Pro Tyr Pro Ser Lys Ser Asn Gln Glu Val Leu Glu
340 345 350

Phe Val Thr Ser Gly Gly Arg Met Asp Pro Pro Lys Asn Cys Pro Gly
355 360 365

Pro Val Tyr Arg Ile Met Thr Gln Cys Trp Gln His Gln Pro Glu Asp
370 375 380

Arg Pro Asn Phe Ala Ile Ile Leu Glu Arg Ile Glu Tyr Cys Thr Gln
385 390 395 400

Asp Pro Asp Val Ile Asn Thr Ala Leu Pro Ile Glu Tyr Gly Pro Leu
405 410 415

Val Glu Glu Glu Glu Lys Val Pro Val Arg Pro Lys Asp Pro Glu Gly
420 425 430

Val Pro Pro Leu Leu Val Ser Gln Gln Ala Lys Arg Glu Glu Glu Arg
435 440 445

Ser Pro Ala Ala Pro Pro Pro Leu Pro Thr Thr Ser Ser Gly Lys Ala
450 455 460

Ala Lys Lys Pro Thr Ala Ala Glu Val Ser Val Arg Val Pro Arg Gly
465 470 475 480

Pro Ala Val Glu Gly Gly His Val Asn Met Ala Phe Ser Gln Ser Asn
485 490 495

Pro Pro Ser Glu Leu His Arg Val His Gly Ser Arg Asn Lys Pro Thr
500 505 510

[0064] Ser Leu Trp Asn Pro Thr Tyr Gly Ser Trp Phe Thr Glu Lys Pro Thr
515 520 525

Lys Lys Asn Asn Pro Ile Ala Lys Lys Glu Pro His Glu Arg Gly Asn
530 535 540

Leu Gly Leu Glu Gly Ser Cys Thr Val Pro Pro Asn Val Ala Thr Gly
545 550 555 560

Arg Leu Pro Gly Ala Ser Leu Leu Leu Glu Pro Ser Ser Leu Thr Ala
565 570 575

Asn Met Lys Glu Val Pro Leu Phe Arg Leu Arg His Phe Pro Cys Gly
580 585 590

Asn Val Asn Tyr Gly Tyr Gln Gln Gln Gly Leu Pro Leu Glu Ala Ala
595 600 605

Thr Ala Pro Gly Ala Gly His Tyr Glu Asp Thr Ile Leu Lys Ser Lys
610 615 620

Asn Ser Met Asn Gln Pro Gly Pro
625 630

<210> 17
<211> 2131
<212> DNA
<213>人类

<400> 17
actctgtcgg tccgctgaat gaagtgcccg cccctctaag cccggagccc ggcgctttcc 60

ccgcaagatg gacggtttcg ccggcagtct cgatgatagt atttctgctg caagtacttc 120

tgatgttcaa gatcgcctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat 180

[0065] cactgtgcta aaggcggctt tggctgatgt tttgaggcgt cttgcaatct ctgaagatca 240

tgtggcctca gtgaaaaaat cagtctcaag taaaggttca gagctcaggg gaggatatgg 300

agatccaggg aggcttcctg taggaagtgg cctgtgtagt gcttcaaggg ccaggctgcc 360

aggccatgtt gcagctgacc acccacctgc agtgtaccgc cggaagcacc aggagctgca 420

agccatgcag atggagctgc agagccctga gtacaagctg agcaagctcc gcacctcgac 480

catcatgacc gactacaacc ccaactactg ctttgctggc aagacctcct ccatcagtga 540

cctgaaggag gtgccgcgga aaaacatcac cctcattcgg ggtctgggcc atggagcctt 600

tggggaggtg tatgaaggcc aggtgtccgg aatgcccaac gacccaagcc cctgcaagt 660

ggctgtgaag acgctgcctg aagtgtgctc tgaacaggac gaactggatt tcctcatgga 720

agccctgatc atcagcaaat tcaaccacca gaacattgtt cgctgcattg gggtgagcct 780

gcaatccctg ccccggttca tcctgctgga gctcatggcg gggggagacc tcaagtcctt 840

cctccgagag acccgcccte gcccgagcca gccctcctcc ctggccatgc tggaccttct 900

gcacgtggct cgggacattg cctgtggctg tcagtatttg gaggaaaacc acttcatcca 960

ccgagacatt gctgccagaa actgcctctt gacctgtcca ggccctggaa gagtggccaa 1020

gattggagac ttcgggatgg cccgagacat ctacagggcg agctactata gaaagggagg 1080

ctgtgccatg ctgccagtta agtggatgcc cccagaggcc ttcatggaag gaatattcac 1140

ttctaaaaca gacacatggt cctttggagt gctgctatgg gaaatctttt ctcttggata 1200

tatgccatac cccagcaaaa gcaaccagga agttctggag tttgtcacca gtggaggccg 1260

gatggaccca cccaagaact gccctgggcc tgtataccgg ataatgactc agtgctggca 1320

acatcagcct gaagacaggc ccaactttgc catcattttg gagaggattg aatactgcac 1380

ccaggacccg gatgtaatca acaccgcttt gccgatagaa tatggtccac ttgtggaaga 1440

ggaagagaaa gtgcctgtga ggcccaagga ccctgagggg gttcctcctc tcctggtctc 1500

[0066]

tcaacaggca aaacgggagg aggagcgcag cccagctgcc ccaccacctc tgcctaccac 1560

ctcctctggc aaggctgcaa agaaacccac agctgcagag gtctctgttc gagtccctag 1620

agggccggcc gtggaagggg gacacgtgaa tatggcattc tctcagtcca accctccttc 1680

ggagttgcac agggtccacg gatccagaaa caagcccacc agcttgtgga acccaacgta 1740

cggctcctgg tttacagaga aacccaccaa aaagaataat cctatagcaa agaaggagcc 1800

acacgagagg ggtaacctgg ggctggaggg aagctgtact gtcccaccta acgttgcaac 1860

tgggagactt ccgggggcct cactgctcct agagccctct tcgctgactg ccaatatgaa 1920

ggaggtacct ctgttcaggc tacgtcactt cccttgtggg aatgtcaatt acggctacca 1980

gcaacagggc ttgcccttag aagccgctac tgcccctgga gctggtcatt acgaggatac 2040

cattctgaaa agcaagaata gcatgaacca gcctgggccc tgagctcggt cgcacactca 2100

cttctcttcc ttgggatccc taagaccgtg g 2131

<210> 18
<211> 671
<212> PRT
<213> 人类

<400> 18
Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser
1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln
20 25 30

Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val
35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys
50 55 60

[0067]

Ser Val Ser Ser Lys Gly Ser Glu Leu Arg Gly Gly Tyr Gly Asp Pro
65 70 75 80

Gly Arg Leu Pro Val Gly Ser Gly Leu Cys Ser Ala Ser Arg Ala Arg
85 90 95

Leu Pro Gly His Val Ala Ala Asp His Pro Pro Ala Val Tyr Arg Arg
100 105 110

Lys His Gln Glu Leu Gln Ala Met Gln Met Glu Leu Gln Ser Pro Glu
115 120 125

Tyr Lys Leu Ser Lys Leu Arg Thr Ser Thr Ile Met Thr Asp Tyr Asn
130 135 140

Pro Asn Tyr Cys Phe Ala Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys

145 150 155 160

Glu Val Pro Arg Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly His Gly
165 170 175

Ala Phe Gly Glu Val Tyr Glu Gly Gln Val Ser Gly Met Pro Asn Asp
180 185 190

Pro Ser Pro Leu Gln Val Ala Val Lys Thr Leu Pro Glu Val Cys Ser
195 200 205

Glu Gln Asp Glu Leu Asp Phe Leu Met Glu Ala Leu Ile Ile Ser Lys
210 215 220

Phe Asn His Gln Asn Ile Val Arg Cys Ile Gly Val Ser Leu Gln Ser
225 230 235 240

[0068]

Leu Pro Arg Phe Ile Leu Leu Glu Leu Met Ala Gly Gly Asp Leu Lys
245 250 255

Ser Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln Pro Ser Ser Leu
260 265 270

Ala Met Leu Asp Leu Leu His Val Ala Arg Asp Ile Ala Cys Gly Cys
275 280 285

Gln Tyr Leu Glu Glu Asn His Phe Ile His Arg Asp Ile Ala Ala Arg
290 295 300

Asn Cys Leu Leu Thr Cys Pro Gly Pro Gly Arg Val Ala Lys Ile Gly
305 310 315 320

Asp Phe Gly Met Ala Arg Asp Ile Tyr Arg Ala Ser Tyr Tyr Arg Lys

325 330 335

Gly Gly Cys Ala Met Leu Pro Val Lys Trp Met Pro Pro Glu Ala Phe
340 345 350

Met Glu Gly Ile Phe Thr Ser Lys Thr Asp Thr Trp Ser Phe Gly Val
355 360 365

Leu Leu Trp Glu Ile Phe Ser Leu Gly Tyr Met Pro Tyr Pro Ser Lys
370 375 380

Ser Asn Gln Glu Val Leu Glu Phe Val Thr Ser Gly Gly Arg Met Asp
385 390 395 400

Pro Pro Lys Asn Cys Pro Gly Pro Val Tyr Arg Ile Met Thr Gln Cys
405 410 415

[0069]

Trp Gln His Gln Pro Glu Asp Arg Pro Asn Phe Ala Ile Ile Leu Glu
420 425 430

Arg Ile Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile Asn Thr Ala Leu
435 440 445

Pro Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu Glu Lys Val Pro Val
450 455 460

Arg Pro Lys Asp Pro Glu Gly Val Pro Pro Leu Leu Val Ser Gln Gln
465 470 475 480

Ala Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro Pro Pro Leu Pro
485 490 495

Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr Ala Ala Glu Val

500 505 510

Ser Val Arg Val Pro Arg Gly Pro Ala Val Glu Gly Gly His Val Asn
515 520 525

Met Ala Phe Ser Gln Ser Asn Pro Pro Ser Glu Leu His Arg Val His
530 535 540

Gly Ser Arg Asn Lys Pro Thr Ser Leu Trp Asn Pro Thr Tyr Gly Ser
545 550 555 560

Trp Phe Thr Glu Lys Pro Thr Lys Lys Asn Asn Pro Ile Ala Lys Lys
565 570 575

Glu Pro His Glu Arg Gly Asn Leu Gly Leu Glu Gly Ser Cys Thr Val
580 585 590

[0070]

Pro Pro Asn Val Ala Thr Gly Arg Leu Pro Gly Ala Ser Leu Leu Leu
595 600 605

Glu Pro Ser Ser Leu Thr Ala Asn Met Lys Glu Val Pro Leu Phe Arg
610 615 620

Leu Arg His Phe Pro Cys Gly Asn Val Asn Tyr Gly Tyr Gln Gln Gln
625 630 635 640

Gly Leu Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala Gly His Tyr Glu
645 650 655

Asp Thr Ile Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly Pro
660 665 670

<210> 19

<211> 3365
<212> DNA
<213> 人类

<400> 19

[0071]

tactctgtcg gtccgctgaa tgaagtgccc gcccctctaa gcccggagcc cggcgctttc 60

cccgcaagat ggacggtttc gccggcagtc tcgatgatag tatttctgct gcaagtactt 120

ctgatgttca agatcgcctg tcagctcttg agtcacgagt tcagcaacaa gaagatgaaa 180

tcactgtgct aaaggcggct ttggctgatg ttttgaggcg tcttgcaatc tctgaagatc 240

atgtggcctc agtgaaaaaa tcagtctcaa gtaaaggcca accaagcccc cgagcagtta 300

ttcccatgtc ctgtataacc aatggaagtg gtgcaaacag aaaaccaagt cataccagtg 360

ctgtctcaat tgcaggaaaa gaaactcttt catctgctgc taaaagtggt acagaaaaaa 420

agaaagaaaa accacaagga cagagagaaa aaaaagagga atctcattct aatgatcaaa 480

gtccacaaat tcgagcatca ccttctcccc agccctcttc acaacctctc caaatacaca 540

gacaaactcc agaaagcaag aatgctactc ccaccaaaag cataaaacga ccatcaccag 600

ctgaaaagtc acataattct tgggaaaatt cagatgatag ccgtaataaa ttgtcgaaaa 660

taccttcaac acccaaatta ataccaaaag ttaccaaaac tgcagacaag cataaagatg 720

tcatcatcaa ccaagaagga gaatatatta aaatgtttat gcgcggtcgg ccaattacca 780

tgttcattcc ttccgatgtt gacaactatg atgacatcag aacggaactg cctcctgaga 840

agctcaaact ggagtgggca tatggttatc gaggaaagga ctgtagagct aatgtttacc 900

ttcttccgac cggggaaata gtttatttca ttgcatcagt agtagtacta tttaattatg 960

aggagagaac tcagcgacac tacctgggcc atacagactg tgtgaaatgc cttgctatac 1020

atcctgacaa aattaggatt gcaactggac agatagctgg cgtggataaa gatggaaggc 1080

ctctacaacc ccacgtcaga gtgtgggatt ctgttactct atccacactg cagattattg 1140

gacttggcac tttgagcgt ggagtaggat gcctggattt ttcaaaagca gattcaggtg 1200

ttcatttatg tgttattgat gactccaatg agcatatgct tactgtatgg gactggcaga 1260

ggaaagcaaa aggagcagaa ataaagacaa caaatgaagt tgttttggct gtggagtttc 1320

acccaacaga tgcaaatacc ataattacat gcggtaaatc tcatattttc ttctggacct 1380

ggagcggcaa ttcactaaca agaaaacagg gaatttttgg gaaatatgaa aagccaaaat 1440

ttgtgcagtg tttagcattc ttggggaatg gagatgttct tactggagac tcaggtggag 1500

tcatgcttat atggagcaaa actactgtag agcccacacc tgggaaagga cctaaaggaa 1560

gtggcctgtg tagtgcttca agggccaggc tgccaggcca tgttgcagct gaccacccac 1620

ctgcagtgta ccgccggaag caccaggagc tgcaagccat gcagatggag ctgcagagcc 1680

ctgagtacaa gctgagcaag ctccgcacct cgaccatcat gaccgactac aaccccaact 1740

actgctttgc tggcaagacc tcctccatca gtgacctgaa ggaggtgccg cggaaaaaca 1800

tcaccctcat tcggggtctg ggccatggag cctttgggga ggtgtatgaa ggccaggtgt 1860

[0072]

