

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200580030863.2

[51] Int. Cl.

C12Q 1/68 (2006.01)

C07H 21/04 (2006.01)

[43] 公开日 2007 年 8 月 15 日

[11] 公开号 CN 101018874A

[22] 申请日 2005.8.12

[21] 申请号 200580030863.2

[30] 优先权

[32] 2004.8.13 [33] US [31] 60/601,413

[86] 国际申请 PCT/US2005/028781 2005.8.12

[87] 国际公布 WO2006/028655 英 2006.3.16

[85] 进入国家阶段日期 2007.3.14

[71] 申请人 千年药品公司

地址 美国马萨诸塞州

[72] 发明人 J·E·莫纳汉 S·卡马特卡

S·霍尔施 B·O·戈巴彻瓦

K·格拉特 D·福特

W·O·恩德格 D·L·安德森

[74] 专利代理机构 北京市金杜律师事务所

代理人 陈文平

权利要求书 4 页 说明书 87 页 序列表 56 页

[54] 发明名称

用于鉴定、评估、预防和治疗前列腺癌的基因、组合物、试剂盒和方法

[57] 摘要

本发明涉及新发现的与前列腺癌包括恶变前状况有关的核酸分子和蛋白质。提供了用于检测、表征、预防和治疗人类前列腺癌的组合物、试剂盒和方法。

1. 一种评估患者是否患有前列腺癌的方法，该方法包括：
 - a) 测定患者样品中一种标志物的表达水平，其中该标志物选自表1中所列的标志物；
 - b) 测定对照样品中该标志物的表达水平；
 - c) 比较患者样品中与对照样品中该标志物的表达水平；和
 - d) 当患者样品中与对照样品中该标志物的表达水平的显著差异指示该患者患有乳腺癌时，确定该患者患有前列腺癌，如此评估患者是否患有前列腺癌。
2. 权利要求1的方法，其中对照样品的表达水平通过选自下组的方法来确定：
 - a) 由来自非癌患者的前列腺细胞测定的水平；
 - b) 由来自良性前列腺增生或无前列腺肿瘤受试者的前列腺细胞测定的水平；和
 - c) 使用具有良性前列腺增生或无前列腺肿瘤的受试者群的平均表达水平预先确定的水平。
3. 权利要求1的方法，其中所述标志物对应于一种分泌的蛋白质。
4. 权利要求1的方法，其中所述标志物包含转录的多核苷酸或其部分。
5. 权利要求1的方法，其中所述样品包括选自下组的样品：
 - a) 从患者获得的细胞；和
 - b) 选自血液、淋巴、尿、前列腺液和精液的液体。
6. 权利要求3的方法，其中所述标志物蛋白质的存在用特异性结合该蛋白质的试剂来检测。
7. 权利要求6的方法，其中所述试剂选自抗体、抗体衍生物和抗体片段。
8. 权利要求4的方法，其中通过检测样品中对应于核酸标志物的

转录多核苷酸或其部分的存在来评估样品中标志物的表达水平。

9. 权利要求 8 的方法，其中检测转录的多核苷酸包括扩增所述转录的多核苷酸。

10. 权利要求 8 的方法，其中通过检测样品中一种转录的多核苷酸的存在来评估样品中标志物的表达水平，该转录的多核苷酸在严格杂交条件下与一种核酸标志物或者其一部分退火。

11. 权利要求 1 的方法，其中样品中所述标志物的表达水平与未患前列腺癌者中该标志物的正常表达水平相差至少大约 2 倍或者至少大约 5 倍。

12. 一种评估患者是否患有前列腺癌的方法，该方法包括：

a) 测定样品中至少两种标志物的表达水平，这些标志物独立地选自表 1 中所列的标志物；

b) 测定对照样品中每种标志物的表达水平；

c) 比较患者样品中与来自对照样品中该标志物的表达水平；和

d) 当至少两种标志物的表达水平相对于这些标志物的相应对照表达水平的显著差异指示该患者患有前列腺癌时，确定该患者患有前列腺癌；

如此评估患者是否患有前列腺癌。

13. 权利要求 12 的方法，其中对照样品的表达水平通过选自下组的方法来确定：

a) 由来自非癌患者的前列腺细胞测定的水平；

b) 由来自良性前列腺增生或无前列腺肿瘤受试者的前列腺细胞测定的水平；和

c) 使用具有良性前列腺增生或无前列腺肿瘤的受试者群的平均表达水平预先确定的水平。

14. 权利要求 12 的方法，其中测定至少三种标志物的表达。

15. 权利要求 12 的方法，其中测定至少五种标志物的表达。

16. 一种监测患者中前列腺癌进展的方法，该方法包括：

a) 在第一时间点测定患者样品中标志物的表达水平，其中该标志

物选自表 1 所列的标志物;

b) 在后续的时间点测定患者样品中标志物的表达水平; 和

c) 比较在步骤 a) 和 b) 中检测的表达水平, 从而监测患者中前列腺癌的进展,

其中, 标志物表达水平的改变指示前列腺癌的进展或消退。

17. 权利要求 16 的方法, 其中所述标志物对应于一种分泌的蛋白质。

18. 权利要求 16 的方法, 其中所述标志物包含转录的多核苷酸或其部分。

19. 权利要求 16 的方法, 其中所述样品包括选自下组的样品:

a) 从患者获得的细胞; 和

b) 选自血液、淋巴、尿、前列腺液和精液的液体。

20. 权利要求 16 的方法, 其中所述患者已经在第一时间点和后续时间点之间进行了切除肿瘤的手术。

21. 一种鉴定抑制患者中前列腺癌的候选试验化合物的方法, 该方法包括:

a) 测定从该患者中获得并且暴露于试验化合物的第一样品中一种标志物的表达, 其中该标志物选自表 1 中所列的标志物,

b) 测定从该患者中获得的第二样品中该标志物的表达, 其中该样品未暴露于该试验化合物,

c) 比较暴露于试验化合物的样品中与未暴露于试验化合物的样品中该标志物的表达; 和

d) 当暴露于试验化合物的样品中该标志物的表达水平显著低于第二样品指示该试验化合物对抑制患者中的前列腺癌有效时, 确定该试验化合物是抑制患者中前列腺癌的候选化合物。

22. 权利要求 21 的方法, 其中所述第一和第二样品是从患者中获得的同一样品的部分, 或者所述第一和第二样品是从患者中获得的合并样品的部分。

23. 一种评估患者是否患有前列腺癌的试剂盒, 该试剂盒包括用

于评估选自表 1 所列标志物的至少一种标志物的表达的试剂;

其中该试剂盒包括选自下组的试剂:

a) 至少一种核酸探针, 其中该探针特异性结合对应于选自表 1 所列标志物的至少一种标志物的转录多核苷酸; 和

b) 至少一种抗体, 其中该抗体特异性结合对应于选自表 1 所列标志物的至少一种标志物的蛋白质。

24. 一种分离的核酸分子, 其包含选自 SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 21, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29 和 SEQ ID NO: 43 的核苷酸序列。

25. 一种由权利要求 24 的核酸编码的分离的多肽, 其包含选自 SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 22, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30 和 SEQ ID NO: 44 的氨基酸序列。

26. 一种选择性结合权利要求 25 的多肽的抗体。

用于鉴定、评估、预防和治疗前列腺癌的 基因、组合物、试剂盒和方法

相关申请

本申请要求于 2004 年 8 月 13 日提交的美国临时申请 60/601,413 的优先权，其完整内容在此引作参考。

发明领域

本发明的领域为前列腺癌，包括前列腺癌的诊断、表征、处理和
治疗。

发明背景

在美国以及实际上在世界范围内报告的增加的癌症病例数都备受关注。当前对于特定类型的癌症仅有少数几种治疗方法有效，并且无法提供绝对成功的保证。为使疗效最好，这些治疗方法不仅需要
对恶性肿瘤早期检测，还需要对恶性肿瘤的严重程度进行可靠的评估。

前列腺癌（PCA）是在美国男性中最常被诊断出的癌症，并且是排第二位的男性癌症死亡原因（Karp 等人, 1996, *Cancer Res.* 56:5547-5556）。男性中该器官对癌症的急性易感性尚未被认识。斯基恩腺是女性中与男性前列腺相对应的组织，但不是观察到明显致瘤性转化的部位。

前列腺癌所表现出的一种不同寻常的挑战是大多数前列腺肿瘤不是危及生命的疾病。尸检调查表明多达 1100 万美国男性患有前列腺癌（Dhom, 1983, *J. Cancer Res. Clin. Oncol.*, 106:210-218）。这些数字与前列腺癌的临床观察结果一致，即通常表现为缓慢和延迟的进展过程。这样的疾病进展导致相对较少的前列腺肿瘤在患者的生命期间发展为临床上关注的病例。如果一旦用有效的方法检测到，癌症就似乎已经是

分化良好、局灶且器官限制的，治疗通常不能延长老年患者的预期寿命。

不幸的是，自然进展的前列腺癌在利用有效方法临床检测到时通常已经转移。患转移性前列腺癌的个体的存活率极低。在这两个极端之间是在生命期间将会转移但是尚未转移的前列腺肿瘤患者。对于这些患者，手术切除前列腺是可治愈的，并且延长预期寿命。因此，准确地确定新诊断的患者属于哪一类对于确定最佳治疗和患者的存活是关键性的。

当前医生至少可以使用一种早期的非侵入性试验来检测无症状的疾病。使用标准的基于抗体的检测试剂盒可以相对容易地检测血样中前列腺特异性抗原(PSA)的存在。患者血清中存在异常高水平的该抗原表明前列腺疾病的可能性，可能是癌症、良性前列腺增生(BPH)或前列腺炎。在大多数病例中，PSA 升高是由于 BPH 或前列腺炎而不是由于癌症。

尽管已经利用临床和病理学分期和组织学分级系统（例如 Gleason's）根据肿瘤分化程度或腺型类型来指示患者组的预后(Carter 和 Coffey, In: J. P. Karr 和 H. Yamanak (eds.), *Prostate Cancer: The Second Tokyo Symposium*, pp. 19-27, New York: Elsevier, 1989.; Diamond 等人, *J. Urol.*, 128: 729-734, 1982)，但是这些系统不能充分地预测癌症的进展速度。尽管已经建议使用原发损伤的组织切片的“核圆度”的计算机系统图像分析来作为处置患者个体的辅助手段(Diamond 等人, 1982, *J. Urol.*, 128:729-734)，但是该方法在研究疾病进展中应用有限。

使用流式细胞术和 FISH 分析 DNA 含量/倍性已经证明可用于预测前列腺癌的侵占性(Pearsons 等人, 1993, *J. Urol.*, 150:120-125; Macoska 等人, 1994, *Cancer Res.*, 54: 3824-3830; Visakorpi 等人, 1994, *Am. J. Pathol.*, 145:1-7; Takahashi 等人, 1994, *Cancer Res.*, 54:3574-3579; Alcaraz 等人, *Cancer Res.*, 55:3998-4002, 1994)，但是这些方法昂贵、费时，并且后一种方法需要构建着丝粒特异性探针进行分析。还存在已经报道其表达与前列腺癌有关的特异性核基质蛋白。但是，这些蛋白

标志物显然不能区别 BPH 和前列腺癌(Partin 等人, 1993, *Cancer Res.*, 53:744-746)。遗憾的是, 不能区别良性和恶性前列腺肿瘤的标志物没有价值。

因此, 提供用于诊断、分期、预后、监测和治疗前列腺癌的方法和试剂是有益的。

发明内容

本发明涉及癌症标志物(下文称为“标志物”或“本发明的标志物”), 它们在表 1 中列出。本发明提供这些标志物编码的或相对应的核酸和蛋白质(下文分别称为“标志物核酸”和“标志物蛋白”)。本发明还提供与这些标志物蛋白和/或标志物蛋白的片段特异性结合的抗体、抗体衍生物和抗体片段。

本发明还涉及用于诊断、分期、预后、监测和治疗前列腺癌的各种方法、试剂和试剂盒。在一个实施方案中, 本发明提供一种评估患者是否患有前列腺癌或者其发展为前列腺癌的危险是否高于正常的诊断方法, 包括比较患者样品中至少一种本发明标志物的表达水平和对照, 如未患前列腺癌者样品中所述标志物的正常表达水平。患者样品中所述标志物的表达水平显著提高指示所述患者患有前列腺癌或具有发展前列腺癌的危险。

在另一个实施方案中, 本发明提供了评估患者是否患有前列腺癌或是否可能发展为前列腺癌的诊断方法, 包括比较患者样品中至少一种本发明标志物的表达水平和患有良性前列腺增生或无前列腺肿瘤的对照受试者样品中该标志物的表达水平。标志物的表达水平升高可指示前列腺癌。

因此, 本发明的方法可以用于鉴定发展前列腺癌的危险升高的患者(如具有前列腺癌家族史的患者、被鉴定为具有突变的癌基因的患者)。该方法也是用于评估患者是否患有前列腺癌或是否可能发展前列腺癌的有用的诊断方法。

本发明的方法也可以用于预测前列腺癌的特定分期，以及评估癌症是否转移（如转移到淋巴结）。此外，本发明的方法也可以用于预测前列腺癌患者或经受根除前列腺癌治疗的患者的临床结果。此外，本发明的方法也用于评估前列腺癌患者的治疗效果（如化疗效果）。另外，本发明的方法也可用于评估物质的与前列腺相关的致癌性。

根据本发明，选择标志物，使得本发明方法的阳性预测值为至少约 10%，优选约 25%，更优选约 50%以及最优选约 90%。本发明的以下实施方案也是优选的，其中与正常非前列腺癌患者相比，在至少约 15%的前列腺癌患者（包括，例如 T1 期前列腺癌患者，T2 期前列腺癌患者，T3 期前列腺癌患者，T4 期前列腺癌患者，N 期前列腺癌患者，M 期前列腺癌患者，以及任何其它类型的与前列腺相关的癌、恶性肿瘤和转化）中所述标志物过量表达至少 5 倍。

本发明进一步提供了用于评估患者是否患有已经转移的或可能转移的前列腺癌的诊断方法，该方法包括比较患者样品中至少一种表 1 所列的标志物的表达水平和患有非转移前列腺肿瘤或无前列腺肿瘤的对照受试者样品中所述标志物的表达水平。与对照受试者样品中的水平相比，患者样品中的表达水平明显较高表明前列腺癌已经转移或可能转移。

本发明还提供了用于预测前列腺癌患者的临床结果的方法，包括比较患者样品中至少一种表 1 所列的标志物的表达水平和具有良好临床结果的对照受试者（如具有大于 5 年的无病存活水平的以前的前列腺癌患者）样品中所述标志物的表达水平。与对照受试者样品中的表达水平相比，患者样品中的表达水平明显较高表明患者具有不良的结果（如小于三年的无病存活）。

本发明也提供用于评估一种疗法抑制患者前列腺癌的疗效的方法。该方法包括将对患者提供至少一部分治疗之前从患者中获得的第一个样品中至少一种本发明标志物的表达与实施该部分治疗后从患者中获得的第二个样品中该标志物的表达进行比较。该标志物在第二个样品中的表达水平显著低于在第一个样品中的表达水平表明该疗法对于抑制患

者前列腺癌是有效的。

应当理解，在这些方法中，“疗法”可以是治疗前列腺癌的任何疗法，包括但不限于化疗、放疗、手术切除肿瘤组织、基因治疗和生物治疗，如施用抗体和趋化因子。因此，本发明的方法可以用于在治疗之前、期间及之后评价患者，例如评价肿瘤负荷的减少。

在一个优选实施方案中，这些方法是针对采用化学剂或生物剂的治疗。这些方法包括将从患者中获得并在化学剂或生物剂存在下保存的第一样品中至少一种本发明标志物的表达与从患者中获得并在不含上述药剂条件下保存的第二样品中该标志物的表达进行比较。该标志物在第二样品中的表达水平显著低于在第一样品中的表达水平表明该药剂在抑制患者的前列腺癌上有效。在某些实施方案中，第一和第二样品可以是患者中获得的同一样品的部分，或者是从患者中获得的合并样品的部分。

本发明另外还提供一种评估患者中前列腺癌进展的监测方法，该方法包括：

检测第一时间点的患者样品中至少一种本发明标志物的表达；

在后一时间点及时重复该表达检测步骤；和

比较第一和第二检测步骤中检测到的表达水平，由此监测患者中前列腺癌的进展。该标志物在后一时间点样品中的表达水平显著高于在第一时间点样品中的表达水平表明患者的前列腺癌已经进展，而表达水平显著较低则表明前列腺癌已经消退。

本发明还提供一种选择用于抑制患者前列腺癌的候选组合物/化合物的试验方法。该方法包括以下步骤：

从患者中获得含癌细胞的样品；

在至少一种试验组合物/化合物存在下分开保存来自患者的至少一种含有癌细胞的样品；

比较至少一种本发明标志物在各等份中的表达；和

选择一种试验组合物/化合物作为抑制前列腺癌的候选组合物/化合物，其中相对于该标志物在其他试验组合物/化合物存在下的表达水

平，所选试验组合物/化合物显著降低至少一种本发明的标志物在含该试验组合物/化合物的等份中的表达水平。

本发明另外提供一种评估化合物的致前列腺癌能力的试验方法。该方法包括以下步骤：在存在及不存在该化合物的条件下分开保存前列腺细胞等份；比较本发明标志物在各等份中的表达。在该化合物存在下保存的等份中该标志物的表达水平显著高于在不含该化合物条件下保存的等份中的表达水平表明该化合物具有致前列腺癌能力。

另外，本发明进一步提供一种抑制患者前列腺癌的方法。该方法包括以下步骤：

从患者中获得含癌细胞的样品；

在试验组合物/化合物存在下分开保存来自患者的至少一种含癌细胞的样品；

比较本发明标志物在各等份中的表达；

确定一种组合物/化合物为前列腺癌的抑制剂，其中相对于该标志物在其他组合物/化合物存在下的表达水平，该组合物/化合物显著降低该标志物在含该组合物/化合物的等份中的表达水平；和

向患者施用至少一种被确定为前列腺癌抑制剂的组合物/化合物。

根据本发明，本发明标志物在样品中的表达水平例如可以通过检测样品中下列物质的存在来评估：

相应的标志物蛋白（例如具有如下序列之一的蛋白质：SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO: 8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16、SEQ ID NO:18、SEQ ID NO:20、SEQ ID NO:22、SEQ ID NO:24、SEQ ID NO:26、SEQ ID NO:30、SEQ ID NO:32、SEQ ID NO:34、SEQ ID NO:36、SEQ ID NO:38、SEQ ID NO:40、SEQ ID NO:42 和 SEQ ID NO:44)或蛋白质片段（例如通过使用特异性结合该蛋白质或蛋白质片段的试剂，例如抗体、抗体衍生物、抗体片段或单链抗体）；

相应的标志物核酸（例如具有如下 SEQ ID NO 序列之一的核苷酸转录物（例如 SEQ ID NO: 1、SEQ ID NO:3、SEQ ID NO:5、SEQ

ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、SEQ ID NO:15、SEQ ID NO:17、SEQ ID NO:19、SEQ ID NO:21、SEQ ID NO:23、SEQ ID NO:25、SEQ ID NO:29、SEQ ID NO:31、SEQ ID NO:33、SEQ ID NO:35、SEQ ID NO:37、SEQ ID NO:39、SEQ ID NO:41 和 SEQ ID NO:43) 或其互补序列), 或核酸片段(例如通过将样品中获得的转录多核苷酸与一种基质相接触, 该基质上附着有一种或多种具有以下任一 SEQ ID NO 序列的完整序列或片段的核酸(例如 SEQ ID NO: 1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、SEQ ID NO:15、SEQ ID NO:17、SEQ ID NO:19、SEQ ID NO:21、SEQ ID NO:23、SEQ ID NO:25、SEQ ID NO:29、SEQ ID NO:31、SEQ ID NO:33、SEQ ID NO:35、SEQ ID NO:37、SEQ ID NO:39、SEQ ID NO:41 和 SEQ ID NO:43)) 或其互补序列; 或

相应标志物蛋白直接(即催化)或间接产生的代谢物。

根据本发明, 上述任何方法都可以使用多种(例如 2、3、5 或 10 种或更多)前列腺癌标志物进行或对其进行检测, 包括本发明提供的标志物与本领域公知的其他前列腺癌标志物的组合。在该方法中, 将样品中多种标志物(其中至少一种是本发明的标志物)各自的表达水平与从未患前列腺癌的对照人体中获得的同类样品中所述多种标志物各自的正常表达水平进行比较。一种或多种本发明标志物或其一定组合在样品中的表达水平相对于该标志物的相应正常或对照水平显著改变(即升高或降低, 如使用一种标志物的上述方法所说明的)表明该患者患有前列腺癌。对于所有上述方法, 优选选择标志物, 使该方法的阳性预测值至少约为 10%。

另一方面, 本发明提供了与标志物蛋白(例如具有如下任一序列的蛋白质: SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16、SEQ ID NO:18、SEQ ID NO:20、SEQ ID NO:22、SEQ ID NO:24、SEQ ID NO:26、SEQ ID NO:30、SEQ ID NO:32、SEQ ID

NO:34、SEQ ID NO:36、SEQ ID NO:38、SEQ ID NO:40、SEQ ID NO:42 和 SEQ ID NO:44) 或该蛋白的片段特异性结合的抗体、抗体衍生物或抗体片段。本发明也提供制备这些抗体、抗体衍生物和抗体片段的方法。这些方法可包括用包含标志物蛋白的全部或 10 个或 10 个以上氨基酸的片段(例如具有如下任一序列的蛋白质: SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16、SEQ ID NO:18、SEQ ID NO:20、SEQ ID NO:22、SEQ ID NO:24、SEQ ID NO:26、SEQ ID NO:30、SEQ ID NO:32、SEQ ID NO:34、SEQ ID NO:36、SEQ ID NO:38、SEQ ID NO:40、SEQ ID NO:42 和 SEQ ID NO:44) 的蛋白质或肽免疫哺乳动物, 其中该蛋白质或肽可以从细胞获得或者可以通过化学合成获得。本发明的方法也包括制备单克隆和单链抗体, 进一步包括从被免疫的哺乳动物中分离脾细胞, 将分离的脾细胞与无限增殖化细胞系相融合, 形成杂交瘤, 筛选各个杂交瘤, 选择产生与标志物蛋白或该蛋白的片段特异性结合的抗体的杂交瘤。

另一方面, 本发明涉及多种诊断和检测试剂盒。在一个实施方案中, 本发明提供一种评估患者是否患有前列腺肿瘤的试剂盒。另一方面, 该试剂盒可以用于评估患者是否具有发展前列腺癌的危险。该试剂盒包括用于评估至少一种本发明标志物的表达的试剂。再另一个实施方案提供一种试剂盒, 其可以用于评估患者是否患有攻击性前列腺肿瘤。该试剂盒包括用于评估至少一种本发明标志物的表达的试剂。在另一个实施方案中, 本发明提供一种评估化学剂或生物剂用于抑制患者前列腺癌的适用性的试剂盒。该试剂盒包括用于评估至少一种本发明标志物的表达的试剂, 也可以包括一种或多种这样的试剂。在另一个实施方案中, 本发明提供用于评估前列腺癌细胞的存在或者用于治疗前列腺癌的试剂盒。该试剂盒可以包括与标志物蛋白或蛋白质片段特异性结合的抗体、抗体衍生物或抗体片段。该试剂盒也可包括多种抗体、抗体衍生物或抗体片段, 其中所述多种抗体物质与标志物蛋白或其片段特异性结合。

在另一个实施方案中，本发明提供用于评估前列腺癌细胞的存在试剂盒，其中该试剂盒包括与至少一种标志物核酸或其片段特异性结合的至少一种核酸探针。该试剂盒还可包括多种探针，其中每种探针都与标志物核酸或其片段特异性结合。

另一方面，本发明涉及治疗前列腺癌患者或具有发展前列腺癌危险的患者的方法。该方法可以包括减少至少一种本发明标志物的表达和/或干扰其生物学功能。在一个实施方案中，该方法包括对患者提供与标志物核酸互补的反义寡核苷酸或多核苷酸或其片段。例如，可以通过给予表达标志物核酸或其片段的反义多核苷酸的载体而对患者提供反义多核苷酸。在另一个实施方案中，该方法包括对患者提供与标志物蛋白或其片段特异性结合的抗体、抗体衍生物或抗体片段。在一个优选实施方案中，抗体、抗体衍生物或抗体片段特异性结合具有 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16、SEQ ID NO:18、SEQ ID NO:20、SEQ ID NO:22、SEQ ID NO:24、SEQ ID NO:26、SEQ ID NO:30、SEQ ID NO:32、SEQ ID NO:34、SEQ ID NO:36、SEQ ID NO:38、SEQ ID NO:40、SEQ ID NO:42 和 SEQ ID NO:44 的序列的蛋白质或该蛋白质的片段。

应当理解，本发明的方法和试剂盒也可以包括已知的癌症标志物，包括已知的前列腺癌标志物。应当进一步理解，该方法和试剂盒可以用于鉴定除前列腺癌以外的癌症。

另一方面，本发明涉及编码标志物蛋白或标志物多肽例如标志物蛋白的生物活性部分的核酸分子。在一个优选实施方案中，分离的核酸分子编码具有如下氨基酸序列的标志物多肽：SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16、SEQ ID NO:18、SEQ ID NO:20、SEQ ID NO:22、SEQ ID NO:24、SEQ ID NO:26、SEQ ID NO:30、SEQ ID NO:32、SEQ ID NO:34、SEQ ID NO:36、SEQ ID NO:38、SEQ ID NO:40、SEQ ID NO:42 和 SEQ ID NO:44。在另外的

实施方案中，本发明提供具有选自以下任一所示核苷酸序列的分离的标志物核酸分子：SEQ ID NO: 1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、SEQ ID NO:15、SEQ ID NO:17、SEQ ID NO:19、SEQ ID NO:21、SEQ ID NO:23、SEQ ID NO:25、SEQ ID NO:29、SEQ ID NO:31、SEQ ID NO:33、SEQ ID NO:35、SEQ ID NO:37、SEQ ID NO:39、SEQ ID NO:41 和 SEQ ID NO:43。在另外的实施方案中，本发明提供与如下所示核苷酸序列基本上相同（例如天然发生的等位基因变体）的核酸分子：SEQ ID NO: 1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、SEQ ID NO:15、SEQ ID NO:17、SEQ ID NO:19、SEQ ID NO:21、SEQ ID NO:23、SEQ ID NO:25、SEQ ID NO:27、SEQ ID NO:29、SEQ ID NO:31、SEQ ID NO:33、SEQ ID NO:35、SEQ ID NO:37、SEQ ID NO:39、SEQ ID NO:41 和 SEQ ID NO:43。在其他实施方案中，本发明提供在如此处所述的严格杂交条件下与包含下列核苷酸序列的核酸分子杂交的核酸分子：SEQ ID NO: 1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、SEQ ID NO:15、SEQ ID NO:17、SEQ ID NO:19、SEQ ID NO:21、SEQ ID NO:23、SEQ ID NO:25、SEQ ID NO:27、SEQ ID NO:29、SEQ ID NO:31、SEQ ID NO:33、SEQ ID NO:35、SEQ ID NO:37、SEQ ID NO:39、SEQ ID NO:41 和 SEQ ID NO:43，其中所述核酸编码全长标志物蛋白或其活性片段。

在一个相关方案中，本发明进一步提供包含本文所述标志物核酸分子的核酸构建体。在某些实施方案中，本发明的核酸分子与天然或异源调节序列可操作地连接。还包括含有本发明标志物核酸分子的载体和宿主细胞，例如适于产生多肽的载体和宿主细胞。

在另一相关方案中，本发明提供适合作为用于检测标志物编码核酸的引物或杂交探针的核酸片段。

在另一相关方案中，提供与标志物编码核酸分子反义的分离的核

酸分子。

在其他实施方案中，本发明提供标志物多肽，例如具有如下氨基酸序列的标志物多肽：SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16、SEQ ID NO:18、SEQ ID NO:20、SEQ ID NO:22、SEQ ID NO:24、SEQ ID NO:26、SEQ ID NO:28、SEQ ID NO:30、SEQ ID NO:32、SEQ ID NO:34、SEQ ID NO:36、SEQ ID NO:38、SEQ ID NO:40、SEQ ID NO:42 和 SEQ ID NO:44 所示的氨基酸序列；与 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16、SEQ ID NO:18、SEQ ID NO:20、SEQ ID NO:22、SEQ ID NO:24、SEQ ID NO:26、SEQ ID NO:28、SEQ ID NO:30、SEQ ID NO:32、SEQ ID NO:34、SEQ ID NO:36、SEQ ID NO:38、SEQ ID NO:40、SEQ ID NO:42 和 SEQ ID NO:44 所示氨基酸序列基本相同的氨基酸序列；或者由具有一种核苷酸序列的核酸分子编码的氨基酸序列，该核苷酸序列在本文所述的严格杂交条件下与包含 SEQ ID NO: 1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、SEQ ID NO:15、SEQ ID NO:17、SEQ ID NO:19、SEQ ID NO:21、SEQ ID NO:23、SEQ ID NO:25、SEQ ID NO:29、SEQ ID NO:31、SEQ ID NO:33、SEQ ID NO:35、SEQ ID NO:37、SEQ ID NO:39、SEQ ID NO:41 和 SEQ ID NO:43 的核苷酸序列的核酸分子杂交，其中该核酸编码全长标志物蛋白或其活性片段。

在一个相关的方案中，本发明进一步提供包含本文所述多肽分子的蛋白质或肽构建体。在某些实施方案中，本发明的标志物多肽或片段与天然或异源非标志物多肽序列可操作连接，形成融合蛋白序列。

另一方面，本发明涉及与标志物多肽反应，或者更优选地，特异性或选择性结合的抗体及其抗原结合片段。

本发明涉及新发现的与前列腺细胞的癌性状态有关的癌症标志物。现发现，任何该标志物或该标志物组合的表达水平高于正常与患

者存在前列腺癌有关。本发明部分地基于新鉴定的标志物，与在正常（即非癌）前列腺细胞中的表达相比这些标志物在前列腺癌细胞中过量表达。一种或多种这样的标志物在前列腺细胞中的表达提高在此与组织的癌性状态有关。表 1 列出了本发明的所有标志物，与在正常（即非癌）前列腺细胞中的表达中相比，它们在前列腺癌细胞中过量表达。

提供了检测样品中前列腺癌的存在、样品中前列腺癌包括恶变前病症如发育异常的不存在、前列腺癌分期，以及利用与预防、诊断、表征和治疗患者前列腺癌有关的前列腺癌其他特性的方法。也提供了鉴定用于治疗前列腺癌的药物方法，并且也提供了治疗前列腺癌的方法。

表 1 列出了本发明的前列腺标志物。给所述标志物指定了名称（“标志物”）、如果适用，公知的基因名称（“基因名称”）、标志物编码或相对应的核苷酸转录物 cDNA 序列的序列表编号（“SEQ ID NO (nts)”）、核苷酸转录物编码的蛋白质氨基酸序列的序列表编号（“SEQ ID NO (AAs)”）和蛋白质编码序列在 cDNA 序列中的位置（“CDS”）。

表 1. 前列腺癌标志物

标志物	基因名称	SEQ ID NO (nts)	SEQ ID NO (AAs)	CDS
M245A	ABCC4: ATP-结合盒，亚家族 C (CFTR/MRP), 成员 4	1	2	116..4093
M683A	AMACR: α -甲基酰基-CoA 消旋酶, 变体 1	3	4	90..1238
M684A	AMACR: α -甲基酰基-CoA 消旋酶, 变体 2	5	6	90..1274
M686	AMACR: α -甲基酰基-CoA 消旋酶, 变体 3	7	8	66..932
M260A	ERG: v-ets 成红细胞增生病毒 E26 癌基因样(禽), 变体 1	9	10	247..1635

M742	ERG: v-ets 成红细胞增生病毒 E26 癌基因样(禽), 变体 2	11	12	247..1626
M743	ERG: v-ets 成红细胞增生病毒 E26 癌基因样(禽), 变体 3	13	14	326..1408
M744	ERG: v-ets 成红细胞增生病毒 E26 癌基因样(禽), 变体 4	15	16	326..1417
M745	ERG: v-ets 成红细胞增生病毒 E26 癌基因样(禽), 变体 5	17	18	104..1471
M261	FABP5: 脂肪酸结合蛋白 5 (牛 皮癣相关)	19	20	49..456
M746	FLJ23153: 肿瘤坏死因子- α -诱导 的脂肪相关蛋白	21	22	103..1482
M271A	GOLPH2: 高尔基体磷蛋白 2	23	24	199...1404
M279A	HSRG1: HSV-1 刺激相关基因 1	25	26	331..1974
M747	LIM: LIM 蛋白(类似于大鼠蛋白 激酶 C-结合 enigma), 变体 2	29	30	394..828
M164A	NET-6: 跨膜 4 超家族成员 tetraspan NET-6	31	32	207...821
M301A	NPY: 神经肽 Y	33	34	87...380
M308A	PLA2G2A: 磷脂酶 A2, IIA 组(血 小板, 滑液)	35	36	273..707
M318A	SMS: 精胺合酶, 亚精胺氨丙基转 移酶	37	38	102..1202
M320	SPON2: spondin 2, 胞外基质蛋 白	39	40	276..1271
M389	TACSTD1: 肿瘤相关钙信号转 导蛋白 1	41	42	179..1123
M322A	TARP: T-细胞受体 γ -链替代 (alternative) 阅读框蛋白	43	44	98..274

定义

本文中使用的下列所有术语具有与其在该章节中有关的含义。

本文中使用的冠词“一种”是指一种或一种以上（即至少一种）的

语法上的对象。例如，“一种成分”是指一种成分或一种以上的成分。

“标志物”是一种基因，其在组织或细胞中的表达水平与在正常或健康组织中的表达水平相比的改变与疾病状态如癌症有关。“标志物核酸”是本发明标志物编码的或相对应的核酸（例如 mRNA、cDNA）。这些标志物核酸包括包含任何 SEQ ID NO (nts)的完整或部分序列或该序列的互补序列的 DNA（例如 cDNA）。标志物核酸也包括包含任何 SEQ ID NO (nts)的完整或部分序列或该序列的互补序列的 RNA，其中所有胸腺嘧啶核苷残基都被尿嘧啶核苷残基取代。“标志物蛋白”是本发明标志物所编码的或相对应的蛋白质。标志物蛋白包含任何 SEQ ID NO (AAs)的完整或部分序列。术语“蛋白质”和“多肽”可互换使用。

术语“探针”是指能够选择性结合特定靶分子例如标志物编码的或对应的核苷酸转录物或蛋白质的任何分子。探针既可以由本领域技术人员合成，也可以来源于适当的生物制品。为了检测靶分子，如本文所述，探针可以特别设计为标记探针。能够用作探针的分子的例子包括但不限于 RNA、DNA、蛋白质、抗体和有机分子。

“前列腺相关的”体液是一种液体，当其在患者体内时接触或者通过前列腺细胞，或者从前列腺细胞脱落的细胞或蛋白质能够进入该液体中。示例性的前列腺相关的体液包括血液（例如全血、血清、已经除去血小板的血液）、淋巴、尿、前列腺液和精液。许多前列腺相关的体液（即通常不包括尿）可能含有前列腺细胞，特别是在前列腺细胞为癌细胞时，更特别的是在前列腺癌转移时。

“样品”或“患者样品”包括从患者中获得的细胞或前列腺相关的体液。该细胞可以从例如通过前列腺组织活检或组织切片或骨髓活检收集的前列腺组织样品中分离、鉴定或发现。或者，患者样品为体内样品。再另一种可替代的样品包括前列腺细胞或前列腺起始细胞的体外细胞或细胞系。

标志物的“正常”表达水平是标志物在未患前列腺癌的人类受试者或患者的前列腺细胞中的表达水平。

标志物的“过量表达”或“显著较高水平的表达”是指在试验样品中的表达水平高于用来评估表达的试验的标准误差，并且比该标志物在对照样品（例如来自未患该标志物相关疾病（即前列腺癌）的健康受试者的样品）中的表达水平，优选比该标志物在几种对照样品中的平均表达水平，优选至少高2倍，更优选3、4、5或10倍。

标志物的“显著较低水平的表达”是指该标志物在试验样品中的表达水平比在对照样品（例如来自未患该标志物相关疾病（即前列腺癌）的健康受试者的样品）中的表达水平，优选比该标志物在几种对照样品中的平均表达水平，至少低2倍，更优选3、4、5或10倍。

此处使用的术语“启动子/调节序列”是指与启动子/调节序列可操作连接的基因产物表达所需的核酸序列。在某些情况下，该序列可以是核心启动子序列，在另外一些情况下，该序列也可以包括增强子序列和基因产物表达所需的其他调节元件。启动子/调节序列例如可以是以组织特异性方式表达基因产物的序列。

“组成型”启动子是一种核苷酸序列，当与编码或确定基因产物的多核苷酸可操作地连接时，它使得在细胞的大多数或所有生理条件下，在活人类细胞中产生基因产物。

“诱导型”启动子是一种核苷酸序列，当其与编码或确定基因产物的多核苷酸可操作地连接时，使得基本上只有在细胞中存在对应于该启动子的诱导物时，才在活人类细胞中产生基因产物。

“组织特异性”启动子是一种核苷酸序列，当其与编码或确定基因产物的多核苷酸可操作地连接时，使得基本上只有在细胞是对应于该启动子的组织类型的细胞时，才在活人类细胞中产生基因产物。

“转录的多核苷酸”或“核苷酸转录物”是与成熟 mRNA 的全部或一部分互补或同源的多核苷酸（例如 mRNA、hnRNA、cDNA 或该 RNA 或 cDNA 的类似物），该成熟 mRNA 是通过本发明标志物的转录及（如果有的话）RNA 转录物的正常转录后加工（例如剪接）和 RNA 转录物的逆转录而获得的。