ccggaatgcc caacgaccca agcccccctgc aagtggctgt gaagacgctg cctgaagtgt 1920

gctctgaaca ggacgaactg gatttcctca tggaagccct gatcatcagc aaattcaacc 1980

accagaacat tgttcgctgc attggggtga gcctgcaatc cctgccccgg ttcatcctgc 2040

tggagctcat ggcgggggga gacctcaagt ccttcctccg agagacccgc cctcgcccga 2100

gccagccctc ctccctggcc atgctggacc ttctgcacgt ggctcgggac attgcctgtg 2160

gctgtcagta tttggaggaa aaccacttca tccaccgaga cattgctgcc agaaactgcc 2220

tcttgacctg tccaggccct ggaagagtgg ccaagattgg agacttcggg atggcccgag 2280

acatctacag ggcgagctac tatagaaagg gaggctgtgc catgctgcca gttaagtgga 2340

tgccccccaga ggccttcatg gaaggaatat tcacttctaa aacagacaca tggtcctttg 2400

gagtgctgct atgggaaatc ttttctcttg gatatatgcc ataccccagc aaaagcaacc 2460

aggaagttct ggagtttgtc accagtggag gccggatgga cccacccaag aactgccctg 2520

ggcctgtata ccggataatg actcagtgct ggcaacatca gcctgaagac aggcccaact 2580

ttgccatcat tttggagagg attgaatact gcacccagga cccggatgta atcaacaccg 2640

ctttgccgat agaatatggt ccacttgtgg aagaggaaga gaaagtgcct gtgaggccca 2700

aggaccctga gggggttcct cctctcctgg tctctcaaca ggcaaaacgg gaggaggagc 2760

gcagcccagc tgccccacca cctctgccta ccacctcctc tggcaaggct gcaaagaaac 2820

ccacagctgc agaggtctct gttcgagtcc ctagagggcc ggccgtggaa gggggacacg 2880

tgaatatggc attctctcag tccaaccctc cttcggagtt gcacagggtc cacggatcca 2940

gaaacaagcc caccagcttg tggaacccaa cgtacggctc ctggtttaca gagaaaccca 3000

ccaaaaagaa taatcctata gcaaagaagg agccacacga gaggggtaac ctggggctgg 3060

agggaagctg tactgtccca cctaacgttg caactgggag acttccgggg gcctcactgc 3120

tcctagagcc ctcttcgctg actgccaata tgaaggaggt acctctgttc aggctacgtc 3180

[0073]

acttcccttg tgggaatgtc aattacggct accagcaaca gggcttgccc ttagaagccg 3240

ctactgcccc tggagctggt cattacgagg ataccattct gaaaagcaag aatagcatga 3300

accagcctgg gccctgagct cggtcgcaca ctcacttctc ttccttggga tccctaagac 3360

cgtgg 3365

<210> 20
<211> 1082
<212> PRT
<213> 人类

<400> 20
Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser
1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln
20 25 30

Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val
35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys
50 55 60

Ser Val Ser Ser Lys Gly Gln Pro Ser Pro Arg Ala Val Ile Pro Met
65 70 75 80

Ser Cys Ile Thr Asn Gly Ser Gly Ala Asn Arg Lys Pro Ser His Thr
85 90 95

Ser Ala Val Ser Ile Ala Gly Lys Glu Thr Leu Ser Ser Ala Ala Lys
100 105 110

[0074] Ser Gly Thr Glu Lys Lys Lys Glu Lys Pro Gln Gly Gln Arg Glu Lys
115 120 125

Lys Glu Glu Ser His Ser Asn Asp Gln Ser Pro Gln Ile Arg Ala Ser
130 135 140

Pro Ser Pro Gln Pro Ser Ser Gln Pro Leu Gln Ile His Arg Gln Thr
145 150 155 160

Pro Glu Ser Lys Asn Ala Thr Pro Thr Lys Ser Ile Lys Arg Pro Ser
165 170 175

Pro Ala Glu Lys Ser His Asn Ser Trp Glu Asn Ser Asp Asp Ser Arg
180 185 190

Asn Lys Leu Ser Lys Ile Pro Ser Thr Pro Lys Leu Ile Pro Lys Val
195 200 205

Thr Lys Thr Ala Asp Lys His Lys Asp Val Ile Ile Asn Gln Glu Gly
210 215 220

Glu Tyr Ile Lys Met Phe Met Arg Gly Arg Pro Ile Thr Met Phe Ile
225 230 235 240

Pro Ser Asp Val Asp Asn Tyr Asp Asp Ile Arg Thr Glu Leu Pro Pro
245 250 255

Glu Lys Leu Lys Leu Glu Trp Ala Tyr Gly Tyr Arg Gly Lys Asp Cys
260 265 270

Arg Ala Asn Val Tyr Leu Leu Pro Thr Gly Glu Ile Val Tyr Phe Ile
275 280 285

[0075] Ala Ser Val Val Val Leu Phe Asn Tyr Glu Glu Arg Thr Gln Arg His
290 295 300

Tyr Leu Gly His Thr Asp Cys Val Lys Cys Leu Ala Ile His Pro Asp
305 310 315 320

Lys Ile Arg Ile Ala Thr Gly Gln Ile Ala Gly Val Asp Lys Asp Gly
325 330 335

Arg Pro Leu Gln Pro His Val Arg Val Trp Asp Ser Val Thr Leu Ser
340 345 350

Thr Leu Gln Ile Ile Gly Leu Gly Thr Phe Glu Arg Gly Val Gly Cys
355 360 365

Leu Asp Phe Ser Lys Ala Asp Ser Gly Val His Leu Cys Val Ile Asp
370 375 380

Asp Ser Asn Glu His Met Leu Thr Val Trp Asp Trp Gln Arg Lys Ala
385 390 395 400

Lys Gly Ala Glu Ile Lys Thr Thr Asn Glu Val Val Leu Ala Val Glu
405 410 415

Phe His Pro Thr Asp Ala Asn Thr Ile Ile Thr Cys Gly Lys Ser His
420 425 430

Ile Phe Phe Trp Thr Trp Ser Gly Asn Ser Leu Thr Arg Lys Gln Gly
435 440 445

Ile Phe Gly Lys Tyr Glu Lys Pro Lys Phe Val Gln Cys Leu Ala Phe
450 455 460

[0076] Leu Gly Asn Gly Asp Val Leu Thr Gly Asp Ser Gly Gly Val Met Leu
465 470 475 480

Ile Trp Ser Lys Thr Thr Val Glu Pro Thr Pro Gly Lys Gly Pro Lys
485 490 495

Gly Ser Gly Leu Cys Ser Ala Ser Arg Ala Arg Leu Pro Gly His Val
500 505 510

Ala Ala Asp His Pro Pro Ala Val Tyr Arg Arg Lys His Gln Glu Leu
515 520 525

Gln Ala Met Gln Met Glu Leu Gln Ser Pro Glu Tyr Lys Leu Ser Lys
530 535 540

Leu Arg Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro Asn Tyr Cys Phe
545 550 555 560

Ala Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu Val Pro Arg Lys
565 570 575

Asn Ile Thr Leu Ile Arg Gly Leu Gly His Gly Ala Phe Gly Glu Val
580 585 590

Tyr Glu Gly Gln Val Ser Gly Met Pro Asn Asp Pro Ser Pro Leu Gln
595 600 605

Val Ala Val Lys Thr Leu Pro Glu Val Cys Ser Glu Gln Asp Glu Leu
610 615 620

Asp Phe Leu Met Glu Ala Leu Ile Ile Ser Lys Phe Asn His Gln Asn
625 630 635 640

[0077] Ile Val Arg Cys Ile Gly Val Ser Leu Gln Ser Leu Pro Arg Phe Ile
645 650 655

Leu Leu Glu Leu Met Ala Gly Gly Asp Leu Lys Ser Phe Leu Arg Glu
660 665 670

Thr Arg Pro Arg Pro Ser Gln Pro Ser Ser Leu Ala Met Leu Asp Leu
675 680 685

Leu His Val Ala Arg Asp Ile Ala Cys Gly Cys Gln Tyr Leu Glu Glu
690 695 700

Asn His Phe Ile His Arg Asp Ile Ala Ala Arg Asn Cys Leu Leu Thr
705 710 715 720

Cys Pro Gly Pro Gly Arg Val Ala Lys Ile Gly Asp Phe Gly Met Ala
725 730 735

Arg Asp Ile Tyr Arg Ala Ser Tyr Tyr Arg Lys Gly Gly Cys Ala Met
740 745 750

Leu Pro Val Lys Trp Met Pro Pro Glu Ala Phe Met Glu Gly Ile Phe
755 760 765

Thr Ser Lys Thr Asp Thr Trp Ser Phe Gly Val Leu Leu Trp Glu Ile
770 775 780

Phe Ser Leu Gly Tyr Met Pro Tyr Pro Ser Lys Ser Asn Gln Glu Val
785 790 795 800

Leu Glu Phe Val Thr Ser Gly Gly Arg Met Asp Pro Pro Lys Asn Cys
805 810 815

[0078] Pro Gly Pro Val Tyr Arg Ile Met Thr Gln Cys Trp Gln His Gln Pro
820 825 830

Glu Asp Arg Pro Asn Phe Ala Ile Ile Leu Glu Arg Ile Glu Tyr Cys
835 840 845

Thr Gln Asp Pro Asp Val Ile Asn Thr Ala Leu Pro Ile Glu Tyr Gly
850 855 860

Pro Leu Val Glu Glu Glu Glu Lys Val Pro Val Arg Pro Lys Asp Pro
865 870 875 880

Glu Gly Val Pro Pro Leu Leu Val Ser Gln Gln Ala Lys Arg Glu Glu
885 890 895

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

Lys Ala Ala Lys Lys Pro Thr Ala Ala Glu Val Ser Val Arg Val Pro
915 920 925

Arg Gly Pro Ala Val Glu Gly Gly His Val Asn Met Ala Phe Ser Gln
930 935 940

Ser Asn Pro Pro Ser Glu Leu His Arg Val His Gly Ser Arg Asn Lys
945 950 955 960

Pro Thr Ser Leu Trp Asn Pro Thr Tyr Gly Ser Trp Phe Thr Glu Lys
965 970 975

Pro Thr Lys Lys Asn Asn Pro Ile Ala Lys Lys Glu Pro His Glu Arg
980 985 990

[0079] Gly Asn Leu Gly Leu Glu Gly Ser Cys Thr Val Pro Pro Asn Val Ala
995 1000 1005

Thr Gly Arg Leu Pro Gly Ala Ser Leu Leu Leu Glu Pro Ser Ser
1010 1015 1020

Leu Thr Ala Asn Met Lys Glu Val Pro Leu Phe Arg Leu Arg His
1025 1030 1035

Phe Pro Cys Gly Asn Val Asn Tyr Gly Tyr Gln Gln Gln Gly Leu
1040 1045 1050

Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala Gly His Tyr Glu Asp
1055 1060 1065

Thr Ile Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly Pro
1070 1075 1080

<210> 21
<211> 3435
<212> DNA
<213>人类

<400> 21
tactctgtcg gtccgctgaa tgaagtgccc gccccctctaa gcccggagcc cggcgctttc 60

cccgcaagat ggacggtttc gccggcagtc tcgatgatag tatttctgct gcaagtactt 120

ctgatgttca agatcgcctg tcagctcttg agtcacgagt tcagcaacaa gaagatgaaa 180

tcactgtgct aaaggcggct ttggctgatg ttttgaggcg tcttgcaatc tctgaagatc 240

atgtggcctc agtgaaaaaa tcagtctcaa gtaaaggcca accaagccct cgagcagtta 300

ttcccatgtc ctgtataacc aatggaagtg gtgcaaacag aaaaccaagt cataccagtg 360

ctgtctcaat tgcaggaaaa gaaactcttt catctgctgc taaaagtggt acagaaaaaa 420

[0080] agaaagaaaa accacaagga cagagagaaa aaaaagagga atctcattct aatgatcaaa 480

gtccacaaat tcgagcatca ccttctcccc agccctcttc acaacctctc caaatacaca 540

gacaaactcc agaaagcaag aatgctactc ccaccaaaag cataaaacga ccatcaccag 600

ctgaaaagtc acataattct tgggaaaatt cagatgatag ccgtaataaa ttgtcgaaaa 660

taccttcaac acccaaatta ataccaaaag ttaccaaaac tgcagacaag cataaagatg 720

tcatcatcaa ccaagaagga gaatatatta aaatgtttat gcgcggtcgg ccaattacca 780

tgttcattcc ttccgatgtt gacaactatg atgacatcag aacggaactg cctcctgaga 840

agctcaaact ggagtgggca tatggttatc gaggaaagga ctgtagagct aatgtttacc 900

ttcttccgac cggggaaata gtttatttca ttgcatcagt agtagtacta tttaattatg 960

aggagagaac tcagcgacac tacctgggcc atacagactg tgtgaaatgc cttgctatac 1020

atcctgacaa aattaggatt gcaactggac agatagctgg cgtggataaa gatggaaggc 1080

ctctacaacc ccacgtcaga gtgtgggatt ctgttactct atccacactg cagattattg 1140

gacttggcac ttttgagcgt ggagtaggat gcctggattt ttcaaaagca gattcaggtg 1200

ttcatttatg tgttattgat gactccaatg agcatatgct tactgtatgg gactggcaga 1260

ggaaagcaaa aggagcagaa ataaagacaa caaatgaagt tgttttggct gtggagtttc 1320

acccaacaga tgcaaatacc ataattacat gcggtaaatc tcatattttc ttctggacct 1380

ggagcggcaa ttcactaaca agaaaacagg gaatttttgg gaaatatgaa aagccaaaat 1440

ttgtgcagtg tttagcattc ttggggaatg gagatgttct tactggagac tcaggtggag 1500

tcatgcttat atggagcaaa actactgtag agcccacacc tgggaaagga cctaaaggtg 1560

tatatcaaat cagcaaacaa atcaaagctc atgatggcag tgtgttcaca ctttgtcaga 1620

tgagaaatgg gatgttatta actggaggag ggaaagacag aaaaataatt ctgtgggatc 1680

atgatctgaa tcctgaaaga gaaatagagc accaggagct gcaagccatg cagatggagc 1740

[0081]

tgcagagccc tgagtacaag ctgagcaagc tccgcacctc gaccatcatg accgactaca 1800

accccaacta ctgctttgct ggcaagacct cctccatcag tgacctgaag gaggtgccgc 1860

ggaaaaacat caccctcatt cggggtctgg gccatggagc ctttggggag gtgtatgaag 1920

gccaggtgtc cggaatgccc aacgacccaa gccccctgca agtggctgtg aagacgctgc 1980

ctgaagtgtg ctctgaacag gacgaactgg atttcctcat ggaagccctg atcatcagca 2040

aattcaacca ccagaacatt gttcgctgca ttggggtgag cctgcaatcc ctgccccggt 2100

tcatcctgct ggagctcatg gcgggggggag acctcaagtc cttcctccga gagacccgcc 2160

ctcgccccgag ccagccctcc tccctggcca tgctggacct tctgcacgtg gctcgggaca 2220

ttgcctgtgg ctgtcagtat ttggaggaaa accacttcat ccaccgagac attgctgcca 2280

gaaactgcct cttgacctgt ccaggccctg gaagagtggc caagattgga gacttcggga 2340

tggcccgaga catctacagg gcgagctact atagaaaggg aggctgtgcc atgctgccag 2400

ttaagtggat gcccccagag gccttcatgg aaggaatatt cacttctaaa acagacacat 2460

ggtcctttgg agtgctgcta tgggaaatct tttctcttgg atatatgcca taccccagca 2520

aaagcaacca ggaagttctg gagtttgtca ccagtggagg ccggatggac ccacccaaga 2580

actgccctgg gcctgtatac cggataatga ctcagtgctg gcaacatcag cctgaagaca 2640

ggcccaactt tgccatcatt ttggagagga ttgaatactg cacccaggac ccggatgtaa 2700

tcaacaccgc tttgccgata gaatatggtc cacttgtgga agaggaagag aaagtgcctg 2760

tgaggcccaa ggaccctgag ggggttcctc ctctcctggt ctctcaacag gcaaaacggg 2820

aggaggagcg cagcccagct gccccaccac ctctgcctac cacctcctct ggcaaggctg 2880

caaagaaacc cacagctgca gaggtctctg ttcgagtccc tagagggccg gccgtggaag 2940

ggggacacgt gaatatggca ttctctcagt ccaaccctcc ttcggagttg cacaaggtcc 3000

acggatccag aaacaagccc accagcttgt ggaacccaac gtacggctcc tggtttacag 3060

[0082] agaaacccac caaaaagaat aatcctatag caaagaagga gccacacgac aggggtaacc 3120

tggggctgga gggaagctgt actgtcccac ctaacgttgc aactgggaga cttccggggg 3180

cctcactgct cctagagccc tcttcgctga ctgccaatat gaaggaggta cctctgttca 3240

ggctacgtca cttcccttgt gggaatgtca attacggcta ccagcaacag ggcttgccct 3300

tagaagccgc tactgcccct ggagctggtc attacgagga taccattctg aaaagcaaga 3360

atagcatgaa ccagcctggg ccctgagctc ggtcgcacac tcacttctct tccttgggat 3420

ccctaagacc gtgga 3435

<210> 22
<211> 1105
<212> PRT
<213>人类

<400> 22
Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser

1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln
20 25 30

Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val
35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys
50 55 60