“互补”是指两条核酸链的区域之间或同一核酸链的两个区域之间

序列互补性的较宽概念。众所周知，第一核酸区的腺嘌呤残基能够与反平行于第一区的第二核酸区的残基（如果该残基是胸腺嘧啶或尿嘧啶的话）形成特定的氢键（“碱基配对”）。同样众所周知，第一核酸链的胞嘧啶残基能够与反平行于第一链的第二核酸链的残基（如果该残基是鸟嘌呤的话）碱基配对。如果当核酸的第一区与相同或不同核酸的第二区以反平行方式排列时，第一区的至少一个核苷酸残基能够与第二区的一个碱基配对，则这两个区互补。优选地，该第一区包含第一部分，第二区包含第二部分，由此当第一部分和第二部分以反平行方式排列时，第一部分中至少约 50%，优选至少约 75%，至少约 90%，或至少约 95%的核苷酸残基能够与第二部分中的核苷酸残基碱基配对。更优选地，第一部分的所有核苷酸残基都能够与第二部分的核苷酸残基碱基配对。

此处使用的“同源”是指同一核酸链的两个区之间或两个不同核苷酸链的区之间的核苷酸序列相似性。当两个区中的一个核苷酸残基位点被相同核苷酸残基占据时，则两个区在该位点处同源。如果第一区和第二区的各自至少一个核苷酸残基位点被相同残基占据，则这两个区同源。两个区之间的同源性以两个区被相同核苷酸残基占据的核苷酸残基位点的比例来表示。例如，具有核苷酸序列 5'-ATTGCC-3'的区和具有核苷酸序列 5'-TATGGC-3'的区具有 50%的同源性。优选地，第一区包含第一部分，第二区包含第二部分，从而每个部分有至少约 50%，优选至少约 75%，至少约 90%，或至少约 95%的核苷酸残基位点被相同的核苷酸残基所占据。更优选地，每个部分的所有核苷酸残基位点都被相同的核苷酸残基所占据。

如果一种分子与基质共价或非共价结合，该基质能够用流体（例如标准枸橼酸盐水，pH 7.4）漂洗，而没有相当一部分的分子从基质上脱离下来，则称该分子被“固定”或“附着”于该基质上。

此处使用的“天然存在的”核酸分子是指具有在自然界发现的生物中存在的核苷酸序列的 RNA 或 DNA 分子。

如果癌症的至少一种症状得到缓解、终止、减慢或阻止，则称该

癌症得到“抑制”。在本文中，如果前列腺癌的复发或转移被减少、减慢、延迟或防止，也称前列腺癌得到“抑制”。

试剂盒是包含至少一种特异性检测本发明标志物表达的试剂例如探针的任何产品（例如包装或容器）。试剂盒可以作为一个用于实施本发明的方法的单元来推广、分发或销售。

“本发明的蛋白质”包括标志物蛋白及其片段；变异标志物蛋白及其片段；包含标志物蛋白或变异标志物蛋白的至少 15 个氨基酸片段的肽和多肽；以及包含标志物蛋白或变异标志物蛋白，或标志物蛋白或变异标志物蛋白的至少 15 个氨基酸片段的融合蛋白。

除非本文中另外指明，术语“抗体”广泛地包括天然存在的抗体形式（例如 IgG、IgA、IgM、IgE）和重组抗体，如单链抗体、嵌合和人源化抗体和多特异性抗体，以及所有上述抗体的片段和衍生物，这些片段和衍生物具有至少一个抗原结合位点。抗体衍生物可以包括与抗体偶联的蛋白质或化学部分。

本发明提供用于评估前列腺细胞（例如从人体获得的细胞，培养的人类细胞，存档或保藏的人类细胞，和体内细胞）的癌性状态以及治疗前列腺癌患者的组合物、试剂盒和方法。

本发明的组合物、试剂盒和方法尤其具有如下用途：

评估患者是否患有前列腺癌；

评估人类患者中前列腺癌的转移能力；

制备用于治疗前列腺癌和/或评估患者是否患有前列腺癌的抗体、抗体片段或抗体衍生物；

评估前列腺癌细胞的存在；

评估一种或多种试验化合物抑制患者的前列腺癌的效果；

评估一种治疗方法抑制患者的前列腺癌的效果；

监测患者中前列腺癌的进展；

筛选用于抑制患者的前列腺癌的组合物或治疗方法；

治疗前列腺癌患者；

抑制患者的前列腺癌；

评估试验化合物的致前列腺癌能力；以及

预防具有发展前列腺癌危险的患者中前列腺癌的发病。

本发明因此包括一种评估患者是否患有前列腺癌的方法，包括评估患者是否具有转移前的前列腺癌。该方法将本发明标志物在患者样品中的表达水平与该标志物在对照例如非前列腺癌样品中的正常表达水平进行比较。该标志物在患者样品中的表达水平显著高于正常水平表明该患者患有前列腺癌。在一个优选实施方案中，该标志物选自表1中的标志物。

本发明也提供了基因递送载体、宿主细胞和组合物（均在此描述），其含有包含 SEQ ID NO (nts)任一序列或这些序列的互补序列的全部或其15个或更多核苷酸片段的核酸，以及包含 SEQ ID NO (AAs)任一序列的全部或10个或更多氨基酸片段的多肽。

如本文所述，患者的前列腺癌与一种或多种本发明标志物的表达水平升高有关。尽管如上所述，有些这样的表达水平改变是前列腺癌的存在引起的，但是另外一些这样的改变诱导、保持和促进了前列腺癌细胞的癌性状态。因此，特征在于一种或多种本发明标志物表达水平升高的前列腺癌能够通过减少和/或干扰该标志物的表达和/或这些标志物编码的蛋白质的功能而得到抑制。

本发明标志物的表达能够以本领域公知的许多方法进行抑制。例如，能够向前列腺癌细胞提供反义寡核苷酸，从而抑制标志物的转录、翻译或转录及翻译。或者，可以向该细胞提供编码特异性结合标志物蛋白的抗体、抗体衍生物或抗体片段并且与适当启动子/调节区可操作连接的多核苷酸，从而产生将抑制该蛋白质的功能或活性的细胞内抗体。用特异性结合标志物蛋白的抗体、抗体衍生物或抗体片段处理前列腺癌细胞也可以抑制标志物的表达和/或功能。使用此处所述的方法，能够筛选多种分子，特别包括小到足以能够穿过细胞膜的分子，以鉴定抑制标志物表达或抑制标志物蛋白功能的分子。可以向患者提供这样鉴定的化合物，从而抑制患者的前列腺癌细胞。

本发明的任何标志物或标志物组合，以及任何已知的标志物与本

发明的标志物组合，都可以在本发明的组合物、试剂盒和方法中使用。通常，优选使用在前列腺癌细胞中的表达水平与在正常前列腺细胞中的表达水平相差尽可能大的标志物。尽管这种差异可以小到用于评估该标志物表达的方法的检测限度，但是优选该差异至少高于该评估方法的标准误差，优选与同一标志物在正常前列腺组织中的表达水平相差至少 2、3、4、5、6、7、8、9、10、15、20、25、100、500、1000 倍或更多。

已经公认某些标志物蛋白从前列腺细胞（即正常细胞和癌细胞之一或两者）分泌到细胞周围的细胞外间隙中。这些标志物优选用于本发明的组合物、试剂盒和方法的某些实施方案，其原因如下事实：在比组织活检样品更易从人类患者采集的前列腺相关体液样品中能够检测到这些标志物蛋白。另外，检测标志物蛋白的优选体内技术包括向受试者中导入标记的抗该蛋白质的抗体。例如，抗体可以用一种放射性标记物来标记，该放射性标记物在受试者中的存在和定位可以用标准成像技术来检测。

对于本领域技术人员而言，确定任一特定标志物蛋白是否是分泌的蛋白质是一件简单的事情。为了进行这种确定，例如在哺乳动物细胞中，优选在人类前列腺细胞系中表达该标志物蛋白，采集细胞外液，评估细胞外液中该蛋白质的存在与否（例如利用特异性结合该蛋白质的标记抗体）。

以下是能够用来检测蛋白质分泌的方法的一个例子。大约 8×10^5 293T 细胞在含有培养基(Dulbecco 改进 Eagle 培养基{DMEM}，补充 10%胎牛血清)的孔中在 5% (v/v) CO₂，95%空气气氛下 37°C 温育，至大约 60-70%汇合。然后用标准转染混合物转染细胞，该转染混合物每孔含有 2 微克含编码该蛋白质的表达载体的 DNA 和 10 微升 LipofectAMINETM(GIBCO/BRL 目录号 18342-012)。转染混合物保持大约 5 小时，然后更换为新鲜培养基，保持在空气气氛中。各孔用不含甲硫氨酸或半胱氨酸的 DMEM (DMEM-MC; ICN 目录号 16-424-54) 轻轻漂洗两次。向每孔加入约 1 毫升 DMEM-MC 和约 50 微居里

Trans-³⁵S™ 试剂(ICN 目录号 51006)。各孔保持在上述 5% CO₂ 气氛下,并在选择的时间里 37°C 温育。温育后,除去 150 微升条件培养基,离心除去漂浮的细胞和碎片。上清液中存在该蛋白质表明该蛋白质被分泌。

应当认识到在本发明的方法中可以使用含有前列腺细胞的患者样品。在这些实施方案中,可以通过评估前列腺细胞样品中,例如从患者获得的前列腺组织活检中该标志物的量(例如绝对量或浓度)来评价标志物的表达水平。在评价样品中该标志物的量之前当然可以对该细胞样品进行多种众所周知的采集后制备和贮存技术(例如核酸和/或蛋白质提取、固定、贮存、冷冻、超滤、浓缩、蒸发、离心等)。同样,前列腺组织活检样品也可以进行采集后制备和贮存技术,例如固定。

本发明的组合物、试剂盒和方法可以用于检测其至少一部分展示于表达它的细胞表面上的标志物蛋白的表达。对于本领域技术人员而言,确定标志物蛋白或其部分是否暴露于细胞表面是一件容易的事情。例如,免疫学方法可以用来检测全细胞上的这些蛋白质,或者众所周知的基于计算机的序列分析方法可以用来预测至少一个胞外域的存在(即,既包括分泌的蛋白质又包括含有至少一个细胞表面域的蛋白质)。不必裂解细胞即可检测其至少一部分展示于表达它的细胞表面上的标志物蛋白的表达(例如使用特异性结合蛋白质的细胞表面域的标记抗体)。

本发明标志物的表达可以采用多种众所周知的用于检测转录核酸或蛋白质表达的方法中的任一种来评估。该方法的非限制性例子包括用于检测分泌的、细胞表面、细胞质或核蛋白质的免疫学方法,蛋白质纯化方法,蛋白质功能或活性测定,核酸杂交方法,核酸逆转录方法,和核酸扩增方法。

在一个优选实施方案中,标志物表达的评估使用特异性结合该标志物蛋白或其片段-包括经历全部或部分正常翻译后修饰的标志物蛋白-的抗体(例如放射性标记的、生色团标记的、荧光团标记的或酶标

记的抗体)、抗体衍生物(例如与底物或与蛋白质-配体对{例如生物素-链霉亲和素}的蛋白质或配体偶联的抗体)或抗体片段(例如单链抗体,分离的抗体超变区等)。

在另一个优选实施方案中,标志物的表达如下评估:从患者样品中的细胞制备 mRNA/cDNA(即转录的多核苷酸),并且将该 mRNA/cDNA 与互补于标志物核酸的参照多核苷酸或其片段进行杂交。任选地,cDNA 在与参照多核苷酸杂交之前可以用多种聚合酶链反应方法中的任何一种进行扩增;优选不进行扩增。一种或多种标志物的表达也可以用定量 PCR 进行检测,以估计该标志物的表达水平。或者,也可以使用检测本发明标志物的突变或变异(例如单核苷酸多态性、缺失等)的任一已知方法来检测患者中标志物的存在。

在一个有关实施方案中,从样品中获得的转录多核苷酸混合物与一个基质接触,该基质上固定了与标志物核酸的至少一部分(例如至少 7、10、15、20、25、30、40、50、100、500 或更多核苷酸残基)互补或同源的多核苷酸。如果互补或同源的多核苷酸在基质上可差别地检测到(例如可用不同生色团或荧光团检测,或者固定于选择的不同位置),则多种标志物的表达水平能够用一种基质(例如固定于选定位置的多核苷酸的“基因芯片”微阵列)同时评估。当使用的评估标志物表达的方法涉及一种核酸与另一种核酸杂交时,优选在严格杂交条件下进行杂交。

因为本发明的组合物、试剂盒和方法依赖于检测一种或多种本发明标志物的表达水平的差异,因此优选标志物的表达水平显著高于用来评估在正常前列腺细胞和癌性前列腺细胞至少之一中之表达的方法的最低检测限值。

应当理解,通过使用一种或多种本发明的标志物常规筛选其它患者样品,将认识到某些标志物在多种类型的癌症中过量表达,包括特定前列腺癌,以及其他癌症。例如,将会证实某些本发明的标志物在大多数(即 50%或以上)或基本上全部(即 80%或以上)前列腺癌中过量表达。此外也证实,某些本发明的标志物与各个分期的前列腺癌

有关。TNM 分期法根据前列腺内原发肿瘤的大小和位置将原发肿瘤 (T) 分为四期中的一个 (以及这些分类中的另外的亚分期)。T1 表示通过数字直肠检查检测不到的显微肿瘤。T2NO 是指在数字直肠检查中可触及但是包括在前列腺囊内的肿瘤 (局部疾病)。在所有形式的 T3 期疾病中, 肿瘤已经通过前列腺囊扩散到周围的结缔组织或精囊。T4 是指已经离开前列腺并且可在盆腔区中发现的肿瘤。N 期涉及原发性肿瘤是否已经扩散到局部淋巴结 (盆腔淋巴结)。M 期涉及肿瘤细胞是否已经转移到远端部位。

另外, 通过对较大数量的患者样品评估本发明标志物的表达的改变, 并且将获得样品的各个患者的结果相关联, 也证实本发明某些标志物的表达改变与恶性肿瘤高度相关。本发明的组合物、试剂盒和方法因此可用于表征患者前列腺癌的分期、分级、组织学类型和良性/恶性性质中的一种或多种。

当本发明的组合物、试剂盒和方法用于表征患者前列腺癌的分期、分级、组织学类型和良性/恶性性质中的一种或多种时, 优选选择本发明的标志物或一组标志物, 以便在至少约 20%, 优选至少约 40%、60%或 80%及更多的, 优选在基本上全部的相应分期、分级、组织学类型或良性/恶性的前列腺癌患者中获得阳性结果。优选选择本发明的标志物或一组标志物, 以便在总人群中获得超过约 10%的阳性预测值(PPV) (更优选地同时有高于 80%的测定特异性)。

当在本发明的组合物、试剂盒和方法中使用多种本发明的标志物时, 可以将患者样品中每种标志物的表达水平与在单一反应混合物中 (即对于每种标志物使用试剂, 例如不同的荧光探针) 或在对应于一种或多种标志物的各反应混合物中的相同类型非癌样品中所述多种标志物之每一种的正常表达水平进行比较。在一个实施方案中, 相对于相应的正常水平, 多种标志物中一种以上的标志物在样品中的表达水平显著增加表明该患者患有前列腺癌。当使用多种标志物时, 优选使用 2、3、4、5、8、10、12、15、20、30 或 50 种或更多的标志物, 其中优选较少的标志物。

为了使本发明的组合物、试剂盒和方法的灵敏度最大化(即通过患者样品中非前列腺来源的细胞引起的干扰),优选此处使用的本发明的标志物是具有受限的组织分布的标志物,例如在非前列腺组织中通常不表达的标志物。

只知道有少量标志物与前列腺癌相关(例如前列腺特异性抗原(PSA, KLK3)、前列腺特异性膜抗原(PSMA FOLH1)、前列腺酸性磷酸酶(PAP, ACPP)、前列腺癌抗原 3 (PCA3, DD3)、前列腺癌肿瘤抗原(PCTA-1 LGALS8)、前列腺干细胞抗原(PSCA)、腺激肽释放酶-1(hGK-1, KLK2)、prostase (KLK4, PRSS17, KLK-L1)、前列腺特异性G-蛋白偶联受体(PSGR)和前列腺六跨膜上皮抗原(STEAP, PRSS24))。在一个实施方案中,这些标志物可以与一组标志物中的一种或多种本发明的标志物一起使用。众所周知,某些基因类型,例如癌基因、肿瘤抑制基因、生长因子样基因、蛋白酶样基因和蛋白激酶样基因,通常与各种类型癌症的发展有关。因此,在本发明的标志物中,优选使用相应于与已知癌基因和肿瘤抑制基因所编码的已知蛋白质类似的蛋白质的标志物,和相应于类似生长因子、蛋白酶和蛋白激酶的蛋白质的标志物。

认为本发明的组合物、试剂盒和方法对发展前列腺癌危险高的患者以及他们的医学顾问特别有用。被认为发展前列腺癌危险高的患者包括,例如,具有前列腺癌家族史的患者,被确定为具有突变癌基因(即至少一个等位基因)的患者,和高龄患者(即大约 50 或 60 岁以上的男性)。

一种标志物在正常(即非癌)人类前列腺组织中的表达水平可以用多种方法评估。在一个实施方案中,这种正常表达水平的评估方法包括:评估该标志物在看来是非癌细胞的前列腺细胞的一部分中的表达水平,将该正常表达水平与在怀疑为癌细胞的前列腺细胞的一部分中的表达水平进行比较。另外并且特别地,由于常规进行本文所述的方法可以获得进一步的信息,本发明标志物的正常表达可以使用人群平均值。在其他实施方案中,标志物的“正常”表达水平可以如下确

定：评估从非癌患者获得的患者样品中、在怀疑患者前列腺癌发病之前从患者中获得的患者样品中、存档的患者样品等中该标志物的表达。

本发明包括用于评估样品（例如存档的组织样品或从患者中获得的样品）中前列腺癌细胞的存在组合物、试剂盒和方法。这些组合物、试剂盒和方法基本与上述相同，除了必要时为了使用患者样品以外的样品而对该组合物、试剂盒和方法作适应性改变。例如，当使用的样品是石蜡处理的存档的人类组织样品时，调节本发明组合物、本发明试剂盒中化合物的比例，或调整用于评估样品中标志物表达水平的方法可能是必要的。这些方法在本领域中公知，并且在本领域技术人员的技能之内。

本发明包括用于评估（例如在样品如患者样品中）前列腺癌细胞的存在试剂盒。该试剂盒包括多种试剂，其中每一种试剂都能够特异性结合一种标志物核酸或蛋白。合适的用于结合标志物蛋白的试剂包括抗体、抗体衍生物、抗体片段等。合适的用于结合标志物核酸（例如基因组 DNA、mRNA、剪接的 mRNA、cDNA 等）的试剂包括互补核酸。例如，核酸试剂可包括固定于基质上的寡核苷酸（标记或非标记的）、未与基质结合的标记寡核苷酸、PCR 引物对、分子信标探针等。

本发明的试剂盒可以任选地包括用于进行本发明方法的其他成分。例如，该试剂盒可包括适用于互补核酸退火或适用于结合抗体和与之特异性结合的蛋白质的流体（例如 SSC 缓冲液），一个或多个样品室，描述本发明方法的操作的说明材料，正常前列腺细胞样品，前列腺癌细胞样品，等等。

本发明也包括一种制备分离的杂交瘤的方法，该杂交瘤产生用于评估患者是否患有前列腺癌的抗体。在该方法中，合成或分离包含标志物蛋白的全部或片段的蛋白质或肽（例如通过从表达细胞中纯化，或者通过应用已知方法在体内或体外转录和翻译编码该蛋白质或肽的核酸）。用该蛋白质或肽免疫脊椎动物，优选哺乳动物，如小鼠、大

鼠、兔或绵羊。该脊椎动物任选地(且优选地)可以用该蛋白质或肽进行至少另一次免疫,以使该脊椎动物表现出对该蛋白质或肽的加强的免疫应答。使用本领域公知的任一方法,从免疫的脊椎动物中分离脾细胞,并且将其与无限增殖化细胞系融合,形成杂交瘤。然后用标准方法筛选这样形成的杂交瘤,从而鉴定出一种或多种产生特异性结合该标志物蛋白或其片段的抗体的杂交瘤。本发明也包括利用该方法制备的杂交瘤和利用该杂交瘤制备的抗体。

本发明也包括一种评估试验化合物抑制前列腺癌细胞的效果的方法。如上所述,本发明标志物的表达水平的差异与前列腺细胞的癌性状态有关。虽然认为本发明某些标志物的表达水平的改变可能是前列腺细胞的癌性状态引起的,但是同样认为,本发明其他标志物的表达水平的改变诱导、保持和促进了这些细胞的癌性状态。因此,抑制患者中前列腺癌的化合物将使一种或多种本发明标志物的表达水平变化为更接近于该标志物的正常表达水平(即该标志物在非癌前列腺细胞中的表达水平)的水平。

该方法因此包括将在试验化合物存在下保存的第一前列腺细胞样品中一种标志物的表达与在不含该试验化合物条件下保存的第二前列腺细胞样品中该标志物的表达进行比较。在试验化合物存在下本发明标志物的表达显著降低表明该试验化合物抑制前列腺癌。前列腺细胞样品例如可以是从患者中获得的一种正常前列腺细胞样品的等份,从患者中获得的正常前列腺细胞的合并样品,正常前列腺细胞系细胞,从患者中获得的一种前列腺癌细胞样品的等份,从患者中获得的前列腺癌细胞的合并样品,前列腺癌细胞系细胞,等等。在一个实施方案中,样品是从患者中获得的前列腺癌细胞,并且检测已知有效抑制各种前列腺癌的多种化合物,以鉴定最有可能抑制患者前列腺癌的化合物。

该方法同样可以用于评估一种疗法抑制患者前列腺癌的疗效。在该方法中,评估一种或多种本发明标志物在一对样品(一个经过治疗,另一个没有经过治疗)中的表达水平。如同评估试验化合物的效

果的方法一样，如果这种疗法诱导明显较低的本发明标志物的表达水平，则这种疗法对于抑制前列腺癌是有效的。如上所述，如果在该方法中使用来自选定患者的样品，则可以在体外评估备选疗法，以选择最可能有效抑制患者前列腺癌的疗法。

如上所述，人前列腺细胞的癌性状态与本发明标志物的表达水平的变化有关。本发明包括用于评估试验化合物致人前列腺细胞癌的能力的方法。该方法包括在存在及不存在试验化合物的条件下分开保存人前列腺细胞等份。比较本发明标志物在各等份中的表达。（与不存在试验化合物条件下保存的等份相比）在试验化合物存在下保存的等份中本发明标志物的表达水平明显更高表明该试验化合物具有致人前列腺细胞癌的能力。各种试验化合物的相对致癌能力的评估可以通过比较有关标志物表达水平提高或抑制的程度，比较表达水平提高或抑制的标志物的数量，或者以上两者都比较。

本发明的各个方面在下面的部分中进一步详细描述。

分离的核酸分子

本发明的一个方案涉及分离的核酸分子，包括编码标志物蛋白或其一部分的核酸。本发明的分离的核酸也包括足以用作杂交探针来鉴定标志物核酸分子和标志物核酸分子的片段的核酸分子，例如适合作为标志物核酸分子扩增或突变的 PCR 引物使用的那些。此处使用的术语“核酸分子”包括 DNA 分子(例如 cDNA 或基因组 DNA)和 RNA 分子(例如 mRNA)和用核苷酸类似物产生的 DNA 或 RNA 的类似物。核酸分子可以是单链或双链的，但优选是双链 DNA。

“分离的”核酸分子是从该核酸分子天然来源中存在的其他核酸分子中分离出来的核酸分子。在一个实施方案中，“分离的”核酸分子不含在作为该核酸来源的生物的基因组 DNA 中天然位于该核酸侧翼的序列(优选蛋白质编码序列)(即位于该核酸 5'和 3'末端的序列)。例如，在各种实施方案中，分离的核酸分子可含有少于约 5 kB、4 kB、3 kB、2 kB、1 kB、0.5 kB 或 0.1 kB 的在作为该核酸来源的细胞的基

基因组 DNA 中天然位于该核酸分子之侧的核苷酸序列。在另外一个实施方案中，“分离的”核酸分子，如 cDNA 分子，基本不含其他细胞材料，或者当利用重组技术产生时基本不含培养基，或者当化学合成时基本不含化学前体或其他化学物质。基本不含细胞材料的核酸分子包括含有少于约 30%、20%、10%或 5%的异源核酸（在此也被称为“污染核酸”）的制剂。

本发明的核酸分子可以利用标准分子生物学技术和本文所述数据库中记录的序列信息来分离。利用这些核酸序列的全部或一部分，可以利用标准杂交和克隆技术（例如，Sambrook 等人，ed., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989 所述的方法）分离本发明的核酸分子。

本发明的核酸分子可以用 cDNA、mRNA 或基因组 DNA 作为模板并且使用合适的寡核苷酸引物按照标准 PCR 扩增技术来扩增。可将这样扩增的核酸克隆到合适的载体中，并且通过 DNA 序列分析进行表征。此外，对应于本发明核酸分子的全部或一部分的核苷酸也可以利用标准合成技术制备，例如用自动 DNA 合成仪制备。

在另一优选实施方案中，本发明的分离的核酸分子包括一种核酸分子，其具有与标志物核酸的核苷酸序列互补或与编码标志物蛋白的核酸的核苷酸序列互补的核苷酸序列。与特定核苷酸序列互补的核酸分子是与特定核苷酸序列完全互补，因此能够与特定核苷酸序列杂交从而形成稳定双链体的核酸分子。

另外，本发明的核酸分子可以只含有核酸序列的一部分，其中全长核酸序列包含一种标志物核酸或编码一种标志物蛋白。该核酸例如可以用作探针或引物。探针/引物一般使用一种或多种基本上纯化的寡核苷酸。该寡核苷酸一般包含一个核苷酸序列区，该区在严格条件下与本发明核酸的至少约 7 个，优选约 15 个，更优选约 25、50、75、100、125、150、175、200、250、300、350 或 400 个或更多的连续核苷酸杂交。

基于本发明核酸分子序列的探针能够用于检测对应于一种或多种本发明标志物的转录物或基因组序列。探针含有与之连接的标记基团，例如放射性同位素、荧光化合物、酶或酶辅因子。这些探针可以用作鉴定错误表达该蛋白质的细胞或组织的诊断检测试剂盒的一部分，例如通过测定在来自受试者的细胞样品中编码该蛋白质的核酸分子的水平，例如检测 mRNA 水平或确定编码该蛋白质的基因是否已经突变或缺失。

本发明进一步包括由于遗传密码简并性而与编码标志物蛋白（即具有 SEQ ID NO (AAs) 的序列的蛋白质）的核酸的核苷酸序列不同，因而编码相同蛋白质的核酸分子。

本领域技术人员应当理解，导致氨基酸序列变化的 DNA 序列多态性可能存在于一个群体（例如人群）中。由于自然等位基因变异，这种遗传多态性可能存在于一个群体内的个体之间。等位基因是在特定基因座处可替代地存在的一组基因中的一个。另外，应当理解，也可能存在影响 RNA 表达水平的 DNA 多态性，其可以影响该基因的总表达水平（例如通过影响调节或降解）。

此处使用的短语“等位基因变体”是指在特定基因座处发生的核苷酸序列，或者是指该核苷酸序列编码的多肽。

此处使用的术语“基因”和“重组基因”是指包含编码对应于本发明标志物的多肽的开放阅读框的核酸分子。这些自然等位基因变异一般可导致特定基因的核苷酸序列中 1-5% 的变化。替代的等位基因可以通过对大量不同个体中的目标基因进行测序来鉴定。使用杂交探针鉴定多个个体中的同一基因座能够容易地进行这种鉴定。自然等位基因变异导致的并且未改变功能活性的任意及所有的这些核苷酸变异和产生的氨基酸多态性或变异，都在本发明的范围之内。

在另一个实施方案中，本发明的分离的核酸分子长度至少为 7、15、20、25、30、40、60、80、100、150、200、250、300、350、400、450、550、650、700、800、900、1000、1200、1400、1600、1800、2000、2200、2400、2600、2800、3000、3500、4000、4500

个或更多的核苷酸，并且在严格条件下与标志物核酸或与编码标志物蛋白的核酸杂交。此处使用的术语“在严格条件下杂交”是指这样的杂交和洗涤条件，在该条件下彼此至少 60% (65%、70%，优选 75%) 相同的核苷酸序列一般保持彼此杂交。这些严格条件为本领域技术人员所公知，并且可见于 *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989) 的第 6.3.1-6.3.6 节。严格杂交条件的一个优选的非限制性实例是在大约 45°C 下在 6×氯化钠/柠檬酸钠(SSC)中杂交，然后在 50-65°C 下用 0.2× SSC, 0.1% SDS 洗涤一次或数次。

除了群体中可能存在的本发明核酸分子的天然存在的等位基因变体之外，本领域技术人员应当进一步理解，可以通过突变引入序列改变，从而使其编码的蛋白质的氨基酸序列改变，而不改变其编码的蛋白质的生物活性。例如，可以进行核苷酸置换，在“非必需”氨基酸残基处进行氨基酸置换。“非必需”氨基酸残基是可能与野生型序列相比发生改变但不改变生物活性的残基，而“必需”氨基酸残基是生物活性所必需的。例如，在不同物种的同源物之间不保守或仅仅半保守的氨基酸残基对于活性可能是非必需的，因此可能是改变的目标。另外，在不同物种（例如鼠和人类）的同源物之间保守的氨基酸残基对于活性可能是必需的，因此不能作为改变的目标。

因此，本发明的另一方案涉及编码变异标志物蛋白的核酸分子，该变异标志物蛋白含有对于活性而言非必需的氨基酸残基的改变。该变异标志物蛋白在氨基酸序列上不同于天然存在的标志物蛋白，但仍保留生物活性。在一个实施方案中，这种变异标志物蛋白具有与标志物蛋白的氨基酸序列至少大约 40%、50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98% 或 99% 相同的氨基酸序列。

编码变异标志物蛋白的分离的核酸分子可以如下产生：向标志物核酸的核苷酸序列中引入一个或多个核苷酸置换、添加或缺失，从而向其编码的蛋白质中引入一个或多个氨基酸残基置换、添加或缺失。可以利用标准技术如定点诱变和 PCR 介导的诱变来引入突变。优选

地，在一个或多个预测的非必需氨基酸残基处进行保守氨基酸置换。“保守氨基酸置换”是氨基酸残基被置换为具有相似侧链的另一氨基酸残基的置换。具有相似侧链的氨基酸残基的家族在本领域中已经定义。这些家族包括具有碱性侧链（例如赖氨酸、精氨酸、组氨酸）、酸性侧链（例如天冬氨酸、谷氨酸）、不带电荷的极性侧链（例如甘氨酸、天冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、酪氨酸、半胱氨酸）、非极性侧链（例如丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、色氨酸）、 β -分支侧链（例如苏氨酸、缬氨酸、异亮氨酸）和芳香族侧链（例如酪氨酸、苯丙氨酸、色氨酸、组氨酸）的氨基酸。或者，也可以通过例如饱和诱变，沿编码序列的全部或一部分随机引入突变，获得的突变体可以根据生物活性进行筛选，以鉴定出保留活性的突变体。诱变后，可以重组表达编码的蛋白质，并且可以测定该蛋白质的活性。

本发明包括反义核酸分子，即与本发明的有义核酸互补的分子，例如与双链标志物 cDNA 分子的编码链互补，或者与标志物 mRNA 序列互补的分子。因此，本发明的反义核酸可以与本发明的有义核酸形成氢键（即退火）。反义核酸可以互补于整个编码链，或者只互补于其中一部分，例如蛋白质编码区（或开放阅读框）的全部或一部分。反义核酸分子也可以对于编码标志物蛋白的核苷酸序列编码链的非编码区的全部或一部分而言是反义的。非编码区（“5'和 3'非翻译区”）是位于编码区之侧的 5'和 3'序列，不翻译为氨基酸。

反义寡核苷酸的长度可以是，例如大约 5、10、15、20、25、30、35、40、45 或 50 个或更多个核苷酸。本发明的反义核酸可以利用化学合成及酶连接反应采用本领域公知的程序来构建。例如，反义核酸（例如反义寡核苷酸）可以利用天然存在的核苷酸或各种修饰的核苷酸化学合成，这些修饰核苷酸被设计为提高该分子的生物学稳定性或提高反义与有义核酸之间形成的双链体的物理稳定性，例如可以使用硫代磷酸酯衍生物和吡啶取代的核苷酸。能够用于产生反义核酸的修饰核苷酸的例子包括 5-氟尿嘧啶、5-溴尿嘧啶、5-氯尿嘧啶、5-

碘尿嘧啶、次黄嘌呤、黄嘌呤、4-乙酰胞嘧啶、5-(羧基羟甲基)尿嘧啶、5-羧甲基氨甲基-2-硫尿核苷, 5-羧甲基氨甲基尿嘧啶、二氢尿嘧啶、 β -D-半乳糖基 Q 核苷、肌苷、N6-异戊烯腺嘌呤、1-甲基鸟嘌呤、1-甲基肌苷、2,2-二甲基鸟嘌呤、2-甲基腺嘌呤、2-甲基鸟嘌呤、3-甲基胞嘧啶、5-甲基胞嘧啶、N6-腺嘌呤、7-甲基鸟嘌呤、5-甲基氨甲基尿嘧啶、5-甲氧基氨甲基-2-硫尿嘧啶、 β -D-甘露糖基 Q 核苷、5'-甲氧基羧甲基尿嘧啶、5-甲氧基尿嘧啶、2-甲硫基-N6-异戊烯腺嘌呤、尿嘧啶-5-氧乙酸(v)、怀丁苷(wybutoxosine)、假尿嘧啶、Q 核苷、2-硫胞嘧啶、5-甲基-2-硫尿嘧啶、2-硫尿嘧啶、4-硫尿嘧啶、5-甲基尿嘧啶、尿嘧啶-5-氧乙酸甲酯、尿嘧啶-5-氧乙酸(v)、5-甲基-2-硫尿嘧啶、3-(3-氨基-3-N-2-羧丙基)尿嘧啶、(acp3)w 和 2,6-二氨基嘌呤。或者, 反义核酸也可以利用表达载体以生物学方法产生, 该表达载体中已经以反义方向亚克隆了核酸(即由插入核酸转录的 RNA 与靶核酸为反义方向, 在以下部分中进一步描述)。

本发明的反义核酸分子一般对受试者施用, 或者在原位产生, 使其与编码标志物蛋白的细胞 mRNA 和/或基因组 DNA 杂交或结合, 从而抑制该标志物的表达, 例如通过抑制转录和/或翻译来实现。杂交可以通过常规核苷酸互补性来形成稳定的双链体, 或者例如在结合 DNA 双链体的反义核酸分子的情况下, 通过双螺旋大沟中的特异性相互作用来实现。本发明反义核酸分子的给药途径的例子包括在组织部位直接注射, 或者将反义核酸输注到前列腺相关体液中。或者, 也可以为了靶向于选定的细胞而修饰反义核酸分子, 然后全身给药。例如, 对于全身给药, 可以修饰反义分子, 使其特异性结合在所选细胞表面表达的受体或抗原, 例如, 通过将反义核酸分子与结合细胞表面受体或抗原的肽或抗体相连接。也可以利用此处所述的载体将反义核酸分子导入细胞中。为了达到反义分子的足够的细胞内浓度, 优选反义核酸分子被置于 pol II 或 pol III 强启动子控制下的载体构建体。

本发明的反义核酸分子可以是 α -异头核酸分子。 α -异头核酸分子与互补 RNA 形成特殊的双链杂合体, 其中与通常的 α -单元不同, 两

条链彼此平行(Gaultier 等人, 1987, *Nucleic Acids Res.* 15:6625-6641)。反义核酸分子也可以包含 2'-O-甲基核糖核苷酸(Inoue 等人, 1987, *Nucleic Acids Res.* 15:6131-6148)或嵌合 RNA-DNA 类似物(Inoue 等人, 1987, *FEBS Lett.* 215:327-330)。

本发明也包括核酶。核酶是具有核糖核酸酶活性的催化性 RNA 分子, 它能够切割与其具有互补区的单链核酸, 如 mRNA。因此, 核酶(例如 Haselhoff 和 Gerlach, 1988, *Nature* 334:585-591 所述的锤头状核酶)能够用于催化性切割 mRNA 转录物, 从而抑制 mRNA 编码的蛋白质的翻译。对编码标志物蛋白的核酸分子具有特异性的核酶可以根据对应于该标志物的 cDNA 的核苷酸序列来设计。例如, 可以构建四膜虫 (*Tetrahymena*) L-19 IVS RNA 的衍生物, 其中活性部位的核苷酸序列与将要切割的核苷酸序列互补(参见 Cech 等人, 美国专利 No. 4,987,071; 和 Cech 等人, 美国专利 No. 5,116,742)。或者, 能够利用编码本发明多肽的 mRNA 从 RNA 分子库中选择具有特异性核糖核酸酶活性的催化性 RNA(参见, 例如 Bartel 和 Szostak, 1993, *Science* 261:1411-1418)。

本发明也包括形成三股螺旋结构的核酸分子。例如, 通过靶向至与编码标志物核酸或蛋白的基因的调节区(例如启动子和/或增强子)互补的核苷酸序列, 形成阻止基因在靶细胞中转录的三股螺旋结构, 能够抑制本发明标志物的表达。一般地参见 Helene (1991) *Anticancer Drug Des.* 6(6):569-84; Helene (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; 和 Maher (1992) *Bioassays* 14(12):807-15。

在各个实施方案中, 可以在碱基部分、糖部分或磷酸骨架上修饰本发明的核酸分子, 以改善例如分子的稳定性、杂交或溶解性。例如, 可以修饰核酸的脱氧核糖磷酸骨架, 产生肽核酸(参见 Hyrup 等人, 1996, *Bioorganic & Medicinal Chemistry* 4(1): 5-23)。此处使用的术语“肽核酸”或“PNA”是指脱氧核糖磷酸骨架被假肽骨架所取代并且只保留了 4 种天然核碱基的核酸模拟物, 例如 DNA 模拟物。已经证实, PNA 的中性骨架允许 DNA 和 RNA 在低离子强度条件下特异性

杂交。PNA 寡聚体的合成可以利用 Hyrup 等人 (1996), 同上; Perry-O'Keefe 等人 (1996) *Proc. Natl. Acad. Sci. USA* 93:14670-675 所述的标准固相肽合成方案进行。