Ser Val Ser Ser Lys Gly Gln Pro Ser Pro Arg Ala Val Ile Pro Met
65 70 75 80

Ser Cys Ile Thr Asn Gly Ser Gly Ala Asn Arg Lys Pro Ser His Thr
85 90 95

[0083]

Ser Ala Val Ser Ile Ala Gly Lys Glu Thr Leu Ser Ser Ala Ala Lys
100 105 110

Ser Gly Thr Glu Lys Lys Lys Glu Lys Pro Gln Gly Gln Arg Glu Lys
115 120 125

Lys Glu Glu Ser His Ser Asn Asp Gln Ser Pro Gln Ile Arg Ala Ser
130 135 140

Pro Ser Pro Gln Pro Ser Ser Gln Pro Leu Gln Ile His Arg Gln Thr
145 150 155 160

Pro Glu Ser Lys Asn Ala Thr Pro Thr Lys Ser Ile Lys Arg Pro Ser
165 170 175

Pro Ala Glu Lys Ser His Asn Ser Trp Glu Asn Ser Asp Asp Ser Arg

180 185 190

Asn Lys Leu Ser Lys Ile Pro Ser Thr Pro Lys Leu Ile Pro Lys Val
195 200 205

Thr Lys Thr Ala Asp Lys His Lys Asp Val Ile Ile Asn Gln Glu Gly
210 215 220

Glu Tyr Ile Lys Met Phe Met Arg Gly Arg Pro Ile Thr Met Phe Ile
225 230 235 240

Pro Ser Asp Val Asp Asn Tyr Asp Asp Ile Arg Thr Glu Leu Pro Pro
245 250 255

Glu Lys Leu Lys Leu Glu Trp Ala Tyr Gly Tyr Arg Gly Lys Asp Cys
260 265 270

[0084]

Arg Ala Asn Val Tyr Leu Leu Pro Thr Gly Glu Ile Val Tyr Phe Ile
275 280 285

Ala Ser Val Val Val Leu Phe Asn Tyr Glu Glu Arg Thr Gln Arg His
290 295 300

Tyr Leu Gly His Thr Asp Cys Val Lys Cys Leu Ala Ile His Pro Asp
305 310 315 320

Lys Ile Arg Ile Ala Thr Gly Gln Ile Ala Gly Val Asp Lys Asp Gly
325 330 335

Arg Pro Leu Gln Pro His Val Arg Val Trp Asp Ser Val Thr Leu Ser
340 345 350

Thr Leu Gln Ile Ile Gly Leu Gly Thr Phe Glu Arg Gly Val Gly Cys

355 360 365

Leu Asp Phe Ser Lys Ala Asp Ser Gly Val His Leu Cys Val Ile Asp
370 375 380

Asp Ser Asn Glu His Met Leu Thr Val Trp Asp Trp Gln Arg Lys Ala
385 390 395 400

Lys Gly Ala Glu Ile Lys Thr Thr Asn Glu Val Val Leu Ala Val Glu
405 410 415

Phe His Pro Thr Asp Ala Asn Thr Ile Ile Thr Cys Gly Lys Ser His
420 425 430

Ile Phe Phe Trp Thr Trp Ser Gly Asn Ser Leu Thr Arg Lys Gln Gly
435 440 445

[0085]

Ile Phe Gly Lys Tyr Glu Lys Pro Lys Phe Val Gln Cys Leu Ala Phe
450 455 460

Leu Gly Asn Gly Asp Val Leu Thr Gly Asp Ser Gly Gly Val Met Leu
465 470 475 480

Ile Trp Ser Lys Thr Thr Val Glu Pro Thr Pro Gly Lys Gly Pro Lys
485 490 495

Gly Val Tyr Gln Ile Ser Lys Gln Ile Lys Ala His Asp Gly Ser Val
500 505 510

Phe Thr Leu Cys Gln Met Arg Asn Gly Met Leu Leu Thr Gly Gly Gly
515 520 525

Lys Asp Arg Lys Ile Ile Leu Trp Asp His Asp Leu Asn Pro Glu Arg

530 535 540

Glu Ile Glu His Gln Glu Leu Gln Ala Met Gln Met Glu Leu Gln Ser
545 550 555 560

Pro Glu Tyr Lys Leu Ser Lys Leu Arg Thr Ser Thr Ile Met Thr Asp
565 570 575

Tyr Asn Pro Asn Tyr Cys Phe Ala Gly Lys Thr Ser Ser Ile Ser Asp
580 585 590

Leu Lys Glu Val Pro Arg Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly
595 600 605

His Gly Ala Phe Gly Glu Val Tyr Glu Gly Gln Val Ser Gly Met Pro
610 615 620

[0086]

Asn Asp Pro Ser Pro Leu Gln Val Ala Val Lys Thr Leu Pro Glu Val
625 630 635 640

Cys Ser Glu Gln Asp Glu Leu Asp Phe Leu Met Glu Ala Leu Ile Ile
645 650 655

Ser Lys Phe Asn His Gln Asn Ile Val Arg Cys Ile Gly Val Ser Leu
660 665 670

Gln Ser Leu Pro Arg Phe Ile Leu Leu Glu Leu Met Ala Gly Gly Asp
675 680 685

Leu Lys Ser Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln Pro Ser
690 695 700

Ser Leu Ala Met Leu Asp Leu Leu His Val Ala Arg Asp Ile Ala Cys

705 710 715 720

Gly Cys Gln Tyr Leu Glu Glu Asn His Phe Ile His Arg Asp Ile Ala
725 730 735

Ala Arg Asn Cys Leu Leu Thr Cys Pro Gly Pro Gly Arg Val Ala Lys
740 745 750

Ile Gly Asp Phe Gly Met Ala Arg Asp Ile Tyr Arg Ala Ser Tyr Tyr
755 760 765

Arg Lys Gly Gly Cys Ala Met Leu Pro Val Lys Trp Met Pro Pro Glu
770 775 780

Ala Phe Met Glu Gly Ile Phe Thr Ser Lys Thr Asp Thr Trp Ser Phe
785 790 795 800

[0087]

Gly Val Leu Leu Trp Glu Ile Phe Ser Leu Gly Tyr Met Pro Tyr Pro
805 810 815

Ser Lys Ser Asn Gln Glu Val Leu Glu Phe Val Thr Ser Gly Gly Arg
820 825 830

Met Asp Pro Pro Lys Asn Cys Pro Gly Pro Val Tyr Arg Ile Met Thr
835 840 845

Gln Cys Trp Gln His Gln Pro Glu Asp Arg Pro Asn Phe Ala Ile Ile
850 855 860

Leu Glu Arg Ile Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile Asn Thr
865 870 875 880

Ala Leu Pro Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu Glu Lys Val

885 890 895

Pro Val Arg Pro Lys Asp Pro Glu Gly Val Pro Pro Leu Leu Val Ser
900 905 910

Gln Gln Ala Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro Pro Pro
915 920 925

Leu Pro Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr Ala Ala
930 935 940

Glu Val Ser Val Arg Val Pro Arg Gly Pro Ala Val Glu Gly Gly His
945 950 955 960

Val Asn Met Ala Phe Ser Gln Ser Asn Pro Pro Ser Glu Leu His Lys
965 970 975

[0088]

Val His Gly Ser Arg Asn Lys Pro Thr Ser Leu Trp Asn Pro Thr Tyr
980 985 990

Gly Ser Trp Phe Thr Glu Lys Pro Thr Lys Lys Asn Asn Pro Ile Ala
995 1000 1005

Lys Lys Glu Pro His Asp Arg Gly Asn Leu Gly Leu Glu Gly Ser
1010 1015 1020

Cys Thr Val Pro Pro Asn Val Ala Thr Gly Arg Leu Pro Gly Ala
1025 1030 1035

Ser Leu Leu Leu Glu Pro Ser Ser Leu Thr Ala Asn Met Lys Glu
1040 1045 1050

Val Pro Leu Phe Arg Leu Arg His Phe Pro Cys Gly Asn Val Asn

1055 1060 1065

Tyr Gly Tyr Gln Gln Gln Gly Leu Pro Leu Glu Ala Ala Thr Ala
1070 1075 1080

Pro Gly Ala Gly His Tyr Glu Asp Thr Ile Leu Lys Ser Lys Asn
1085 1090 1095

Ser Met Asn Gln Pro Gly Pro
1100 1105

<210> 23
<211> 4479
<212> DNA
<213>人类

<400> 23

[0089]

tgcgagaaag atggcggacc tggccgagtg caacatcaaa gtgatgtgtc gcttcagacc 60

tctcaacgag tctgaagtga accgcggcga caagtacatc gccaagtttc agggagaaga 120

cacggtcgtg atcgcgtcca agccttatgc atttgatcgg gtgttccagt caagcacatc 180

tcaagagcaa gtgtataatg actgtgcaaa gaagattgtt aaagatgtac ttgaaggata 240

taatggaaca atatttgcat atggacaaac atcctctggg aagacacaca caatggaggg 300

taaacttcat gatccagaag gcatgggaat tattccaaga atagtgcaag atatttttaa 360

ttatatttac tccatggatg aaaatttgga atttcatatt aaggtttcat attttgaaat 420

atatttggat aagataaggg acctgttaga tgtttcaaag accaaccttt cagttcatga 480

agacaaaaac cgagttccct atgtaaaggg gtgcacagag cgttttgtat gtagtccaga 540

tgaagttatg gataccatag atgaaggaaa atccaacaga catgtagcag ttacaaatat 600

gaatgaacat agctctagga gtcacagtat atttcttatt aatgtcaaac aagagaacac 660

acaaacggaa caaaagctga gtggaaaact ttatctggtt gatttagctg gtagtgaaaa 720

ggttagtaaa actggagctg aaggtgctgt gctggatgaa gctaaaaaca tcaacaagtc 780

actttctgct cttggaaatg ttatttctgc tttggctgag ggtagtacat atgttccata 840

tcgagatagt aaaatgacaa gaatccttca agattcatta ggtggcaact gtagaaccac 900

tattgtaatt tgctgctctc catcatcata caatgagtct gaaacaaaat ctacactctt 960

atttggccaa agggccaaaa caattaagaa cacagtttgt gtcaatgtgg agttaactgc 1020

agaacagtgg aaaaagaagt atgaaaaaga aaaagaaaaa aataagatcc tgcggaacac 1080

tattcagtgg cttgaaaatg agctcaacag atggcgtaat ggggagacgg tgcctattga 1140

tgaacagttt gacaaagaga aagccaactt ggaagctttc acagtggata aagatattac 1200

tcttaccaat gataaaccag caaccgcaat tggagttata ggaaatttta ctgatgctga 1260

aagaagaaag tgtgaagaag aaattgctaa attatacaaa cagcttgatg acaaggatga 1320

agaaattaac cagcaaagtc aactggtaga gaaactgaag acgcaaatgt tggatcagga 1380

[0090]

ggagcttttg gcatctacca gaagggatca agacaatatg caagctgagc tgaatcgcct 1440

tcaagcagaa aatgatgcct ctaaagaaga agtgaaagaa gttttacagg ccctagaaga 1500

acttgctgtc aattatgatc agaagtctca ggaagttgaa gacaaaacta aggaatatga 1560

attgcttagt gatgaattga atcagaaatc ggcaacttta gcgagtatag atgctgagct 1620

tcagaaactt aaggaaatga ccaaccacca gaaaaaacga gcagctgaga tgatggcatc 1680

tttactaaaa gaccttgcag aaataggaat tgctgtggga aataatgatg taaagcagcc 1740

tgagggaact ggcatgatag atgaagagtt cactgttgca agactctaca ttagcaaaat 1800

gaagtcagaa gtaaaaacca tggtgaaacg ttgcaagcag ttagaaagca cacaaactga 1860

gagcaacaaa aaaatggaag aaaatgaaaa ggagttagca gcatgtcagc ttcgtatctc 1920

tcaacatgaa gccaaaatca agtcattgac tgaatacctt caaaatgtgg aacaaaagaa 1980

aagacagttg gaggaatctg tcgatgccct cagtgaagaa ctagtccagc ttcgagcaca 2040

agagaaagtc catgaaatgg aaaaggagca cttaaataag gttcagactg caaatgaagt 2100

taagcaagct gttgaacagc agatccagag ccatagagaa actcatcaaa aacagatcag 2160

tagtttgaga gatgaagtag aagcaaaagc aaaacttatt actgatcttc aagaccaaaa 2220

ccagaaaatg atgttagagc aggaacgtct aagagtagaa catgagaagt tgaaagccac 2280

agatcaggaa aagagcagaa aactacatga acttacggtt atgcaagata gacgagaaca 2340

agcaagacaa gacttgaagg gtttggaaga gacagtggca aaagaacttc agactttaca 2400

caacctgcgc aaactctttg ttcaggacct ggctacaaga gttaaaaaga gtgctgagat 2460

tgattctgat gacaccggag gcagcgctgc tcagaagcaa aaaatctcct ttcttgaaaa 2520

taatcttgaa cagctcacta aagtgcacaa acagttggta cgtgataatg cagatctccg 2580

ctgtgaactt cctaagttgg aaaagcgact tcgagctaca gctgagagag tgaaagcttt 2640

ggaatcagca ctgaaagaag ctaaagaaaa tgcatctcgt gatcgcaaac gctatcagca 2700

[0091]

agaagtagat cgcataaagg aagcagtcag gtcaaagaat atggccagaa gagggcattc 2760

tgcacagatt gtgtaccgcc ggaagcacca ggagctgcaa gccatgcaga tggagctgca 2820

gagccctgag tacaagctga gcaagctccg cacctcgacc atcatgaccg actacaaccc 2880

caactactgc tttgctggca agacctcctc catcagtgac ctgaaggagg tgccgcggaa 2940

aaacatcacc ctcattcggg gtctgggcca tggcgccttt ggggaggtgt atgaaggcca 3000

ggtgtccgga atgcccaacg acccaagccc cctgcaagtg gctgtgaaga cgctgcctga 3060

agtgtgctct gaacaggacg aactggattt cctcatggaa gccctgatca tcagcaaatt 3120

caaccaccag aacattgttc gctgcattgg ggtgagcctg caatccctgc cccggttcat 3180

cctgctggag ctcatggcgg ggggagacct caagtccttc ctccgagaga cccgccctcg 3240

cccgagccag ccctcctccc tggccatgct ggaccttctg cacgtggctc gggacattgc 3300

ctgtggctgt cagtatttgg aggaaaacca cttcatccac cgagacattg ctgccagaaa 3360

ctgcctcttg acctgtccag gccctggaag agtggccaag attggagact tcgggatggc 3420

ccgagacatc tacagggcga gctactatag aaagggaggc tgtgccatgc tgccagttaa 3480

gtggatgccc ccagaggcct tcatggaagg aatattcact tctaaaacag acacatggtc 3540

ctttggagtg ctgctatggg aaatcttttc tcttggatat atgccatacc ccagcaaaag 3600

caaccaggaa gttctggagt ttgtcaccag tggaggccgg atggacccac ccaagaactg 3660

ccctgggcct gtataccgga taatgactca gtgctggcaa catcagcctg aagacaggcc 3720

caactttgcc atcattttgg agaggattga atactgcacc caggacccgg atgtaatcaa 3780

caccgctttg ccgatagaat atggtccact tgtggaagag gaagagaaag tgcctgtgag 3840

gcccaaggac cctgaggggg ttcctcctct cctggtctct caacaggcaa aacgggagga 3900

ggagcgcagc ccagctgccc caccacctct gcctaccacc tcctctggca aggctgcaaa 3960

gaaacccaca gctgcagagg tctctgttcg agtccctaga gggccggccg tggaaggggg 4020

[0092]

acacgtgaat atggcattct ctcagtccaa ccctccttcg gagttgcaca aggtccacgg 4080

atccagaaac aagcccacca gcttgtggaa cccaacgtac ggctcctggt ttacagagaa 4140

acccaccaaa aagaataatc ctatagcaaa gaaggagcca cacgacaggg gtaacctggg 4200

gctggaggga agctgtactg tcccacctaa cgttgcaact gggagacttc cgggggcctc 4260

actgctccta gagccctctt cgctgactgc caatatgaag gaggtacctc tgttcaggct 4320

acgtcacttc ccttgtggga atgtcaatta cggctaccag caacagggct tgcccttaga 4380

agccgctact gccccctggag ctggtcatta cgaggatacc attctgaaaa gcaagaatag 4440

catgaaccag cctgggcccc gagctcggtc gcacactca 4479

<210> 24
<211> 1483
<212> PRT
<213>人类

<400> 24

Met Ala Asp Leu Ala Glu Cys Asn Ile Lys Val Met Cys Arg Phe Arg
1 5 10 15

Pro Leu Asn Glu Ser Glu Val Asn Arg Gly Asp Lys Tyr Ile Ala Lys
20 25 30

Phe Gln Gly Glu Asp Thr Val Val Ile Ala Ser Lys Pro Tyr Ala Phe
35 40 45

Asp Arg Val Phe Gln Ser Ser Thr Ser Gln Glu Gln Val Tyr Asn Asp
50 55 60

Cys Ala Lys Lys Ile Val Lys Asp Val Leu Glu Gly Tyr Asn Gly Thr
65 70 75 80

[0093] Ile Phe Ala Tyr Gly Gln Thr Ser Ser Gly Lys Thr His Thr Met Glu
85 90 95

Gly Lys Leu His Asp Pro Glu Gly Met Gly Ile Ile Pro Arg Ile Val
100 105 110

Gln Asp Ile Phe Asn Tyr Ile Tyr Ser Met Asp Glu Asn Leu Glu Phe
115 120 125

His Ile Lys Val Ser Tyr Phe Glu Ile Tyr Leu Asp Lys Ile Arg Asp
130 135 140

Leu Leu Asp Val Ser Lys Thr Asn Leu Ser Val His Glu Asp Lys Asn
145 150 155 160

Arg Val Pro Tyr Val Lys Gly Cys Thr Glu Arg Phe Val Cys Ser Pro
165 170 175

Asp Glu Val Met Asp Thr Ile Asp Glu Gly Lys Ser Asn Arg His Val
180 185 190

Ala Val Thr Asn Met Asn Glu His Ser Ser Arg Ser His Ser Ile Phe
195 200 205

Leu Ile Asn Val Lys Gln Glu Asn Thr Gln Thr Glu Gln Lys Leu Ser
210 215 220

Gly Lys Leu Tyr Leu Val Asp Leu Ala Gly Ser Glu Lys Val Ser Lys
225 230 235 240

Thr Gly Ala Glu Gly Ala Val Leu Asp Glu Ala Lys Asn Ile Asn Lys
245 250 255

[0094] Ser Leu Ser Ala Leu Gly Asn Val Ile Ser Ala Leu Ala Glu Gly Ser
260 265 270

Thr Tyr Val Pro Tyr Arg Asp Ser Lys Met Thr Arg Ile Leu Gln Asp
275 280 285

Ser Leu Gly Gly Asn Cys Arg Thr Thr Ile Val Ile Cys Cys Ser Pro
290 295 300

Ser Ser Tyr Asn Glu Ser Glu Thr Lys Ser Thr Leu Leu Phe Gly Gln
305 310 315 320

Arg Ala Lys Thr Ile Lys Asn Thr Val Cys Val Asn Val Glu Leu Thr
325 330 335

Ala Glu Gln Trp Lys Lys Lys Tyr Glu Lys Glu Lys Glu Lys Asn Lys
340 345 350

Ile Leu Arg Asn Thr Ile Gln Trp Leu Glu Asn Glu Leu Asn Arg Trp
355 360 365

Arg Asn Gly Glu Thr Val Pro Ile Asp Glu Gln Phe Asp Lys Glu Lys
370 375 380

Ala Asn Leu Glu Ala Phe Thr Val Asp Lys Asp Ile Thr Leu Thr Asn
385 390 395 400

Asp Lys Pro Ala Thr Ala Ile Gly Val Ile Gly Asn Phe Thr Asp Ala
405 410 415

Glu Arg Arg Lys Cys Glu Glu Glu Ile Ala Lys Leu Tyr Lys Gln Leu
420 425 430

[0095] Asp Asp Lys Asp Glu Glu Ile Asn Gln Gln Ser Gln Leu Val Glu Lys
435 440 445

Leu Lys Thr Gln Met Leu Asp Gln Glu Glu Leu Leu Ala Ser Thr Arg
450 455 460

Arg Asp Gln Asp Asn Met Gln Ala Glu Leu Asn Arg Leu Gln Ala Glu
465 470 475 480

Asn Asp Ala Ser Lys Glu Glu Val Lys Glu Val Leu Gln Ala Leu Glu
485 490 495

Glu Leu Ala Val Asn Tyr Asp Gln Lys Ser Gln Glu Val Glu Asp Lys
500 505 510

Thr Lys Glu Tyr Glu Leu Leu Ser Asp Glu Leu Asn Gln Lys Ser Ala
515 520 525

Thr Leu Ala Ser Ile Asp Ala Glu Leu Gln Lys Leu Lys Glu Met Thr
530 535 540

Asn His Gln Lys Lys Arg Ala Ala Glu Met Met Ala Ser Leu Leu Lys
545 550 555 560

Asp Leu Ala Glu Ile Gly Ile Ala Val Gly Asn Asn Asp Val Lys Gln
565 570 575

Pro Glu Gly Thr Gly Met Ile Asp Glu Glu Phe Thr Val Ala Arg Leu
580 585 590

Tyr Ile Ser Lys Met Lys Ser Glu Val Lys Thr Met Val Lys Arg Cys
595 600 605

[0096] Lys Gln Leu Glu Ser Thr Gln Thr Glu Ser Asn Lys Lys Met Glu Glu
610 615 620

Asn Glu Lys Glu Leu Ala Ala Cys Gln Leu Arg Ile Ser Gln His Glu
625 630 635 640

Ala Lys Ile Lys Ser Leu Thr Glu Tyr Leu Gln Asn Val Glu Gln Lys
645 650 655

Lys Arg Gln Leu Glu Glu Ser Val Asp Ala Leu Ser Glu Glu Leu Val
660 665 670

Gln Leu Arg Ala Gln Glu Lys Val His Glu Met Glu Lys Glu His Leu
675 680 685

Asn Lys Val Gln Thr Ala Asn Glu Val Lys Gln Ala Val Glu Gln Gln
690 695 700

Ile Gln Ser His Arg Glu Thr His Gln Lys Gln Ile Ser Ser Leu Arg
705 710 715 720

Asp Glu Val Glu Ala Lys Ala Lys Leu Ile Thr Asp Leu Gln Asp Gln
725 730 735

Asn Gln Lys Met Met Leu Glu Gln Glu Arg Leu Arg Val Glu His Glu
740 745 750

Lys Leu Lys Ala Thr Asp Gln Glu Lys Ser Arg Lys Leu His Glu Leu
755 760 765

Thr Val Met Gln Asp Arg Arg Glu Gln Ala Arg Gln Asp Leu Lys Gly
770 775 780

[0097] Leu Glu Glu Thr Val Ala Lys Glu Leu Gln Thr Leu His Asn Leu Arg
785 790 795 800

Lys Leu Phe Val Gln Asp Leu Ala Thr Arg Val Lys Lys Ser Ala Glu
805 810 815

Ile Asp Ser Asp Asp Thr Gly Gly Ser Ala Ala Gln Lys Gln Lys Ile
820 825 830

Ser Phe Leu Glu Asn Asn Leu Glu Gln Leu Thr Lys Val His Lys Gln
835 840 845

Leu Val Arg Asp Asn Ala Asp Leu Arg Cys Glu Leu Pro Lys Leu Glu
850 855 860

Lys Arg Leu Arg Ala Thr Ala Glu Arg Val Lys Ala Leu Glu Ser Ala
865 870 875 880

Leu Lys Glu Ala Lys Glu Asn Ala Ser Arg Asp Arg Lys Arg Tyr Gln
885 890 895

Gln Glu Val Asp Arg Ile Lys Glu Ala Val Arg Ser Lys Asn Met Ala
900 905 910

Arg Arg Gly His Ser Ala Gln Ile Val Tyr Arg Arg Lys His Gln Glu
915 920 925

Leu Gln Ala Met Gln Met Glu Leu Gln Ser Pro Glu Tyr Lys Leu Ser
930 935 940

Lys Leu Arg Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro Asn Tyr Cys
945 950 955 960

[0098] Phe Ala Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu Val Pro Arg
965 970 975

Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly His Gly Ala Phe Gly Glu
980 985 990

Val Tyr Glu Gly Gln Val Ser Gly Met Pro Asn Asp Pro Ser Pro Leu
995 1000 1005

Gln Val Ala Val Lys Thr Leu Pro Glu Val Cys Ser Glu Gln Asp
1010 1015 1020

Glu Leu Asp Phe Leu Met Glu Ala Leu Ile Ile Ser Lys Phe Asn
1025 1030 1035

His Gln Asn Ile Val Arg Cys Ile Gly Val Ser Leu Gln Ser Leu
1040 1045 1050

Pro Arg Phe Ile Leu Leu Glu Leu Met Ala Gly Gly Asp Leu Lys
1055 1060 1065

Ser Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln Pro Ser Ser
1070 1075 1080

Leu Ala Met Leu Asp Leu Leu His Val Ala Arg Asp Ile Ala Cys
1085 1090 1095

Gly Cys Gln Tyr Leu Glu Glu Asn His Phe Ile His Arg Asp Ile
1100 1105 1110

Ala Ala Arg Asn Cys Leu Leu Thr Cys Pro Gly Pro Gly Arg Val
1115 1120 1125

[0099] Ala Lys Ile Gly Asp Phe Gly Met Ala Arg Asp Ile Tyr Arg Ala
1130 1135 1140

Ser Tyr Tyr Arg Lys Gly Gly Cys Ala Met Leu Pro Val Lys Trp
1145 1150 1155

Met Pro Pro Glu Ala Phe Met Glu Gly Ile Phe Thr Ser Lys Thr
1160 1165 1170

Asp Thr Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Phe Ser Leu
1175 1180 1185

Gly Tyr Met Pro Tyr Pro Ser Lys Ser Asn Gln Glu Val Leu Glu
1190 1195 1200

Phe Val Thr Ser Gly Gly Arg Met Asp Pro Pro Lys Asn Cys Pro
1205 1210 1215

Gly Pro Val Tyr Arg Ile Met Thr Gln Cys Trp Gln His Gln Pro
1220 1225 1230

Glu Asp Arg Pro Asn Phe Ala Ile Ile Leu Glu Arg Ile Glu Tyr
1235 1240 1245

Cys Thr Gln Asp Pro Asp Val Ile Asn Thr Ala Leu Pro Ile Glu
1250 1255 1260

Tyr Gly Pro Leu Val Glu Glu Glu Glu Lys Val Pro Val Arg Pro
1265 1270 1275

Lys Asp Pro Glu Gly Val Pro Pro Leu Leu Val Ser Gln Gln Ala
1280 1285 1290

[0100] Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro Pro Pro Leu Pro
1295 1300 1305

Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr Ala Ala Glu
1310 1315 1320

Val Ser Val Arg Val Pro Arg Gly Pro Ala Val Glu Gly Gly His
1325 1330 1335

Val Asn Met Ala Phe Ser Gln Ser Asn Pro Pro Ser Glu Leu His
1340 1345 1350

Lys Val His Gly Ser Arg Asn Lys Pro Thr Ser Leu Trp Asn Pro
1355 1360 1365

Thr Tyr Gly Ser Trp Phe Thr Glu Lys Pro Thr Lys Lys Asn Asn
1370 1375 1380

Pro Ile Ala Lys Lys Glu Pro His Asp Arg Gly Asn Leu Gly Leu
1385 1390 1395

Glu Gly Ser Cys Thr Val Pro Pro Asn Val Ala Thr Gly Arg Leu
1400 1405 1410

Pro Gly Ala Ser Leu Leu Leu Glu Pro Ser Ser Leu Thr Ala Asn
1415 1420 1425

Met Lys Glu Val Pro Leu Phe Arg Leu Arg His Phe Pro Cys Gly
1430 1435 1440

Asn Val Asn Tyr Gly Tyr Gln Gln Gln Gly Leu Pro Leu Glu Ala
1445 1450 1455

[0101] Ala Thr Ala Pro Gly Ala Gly His Tyr Glu Asp Thr Ile Leu Lys
1460 1465 1470