PNA 可用于治疗和诊断用途。例如, PNA 可以作为反义试剂或抗基因剂用于序列特异性地调节基因表达, 例如通过诱导转录或翻译停滞或抑制复制。PNA 也可以用于例如基因中单碱基对突变的分析, 例如通过 PNA 引导的 PCR 锁止技术 (PCR clamping); 在与其他酶例如 S1 核酸酶联合使用时作为人工限制性酶(Hyrup (1996), 同上); 或者作为 DNA 测序和杂交的探针或引物(Hyrup, 1996, 同上; Perry-O'Keefe 等人, 1996, *Proc. Natl. Acad. Sci. USA* 93:14670-675)。

在另一个实施方案中, 通过将亲脂性或其他辅助基团与 PNA 连接, 通过形成 PNA-DNA 嵌合体, 或者通过使用脂质体或本领域公知的其他给药技术, 可以修饰 PNA, 例如提高其稳定性或细胞摄取。例如, 可以生成 PNA-DNA 嵌合体, 该嵌合体能够结合 PNA 和 DNA 的有利特性。该嵌合体允许 DNA 识别酶, 例如 RNase H 和 DNA 聚合酶, 与 DNA 部分相互作用, 而 PNA 部分提供高结合亲和力和特异性。PNA-DNA 嵌合体可以用根据碱基堆积、核碱基间键的数量和方向选择的适当长度的接头连接(Hyrup, 1996, 同上)。PNA-DNA 嵌合体的合成可以如 Hyrup (1996), 同上, 和 Finn 等人 (1996) *Nucleic Acids Res.* 24(17):3357-63 所述进行。例如, 可以利用标准亚磷酰胺偶合化学法和修饰的核苷类似物在固体支持体上合成 DNA 链。化合物, 如 5'-(4-甲氧三苯甲基)氨基-5'-脱氧-胸苷亚磷酰胺, 可以用作 PNA 与 DNA 的 5'末端之间的连接物(Mag 等人, 1989, *Nucleic Acids Res.* 17:5973-88)。PNA 单体然后以逐步的方式偶合, 产生具有 5' PNA 片段和 3' DNA 片段的嵌合分子(Finn 等人, 1996, *Nucleic Acids Res.* 24(17):3357-63)。或者, 也可以合成具有 5' DNA 片段和 3' PNA 片段的嵌合分子(Peterser 等人, 1975, *Bioorganic Med. Chem. Lett.* 5:1119-11124)。

在另外的实施方案中, 寡核苷酸可以包括其他附加基团, 例如肽

(例如用于在体内靶向至宿主细胞受体), 或有利于穿过细胞膜(参见, 例如 Letsinger 等人, 1989, *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre 等人, 1987, *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT 公布号 WO 88/09810)或血脑屏障 (参见, 例如 PCT 公布号 WO 89/10134) 转运的试剂。另外, 寡核苷酸也可以用杂交触发的切割剂(参见, 例如 Krol 等人, 1988, *Bio/Techniques* 6:958-976)或嵌入剂(参见, 例如 Zon, 1988, *Pharm. Res.* 5:539-549)修饰。为此, 寡核苷酸可以偶联到另一分子, 例如肽、杂交触发的交联剂、转运载体、杂交触发的切割剂等。

本发明也包括分子信标核酸, 其具有至少一个与本发明的核酸互补的区, 因此该分子信标可用于定量样品中本发明的核酸的存在。“分子信标”核酸是包含一对互补区并且具有与之结合的荧光团和荧光猝灭剂的核酸。荧光团和猝灭剂与核酸的不同部分结合, 其方向为, 当互补区彼此退火时, 荧光团的荧光被猝灭剂猝灭。当核酸的互补区彼此不退火时, 荧光团的荧光猝灭程度较低。分子信标核酸例如在美国专利 5,876,930 中记载。

分离的蛋白质和抗体

本发明的一个方案涉及分离的标志物蛋白及其生物活性部分, 以及适于作为免疫原来产生针对标志物蛋白或其片段的抗体的多肽片段。在一个实施方案中, 可以采用标准蛋白质纯化技术按照适当的纯化方案从细胞或组织来源中分离天然标志物蛋白。在另一个实施方案中, 含有标志物蛋白的全部或片段的蛋白质或肽通过重组 DNA 技术产生。作为重组表达的替代方法, 该蛋白质或肽可以采用标准肽合成技术化学合成。

“分离的”或“纯化的”蛋白质或其生物活性部分基本不含来自衍生该蛋白质的细胞或组织来源的细胞材料或其他污染蛋白质, 或者在化学合成时基本不含化学前体或其他化学物质。短语“基本不含细胞材料”包括其中蛋白质与分离或重组产生该蛋白质的细胞的细胞成分相

分离的蛋白质制品。因此，基本不含细胞材料的蛋白质包括含有少于约 30%、20%、10%或 5% (基于干重)异源蛋白质 (在此也称为“污染蛋白质”)的蛋白质制品。当蛋白质或其生物活性部分重组产生时，也优选基本不含培养基，即培养基占蛋白质制品体积的不到约 20%、10%或 5%。当蛋白质通过化学合成产生时，优选基本不含化学前体或其他化学物质，即与参与蛋白质合成的化学前体或其他化学物质分离。因此，这些蛋白质制品含有低于约 30%、20%、10%、5% (基于干重)的化学前体或目标多肽以外的化合物。

标志物蛋白的生物活性部分包括含有充分等同于或来源于该标志物蛋白的氨基酸序列的氨基酸序列的多肽，其包含比全长蛋白质更少的氨基酸，并且表现出相应全长蛋白质的至少一种活性。一般来说，生物活性部分包含具有相应全长蛋白质的至少一种活性的结构域或基序。本发明标志物蛋白的生物活性部分可以是长度为例如 10、25、50、100 个或更多氨基酸的多肽。而且，其中标志物蛋白的其他区缺失的其他生物活性部分可以利用重组技术制备，并且针对标志物蛋白天然形式的一种或多种功能活性进行评价。

优选的标志物蛋白由包含 SEQ ID NO (nts)任一序列的核苷酸序列编码。其他有用的蛋白质与这些序列之一具有基本相同(例如至少约 40%，优选 50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%相同)，并且保留天然存在的相应标志物蛋白的功能活性，然而由于天然等位基因变异或诱变其在氨基酸序列上不同。

为了确定两个氨基酸序列或两种核酸的百分同一性，将序列进行比对以便进行最佳比较 (例如，为了与第二氨基酸或核酸序列进行最佳比对，可以在第一氨基酸或核酸序列中引入缺口)。然后比较相应氨基酸位点或核苷酸位点处的氨基酸残基或核苷酸。当第一序列中的一个位点与第二序列中的相应位点被相同的氨基酸残基或核苷酸占据时，该分子在该位点处相同。两个序列之间的百分同一性是序列共有的相同位点数的函数 (即 %同一性 = 相同位点数/位点 (例如重叠位

点) 总数 $\times 100$)。在一个实施方案中, 两个序列长度相同。

两个序列之间百分同一性的确定可以用数学算法实现。用于比较两个序列的数学算法的一个优选的非限制性例子是 Karlin 和 Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268 的算法, Karlin 和 Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877 对其进行了改良。该算法并入 Altschul 等人 (1990) *J. Mol. Biol.* 215:403-410 的 BLASTN 和 BLASTX 程序中。BLAST 核苷酸检索可以用 BLASTN 程序进行, 得分 = 100, 字长 = 12, 以获得与本发明的核酸分子同源的核苷酸序列。BLAST 蛋白质检索可以用 BLASTP 程序进行, 得分 = 50, 字长 = 3, 以获得与本发明的蛋白质分子同源的氨基酸序列。为了获得用于比较目的的有缺口的比对, 可以如 Altschul 等人 (1997) *Nucleic Acids Res.* 25:3389-3402 所述使用较新版本的 BLAST 算法, 称为缺口 BLAST, 它能够进行缺口局部比对, 用于 BLASTN、BLASTP 和 BLASTX 程序。或者, 也能够利用 PSI-Blast 进行迭代检索, 该迭代检索检测分子之间的远缘关系。当使用 BLAST、缺口 BLAST 和 PSI-Blast 程序时, 可以使用相应程序(例如 BLASTX 和 BLASTN)的缺省参数。参见 <http://www.ncbi.nlm.nih.gov>。用于序列比较的数学算法的另一优选的非限制性例子是 Myers 和 Miller, (1988) *CABIOS* 4:11-17 的算法。该算法被并入 ALIGN 程序(版本 2.0)中, 该程序是 GCG 序列比对软件包的一部分。当使用 ALIGN 程序比较氨基酸序列时, 可以使用 PAM120 重量残基表、空位长度罚分为 12、空位罚分为 4。用于鉴定局部序列相似性区和用于比对的另一种有用的算法是如 Pearson 和 Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85:2444-2448 所述的 FASTA 算法。当使用 FASTA 算法比较核苷酸或氨基酸序列时, PAM120 重量残基表例如可以使用为 2 的 k -元组值。

两个序列之间的百分同一性可以用与以上所述类似、允许或不允许缺口的技术来确定。在计算百分同一性时, 只计数准确的匹配。

本发明也提供包含标志物蛋白或其片段的嵌合或融合蛋白。此处使用的“嵌合蛋白”或“融合蛋白”包含与异源多肽(即标志物蛋白以外

的多肽)可操作连接的标志物蛋白的全部或部分(优选生物活性部分)。用于融合蛋白时,术语“可操作连接”是指标志物蛋白或其片段和异源多肽彼此在阅读框内融合。异源多肽可以与标志物蛋白或片段的氨基端或羧基端融合。

一种有用的融合蛋白是 GST 融合蛋白,其中标志物蛋白或片段与 GST 序列的羧基端融合。该融合蛋白有利于本发明的重组多肽的纯化。

在另一个实施方案中,融合蛋白在其氨基端含有异源信号序列。例如,可以将标志物蛋白的天然信号序列除去并替换为来自另一种蛋白质的信号序列。例如,杆状病毒包膜蛋白的 gp67 分泌序列可以用作异源信号序列(Ausubel 等人, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, NY, 1992)。真核生物异源信号序列的其他例子包括蜂毒肽和人胎盘碱性磷酸酶的分泌序列(Stratagene; La Jolla, California)。在另一个例子中,有用的原核生物异源信号序列包括 phoA 分泌信号(Sambrook 等人,同上)和蛋白 A 分泌信号(Pharmacia Biotech; Piscataway, New Jersey)。

在另一个实施方案中,所述融合蛋白是免疫球蛋白融合蛋白,其中标志物蛋白的全部或一部分与来源于该免疫球蛋白家族成员的序列融合。本发明的免疫球蛋白融合蛋白可以加入到药物组合物中,对受试者施用,来抑制(可溶性的或膜结合的)配体与细胞表面上的蛋白质(受体)之间的相互作用,从而抑制体内信号转导。免疫球蛋白融合蛋白可以用来影响标志物蛋白的相应配体的生物利用度。配体/受体相互作用的抑制在治疗上可能有用,既可用于治疗增生性和分化性疾病,又可用于调节(例如促进或抑制)细胞存活率。而且,本发明的免疫球蛋白融合蛋白能够作为免疫原用于在受试者中产生抗标志物蛋白的抗体,用于纯化配体,以及在筛选试验中用来鉴定抑制标志物蛋白与配体相互作用的分子。

本发明的嵌合蛋白和融合蛋白可以通过标准重组 DNA 技术制备。在另一个实施方案中,融合基因可以利用常规技术合成,包括使

用自动 DNA 合成仪。或者，可以用锚引物进行基因片段的 PCR 扩增，该锚引物在两个连续基因片段之间产生互补的突出部分（overhang），之后可以退火并再次扩增，产生嵌合基因序列（参见，例如 Ausubel 等人,同上）。另外，许多已经编码融合部分（例如 GST 多肽）的表达载体在商业上可以获得。可以将编码本发明多肽的核酸克隆到这种表达载体中，使融合部分与本发明的多肽在阅读框内连接。

可以利用信号序列促进标志物蛋白的分泌和分离。信号序列的特征一般在于疏水性氨基酸的核心，通常在一个或多个切割事件中在分泌过程中从成熟蛋白上切下。该信号肽含有加工位点，该加工位点允许在其通过分泌途径时从成熟蛋白上切下信号序列。因此，本发明涉及具有信号序列的标志物蛋白、融合蛋白或其片段，以及已经蛋白水解切割下信号序列的蛋白质（即切割产物）。在一个实施方案中，编码信号序列的核酸序列在表达载体中可以与目标蛋白质如标记蛋白或其片段可操作地连接。信号序列引导例如从用表达载体转化的原核宿主中分泌蛋白质，该信号序列后来或者同时被切下。然后通过本领域公知的方法容易地从胞外基质中纯化该蛋白质。或者，可以利用有利于纯化的序列，如利用 GST 域，将信号序列与目标蛋白质连接。

本发明也涉及标志物蛋白的变体。该变体具有改变的氨基酸序列，其可以作为激动剂（模拟物）或拮抗剂。可以通过诱变，例如不连续的点突变或截短来产生变体。激动剂可保留与天然存在的蛋白质形式基本相同的生物活性，或其一部分活性。蛋白质的拮抗剂能够抑制该蛋白质天然存在形式的一种或多种活性，例如通过竞争性结合到包括目标蛋白质的细胞信号级联的下游或上游成员上产生这种抑制。因此，用有限功能的变体处理能够引发特定的生物效应。相对于用蛋白质的天然存在形式处理，用具有该蛋白质天然存在形式的一部分生物活性的变体处理受试者可能在受试者中具有更少的副作用。

作为激动剂（模拟物）或拮抗剂的标志物蛋白的变体可以通过对激动剂或拮抗剂活性筛查本发明蛋白质的突变体例如截短突变体的组

合文库来鉴定。在一个实施方案中，多样化的（variegated）变体文库通过在核酸水平上的组合诱变而产生，并且由多样化的基因文库编码。例如，通过将合成寡核苷酸的混合物酶学连接到基因序列中，使简并的一组可能的蛋白质序列表达为各种多肽，或者表达为一组较大的融合蛋白（例如用于噬菌体展示），能够产生多样化的变体文库。有多种方法能够用于由简并寡核苷酸序列产生标志物蛋白的可能变体的文库。合成简并寡核苷酸的方法在本领域公知（参见，例如 Narang, 1983, *Tetrahedron* 39:3; Itakura 等人, 1984, *Annu. Rev. Biochem.* 53:323; Itakura 等人, 1984, *Science* 198:1056; Ike 等人, 1983 *Nucleic Acid Res.* 11:477)。

另外，可以利用标志物蛋白片段文库产生多样化的一群多肽，用于筛选及随后选择变异标志物蛋白或其片段。例如，编码序列片段文库可以如下产生：在每个分子仅发生大约一次缺刻的条件下用核酸酶处理目标编码序列的双链 PCR 片段，变性该双链 DNA，将该 DNA 复性以形成可以包括来自不同缺刻产物的有义/反义对的双链 DNA，通过 S1 核酸酶处理从重新形成的双链体中除去单链部分，以及将产生的片段文库连接到表达载体中。利用该方法能够产生表达文库，其编码目标蛋白质的各种大小的氨基末端和内部片段。

有几种技术在本领域中公知用于筛查通过点突变或截短制备的组合文库的基因产物，以及为了具有选择的性质的基因产物而筛查 cDNA 文库。适用于高通量分析、用于筛查大基因文库的最广泛应用的技术通常包括，将基因文库克隆到复制型表达载体中，用得到的载体文库转化合适的细胞，以及在一定条件下表达该组合基因，其中对所需活性的检测有利于分离编码其产物被检测的基因的载体。递归总体诱变(REM)是一种提高文库中功能突变体频率的技术，可以与筛查试验联合应用，用来鉴定本发明的蛋白质的变体(Arkin 和 Yourvan, 1992, *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave 等人, 1993, *Protein Engineering* 6(3):327-331)。

本发明的另一个方案涉及抗本发明蛋白质的抗体。在优选实施方

案中，所述抗体特异性结合标志物蛋白或其片段。此处使用的术语“抗体”是指免疫球蛋白分子及其包含免疫球蛋白分子的免疫活性部分的片段和衍生物（即该部分含有抗原结合位点，该位点特异性结合抗原，如标志物蛋白，例如标志物蛋白质的表位）。特异性结合本发明蛋白质的抗体是与该蛋白质结合但是基本不与天然含有该蛋白质的样品例如生物样品中的其他分子结合的抗体。免疫球蛋白分子的免疫活性部分的例子包括但不限于单链抗体(scAb)、F(ab)和 F(ab')₂ 片段。

分离的本发明的蛋白质或其片段可以作为免疫原用于产生抗体。可以使用全长蛋白质，或者本发明提供抗原性肽片段用作免疫原。本发明蛋白质的抗原性肽包含本发明蛋白质之一的氨基酸序列的至少 8 个(优选 10、15、20 或 30 个或更多)氨基酸残基，并且包含该蛋白质的至少一个表位，从而抗该肽的抗体与该蛋白质形成特定的免疫复合物。该抗原性肽包含的优选表位是位于蛋白质表面上的区域，例如亲水区。疏水性序列分析，亲水性序列分析，或类似的分析，都能够用来鉴定亲水区。在优选实施方案中，用分离的标志物蛋白或其片段作为免疫原。

一般利用免疫原，通过免疫合适的（即免疫活性）受试者，如兔、山羊、小鼠或其他哺乳动物或脊椎动物，来制备抗体。合适的免疫原性制品可能包含，例如重组表达的或化学合成的蛋白质或肽。该制品可进一步包含佐剂，如弗氏完全或不完全佐剂，或类似的免疫刺激剂。优选的免疫原组合物不含其他人类蛋白质，例如利用重组表达本发明蛋白质的非人类宿主细胞制备的免疫原组合物。在这种情况下，产生的抗体组合物没有或者具有减少的本发明蛋白质以外的人类蛋白质的结合。

本发明提供多克隆和单克隆抗体。此处使用的术语“单克隆抗体”或“单克隆抗体组合物”是指一组抗体分子，它们只含一种能够与特定表位免疫反应的抗原结合位点。优选的多克隆和单克隆抗体组合物是已经为抗本发明蛋白质的抗体而选择的那些。特别优选的多克隆和单克隆抗体制剂是只含抗标志物蛋白或其片段的抗体的那些。

多克隆抗体可以通过用本发明的蛋白质作为免疫原免疫合适的受试者来制备。被免疫的受试者中的抗体效价可以利用标准技术随时间监测,例如采用应用固定多肽的酶联免疫吸附测定(ELISA)。在免疫后适当的时间,例如当特异性抗体效价最高时,可以从受试者中获得抗体产生细胞,并且通过标准技术用其制备单克隆抗体(mAb),例如最早由 Kohler 和 Milstein (1975) *Nature* 256:495-497 描述的杂交瘤技术,人 B 细胞杂交瘤技术(参见 Kozbor 等人, 1983, *Immunol. Today* 4:72), EBV-杂交瘤技术(参见 Cole 等人, pp. 77-96 In *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., 1985)或三源杂交瘤(trioma)技术。产生杂交瘤的技术众所周知(一般参见 *Current Protocols in Immunology*, Coligan 等人 ed., John Wiley & Sons, New York, 1994)。通过例如利用标准 ELISA 试验针对可结合目标多肽的抗体筛查杂交瘤培养上清液,检测产生本发明的单克隆抗体的杂交瘤细胞。

作为制备分泌单克隆抗体的杂交瘤的替代方法,也可以通过用目标多肽筛查重组组合免疫球蛋白文库(例如抗体噬菌体展示文库)来鉴定和分离抗本发明蛋白质的单克隆抗体。用于制备和筛查噬菌体展示文库的试剂盒在商业上可以获得(例如 Pharmacia 重组噬菌体抗体系统,目录号 27-9400-01;和 Stratagene *SurfZAP* 噬菌体展示试剂盒,目录号 240612)。另外,特别适用于制备和筛查抗体展示文库的方法和试剂的例子可见于例如美国专利 No. 5,223,409; PCT 公布号 WO 92/18619; PCT 公布号 WO 91/17271; PCT 公布号 WO 92/20791; PCT 公布号 WO 92/15679; PCT 公布号 WO 93/01288; PCT 公布号 WO 92/01047; PCT 公布号 WO 92/09690; PCT 公布号 WO 90/02809; Fuchs 等人 (1991) *Bio/Technology* 9:1370-1372; Hay 等人 (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse 等人 (1989) *Science* 246:1275-1281; Griffiths 等人 (1993) *EMBO J.* 12:725-734。

本发明也提供特异性结合本发明蛋白质的重组抗体。在优选实施方案中,该重组抗体特异性结合标志物蛋白或其片段。重组抗体包括

但不限于包含人类和非人类部分的嵌合和人源化单克隆抗体、单链抗体和多特异性抗体。嵌合抗体是不同部分来源于不同动物种的分子，例如具有来源于鼠 mAb 的可变区和人免疫球蛋白恒定区的抗体。(参见，例如 Cabilly 等人，美国专利 No. 4,816,567;和 Boss 等人，美国专利 No. 4,816,397，其全文在此引作参考)。单链抗体具有一个抗原结合位点，并且由一种多肽组成。它们可以利用本领域公知的技术制备，例如利用 Ladner 等人的美国专利 No. 4,946,778 (其全文在此引作参考); Bird 等人，(1988) *Science* 242:423-426; Whitlow 等人，(1991) *Methods in Enzymology* 2:1-9; Whitlow 等人，(1991) *Methods in Enzymology* 2:97-105;和 Huston 等人，(1991) *Methods in Enzymology Molecular Design and Modeling: Concepts and Applications* 203:46-88 所述的方法制备。多特异性抗体是具有至少两个抗原结合位点的抗体分子，这些位点特异性可结合不同的抗原。这类分子可以利用本领域公知的技术制备，例如利用 Segal 的美国专利 No. 4,676,980 (其全文在此引作参考); Holliger 等人，(1993) *Proc. Natl. Acad. Sci. USA* 90:6444-6448; Whitlow 等人，(1994) *Protein Eng.* 7:1017-1026 和美国专利 No. 6,121,424 所述的方法制备。

人源化抗体是这样的抗体分子：其来自非人类物种，具有一个或多个来自非人类物种的互补决定区(CDR)和来自人免疫球蛋白分子的框架区。(参见，例如 Queen，美国专利 No. 5,585,089，其全文在此引作参考)。人源化单克隆抗体可以利用本领域公知的重组 DNA 技术制备，例如采用 PCT 公布号 WO 87/02671; 欧洲专利申请 184,187; 欧洲专利申请 171,496; 欧洲专利申请 173,494; PCT 公布号 WO 86/01533; 美国专利 No. 4,816,567; 欧洲专利申请 125,023; Better 等人 (1988) *Science* 240:1041-1043; Liu 等人 (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu 等人 (1987) *J. Immunol.* 139:3521- 3526; Sun 等人 (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura 等人 (1987) *Cancer Res.* 47:999-1005; Wood 等人 (1985) *Nature* 314:446-449; Shaw 等人 (1988) *J. Natl. Cancer Inst.* 80:1553-1559; Morrison (1985) *Science*

229:1202-1207; Oi 等人 (1986) *Bio/Techniques* 4:214; 美国专利 5,225,539; Jones 等人 (1986) *Nature* 321:552-525; Verhoeyan 等人 (1988) *Science* 239:1534; 和 Beidler 等人 (1988) *J. Immunol.* 141:4053-4060 中所述的方法制备。

更特别地,人源化抗体例如可以用转基因小鼠产生,该转基因小鼠不能表达内源免疫球蛋白重链和轻链基因,但是能够表达人重链和轻链基因。用选定的抗原,例如对应于本发明标志物的多肽的全部或一部分,按照正常方式免疫该转基因小鼠。可以用常规杂交瘤技术获得抗该抗原的单克隆抗体。转基因小鼠携带的人免疫球蛋白转基因在 B 细胞分化过程中重排,随后经历类别转换和体细胞突变。因此,采用该技术,能够产生在治疗上有用的 IgG、IgA 和 IgE 抗体。关于这种制备人抗体的技术的综述,参见 Lonberg 和 Huszar (1995) *Int. Rev. Immunol.* 13:65-93。关于这种制备人抗体和人单克隆抗体的技术和制备该抗体的方案的详细记载,参见,例如美国专利 5,625,126; 美国专利 5,633,425; 美国专利 5,569,825; 美国专利 5,661,016; 和美国专利 5,545,806。另外,诸如 Abgenix, Inc. (Freemont, CA) 等公司能够接受委托利用与上述类似的技术提供抗选定抗原的人抗体。

识别选定表位的完全人抗体可以利用一种被称为“指导选择”的技术来制备。在该方法中,利用选择的非人单克隆抗体,例如鼠抗体,指导对识别相同表位的完全人抗体的选择(Jespers 等人, 1994, *Bio/technology* 12:899-903)。

本发明的抗体可以在(例如从受试者血液或血清中)产生或合成后分离,并进一步用众所周知的技术纯化。例如, IgG 抗体可以利用蛋白 A 层析法纯化。本发明蛋白质的特异性抗体可以利用例如亲和层析法来选择或(例如部分纯化)或纯化。例如,重组表达并纯化的(或部分纯化的)本发明的蛋白质如此处所述制备,并且共价或非共价偶联到固体支持物,例如层析柱上。然后可以用该柱从含有抗大量不同表位的抗体的样品中亲和纯化本发明蛋白质的特异性抗体,从而产生基本纯化的抗体组合物,即基本不含污染抗体的组合物。基本纯

化的抗体组合物在本文中是指抗体样品最多只含有 30% (基于干重) 抗本发明所需蛋白质表位之外的表位的污染抗体, 该样品中优选最多 20%, 更优选最多 10%, 最优选最多 5% (基于干重) 是污染抗体。纯化的抗体组合物是指该组合物中至少 99% 的抗体是抗本发明所需蛋白质的抗体。

在一个优选实施方案中, 基本纯化的本发明的抗体可以特异性结合本发明蛋白质的信号肽、分泌序列、胞外域、跨膜区或胞质域或细胞质膜。在一个特别优选的实施方案中, 基本纯化的本发明的抗体特异性结合本发明蛋白质的氨基酸序列的分泌序列或胞外域。在一个更优选的实施方案中, 基本纯化的本发明的抗体特异性结合标志物蛋白的氨基酸序列的分泌序列或胞外域。

可以通过标准技术, 例如亲和层析或免疫沉淀法, 利用抗本发明蛋白质的抗体分离该蛋白质。另外, 也可以利用这种抗体检测 (例如细胞裂解液或细胞上清液中的) 标志物蛋白或其片段, 从而评价该标志物的表达水平和模式。也可以在诊断上应用这些抗体来监测组织或体液 (例如前列腺相关体液) 中的蛋白质水平, 作为临床检验程序的一部分, 例如确定特定治疗方案的效果。利用包含偶联到可检测物质上的本发明抗体的抗体衍生物可有利于检测。可检测物质的例子包括各种酶、辅基、荧光物质、发光物质、生物发光物质和放射性物质。合适的酶的例子包括辣根过氧化物酶、碱性磷酸酶、 β -半乳糖苷酶或乙酰胆碱酯酶; 合适的辅基复合物的例子包括链霉亲和素/生物素和亲和素/生物素; 合适的荧光物质的例子包括伞形酮、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; 发光物质的例子包括鲁米诺; 生物发光物质的例子包括萤光素酶、萤光素和水母蛋白; 合适的放射性物质的例子包括 ^{125}I 、 ^{131}I 、 ^{35}S 或 ^3H 。

本发明的抗体也可以用作治疗癌症的治疗剂。在一个优选实施方案中, 本发明的完全人类抗体用于治疗人类癌症患者, 特别是前列腺癌患者。在另一优选实施方案中, 能与标志物蛋白或其片段特异性结合的抗体用于治疗。另外, 这些治疗性抗体也可以是抗体衍生物或免

疫毒素，其含有与治疗性部分如细胞毒素、治疗剂或放射性金属离子偶联的抗体。细胞毒素或细胞毒性剂包括对细胞有害的任何药剂。例子包括紫杉醇、松胞菌素 B、短杆菌肽 D、溴乙锭、依米丁、丝裂霉素、依托泊苷、鬼臼噻吩甙 (tenoposide)、长春新碱、长春碱、秋水仙素、阿霉素、柔红霉素、二羟基炭疽菌素二酮、米托蒽醌、光辉霉素、放线菌素 D、1-去氢睾酮、糖皮质激素类、普鲁卡因、丁卡因、利多卡因、普萘洛尔和嘌呤霉素及其类似物或同源物。治疗剂包括但不限于：抗代谢物(例如，氨甲喋呤、6-巯基嘌呤、6-硫代鸟嘌呤、阿糖胞苷、5-氟尿嘧啶甲氮咪胺)、烷化剂(例如，氮芥、三胺硫磷(thioepa)、苯丁酸氮芥、苯丙氨酸氮芥、卡莫司汀(BSNU)和洛莫司汀(CCNU)、环磷酰胺、白消安、二溴甘露醇、链脲霉素、丝裂霉素 C 和顺-二氯二胺铂(II) (DDP) 顺铂)、蒽环霉素类(例如柔红霉素(以前称为道诺霉素)和多柔比星)、抗生素(例如放线菌素 D (以前称为放线菌素)、博来霉素、光辉霉素和安曲霉素(AMC))，和抗有丝分裂剂(例如长春新碱和长春碱)。

偶联的本发明的抗体可以用于改变特定生物学应答，对于药物部分而言，不应理解为限于经典的化学治疗剂。例如，药物部分可以是具有需要的生物活性的蛋白质或多肽。这样的蛋白质可包括，例如：毒素，如核糖体抑制蛋白(参见 Better 等人的美国专利 No. 6,146,631，其公开内容在此全文引作参考)、相思豆毒蛋白、蓖麻毒蛋白 A、假单胞菌外毒素或白喉毒素；蛋白质，如肿瘤坏死因子、 α -干扰素、 β -干扰素、神经生长因子、血小板衍生生长因子、组织纤维蛋白溶酶原激活剂；或生物反应调节剂，例如淋巴因子、白介素-1 ("IL-1")、白介素-2 ("IL-2")、白介素-6 ("IL-6")、粒细胞巨噬细胞集落刺激因子("GM-CSF")、粒细胞集落刺激因子("G-CSF")或其他生长因子。

将治疗性部分与抗体进行偶联的技术众所周知，参见，例如 Arnon 等人，"Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy" (癌症治疗中用于药物免疫靶向的单克隆抗体)，in

Monoclonal Antibodies And Cancer Therapy, Reisfeld 等人 (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom 等人, "Antibodies For Drug Delivery" (用于药物递送的抗体), in Controlled Drug Delivery (2nd Ed.), Robinson 等人 (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review" (癌症治疗中细胞毒性剂的抗体载体: 综述), in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera 等人 (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy" (放射性标记抗体在癌症治疗中的治疗应用的分析、结果和前景), in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin 等人 (eds.), pp. 303-16 (Academic Press 1985), 和 Thorpe 等人, "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates" (抗体-毒素偶联物的制备和细胞毒性性质), Immunol. Rev., 62:119-58 (1982)。

因此, 在一个方面, 本发明提供基本纯化的抗体、抗体片段和衍生物, 它们全都特异性结合本发明的蛋白质, 优选标志物蛋白。在各个实施方案中, 基本纯化的本发明的抗体或其片段或衍生物可以是人、非人、嵌合和/或人源化抗体。在另一个方面, 本发明提供非人抗体、抗体片段和衍生物, 它们全都特异性结合本发明的蛋白质, 优选标志物蛋白。这样的非人抗体可以是山羊、小鼠、绵羊、马、鸡、兔或大鼠抗体。或者, 本发明的非人抗体可以是嵌合和/或人源化抗体。另外, 本发明的非人抗体可以是多克隆抗体或单克隆抗体。在又另一个方面, 本发明提供单克隆抗体、抗体片段和衍生物, 它们全都特异性结合本发明的蛋白质, 优选标志物蛋白。该单克隆抗体可以是人、人源化、嵌合和/或非人抗体。

本发明也提供一种试剂盒, 其包括与可检测物质偶联的本发明的抗体, 和使用说明。本发明的再一方案是一种含有本发明的抗体的药物组合物。在一个实施方案中, 药物组合物含有本发明的抗体和药学

可接受的载体。

重组表达载体和宿主细胞

本发明的另一方案涉及载体，优选表达载体，其含有编码标志物蛋白(或该蛋白的一部分)的核酸。此处使用的术语“载体”是指能够转运与之连接的另一种核酸的核酸分子。一种类型的载体是“质粒”，它是指可以向其中连接其他 DNA 片段的环状双链 DNA 环。另一种类型的载体是病毒载体，其中其他 DNA 片段可以连接到病毒基因组中。某些载体能够在引入它们的宿主细胞中自主复制（例如具有细菌复制起点的细菌载体和附加型哺乳动物载体）。其他载体（例如非附加型哺乳动物载体）在引入宿主细胞后整合到宿主细胞基因组中，从而与宿主基因组一起复制。而且，某些载体，即表达载体，能够引导与它们可操作连接的基因的表达。通常，在重组 DNA 技术中使用的表达载体常常为质粒形式（载体）。但是，本发明也包括行使同等功能的其他形式的表达载体，例如病毒载体（例如复制缺陷型逆转录病毒、腺病毒和腺伴随病毒）。

本发明的重组表达载体包含本发明的核酸，其形式适于在宿主细胞中表达该核酸。意思是重组表达载体包括一个或多个调节序列，该调节序列是根据用于表达的宿主细胞选择的，与将要表达的核酸序列可操作地连接。在重组表达载体内，“可操作地连接”是指目标核苷酸序列以允许该核苷酸序列（例如在体外转录/翻译系统中，或者当向宿主细胞中引入该载体时在宿主细胞中）表达的方式与调节序列连接。术语“调节序列”包括启动子、增强子和其他表达控制元件（例如聚腺苷酸化信号）。这些调节序列例如在 Goeddel, *Methods in Enzymology: Gene Expression Technology* vol.185, Academic Press, San Diego, CA (1991)中记载。调节序列包括引导核苷酸序列在多种类型宿主细胞中组成型表达的那些，和引导核苷酸序列只在特定宿主细胞中表达的那些(例如组织特异性调节序列)。本领域技术人员应当理解，表达载体的设计将取决于诸如将要转化的宿主细胞的选择、需要的蛋白质表达

水平等因素。本发明的表达载体可以导入宿主细胞中，从而产生如本文所述的由核酸编码的蛋白质或肽，包括融合蛋白或肽。

本发明的重组表达载体可以设计用于标志物蛋白或其片段在原核（例如大肠杆菌）或真核细胞（例如昆虫细胞{利用杆状病毒表达载体}、酵母细胞或哺乳动物细胞）中的表达。合适的宿主细胞进一步记载在 Goeddel,同上中。或者，重组表达载体也可以在体外转录和翻译，例如应用 T7 启动子调节序列和 T7 聚合酶。

蛋白质在原核生物中的表达最常使用含有引导融合或非融合蛋白表达的组成型或诱导型启动子的载体在大肠杆菌中进行。融合载体向本文所述编码的蛋白质中，通常向重组蛋白的氨基端添加许多氨基酸。该融合载体一般用于三个目的：1)提高重组蛋白的表达；2) 提高重组蛋白的可溶性；和 3) 通过在亲和纯化中作为配体来帮助重组蛋白的纯化。在融合表达载体中，通常在融合部分和重组蛋白的连接处引入一个蛋白水解切割位点，从而能够在融合蛋白纯化后将重组蛋白与融合部分分离。这样的酶，及其相应的识别序列，包括因子 Xa、凝血酶和肠激酶。典型的融合表达载体包括 pGEX (Pharmacia Biotech Inc; Smith 和 Johnson, 1988, *Gene* 67:31-40)、pMAL (New England Biolabs, Beverly, MA)和 pRIT5 (Pharmacia, Piscataway, NJ)，它们分别将谷胱甘肽 S-转移酶(GST)、麦芽糖 E 结合蛋白或蛋白 A 与靶重组蛋白相融合。

合适的诱导型非融合大肠杆菌表达载体的例子包括 pTrc (Amann 等人, 1988, *Gene* 69:301-315)和 pET 11d (Studier 等人, p. 60-89, In *Gene Expression Technology: Methods in Enzymology* vol.185, Academic Press, San Diego, CA, 1991)。由 pTrc 载体表达靶基因依赖于从杂合 trp-lac 融合启动子开始的宿主 RNA 聚合酶转录。由 pET 11d 载体表达靶基因依赖于由共表达的病毒 RNA 聚合酶(T7 gn1)介导的从 T7 gn10-lac 融合启动子开始的转录。该病毒聚合酶由宿主株 BL21(DE3)或 HMS174(DE3)从驻留的原噬菌体提供，该原噬菌体携带处于 lacUV 5 启动子转录控制下的 T7 gn1 基因。

一种使重组蛋白在大肠杆菌中的表达最大化的策略是在蛋白水解该重组蛋白能力受损的宿主细菌中表达该蛋白质(Gottesman, p. 119-128, In *Gene Expression Technology: Methods in Enzymology* vol. 185, Academic Press, San Diego, CA, 1990)。另一种策略是改变将要插入到表达载体内的核酸的核酸序列,使每个氨基酸的各个密码子是大肠杆菌偏好使用的(Wada 等人, 1992, *Nucleic Acids Res.* 20:2111-2118)。本发明的核酸序列的这种改变可以采用标准 DNA 合成技术进行。

在另一个实施方案中,表达载体是酵母表达载体。用于在酿酒酵母 (*S. cerevisiae*) 中表达的载体的例子包括 pYepSec1 (Baldari 等人, 1987, *EMBO J.* 6:229-234)、pMFa (Kurjan 和 Herskowitz, 1982, *Cell* 30:933-943)、pJRY88 (Schultz 等人, 1987, *Gene* 54:113-123)、pYES2 (Invitrogen Corporation, San Diego, CA)和 pPicZ (Invitrogen Corp, San Diego, CA)。

或者,表达载体是杆状病毒表达载体。可以获得的用于在培养的昆虫细胞(例如 Sf 9 细胞)中表达蛋白质的杆状病毒载体包括 pAc 系列(Smith 等人, 1983, *Mol. Cell Biol.* 3:2156-2165)和 pVL 系列(Lucklow 和 Summers, 1989, *Virology* 170:31-39)。

在另一个实施方案中,使用哺乳动物表达载体在哺乳动物细胞中表达本发明的核酸。哺乳动物表达载体的例子包括 pCDM8 (Seed, 1987, *Nature* 329:840)和 pMT2PC (Kaufman 等人, 1987, *EMBO J.* 6:187-195)。当在哺乳动物细胞中使用时,表达载体的控制功能通常由病毒调节元件提供。例如,常用的启动子来源于多瘤、腺病毒 2、巨细胞病毒和猿猴病毒 40。关于用于原核和真核细胞的其他合适的表达系统,见 Sambrook 等人,同上的第 16、17 章。