Ser Lys Asn Ser Met Asn Gln Pro Gly Pro
1475 1480

<210> 25
<211> 2412
<212> DNA
<213>人类

<400> 25
atgaacggac agttggatct aagtgggaag ctaatcatca aagctcaact tggggaggat 60

attcggcgaa ttcctattca taatgaagat attacttatg atgaattagt gctaatgatg 120

caacgagttt tcagaggaaa acttctgagt aatgatgaag taacaataaa gtataaagat 180

gaagatggag atcttataac aatttttgat agttctgacc tttcctttgc aattcagtgc 240

agtaggatac tgaaactgac attatttgtt aatggccagc caagacccct tgaatcaagt 300

caggtgaaat atctccgtcg agaactgata gaacttcgaa ataaagtgaa tcgtttattg 360

gatagcttgg aaccacctgg agaaccagga ccttccacca atattcctga aaatgatact 420

gtggatggta gggaagaaaa gtctgcttct gattcttctg gaaaacagtc tactcaggtt 480

atggcagcaa gtatgtctgc ttttgatcct ttaaaaaacc aagatgaaat caataaaaat 540

gttatgtcag cgtttggctt aacagatgat caggtttcag ggccacccag tgctcctgca 600

gaagatcgtt caggaacacc cgacagcatt gcttcctcct cctcagcagc tcacccacca 660

ggcgttcagc cacagcagcc accatataca ggagctcaga ctcaagcagg tcagattgaa 720

gtgtaccgcc ggaagcacca ggagctgcaa gccatgcaga tggagctgca gagccctgag 780

tacaagctga gcaagctccg cacctcgacc atcatgaccg actacaaccc caactactgc 840

tttgctggca agacctcctc catcagtgac ctgaaggagg tgccgcggaa aaacatcacc 900

[0102]

ctcattcggg gtctgggcca tggcgccttt ggggaggtgt atgaaggcca ggtgtccgga 960

atgcccaacg acccaagccc cctgcaagtg gctgtgaaga cgctgcctga agtgtgctct 1020

gaacaggacg aactggattt cctcatggaa gccctgatca tcagcaaatt caaccaccag 1080

aacattgttc gctgcattgg ggtgagcctg caatccctgc cccggttcat cctgctggag 1140

ctcatggcgg ggggagacct caagtccttc ctccgagaga cccgccctcg cccgagccag 1200

cctcctcccc tggccatgct ggaccttctg cacgtggctc gggacattgc ctgtggctgt 1260

cagtatttgg aggaaaacca cttcatccac cgagacattg ctgccagaaa ctgcctcttg 1320

acctgtccag gccctggaag agtggccaag attggagact tcgggatggc ccgagacatc 1380

tacagggcga gctactatag aaagggaggc tgtgccatgc tgccagttaa gtggatgccc 1440

ccagaggcct tcatggaagg aatattcact tctaaaacag acacatggtc ctttggagtg 1500

ctgctatggg aaatcttttc tcttggatat atgccatacc ccagcaaaag caaccaggaa 1560

gttctggagt ttgtcaccag tggaggccgg atggacccac ccaagaactg ccctgggcct 1620

gtataccgga taatgactca gtgctggcaa catcagcctg aagacaggcc caactttgcc 1680

atcattttgg agaggattga atactgcacc caggacccgg atgtaatcaa caccgctttg 1740

ccgatagaat atggtccact tgtggaagag gaagagaaag tgcctgtgag gcccaaggac 1800

cctgaggggg ttcctcctct cctggtctct caacaggcaa aacgggagga ggagcgcagc 1860

ccagctgccc caccacctct gcctaccacc tcctctggca aggctgcaaa gaaacccaca 1920

gctgcagagg tctctgttcg agtccctaga gggccggccg tggaaggggg acacgtgaat 1980

atggcattct ctcagtccaa ccctccttcg gagttgcaca aggtccacgg atccagaaac 2040

aagcccacca gcttgtggaa cccaacgtac ggctcctggt ttacagagaa acccaccaaa 2100

aagaataatc ctatagcaaa gaaggagcca cacgacaggg gtaacctggg gctggaggga 2160

agctgtactg tcccacctaa cgttgcaact gggagacttc cgggggccta actgctccta 2220

[0103]

gagccctctt cgctgactgc caatatgaag gaggtacctc tgttcaggct acgtcacttc 2280

ccttgtggga atgtcaatta cggctaccag caacagggct tgcccttaga agccgctact 2340

gcccctggag ctggtcatta cgaggatacc attctgaaaa gcaagaatag catgaaccag 2400

cctgggccct ga 2412

<210> 26
<211> 803
<212> PRT
<213>人类

<400> 26
Met Asn Gly Gln Leu Asp Leu Ser Gly Lys Leu Ile Ile Lys Ala Gln
1 5 10 15

Leu Gly Glu Asp Ile Arg Arg Ile Pro Ile His Asn Glu Asp Ile Thr
20 25 30

Tyr Asp Glu Leu Val Leu Met Met Gln Arg Val Phe Arg Gly Lys Leu
35 40 45

Leu Ser Asn Asp Glu Val Thr Ile Lys Tyr Lys Asp Glu Asp Gly Asp
50 55 60

Leu Ile Thr Ile Phe Asp Ser Ser Asp Leu Ser Phe Ala Ile Gln Cys
65 70 75 80

Ser Arg Ile Leu Lys Leu Thr Leu Phe Val Asn Gly Gln Pro Arg Pro
85 90 95

Leu Glu Ser Ser Gln Val Lys Tyr Leu Arg Arg Glu Leu Ile Glu Leu
100 105 110

[0104] Arg Asn Lys Val Asn Arg Leu Leu Asp Ser Leu Glu Pro Pro Gly Glu
115 120 125

Pro Gly Pro Ser Thr Asn Ile Pro Glu Asn Asp Thr Val Asp Gly Arg
130 135 140

Glu Glu Lys Ser Ala Ser Asp Ser Ser Gly Lys Gln Ser Thr Gln Val
145 150 155 160

Met Ala Ala Ser Met Ser Ala Phe Asp Pro Leu Lys Asn Gln Asp Glu
165 170 175

Ile Asn Lys Asn Val Met Ser Ala Phe Gly Leu Thr Asp Asp Gln Val
180 185 190

Ser Gly Pro Pro Ser Ala Pro Ala Glu Asp Arg Ser Gly Thr Pro Asp
195 200 205

Ser Ile Ala Ser Ser Ser Ser Ala Ala His Pro Pro Gly Val Gln Pro
210 215 220

Gln Gln Pro Pro Tyr Thr Gly Ala Gln Thr Gln Ala Gly Gln Ile Glu
225 230 235 240

Val Tyr Arg Arg Lys His Gln Glu Leu Gln Ala Met Gln Met Glu Leu
245 250 255

Gln Ser Pro Glu Tyr Lys Leu Ser Lys Leu Arg Thr Ser Thr Ile Met
260 265 270

Thr Asp Tyr Asn Pro Asn Tyr Cys Phe Ala Gly Lys Thr Ser Ser Ile
275 280 285

[0105] Ser Asp Leu Lys Glu Val Pro Arg Lys Asn Ile Thr Leu Ile Arg Gly
290 295 300

Leu Gly His Gly Ala Phe Gly Glu Val Tyr Glu Gly Gln Val Ser Gly
305 310 315 320

Met Pro Asn Asp Pro Ser Pro Leu Gln Val Ala Val Lys Thr Leu Pro
325 330 335

Glu Val Cys Ser Glu Gln Asp Glu Leu Asp Phe Leu Met Glu Ala Leu
340 345 350

Ile Ile Ser Lys Phe Asn His Gln Asn Ile Val Arg Cys Ile Gly Val
355 360 365

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

Gly Asp Leu Lys Ser Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln
385 390 395 400

Pro Ser Ser Leu Ala Met Leu Asp Leu Leu His Val Ala Arg Asp Ile
405 410 415

Ala Cys Gly Cys Gln Tyr Leu Glu Glu Asn His Phe Ile His Arg Asp
420 425 430

Ile Ala Ala Arg Asn Cys Leu Leu Thr Cys Pro Gly Pro Gly Arg Val
435 440 445

Ala Lys Ile Gly Asp Phe Gly Met Ala Arg Asp Ile Tyr Arg Ala Ser
450 455 460

[0106] Tyr Tyr Arg Lys Gly Gly Cys Ala Met Leu Pro Val Lys Trp Met Pro
465 470 475 480

Pro Glu Ala Phe Met Glu Gly Ile Phe Thr Ser Lys Thr Asp Thr Trp
485 490 495

Ser Phe Gly Val Leu Leu Trp Glu Ile Phe Ser Leu Gly Tyr Met Pro
500 505 510

Tyr Pro Ser Lys Ser Asn Gln Glu Val Leu Glu Phe Val Thr Ser Gly
515 520 525

Gly Arg Met Asp Pro Pro Lys Asn Cys Pro Gly Pro Val Tyr Arg Ile
530 535 540

Met Thr Gln Cys Trp Gln His Gln Pro Glu Asp Arg Pro Asn Phe Ala
545 550 555 560

Ile Ile Leu Glu Arg Ile Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile
565 570 575

Asn Thr Ala Leu Pro Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu Glu
580 585 590

Lys Val Pro Val Arg Pro Lys Asp Pro Glu Gly Val Pro Pro Leu Leu
595 600 605

Val Ser Gln Gln Ala Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro
610 615 620

Pro Pro Leu Pro Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr
625 630 635 640

[0107] Ala Ala Glu Val Ser Val Arg Val Pro Arg Gly Pro Ala Val Glu Gly
645 650 655

Gly His Val Asn Met Ala Phe Ser Gln Ser Asn Pro Pro Ser Glu Leu
660 665 670

His Lys Val His Gly Ser Arg Asn Lys Pro Thr Ser Leu Trp Asn Pro
675 680 685

Thr Tyr Gly Ser Trp Phe Thr Glu Lys Pro Thr Lys Lys Asn Asn Pro
690 695 700

Ile Ala Lys Lys Glu Pro His Asp Arg Gly Asn Leu Gly Leu Glu Gly
705 710 715 720

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

Ser Leu Leu Leu Glu Pro Ser Ser Leu Thr Ala Asn Met Lys Glu Val
740 745 750

Pro Leu Phe Arg Leu Arg His Phe Pro Cys Gly Asn Val Asn Tyr Gly
755 760 765

Tyr Gln Gln Gln Gly Leu Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala
770 775 780

Gly His Tyr Glu Asp Thr Ile Leu Lys Ser Lys Asn Ser Met Asn Gln
785 790 795 800

Pro Gly Pro

[0108] <210> 27
<211> 2779
<212> DNA
<213>人类

<400> 27
cctccgcaag ccgtctttct ctagagttgt atatatagaa catcctggag tccaccatga 60

acggacagtt ggatctaagt gggaagctaa tcatcaaagc tcaacttggg gaggatattc 120

ggcgaattcc tattcataat gaagatatta cttatgatga attagtgcta atgatgcaac 180

gagttttcag aggaaaactt ctgagtaatg atgaagtaac aataaagtat aaagatgaag 240

atggagatct tataacaatt tttgatagtt ctgacctttc ctttgcaatt cagtgcagta 300

ggatactgaa actgacatta tttgttaatg gccagccaag acccccttgaa tcaagtcagg 360

tgaaatatct ccgtcgagaa ctgatagaac ttcgaaataa agtgaatcgt ttattggata 420

gcttggaacc acctggagaa ccaggacctt ccaccaatat tcctgaaaat gatactgtgg 480

atggtaggga agaaaagtct gcttctgatt cttctggaaa acagtctact caggttatgg 540

cagcaagtat gtctgctttt gatcctttaa aaaaccaaga tgaaatcaat aaaaatgtta 600

tgtcagcgtt tggcttaaca gatgatcagg tttcagtgta ccgccggaag caccaggagc 660

tgcaagccat gcagatggag ctgcagagcc ctgagtacaa gctgagcaag ctccgcacct 720

cgaccatcat gaccgactac aaccccaact actgctttgc tggcaagacc tcctccatca 780

gtgacctgaa ggaggtgccg cggaaaaaca tcaccctcat tcggggtctg ggccatggcg 840

cctttgggga ggtgtatgaa ggccaggtgt ccggaatgcc caacgaccca agccccctgc 900

aagtggctgt gaagacgctg cctgaagtgt gctctgaaca ggacgaactg gatttcctca 960

tggaagccct gatcatcagc aaattcaacc accagaacat tgttcgctgc attggggtga 1020

gcctgcaatc cctgccccgg ttcatcctgc tggagctcat ggcgggggga gacctcaagt 1080

ccttcctccg agagacccgc cctcgcccga gccagccctc ctccctggcc atgctggacc 1140

ttctgcacgt ggctcgggac attgcctgtg gctgtcagta tttggaggaa aaccacttca 1200

[0109]

tccaccgaga cattgctgcc agaaactgcc tcttgacctg tccaggccct ggaagagtgg 1260

ccaagattgg agacttcggg atggcccgag acatctacag ggcgagctac tatagaaagg 1320

gaggctgtgc catgctgcca gttaagtgga tgccccccaga ggccttcatg gaaggaatat 1380

tcacttctaa aacagacaca tggtcctttg gagtgctgct atgggaaatc ttttctcttg 1440

gatatatgcc ataccccagc aaaagcaacc aggaagttct ggagtttgtc accagtggag 1500

gccggatgga cccacccaag aactgccctg ggcctgtata ccggataatg actcagtgct 1560

ggcaacatca gcctgaagac aggcccaact ttgccatcat tttggagagg attgaatact 1620

gcacccagga cccggatgta atcaacaccg ctttgccgat agaatatggt ccacttgtgg 1680

aagaggaaga gaaagtgcct gtgaggccca aggaccctga gggggttcct cctctcctgg 1740

tctctcaaca ggcaaaacgg gaggaggagc gcagcccagc tgccccacca cctctgccta 1800

ccacctcctc tggcaaggct gcaaagaaac ccacagctgc agaggtctct gttcgagtcc 1860

ctagagggcc ggccgtggaa gggggacacg tgaatatggc attctctcag tccaaccctc 1920

cttcggagtt gcacaaggtc cacggatcca gaaacaagcc caccagcttg tggaacccaa 1980

cgtacggctc ctggtttaca gagaaaccca ccaaaaagaa taatcctata gcaaagaagg 2040

agccacacga caggggtaac ctggggctgg agggaagctg tactgtccca cctaacgttg 2100

caactgggag acttccgggg gcctcactgc tcctagagcc ctcttcgctg actgccaata 2160

tgaaggaggt acctctgttc aggctacgtc acttcccttg tgggaatgtc aattacggct 2220

accagcaaca gggcttgccc ttagaagccg ctactgcccc tggagctggt cattacgagg 2280

ataccattct gaaaagcaag aatagcatga accagcctgg gccctgagct cggtcgcaca 2340

ctcacttctc ttccttggga tccctaagac cgtggaggag agagaggcaa tggctccttc 2400

acaaaccaga gaccaaatgt cacgttttgt tttgtgccaa cctattttga agtaccacca 2460

[0110] aaaaagctgt attttgaaaa tgctttagaa aggttttgag catgggttca tcctattctt 2520

tcgaaagaag aaaatatcat aaaaatgagt gataaataca aggcccagat gtggttgcat 2580

aaggttttta tgcatgtttg ttgtatactt ccttatgctt cttttaaatt gtgtgtgctc 2640

tgcttcaatg tagtcagaat tagctgcttc tatgtttcat agttggggtc atagatgttt 2700

ccttgccttg ttgatgtgga catgagccat ttgaggggag agggaacgga aataaaggag 2760

ttatttgtaa tgactaaaa 2779

<210> 28
<211> 756
<212> PRT
<213>人类

<400> 28
Met Asn Gly Gln Leu Asp Leu Ser Gly Lys Leu Ile Ile Lys Ala Gln
1 5 10 15