在另一个实施方案中,重组哺乳动物表达载体能够引导核酸优先在特定细胞型中表达(例如利用组织特异性调节元件表达核酸)。组织特异性调节元件在本领域中公知。合适的组织特异性启动子的例子包括但不限于白蛋白启动子(肝特异性的; Pinkert 等人, 1987, *Genes Dev.* 1:268-277),淋巴特异性启动子(Calame 和 Eaton, 1988, *Adv.*

Immunol. 43:235-275), 特别是 T 细胞受体的启动子(Winoto 和 Baltimore, 1989, *EMBO J.* 8:729-733)和免疫球蛋白的启动子(Banerji 等人, 1983, *Cell* 33:729-740; Queen 和 Baltimore, 1983, *Cell* 33:741-748), 神经元特异性启动子(例如神经丝启动子; Byrne 和 Ruddle, 1989, *Proc. Natl. Acad. Sci. USA* 86:5473-5477), 胰特异性启动子(Edlund 等人, 1985, *Science* 230:912-916), 和乳腺特异性启动子(例如乳清启动子; 美国专利 No. 4,873,316 和欧洲申请公布号 264,166)。也包括发育调节的启动子, 例如鼠 *hox* 启动子(Kessel 和 Gruss, 1990, *Science* 249:374-379)和甲胎蛋白启动子(Camper 和 Tilghman, 1989, *Genes Dev.* 3:537-546)。

本发明还提供一种重组表达载体, 其包含以反义方向克隆到该表达载体内的本发明的 DNA 分子。即, 该 DNA 分子与调节序列可操作地连接, 其连接方式允许相对于编码本发明多肽的 mRNA 而言为反义的 RNA 分子表达(通过 DNA 分子的转录)。可以选择以反义方向与克隆的核酸可操作连接的调节序列, 其引导反义 RNA 分子在多种细胞型中的连续表达, 例如病毒启动子和/或增强子, 或者可以选择引导反义 RNA 的组成型、组织特异性或细胞型特异性表达的调节序列。反义表达载体可以是重组质粒、噬菌粒或减毒病毒的形式, 其中反义核酸在高效调节区控制下产生, 其活性取决于该载体所导入的细胞型。关于使用反义基因调节基因表达的讨论, 参见 Weintraub 等人, 1986, *Trends in Genetics*, Vol. 1(1)。

本发明的另一方面涉及已经导入本发明的重组表达载体的宿主细胞。术语“宿主细胞”和“重组宿主细胞”在此可交换使用。应当理解, 这些术语不仅指特定受试细胞, 而且指该细胞的后代或可能的后代。由于突变或环境的影响在连续几代中可能发生某些变化, 因此这些后代实际上可能与母细胞不同, 但是仍然包括在此处所用术语的范围内。

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

可以通过常规转化或转染技术将载体 DNA 导入原核或真核细胞中。此处使用的术语“转化”和“转染”是指多种本领域公知的向宿主细胞中引入外源核酸的技术，包括磷酸钙或氯化钙共沉淀、DEAE-葡聚糖-介导的转染、脂转染或电穿孔。用于转化或转染宿主细胞的合适的方法可见 Sambrook, 等人 (同上) 和其他实验室手册。

对于哺乳动物细胞的稳定转染，众所周知，根据使用的表达载体和转染技术，只有少部分细胞可以将外源 DNA 整合到其基因组中。为了鉴定和选择这些整合体，通常与目标基因一起向宿主细胞中引入编码选择性标记（例如抗生素抗性）的基因。优选的选择性标记包括提供药物抗性，例如 G418、潮霉素和氨甲喋呤抗性的标记。用引入的核酸稳定转染的细胞可以通过药物筛选来鉴定（例如，已引入选择性标记的细胞将存活，而其他细胞将死亡）。

本发明的宿主细胞，如培养的原核或真核宿主细胞，可以用来产生标志物蛋白或其片段。因此，本发明进一步提供利用本发明的宿主细胞产生标志物蛋白或其片段的方法。在一个实施方案中，该方法包括在适当的培养基中培养本发明的宿主细胞（其中已经引入了编码标志物蛋白或其片段的重组表达载体），从而使其产生。在另一个实施方案中，该方法进一步包括从培养基或宿主细胞中分离该标志物蛋白或其片段。

也可以利用本发明的宿主细胞产生非人转基因动物。例如，在一个实施方案中，本发明的宿主细胞是已经引入编码标志物蛋白或其片段的序列的受精卵母细胞或胚胎干细胞。然后用该宿主细胞产生已经向其基因组中导入编码本发明标志物蛋白的外源序列的非人转基因动物，或编码标志物蛋白的内源基因已经改变的同源重组动物。这些动物可用于研究该标志物蛋白的功能和/或活性，以及鉴定和/或评价该标志物蛋白的调节剂。此处使用的“转基因动物”是非人动物，优选哺乳动物，更优选啮齿类动物，如大鼠或小鼠，其中该动物的一个或多个细胞包含转基因。转基因动物的其他例子包括非人灵长类动物、绵羊、狗、牛、山羊、鸡、两栖动物等。转基因是这样的外源

DNA，其整合到将发育为转基因动物的细胞的基因组内，并且保留在成熟动物的基因组内，从而引导在该转基因动物的一种或多种细胞型或组织中表达编码的基因产物。此处使用的“同源重组动物”是非人类动物，优选哺乳动物，更优选小鼠，其中在发育为动物之前，通过内源基因与引入动物细胞例如动物胚胎细胞内的外源 DNA 分子之间的同源重组，该内源基因已经改变。

本发明的转基因动物可以如下产生：例如通过显微注射、逆转录病毒感染将编码标志物蛋白的核酸导入受精卵母细胞的雄性原核内，并且使该卵母细胞在假孕雌性代孕动物中发育。内含子序列和聚腺苷酸化信号也可以包含在转基因中来提高转基因的表达效率。组织特异性调节序列可以与转基因可操作连接，以使本发明多肽的表达针对特定细胞。通过胚胎操作和显微注射产生转基因动物，特别是如小鼠等动物的方法在本领域中已经成为常规，例如在美国专利 No. 4,736,866 和 4,870,009，美国专利 No. 4,873,191 和 Hogan, *Manipulating the Mouse Embryo*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986 中描述。类似的方法用于其他转基因动物的产生。可以根据转基因在其基因组中的存在和/或编码转基因的 mRNA 在动物组织或细胞中的表达来鉴定转基因起始（founder）动物。然后可以用转基因起始动物繁殖携带该转基因的其他动物。另外，携带该转基因的转基因动物可以进一步与携带其他转基因的其他转基因动物繁殖。

为了产生同源重组动物，制备一种载体，该载体含有编码标志物蛋白的基因的至少一部分，其中已经引入了缺失、添加或置换，从而改变，例如功能破坏了该基因。在一个优选实施方案中，该载体被设计为内源基因在同源重组后被功能破坏（即不再编码功能蛋白质；也被称为“敲除”载体）。或者，该载体也可以被设计为内源基因在同源重组后突变或者以其他方式改变，但是仍然编码功能蛋白质（例如可以改变上游调节区，从而改变内源蛋白质的表达）。在该同源重组载体中，改变的基因部分在 5'端和 3'端之侧为该基因的其他核酸，允许

在该载体携带的外源基因与胚胎干细胞中的内源基因之间发生同源重组。其他侧翼核酸序列的长度足以成功地与内源基因同源重组。载体中一般包含几千个碱基的侧翼 DNA（在 5'端和 3'端）（关于同源重组载体的描述，参见，例如，Thomas 和 Capecchi, 1987, *Cell* 51:503）。将该载体导入胚胎干细胞系中（例如通过电穿孔），选择被导入的基因已经与内源基因同源重组的细胞（参见，例如 Li 等人, 1992, *Cell* 69:915）。然后将选择的细胞注射到动物（例如小鼠）的胚囊中，形成聚集嵌合体（参见，例如 Bradley, *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, Ed., IRL, Oxford, 1987, pp. 113-152）。然后将嵌合胚植入合适的假孕雌性代孕动物中，使胚胎发育。在生殖细胞中携带同源重组的 DNA 的后代可以用来繁殖动物，其中通过转基因的种系传递，动物的所有细胞都含有同源重组的 DNA。构建同源重组载体和同源重组动物的方法在 Bradley (1991) *Current Opinion in Bio/Technology* 2:823-829 和 PCT 公布号 WO 90/11354、WO 91/01140、WO 92/0968 和 WO 93/04169 中记载。

在另一个实施方案中，可以产生转基因非人动物，其含有选择的允许调节转基因表达的系统。该系统的一个例子是噬菌体 P1 的 *cre/loxP* 重组酶系统。关于 *cre/loxP* 重组酶系统的描述，参见，例如 Lakso 等人 (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236。重组酶系统的另一个例子是酿酒酵母 FLP 重组酶系统 (O'Gorman 等人, 1991, *Science* 251:1351-1355)。如果利用 *cre/loxP* 重组酶系统调节转基因的表达，则需要含有编码 *Cre* 重组酶和选定蛋白质这两者的转基因的动物。可以通过构建“双”转基因动物提供这样的动物，例如通过使两个转基因动物交配，其中一个含有编码选定蛋白质的转基因，另一个含有编码重组酶的转基因。

本文所述的非人转基因动物的克隆也可以按照 Wilmut 等人 (1997) *Nature* 385:810-813 和 PCT 公布号 WO 97/07668 和 WO 97/07669 所述的方法产生。

药物组合物

本发明的核酸分子、多肽和抗体（在此也被称为“活性化合物”）可以掺入适于施用的药物组合物中。该组合物一般含有核酸分子、蛋白质或抗体和药学可接受的载体。此处使用的短语“药学可接受的载体”包括与给药相容的任何和所有溶剂、分散介质、包衣、抗细菌剂和抗真菌剂、等渗剂和吸收延迟剂，等等。这些介质和药剂用于药学活性物质的用途在本领域公知。除非任一常规介质或药剂与活性化合物不相容，涉及其在组合物中的应用。也可以在组合物中掺入补充活性化合物。

本发明包括制备用于调节标志物核酸或蛋白的表达或活性的药物组合物的方法。该方法包括将药学可接受的载体与调节标志物核酸或蛋白的表达或活性的药剂进行配制。该组合物可进一步包含其他活性剂。因此，本发明进一步包括通过将药学可接受的载体与调节标志物核酸或蛋白的表达或活性的药剂和一种或多种其他活性化合物进行配制而制备药物组合物的方法。

本发明也提供用于鉴定调节剂，即候选或试验化合物或药剂（例如肽、拟肽、类肽、小分子或其他药物）的方法（在此也称为“筛选试验”），所述调节剂：(a)能与标志物结合，或者(b)对标志物的活性具有调节（例如刺激或抑制）作用，或者更特别地，(c)对标志物与一种或多种其天然底物（例如肽、蛋白质、激素、辅因子或核酸）的相互作用具有调节作用，或者(d)对标志物的表达具有调节作用。该试验一般包括标志物与一种或多种测定成分的反应。其他成分可以是试验化合物本身，或试验化合物与标志物的天然结合配偶体的组合。

本发明的试验化合物可以从任何可获得的来源中获得，包括天然和/或合成化合物的系统文库。试验化合物也可以利用本领域公知的任何一种组合文库方法而获得，包括：生物学文库；类肽文库(分子文库，该分子具有肽的功能，但是具有新的非肽骨架，其耐受酶的降解，但是仍然具有生物活性；参见，例如，Zuckermann 等人，1994, *J. Med. Chem.* 37:2678-85)；可空间定位的平行固相或溶液相文库；需要重叠

合 (deconvolution) 的合成文库方法; “一珠一化合物”文库法; 和采用亲和层析筛选的合成文库法。生物文库和类肽文库法仅限于肽库, 而其他四种方法适于肽库、非肽寡聚体文库或小分子化合物文库(Lam, 1997, *Anticancer Drug Des.* 12:145)。

用于合成分子文库的方法的例子在本领域中可见于, 例如: DeWitt 等人 (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb 等人 (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann 等人 (1994). *J. Med. Chem.* 37:2678; Cho 等人 (1993) *Science* 261:1303; Carrell 等人 (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell 等人 (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; 和 Gallop 等人 (1994) *J. Med. Chem.* 37:1233。

化合物文库可以存在于溶液中 (例如 Houghten, 1992, *Biotechniques* 13:412-421) 或珠 (Lam, 1991, *Nature* 354:82-84)、芯片 (Fodor, 1993, *Nature* 364:555-556)、细菌和/或孢子 (Ladner, USP 5,223,409)、质粒 (Cull 等人, 1992, *Proc Natl Acad Sci USA* 89:1865-1869) 或噬菌体 (Scott 和 Smith, 1990, *Science* 249:386-390; Devlin, 1990, *Science* 249:404-406; Cwirla 等人, 1990, *Proc. Natl. Acad. Sci.* 87:6378-6382; Felici, 1991, *J. Mol. Biol.* 222:301-310; Ladner, 同上) 上。

在一个实施方案中, 本发明提供用于筛选候选或试验化合物的试验, 该化合物是一种标志物或其生物活性部分所编码的或对应的蛋白质的底物。在另一个实施方案中, 本发明提供用于筛选候选或试验化合物的试验, 该化合物与一种标志物或其生物活性部分所编码的或对应的蛋白质结合。确定试验化合物直接结合一种蛋白质的能力可以如下实现, 例如, 通过将该化合物与放射性同位素或酶标记物偶联, 使该化合物与标志物的结合可以通过检测复合物中的标记的标志物化合物来测定。例如, 化合物 (例如标志物底物) 可以用 ^{125}I 、 ^{35}S 、 ^{14}C 或 ^3H 直接或间接标记, 通过直接计数放射性发射或通过闪烁计数来检测放射性同位素。或者, 测定成分可以用例如辣根过氧化物酶、碱性磷

酸酶或萤光素酶进行酶标记，通过测定适当底物向产物的转化检测酶标记。

在另一个实施方案中，本发明提供用于筛选候选或试验化合物的试验，该化合物调节标志物的表达或标志物或其生物活性部分所编码的或对应的蛋白质的活性。在所有的可能性中，该标志物所编码的或对应的蛋白质能够在体内与一种或多种分子相互作用，这些分子例如但不限于肽、蛋白质、激素、辅因子和核酸。为进行描述，这些细胞和细胞外分子在此被称为“结合配偶体”或标志物“底物”。

为了有利于这种筛选，本发明的一个必要的实施方案是使用标志物所编码的或对应的蛋白质鉴定蛋白质的天然体内结合配偶体。实现这一点可使用本领域技术人员公知的多种方法。一个例子是在双杂交试验或三杂交试验中使用标志物蛋白作为“饵蛋白”(参见，例如，美国专利 No. 5,283,317; Zervos 等人, 1993, *Cell* 72:223-232; Madura 等人, 1993, *J. Biol. Chem.* 268:12046-12054; Bartel 等人, 1993, *Biotechniques* 14:920-924; Iwabuchi 等人, 1993 *Oncogene* 8:1693-1696; Brent WO94/10300)，以鉴定与该标志物结合或相互作用（结合配偶体），因此可能与该标志物的天然功能有关的其他蛋白质。该标志物结合配偶体也可能参与通过标志物蛋白或标志物蛋白介导的信号传导途径的下游元件的信号传送。此外也可能发现该标志物蛋白结合配偶体是标志物蛋白的抑制剂。

双杂交系统基于大多数转录因子的模块性质，由可分开的 DNA 结合域和活化域组成。简而言之，该试验使用两种不同的 DNA 构建体。在一种构建体中，编码标志物蛋白的基因与编码已知转录因子（例如 GAL-4）的 DNA 结合域的基因融合。在另一构建体中，来自 DNA 序列文库的编码未确定蛋白质（“猎物”或“样品”）的 DNA 序列与编码已知转录因子活化域的基因融合。如果“诱饵”和“猎物”蛋白能够在体内相互作用，形成依赖于标志物的复合物，则该转录因子的 DNA 结合域和活化域紧密接近。这种接近性允许与转录因子反应性的转录调节位点可操作连接的报道基因(例如 LacZ)转录。报道基因的表

达能够容易地检测，含有功能转录因子的细胞集落能够分离，并且用来获得编码与标志物蛋白相互作用的蛋白质的克隆基因。

在另一个实施方案中，可以利用本发明设计试验，用于鉴定调节（例如正或负影响）标志物蛋白与其底物和/或结合配偶体之间相互作用的化合物。这样的化合物包括但不限于诸如抗体、肽、激素、寡核苷酸、核酸及其类似物的分子。这样的化合物也可以从任何可获得的来源获得，包括天然和/或合成化合物的系统文库。用于该实施方案的优选试验成分是此处鉴定的前列腺癌标志物蛋白，其已知的结合配偶体和/或底物，和试验化合物。试验化合物可以来自任何来源。

用来鉴定干扰标志物蛋白与其结合配偶体之间相互作用的化合物的试验系统的基本原则包括在足以使两种产物相互作用和结合的条件下及时间里制备含有该标志物蛋白及其结合配偶体的反应混合物，从而形成复合物。为了检测药剂的抑制活性，在存在及不存在试验化合物的情况下制备反应混合物。试验化合物可以在开始时即包含在反应混合物中，或者可以在添加标志物蛋白及其结合配偶体之后加入。对照反应混合物不含试验化合物或者与安慰剂一起温育。然后检测标志物蛋白与其结合配偶体之间任何复合物的形成。若在对照反应中形成复合物，但在含有试验化合物的反应混合物中不形成或者形成很少，则表明该化合物干扰标志物蛋白与其结合配偶体的相互作用。相反，若在化合物存在下比在对照反应中形成更多复合物，则表明该化合物可增强标志物蛋白与其结合配偶体的相互作用。

对干扰标志物蛋白与其结合配偶体相互作用的化合物的测定可以以非均相或均相形式进行。非均相测定包括将标志物蛋白或其结合配偶体锚定在固相上，在反应结束时检测锚定在固相上的复合物。在均相测定中，整个反应都在液相中进行。在每一种方法中，为了获得关于所测化合物的不同信息，反应物的添加顺序可以不同。例如，通过在试验物质的存在下进行反应，即通过在标志物及其反应性结合配偶体之前或同时将试验物质加入反应混合物中，来鉴定干扰（例如通过竞争）标志物蛋白与其结合配偶体的相互作用的试验化合物。或者，

可以通过在形成复合物之后将试验化合物加入反应混合物中，来检测破坏预先形成的复合物的试验化合物，例如具有替换复合物中一种成分的较高结合常数的化合物。各种形式在下面简述。

在一种非均相测定系统中，标志物蛋白或其结合配偶体锚定于固体表面或基质上，而另一相应的非锚定成分可以被直接或间接标记。在实际中，该方法通常使用微量滴定板。可以利用许多方法非共价或共价固定锚定的物质，这些方法一般是实施该技术的人员所公知的。非共价连接通常可以通过用标志物蛋白或其结合配偶体的溶液包被固体表面并且干燥而简单地实现。另外，对于将要锚定的测定成分特异性的固定抗体也可以用于该目的。所述表面通常可以提前准备并贮存起来。

在相关实施方案中，提供一种融合蛋白，该融合蛋白增加了允许一种或两种测定成分锚定在基质上的结构域。例如，谷胱甘肽-S-转移酶/标志物融合蛋白或谷胱甘肽-S-转移酶/结合配偶体可以吸附到谷胱甘肽 sepharose 珠(Sigma Chemical, St. Louis, MO)或谷胱甘肽衍化的微量滴定板上，然后与试验化合物，或者与试验化合物和未吸附的标志物或其结合配偶体组合，该混合物在有利于复合物形成的条件（例如生理条件）下温育。温育后，洗涤珠或微量滴定板的孔，除去任何未结合的测定成分，直接或间接评估固定的复合物，例如如上所述。或者，复合物也可以从基质上解离下来，并用标准技术测定标志物的结合或活性水平。

在本发明的筛选试验中也可以使用其他用于将蛋白质固定在基质上的技术。例如，标志物蛋白或标志物蛋白结合配偶体可以利用生物素和链霉亲和素的偶联来固定。生物素化标志物蛋白或靶分子可以利用本领域公知的技术（例如生物素化试剂盒，Pierce Chemicals, Rockford, IL）由生物素-NHS (N-羟基-琥珀酰亚胺)制备，并且固定在链霉亲和素包被的 96 孔板(Pierce Chemical)的孔中。在一些实施方案中，蛋白质固定的表面可以提前准备并贮存起来。

为了进行该试验，使固定的测定成分的相应配偶体在含有或不含

试验化合物的情况下暴露于包被的表面。反应结束后，除去（例如通过洗涤）未反应的测定成分，而形成的任何复合物仍将固定在固体表面上。锚定在固体表面上的复合物的检测可以用多种方法来完成。当预先标记未固定的成分时，检测到固定于表面上的标记则表明形成了复合物。当未固定的成分未预先标记时，可以利用间接标记检测锚定在该表面上的复合物；例如，使用对于开始时未固定的物质具有特异性的标记抗体（该抗体可以直接标记或者利用例如标记的抗 Ig 抗体间接标记）。根据反应成分的添加顺序，可以检测调节（抑制或增强）复合物的形成或破坏预先形成的复合物的试验化合物。

在本发明的一个替代实施方案中，可以采用均相测定。这一般是同以上所述类似的反应，其在存在或不存在试验化合物的液相中进行。然后将形成的复合物与未反应的成分分离，测定形成的复合物的量。如非均相测定系统所述，向液相中添加反应物的顺序可以获得关于哪种试验化合物调节（抑制或增强）复合物的形成以及哪种破坏预先形成的复合物的信息。

在这样的均相测定中，可以利用大量标准技术之一将反应产物与未反应的测定成分分离，这些技术包括但不限于：差速离心、层析法、电泳和免疫沉淀。在差速离心中，由于复合物根据其大小和密度不同而具有不同的沉降平衡，因此通过一系列离心步骤，分子的复合物可以与未复合的分子分离（参见，例如 Rivas, G.和 Minton, A.P., *Trends Biochem Sci* 1993 Aug;18(8):284-7）。也可以利用标准层析技术将复合的分子与未复合的分子分离。例如，凝胶过滤层析利用适当的凝胶过滤树脂以柱的形式基于大小分离分子，例如相对较大的复合物可以与相对较小的未复合的成分分离。类似地，可以利用复合物与未复合分子相比相对不同的电荷性质将该复合物与剩余的各种反应物差异分离，例如，使用离子交换层析树脂。这些树脂和层析技术是本领域技术人员公知的（参见，例如，Heegaard, 1998, *J Mol. Recognit.* 11:141-148; Hage 和 Tweed, 1997, *J. Chromatogr. B. Biomed. Sci. Appl.*, 699:499-525）。也可以利用凝胶电泳将复合的分子与未结合分子分

离(参见, 例如, Ausubel 等人(编), *Current Protocols in Molecular Biology*, J. Wiley & Sons, New York. 1999)。在该技术中, 例如根据大小或电荷分离蛋白质或核酸复合物。为了在电泳过程中保持结合相互作用, 通常优选不含还原剂的非变性凝胶, 但是适合特定相互作用物的条件是本领域技术人员公知的。免疫沉淀法是用从溶液中分离蛋白质-蛋白质复合物的另一种常用技术(参见, 例如, Ausubel 等人(eds.), In: *Current Protocols in Molecular Biology*, J. Wiley & Sons, New York. 1999)。在该技术中, 通过将一种结合分子的特异性抗体与易于离心收集的聚合物珠相偶联, 从溶液中沉淀与该抗体结合的所有蛋白质。从珠上释放结合的测定成分(通过特异性蛋白质水解或本领域周知的不干扰复合物中蛋白质-蛋白质相互作用的其他技术), 进行第二个免疫沉淀步骤, 这次使用对相应不同相互作用的测定成分具有特异性的抗体。这样, 只有形成的复合物保持连接到珠上。可以比较在存在及不存在试验化合物条件下复合物形成的变化, 从而提供关于该化合物调节标志物蛋白与其结合配偶体间相互作用的能力的信息。

本发明的范围内也包括不需进一步样品处理即在均相或非均相测定系统中直接检测标志物蛋白与其结合配偶体和/或试验化合物之间相互作用的方法。例如, 可以采用荧光能量转移技术(参见, 例如 Lakowicz 等人的美国专利 No. 5,631,169; Stavrianopoulos 等人的美国专利 No. 4,868,103)。该技术通常包括在第一“供体”分子(例如标志物或试验化合物)上添加荧光团标记, 使其发射的荧光能量被第二“受体”分子(例如标志物或试验化合物)上的荧光标记吸收, 第二“受体”分子又由于吸收了能量而能够发荧光。或者, “供体”蛋白质分子可以只利用色氨酸残基的天然荧光能量。选择发出不同波长的光的标记, 使“受体”分子标记可以与“供体”分子标记区别开来。由于标记之间的能量转移效率与分子间距有关, 因此能够估测分子之间的空间关系。在分子之间发生结合的情况下, 试验中“受体”分子标记的荧光发射应当最大。可以通过本领域公知的标准荧光检测手段(例如使用荧光计)常规测定 FET 结合事件。能增强或阻碍一种物质加入预先形成的

复合物中的试验物质将导致产生不同于背景的信号。这样，可以用受控制的试验鉴定能调节标志物与其结合配偶体之间的相互作用的试验物质。

在另一个实施方案中，用一种方法鉴定标志物表达的调节剂，在该方法中，使细胞与一种候选化合物接触，并且测定细胞中标志物 mRNA 或蛋白质的表达。将在该候选化合物存在下标志物 mRNA 或蛋白质的表达水平与不存在该候选化合物条件下标志物 mRNA 或蛋白质的表达水平进行比较。然后可以基于这种比较将该候选化合物鉴定为标志物表达的调节剂。例如，当在该候选化合物存在下标志物 mRNA 或蛋白质的表达高于（统计学上显著性地高于）不存在该候选化合物的情况时，该候选化合物被鉴定为标志物 mRNA 或蛋白质表达的刺激剂。相反，当在该候选化合物存在下标志物 mRNA 或蛋白质的表达低于（统计学上显著性地低于）不存在该候选化合物的情况时，该候选化合物被鉴定为标志物 mRNA 或蛋白质表达的抑制剂。细胞中标志物 mRNA 或蛋白质的表达水平可以用此处所述的用于检测标志物 mRNA 或蛋白质的方法来测定。

在另一方面，本发明涉及两种或多种本文所述试验的组合。例如，可以利用基于细胞的试验或不含细胞的试验鉴定调节剂，并且调节剂调节标志物蛋白的活性的能力可以进一步在体内证实，例如在用于细胞转化和/或肿瘤发生的完整动物模型中证实。

本发明进一步涉及通过上述筛选试验鉴定的新药剂。因此，本发明的范围内包括在合适的动物模型中进一步使用如本文所述鉴定的药剂。例如，如此处所述鉴定的药剂（例如标志物调节剂，反义标志物核酸分子，标志物特异性抗体，或标志物结合配偶体）可以在动物模型中用来确定使用该药剂治疗的效果、毒性或副作用。另外，如此处所述鉴定的药剂也可以在动物模型中用于确定该药剂的作用机理。另外，本发明涉及通过上述筛选试验鉴定的新药剂用于此处所述治疗的用途。

应当理解，小分子药物和蛋白质或多肽剂的适当剂量取决于普通

医生、兽医或研究人员所知的许多因素。例如，根据所治疗的受试者或样品的特性、大小和状况，如果可行，进一步根据组合物的施用途径，以及医生所希望的该药剂对本发明核酸或多肽的作用，这些药剂的剂量将会变化。小分子剂量的例子包括每千克受试者或样品重量毫克或微克的量（例如约每千克 1 微克至约每千克 500 毫克，约每千克 100 微克至约每千克 5 毫克，或者约每千克 1 微克至约每千克 50 微克）。蛋白质或多肽的剂量的例子包括每千克受试者或样品重量克、毫克或微克的量（例如约每千克 1 微克至约每千克 5 克，约每千克 100 微克至约每千克 500 毫克，或者约每千克 1 毫克至约每千克 50 毫克）。还应当理解，一种这样的药剂的适当剂量取决于该药剂与调节表达或活性有关的能力。这种适当的剂量可以用此处所述的试验来确定。当对动物（例如人）施用一种或多种这样的药剂，以调节本发明的多肽或核酸的表达或活性时，医生、兽医或研究人员可以，例如，先给出较低剂量的处方，然后再提高剂量，直到获得合适的应答。另外，应当理解，用于任何特定受试动物的特定剂量水平将取决于多种因素，包括所用特定药剂的活性，受试者的年龄、体重、一般健康情况、性别和饮食，给药时间，给药途径，排泄率，任何药物联用，以及要调节的表达或活性的程度。

本发明的药物组合物配制成与预期给药途径相适应的形式。给药途径的例子包括肠胃外给药，例如静脉、皮内、皮下、经口（例如吸入）、经皮（局部）、经粘膜和直肠给药。用于肠胃外、皮内或皮下施用的溶液或悬浮液可包括如下成分：无菌稀释液，如注射用水、盐水溶液、非挥发油、聚乙二醇、甘油、丙二醇或其他合成溶剂；抗菌剂，如苯甲醇或对羟基苯甲酸甲酯；抗氧化剂，如抗坏血酸或亚硫酸氢钠；螯合剂，如乙二胺四乙酸；缓冲剂，如乙酸盐、柠檬酸盐或磷酸盐缓冲液，和用于调节张度的试剂，如氯化钠或葡萄糖。pH 可以用酸或碱，如盐酸或氢氧化钠来调节。肠胃外制剂可以包封在由玻璃或塑料制成的安瓿、一次性注射器或多剂量瓶中。

适合注射使用的药物组合物包括无菌水溶液（水溶性时）或分散

液，和用于即时配制无菌注射溶液或分散液的无菌粉末。对于静脉给药，合适的载体包括生理盐水、抑菌水、聚氧乙烯蓖麻油（Cremophor EL）（BASF; Parsippany, NJ）或磷酸盐缓冲液（PBS）。在所有情况下，组合物必须是无菌的，并且其流动性应当达到易于注射的程度。在生产和贮存条件下必须稳定，必须能够抵御微生物如细菌和真菌的污染作用。载体可以是溶剂或分散介质，其含有例如水、乙醇、多元醇（例如甘油、丙二醇和液体聚乙二醇等）及其适当的混合物。例如，通过使用包衣如卵磷脂，对于分散液的情况通过保持需要的颗粒大小，和通过使用表面活性剂，可以保持适当的流动性。微生物作用的防止可以用各种抗细菌剂和抗真菌剂来实现，例如对羧苯甲酸酯、三氯叔丁醇、苯酚、抗坏血酸、硫柳汞等。在许多情况下，组合物中优选包含等渗剂，例如糖、多元醇，如甘露醇、山梨醇，或氯化钠。注射组合物的延时吸收可以通过在该组合物中包含延迟吸收的药剂例如单硬脂酸铝和明胶来实现。

无菌注射溶液可以如下制备：将活性化合物（例如多肽或抗体）以所需的量加入适当的溶剂中，需要时加入以上列举的一种成分或成分组合，接下来过滤除菌。分散液通常如下制备：将活性化合物掺入含有基本分散介质的无菌载体中，然后掺入选自以上列举的所需的其他成分。对于用于制备无菌注射溶液的无菌粉末而言，优选的制备方法是真空干燥和冷冻干燥，该方法由预先无菌过滤的溶液产生活性成分加需要的任何其他成分的粉末。

口服组合物通常包含惰性稀释剂或可食用载体。它们可以包封在明胶胶囊中或者压制成片剂。为了治疗性口服给药，活性化合物可以与赋形剂掺合，并以片剂、糖锭或胶囊的形式使用。口服组合物也可以用一种流体载体制备用作漱口剂，其中该液体载体中的化合物经口施用、漱口（swish）和吐出或吞咽。

可以包含药学上相容的结合剂，和/或佐剂物质，作为组合物的一部分。片剂、丸剂、胶囊、糖锭等可以含有任何下列成分，或具有类似性质的化合物：粘合剂，如微晶纤维素、西黄蓍胶或明胶；赋形

剂，如淀粉或乳糖；崩解剂，如海藻酸、Primogel 或玉米淀粉；润滑剂，如硬脂酸镁或 Sterotes；助流剂，如二氧化硅胶体，甜味剂，如蔗糖或糖精；或调味剂，如薄荷、水杨酸甲酯或橙味调味剂。

为了通过吸入给药，化合物以喷雾剂的形式施用，该喷雾剂来自压力容器或含有适当的推进剂，例如气体，如二氧化碳的分液器或雾化器。

全身给药也可以通过经粘膜或经皮方式实现。为了经粘膜或经皮给药，在制剂中使用适合于所要渗透的障碍物的渗透剂。该渗透剂在本领域中公知，例如对于经粘膜给药而言，包括去污剂、胆汁盐和夫西地酸衍生物。经粘膜给药可以使用鼻腔喷剂或栓剂来实现。对于经皮给药而言，活性化合物配制成本领域公知的软膏、药膏、凝胶或乳膏。

该化合物也可以制备成用于直肠给药的栓剂（例如含有常规栓剂基质，如可可脂和其他甘油酯）或保留灌肠剂的形式。

在一个实施方案中，活性化合物用将保护化合物免遭机体快速清除的载体制备，例如控释制剂，包括埋植剂和微胶囊化的递送系统。可以使用生物可降解的、生物相容的聚合物，例如乙烯乙酸乙烯酯、聚酐、聚乙醇酸、胶原、聚正酯和聚乳酸。制备这些制剂的方法对于本领域技术人员来说是显而易见的。这些材料也可以商业上从 Alza Corporation 和 Nova Pharmaceuticals, Inc. 获得。脂质体悬浮液（包括向其中或其上掺入单克隆抗体的脂质体）也可以用作药学可接受的载体。它们可以按照本领域技术人员公知的方法制备，例如美国专利 No. 4,522,811 所述。

特别有利的是以剂量单位形式配制口服或肠胃外组合物，以利于给药和剂量的均匀性。此处使用的剂量单位形式是指适于作为单一剂量用于所治疗的受试者的物理上分散的单位；每单位含有预定量的活性化合物，该预定量是计算出能与需要的药物载体相结合产生理想的疗效的量。本发明的剂量单位形式的规格直接取决于活性化合物的独特性质和所要达到的特定疗效，以及在制备用于个体治疗的活性化合

物的技术中固有的限制。

对于抗体，优选剂量是 0.1 mg/kg 至 100 mg/kg 体重(通常 10 mg/kg 至 20 mg/kg)。如果抗体在脑中起作用，50 mg/kg 至 100 mg/kg 的剂量通常是合适的。一般来说，部分人抗体和完全人抗体在人体内比其他抗体具有更长的半衰期。因此，较低的剂量和较低频率的给药常常是可能的。可以利用诸如脂化的修饰来稳定抗体以及提高摄入和组织渗透(例如向前列腺上皮内渗透)。Cruikshank 等人 (1997) *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193 描述了一种脂化抗体的方法。

本发明也提供用于预防和/或治疗前列腺癌的疫苗组合物。本发明提供一种前列腺癌疫苗组合物，其中为了刺激针对前列腺癌的免疫应答，向受试者中导入表 1 中的标志物蛋白或表 1 的标志物蛋白的组合。本发明也提供一种前列腺癌疫苗组合物，其中向受试者中导入表达表 1 所列标志物或标志物片段的基因表达构建体，从而使受试者中被转染的细胞以高于正常的水平产生表 1 的标志物所编码的蛋白质或蛋白质片段，并且引发免疫应答。

在一个实施方案中，提供一种前列腺癌疫苗并将其作为用于预防前列腺癌的免疫治疗剂使用。在另一个实施方案中，提供一种前列腺癌疫苗并将其作为用于治疗前列腺癌的免疫治疗剂使用。

例如，一种由表 1 的标志物蛋白组成的前列腺癌疫苗可以通过经例如皮内、皮下或肌肉内的多种途径施用该疫苗用于预防和/或治疗受试者中的前列腺癌。另外，前列腺癌疫苗可以与佐剂和/或免疫调节剂一起施用，以加强疫苗的活性和受试者的应答。在一个实施方案中，可以将适合持续或间歇释放的含有该疫苗的装置和/或组合物植入体内，或者局部应用，以便相对较慢地向体内释放该物质。为了产生在消除癌症上将会更有效的应答，前列腺癌疫苗可以与能够改变所产生的免疫应答类型的免疫调节化合物一起引入。

在另一个实施方案中，通过注射到肌肉内，或者通过包衣到微粒上并且使用为将颗粒高速射入皮肤内而设计的装置，可以引入由表 1

中标志物的表达构建体组成的前列腺癌疫苗。受试者的细胞于是表达表 1 的标志物蛋白的蛋白质或片段，并且诱导免疫应答。另外，为了产生在消除癌症上将会更有效的应答，前列腺癌疫苗可以与免疫调节分子如细胞因子的表达构建体一起引入，该免疫调节分子可以增强免疫应答或调节所产生的免疫应答的类型。

标志物核酸分子可以插入载体中，并且用作基因治疗载体。基因治疗载体可以通过例如静脉注射、局部给药(美国专利 5,328,470)或立体定向注射(参见，例如，Chen 等人, 1994, *Proc. Natl. Acad. Sci. USA* 91:3054-3057)递送给受试者。基因治疗载体的药物制剂可包括在可接受的稀释剂中的基因治疗载体，或者可包含包埋有基因递送载体的缓释基质。或者，当能够从重组细胞中完整产生整个基因递送载体，例如逆转录病毒载体时，该药物制剂可包含一种或多种产生该基因递送系统的细胞。

药物组合物可以与使用说明一起包含在容器、包装或分配器中。

预测医学

本发明涉及预测医学领域，其中诊断试验、预后试验、药物基因组学和监测临床试验用于预后（预测）目的，从而预防性地处置个体。因此，本发明的一个方案涉及一种诊断试验，其用于测定一种或多种标志物蛋白或核酸的表达水平，从而确定个体是否具有发展前列腺癌的危险。该试验可以用于预后或预测目的，从而在癌症发病之前预防性地处置个体。

本发明的另一个方案涉及在临床试验中监测药剂（例如为了抑制前列腺癌或治疗或预防其他任何疾病{即为了理解这种治疗可能具有的任何致前列腺癌作用}而施用的药物或其他化合物）对本发明标志物的表达或活性的影响。这些及其他药剂在以下部分中进一步详细描述。

诊断试验

检测生物样品中是否存在标志物蛋白或核酸的方法的一个例子包

括从受试者中获得生物样品（例如前列腺相关体液），使该生物样品与能够检测该多肽或核酸（例如 mRNA、基因组 DNA 或 cDNA）的化合物或试剂接触。本发明的检测方法因此可以用于在体内及体外检测例如生物样品中的 mRNA、蛋白质、cDNA 或基因组 DNA。例如，用于检测 mRNA 的体外技术包括 Northern 杂交和原位杂交。用于检测标志物蛋白的体外技术包括酶联免疫吸附测定(ELISA)、蛋白质印迹分析、免疫沉淀和免疫荧光法。用于检测基因组 DNA 的体外技术包括 Southern 杂交。用于检测 mRNA 的体内技术包括聚合酶链反应（PCR）、Northern 杂交和原位杂交。此外，用于检测标志物蛋白的体内技术包括向受试者中导入抗该蛋白质或其片段的标记抗体。例如，抗体可以用一种放射性标记进行标记，该标记在受试者中的存在和定位可以通过标准成像技术来检测。

这些诊断和预后试验的一个共同原则包括制备可能含有标志物的样品或反应混合物，和探针，在适当的条件下和足够的时间里使标志物和探针相互作用及结合，从而在反应混合物中形成能够取出和/或检测的复合物。这些试验可以用多种方式进行。

例如，一种进行这种测定的方法包括将标志物或探针锚定在固相支持体上，该固相支持体也被称为基质，并且在反应结束时检测锚定在固相上的靶标志物/探针复合物。在这种方法的一个实施方案中，将要测定标志物的存在和/或浓度的来自受试者的样品可以锚定在载体或固相支持体上。在另一个实施方案中，可以是相反的情况，即其中探针可以锚定在固相上，而来自受试者的样品可以作为未锚定的测定成分进行反应。

有多种已建立的方法用于将测定成分锚定在固相上。包括但不限于通过生物素和链霉亲和素偶联而固定的标志物或探针分子。这些生物素化的测定成分可以利用本领域公知的方法（例如生物素化试剂盒，Pierce Chemicals, Rockford, IL）由生物素-NHS（N-羟基-琥珀酰亚胺）制备，并且固定在链霉亲和素包被的 96 孔板(Pierce Chemical)的孔中。在某些实施方案中，具有固定的测定成分的表面可以提前制备

好并贮存起来。

用于这些试验的其他合适的载体或固相支持体包括能够与标志物或探针所属的分子类别结合的任何材料。众所周知的支持体或载体包括但不限于玻璃、聚苯乙烯、尼龙、聚丙烯、尼龙、聚乙烯、葡聚糖、淀粉酶、天然及改性纤维素、聚丙烯酰胺、辉长岩和磁铁矿石。

为了用上述方法进行试验，将未固定的成分加到上面锚定了第二成分的固相上。反应结束后，可以在使得形成的任何复合物将保持固定在固相上的条件下除去（例如通过洗涤）未复合的成分。锚定在固相上的标志物/探针复合物的检测可以用此处所述的许多方法进行。

在一个优选实施方案中，当探针是未锚定的测定成分时，可以为了检测和试验的读数目的而用此处所述的及本领域技术人员公知的可检测标记直接或间接标记该探针。

也可以直接检测标志物/探针复合物的形成，而无需进一步操作或标记每种成分（标志物或探针），例如，利用荧光能量转移技术（参见，例如，Lakowicz 等人的美国专利 No. 5,631,169; Stavrianopoulos 等人的美国专利 No. 4,868,103）。选择第一“供体”分子上的荧光团标记，使其在用适当波长的入射光激发后，发射的荧光能量将被第二“受体”分子上的荧光标记吸收，第二“受体”分子由于吸收了能量又能够发出荧光。或者，“供体”蛋白质分子可以只利用色氨酸残基的天然荧光能量。选择发出不同波长的光的标记，使“受体”分子标记可以与“供体”分子标记区别开来。由于标记之间的能量转移效率与分子间距有关，因此能够估测分子之间的空间关系。在分子之间发生结合的情况下，试验中“受体”分子标记的荧光发射应当最大。可以通过本领域公知的标准荧光检测手段（例如使用荧光计）常规测定 FET 结合事件。

在另一个实施方案中，可以不标记每种测定成分（探针或标志物），而利用诸如实时生物分子相互作用分析(BIA)（参见，例如，Sjolander, S.和 Urbaniczky, C., 1991, *Anal. Chem.* 63:2338-2345, 和 Szabo 等人, 1995, *Curr. Opin. Struct. Biol.* 5:699-705)等技术来确定探

针识别标志物的能力。如此处使用的“BIA”或“表面等离子共振（surface plasmon resonance）”是用于实时研究生物特异性相互作用的技术，无需标记任何相互作用物（例如 BIAcore）。结合表面的质量的改变（指示结合事件）导致靠近表面的光折射率改变（表面等离子共振（SPR）的光学现象），产生可检测的信号，该信号可以用作生物分子之间实时反应的指示。

此外，在另一个实施方案中，可以用标志物和探针作为液相中的溶质进行类似的诊断和预后试验。在这种试验中，利用许多标准技术中的任一种将复合的标志物和探针与未复合的成分分离开来，这些标准技术包括但不限于：差速离心、层析、电泳和免疫沉淀。在差速离心中，由于复合物根据不同的大小和密度而有不同的沉降平衡，因此可以通过一系列离心步骤将标志物/探针复合物与未复合的测定成分分离开来（参见，例如，Rivas, G.和 Minton, A.P., 1993, *Trends Biochem Sci.* 18(8):284-7）。标准层析技术也可以用来将复合的分子与未复合的分子分离开来。例如，凝胶过滤层析法通过以柱形式使用适当的凝胶过滤树脂基于大小分离分子，例如，相对较大的复合物可以与相对较小的未复合的成分相分离。类似地，可以利用标志物/探针复合物与未复合的成分相比相对不同的电荷性质来区别复合物与未复合的成分，例如通过使用离子交换层析树脂。这些树脂和层析技术是本领域技术人员公知的（参见，例如，Heegaard, N.H., 1998, *J. Mol. Recognit.* Winter 11(1-6):141-8; Hage, D.S.和 Tweed, S.A. *J Chromatogr B Biomed Sci Appl* 1997 Oct 10;699(1-2):499-525）。也可以利用凝胶过滤将复合的测定成分与未结合的成分分离（参见，例如 Ausubel 等人, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1987-1999）。在该技术中，例如基于大小或电荷分离蛋白质或核酸复合物。为了在电泳过程中保持结合的相互作用，一般优选不含还原剂的非变性凝胶基质材料和条件。用于特定试验及其成分的适当条件是本领域技术人员公知的。

在一个特定实施方案中，可以使用本领域公知的方法通过原位和

体外形式测定生物样品中标志物 mRNA 的水平。术语“生物样品”包括从受试者中分离的组织、细胞、生物液体及其分离物，以及存在于受试者体内的组织、细胞和液体。有许多表达检测方法使用分离的 RNA。对于体外方法而言，任何不是针对分离 mRNA 而选择的 RNA 分离技术都可以用于从前列腺细胞中纯化 RNA (参见，例如，Ausubel 等人, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York 1987-1999)。另外，能够用本领域技术人员公知的技术，例如 Chomczynski (1989, 美国专利 No. 4,843,155) 的一步 RNA 分离法来容易地处理大量组织样品。

分离的 mRNA 可以在杂交或扩增试验中使用，这些试验包括但不限于 Southern 或 Northern 分析、聚合酶链反应分析和探针阵列。用于检测 mRNA 水平的一种优选诊断方法包括使分离的 mRNA 与一种核酸分子（探针）接触，该核酸分子能够与被检测的基因所编码的 mRNA 杂交。该核酸分子探针可以是，例如，全长 cDNA 或其部分，如长度为至少 7、15、30、50、100、250 或 500 个核苷酸且在严格条件下足以与编码本发明标志物的 mRNA 或基因组 DNA 特异性杂交的寡核苷酸。此处描述了其他适合在本发明的诊断试验中使用的探针。mRNA 与探针杂交表明所述标志物得到表达。

在一种形式中，将 mRNA 固定在固体表面上，并且与探针接触，例如通过在琼脂糖凝胶上电泳分离的 mRNA，并将 mRNA 从凝胶转移到诸如硝酸纤维素膜的膜上。在一种替代形式中，将探针固定在固体表面上，mRNA 与例如 Affymetrix 基因芯片阵列中的探针接触。本领域技术人员能够容易地改变已知的 mRNA 检测方法，以用于检测本发明标志物编码的 mRNA 的水平。

测定样品中 mRNA 标志物水平的一种替代方法包括如下步骤：核酸扩增，例如通过 RT-PCR (Mullis, 1987, 美国专利 No. 4,683,202 中所述的实验实施方案)，连接酶链反应(Barany, 1991, *Proc. Natl. Acad. Sci. USA*, 88:189-193)，自动维持序列扩增(Guatelli 等人, 1990, *Proc. Natl. Acad. Sci. USA* 87:1874-1878)，转录扩增系统(Kwoh 等人, 1989,

Proc. Natl. Acad. Sci. USA 86:1173-1177), Q-Beta 复制酶(Lizardi 等人, 1988, *Bio/Technology* 6:1197), 滚环复制(Lizardi 等人的美国专利 No. 5,854,033), 或其他任何核酸扩增方法, 之后使用本领域技术人员公知的技术检测扩增的分子。这些检测方案尤其可用于检测以极低量存在的核酸分子。此处使用的扩增引物被定义为一对核酸分子, 它可以与基因的 5'或 3'区退火(分别是正链和负链, 或者相反), 并且在之间含有一个短区。通常, 扩增引物的长度为约 10 到 30 个核苷酸, 并且位于长度约 50 到 200 个核苷酸的区域之侧。在适当的条件下及使用适当的试剂, 这些引物允许扩增包含侧翼为该引物的核苷酸序列的核酸分子。

对于原位方法而言, mRNA 不需要在检测前从前列腺细胞中分离出来。在这些方法中, 利用已知的组织学方法制备/处理细胞或组织样品。然后将样品固定于支持物上, 一般是载玻片, 然后使其接触能够与编码该标志物的 mRNA 杂交的探针。

作为根据标志物的绝对表达水平进行确定的替代方法, 也可以根据标准化的标志物表达水平进行确定。通过将其表达与不是标志物的基因, 例如组成型表达的管家基因的表达进行比较, 修正标志物的绝对表达水平, 将表达水平标准化。用于标准化的合适的基因包括管家基因, 如肌动蛋白基因, 或上皮细胞特异性的基因。这种标准化允许将一种样品例如患者样品中的表达水平与另一种样品例如非前列腺癌样品中的表达水平进行比较, 或者在来自不同来源的样品之间进行比较。

另外, 表达水平也可以作为相对表达水平来提供。为了确定标志物的相对表达水平, 在测定所述样品的表达水平之前, 测定正常细胞相对于癌细胞分离物的各 10 个或更多样品中, 优选 50 个或更多样品中标志物的表达水平。确定在大量样品中测定的每种基因的平均表达水平, 用其作为该标志物的基线表达水平。然后将测定的试验样品中标志物的表达水平(绝对表达水平)除以获得的该标志物的平均表达值。从而得到相对表达水平。

优选地，在基线测定中使用的样品来自于前列腺癌或来自于前列腺组织的非前列腺癌细胞。细胞来源的选择取决于相对表达水平的使用。利用在正常组织中发现的表达作为平均表达得分有助于确认测定的标志物是否是前列腺特异性的（相对于正常细胞）。另外，随着更多数据的积累，可以修正平均表达值，提供基于积累的数据修正的相对表达值。来自前列腺细胞的表达数据提供用于对前列腺癌状态的严重程度进行分级的手段。

在本发明的另一个实施方案中，检测标志物蛋白。用于检测本发明标志物蛋白的优选试剂是能够结合该蛋白质或其片段的抗体，优选具有可检测标记的抗体。抗体可以是多克隆抗体，或者更优选地，可以是单克隆抗体。可以使用完整的抗体，或其片段或衍生物（例如 Fab 或 $F(ab')_2$ ）。术语“标记”用于探针或抗体时包括通过偶联（即物理连接）可检测物质与探针或抗体来直接标记探针或抗体，以及通过与直接标记的另一试剂反应来间接标记探针或抗体。间接标记的例子包括利用荧光标记的第二抗体检测第一抗体，以及用生物素末端标记 DNA 探针，使其能够用荧光标记的链霉亲和素检测。

来源于前列腺细胞的蛋白质可以用本领域公知的技术来分离。采用的蛋白质分离方法可以是，例如 Harlow 和 Lane (Harlow 和 Lane, 1988, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York)所述的方法。

可以利用多种形式确定一种样品是否含有能够结合特定抗体的蛋白质。这类形式的例子包括但不限于酶免疫测定(EIA)、放射免疫测定(RIA)、蛋白质印迹分析和酶联免疫吸附测定(ELISA)。本领域技术人员能够容易地修改已知的蛋白质/抗体检测方法，用于确定前列腺细胞是否表达本发明的标志物。

在一种形式中，抗体或抗体片段或衍生物可以在诸如蛋白质印迹或免疫荧光技术等方法中使用，用来检测表达的蛋白质。在这些应用中，通常优选将抗体或蛋白质固定在固体支持体上。合适的固相支持体或载体包括能够结合抗原或抗体的任何支持体。众所周知的支持体

或载体包括玻璃、聚苯乙烯、聚丙烯、聚乙烯、葡聚糖、尼龙、淀粉酶、天然及改性纤维素、聚丙烯酰胺、辉长岩和磁铁矿石。

本领域技术人员应当知道许多其他的用于结合抗体或抗原的合适的载体，并且能够将这些支持体用于本发明。例如，从前列腺细胞中分离的蛋白质可以进行聚丙烯酰胺凝胶电泳，并固定在固相支持体如硝酸纤维素上。然后可以用适当的缓冲液洗涤该支持体，接着用可检测标记的抗体处理。然后可以用缓冲液第二次洗涤该固相支持体，以除去未结合的抗体。然后可以用常规方法检测该固相支持体上结合的标记的量。

本发明也包括用于检测生物样品中标志物蛋白或核酸的存在的试剂盒。该试剂盒能够用来确定受试者是否患有前列腺癌或者具有提高的发展前列腺癌的危险。例如，该试剂盒可包括能够检测生物样品中的标志物蛋白或核酸的标记化合物或试剂，以及用于确定样品中蛋白质或 mRNA 的量的工具(例如，能够结合该蛋白质或其片段的抗体，或能够结合编码该蛋白质的 DNA 或 mRNA 的寡核苷酸探针)。试剂盒也可包括用于说明使用该试剂盒获得的结果的说明书。

对于基于抗体的试剂盒，该试剂盒可包括，例如：(1) 能够结合标志物蛋白的第一抗体（例如连接到固体支持体上）；和任选地，(2) 不同的第二抗体，其能够结合该蛋白质或第一抗体，并且与可检测标记偶联。

对于基于寡核苷酸的试剂盒，该试剂盒可包括，例如：(1)寡核苷酸，例如可检测标记的寡核苷酸，它与编码标志物蛋白的核酸序列杂交，或(2)用于扩增标志物核酸分子的一对引物。该试剂盒也可包括，例如缓冲剂、防腐剂或蛋白质稳定剂。该试剂盒可进一步包括用于检测可检测标记所必需的成分（例如酶或底物）。该试剂盒也可含有可以测定并与受试样品进行比较的对照样品或一系列对照样品。该试剂盒的每种成分都可以包含在各自的容器内，所有各容器都可以与用于说明使用该试剂盒进行试验的结果的说明书一起置于一个包装内。

药物基因组学和药效学

本发明的标志物也可用作药物基因组学标志物，这里所用的“药物基因组学标志物”指客观的生化标志物，其表达水平与患者的特定临床药物反应或易感性相关（参见，例如，McLeod 等人 (1999) *Eur. J. Cancer* 35(12): 1650-1652）。药物基因组学标志物表达的存在或量与预测的患者，更具体地患者的肿瘤，对特定药物或药物种类治疗的反应有关。通过评估患者中的一个或多个药物基因组学标志物的表达的存在或量，可以选出最适合于该患者的或预计有更大成功程度的药物疗法。例如，根据患者中特定肿瘤标志物编码的 RNA 或蛋白质的存在或数量，可以选择最优化的用于治疗患者中可能存在的特定肿瘤的药物或疗程。因此应用药物基因组学标志物可以在不用尝试不同药物或方案的情况下，对每个癌症患者选择出或设计出最合适的疗法。

药物基因组学的另一方面涉及改变身体对药物的作用方式的遗传状况。这些药理遗传状况可以作为罕见的缺陷或作为多态性发生。例如，葡萄糖-6-磷酸脱氢酶(G6PD)缺陷是一种常见的遗传性酶病，其主要的临床并发症是摄入氧化剂药物（抗疟药、磺胺类药物、镇痛药、硝基呋喃类药物）及食用蚕豆后溶血。

作为一个说明性的实施方案，药物代谢酶的活性是药物作用强度和持续时间的一个主要决定因素。药物代谢酶（例如 N-乙酰转移酶 2 (NAT 2) 和细胞色素 P450 酶 CYP2D6 和 CYP2C19）的遗传多态性的发现已经对为什么一些患者无法获得预期的药物效果或者在使用标准且安全剂量的药物后显示过度的药物应答和严重的毒性提供了一种解释。这些多态性在人群中以两种表型表达，强代谢者（EM）和弱代谢者（PM）。PM 的普遍程度在不同群体中不同。例如，编码 CYP2D6 的基因具有高度多态性，并且在 PM 中已经鉴定出几个突变，这些突变均导致功能性 CYP2D6 的缺乏。CYP2D6 和 CYP2C19 的弱代谢者当接受标准剂量时非常经常地表现出过度的药物应答和副作用。如果代谢物是活性治疗性部分，则 PM 将不表现治疗响应，正如通过其 CYP2D6-形成的代谢物吗啡介导的可待因的镇痛作用所证实的。另一极端情况是所谓的超快速代谢者，他们对标准剂量没有响

应。最近，超快速代谢的分子基础已经被确定为是由于 CYP2D6 基因扩增。

因此，可以确定个体中本发明标志物的表达水平，从而选择用于治疗或预防性处理个体的合适的药物。另外，药物遗传学研究也能够用来将编码药物代谢酶的多态性等位基因的基因分型应用于鉴定个体的药物响应表型。当用本发明标志物的表达的调节剂处理受试者时，这些知识在用于选择剂量或药物时可以避免不良反应或治疗失败，因而提高了治疗或预防的效果。

本发明的标志物可以作为一种或多种紊乱或疾病状态或可导致疾病状态特别是前列腺癌的状况的替代标志物。此处使用的“替代标志物”是一种客观生化标志物，它与是否患有疾病或紊乱有关，或者与疾病或紊乱的进展有关（例如与肿瘤的存在与否有关）。该标志物的存在或量独立于疾病。因此，这些标志物可用来说明特定治疗过程是否可有效缓解疾病状态或紊乱。当疾病状态或紊乱的存在或程度难以用标准方法评估时（例如早期肿瘤），或者当在达到可能危险的临床终点之前希望评价疾病进展时，替代标志物特别有用（例如，在心肌梗塞或完全发展的 AIDS 的不良临床结果之前，可以用胆固醇水平作为替代标志物进行心血管疾病的评价，以及可以用 HIV RNA 水平作为替代标志物进行 HIV 感染的分析）。本领域中替代标志物的应用的例子包括：Koomen 等人 (2000) *J. Mass. Spectrom.* 35: 258-264; 和 James (1994) *AIDS Treatment News Archive* 209。

本发明的标志物也可用作药效学标志物。此处使用的“药效学标志物”是与药物疗效特别相关的客观生物化学标志物。药效学标志物的存在或量与给药治疗的疾病状态或紊乱无关；因此，该标志物的存在或量指示该药物在受试者中的存在或活性。例如，药效学标志物可以指示药物在生物组织中的浓度，因为该标志物在该组织中表达或转录或不表达或不转录与药物水平有关。以这种方式，可以用药效学标志物监测药物的分布或摄取。同样，药效学标志物的存在或量可能与药物代谢产物的存在或量有关，因此标志物的存在或量表示药物在体

内的相对分解率。药效学标志物尤其用于提高药物疗效检测的灵敏度，特别是在低剂量给药时。由于至少少量药物也可能足以激活多轮标志物转录或表达，扩增的标志物的量可能比药物本身更容易检测到。另外由于标志物本身的性质，标志物也可能更容易检测；例如，使用此处所述的方法，可以在基于免疫的蛋白质标志物检测系统中使用抗体，或者可以利用标志物特异性的放射性标记探针检测 mRNA 标志物。而且，使用药效学标志物还可以允许基于机理地预测超出可直接观测范围以外的药物治疗引起的危险。本领域中药效学标志物的应用的例子包括：Matsuda 等人 US 6,033,862; Hattis 等人 (1991) *Env. Health Perspect.* 90: 229-238; Schentag (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S21-S24; 和 Nicolau (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S16-S20。

监测临床试验

监测药剂（例如药物化合物）对本发明标志物的表达水平的影响不仅可以在基本药物筛选中使用，而且可以在临床试验中使用。例如，可以在接受前列腺癌治疗的受试者的临床试验中监测一种药剂影响标志物表达的有效性。在一个优选实施方案中，本发明提供一种监测用药剂（例如激动剂、拮抗剂、拟肽、蛋白质、肽、核酸、小分子或其他候选药物）治疗受试者的效果的方法，该方法包括如下步骤：(i) 在施用所述药剂前从受试者中获得给药前样品；(ii) 检测给药前样品中一种或多种选择的本发明标志物的表达水平；(iii) 从受试者获得一种或多种给药后样品；(iv) 检测给药后样品中该标志物的表达水平；(v) 将给药前样品中该标志物的表达水平与给药后样品中该标志物的表达水平进行比较；和 (vi) 相应地改变对受试者的给药。例如，治疗过程中标志物基因的表达提高可能表明剂量无效及需要提高剂量。相反，标志物基因的表达降低可能表明治疗有效，并且不需要改变剂量。

电子设备可读介质和阵列

也提供了包含本发明标志物的电子设备可读介质。此处使用的

“电子设备可读介质”是指用于存储、保存或容纳数据或信息的任何合适的介质，这些数据或信息可以被电子设备直接读取和访问。这类介质包括但不限于：磁性存储介质，如软盘、硬盘存储介质和磁带；光存储介质，如光盘；电子存储介质，如 RAM、ROM、EPROM、EEPROM 等；普通硬盘，和这些种类的混合体，例如磁性/光存储介质。改变或构造该介质用于在上面记录本发明的标志物。

此处使用的术语“电子设备”包括为存储数据或信息而构造或改变的任何合适的计算或处理设备或其他装置。适合本发明使用的电子设备的例子包括单机计算设备；网络，包括局域网(LAN)、广域网(WAN)、互联网、内联网和外联网；电子用品，如个人数字助理(PDA)、移动电话、寻呼机等；以及本地的和分布式处理系统。

此处使用的“记录”是指在电子设备可读介质上存储或编码信息的过程。本领域技术人员可以容易地采用任何当前已知的方法在已知介质上记录信息，从而产生包含本发明标志物的产品。

有多种软件程序和格式可以用于在电子设备可读介质上存储本发明的标志物信息。例如，标志物核酸序列可以用字处理文本文件来表示，其格式为商业上可获得的软件如 WordPerfect 和 MicroSoft Word，或者用 ASCII 文件的形式来表示，存储在数据库中，如 DB2、Sybase、Oracle 等，以及用其他形式。为了获得或产生上面记录着本发明标志物的介质，可以使用任意数量的数据处理器结构格式（例如文本文件或数据库）。

通过以可读形式提供本发明的标志物，人们能够为了多种目的常规获取标志物序列信息。例如，本领域技术人员可以使用可读形式的本发明的核苷酸或氨基酸序列，将靶序列或靶结构基序与存储在数据存储工具内的序列信息进行比较。利用检索工具鉴定与特定靶序列或靶基序匹配的本发明序列的片段或区域。

本发明因此提供一种保存运行一种方法的指令的介质，该方法用于确定受试者是否患有前列腺癌或者易患前列腺癌，其中该方法包括以下步骤：确定标志物的存在与否，根据该标志物的存在与否确定受

试者是否患有前列腺癌或易患前列腺癌，和/或推荐用于前列腺癌或前列腺癌前病变的具体治疗方法。

本发明进一步在电子系统和/或网络中提供一种确定受试者是否患有前列腺癌或者易患与一种标志物有关的前列腺癌的方法，其中该方法包括以下步骤：确定标志物的存在与否，根据该标志物的存在与否确定受试者是否患有前列腺癌或易患前列腺癌，和/或推荐用于前列腺癌或前列腺癌前病变的具体治疗方法。该方法可以进一步包括以下步骤：接受与该受试者有关的表型信息，和/或从网络上获取与该受试者有关的表型信息。

本发明也在网络上提供一种确定受试者是否患有前列腺癌或者易患与一种标志物有关的前列腺癌的方法，该方法包括以下步骤：接受与该标志物有关的信息，接受与该受试者有关的表型信息，从网络上获取对应于该标志物和/或前列腺癌的信息，并且根据一种或多种表型信息、标志物和获得的信息，确定该受试者是否患有前列腺癌或易患前列腺癌。该方法可进一步包括推荐用于前列腺癌或前列腺癌前病变的具体治疗方法的步骤。

本发明也提供一种确定受试者是否患有前列腺癌或者易患前列腺癌的商业方法，该方法包括以下步骤：接受与标志物有关的信息，接受与该受试者有关的表型信息，从网络上获取对应于该标志物和/或前列腺癌的信息，并且根据一种或多种表型信息、标志物和获得的信息，确定该受试者是否患有前列腺癌或易患前列腺癌。该方法可进一步包括推荐用于前列腺癌或前列腺癌前病变的具体治疗方法的步骤。

本发明也包括一种包含本发明标志物的阵列。该阵列可以用来在阵列中测定一个或多个基因的表达。在一个实施方案中，该阵列可以用来在阵列中测定组织中的基因表达，从而确定基因的组织特异性。利用这种方法，可以同时测定最多大约 7600 个基因的表达。这可以形成一个图谱，显示在一个或多个组织中特异性表达的一组基因。

除了这些定性检测以外，本发明还允许对基因表达定量。因此，不仅一组基因的组织特异性，而且其在组织中的表达水平，都可以确

定。因此,可以根据其组织表达本身以及在该组织中的表达水平来对基因分组。例如,在确定组织之间基因表达的关系时,这是有用的。因此,可以干扰一种组织,而测定对第二组织中基因表达的影响。在本文中,可以测定一种细胞型响应生物学刺激对另一种细胞型的作用。这种测定例如可用于了解在基因表达水平上细胞-细胞相互作用的效应。如果治疗性地施用一种药剂来处理一种细胞型但是对另一种细胞型有不希望的作用,则本发明提供一种确定这种不希望的作用的分子基础的试验,从而提供同时施用对抗剂或者以其他方式处理这种不希望的作用的机会。同样,甚至在单细胞型内,也可以在分子水平上确定不希望的生物效应。因此,可以确定并抵消一种药剂对靶基因以外的基因表达的影响。

在另一个实施方案中,可以利用阵列监测一种或多种基因的表达的时程。这可以在如此处所述的各种生物事件中发生,例如,前列腺癌的发展,前列腺癌的进展,和过程,与前列腺癌有关的细胞转化。

所述阵列也可以用于确定一种基因的表达对同一细胞中或不同细胞中其他基因表达的影响。这提供了,例如,如果最终或下游靶标不能调节,则为治疗性干预选择替代分子靶标。

所述阵列也可用于确定一种或多种基因在正常和异常细胞中的不同表达模式。这提供了一组基因,它们能用作诊断或治疗性干预的分子靶标。

实验

实施例 1: 前列腺癌标志物的鉴定

材料与方法

RNA 制备

从来源于前列腺肿瘤标本、正常前列腺标本、非前列腺肿瘤标本和非前列腺正常标本的样品制备 RNA。将冷冻的组织块切片,从样品中分离 RNA。使用 Trizol 试剂(Invitrogen, San Diego, CA)从冷冻的组织中提取总 RNA,然后用 Qiagen's RNeasy 试剂盒进行第二次纯化步

骤, 以提高 RNA 探针的标记效率 (Qiagen, Valencia CA)。该研究中只使用根据琼脂糖凝胶电泳后 RNA 的乙锭染色计算 28S/18S 核糖体 RNA 比值至少为 1.0 的 RNA。

cDNA 微阵列杂交

在尼龙滤膜上形成含有来自 Research Genetics (Huntsville, AL) 的 30,732 个单基因克隆的 cDNA 微阵列。用总共 4-6 μg 总 RNA 作为模板, 使用 ^{33}P -dCTP、oligo dT-30 引物和 Superscript II 逆转录酶 (Life Technologies) 通过逆转录产生放射性标记的 cDNA。为了减少非特异性杂交, ^{33}P -标记的第一链 cDNA 与 cot-1 DNA 和聚-dA 40-60 (Pharmacia, Peapack, NJ) 预退火。在 65°C 下每个滤膜与约 6×10^6 计数的标记探针在含 7% 十二烷基硫酸钠 (SDS)、250mM Na_3PO_4 (pH 7.2)、1 mM EDTA、0.5% 酪蛋白-Hammerstein 和 0.1mg/ml 变性鲑精 DNA 的缓冲液中杂交 16 小时。用 4% 和 1% SDS 洗涤缓冲液 (20mM Na_3PO_4 (pH 7.2), 1 mM EDTA 和 4% 或 1% SDS) 洗涤滤膜后, 将其暴露于 Fuji 磷成像仪屏, 并用 Fuji 扫描仪 BAS 2500 扫描。用 Millennium Pharmaceuticals, Inc. 开发的自动阵列分析程序 Grid Guru v1.0 对斑点进行定量。

标志物评分算法和数据分析

为了修正杂交效率的差异, 用滤膜上所有斑点的密度中位数标准化来自每个微阵列滤膜的数字化数据。进行基于阵列的和基于基因的分级群聚, 并用 Stanford's 基因簇和基因树显示软件显现出来。通过计算每个基因的标志物得分将差异表达的基因排序。

为了计算标志物得分, 将样品分为对照组和试验组。标志物得分的起点是试验样品高于对照样品的平均倍数改变 (比值)。得分设计为反映显示差异表达的试验样品的改变程度 (表达比) 和数量, 而不是被具有极高数值的少部分试验样品所主导。为了减少这种“离群值”效应, 将表达比高于 10 的基因处理为与表达比为 10 的基因没有实质不同。在得到检测样品间的平均倍数改变之前, 用渐近压缩函数转化试验样品: 对照表达比, 来实现标志物得分的这种理想的表现。当试验

样品看起来没有比对照更高表达时,标志物得分的值为 1,在其他情况下该值大于 1。任何基因的标志物得分都不会超过 10。

因此基因 g 的标志物得分 S_g 如下计算为各试验样品和对照水平的加权强度的比值的平均值:

$$S_g = (\sum S_{gs}) / N_{\text{试验}}$$

$S_{gs} = C(x_{gs}/(k+x_g^Q))$, 其中 S_{gs} 代表基因 g 和样品 s 的标志物得分,

$C(r)$ $r \geq 1$ 时, 是压缩函数 $C(r) = A(1-e^{-r/A})$, $r < 1$ 时 $C(r) = 1$,

A 是倍数改变值上的上渐进数(我们使用 10),

x_{gs} 是样品 s 上的基因 g 的表达值,

x_g^Q 是对照样品表达值的第 Q 百分位数; 一般 $Q = 50$,

k 是一个常数, 其反映数据中的可加噪声, 即重复测量中变异的固定成分。使用微阵列技术由校准实验得到的该参数的值为 0.25。

$N_{\text{试验}}$ 试验样品数。

利用定量 PCR 的基因表达分析

通过 TAQMAN®定量 PCR (Applied Biosystems) 测定从正常和患病(如癌)人组织样品制备的 cDNA 中的基因表达。简言之, 按照厂商的说明书 (Invitrogen) 利用 TRIZOL 试剂, 用单步提取法从患者样品制备总 RNA。用 DNase I (Ambion) 在 37°C 下处理每个 RNA 制备物 1 小时。利用 $\beta 2$ 微球蛋白作为内部扩增对照, 样品如果需要至少 38 个 PCR 扩增循环以达到荧光阈值水平(或 18s 核糖体基因进行 35 个 PCR 扩增循环), 则确定 DNase I 处理已完成。DNase I 处理后, 通过琼脂糖凝胶电泳和溴化乙锭染色来确认 RNA 样品的完整性。用苯酚提取 RNA 后, 根据厂商的说明书 (Applied Biosystems) 利用 Taqman 逆转录试剂从样品中制备 cDNA。未加逆转录酶的 RNA 阴性对照针对每个 RNA 样品进行模拟逆转录。

基于特定基因及其相关转录物的序列, 用 PrimerExpress 软件 (Applied Biosystems) 设计探针。每个靶基因探针用 FAM (6-羧基荧光素) 标记, 而 18s 参照探针用不同的荧光染料 VIC 来标记。因此靶

基因和内部对照基因不一样的标记能使测定在同一个孔中进行。检测引物和探针对特定基因的每个转录物的灵敏度和特异性。把用于 18s 和靶基因的正向和反向引物和探针加入到 TAQMAN®通用 PCR 主混合物 (Applied Biosystems) 中。尽管引物和探针的终浓度可能发生变化,但是每个在特定实验中都是内在一致的。对于 18s,典型的实验包含 100nM 正向和反向引物加上 200 nM 探针,对于靶基因,包含 900nM 正向和反向引物加上 250 nM 探针。在 ABI PRISM 7700 型序列检测系统 (Applied Biosystems) 上进行 TAQMAN®基质实验。热循环条件如下: 50℃下维持 2 分钟并在 95℃下维持 10 分钟,接着进行两步 PCR: 95℃下 15 秒,接着 60℃下 1 分钟,40 个循环。

应用以下方法定量计算相对于相同组织中的 18s 表达,各种组织中的基因表达。阈循环 (Ct) 值定义为检测到荧光有统计学显著增加的循环。较低的 Ct 值指示了较高的 mRNA 浓度。应用以下公式: $\Delta Ct = Ct(\text{靶转录物}) - Ct(18s)$, 减去 18s 核糖体基因的 Ct 值获得 ΔCt 值,使基因的 Ct 值标准化。然后用 $2^{-\Delta Ct}$ 给出的算术公式计算相对表达。

结果

前列腺标志物选择

通过在上文材料与方法部分定义的转录谱分析,使用来自于前列腺筛选组的 mRNA,鉴定了表 1 中列出的所有标志物,所述筛选组由“前列腺肿瘤库”(3 个前列腺肿瘤患者样品)、“前列腺正常库”(5 个正常前列腺上皮患者样品)和“其它正常库”(来自于正常心脏、肾、小肠、脾、白细胞、肺、肝、脑、骨髓和结肠患者组织样品每种的一个样品)和“其它肿瘤库”(4 个宫颈癌、5 个结肠肿瘤、8 个各种类型的肺癌患者样品、4 个卵巢肿瘤样品和 5 个前列腺肿瘤样品)的患者样品组成。与前列腺正常、其它正常或其它肿瘤中的表达相比,在至少 25%的前列腺肿瘤中表达至少高 3 倍的克隆被称作前列腺肿瘤特异性筛选标志物。选择这些 cDNA 克隆以确定它们的蛋白编码转录物序

列。

为了确定所选 cDNA 克隆的全长蛋白编码转录物，用所选克隆的序列查询公共和私有的序列数据库，以便鉴定其它具有显著重叠的 EST 序列或簇。这样，将邻接的 EST 序列和 / 或簇组装到蛋白编码转录物中。按如下描述进行所有要求保护的标志物的可变转录物分析：

采用现有的将任意给定标志物基因的已知核苷酸序列对人基因组序列作图的方法，并且通过额外将任意给定标志物基因的新核苷酸序列在人基因组序列上作图（如，使用诸如“UCSC 基因组浏览器”等资源或具有相似功能的内部资源，结合允许快速和精确地将检索序列在基因组序列上作图的诸如 BLAT 等算法），确定标志物基因的外显子-内含子结构，其中还考虑了匹配相同基因的 EST 序列。

设计 PCR 引物以便从感兴趣的组织和对照样品扩增给定标志物基因的编码序列。由此分析得到的具有改变编码序列潜力的标志物基因的替代的 5' 或 3' 末端导致设计出了另外一种对此替代的末端具有特异性的引物。

将用来源于前列腺肿瘤标本的 cDNA 模板获得的 PCR 产物克隆到质粒载体中，并且通过 DNA 序列分析进行表征。典型地，通过 PCR 产物的限制消化和凝胶电泳，或通过 DNA 序列分析，分析 96 个克隆。

对以 2% 或更高的频率存在的代表可变基因转录物的克隆进行测序。通过基于所有可获得的序列鉴定在人工管理的组装体（重叠群）中包含的可读框（ORF）而完成相应于这些可变转录物的蛋白质序列的鉴定。在表 1 和序列表中指定了这些被鉴定序列的归属。

在从患者组织标本制备的 cDNA 中通过 TAQMAN® 定量 PCR（Applied Biosystems）证实了具有鉴定的可变转录物的基因的差别基因表达。在某些情况下（如 AMACR, LIM）制备对给定基因鉴定的所有转录物敏感的基因特异性 TAQMAN® 试剂。此外，对某些标志物的每种转录物或转录物组（如 M683A, M684A, M686; M283A, M747; M742, M743, M744, M745, M260A; M322A）分别开发剪接形式特异

性 TAQMAN®试剂。发现了具有相似扩增效率的、证明有不同的肿瘤-正常表达的基因特异性以及转录物特异性表达谱。与正常或其他肿瘤样品相比，如表 1 所示的标志物 M683A、M684A、M686、M283A、M747、M742、M743、M744、M745、M260A 和 M322A 中的每一个都证明有高于正常的表达水平。M747 和 M283A 为 LIM 基因的可变剪接变体(见表 2)。但是只证明 M747 有差别表达，导致与正常或其他肿瘤样品相比其表达上调。M747 代表比 M283A 实质上更短的 LIM 转录物，其特征在于独特的可变转录起始位点。M747 比 M283A 短，但是与 M283A 蛋白序列的 C 末端部分相同(见表 2; SEQ ID NO: 27, SEQ ID NO: 29; 和 SEQ ID NO: 28, SEQ ID NO: 30)。