Leu Gly Glu Asp Ile Arg Arg Ile Pro Ile His Asn Glu Asp Ile Thr
20 25 30

Tyr Asp Glu Leu Val Leu Met Met Gln Arg Val Phe Arg Gly Lys Leu
35 40 45

Leu Ser Asn Asp Glu Val Thr Ile Lys Tyr Lys Asp Glu Asp Gly Asp
50 55 60

Leu Ile Thr Ile Phe Asp Ser Ser Asp Leu Ser Phe Ala Ile Gln Cys
65 70 75 80

Ser Arg Ile Leu Lys Leu Thr Leu Phe Val Asn Gly Gln Pro Arg Pro
85 90 95

[0111] Leu Glu Ser Ser Gln Val Lys Tyr Leu Arg Arg Glu Leu Ile Glu Leu
100 105 110

Arg Asn Lys Val Asn Arg Leu Leu Asp Ser Leu Glu Pro Pro Gly Glu
115 120 125

Pro Gly Pro Ser Thr Asn Ile Pro Glu Asn Asp Thr Val Asp Gly Arg
130 135 140

Glu Glu Lys Ser Ala Ser Asp Ser Ser Gly Lys Gln Ser Thr Gln Val
145 150 155 160

Met Ala Ala Ser Met Ser Ala Phe Asp Pro Leu Lys Asn Gln Asp Glu
165 170 175

Ile Asn Lys Asn Val Met Ser Ala Phe Gly Leu Thr Asp Asp Gln Val
180 185 190

Ser Val Tyr Arg Arg Lys His Gln Glu Leu Gln Ala Met Gln Met Glu
195 200 205

Leu Gln Ser Pro Glu Tyr Lys Leu Ser Lys Leu Arg Thr Ser Thr Ile
210 215 220

Met Thr Asp Tyr Asn Pro Asn Tyr Cys Phe Ala Gly Lys Thr Ser Ser
225 230 235 240

Ile Ser Asp Leu Lys Glu Val Pro Arg Lys Asn Ile Thr Leu Ile Arg
245 250 255

Gly Leu Gly His Gly Ala Phe Gly Glu Val Tyr Glu Gly Gln Val Ser
260 265 270

[0112] Gly Met Pro Asn Asp Pro Ser Pro Leu Gln Val Ala Val Lys Thr Leu
275 280 285

Pro Glu Val Cys Ser Glu Gln Asp Glu Leu Asp Phe Leu Met Glu Ala
290 295 300

Leu Ile Ile Ser Lys Phe Asn His Gln Asn Ile Val Arg Cys Ile Gly
305 310 315 320

Val Ser Leu Gln Ser Leu Pro Arg Phe Ile Leu Leu Glu Leu Met Ala
325 330 335

Gly Gly Asp Leu Lys Ser Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser
340 345 350

Gln Pro Ser Ser Leu Ala Met Leu Asp Leu Leu His Val Ala Arg Asp
355 360 365

Ile Ala Cys Gly Cys Gln Tyr Leu Glu Glu Asn His Phe Ile His Arg
370 375 380

Asp Ile Ala Ala Arg Asn Cys Leu Leu Thr Cys Pro Gly Pro Gly Arg
385 390 395 400

Val Ala Lys Ile Gly Asp Phe Gly Met Ala Arg Asp Ile Tyr Arg Ala
405 410 415

Ser Tyr Tyr Arg Lys Gly Gly Cys Ala Met Leu Pro Val Lys Trp Met
420 425 430

Pro Pro Glu Ala Phe Met Glu Gly Ile Phe Thr Ser Lys Thr Asp Thr
435 440 445

[0113] Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Phe Ser Leu Gly Tyr Met
450 455 460

Pro Tyr Pro Ser Lys Ser Asn Gln Glu Val Leu Glu Phe Val Thr Ser
465 470 475 480

Gly Gly Arg Met Asp Pro Pro Lys Asn Cys Pro Gly Pro Val Tyr Arg
485 490 495

Ile Met Thr Gln Cys Trp Gln His Gln Pro Glu Asp Arg Pro Asn Phe
500 505 510

Ala Ile Ile Leu Glu Arg Ile Glu Tyr Cys Thr Gln Asp Pro Asp Val
515 520 525

Ile Asn Thr Ala Leu Pro Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu
530 535 540

Glu Lys Val Pro Val Arg Pro Lys Asp Pro Glu Gly Val Pro Pro Leu
545 550 555 560

Leu Val Ser Gln Gln Ala Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala
565 570 575

Pro Pro Pro Leu Pro Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro
580 585 590

Thr Ala Ala Glu Val Ser Val Arg Val Pro Arg Gly Pro Ala Val Glu
595 600 605

Gly Gly His Val Asn Met Ala Phe Ser Gln Ser Asn Pro Pro Ser Glu
610 615 620

[0114]

Leu His Lys Val His Gly Ser Arg Asn Lys Pro Thr Ser Leu Trp Asn
625 630 635 640

Pro Thr Tyr Gly Ser Trp Phe Thr Glu Lys Pro Thr Lys Lys Asn Asn
645 650 655

Pro Ile Ala Lys Lys Glu Pro His Asp Arg Gly Asn Leu Gly Leu Glu
660 665 670

Gly Ser Cys Thr Val Pro Pro Asn Val Ala Thr Gly Arg Leu Pro Gly
675 680 685

Ala Ser Leu Leu Leu Glu Pro Ser Ser Leu Thr Ala Asn Met Lys Glu
690 695 700

Val Pro Leu Phe Arg Leu Arg His Phe Pro Cys Gly Asn Val Asn Tyr
705 710 715 720

Gly Tyr Gln Gln Gln Gly Leu Pro Leu Glu Ala Ala Thr Ala Pro Gly
725 730 735

Ala Gly His Tyr Glu Asp Thr Ile Leu Lys Ser Lys Asn Ser Met Asn
740 745 750

Gln Pro Gly Pro
755

<210> 29
<211> 2614
<212> DNA
<213>人类

<400> 29
cctccgcaag ccgtctttct ctagagttgt atatatagaa catcctggag tccaccatga 60

acggacagtt ggatctaagt gggaagctaa tcatcaaagc tcaacttggg gaggatattc 120

[0115]

ggcgaattcc tattcataat gaagatatta cttatgatga attagtgcta atgatgcaac 180

gagttttcag aggaaaactt ctgagtaatg atgaagtaac aataaagtat aaagatgaag 240

atggagatct tataacaatt tttgatagtt ctgacctttc ctttgcaatt cagtgcagta 300

ggatactgaa actgacatta tttgttaatg gccagccaag acccccttgaa tcaagtcagg 360

tgaaatatct ccgtcgagaa ctgatagaac ttcgaaataa agtgaatcgt ttattggata 420

gcttggaacc acctggagaa ccaggacctt ccaccaatat tcctgaaaat gtgtaccgcc 480

ggaagcacca ggagctgcaa gccatgcaga tggagctgca gagccctgag tacaagctga 540

gcaagctccg cacctcgacc atcatgaccg actacaaccc caactactgc tttgctggca 600

agacctcctc catcagtgac ctgaaggagg tgccgcggaa aaacatcacc ctcattcggg 660

gtctgggcca tggcgccttt ggggaggtgt atgaaggcca ggtgtccgga atgcccaacg 720

acccaagccc cctgcaagtg gctgtgaaga cgctgcctga agtgtgctct gaacaggacg 780

aactggattt cctcatggaa gccctgatca tcagcaaatt caaccaccag aacattgttc 840

gctgcattgg ggtgagcctg caatccctgc cccggttcat cctgctggag ctcatggcgg 900

ggggagacct caagtccttc ctccgagaga cccgccctcg cccgagccag ccctcctccc 960

tggccatgct ggaccttctg cacgtggctc gggacattgc ctgtggctgt cagtatttgg 1020

aggaaaacca cttcatccac cgagacattg ctgccagaaa ctgcctcttg acctgtccag 1080

gccctggaag agtggccaag attggagact tcgggatggc ccgagacatc tacagggcga 1140

gctactatag aaagggaggc tgtgccatgc tgccagttaa gtggatgccc ccagaggcct 1200

tcatggaagg aatattcact tctaaaacag acacatggtc ctttggagtg ctgctatggg 1260

aaatcttttc tcttggatat atgccatacc ccagcaaaag caaccaggaa gttctggagt 1320

ttgtcaccag tggaggccgg atggacccac ccaagaactg ccctgggcct gtataccgga 1380

taatgactca gtgctggcaa catcagcctg aagacaggcc caactttgcc atcattttgg 1440

[0116]

agaggattga atactgcacc caggacccgg atgtaatcaa caccgctttg ccgatagaat 1500

atggtccact tgtggaagag gaagagaaag tgcctgtgag gcccaaggac cctgaggggg 1560

ttcctcctct cctggtctct caacaggcaa aacgggagga ggagcgcagc ccagctgccc 1620

caccacctct gcctaccacc tcctctggca aggctgcaaa gaaacccaca gctgcagagg 1680

tctctgttcg agtccctaga gggccggccg tggaaggggg acacgtgaat atggcattct 1740

ctcagtccaa ccctccttcg gagttgcaca aggtccacgg atccagaaac aagcccacca 1800

gcttgtggaa cccaacgtac ggctcctggt ttacagagaa acccaccaaa aagaataatc 1860

ctatagcaaa gaaggagcca cacgacaggg gtaacctggg gctggaggga agctgtactg 1920

tcccacctaa cgttgcaact gggagacttc cgggggcctc actgctccta gagccctctt 1980

cgctgactgc caatatgaag gaggtacctc tgttcaggct acgtcacttc ccttgtggga 2040

atgtcaatta cggctaccag caacagggct tgcccttaga agccgctact gcccctggag 2100

ctggtcatta cgaggatacc attctgaaaa gcaagaatag catgaaccag cctgggccct 2160

gagctcggtc gcacactcac ttctcttcct tgggatccct aagaccgtgg aggagagaga 2220

ggcaatggct ccttcacaaa ccagagacca aatgtcacgt tttgttttgt gccaacctat 2280

tttgaagtac caccaaaaaa gctgtatttt gaaaatgctt tagaaaggtt ttgagcatgg 2340

gttcatccta ttctttcgaa agaagaaaat atcataaaaa tgagtgataa atacaaggcc 2400

cagatgtggt tgcataaggt ttttatgcat gtttgttgta tacttcctta tgcttctttt 2460

aaattgtgtg tgctctgctt caatgtagtc agaattagct gcttctatgt ttcatagttg 2520

gggtcataga tgtttccttg ccttgttgat gtggacatga gccatttgag gggagaggga 2580

acggaaataa aggagttatt tgtaatgact aaaa 2614

<210> 30
<211> 701
[0117] <212> PRT
<213>人类

<400> 30
Met Asn Gly Gln Leu Asp Leu Ser Gly Lys Leu Ile Ile Lys Ala Gln
1 5 10 15

Leu Gly Glu Asp Ile Arg Arg Ile Pro Ile His Asn Glu Asp Ile Thr
20 25 30

Tyr Asp Glu Leu Val Leu Met Met Gln Arg Val Phe Arg Gly Lys Leu
35 40 45

Leu Ser Asn Asp Glu Val Thr Ile Lys Tyr Lys Asp Glu Asp Gly Asp
50 55 60

Leu Ile Thr Ile Phe Asp Ser Ser Asp Leu Ser Phe Ala Ile Gln Cys
65 70 75 80

Ser Arg Ile Leu Lys Leu Thr Leu Phe Val Asn Gly Gln Pro Arg Pro
85 90 95

Leu Glu Ser Ser Gln Val Lys Tyr Leu Arg Arg Glu Leu Ile Glu Leu
100 105 110

Arg Asn Lys Val Asn Arg Leu Leu Asp Ser Leu Glu Pro Pro Gly Glu
115 120 125

Pro Gly Pro Ser Thr Asn Ile Pro Glu Asn Val Tyr Arg Arg Lys His
130 135 140

Gln Glu Leu Gln Ala Met Gln Met Glu Leu Gln Ser Pro Glu Tyr Lys
145 150 155 160

[0118] Leu Ser Lys Leu Arg Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro Asn
165 170 175

Tyr Cys Phe Ala Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu Val
180 185 190

Pro Arg Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly His Gly Ala Phe
195 200 205

Gly Glu Val Tyr Glu Gly Gln Val Ser Gly Met Pro Asn Asp Pro Ser
210 215 220

Pro Leu Gln Val Ala Val Lys Thr Leu Pro Glu Val Cys Ser Glu Gln
225 230 235 240

Asp Glu Leu Asp Phe Leu Met Glu Ala Leu Ile Ile Ser Lys Phe Asn
245 250 255

His Gln Asn Ile Val Arg Cys Ile Gly Val Ser Leu Gln Ser Leu Pro
260 265 270

Arg Phe Ile Leu Leu Glu Leu Met Ala Gly Gly Asp Leu Lys Ser Phe
275 280 285

Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln Pro Ser Ser Leu Ala Met
290 295 300

Leu Asp Leu Leu His Val Ala Arg Asp Ile Ala Cys Gly Cys Gln Tyr
305 310 315 320

Leu Glu Glu Asn His Phe Ile His Arg Asp Ile Ala Ala Arg Asn Cys
325 330 335

[0119] Leu Leu Thr Cys Pro Gly Pro Gly Arg Val Ala Lys Ile Gly Asp Phe
340 345 350