表 2

标志物	基因名称	SEQ ID NO (nts)	SEQ ID NO (AAs)	CDS
M747	LIM: LIM 蛋白(类似于大鼠蛋白激酶 C 结合 enigma), 变体 2	29	30	394..828
M283A	LIM: LIM 蛋白(类似于大鼠蛋白激酶 C 结合 enigma), 变体 1	27	28	121..1911

实施例 2: 通过终点 PCR 的基因表达分析

材料与amp;方法

终点 PCR 分析

简言之，合并来自不同样品的总 RNA，将其用作产生第一链 cDNA 的模板。库中包括等量的每种样品。前列腺筛选组由“前列腺肿瘤库”(3 个前列腺肿瘤患者样品)、“前列腺正常库”(5 个正常前列腺上皮患者样品)、“其他正常库”(来自正常心脏、肾、小肠、脾脏、WBC、肺、肝脏、脑、骨髓和结肠组织患者组织样品的各 1 个样

品)、和“其他肿瘤库”(4个子宫颈癌,5个结肠肿瘤,8个不同类型的肺癌患者样品,4个卵巢肿瘤样品,和5个前列腺肿瘤样品)的患者样品组成(例如,见表3)。

使用 TRIZOL 试剂按照使用说明书(Invitrogen)通过一步提取法从患者样品制备总 RNA。每种 RNA 制品都用 DNase I (Ambion)在 37°C 下处理 1 小时。将来自每个患者样品的 RNA 并入四个库之一,例如前列腺正常库、前列腺肿瘤库、其他正常库和其他肿瘤库。

利用 ThermoScript RT-PCR 系统(Invitrogen, San Diego, CA)从 4 个库中的每一个获得 cDNA。简言之,按照使用说明书所述,1 μ g RNA 在 65°C 下在 10 μ l 体积中与 1 μ l 50 μ M oligo (dT)20 引物变性 5 分钟。在 85°C 下温育 5 分钟终止反应。终产物用水稀释至终体积为 100 μ l。

基因特异性引物设计为恰好在开放阅读框之外(如表 3 中项目“终点 PCR 引物 1”和“终点 PCR 引物 2”所示)。对引物和预期的产物大小优化 PCR 条件。应用降落(touchdown)循环条件在 20 μ l 反应中使用 2 μ l cDNA。产物在含溴化乙锭的琼脂糖凝胶上电泳,并对得到的凝胶图进行半定量分析和评分。对组织组进行的终点 PCR 的凝胶图按照 1-5 的等级进行评分。每张图由三个人根据目测条带强度独立评分,并汇总结果,将得分进行比较,以证实全部三个人在条带相对强度上一致,必要时进行修正。然后记录三个得分的中位数作为最终得分。

II. 结果

表 3. 前列腺筛选终点 PCR 数据

标志物	基因名称	PCR 引物 1	PCR 引物 2	正常库	肿瘤库	前列腺正常	前列腺肿瘤
M245A	ABCC4: ATP-结合盒, 亚家族 C (CFTR/MRP), 成员 4	3225...3247	3611...3634	1	2	3	4
M683A	AMACR: α -甲基酰基-CoA 消旋酶, 变体 1	60..79	2103...2127	1	1	2	2

M684A	AMACR: α -甲基酰基-CoA 消旋酶, 变体 2	60..79	1354....1378	0	0	0	1
M686	AMACR: α -甲基酰基-CoA 消旋酶, 变体 3	36..55	1019...1042	n/d	n/d	1	3
M261	FABP5: 脂肪酸结合蛋白 5 (牛皮癣相关)	45..65	463.....488	1	3	2	4
M746	FLJ23153: 肿瘤坏死因子- α -诱导的脂肪相关蛋白	85..105	1606....1626	0	0	4	2
M271A	GOLPH2: 高尔基体磷蛋白 2	185.....205	1426....1445	1	2	3	5
M279A	HSRG1: HSV-1 刺激相关基因 1	587.....606	1178....1197	2	0	3	1
M283A	LIM: LIM 蛋白(类似于大鼠蛋白激酶 C-结合 enigma), 变体 1	93..115	2081....2104	2	1	5	3
M747	LIM: LIM 蛋白(类似于大鼠蛋白激酶 C-结合 enigma), 变体 2	232.....255	995.....1019	0	1	1	3
M164A	NET-6: 跨膜 4 超家族成员 tetraspan NET-6	130.....149	984.....1004	1	1	3	5
M301A	NPY: 神经肽 Y	34..52	446.....465	1	3	5	5
M308A	PLA2G2A: 磷脂酶 A2, IIA 组(血小板, 滑液)	206.....227	832.....853	2	3	5	5
M318A	SMS: 精胺合酶, 亚精胺氨丙基转移酶	102.....119	1178....1199	2	3	5	5
M320	SPON2: spondin 2, 胞外基质蛋白	191.....211	1364...1383	n/d	n/d	0	3
M389	TACSTD1: 肿瘤相关钙信号转导蛋白 1	146.....163	1185...1204	0	1	1	3
M322A	TARP: T-细胞受体 γ -链替代阅读框蛋白	1..24	141.....160	2	5	0	3

标志物在前列腺肿瘤样品中的表达水平高于在其他肿瘤样品或正常样品组中获得的表达水平(表 3)。对于 M245A, M686, M261, M271A, M747, M164A, M320 和 M389, 与前列腺正常、其他正常或其他肿瘤组相比, 在前列腺肿瘤组中观察到特别强的表达。

此处引用的参考文献, 包括本申请全文中引用的期刊文章、专利、公开的专利申请和数据库记录, 包括 GenBank, IMAGE 联盟和 Derwent, 均引入作为参考。

其他实施方案

本领域技术人员应当认识到，或者仅利用常规实验就能够确定本文所述的本发明具体实施方案的等同实施方案。这些等同实施方案包括在权利要求书的范围内。

<110> 千年药品公司等

<120> 用于鉴定、评价、预防和治疗前列腺癌的基因、组合物、试剂盒和方法

<130> MRI-066PC

<150> 60/601413

<151> 2004-08-13

<160> 44

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 5871

<212> DNA

<213> 人

<220>

<221> CDS

<222> (116)... (4093)

<400> 1

```

ggacaggcgt ggcgccgga gccccagcat ccctgcttga ggtccaggag cggagcccgc 60
ggccaccgcc gcctgatcag cgcgaccccg gcccgcgccc gccccgccc gcaag atg 118
                                         Met
                                         1

```

```

ctg ccc gtg tac cag gag gtg aag ccc aac ccg ctg cag gac gcg aac 166
Leu Pro Val Tyr Gln Glu Val Lys Pro Asn Pro Leu Gln Asp Ala Asn
                    5                      10                      15

```

```

atc tgc tca cgc gtg ttc ttc tgg tgg ctc aat ccc ttg ttt aaa att 214
Ile Cys Ser Arg Val Phe Phe Trp Trp Leu Asn Pro Leu Phe Lys Ile
                20                      25                      30

```

```

ggc cat aaa cgg aga tta gag gaa gat gat atg tat tca gtg ctg cca 262
Gly His Lys Arg Arg Leu Glu Glu Asp Asp Met Tyr Ser Val Leu Pro
                35                      40                      45

```

```

gaa gac cgc tca cag cac ctt gga gag gag ttg caa ggg ttc tgg gat 310
Glu Asp Arg Ser Gln His Leu Gly Glu Glu Leu Gln Gly Phe Trp Asp
                50                      55                      60                      65

```

```

aaa gaa gtt tta aga gct gag aat gac gca cag aag cct tct tta aca 358
Lys Glu Val Leu Arg Ala Glu Asn Asp Ala Gln Lys Pro Ser Leu Thr
                70                      75                      80

```

```

aga gca atc ata aag tgt tac lgg aaa tct tat tta gtt ttg gga att 406
Arg Ala Ile Ile Lys Cys Tyr Trp Lys Ser Tyr Leu Val Leu Gly Ile
                85                      90                      95

```

```

ttt acg tta att gag gaa agt gcc aaa gta atc cag ccc ata ttt ttg 454
Phe Thr Leu Ile Glu Glu Ser Ala Lys Val Ile Gln Pro Ile Phe Leu
                100                      105                      110

```

```

gga aaa att att aat tat ttt gaa aat tat gat ccc atg gat tct gtg 502
Gly Lys Ile Ile Asn Tyr Phe Glu Asn Tyr Asp Pro Met Asp Ser Val
                115                      120                      125

```

```

gct ttg aac aca gcg tac gcc tat gcc acg gtg ctg act ttt tgc acg 550
Ala Leu Asn Thr Ala Tyr Ala Tyr Ala Thr Val Leu Thr Phe Cys Thr
                130                      135                      140                      145

```

```

ctc att ttg gct ata ctg cat cac tta tat ttt tat cac gtt cag tgt 598
Leu Ile Leu Ala Ile Leu His His Leu Tyr Phe Tyr His Val Gln Cys
                150                      155                      160

```

```

gct ggg atg agg tta cga gta gcc atg tgc cat atg att tat cgg aag 646
Ala Gly Met Arg Leu Arg Val Ala Met Cys His Met Ile Tyr Arg Lys
                165                      170                      175

```

```

gca ctt cgt ctt agt aac atg gcc atg ggg aag aca acc aca ggc cag 694
Ala Leu Arg Leu Ser Asn Met Ala Met Gly Lys Thr Thr Thr Gly Gln

```

180	185	190	
ata gtc aat ctg ctg tcc aat gat gtg aac aag ttt gat cag gtg aca Ile Val Asn Leu Leu Ser Asn Asp Val Asn Lys Phe Asp Gln Val Thr 195 200 205			742
gtg ttc tta cac ttc ctg tgg gca gga cca ctg cag gcg atc gca gtg Val Phe Leu His Phe Leu Trp Ala Gly Pro Leu Gln Ala Ile Ala Val 210 215 220 225			790
act gcc cta ctc tgg atg gag ata gga ata tcg tgc ctt gct ggg atg Thr Ala Leu Leu Trp Met Glu Ile Gly Ile Ser Cys Leu Ala Gly Met 230 235 240			838
gca gtt cta atc att ctc ctg ccc ttg caa agc tgt ttt ggg aag ttg Ala Val Leu Ile Ile Leu Leu Pro Leu Gln Ser Cys Phe Gly Lys Leu 245 250 255			886
ttc tca tca ctg agg agt aaa act gca act ttc acg gat gcc agg atc Phe Ser Ser Leu Arg Ser Lys Thr Ala Thr Phe Thr Asp Ala Arg Ile 260 265 270			934
agg acc atg aat gaa gtt ata act ggt ata agg ata ata aaa atg tac Arg Thr Met Asn Glu Val Ile Thr Gly Ile Arg Ile Ile Lys Met Tyr 275 280 285			982
gcc tgg gaa aag tca ttt tca aat ctt att acc aat ttg aga aag aag Ala Trp Glu Lys Ser Phe Ser Asn Leu Ile Thr Asn Leu Arg Lys Lys 290 295 300 305			1030
gag att tcc aag att ctg aga agt tcc tgc ctc agg ggg atg aat ttg Glu Ile Ser Lys Ile Leu Arg Ser Ser Cys Leu Arg Gly Met Asn Leu 310 315 320			1078
gct tcg ttt ttc agt gca agc aaa atc atc gtg ttt gtg acc ttc acc Ala Ser Phe Phe Ser Ala Ser Lys Ile Ile Val Phe Val Thr Phe Thr 325 330 335			1126
acc tac gtg ctc ctc ggc agt gtg atc aca gcc agc cgc gtg ttc gtg Thr Tyr Val Leu Leu Gly Ser Val Ile Thr Ala Ser Arg Val Phe Val 340 345 350			1174
gca gtg acg ctg tat ggg gct gtg cgg ctg acg gtt acc ctc ttc ttc Ala Val Thr Leu Tyr Gly Ala Val Arg Leu Thr Val Thr Leu Phe Phe 355 360 365			1222
ccc tca gcc att gag agg gtg tca gag gca atc gtc agc atc cga aga Pro Ser Ala Ile Glu Arg Val Ser Glu Ala Ile Val Ser Ile Arg Arg 370 375 380 385			1270
atc cag acc ttt ttg cta ctt gat gag ata tca cag cgc aac cgt cag Ile Gln Thr Phe Leu Leu Leu Asp Glu Ile Ser Gln Arg Asn Arg Gln 390 395 400			1318
ctg ccg tca gat ggt aaa aag atg gtg cat gtg cag gat ttt act gct Leu Pro Ser Asp Gly Lys Lys Met Val His Val Gln Asp Phe Thr Ala 405 410 415			1366
ttt tgg gat aag gca tca gag acc cca act cta caa ggc ctt tcc ttt Phe Trp Asp Lys Ala Ser Glu Thr Pro Thr Leu Gln Gly Leu Ser Phe 420 425 430			1414
act gtc aga cct ggc gaa ttg tta gct gtg gtc ggc ccc gtg gga gca Thr Val Arg Pro Gly Glu Leu Leu Ala Val Val Gly Pro Val Gly Ala 435 440 445			1462
ggg aag tca tca ctg tta agt gcc gtg ctc ggg gaa ttg gcc cca agt Gly Lys Ser Ser Leu Leu Ser Ala Val Leu Gly Glu Leu Ala Pro Ser 450 455 460 465			1510
cac ggg ctg gtc agc gtg cat gga aga att gcc tat gtg tct cag cag His Gly Leu Val Ser Val His Gly Arg Ile Ala Tyr Val Ser Gln Gln 470 475 480			1558
ccc tgg gtg ttc tcg gga act ctg agg agt aat att tta ttt ggg aag Pro Trp Val Phe Ser Gly Thr Leu Arg Ser Asn Ile Leu Phe Gly Lys 485 490 495			1606

aaa tat gaa aag gaa cga tat gaa aaa gtc ata aag gct tgt gct ctg Lys Tyr Glu Lys Glu Arg Tyr Glu Lys Val Ile Lys Ala Cys Ala Leu 500 505 510	1654
aaa aag gat tta cag ctg ttg gag gat ggt gat ctg act gtg ata gga Lys Lys Asp Leu Gln Leu Leu Glu Asp Gly Asp Leu Thr Val Ile Gly 515 520 525	1702
gat cgg gga acc acg ctg agt gga ggg cag aaa gca cgg gta aac ctt Asp Arg Gly Thr Thr Leu Ser Gly Gly Gln Lys Ala Arg Val Asn Leu 530 535 540 545	1750
gca aga gca gtg tat caa gat gct gac atc tat ctc ctg gac gat cct Ala Arg Ala Val Tyr Gln Asp Ala Asp Ile Tyr Leu Leu Asp Asp Pro 550 555 560	1798
ctc agt gca gta gat gcg gaa gtt agc aga cac ttg ttc gaa ctg tgt Leu Ser Ala Val Asp Ala Glu Val Ser Arg His Leu Phe Glu Leu Cys 565 570 575	1846
att tgt caa att ttg cat gag aag atc aca att tta gtg act cat cag Ile Cys Gln Ile Leu His Glu Lys Ile Thr Ile Leu Val Thr His Gln 580 585 590	1894
ttg cag tac ctc aaa gct gca agt cag att ctg ata ttg aaa gat ggt Leu Gln Tyr Leu Lys Ala Ala Ser Gln Ile Leu Ile Leu Lys Asp Gly 595 600 605	1942
aaa atg gtg cag aag ggg act tac act gag ttc cta aaa tct ggt ata Lys Met Val Gln Lys Gly Thr Tyr Thr Glu Phe Leu Lys Ser Gly Ile 610 615 620 625	1990
gat ttt ggc tcc ctt tta aag aag gat aat gag gaa agt gaa caa cct Asp Phe Gly Ser Leu Leu Lys Lys Asp Asn Glu Glu Ser Glu Gln Pro 630 635 640	2038
cca gtt cca gga act ccc aca cta agg aat cgt acc ttc tca gag tct Pro Val Pro Gly Thr Pro Thr Leu Arg Asn Arg Thr Phe Ser Glu Ser 645 650 655	2086
tcg gtt tgg tct caa caa tct tct aga ccc tcc ttg aaa gat ggt gct Ser Val Trp Ser Gln Gln Ser Ser Arg Pro Ser Leu Lys Asp Gly Ala 660 665 670	2134
ctg gag agc caa gat aca gag aat gtc cca gtt aca cta tca gag gag Leu Glu Ser Gln Asp Thr Glu Asn Val Pro Val Thr Leu Ser Glu Glu 675 680 685	2182
aac cgt tct gaa gga aaa gtt ggt ttt cag gcc tat aag aat tac ttc Asn Arg Ser Glu Gly Lys Val Gly Phe Gln Ala Tyr Lys Asn Tyr Phe 690 695 700 705	2230
aga gct ggt gct cac tgg att gtc ttc att ttc ctt att ctc cta aac Arg Ala Gly Ala His Trp Ile Val Phe Ile Phe Leu Ile Leu Leu Asn 710 715 720	2278
act gca gct cag gtt gcc tat gtg ctt caa gat tgg tgg ctt tca tac Thr Ala Ala Gln Val Ala Tyr Val Leu Gln Asp Trp Trp Leu Ser Tyr 725 730 735	2326
tgg gca aac aaa caa agt atg cta aat gtc act gta aat gga gga gga Trp Ala Asn Lys Gln Ser Met Leu Asn Val Thr Val Asn Gly Gly Gly 740 745 750	2374
aat gta acc gag aag cta gat ctt aac tgg tac tta gga att tat tca Asn Val Thr Glu Lys Leu Asp Leu Asn Trp Tyr Leu Gly Ile Tyr Ser 755 760 765	2422
ggt tta act gta gct acc gtt ctt ttt ggc ata gca aga tct cta ttg Gly Leu Thr Val Ala Thr Val Leu Phe Gly Ile Ala Arg Ser Leu Leu 770 775 780 785	2470
gta ttc tac gtc ctt gtt aac tct tca caa act ttg cac aac aaa atg Val Phe Tyr Val Leu Val Asn Ser Ser Gln Thr Leu His Asn Lys Met 790 795 800	2518

ttt gag tca att ctg aaa gct ccg gta tta ttc ttt gat aga aat cca Phe Glu Ser Ile Leu Lys Ala Pro Val Leu Phe Phe Asp Arg Asn Pro 805 810 815	2566
ata gga aga att tta aat cgt ttc tcc aaa gac att gga cac ttg gat Ile Gly Arg Ile Leu Asn Arg Phe Ser Lys Asp Ile Gly His Leu Asp 820 825 830	2614
gat ttg ctg ccg ctg acg ttt tta gat ttc atc cag aca ttg cta caa Asp Leu Leu Pro Leu Thr Phe Leu Asp Phe Ile Gln Thr Leu Leu Gln 835 840 845	2662
gtg gtt ggt gtg gtc tct gtg gct gtg gcc gtg att cct tgg atc gca Val Val Gly Val Val Ser Val Ala Val Ala Val Ile Pro Trp Ile Ala 850 855 860 865	2710
ata ccc ttg gtt ccc ctt gga atc att ttc att ttt ctt cgg cga tat Ile Pro Leu Val Pro Leu Gly Ile Ile Phe Ile Phe Leu Arg Arg Tyr 870 875 880	2758
ttt ttg gaa acg tca aga gat gtg aag cgc ctg gaa tct aca act cgg Phe Leu Glu Thr Ser Arg Asp Val Lys Arg Leu Glu Ser Thr Thr Arg 885 890 895	2806
agt cca gtg ttt tcc cac ttg tca tct tct ctc cag ggg ctc tgg acc Ser Pro Val Phe Ser His Leu Ser Ser Ser Leu Gln Gly Leu Trp Thr 900 905 910	2854
atc cgg gca tac aaa gca gaa gag agg tgt cag gaa ctg ttt gat gca Ile Arg Ala Tyr Lys Ala Glu Glu Arg Cys Gln Glu Leu Phe Asp Ala 915 920 925	2902
cac cag gat tta cat tca gag gct tgg ttc ttg ttt ttg aca acg tcc His Gln Asp Leu His Ser Glu Ala Trp Phe Leu Phe Leu Thr Thr Ser 930 935 940 945	2950
cgc tgg ttc gcc gtc cgt ctg gat gcc atc tgt gcc atg ttt gtc atc Arg Trp Phe Ala Val Arg Leu Asp Ala Ile Cys Ala Met Phe Val Ile 950 955 960	2998
atc gtt gcc ttt ggg tcc ctg att ctg gca aaa act ctg gat gcc ggg Ile Val Ala Phe Gly Ser Leu Ile Leu Ala Lys Thr Leu Asp Ala Gly 965 970 975	3046
cag gtt ggt ttg gca ctg tcc tat gcc ctc acg ctc atg ggg atg ttt Gln Val Gly Leu Ala Leu Ser Tyr Ala Leu Thr Leu Met Gly Met Phe 980 985 990	3094
cag tgg tgt gtt cga caa agt gct gaa gtt gag aat atg atg atc tca Gln Trp Cys Val Arg Gln Ser Ala Glu Val Glu Asn Met Met Ile Ser 995 1000 1005	3142
gla gaa agg gtc att gaa tac aca gac ctt gaa aaa gaa gca cct tgg Val Glu Arg Val Ile Glu Tyr Thr Asp Leu Glu Lys Glu Ala Pro Trp 1010 1015 1020 1025	3190
gaa tai cag aaa cgc cca cca cca gcc tgg ccc cat gaa gga gtg ata Glu Tyr Gln Lys Arg Pro Pro Pro Ala Trp Pro His Glu Gly Val Ile 1030 1035 1040	3238
atc ttt gac aat gtg aac ttc atg tac agt cca ggt ggg cct ctg gta Ile Phe Asp Asn Val Asn Phe Met Tyr Ser Pro Gly Gly Pro Leu Val 1045 1050 1055	3286
ctg aag cat ctg aca gca ctc att aaa tca caa gaa aag gtt ggc att Leu Lys His Leu Thr Ala Leu Ile Lys Ser Gln Glu Lys Val Gly Ile 1060 1065 1070	3334
gtg gga aga acc gga gct gga aaa agt tcc ctc atc tca gcc ctt ttt Val Gly Arg Thr Gly Ala Gly Lys Ser Ser Leu Ile Ser Ala Leu Phe 1075 1080 1085	3382
aga ttg tca gaa ccc gaa ggt aaa att tgg att gat aag atc ttg aca Arg Leu Ser Glu Pro Glu Gly Lys Ile Trp Ile Asp Lys Ile Leu Thr 1090 1095 1100 1105	3430
act gaa att gga ctt cac gat tta agg aag aaa atg tca atc ata cct	3478

Thr Glu Ile Gly Leu His Asp Leu Arg Lys Lys Met Ser Ile Ile Pro	
1110 1115 1120	
cag gaa cct gtt ttg ttc act gga aca atg agg aaa aac ctg gat ccc	3526
Gln Glu Pro Val Leu Phe Thr Gly Thr Met Arg Lys Asn Leu Asp Pro	
1125 1130 1135	
ttt aag gag cac acg gat gag gaa ctg tgg aat gcc tta caa gag gta	3574
Phe Lys Glu His Thr Asp Glu Glu Leu Trp Asn Ala Leu Gln Glu Val	
1140 1145 1150	
caa ctt aaa gaa acc att gaa gat ctt cct ggt aaa atg gat act gaa	3622
Gln Leu Lys Glu Thr Ile Glu Asp Leu Pro Gly Lys Met Asp Thr Glu	
1155 1160 1165	
tta gca gaa tca gga tcc aat ttt agt gtt gga caa aga caa ctg gtg	3670
Leu Ala Glu Ser Gly Ser Asn Phe Ser Val Gly Gln Arg Gln Leu Val	
1170 1175 1180 1185	
tgc ctt gcc agg gca att ctc agg aaa aat cag ata ttg att att gat	3718
Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn Gln Ile Leu Ile Ile Asp	
1190 1195 1200	
gaa gcg acg gca aat gtg gat cca aga act gat gag tta ata caa aaa	3766
Glu Ala Thr Ala Asn Val Asp Pro Arg Thr Asp Glu Leu Ile Gln Lys	
1205 1210 1215	
aaa atc cgg gag aaa ttt gcc cac tgc acc gtg cta acc att gca cac	3814
Lys Ile Arg Glu Lys Phe Ala His Cys Thr Val Leu Thr Ile Ala His	
1220 1225 1230	
aga ttg aac acc att att gac agc gac aag ata atg gtt tta gat tca	3862
Arg Leu Asn Thr Ile Ile Asp Ser Asp Lys Ile Met Val Leu Asp Ser	
1235 1240 1245	
gga aga ctg aaa gaa tat gat gag ccg tat gtt ttg ctg caa aat aaa	3910
Gly Arg Leu Lys Glu Tyr Asp Glu Pro Tyr Val Leu Leu Gln Asn Lys	
1250 1255 1260 1265	
gag agc cta ttt tac aag atg gtg caa caa ctg ggc aag gca gaa gcc	3958
Glu Ser Leu Phe Tyr Lys Met Val Gln Gln Leu Gly Lys Ala Glu Ala	
1270 1275 1280	
gct gcc ctc act gaa aca gca aaa cag gta tac ttc aaa aga aat tat	4006
Ala Ala Leu Thr Glu Thr Ala Lys Gln Val Tyr Phe Lys Arg Asn Tyr	
1285 1290 1295	
cca cat att ggt cac act gac cac atg gtt aca aac act tcc aat gga	4054
Pro His Ile Gly His Thr Asp His Met Val Thr Asn Thr Ser Asn Gly	
1300 1305 1310	
cag ccc tcg acc tta act att ttc gag aca gca ctg tga atccaaccaa	4103
Gln Pro Ser Thr Leu Thr Ile Phe Glu Thr Ala Leu *	
1315 1320 1325	
aatgtcaagt ccgttccgaa ggcattttcc actagttttt ggactatgta aaccacattg	4163
tacttttttt tacttttgca acaaatattt atacatacaa gatgctagtt catttgaata	4223
tttctcccaa tttatccaag gatctccagc tctaacaataa tggtttattt ttattttaat	4283
gtcaatagtt gttttttaaa atccaaatca gagggtgcagg ccaccagtta aatgccgtct	4343
atcagggttt gtgccttaag agactacaga gtcaaagctc atttttaaag gagtaggaca	4403
aagtgtcac aggtttttgt tgtgttttt atlgccccc aaattacatg ttaatttcca	4463
tttatatcag ggatttctatt tacilgaaga ctglgaagtt gccattttgt ctcatgttt	4523
tcitlgacat aactaggatc cattatttcc ccigaagget tcttgttaga aaatagtaca	4583
gttacaacca ataggaacaa caaaaagaaa aagtttttga cattgttagta gggagtgtgt	4643
accccttact ccccatcaaa aaaaaaatg gatacatggt taaaggatag aagggaata	4703
ttttatcata tgttctaaaa gagaaggag agaaaatact actttctcaa aatggaagcc	4763
cttaaagggtg ctltgatact gaaggacaca aatgtgaccg tccatcctcc tttagagtig	4823
catgacttgg acacggtaac tglitgagtt ttagactcag cattgtgaca ctcccaaga	4883
aggccaaacc tctaaccgac attcctgaaa tacgtggcat tattcttttt tggatttctc	4943
atttatggaa ggctaacct ctgttgactg taagcctttt ggtttgggct gtattgaaat	5003
ccitttctaaa ttgcatgaat aggcctctgct aacgtgatga gacaaactga aaattattgc	5063
aagcattgac tataattatg cagtacgttc tcaggatgca tccaggggtt cattttcatg	5123
agcctgtcca ggtagtttta ctcttgacca ctaatagcat tgtcatttgg gctttctgtt	5183
gaatgaatca acaaacacaca atacttctg ggacctttt tactttattt gaactatgag	5243
tcittaatit ttcttgatga tgggtgctgt aatatgttga gttcagttta ctaaaggttt	5303
tactattatg ttttgaagtg gagtctcatg acctctcaga ataagggtgc acctccctga	5363
aattgcatai atgtatatag acatgcacac gtgtgcattt gttgtatatac atatatittg	5423

```

ccttcgtata gcaagttttt tgctcatcag cagagagcaa cagatgtttt attgagttaa 5483
gccttaaaaa gcacacacca cacacagcta actgccaaaa tacattgacc gtagtagctg 5543
ttcaactcct agtacttaga aatacacgta tggttaatgt tcagtccaac aaaccacaca 5603
cagtaaatgt ttattaatag tcatggttcg tatttttaggt gactgaaatt gcatcagtga 5663
tcataatgag gtttgttaaa acgataagcta tattcaaaat gtctatatgt ttatttggac 5723
ttttgaggtt aaagacagtc atataaacgt cctgtttctg ttttaatgtt atcatagaat 5783
tttttaatga aactaaattc aattgaaata aatgatatgt ttcattctcca aaaaaaaaaa 5843
aaaaaaaaaa aaaaaaaaaa aaaaaaaa 5871

```

<210> 2
 <211> 1325
 <212> PRT
 <213> 人

<400> 2

```

Met Leu Pro Val Tyr Gln Glu Val Lys Pro Asn Pro Leu Gln Asp Ala
1      5      10      15
Asn Ile Cys Ser Arg Val Phe Phe Trp Trp Leu Asn Pro Leu Phe Lys
20     25     30
Ile Gly His Lys Arg Arg Leu Glu Glu Asp Asp Met Tyr Ser Val Leu
35     40     45
Pro Glu Asp Arg Ser Gln His Leu Gly Glu Glu Leu Gln Gly Phe Trp
50     55     60
Asp Lys Glu Val Leu Arg Ala Glu Asn Asp Ala Gln Lys Pro Ser Leu
65     70     75     80
Thr Arg Ala Ile Ile Lys Cys Tyr Trp Lys Ser Tyr Leu Val Leu Gly
85     90     95
Ile Phe Thr Leu Ile Glu Glu Ser Ala Lys Val Ile Gln Pro Ile Phe
100    105    110
Leu Gly Lys Ile Ile Asn Tyr Phe Glu Asn Tyr Asp Pro Met Asp Ser
115    120    125
Val Ala Leu Asn Thr Ala Tyr Ala Tyr Ala Thr Val Leu Thr Phe Cys
130    135    140
Thr Leu Ile Leu Ala Ile Leu His His Leu Tyr Phe Tyr His Val Gln
145    150    155    160
Cys Ala Gly Met Arg Leu Arg Val Ala Met Cys His Met Ile Tyr Arg
165    170    175
Lys Ala Leu Arg Leu Ser Asn Met Ala Met Gly Lys Thr Thr Thr Gly
180    185    190
Gln Ile Val Asn Leu Leu Ser Asn Asp Val Asn Lys Phe Asp Gln Val
195    200    205
Thr Val Phe Leu His Phe Leu Trp Ala Gly Pro Leu Gln Ala Ile Ala
210    215    220
Val Thr Ala Leu Leu Trp Met Glu Ile Gly Ile Ser Cys Leu Ala Gly
225    230    235    240
Met Ala Val Leu Ile Ile Leu Leu Pro Leu Gln Ser Cys Phe Gly Lys
245    250    255
Leu Phe Ser Ser Leu Arg Ser Lys Thr Ala Thr Phe Thr Asp Ala Arg
260    265    270
Ile Arg Thr Met Asn Glu Val Ile Thr Gly Ile Arg Ile Ile Lys Met
275    280    285
Tyr Ala Trp Glu Lys Ser Phe Ser Asn Leu Ile Thr Asn Leu Arg Lys
290    295    300
Lys Glu Ile Ser Lys Ile Leu Arg Ser Ser Cys Leu Arg Gly Met Asn
305    310    315    320
Leu Ala Ser Phe Phe Ser Ala Ser Lys Ile Ile Val Phe Val Thr Phe
325    330    335
Thr Thr Tyr Val Leu Leu Gly Ser Val Ile Thr Ala Ser Arg Val Phe
340    345    350
Val Ala Val Thr Leu Tyr Gly Ala Val Arg Leu Thr Val Thr Leu Phe
355    360    365
Phe Pro Ser Ala Ile Glu Arg Val Ser Glu Ala Ile Val Ser Ile Arg
370    375    380
Arg Ile Gln Thr Phe Leu Leu Asp Glu Ile Ser Gln Arg Asn Arg
385    390    395    400
Gln Leu Pro Ser Asp Gly Lys Lys Met Val His Val Gln Asp Phe Thr
405    410    415
Ala Phe Trp Asp Lys Ala Ser Glu Thr Pro Thr Leu Gln Gly Leu Ser
420    425    430
Phe Thr Val Arg Pro Gly Glu Leu Leu Ala Val Val Gly Pro Val Gly
435    440    445
Ala Gly Lys Ser Ser Leu Leu Ser Ala Val Leu Gly Glu Leu Ala Pro
450    455    460
Ser His Gly Leu Val Ser Val His Gly Arg Ile Ala Tyr Val Ser Gln
465    470    475    480
Gln Pro Trp Val Phe Ser Gly Thr Leu Arg Ser Asn Ile Leu Phe Gly
485    490    495

```

Lys Lys Tyr Glu Lys Glu Arg Tyr Glu Lys Val Ile Lys Ala Cys Ala
 500 505 510
 Leu Lys Lys Asp Leu Gln Leu Leu Glu Asp Gly Asp Leu Thr Val Ile
 515 520 525
 Gly Asp Arg Gly Thr Thr Leu Ser Gly Gly Gln Lys Ala Arg Val Asn
 530 535 540
 Leu Ala Arg Ala Val Tyr Gln Asp Ala Asp Ile Tyr Leu Leu Asp Asp
 545 550 555 560
 Pro Leu Ser Ala Val Asp Ala Glu Val Ser Arg His Leu Phe Glu Leu
 565 570 575
 Cys Ile Cys Gln Ile Leu His Glu Lys Ile Thr Ile Leu Val Thr His
 580 585 590
 Gln Leu Gln Tyr Leu Lys Ala Ala Ser Gln Ile Leu Ile Leu Lys Asp
 595 600 605
 Gly Lys Met Val Gln Lys Gly Thr Tyr Thr Glu Phe Leu Lys Ser Gly
 610 615 620
 Ile Asp Phe Gly Ser Leu Leu Lys Lys Asp Asn Glu Glu Ser Glu Gln
 625 630 635 640
 Pro Pro Val Pro Gly Thr Pro Thr Leu Arg Asn Arg Thr Phe Ser Glu
 645 650 655
 Ser Ser Val Trp Ser Gln Gln Ser Ser Arg Pro Ser Leu Lys Asp Gly
 660 665 670
 Ala Leu Glu Ser Gln Asp Thr Glu Asn Val Pro Val Thr Leu Ser Glu
 675 680 685
 Glu Asn Arg Ser Glu Gly Lys Val Gly Phe Gln Ala Tyr Lys Asn Tyr
 690 695 700
 Phe Arg Ala Gly Ala His Trp Ile Val Phe Ile Phe Leu Ile Leu Leu
 705 710 715 720
 Asn Thr Ala Ala Gln Val Ala Tyr Val Leu Gln Asp Trp Trp Leu Ser
 725 730 735
 Tyr Trp Ala Asn Lys Gln Ser Met Leu Asn Val Thr Val Asn Gly Gly
 740 745 750
 Gly Asn Val Thr Glu Lys Leu Asp Leu Asn Trp Tyr Leu Gly Ile Tyr
 755 760 765
 Ser Gly Leu Thr Val Ala Thr Val Leu Phe Gly Ile Ala Arg Ser Leu
 770 775 780
 Leu Val Phe Tyr Val Leu Val Asn Ser Ser Gln Thr Leu His Asn Lys
 785 790 795 800
 Met Phe Glu Ser Ile Leu Lys Ala Pro Val Leu Phe Phe Asp Arg Asn
 805 810 815
 Pro Ile Gly Arg Ile Leu Asn Arg Phe Ser Lys Asp Ile Gly His Leu
 820 825 830
 Asp Asp Leu Leu Pro Leu Thr Phe Leu Asp Phe Ile Gln Thr Leu Leu
 835 840 845
 Gln Val Val Gly Val Val Ser Val Ala Val Ala Val Ile Pro Trp Ile
 850 855 860
 Ala Ile Pro Leu Val Pro Leu Gly Ile Ile Phe Ile Phe Leu Arg Arg
 865 870 875 880
 Tyr Phe Leu Glu Thr Ser Arg Asp Val Lys Arg Leu Glu Ser Thr Thr
 885 890 895
 Arg Ser Pro Val Phe Ser His Leu Ser Ser Ser Leu Gln Gly Leu Trp
 900 905 910
 Thr Ile Arg Ala Tyr Lys Ala Glu Glu Arg Cys Gln Glu Leu Phe Asp
 915 920 925
 Ala His Gln Asp Leu His Ser Glu Ala Trp Phe Leu Phe Leu Thr Thr
 930 935 940
 Ser Arg Trp Phe Ala Val Arg Leu Asp Ala Ile Cys Ala Met Phe Val
 945 950 955 960
 Ile Ile Val Ala Phe Gly Ser Leu Ile Leu Ala Lys Thr Leu Asp Ala
 965 970 975
 Gly Gln Val Gly Leu Ala Leu Ser Tyr Ala Leu Thr Leu Met Gly Met
 980 985 990
 Phe Gln Trp Cys Val Arg Gln Ser Ala Glu Val Glu Asn Met Met Ile
 995 1000 1005
 Ser Val Glu Arg Val Ile Glu Tyr Thr Asp Leu Glu Lys Glu Ala Pro
 1010 1015 1020
 Trp Glu Tyr Gln Lys Arg Pro Pro Pro Ala Trp Pro His Glu Gly Val
 1025 1030 1035 1040
 Ile Ile Phe Asp Asn Val Asn Phe Met Tyr Ser Pro Gly Gly Pro Leu
 1045 1050 1055
 Val Leu Lys His Leu Thr Ala Leu Ile Lys Ser Gln Glu Lys Val Gly
 1060 1065 1070
 Ile Val Gly Arg Thr Gly Ala Gly Lys Ser Ser Leu Ile Ser Ala Leu
 1075 1080 1085
 Phe Arg Leu Ser Glu Pro Glu Gly Lys Ile Trp Ile Asp Lys Ile Leu
 1090 1095 1100
 Thr Thr Glu Ile Gly Leu His Asp Leu Arg Lys Lys Met Ser Ile Ile

```

1105          1110          1115          1120
Pro Gln Glu Pro Val Leu Phe Thr Gly Thr Met Arg Lys Asn Leu Asp
          1125          1130          1135
Pro Phe Lys Glu His Thr Asp Glu Glu Leu Trp Asn Ala Leu Gln Glu
          1140          1145          1150
Val Gln Leu Lys Glu Thr Ile Glu Asp Leu Pro Gly Lys Met Asp Thr
          1155          1160          1165
Glu Leu Ala Glu Ser Gly Ser Asn Phe Ser Val Gly Gln Arg Gln Leu
          1170          1175          1180
Val Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn Gln Ile Leu Ile Ile
1185          1190          1195          1200
Asp Glu Ala Thr Ala Asn Val Asp Pro Arg Thr Asp Glu Leu Ile Gln
          1205          1210          1215
Lys Lys Ile Arg Glu Lys Phe Ala His Cys Thr Val Leu Thr Ile Ala
          1220          1225          1230
His Arg Leu Asn Thr Ile Ile Asp Ser Asp Lys Ile Met Val Leu Asp
          1235          1240          1245
Ser Gly Arg Leu Lys Glu Tyr Asp Glu Pro Tyr Val Leu Leu Gln Asn
          1250          1255          1260
Lys Glu Ser Leu Phe Tyr Lys Met Val Gln Gln Leu Gly Lys Ala Glu
1265          1270          1275          1280
Ala Ala Ala Leu Thr Glu Thr Ala Lys Gln Val Tyr Phe Lys Arg Asn
          1285          1290          1295
Tyr Pro His Ile Gly His Thr Asp His Met Val Thr Asn Thr Ser Asn
          1300          1305          1310
Gly Gln Pro Ser Thr Leu Thr Ile Phe Glu Thr Ala Leu
          1315          1320          1325

```

<210> 3
 <211> 3184
 <212> DNA
 <213> 人

<220>
 <221> CDS
 <222> (90)... (1238)

```

<400> 3
gggcgccggg attgggaggg cttcttgacg gctgctgggc tggggctaag ggctgctcag 60
tttccttcag cggggcactg ggaagcgcc atg gca ctg cag ggc atc tcg gtc 113
          Met Ala Leu Gln Gly Ile Ser Val
          1          5

gtg gag ctg tcc ggc ctg gcc ccg ggc ccg ttc tgt gct atg gtc ctg 161
Val Glu Leu Ser Gly Leu Ala Pro Gly Pro Phe Cys Ala Met Val Leu
          10          15          20

gct gac ttc ggg gcg cgt gtg gta cgc gtg gac cgg ccc ggc tcc cgc 209
Ala Asp Phe Gly Ala Arg Val Val Arg Val Asp Arg Pro Gly Ser Arg
          25          30          35          40

tac gac gtg agc cgc ttg ggc cgg ggc aag cgc tcg cta gtg ctg gac 257
Tyr Asp Val Ser Arg Leu Gly Arg Gly Lys Arg Ser Leu Val Leu Asp
          45          50          55

ctg aag cag ccg cgg gga gcc gcc gtg ctg cgg cgt ctg tgc aag cgg 305
Leu Lys Gln Pro Arg Gly Ala Ala Val Leu Arg Arg Leu Cys Lys Arg
          60          65          70

tcg gat gtg ctg ctg gag ccc ttc cgc cgc ggt gtc atg gag aaa ctc 353
Ser Asp Val Leu Leu Glu Pro Phe Arg Arg Gly Val Met Glu Lys Leu
          75          80          85

cag ctg ggc cca gag att ctg cag cgg gaa aat cca agg ctt att tat 401
Gln Leu Gly Pro Glu Ile Leu Gln Arg Glu Asn Pro Arg Leu Ile Tyr
          90          95          100

gcc agg ctg agt gga ttt ggc cag tca gga agc ttc tgc cgg tta gct 449
Ala Arg Leu Ser Gly Phe Gly Gln Ser Gly Ser Phe Cys Arg Leu Ala
          105          110          115          120

ggc cac gat atc aac tat ttg gct ttg tca ggt gtt ctc tca aaa att 497
Gly His Asp Ile Asn Tyr Leu Ala Leu Ser Gly Val Leu Ser Lys Ile
          125          130          135

ggc aga agt ggt gag aat ccg tat gcc ccg ctg aat ctc ctg gct gac 545

```


Gly	Arg	Ser	Gly	Glu	Asn	Pro	Tyr	Ala	Pro	Leu	Asn	Leu	Leu	Ala	Asp	
			140					145					150			
ttt	gct	ggt	ggt	ggc	ctt	atg	tgt	gca	ctg	ggc	att	ata	atg	gct	ctt	593
Phe	Ala	Gly	Gly	Gly	Leu	Met	Cys	Ala	Leu	Gly	Ile	Ile	Met	Ala	Leu	
		155					160					165				
ttt	gac	cgc	aca	cgc	act	ggc	aag	ggt	cag	gtc	att	gat	gca	aat	atg	641
Phe	Asp	Arg	Thr	Arg	Thr	Gly	Lys	Gly	Gln	Val	Ile	Asp	Ala	Asn	Met	
	170					175					180					
gtg	gaa	gga	aca	gca	tat	tta	agt	tct	ttt	ctg	tgg	aaa	act	cag	aaa	689
Val	Glu	Gly	Thr	Ala	Tyr	Leu	Ser	Ser	Phe	Leu	Trp	Lys	Thr	Gln	Lys	
185					190					195					200	
tcg	agt	ctg	tgg	gaa	gca	cct	cga	gga	cag	aac	atg	ttg	gat	ggt	gga	737
Ser	Ser	Leu	Trp	Glu	Ala	Pro	Arg	Gly	Gln	Asn	Met	Leu	Asp	Gly	Gly	
			205					210						215		
gca	cct	ttc	tat	acg	act	tac	agg	aca	gca	gat	ggg	gaa	ttc	atg	gct	785
Ala	Pro	Phe	Tyr	Thr	Thr	Tyr	Arg	Thr	Ala	Asp	Gly	Glu	Phe	Met	Ala	
		220					225						230			
gtt	gga	gca	ata	gaa	ccc	cag	ttc	tac	gag	ctg	ctg	atc	aaa	gga	ctt	833
Val	Gly	Ala	Ile	Glu	Pro	Gln	Phe	Tyr	Glu	Leu	Leu	Ile	Lys	Gly	Leu	
	235					240						245				
gga	cta	aag	tct	gat	gaa	ctt	ccc	aat	cag	atg	agc	atg	gat	gat	tgg	881
Gly	Leu	Lys	Ser	Asp	Glu	Leu	Pro	Asn	Gln	Met	Ser	Met	Asp	Asp	Trp	
	250					255					260					
cca	gaa	atg	aag	aag	aag	ttt	gca	gat	gta	ttt	gca	aag	aag	acg	aag	929
Pro	Glu	Met	Lys	Lys	Lys	Phe	Ala	Asp	Val	Phe	Ala	Lys	Lys	Thr	Lys	
265					270					275					280	
gca	gag	tgg	tgt	caa	atc	ttt	gac	ggc	aca	gat	gcc	tgt	gtg	act	ccg	977
Ala	Glu	Trp	Cys	Gln	Ile	Phe	Asp	Gly	Thr	Asp	Ala	Cys	Val	Thr	Pro	
			285					290						295		
gtt	ctg	act	ttt	gag	gag	gtt	gtt	cat	cat	gat	cac	aac	aag	gaa	cgg	1025
Val	Leu	Thr	Phe	Glu	Glu	Val	Val	His	His	Asp	His	Asn	Lys	Glu	Arg	
		300						305					310			
ggc	tcg	ttt	atc	acc	agt	gag	gag	cag	gac	gtg	agc	ccc	cgc	cct	gca	1073
Gly	Ser	Phe	Ile	Thr	Ser	Glu	Glu	Gln	Asp	Val	Ser	Pro	Arg	Pro	Ala	
	315						320					325				
cct	ctg	ctg	tta	aac	acc	cca	gcc	atc	cct	tct	ttc	aaa	agg	gat	cct	1121
Pro	Leu	Leu	Leu	Asn	Thr	Pro	Ala	Ile	Pro	Ser	Phe	Lys	Arg	Asp	Pro	
	330					335					340					
ttc	ata	gga	gaa	cac	act	gag	gag	ata	ctt	gaa	gaa	ttt	gga	ttc	agc	1169
Phe	Ile	Gly	Glu	His	Thr	Glu	Glu	Ile	Leu	Glu	Glu	Phe	Gly	Phe	Ser	
345					350					355					360	
cgc	gaa	gag	att	tat	cag	ctt	aac	tca	gat	aaa	atc	att	gaa	agt	aat	1217
Arg	Glu	Glu	Ile	Tyr	Gln	Leu	Asn	Ser	Asp	Lys	Ile	Ile	Glu	Ser	Asn	
			365					370						375		
aag	gta	aaa	gct	agt	ctc	taa	cttccagggc	cacggctcaa	gigaatttga							1268
Lys	Val	Lys	Ala	Ser	Leu	*										
			380													
atactgcatt	tacagtgtag	agtaacacat	aacattgtat	gcatggaac	atggaggaac											1328
agtattacag	tgctctacca	ctctaatcaa	gaaaagaatt	acagactctg	attctacagt											1388
gatgattgaa	tictaaaaat	ggttatcatt	agggcttttg	atttataaaa	ctttgggtac											1448
ttatactaaa	ttatggtagt	tattctgcct	tccagtittgc	ttgatataat	tggtgatatt											1508
aagattcttg	acttatattt	tgaaatgggtt	ctagtgaata	aggaatgata	tattcttgaa											1568
gacatcgata	tacatttatt	tacactcttg	attctacaat	gtagaaaatg	aggaaatgcc											1628
acaaattgta	tggtgataaa	agtcacgtga	aacagagtga	ttggttgcat	ccaggccttt											1688
tgtcttggtg	tccatgatct	ccctctaagc	acattccaaa	ctttagcaac	agttatcaca											1748
ctttgtaatt	tgcaaaagaaa	agtttcacct	gtattgaatc	agaatgcctt	caactgaaaa											1808
aaacatatcc	aaaataatga	ggaaatgtgt	tggtcacta	cgtagagtcc	agagggacag											1868
tcagttttag	gggtgcctgt	atccagtaac	tcggggcctg	ttccccgtg	gggtctctggg											1928
ctgtcagctt	tcccttctcc	atgtgtttga	tttctctca	ggctggtagc	aagttcttga											1988
tcctatcccc	aacacacagc	aacatccaga	aataaagatc	tcaggacccc	ccagcaagtc											2048
gttttgltc	tccitggact	gagtttaagt	acaagccttt	cttataacctg	tccttgacaa											2108

```

agaagacggg attgtcttta cataaaacca gcctgctcct ggagcttccc tggactcaac 2168
ttcctaagg catgtgagga aggggtagat tccacaatct aatccgggtg ccatcagagt 2228
agagggagta gagaatggat gttgggtagg ccatcaataa ggtccattct gcgcagtatc 2288
tcaactgccg ttcaacaatc gcaagaggaa ggtggagcag gtttcttcat cttacagtig 2348
agaaaacaga gactcagaag ggcttcttag ttcatgtttc ccttagcgcc tcagtgatit 2408
tttcatggtg gcttaggcc aagaaaatat ctaaccattc aatttataaa taattaggtc 2468
cccaacgaat taaatattat gtctaccac cttattagct gcttgaaaaa tataatacac 2528
ataaataaaa aaatatattt ttcatttcta tticattgtt aatcacaaact acttactaag 2588
gagatgtatg cacctatigg acactgtgca acttctcacc tggaatgaga ttggacactg 2648
ctgccctcat ttctgtctcc atgttgggtg ccatatagta ctlgatitit tatcagatgg 2708
cctggaaaac ccagctctac aaaaatatga aattatcaga aggattatag tgcaatctta 2768
tgttgaaaga atgaaciacc tcactagtag ttcacgtgat gctgacaga tgttgagtitt 2828
catgtgtit gtgtgttcaa atttttaaat attctgagat actctgtga ggtcactcta 2888
atgccctggg tgccttggca cagttttaga aataccagtt gaaaaatatt gctcaggaa 2948
atgcaactag gaaggggcag aatcagaatt taagctttca tattctagcc ttcagtctig 3008
ttcttcaacc atttttagga actttcccat tttccagccc aggcattggag 3068
gatcacttga ggccaagagt tgcagaccag cctggggaac ttggctggac ctccgtttct 3128
acgaaataaa aataaaaaa ttatccagga aaaaaaaaaa aaaaaaa 3184

```

<210> 4
 <211> 382
 <212> PRT
 <213> 人

```

<400> 4
Met Ala Leu Gln Gly Ile Ser Val Val Glu Leu Ser Gly Leu Ala Pro
1      5      10      15
Gly Pro Phe Cys Ala Met Val Leu Ala Asp Phe Gly Ala Arg Val Val
20     25     30
Arg Val Asp Arg Pro Gly Ser Arg Tyr Asp Val Ser Arg Leu Gly Arg
35     40     45
Gly Lys Arg Ser Leu Val Leu Asp Leu Lys Gln Pro Arg Gly Ala Ala
50     55     60
Val Leu Arg Arg Leu Cys Lys Arg Ser Asp Val Leu Leu Glu Pro Phe
65     70     75     80
Arg Arg Gly Val Met Glu Lys Leu Gln Leu Gly Pro Glu Ile Leu Gln
85     90     95
Arg Glu Asn Pro Arg Leu Ile Tyr Ala Arg Leu Ser Gly Phe Gly Gln
100    105    110
Ser Gly Ser Phe Cys Arg Leu Ala Gly His Asp Ile Asn Tyr Leu Ala
115    120    125
Leu Ser Gly Val Leu Ser Lys Ile Gly Arg Ser Gly Glu Asn Pro Tyr
130    135    140
Ala Pro Leu Asn Leu Leu Ala Asp Phe Ala Gly Gly Gly Leu Met Cys
145    150    155    160
Ala Leu Gly Ile Ile Met Ala Leu Phe Asp Arg Thr Arg Thr Gly Lys
165    170    175
Gly Gln Val Ile Asp Ala Asn Met Val Glu Gly Thr Ala Tyr Leu Ser
180    185    190
Ser Phe Leu Trp Lys Thr Gln Lys Ser Ser Leu Trp Glu Ala Pro Arg
195    200    205
Gly Gln Asn Met Leu Asp Gly Gly Ala Pro Phe Tyr Thr Thr Tyr Arg
210    215    220
Thr Ala Asp Gly Glu Phe Met Ala Val Gly Ala Ile Glu Pro Gln Phe
225    230    235    240
Tyr Glu Leu Leu Ile Lys Gly Leu Gly Leu Lys Ser Asp Glu Leu Pro
245    250    255
Asn Gln Met Ser Met Asp Asp Trp Pro Glu Met Lys Lys Lys Phe Ala
260    265    270
Asp Val Phe Ala Lys Lys Thr Lys Ala Glu Trp Cys Gln Ile Phe Asp
275    280    285
Gly Thr Asp Ala Cys Val Thr Pro Val Leu Thr Phe Glu Glu Val Val
290    295    300
His His Asp His Asn Lys Glu Arg Gly Ser Phe Ile Thr Ser Glu Glu
305    310    315    320
Gln Asp Val Ser Pro Arg Pro Ala Pro Leu Leu Leu Asn Thr Pro Ala
325    330    335
Ile Pro Ser Phe Lys Arg Asp Pro Phe Ile Gly Glu His Thr Glu Glu
340    345    350
Ile Leu Glu Glu Phe Gly Phe Ser Arg Glu Glu Ile Tyr Gln Leu Asn
355    360    365
Ser Asp Lys Ile Ile Glu Ser Asn Lys Val Lys Ala Ser Leu
370    375    380

```

<210> 5
 <211> 3654

<212> DNA

<213> 人

<220>

<221> CDS

<222> (90)... (1274)

<400> 5

```

gggcgcgcggg attgggaggg cttcttcag gctgctgggc tggggctaag ggctgctcag 60
tttccttcag cggggcactg ggaagcgcc atg gca ctg cag ggc atc tcg gtc 113
                               Met Ala Leu Gln Gly Ile Ser Val
                               1                               5

gtg gag ctg tcc ggc ctg gcc ccg ggc ccg ttc tgt gct atg gtc ctg 161
Val Glu Leu Ser Gly Leu Ala Pro Gly Pro Phe Cys Ala Met Val Leu
10                               15                               20

gct gac ttc ggg gcg cgt gtg gta cgc gtg gac cgg ccc ggc tcc cgc 209
Ala Asp Phe Gly Ala Arg Val Val Arg Val Asp Arg Pro Gly Ser Arg
25                               30                               35                               40

tac gac gtg agc cgc ttg ggc cgg ggc aag cgc tcg cta gtg ctg gac 257
Tyr Asp Val Ser Arg Leu Gly Arg Gly Lys Arg Ser Leu Val Leu Asp
45                               50                               55

ctg aag cag ccg cgg gga gcc gcc gtg ctg cgg cgt ctg tgc aag cgg 305
Leu Lys Gln Pro Arg Gly Ala Ala Val Leu Arg Arg Leu Cys Lys Arg
60                               65                               70

tcg gat gtg ctg ctg gag ccc ttc cgc cgc ggt gtc atg gag aaa ctc 353
Ser Asp Val Leu Leu Glu Pro Phe Arg Arg Gly Val Met Glu Lys Leu
75                               80                               85

cag ctg ggc cca gag att ctg cag cgg gaa aat cca agg ctt att tat 401
Gln Leu Gly Pro Glu Ile Leu Gln Arg Glu Asn Pro Arg Leu Ile Tyr
90                               95                               100

gcc agg ctg agt gga ttt ggc cag tca gga agc ttc tgc cgg tta gct 449
Ala Arg Leu Ser Gly Phe Gly Gln Ser Gly Ser Phe Cys Arg Leu Ala
105                               110                               115                               120

ggc cac gat atc aac tat ttg gct ttg tca ggt gtt ctc tca aaa att 497
Gly His Asp Ile Asn Tyr Leu Ala Leu Ser Gly Val Leu Ser Lys Ile
125                               130                               135

ggc aga agt ggt gag aat ccg tat gcc ccg ctg aat ctc ctg gct gac 545
Gly Arg Ser Gly Glu Asn Pro Tyr Ala Pro Leu Asn Leu Leu Ala Asp
140                               145                               150

ttt gct ggt ggt ggc ctt atg tgt gca ctg ggc att ata atg gct ctt 593
Phe Ala Gly Gly Gly Leu Met Cys Ala Leu Gly Ile Ile Met Ala Leu
155                               160                               165

ttt gac cgc aca cgc act ggc aag ggt cag gtc att gat gca aat atg 641
Phe Asp Arg Thr Arg Thr Gly Lys Gly Gln Val Ile Asp Ala Asn Met
170                               175                               180

gtg gaa gga aca gca tat tta agt tct ttt ctg tgg aaa act cag aaa 689
Val Glu Gly Thr Ala Tyr Leu Ser Ser Phe Leu Trp Lys Thr Gln Lys
185                               190                               195                               200

tcg agt ctg tgg gaa gca cct cga gga cag aac atg ttg gat ggt gga 737
Ser Ser Leu Trp Glu Ala Pro Arg Gly Gln Asn Met Leu Asp Gly Gly
205                               210                               215

gca cct ttc tat acg act tac agg aca gca gat ggg gaa ttc atg gct 785
Ala Pro Phe Tyr Thr Thr Tyr Arg Thr Ala Asp Gly Glu Phe Met Ala
220                               225                               230

gtt gga gca ata gaa ccc cag ttc tac gag ctg ctg atc aaa gga ctt 833
Val Gly Ala Ile Glu Pro Gln Phe Tyr Glu Leu Leu Ile Lys Gly Leu
235                               240                               245

gga cta aag tct gat gaa ctt ccc aat cag atg agc atg gat gat tgg 881
Gly Leu Lys Ser Asp Glu Leu Pro Asn Gln Met Ser Met Asp Asp Trp
250                               255                               260

```

```

cca gaa atg aag aag aag ttt gca gat gta ttt gca aag aag acg aag 929
Pro Glu Met Lys Lys Lys Phe Ala Asp Val Phe Ala Lys Lys Thr Lys
265 270 275 280

gca gag tgg tgt caa atc ttt gac ggc aca gat gcc tgt gtg act ccg 977
Ala Glu Trp Cys Gln Ile Phe Asp Gly Thr Asp Ala Cys Val Thr Pro
285 290 295

gtt ctg act ttt gag gag gtt gtt cat cat gat cac aac aag gaa cgg 1025
Val Leu Thr Phe Glu Glu Val Val His His Asp His Asn Lys Glu Arg
300 305 310

ggc tcg ttt atc acc agt gag gag cag gac gtg agc ccc cgc cct gca 1073
Gly Ser Phe Ile Thr Ser Glu Glu Gln Asp Val Ser Pro Arg Pro Ala
315 320 325

cct ctg ctg tta aac acc cca gcc atc cct tct ttc aaa agg gat cct 1121
Pro Leu Leu Leu Asn Thr Pro Ala Ile Pro Ser Phe Lys Arg Asp Pro
330 335 340

ttc ata gga gaa cac act gag gag ata ctt gaa gaa ttt gga ttc agc 1169
Phe Ile Gly Glu His Thr Glu Glu Ile Leu Glu Glu Phe Gly Phe Ser
345 350 355 360

cgc gaa gag att tat cag ctt aac tca gat aaa atc att gaa agt aat 1217
Arg Glu Glu Ile Tyr Gln Leu Asn Ser Asp Lys Ile Ile Glu Ser Asn
365 370 375

aag gct ggt agc aag ttc tgg atc tta tac cca aca cac agc aac atc 1265
Lys Ala Gly Ser Lys Phe Trp Ile Leu Tyr Pro Thr His Ser Asn Ile
380 385 390

cag aaa taa agatctcagg accccccagc aagtcgtttt gtgtctcctt 1314
Gln Lys *

ggactgagtt aagttacaag cctttcttat acctgicctt gacaaagaag acgggattgt 1374
ctttacataa aaccagcctg ctccctggagc ttccctggac tcaacttcct aaaggcatgt 1434
gaggaagggg tagattccac aatctaatac ggggtgccac agagtagagg gagtagagaa 1494
tggatgttgg gtaggccatc aataaggtcc attctgcgca gtatctcaac tgccgttcaa 1554
caatcgcaag aggaagggtg agcagggttc ttcattctac agttgagaaa acagagactc 1614
agaagggttt cttagttcat gtttccctta gcgcctcagt gattttttca tgggtggctta 1674
ggccaaaaga aatatctaac catccaattt ataaataatt aggtccccaa cgaattaaat 1734
attatgtcct accaacttat tagctgcttg aaaaataata tacacataaa taaaaaata 1794
tatttttcat ttctatttca ttgttaatca caactactta ctaaggagat gtatgcacct 1854
attggacact gtgcaacttc tcacctggaa tgagattgga cactgctgcc ctcatcttct 1914
gctccatggt ggtgtccata tagtacttga ttttttatca gatggcctgg aaaaccctgt 1974
ctcacaaaaa taigaaatta tcagaaggat tatagtgaac tcttatgttg aaagaatgaa 2034
ctacctcact agtagttcac gtgatgtctg acagatgttg agtttcattg tgtttgtgtg 2094
ttcaaatiii taaatattct gagatactct tgtgaggtca ctctaatagcc ctgggtgcct 2154
tggcacagtt ttagaaatac cagttgaaaa tatttgcctc ggaatatgca actaggaagg 2214
ggcagaatca gaatttaagc ttcatatttc tagccttcag tcttgttctt caaccatttt 2274
taggaacttt cccataaggt tatgttttcc agcccaggca tggaggatca cttagggcca 2334
agagttcgag accagcctgg ggaacttggc tggacctccg ttctacgaa ataaaaataa 2394
aaaaattatc caggtatggt ggtgtgtgcc tgtagtccca tctactcaag ggtggggcag 2454
gaggatcact tgagcccagg aattttaggc cacagtgaat taggattgca ccactgcact 2514
ctagcccagg caacagaaca agaacctgic tctaaataaa taataaaaaa taataataat 2574
aaaaaagatg ttttccctac aaaaaagact ttcatattga actcgttcca gcaaggaaaa 2634
tataaccacac tcgaagtctt taaaacagag gaaatttaat ataaagaatt ccactgggtg 2694
cgaaagagca gagaagccca gaagatagtg aggcaacctt galaggaaca taactaggaa 2754
gccaagacca ctctatggt tgcaggggtg atgggaaagc tgggtgtact ggaccagaa 2814
gccaagattg ctgcaccac cttaggagaca tagacacttg caglaataac tcaggagaa 2874
gaaagaaatc tagggaaata tcttggttcc tttcttctc tctctccctt agtcttctta 2934
ccagtgtctc ccattagcca aatctaccta gaagccagaa aacaagggaa ccctggaaat 2994
gtagcccatc aagataaaga gcaccaagg aatatagatc gagcagacag gcagcacaaa 3054
atgcagtggt tatggtttat tcactcagta attcctttag caaatgttta ttgaggatct 3114
actaggtgcc aggtatcatg atacttctg gggataccat aatgaacaaa acagacctgt 3174
tctccgtctt tgaggaaatc aaagacaaac acaggatatg gaataaaccc agaattatct 3234
cattgtaaaa tgtgttaagt accacgagga gaaatatcag ggccatctga cacagctaat 3294
gatttgaaga aggggtgtgac ctgccaccat ttiaaatcta gttatttcac tcttgagctg 3354
tgtgtgtgga aaacttgtag taaaaaatag aatgtctata ttataaaaa gtttatgaaa 3414
agatatcaat ttatttacct ttgacaaac tctatgtaat aaggctttat tactcacggc 3474
catgtgtgtg atcatgtgta atagcatgtg tgtatgagag agagaagcca tatgtaatta 3534
tgtgtataaa cgtctgtgag agagaagcca tgtgtgtgat catgtaaaa aacgtgtgtg 3594
agagaagcca tgtgtgtgat cgtgtaaaa aacgtgtgtg agaagccgtg tgtgatgtgt 3654

```

<210> 6
 <211> 394
 <212> PRT
 <213> 人

<400> 6
 Met Ala Leu Gln Gly Ile Ser Val Val Glu Leu Ser Gly Leu Ala Pro
 1 5 10 15
 Gly Pro Phe Cys Ala Met Val Leu Ala Asp Phe Gly Ala Arg Val Val
 20 25 30
 Arg Val Asp Arg Pro Gly Ser Arg Tyr Asp Val Ser Arg Leu Gly Arg
 35 40 45
 Gly Lys Arg Ser Leu Val Leu Asp Leu Lys Gln Pro Arg Gly Ala Ala
 50 55 60
 Val Leu Arg Arg Leu Cys Lys Arg Ser Asp Val Leu Leu Glu Pro Phe
 65 70 75 80
 Arg Arg Gly Val Met Glu Lys Leu Gln Leu Gly Pro Glu Ile Leu Gln
 85 90 95
 Arg Glu Asn Pro Arg Leu Ile Tyr Ala Arg Leu Ser Gly Phe Gly Gln
 100 105 110
 Ser Gly Ser Phe Cys Arg Leu Ala Gly His Asp Ile Asn Tyr Leu Ala
 115 120 125
 Leu Ser Gly Val Leu Ser Lys Ile Gly Arg Ser Gly Glu Asn Pro Tyr
 130 135 140
 Ala Pro Leu Asn Leu Leu Ala Asp Phe Ala Gly Gly Leu Met Cys
 145 150 155 160
 Ala Leu Gly Ile Ile Met Ala Leu Phe Asp Arg Thr Arg Thr Gly Lys
 165 170 175
 Gly Gln Val Ile Asp Ala Asn Met Val Glu Gly Thr Ala Tyr Leu Ser
 180 185 190
 Ser Phe Leu Trp Lys Thr Gln Lys Ser Ser Leu Trp Glu Ala Pro Arg
 195 200 205
 Gly Gln Asn Met Leu Asp Gly Gly Ala Pro Phe Tyr Thr Thr Tyr Arg
 210 215 220
 Thr Ala Asp Gly Glu Phe Met Ala Val Gly Ala Ile Glu Pro Gln Phe
 225 230 235 240
 Tyr Glu Leu Leu Ile Lys Gly Leu Gly Leu Lys Ser Asp Glu Leu Pro
 245 250 255
 Asn Gln Met Ser Met Asp Asp Trp Pro Glu Met Lys Lys Lys Phe Ala
 260 265 270
 Asp Val Phe Ala Lys Lys Thr Lys Ala Glu Trp Cys Gln Ile Phe Asp
 275 280 285
 Gly Thr Asp Ala Cys Val Thr Pro Val Leu Thr Phe Glu Glu Val Val
 290 295 300
 His His Asp His Asn Lys Glu Arg Gly Ser Phe Ile Thr Ser Glu Glu
 305 310 315 320
 Gln Asp Val Ser Pro Arg Pro Ala Pro Leu Leu Leu Asn Thr Pro Ala
 325 330 335
 Ile Pro Ser Phe Lys Arg Asp Pro Phe Ile Gly Glu His Thr Glu Glu
 340 345 350
 Ile Leu Glu Glu Phe Gly Phe Ser Arg Glu Glu Ile Tyr Gln Leu Asn
 355 360 365
 Ser Asp Lys Ile Ile Glu Ser Asn Lys Ala Gly Ser Lys Phe Trp Ile
 370 375 380
 Leu Tyr Pro Thr His Ser Asn Ile Gln Lys
 385 390

<210> 7
 <211> 1316
 <212> DNA
 <213> 人

<220>
 <221> CDS
 <222> (66)... (932)

<400> 7
 ttgcaggctg ctgggctggg gctaagggt gctcagtttc cttcagcggg gcactgggaa 60
 gcgcc atg gca ctg cag ggc atc tcg gtc gtg gag ctg tcc ggc ctg gcc 110
 Met Ala Leu Gln Gly Ile Ser Val Val Glu Leu Ser Gly Leu Ala
 1 5 10 15
 ccg ggc ccg ttc tgt gct atg gtc ctg gct gac ttc ggg gcg cgt gtg 158
 Pro Gly Pro Phe Cys Ala Met Val Leu Ala Asp Phe Gly Ala Arg Val
 20 25 30

gta cgc gtg gac cgg ccc ggc tcc cgc tac gac gtg agc cgc ttg ggc 206
 Val Arg Val Asp Arg Pro Gly Ser Arg Tyr Asp Val Ser Arg Leu Gly
 35 40 45

cgg ggc aag cgc tgc cta gtg ctg gac ctg aag cag ccg cgg gga gcc 254
 Arg Gly Lys Arg Ser Leu Val Leu Asp Leu Lys Gln Pro Arg Gly Ala
 50 55 60

gcc gtg ctg cgg cgt ctg tgc aag cgg tgc gat gtg ctg ctg gag ccc 302
 Ala Val Leu Arg Arg Leu Cys Lys Arg Ser Asp Val Leu Leu Glu Pro
 65 70 75

ttc cgc cgc ggt gtc atg gag aaa ctc cag ctg ggc cca gag att ctg 350
 Phe Arg Arg Gly Val Met Glu Lys Leu Gln Leu Gly Pro Glu Ile Leu
 80 85 90 95

cag cgg gaa aat cca agg ctt att tat gcc agg ctg agt gga ttt ggc 398
 Gln Arg Glu Asn Pro Arg Leu Ile Tyr Ala Arg Leu Ser Gly Phe Gly
 100 105 110

cag tca gga agc ttc tgc cgg tta gct ggc cac gat atc aac tat ttg 446
 Gln Ser Gly Ser Phe Cys Arg Leu Ala Gly His Asp Ile Asn Tyr Leu
 115 120 125

gct ttg tca ggt gtt ctc tca aaa att ggc aga agt ggt gag aat ccg 494
 Ala Leu Ser Gly Val Leu Ser Lys Ile Gly Arg Ser Gly Glu Asn Pro
 130 135 140

tat gcc ccg ctg aat ctc ctg gct gac ttt gct ggt ggt ggc ctt atg 542
 Tyr Ala Pro Leu Asn Leu Leu Ala Asp Phe Ala Gly Gly Gly Leu Met
 145 150 155

tgt gca ctg ggc att ata atg gct ctt ttt gac cgc aca cgc act ggc 590
 Cys Ala Leu Gly Ile Met Ala Leu Phe Asp Arg Thr Arg Thr Gly
 160 165 170 175

aag ggt cag gtc att gat gca aat atg gtg gaa gga aca gca tat tta 638
 Lys Gly Gln Val Ile Asp Ala Asn Met Val Glu Gly Thr Ala Tyr Leu
 180 185 190

agt tct ttt ctg tgg aaa act cag aaa tgc agt ctg tgg gaa gca cct 686
 Ser Ser Phe Leu Trp Lys Thr Gln Lys Ser Ser Leu Trp Glu Ala Pro
 195 200 205

cga gga cag aac atg ttg gat ggt gga gca cct ttc tat acg act tac 734
 Arg Gly Gln Asn Met Leu Asp Gly Gly Ala Pro Phe Tyr Thr Thr Tyr
 210 215 220

agg aca gca gat ggg gaa ttc atg gct gtt gga gca ata gaa ccc cag 782
 Arg Thr Ala Asp Gly Glu Phe Met Ala Val Gly Ala Ile Glu Pro Gln
 225 230 235

ttc tac gag ctg ctg atc aaa ggt ctg gga gaa ctg atc ttg ctg aaa 830
 Phe Tyr Glu Leu Leu Ile Lys Gly Leu Gly Glu Leu Ile Leu Leu Lys
 240 245 250 255

ata caa cag gaa gca gta tgc tgc cag gca agg caa acc ctc gtc agt 878
 Ile Gln Gln Glu Ala Val Ser Cys Gln Ala Arg Gln Thr Leu Val Ser
 260 265 270

gig aag caa tgg cca tgc ttg cag ccc aag tca tgg ggt ttt gtg tgg 926
 Val Lys Gln Trp Pro Ser Leu Gln Pro Lys Ser Trp Gly Phe Val Trp
 275 280 285

cag taa ccgttggagg cggcaatggc cattttgagt tggatgttt caagccaatg 982
 Gln *

atgattaaaa atttatgita cactcaggct gctgggggat gcttcagttt ccttcacaga 1042
 aaactaaata gtgggaatcc ggaccaataa agaaaggatc aagaagctaa tgaattagtc 1102
 gctaatgctg gtgacagctc tcaacctca tatagggtat gacaaaacag cagaattgct 1162
 aagacaacac cccaaaatgg atcaacctta aaggaaactg ccgttgaact tggctctctc 1222
 acagcagagc agtttgatga atgggtaaaa cctaaggaca tgctgggtcc aaagtgactt 1282
 aaataagttt taatgaaaat aaacgtgtaa aatt 1316

<210> 8

<211> 288

<212> PRT

<213> 人

<400> 8

```

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

```

<210> 9

<211> 3468

<212> DNA

<213> 人

<220>

<221> CDS

<222> (247)... (1632)

<400> 9

```

ccgtcggcgc cgaggagtt agtgcgaccc ggctcggcgc gcacggccaa ggcacgcgcg 60
ctggcacacg cgggcgcgga cagcgcgga cacacacgtt attccaggat ctttgagac 120
ccgaggaaag ccgtgttgc caaaagcaag acaaatgact cacagagaaa aaagatggca 180
gaaccaaggg caactaaagc cgtcagggtc tgaacagctg gtagatgggc tggcttactg 240
aaggac atg att cag act gtc ccg gac cca gca gct cat atc aag gaa 288
      Met Ile Gln Thr Val Pro Asp Pro Ala Ala His Ile Lys Glu
      1      5      10

gcc tta tca gtt gtg agt gag gac cag tcg ttg ttt gag tgt gcc tac 336
Ala Leu Ser Val Val Ser Glu Asp Gln Ser Leu Phe Glu Cys Ala Tyr
15      20      25      30

gga acg cca cac ctg gct aag aca gag atg acc gcg tcc tcc tcc agc 384
Gly Thr Pro His Leu Ala Lys Thr Glu Met Thr Ala Ser Ser Ser Ser
35      40      45

gac tai gga cag act tcc aag atg agc cca cgc gtc cct cag cag gat 432
Asp Tyr Gly Gln Thr Ser Lys Met Ser Pro Arg Val Pro Gln Gln Asp
50      55      60

tgg ctg tct caa ccc cca gcc agg gtc acc atc aaa atg gaa tgt aac 480
Trp Leu Ser Gln Pro Pro Ala Arg Val Thr Ile Lys Met Glu Cys Asn
65      70      75

cct agc cag gtg aat ggc tca agg aac tct cct gat gaa tgc agt gtg 528