Gly Met Ala Arg Asp Ile Tyr Arg Ala Ser Tyr Tyr Arg Lys Gly Gly
355 360 365

Cys Ala Met Leu Pro Val Lys Trp Met Pro Pro Glu Ala Phe Met Glu
370 375 380

Gly Ile Phe Thr Ser Lys Thr Asp Thr Trp Ser Phe Gly Val Leu Leu
385 390 395 400

Trp Glu Ile Phe Ser Leu Gly Tyr Met Pro Tyr Pro Ser Lys Ser Asn
405 410 415

Gln Glu Val Leu Glu Phe Val Thr Ser Gly Gly Arg Met Asp Pro Pro
420 425 430

Lys Asn Cys Pro Gly Pro Val Tyr Arg Ile Met Thr Gln Cys Trp Gln
435 440 445

His Gln Pro Glu Asp Arg Pro Asn Phe Ala Ile Ile Leu Glu Arg Ile
450 455 460

Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile Asn Thr Ala Leu Pro Ile
465 470 475 480

Glu Tyr Gly Pro Leu Val Glu Glu Glu Glu Lys Val Pro Val Arg Pro
485 490 495

Lys Asp Pro Glu Gly Val Pro Pro Leu Leu Val Ser Gln Gln Ala Lys
500 505 510

[0120] Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro Pro Pro Leu Pro Thr Thr
515 520 525

Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr Ala Ala Glu Val Ser Val
530 535 540

Arg Val Pro Arg Gly Pro Ala Val Glu Gly Gly His Val Asn Met Ala
545 550 555 560

Phe Ser Gln Ser Asn Pro Pro Ser Glu Leu His Lys Val His Gly Ser
565 570 575

Arg Asn Lys Pro Thr Ser Leu Trp Asn Pro Thr Tyr Gly Ser Trp Phe
580 585 590

Thr Glu Lys Pro Thr Lys Lys Asn Asn Pro Ile Ala Lys Lys Glu Pro
595 600 605

His Asp Arg Gly Asn Leu Gly Leu Glu Gly Ser Cys Thr Val Pro Pro
610 615 620

Asn Val Ala Thr Gly Arg Leu Pro Gly Ala Ser Leu Leu Leu Glu Pro
625 630 635 640

Ser Ser Leu Thr Ala Asn Met Lys Glu Val Pro Leu Phe Arg Leu Arg
645 650 655

His Phe Pro Cys Gly Asn Val Asn Tyr Gly Tyr Gln Gln Gln Gly Leu
660 665 670

Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala Gly His Tyr Glu Asp Thr
675 680 685

Ile Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly Pro
[0121] 690 695 700

<210> 31
<211> 1607
<212> DNA
<213>人类

<400> 31
aactccgctg cctttgccgc caccatgccc aaaacgatca gtgtgcgtgt gaccaccatg 60

gatgcagagc tggagtttgc catccagccc aacaccaccg ggaagcagct atttgaccag 120

gtggtgaaaa ctattggctt gagggaagtt tggttctttg gtctgcagta ccaggacact 180

aaaggtttct ccacctggct gaaactcaat aagaaggtga ctgcccagga tgtgcggaag 240

gaaagccccc tgctctttaa gttccgtgcc aagttctacc ctgaggatgt gtccgaggaa 300

ttgattcagg acatcactca gcgcctgttc tttctgcaag tgaaagaggg cattctcaat 360

gatgatattt actgcccgcc tgagaccgct gtgctgctgg cctcgtatgc tgtccagtct 420

aagtatggcg acttcaataa ggaagtgcat aagtctggct acctggccgg agacaagttg 480

ctcccgcaga gagtcctgga acagcacaaa ctcaacaagg accagtggga ggagcggatc 540

caggtgtggc atgaggaaca ccgtggcatg ctcagggagg atgctgtcct ggaatatctg 600

aagattgctc aagatctgga gatgtatggt gtgaactact tcagcatcaa gaacaagaaa 660

ggctcagagc tgtggctggg ggtggatgcc ctgggtctca acatctatga gcagaatgac 720

agactaactc ccaagatagg cttcccctgg agtgaaatca ggaacatctc tttcaatgat 780

aagaaatttg tcatcaagcc cattgacaaa aaagccccgg acttcgtctt ctatgctccc 840

cggctgcgga ttaacaagcg gatcttggcc ttgtgcatgg ggaaccatga actatacatg 900

cgccgtcgca agcctgatac cattgaggtg cagcagatga aggcacaggc ccgggaggag 960

aagcaccaga agcagatgga gcgtgctatg ctggaaaatg agaagaagaa gcgtgaaatg 1020

gcagagaagg agaaagagaa gattgaacgg gagaaggagg agctgatgga gaggctgaag 1080

[0122]

cagatcgagg aacagactaa gaaggctcag caagaactgg aagaacagac ccgtagggct 1140

ctggaacttg agcaggaacg gaagcgtgcc cagagcgagg ctgaaaagct ggccaaggag 1200

cgtcaagaag ctgaagaggc caaggaggcc ttgctgcagg cctcccggga ccagaaaaag 1260

actcaggaac agctggcctt ggaaatggca gagctgacag ctcgaatctc ccagctggag 1320

atggcccgac agaagaagga gagtgaggct gtggagtggc agcagaagca ggagctgcaa 1380

gccatgcaga tggagctgca gagccctgag tacaagctga gcaagctccg cacctcgacc 1440

atcatgaccg actacaaccc caactactgc tttgctggca agacctcctc catcagtgac 1500

ctgaaggagg tgccgcggaa aaacatcacc ctcattcggg gtctgggcca tggcgccttt 1560

ggggaggtgt atgaaggcca ggtgtccgga atgcccaacg acccaag 1607

<210> 32

<211> 527

<212> PRT
<213>人类

<400> 32
Met Pro Lys Thr Ile Ser Val Arg Val Thr Thr Met Asp Ala Glu Leu
1 5 10 15

Glu Phe Ala Ile Gln Pro Asn Thr Thr Gly Lys Gln Leu Phe Asp Gln
20 25 30

Val Val Lys Thr Ile Gly Leu Arg Glu Val Trp Phe Phe Gly Leu Gln
35 40 45

Tyr Gln Asp Thr Lys Gly Phe Ser Thr Trp Leu Lys Leu Asn Lys Lys
50 55 60

[0123] Val Thr Ala Gln Asp Val Arg Lys Glu Ser Pro Leu Leu Phe Lys Phe
65 70 75 80

Arg Ala Lys Phe Tyr Pro Glu Asp Val Ser Glu Glu Leu Ile Gln Asp
85 90 95

Ile Thr Gln Arg Leu Phe Phe Leu Gln Val Lys Glu Gly Ile Leu Asn
100 105 110

Asp Asp Ile Tyr Cys Pro Pro Glu Thr Ala Val Leu Leu Ala Ser Tyr
115 120 125

Ala Val Gln Ser Lys Tyr Gly Asp Phe Asn Lys Glu Val His Lys Ser
130 135 140

Gly Tyr Leu Ala Gly Asp Lys Leu Leu Pro Gln Arg Val Leu Glu Gln
145 150 155 160

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

Glu Glu His Arg Gly Met Leu Arg Glu Asp Ala Val Leu Glu Tyr Leu
180 185 190

Lys Ile Ala Gln Asp Leu Glu Met Tyr Gly Val Asn Tyr Phe Ser Ile
195 200 205

Lys Asn Lys Lys Gly Ser Glu Leu Trp Leu Gly Val Asp Ala Leu Gly
210 215 220

Leu Asn Ile Tyr Glu Gln Asn Asp Arg Leu Thr Pro Lys Ile Gly Phe
225 230 235 240

[0124] Pro Trp Ser Glu Ile Arg Asn Ile Ser Phe Asn Asp Lys Lys Phe Val
245 250 255

Ile Lys Pro Ile Asp Lys Lys Ala Pro Asp Phe Val Phe Tyr Ala Pro
260 265 270

Arg Leu Arg Ile Asn Lys Arg Ile Leu Ala Leu Cys Met Gly Asn His
275 280 285

Glu Leu Tyr Met Arg Arg Arg Lys Pro Asp Thr Ile Glu Val Gln Gln
290 295 300

Met Lys Ala Gln Ala Arg Glu Glu Lys His Gln Lys Gln Met Glu Arg
305 310 315 320

Ala Met Leu Glu Asn Glu Lys Lys Lys Arg Glu Met Ala Glu Lys Glu
325 330 335

Lys Glu Lys Ile Glu Arg Glu Lys Glu Glu Leu Met Glu Arg Leu Lys
340 345 350

Gln Ile Glu Glu Gln Thr Lys Lys Ala Gln Gln Glu Leu Glu Glu Gln
355 360 365

Thr Arg Arg Ala Leu Glu Leu Glu Gln Glu Arg Lys Arg Ala Gln Ser
370 375 380

Glu Ala Glu Lys Leu Ala Lys Glu Arg Gln Glu Ala Glu Glu Ala Lys
385 390 395 400

Glu Ala Leu Leu Gln Ala Ser Arg Asp Gln Lys Lys Thr Gln Glu Gln
405 410 415

[0125] Leu Ala Leu Glu Met Ala Glu Leu Thr Ala Arg Ile Ser Gln Leu Glu
420 425 430

Met Ala Arg Gln Lys Lys Glu Ser Glu Ala Val Glu Trp Gln Gln Lys
435 440 445

Gln Glu Leu Gln Ala Met Gln Met Glu Leu Gln Ser Pro Glu Tyr Lys
450 455 460

Leu Ser Lys Leu Arg Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro Asn
465 470 475 480

Tyr Cys Phe Ala Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu Val
485 490 495

Pro Arg Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly His Gly Ala Phe
500 505 510

Gly Glu Val Tyr Glu Gly Gln Val Ser Gly Met Pro Asn Asp Pro
515 520 525

<210> 33
<211> 426
<212> DNA
<213>人类

<400> 33
cgagaagttg agggagaaag gcgggcccgg gaacaggctg aggctgaggt ggcctccttg 60

aaccgtagga tccagctggt tgaagaagag ctggaccgtg ctcaggagcg tgcggaggtg 120

tctgaactaa aatgtggtga cctggaagaa gaactcaaga atgttactaa caatctgaaa 180

tctctggagg ctgcatctga aaagtattct gaaaaggagg acaaatatga agaagaaatt 240

aaacttctgt ctgacaaact gaaagaggct gagacccgtg ctgaatttgc agagagaacg 300

gttgcaaaac tggaaaagac aattgatgac ctggaagtgt acctccggaa gcaccaagag 360

[0126]

ctgcaagcca tgcagatgga gctgcagagc cctgagtaca agctgagcaa gctccgcacc 420

ctcgac 426

<210> 34
<211> 142
<212> PRT
<213>人类

<400> 34
Arg Glu Val Glu Gly Glu Arg Arg Ala Arg Glu Gln Ala Glu Ala Glu
1 5 10 15

Val Ala Ser Leu Asn Arg Arg Ile Gln Leu Val Glu Glu Glu Leu Asp
20 25 30

Arg Ala Gln Glu Arg Ala Glu Val Ser Glu Leu Lys Cys Gly Asp Leu
35 40 45

Glu Glu Glu Leu Lys Asn Val Thr Asn Asn Leu Lys Ser Leu Glu Ala
50 55 60

Ala Ser Glu Lys Tyr Ser Glu Lys Glu Asp Lys Tyr Glu Glu Glu Ile
65 70 75 80

Lys Leu Leu Ser Asp Lys Leu Lys Glu Ala Glu Thr Arg Ala Glu Phe
85 90 95

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

Val Tyr Leu Arg Lys His Gln Glu Leu Gln Ala Met Gln Met Glu Leu
115 120 125

[0127] Gln Ser Pro Glu Tyr Lys Leu Ser Lys Leu Arg Thr Leu Asp
130 135 140

<210> 35
<211> 308
<212> DNA
<213>人类

<400> 35
ctggcagagt cccgttgccg agagatggat gagcagatta gactgatgga ccagaacctg 60

aagtgtctga gtgctgctga agaaaagtac tctcaaaaag aagataaata tgaggaagaa 120

atcaagattc ttactgataa actcaaggag gcagagaccc gtgctgaatt tgcagagaga 180

acggttgcaa aactggaaaa gacaattgat gacctggaag tgtaccgccg gaagcaccag 240

gagctgcaag ccatgcagat ggagctgcag agccctgagt acaagctgag caagctccgc 300

acctcgac 308

<210> 36
<211> 103
<212> PRT
<213>人类

<400> 36
Leu Ala Glu Ser Arg Cys Arg Glu Met Asp Glu Gln Ile Arg Leu Met
1 5 10 15

Asp Gln Asn Leu Lys Cys Leu Ser Ala Ala Glu Glu Lys Tyr Ser Gln
20 25 30

Lys Glu Asp Lys Tyr Glu Glu Glu Ile Lys Ile Leu Thr Asp Lys Leu
35 40 45

Lys Glu Ala Glu Thr Arg Ala Glu Phe Ala Glu Arg Thr Val Ala Lys
50 55 60

[0128]

Leu Glu Lys Thr Ile Asp Asp Leu Glu Val Tyr Arg Arg Lys His Gln
65 70 75 80

Glu Leu Gln Ala Met Gln Met Glu Leu Gln Ser Pro Glu Tyr Lys Leu
85 90 95

Ser Lys Leu Arg Thr Ser Thr
100

<210> 37
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 37
tacaacccca actactgctt tgct 24

<210> 38
<211> 24
<212> DNA
<213> 人工序列

<220>
<223>人工序列的描述：合成引物

<400> 38
aggcactttc tcttcctctt ccac 24

<210> 39
<211> 24
<212> DNA
<213> 人工序列

[0129]

<220>
<223>人工序列的描述：合成引物

<400> 39
ccacacctgg gaaaggacct aaag 24

<210> 40
<211> 24
<212> DNA
<213> 人工序列

<220>
<223>人工序列的描述：合成引物

<400> 40
cctccaaata ctgacagcca cagg 24

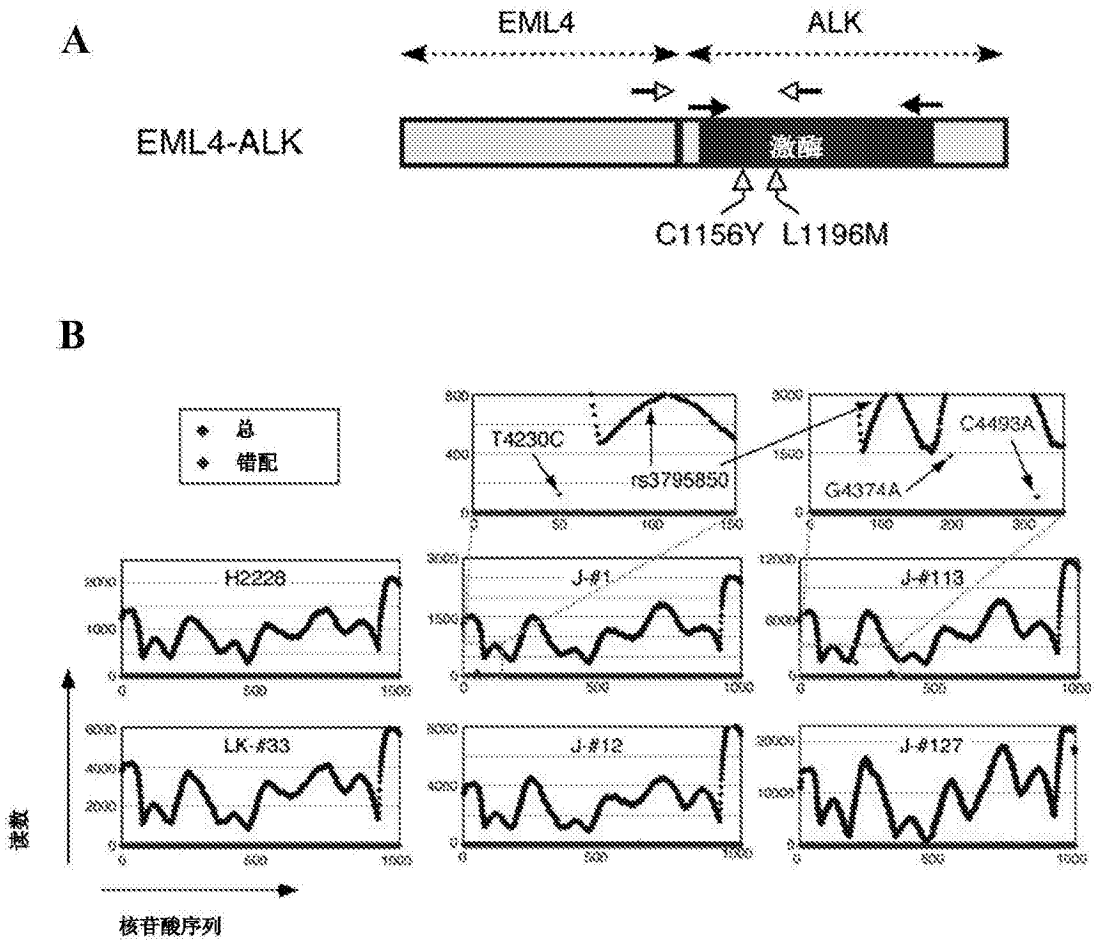

A
EML4
ALK
EML4-ALK
激酶
C1156Y L1196M
B
总
错配
T4230C
rs3795850
C4493A
G4374A
H2228
J-#1
J-#113
LK-#33
J-#12
J-#127
核苷酸序列
C
开始
恶化
G4374A
T G T G C T C
T G T A C T C
C4493A
C T G C T G G
C T G A T G G

图1

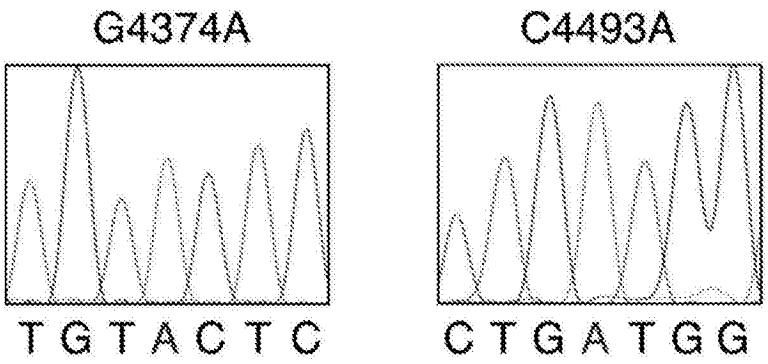


图2

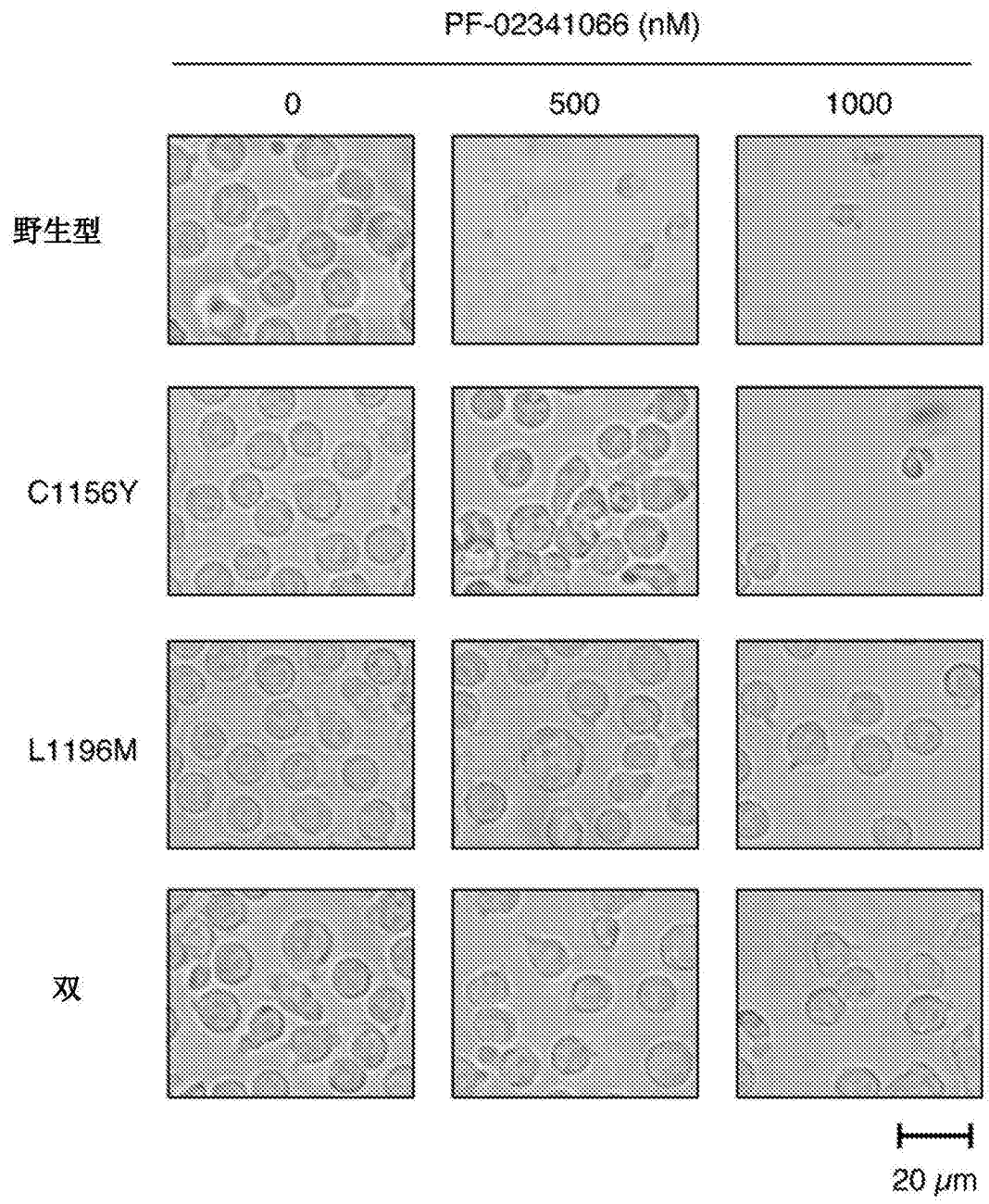
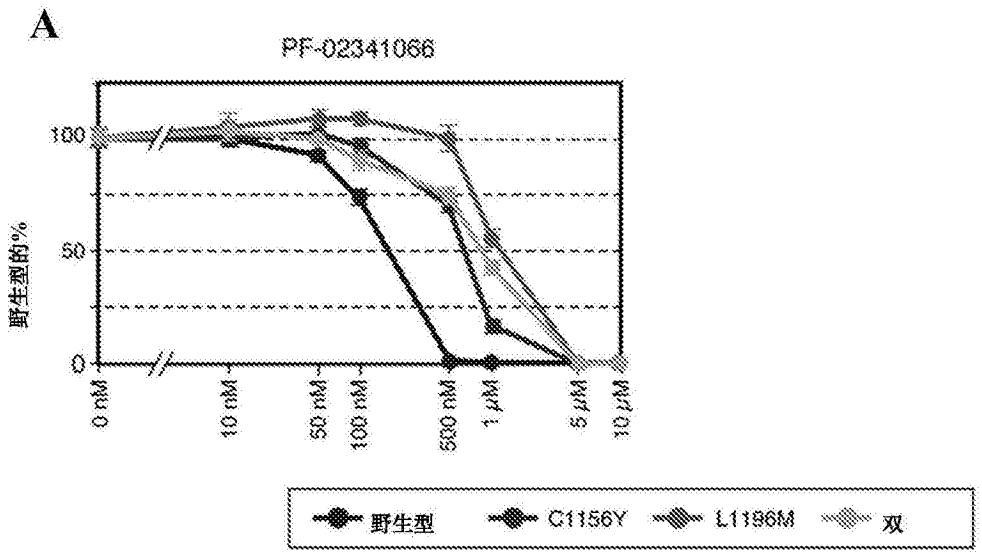
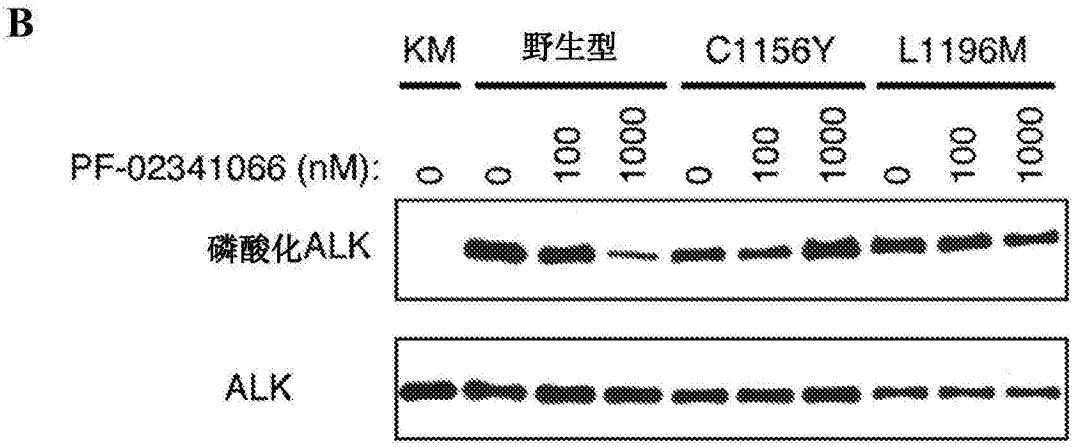


图3


A
PF-02341066
野生型的%
100
50
0
0 nM
10 nM
50 nM
100 nM
500 nM
1 μM
5 μM
10 μM
野生型
C1156Y
L1196M
双


B
KM
野生型
C1156Y
L1196M
PF-02341066 (nM):
0
0
100
1000
0
100
1000
0
100
1000
磷酸化ALK
ALK

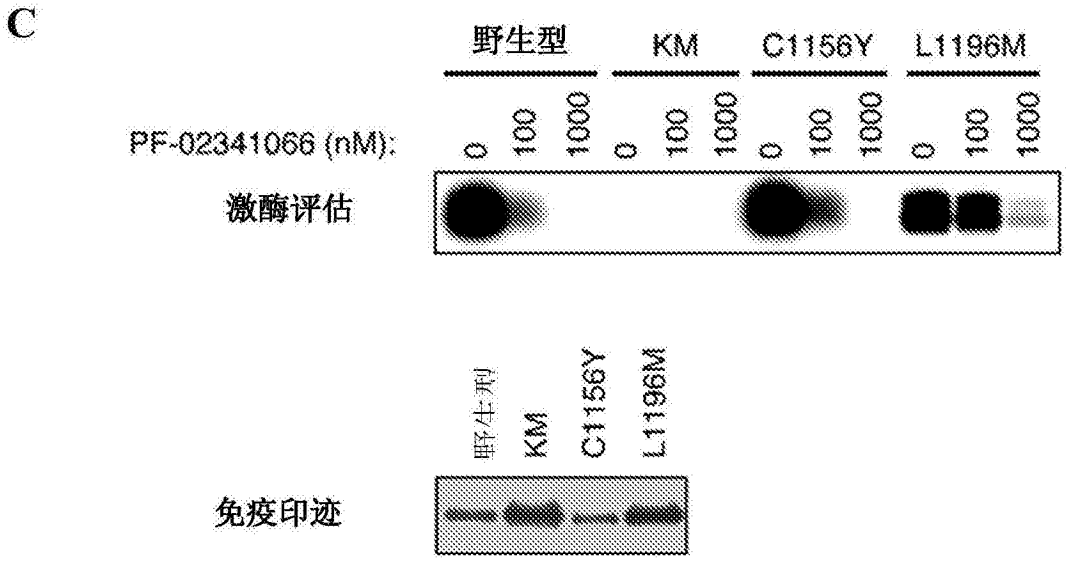

C
野生型
KM
C1156Y
L1196M
PF-02341066 (nM):
0
100
1000
0
100
1000
0
100
1000
0
100
1000
激酶评估
野生型
KM
C1156Y
L1196M
免疫印迹

图4

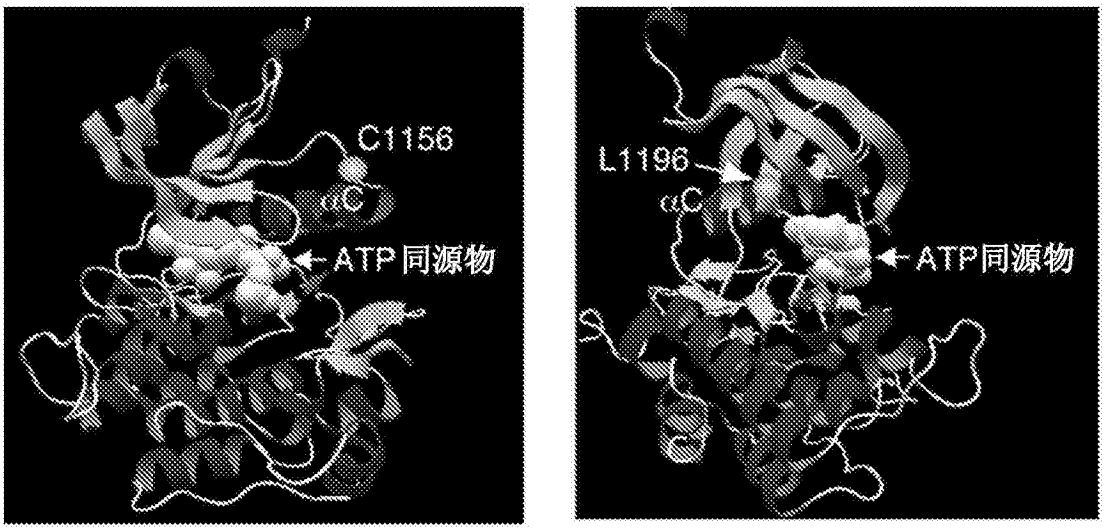


图5