```

Pro	Ser	Gln	Val	Asn	Gly	Ser	Arg	Asn	Ser	Pro	Asp	Glu	Cys	Ser	Val	
80						85				90						
gcc	aaa	ggc	ggg	aag	atg	gtg	ggc	agc	cca	gac	acc	gtt	ggg	atg	aac	576
Ala	Lys	Gly	Gly	Lys	Met	Val	Gly	Ser	Pro	Asp	Thr	Val	Gly	Met	Asn	
95				100					105						110	
tac	ggc	agc	tac	atg	gag	gag	aag	cac	atg	cca	ccc	cca	aac	atg	acc	624
Tyr	Gly	Ser	Tyr	Met	Glu	Glu	Lys	His	Met	Pro	Pro	Pro	Asn	Met	Thr	
				115					120					125		
acg	aac	gag	cgc	aga	gtt	atc	gtg	cca	gca	gat	cct	acg	cta	tgg	agt	672
Thr	Asn	Glu	Arg	Arg	Val	Ile	Val	Pro	Ala	Asp	Pro	Thr	Leu	Trp	Ser	
			130					135					140			
aca	gac	cat	gtg	cgg	cag	tgg	ctg	gag	tgg	gcg	gtg	aaa	gaa	tat	ggc	720
Thr	Asp	His	Val	Arg	Gln	Trp	Leu	Glu	Trp	Ala	Val	Lys	Glu	Tyr	Gly	
		145				150						155				
ctt	cca	gac	gtc	aac	atc	tig	tta	ttc	cag	aac	atc	gat	ggg	aag	gaa	768
Leu	Pro	Asp	Val	Asn	Ile	Leu	Leu	Phe	Gln	Asn	Ile	Asp	Gly	Lys	Glu	
	160					165					170					
ctg	tgc	aag	atg	acc	aag	gac	gac	ttc	cag	agg	ctc	acc	ccc	agc	tac	816
Leu	Cys	Lys	Met	Thr	Lys	Asp	Asp	Phe	Gln	Arg	Leu	Thr	Pro	Ser	Tyr	
	175				180				185						190	
aac	gcc	gac	atc	ctt	ctc	tca	cat	ctc	cac	tac	ctc	aga	gag	act	cct	864
Asn	Ala	Asp	Ile	Leu	Leu	Ser	His	Leu	His	Tyr	Leu	Arg	Glu	Thr	Pro	
			195						200					205		
ctt	cca	cat	tig	act	tca	gat	gat	gtt	gat	aaa	gcc	tta	caa	aac	tct	912
Leu	Pro	His	Leu	Thr	Ser	Asp	Asp	Val	Asp	Lys	Ala	Leu	Gln	Asn	Ser	
		210						215					220			
cca	cgg	tta	atg	cat	gct	aga	aac	aca	gat	tta	cca	tat	gag	ccc	ccc	960
Pro	Arg	Leu	Met	His	Ala	Arg	Asn	Thr	Asp	Leu	Pro	Tyr	Glu	Pro	Pro	
	225						230					235				
agg	aga	tca	gcc	tgg	acc	ggt	cac	ggc	cac	ccc	acg	ccc	cag	tcg	aaa	1008
Arg	Arg	Ser	Ala	Trp	Thr	Gly	His	Gly	His	Pro	Thr	Pro	Gln	Ser	Lys	
	240				245					250						
gct	gct	caa	cca	tct	cct	tcc	aca	gtg	ccc	aaa	act	gaa	gac	cag	cgt	1056
Ala	Ala	Gln	Pro	Ser	Pro	Ser	Thr	Val	Pro	Lys	Thr	Glu	Asp	Gln	Arg	
	255				260				265						270	
cct	cag	tta	gat	cct	tat	cag	att	ctt	gga	cca	aca	agt	agc	cgc	ctt	1104
Pro	Gln	Leu	Asp	Pro	Tyr	Gln	Ile	Leu	Gly	Pro	Thr	Ser	Ser	Arg	Leu	
				275					280					285		
gca	aat	cca	ggc	agt	ggc	cag	atc	cag	ctt	tgg	cag	ttc	ctc	ctg	gag	1152
Ala	Asn	Pro	Gly	Ser	Gly	Gln	Ile	Gln	Leu	Trp	Gln	Phe	Leu	Leu	Glu	
		290						295					300			
ctc	ctg	tcg	gac	agc	tcc	aac	tcc	agc	tgc	atc	acc	tgg	gaa	ggc	acc	1200
Leu	Leu	Ser	Asp	Ser	Ser	Asn	Ser	Ser	Cys	Ile	Thr	Trp	Glu	Gly	Thr	
		305					310					315				
aac	ggg	gag	ttc	aag	atg	acg	gat	ccc	gac	gag	gtg	gcc	cgg	cgc	tgg	1248
Asn	Gly	Glu	Phe	Lys	Met	Thr	Asp	Pro	Asp	Glu	Val	Ala	Arg	Arg	Trp	
	320					325					330					
gga	gag	cgg	aag	agc	aaa	ccc	aac	atg	aac	tac	gat	aag	ctc	agc	cgc	1296
Gly	Glu	Arg	Lys	Ser	Lys	Pro	Asn	Met	Asn	Tyr	Asp	Lys	Leu	Ser	Arg	
	335				340					345					350	
gcc	ctc	cgt	tac	tac	tat	gac	aag	aac	atc	atg	acc	aag	gtc	cat	ggg	1344
Ala	Leu	Arg	Tyr	Tyr	Tyr	Asp	Lys	Asn	Ile	Met	Thr	Lys	Val	His	Gly	
			355						360					365		
aag	cgc	tac	gcc	tac	aag	ttc	gac	ttc	cac	ggg	atc	gcc	cag	gcc	ctc	1392
Lys	Arg	Tyr	Ala	Tyr	Lys	Phe	Asp	Phe	His	Gly	Ile	Ala	Gln	Ala	Leu	
		370						375					380			
cag	ccc	cac	ccc	ccg	gag	tca	tct	ctg	tac	aag	tac	ccc	tca	gac	ctc	1440
Gln	Pro	His	Pro	Pro	Glu	Ser	Ser	Leu	Tyr	Lys	Tyr	Pro	Ser	Asp	Leu	


```

      385              390              395
ccg tac atg ggc tcc tat cac gcc cac cca cag aag atg aac ttt gtg 1488
Pro Tyr Met Gly Ser Tyr His Ala His Pro Gln Lys Met Asn Phe Val
  400              405              410

gcg ccc cac cct cca gcc ctc ccc gtg aca tct tcc agt ttt ttt gct 1536
Ala Pro His Pro Pro Ala Leu Pro Val Thr Ser Ser Ser Phe Phe Ala
  415              420              425              430

gcc cca aac cca tac tgg aat tca cca act ggg ggt ata tac ccc aac 1584
Ala Pro Asn Pro Tyr Trp Asn Ser Pro Thr Gly Gly Ile Tyr Pro Asn
      435              440              445

act agg ctc ccc acc agc cat atg cct tct cat ctg ggc act tac tac 1632
Thr Arg Leu Pro Thr Ser His Met Pro Ser His Leu Gly Thr Tyr Tyr
      450              455              460

taaagacctg gcggaggctt ttcccatcag cgtgcattca ccagcccatc gccacaaact 1692
ctatcggaga acatgaatca aaagtgcctc aagaggaatg aaaaaagctt tactggggct 1752
ggggaaggaa gccggggaag agatccaaag actcttggga gggagtact gaagtcctac 1812
tacagaaatg aggaggatgc taataaatgc acgaatatgg acatatcatc tgtggactga 1872
ccttgtaaaa gacagtgtat gtagaagcat gaagtcctaa ggacaaagt ccaagaaaag 1932
tggctttaaag aaatgtataa acttttagagt agagtttggga atcccactaa tgcaaaactgg 1992
gatgaaacta aagcaataga aacaacacag tttagacctt acataccgtt tataatgcc 2052
ttttaaggaa aactacctgt atttaaaaat agaaacatat caaaaacaag agaaaagaca 2112
cgagagagac tgtggcccat caacagacgt tgatatgcaa ctgcatggca tgtgctgttt 2172
tggttgaaat caaatacatt ccgtttgatg gacagctgtc agctttctca aactgtgaag 2232
atgacccaaa gtttccaact cctttacagt attaccggga ctatgaacta aaaggtggga 2292
ctgaggatgt gtatagagtg agcgtgtgat tgtagacaga ggggtgaaga aggaggagga 2352
agaggcagag aaggaggaga ccagggtgg gaaagaaact tctcaagcaa tgaagactgg 2412
actcaggaca ttgggggact gtgtacaatg agttatggag actcaggagg tcatgcagtc 2472
agtggttatac caaaccagt gttaggagaa aggacacagc gtaatggaga aaggggaggt 2532
agtagaatic agaaacaaaa atgcgcatct ctittcttgt ttgtcaaatg aaaattttaa 2592
ctggaattgt ctgatattta agagaaacat tcaggacctc atcattatgt gggggctttg 2652
ttctccacag ggtcaggtaa gagatggcct tcttggctgc cacaatcaga aatcacgcag 2712
gcattttggg taggcggcct ccagttttcc tttagctgc gaacgctgtg cgtttgtcag 2772
aatgaagtat acaagcgaat gtttttcccc ctittttatat aataatata taacttatgc 2832
atttatacac tacgagttga tctcggccag ccaaagacac acgacaaaag agacaatcga 2892
tataatgtgg ccttgaattt taactctgta tgccttaagt ttacaatatg aagttattag 2952
ttcttagaat gcagaatgta tgaataaaa taagcttggc ctagcatggc aaatcagatt 3012
tatacaggag tctgcatttg cacttttttt agtgactaaa gttgcttaat gaaaacatgt 3072
gttgaatgtt gtggattttg tgttataatt tacittgtcc aggaacttgt gcaagggaga 3132
gccaaggaaa taggatgttt ggccacccaa tggcgctcagc ctctccaggt ccttcttgcc 3192
tcccctcctg tcttttattt ctgacccctt ttggaacaga aggacccggg gtttcacatt 3252
ggagcctcca latttatgcc tggaaalgaa agaggcctat gaagctgggg ttgtcattga 3312
gaaattctag ttcagacctt ggtcacaaat cacccttaat tcctgctatg attaaaatac 3372
atttgttgaa cagtgaacaa gctaccactc gtaaggcaaa ctgtattatt actggcaaat 3432
aaagcgtcat ggatagctgc aatttctcac ttitaca 3468

```

<210> 10
 <211> 462
 <212> PRT
 <213> 人

<400> 10
 Met Ile Gln Thr Val Pro Asp Pro Ala Ala His Ile Lys Glu Ala Leu
 1 5 10 15
 Ser Val Val Ser Glu Asp Gln Ser Leu Phe Glu Cys Ala Tyr Gly Thr
 20 25 30
 Pro His Leu Ala Lys Thr Glu Met Thr Ala Ser Ser Ser Ser Asp Tyr
 35 40 45
 Gly Gln Thr Ser Lys Met Ser Pro Arg Val Pro Gln Gln Asp Trp Leu
 50 55 60
 Ser Gln Pro Pro Ala Arg Val Thr Ile Lys Met Glu Cys Asn Pro Ser
 65 70 75 80
 Gln Val Asn Gly Ser Arg Asn Ser Pro Asp Glu Cys Ser Val Ala Lys
 85 90 95
 Gly Gly Lys Met Val Gly Ser Pro Asp Thr Val Gly Met Asn Tyr Gly
 100 105 110
 Ser Tyr Met Glu Glu Lys His Met Pro Pro Pro Asn Met Thr Thr Asn
 115 120 125
 Glu Arg Arg Val Ile Val Pro Ala Asp Pro Thr Leu Trp Ser Thr Asp
 130 135 140
 His Val Arg Gln Trp Leu Glu Trp Ala Val Lys Glu Tyr Gly Leu Pro
 145 150 155 160
 Asp Val Asn Ile Leu Leu Phe Gln Asn Ile Asp Gly Lys Glu Leu Cys

```

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

```

<210> 11

<211> 3459

<212> DNA

<213> 人

<220>

<221> CDS

<222> (247)... (1626)

<400> 11

```

ccgtcggcgc cgaggagatt agigcgaccc ggctcggcgc gcacggccaa ggcacgcgcg 60
ctggcacacg cgggcgcgga caccgcgcga cacacacggt attccaggat ctttggagac 120
ccgaggaaag ccgtgttgac caaaagcaag acaaatgact cacagagaaa aaagaaggca 180
gaaccaaggg caactaaagc cgtcagggtc tgaacagctg gtagatgggc tggcttactg 240
aaggac atg att cag act gtc ccg gac cca gca gct cat atc aag gaa 288

```

```

Met Ile Gln Thr Val Pro Asp Pro Ala Ala His Ile Lys Glu
1 5 10

```

```

gcc tta tca gtt gtg agt gag gac cag tcg ttg ttt gag tgt gcc tac 336
Ala Leu Ser Val Val Ser Glu Asp Gln Ser Leu Phe Glu Cys Ala Tyr
15 20 25 30

```

```

gga acg cca cac ctg gct aag aca gag atg acc gcg tcc tcc tcc agc 384
Gly Thr Pro His Leu Ala Lys Thr Glu Met Thr Ala Ser Ser Ser Ser
35 40 45

```

```

gac tat gga cag act tcc aag atg agc cca cgc gtc cct cag cag gat 432
Asp Tyr Gly Gln Thr Ser Lys Met Ser Pro Arg Val Pro Gln Gln Asp
50 55 60

```

```

tgg ctg tct caa ccc cca gcc agg gtc acc atc aaa atg gaa tgt aac 480
Trp Leu Ser Gln Pro Pro Ala Arg Val Thr Ile Lys Met Glu Cys Asn
65 70 75

```

```

cct agc cag gtc aat ggc tca agg aac tct cct gat gaa tgc agt gtg 528
Pro Ser Gln Val Asn Gly Ser Arg Asn Ser Pro Asp Glu Cys Ser Val
80 85 90

```

gcc aaa ggc ggg aag atg gtg ggc agc cca gac acc gtt ggg atg aac Ala Lys Gly Gly Lys Met Val Gly Ser Pro Asp Thr Val Gly Met Asn 95 100 105 110	576
tac ggc agc tac atg gag gag aag cac atg cca ccc cca aac atg acc Tyr Gly Ser Tyr Met Glu Glu Lys His Met Pro Pro Pro Asn Met Thr 115 120 125	624
acg aac gag cgc aga gtt atc gtg cca gca gat cct acg cta tgg agt Thr Asn Glu Arg Arg Val Ile Val Pro Ala Asp Pro Thr Leu Trp Ser 130 135 140	672
aca gac cat gtg cgg cag tgg ctg gag tgg gcg gtg aaa gaa tat ggc Thr Asp His Val Arg Gln Trp Leu Glu Trp Ala Val Lys Glu Tyr Gly 145 150 155	720
ctt cca gac gtc aac atc ttg tta ttc cag aac atc gat ggg aag gaa Leu Pro Asp Val Asn Ile Leu Leu Phe Gln Asn Ile Asp Gly Lys Glu 160 165 170	768
ctg tgc aag atg acc aag gac gac ttc cag agg ctc acc ccc agc tac Leu Cys Lys Met Thr Lys Asp Asp Phe Gln Arg Leu Thr Pro Ser Tyr 175 180 185 190	816
aac gcc gac atc ctt ctc tca cat ctc cac tac ctc aga gag agg ggt Asn Ala Asp Ile Leu Leu Ser His Leu His Tyr Leu Arg Glu Arg Gly 195 200 205	864
gca gct ttt att ttc cca aat act tca gla tat cct gaa gct acg caa Ala Ala Phe Ile Phe Pro Asn Thr Ser Val Tyr Pro Glu Ala Thr Gln 210 215 220	912
aga att aca act agg cca gat tta cca tat gag ccc ccc agg aga tca Arg Ile Thr Thr Arg Pro Asp Leu Pro Tyr Glu Pro Pro Arg Arg Ser 225 230 235	960
gcc tgg acc ggt cac ggc cac ccc acg ccc cag tcg aaa gct gct caa Ala Trp Thr Gly His Gly His Pro Thr Pro Gln Ser Lys Ala Ala Gln 240 245 250	1008
cca tct cct tcc aca gtg ccc aaa act gaa gac cag cgt cct cag tta Pro Ser Pro Ser Thr Val Pro Lys Thr Glu Asp Gln Arg Pro Gln Leu 255 260 265 270	1056
gat cct tat cag att ctt gga cca aca agt agc cgc ctt gca aat cca Asp Pro Tyr Gln Ile Leu Gly Pro Thr Ser Ser Arg Leu Ala Asn Pro 275 280 285	1104
ggc agt ggc cag atc cag ctt tgg cag ttc ctc ctg gag ctc ctg tcg Gly Ser Gly Gln Ile Gln Leu Trp Gln Phe Leu Leu Glu Leu Leu Ser 290 295 300	1152
gac agc tcc aac tcc agc tgc atc acc tgg gaa ggc acc aac ggg gag Asp Ser Ser Asn Ser Ser Cys Ile Thr Trp Glu Gly Thr Asn Gly Glu 305 310 315	1200
ttc aag atg acg gat ccc gac gag glg gcc cgg cgc tgg gga gag cgg Phe Lys Met Thr Asp Pro Asp Glu Val Ala Arg Arg Trp Gly Glu Arg 320 325 330	1248
aag agc aaa ccc aac atg aac tac gat aag ctc agc cgc gcc ctc cgt Lys Ser Lys Pro Asn Met Asn Tyr Asp Lys Leu Ser Arg Ala Leu Arg 335 340 345 350	1296
tac tac tat gac aag aac atc atg acc aag gtc cat ggg aag cgc tac Tyr Tyr Tyr Asp Lys Asn Ile Met Thr Lys Val His Gly Lys Arg Tyr 355 360 365	1344
gcc tac aag ttc gac ttc cac ggg atc gcc cag gcc ctc cag ccc cac Ala Tyr Lys Phe Asp Phe His Gly Ile Ala Gln Ala Leu Gln Pro His 370 375 380	1392
ccc ccg gag tca tct ctg tac aag tac ccc tca gac ctc ccg tac atg Pro Pro Glu Ser Ser Leu Tyr Lys Tyr Pro Ser Asp Leu Pro Tyr Met 385 390 395	1440
ggc tcc tat cac gcc cac cca cag aag atg aac ttt gtg gcg ccc cac	1488

```

Gly Ser Tyr His Ala His Pro Gln Lys Met Asn Phe Val Ala Pro His
400                               405                               410

ccT cca gcc ctc ccc gTg aca tct tcc agt ttt ttt gct gcc cca aac 1536
Pro Pro Ala Leu Pro Val Thr Ser Ser Ser Phe Phe Ala Ala Pro Asn
415                               420                               425                               430

cca tac tgg aat tca cca act ggg ggt ata tac ccc aac act agg ctc 1584
Pro Tyr Trp Asn Ser Pro Thr Gly Gly Ile Tyr Pro Asn Thr Arg Leu
435                               440                               445

ccc acc agc cat atg cct tct cat ctg ggc act tac tac taa 1626
Pro Thr Ser His Met Pro Ser His Leu Gly Thr Tyr Tyr *
450                               455

agacctggcg gaggcttttc ccatcagcgt gcattcacca gcccacgcc acaaactcta 1686
tcggagaaca tgaatcaaaa gtgcctcaag aggaatgaaa aaagctttac tggggctggg 1746
gaagggaagcc ggggaagaga tccaaagact ctiggaggagg agttactgaa gtcttactac 1806
agaaatgagg aggatgctaa aaatgtcacg aatatggaca tatcatctgt ggactgacct 1866
tgtaaaagac agtgtatgta gaagcatgaa gtcttaagga caaagtgcc aagaaagtgg 1926
tcttaagaaa tgtataaact ttagagtaga gtttggaaac ccactaatgc aaactgggat 1986
gaaactaaag caatagaaac aacacagitt tgacctaaac taccgtttat aatgccattt 2046
taaggaaaac tactgtattt taaaaataga aacatatcaa aaacaagaga aaagacacga 2106
gagagactgt ggcacctcaa cagacgttga tatgcaactg catggcatgt gctgttttgg 2166
ttgaaatcaa atacattccg ttgatggac agctgtcagc tttctcaaac tgtgaagatg 2226
acccaaagtt tccaaactct ttacagtatt accgggacta tgaactaaaa ggtgggactg 2286
aggatgtgta tagagtgagc gtgtgattgt agacagaggg gtgaagaagg aggaggaaga 2346
ggcagagaag gaggagacca gggctgggaa agaaacttct caagcaatga agactggact 2406
caggacattt ggggactgtg tacaatgagt tatggagact cgagggttca tgcagtcagt 2466
gttatacca acccagtggt aggagaaagg acacagcgta atggagaaa gggagtagt 2526
agaattcaga aacaaaaaat cgcatctctt tctttgtttg tcaaatgaaa attttaactg 2586
gaattgtctg atatttaaga gaaacattca ggaccctatc attatgtggg ggctttgttc 2646
tccacagggt caggtaagag atggccttct tggctgccac aatcagaaat cagcaggca 2706
ttttgggtag ggggcttcca gtttctctt gagtcgcgaa cgctgtgcgt ttgtcagaat 2766
gaagtataca agtcaatggt ttccccctt ttatataaat aattatataa cttatgcatt 2826
tatacactac gattgatcti cggccagcca aagacacacg acaaaagaga caatcgatat 2886
aatgtggcct tgaattttaa ctctgtatgc ttaattgtta caatatgaag ttattagttc 2946
ttagaatgca gaatglatgt aataaaaataa gcttggccta gcatggcaaa tcagatttat 3006
acaggagtct gcatttgcac tttttttagt gactaaagti gcttaatgaa aacatgtgct 3066
gaatgttgtg gatttgtgtt tataatttac ttgtccagg aacttgtgca agggagagcc 3126
aaggaaaatag gatgtttggc acccaaatgg cgtcagcctc tccaggtcct tcttgcctcc 3186
cctcctgtct tttatttcta gcccctttg gaacagaagg accccgggtt tcacattgga 3246
gcciccatat ttaTgcctgg aatggaaaga ggcctatgaa gctgggggtt tcattgagaa 3306
attctagttc agcacttggt cacaatcac ccttaattcc tgcTatgatt aaaatacatt 3366
tgttgacag tgaacaagct accactcgta aggcaaaactg tattattact ggcaaataaa 3426
cgctcalgga tagctgcaat ttctcacttt aca 3459

```

<210> 12

<211> 459

<212> PRT

<213> 人

<400> 12

```

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

```

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

<210> 13

<211> 3241

<212> DNA

<213> 人

<220>

<221> CDS

<222> (326)... (1408)

<400> 13

ccgtcggcgc cgaggagatt agtgcgaccc ggctcggcgc gcacggccaa ggcacgcgcg 60
 ctggcacacg cgggcgcgga caccgcgcga cacacacggt attccaggat cttggagac 120
 ccgaggaaag ccgtgttgac caaaagcaag acaaatgact cacagagaaa aaagatggca 180
 gaaccaaggg caactaaagc cgtcagggtc tgaacagctg gtagatgggc tggcttactg 240
 aaggacatga ttacagactgt cccggaccca gcagctcata tcaaggaact ctctgatga 300
 atgcagtgtg gccaaaggcg ggaag atg gtg ggc agc cca gac acc gtt ggg 352
 Met Val Gly Ser Pro Asp Thr Val Gly
 1 5
 atg aac tac ggc agc tac atg gag gag aag cac atg cca ccc cca aac 400
 Met Asn Tyr Gly Ser Tyr Met Glu Glu Lys His Met Pro Pro Pro Asn
 10 15 20 25
 atg acc acg aac gag cgc aga gtt atc gtg cca gca gat cct acg cta 448
 Met Thr Thr Asn Glu Arg Arg Val Ile Val Pro Ala Asp Pro Thr Leu
 30 35 40
 tgg agt aca gac cat gtg cgg cag tgg ctg gag tgg gcg gtg aaa gaa 496
 Trp Ser Thr Asp His Val Arg Gln Trp Leu Glu Trp Ala Val Lys Glu
 45 50 55
 tat ggc ctt cca gac gtc aac atc ttg tta ttc cag aac atc gat ggg 544
 Tyr Gly Leu Pro Asp Val Asn Ile Leu Leu Phe Gln Asn Ile Asp Gly
 60 65 70
 aag gaa ctg tgc aag atg acc aag gac gac ttc cag agg ctc acc ccc 592
 Lys Glu Leu Cys Lys Met Thr Lys Asp Asp Phe Gln Arg Leu Thr Pro
 75 80 85
 agc tac aac gcc gac atc ctt ctc tca cat ctc cac tac ctc aga gag 640
 Ser Tyr Asn Ala Asp Ile Leu Leu Ser His Leu His Tyr Leu Arg Glu

90	95	100	105	
agg ggt gca gct ttt att ttc cca aat act tca gta tat cct gaa gct Arg Gly Ala Ala Phe Ile Phe Pro Asn Thr Ser Val Tyr Pro Glu Ala	110	115	120	688
acg caa aga att aca act agg cca gat tta cca tat gag ccc ccc agg Thr Gln Arg Ile Thr Thr Arg Pro Asp Leu Pro Tyr Glu Pro Pro Arg	125	130	135	736
aga tca gcc tgg acc ggt cac ggc cac ccc acg ccc cag tcg aaa gct Arg Ser Ala Trp Thr Gly His Gly His Pro Thr Pro Gln Ser Lys Ala	140	145	150	784
gct caa cca tct cct tcc aca gtg ccc aaa act gaa gac cag cgt cct Ala Gln Pro Ser Pro Ser Thr Val Pro Lys Thr Glu Asp Gln Arg Pro	155	160	165	832
cag tta gat cct tat cag att ctt gga cca aca agt agc cgc ctt gca Gln Leu Asp Pro Tyr Gln Ile Leu Gly Pro Thr Ser Ser Arg Leu Ala	170	175	180	880
aat cca ggc agt ggc cag atc cag ctt tgg cag ttc ctc ctg gag ctc Asn Pro Gly Ser Gly Gln Ile Gln Leu Trp Gln Phe Leu Leu Glu Leu	190	195	200	928
ctg tcg gac agc tcc aac tcc agc tgc atc acc tgg gaa ggc acc aac Leu Ser Asp Ser Ser Asn Ser Ser Cys Ile Thr Trp Glu Gly Thr Asn	205	210	215	976
ggg gag ttc aag atg acg gat ccc gac gag gtg gcc cgg cgc tgg gga Gly Glu Phe Lys Met Thr Asp Pro Asp Glu Val Ala Arg Arg Trp Gly	220	225	230	1024
gag cgg aag agc aaa ccc aac atg aac tac gat aag ctc agc cgc gcc Glu Arg Lys Ser Lys Pro Asn Met Asn Tyr Asp Lys Leu Ser Arg Ala	235	240	245	1072
ctc cgt tac tac tat gac aag aac atc atg acc aag gtc cat ggg aag Leu Arg Tyr Tyr Tyr Asp Lys Asn Ile Met Thr Lys Val His Gly Lys	250	255	260	1120
cgc tac gcc tac aag ttc gac ttc cac ggg atc gcc cag gcc ctc cag Arg Tyr Ala Tyr Lys Phe Asp Phe His Gly Ile Ala Gln Ala Leu Gln	270	275	280	1168
ccc cac ccc ccg gag tca tct ctg tac aag tac ccc tca gac ctc ccg Pro His Pro Pro Glu Ser Ser Leu Tyr Lys Tyr Pro Ser Asp Leu Pro	285	290	295	1216
tac atg ggc tcc tat cac gcc cac cca cag aag atg aac ttt gtg gcg Tyr Met Gly Ser Tyr His Ala His Pro Gln Lys Met Asn Phe Val Ala	300	305	310	1264
ccc cac cct cca gcc ctc ccc gtg aca tct tcc agt ttt ttt gct gcc Pro His Pro Pro Ala Leu Pro Val Thr Ser Ser Ser Phe Phe Ala Ala	315	320	325	1312
cca aac cca tac tgg aat tca cca act ggg ggt ata tac ccc aac act Pro Asn Pro Tyr Trp Asn Ser Pro Thr Gly Gly Ile Tyr Pro Asn Thr	330	335	340	1360
agg ctc ccc acc agc cat atg cct tct cat ctg ggc act tac tac taa Arg Leu Pro Thr Ser His Met Pro Ser His Leu Gly Thr Tyr Tyr *	350	355	360	1408
agacctggcg gaggttttc ccatcagcgt gcattcacca gcccatcgcc acaaaactcta tcggagaaca tgaatcaaaa gtgcctcaag aggaatgaaa aaagctttac tggggctggg				1468
gaaggaaacc gggaagaga tccaaagact cttgggaggg agttactgaa gtcttactac				1528
agaaatgagg aggatgctaa aaatgtcacg aatatggaca tatcatctgt ggactgacct				1588
tgtaaaagac agtgtatgla gaagcatgaa gtcttaagga caaagtgcc aagaaagtgg				1648
tcitaaagaa tgtataaact ttagagtaga gtttgaate ccactaatgc aaactgggat				1708
gaaactaaag caatagaaac aacacagtgt tgacctaac taccgtttat aatgccattt				1768
taaggaaaac tacctgtatt taaaaataga aacatatcaa aaacaagaga aaagacacga				1828
gagagactgt ggcccatcaa cagacgttga tatgcaactg catggcatgt gctgttttgg				1888
tigaaaatcaa atacattccg ttgatggac agctgtcagc ttctcaaac tgtgaagatg				1948
acccaaagt tccaactcc ttaacgtatt accgggacta tgaactaaaa ggtgggactg				2008
				2068

```

aggatgtgta tagagtgagc gtgtgattgt agacagaggg gtgaagaagg aggaggaaga 2128
ggcagagaag gaggagacca gggctgggaa agaaacttct caagcaatga agactggact 2188
caggacattt ggggactgtg tacaatgagt tatggagact cgagggttca tgcagtcagt 2248
gttatacca aaccagtgtt aggagaaagg acacagcgta atggagaaag gggaagtagt 2308
agaattcaga aacaaaaatg cgcattctct tctttgtttg tcaaatgaaa attttaactg 2368
gaattgtctg atatttaaga gaaacattca ggacctcatc attatgtggg ggctttgttc 2428
tccacagggt caggtaagag atggccttct tggctgccac aatcagaaat cacgcaggca 2488
ttttgggtag gcgccctcca gttttcctt gagtcgcgaa cgcigtgcgt ttgtcagaat 2548
gaagtataca agtcaatgtt tttccccctt ttatataaat aattatataa cttatgcatt 2608
tatacactac gatttgatct cggccagcca aagacacacg acaaaagaga caatcgatat 2668
aatgtggcct tgaattttta cctctgtatg ttaatgttta caatatgaag ttattagttc 2728
ttagaalgca gaatgtatgt aataaaataa gcttggccta gcatggcaaa tcagatttat 2788
acaggagtct gcatttgcac tttttttagt gactaaagtt gcttaatgaa aacatgtgct 2848
gaatgttgtg gatttttgtg tataatttac tttgtccagg aacttgtgca agggagagcc 2908
aaggaaatag gatgtttggc acccaaatgg cgtcagcctc tccaggtcct tcttgccctc 2968
cctctgtctt ttattttcta gcccttttg gaacagaagg accccgggtt tcacattgga 3028
gcctccatat ttatgcctgg aatggaaaga ggcctatgaa gctgggggtg tcattgagaa 3088
atictagttc agcacctggt cacaaatcac ccttaattcc tgcctatgatt aaaatacatt 3148
tgttgaacag tgaacaagct accactcgta aggcaaactg tattattact ggcaaataaa 3208
gcgtcatgga tagctgcaat ttctcacttt aca 3241

```

<210> 14
 <211> 360
 <212> PRT
 <213> 人

```

<400> 14
Met Val Gly Ser Pro Asp Thr Val Gly Met Asn Tyr Gly Ser Tyr Met
1      5      10      15
Glu Glu Lys His Met Pro Pro Pro Asn Met Thr Thr Asn Glu Arg Arg
20     25     30
Val Ile Val Pro Ala Asp Pro Thr Leu Trp Ser Thr Asp His Val Arg
35     40     45
Gln Trp Leu Glu Trp Ala Val Lys Glu Tyr Gly Leu Pro Asp Val Asn
50     55     60
Ile Leu Leu Phe Gln Asn Ile Asp Gly Lys Glu Leu Cys Lys Met Thr
65     70     75     80
Lys Asp Asp Phe Gln Arg Leu Thr Pro Ser Tyr Asn Ala Asp Ile Leu
85     90     95
Leu Ser His Leu His Tyr Leu Arg Glu Arg Gly Ala Ala Phe Ile Phe
100    105    110
Pro Asn Thr Ser Val Tyr Pro Glu Ala Thr Gln Arg Ile Thr Thr Arg
115    120    125
Pro Asp Leu Pro Tyr Glu Pro Pro Arg Arg Ser Ala Trp Thr Gly His
130    135    140
Gly His Pro Thr Pro Gln Ser Lys Ala Ala Gln Pro Ser Pro Ser Thr
145    150    155    160
Val Pro Lys Thr Glu Asp Gln Arg Pro Gln Leu Asp Pro Tyr Gln Ile
165    170    175
Leu Gly Pro Thr Ser Ser Arg Leu Ala Asn Pro Gly Ser Gly Gln Ile
180    185    190
Gln Leu Trp Gln Phe Leu Leu Glu Leu Leu Ser Asp Ser Ser Asn Ser
195    200    205
Ser Cys Ile Thr Trp Glu Gly Thr Asn Gly Glu Phe Lys Met Thr Asp
210    215    220
Pro Asp Glu Val Ala Arg Arg Trp Gly Glu Arg Lys Ser Lys Pro Asn
225    230    235    240
Met Asn Tyr Asp Lys Leu Ser Arg Ala Leu Arg Tyr Tyr Tyr Asp Lys
245    250    255
Asn Ile Met Thr Lys Val His Gly Lys Arg Tyr Ala Tyr Lys Phe Asp
260    265    270
Phe His Gly Ile Ala Gln Ala Leu Gln Pro His Pro Pro Glu Ser Ser
275    280    285
Leu Tyr Lys Tyr Pro Ser Asp Leu Pro Tyr Met Gly Ser Tyr His Ala
290    295    300
His Pro Gln Lys Met Asn Phe Val Ala Pro His Pro Pro Ala Leu Pro
305    310    315    320
Val Thr Ser Ser Ser Phe Phe Ala Ala Pro Asn Pro Tyr Trp Asn Ser
325    330    335
Pro Thr Gly Gly Ile Tyr Pro Asn Thr Arg Leu Pro Thr Ser His Met
340    345    350
Pro Ser His Leu Gly Thr Tyr Tyr
355    360

```

<210> 15
 <211> 3250

<212> DNA

<213> 人

<220>

<221> CDS

<222> (326)... (1417)

<400> 15

```

ccgtcggcgc cgagggagtt agtgcgaccc ggctcggcgc gcacggccaa ggcacgcgcg 60
ctggcacacg cgggcgcgga cacgcgcgga cacacacgtt attccaggat ctttggagac 120
ccgaggaaag ccgtgttgac caaaagcaag acaaatgact cacagagaaa aaagatggca 180
gaaccaaggc caactaaagc cgtcagggtc tgaacagctg gtagatgggc tggcttactg 240
aaggacaatg ttacagactgt cccggaccca gcagctcata tcaaggaact ctccatgatg 300
atgcagtgtg gccaaaggcg ggaag atg gtg ggc agc cca gac acc gtt ggg 352
                               Met Val Gly Ser Pro Asp Thr Val Gly
                               1           5

atg aac tac ggc agc tac atg gag gag aag cac atg cca ccc cca aac 400
Met Asn Tyr Gly Ser Tyr Met Glu Glu Lys His Met Pro Pro Pro Asn
10           15           20           25

atg acc acg aac gag cgc aga gtt atc gtg cca gca gat cct acg cta 448
Met Thr Thr Asn Glu Arg Arg Val Ile Val Pro Ala Asp Pro Thr Leu
30           35           40

tgg agt aca gac cat gtg cgg cag tgg ctg gag tgg gcg gtg aaa gaa 496
Trp Ser Thr Asp His Val Arg Gln Trp Leu Glu Trp Ala Val Lys Glu
45           50           55

tat ggc ctt cca gac gtc aac atc ttg tta ttc cag aac atc gat ggg 544
Tyr Gly Leu Pro Asp Val Asn Ile Leu Leu Phe Gln Asn Ile Asp Gly
60           65           70

aag gaa ctg tgc aag atg acc aag gac gac ttc cag agg ctc acc ccc 592
Lys Glu Leu Cys Lys Met Thr Lys Asp Asp Phe Gln Arg Leu Thr Pro
75           80           85

agc tac aac gcc gac atc ctt ctc tca cat ctc cac tac ctc aga gag 640
Ser Tyr Asn Ala Asp Ile Leu Leu Ser His Leu His Tyr Leu Arg Glu
90           95           100           105

act cct ctt cca cat ttg act tca gat gat gtt gat aaa gcc tta caa 688
Thr Pro Leu Pro His Leu Thr Ser Asp Asp Val Asp Lys Ala Leu Gln
110           115           120

aac tct cca cgg tta atg cat gct aga aac aca gat tta cca tat gag 736
Asn Ser Pro Arg Leu Met His Ala Arg Asn Thr Asp Leu Pro Tyr Glu
125           130           135

ccc ccc agg aga tca gcc tgg acc ggt cac ggc cac ccc acg ccc cag 784
Pro Pro Arg Arg Ser Ala Trp Thr Gly His Gly His Pro Thr Pro Gln
140           145           150

tcg aaa gct gct caa cca tct cct tcc aca gtg ccc aaa act gaa gac 832
Ser Lys Ala Ala Gln Pro Ser Pro Ser Thr Val Pro Lys Thr Glu Asp
155           160           165

cag cgt cct cag tta gat cct tat cag att ctt gga cca aca agt agc 880
Gln Arg Pro Gln Leu Asp Pro Tyr Gln Ile Leu Gly Pro Thr Ser Ser
170           175           180           185

cgc ctt gca aat cca ggc agt ggc cag atc cag ctt tgg cag ttc ctc 928
Arg Leu Ala Asn Pro Gly Ser Gly Gln Ile Gln Leu Trp Gln Phe Leu
190           195           200

ctg gag ctc ctg tcg gac agc tcc aac tcc agc tgc atc acc tgg gaa 976
Leu Glu Leu Leu Ser Asp Ser Ser Asn Ser Ser Cys Ile Thr Trp Glu
205           210           215

ggc acc aac ggg gag ttc aag atg acg gat ccc gac gag gtg gcc cgg 1024
Gly Thr Asn Gly Glu Phe Lys Met Thr Asp Pro Asp Glu Val Ala Arg
220           225           230

cgc tgg gga gag cgg aag agc aaa ccc aac atg aac tac gat aag ctc 1072
Arg Trp Gly Glu Arg Lys Ser Lys Pro Asn Met Asn Tyr Asp Lys Leu
235           240           245

```



```

agc cgc gcc ctc cgt tac tac tat gac aag aac atc atg acc aag gtc 1120
Ser Arg Ala Leu Arg Tyr Tyr Tyr Asp Lys Asn Ile Met Thr Lys Val
250 255 260 265

cat ggg aag cgc tac gcc tac aag ttc gac ttc cac ggg atc gcc cag 1168
His Gly Lys Arg Tyr Ala Tyr Lys Phe Asp Phe His Gly Ile Ala Gln
270 275 280

gcc ctc cag ccc cac ccc ccg gag tca tct ctg tac aag tac ccc tca 1216
Ala Leu Gln Pro His Pro Pro Glu Ser Ser Leu Tyr Lys Tyr Pro Ser
285 290 295

gac ctc ccg tac atg ggc tcc tat cac gcc cac cca cag aag atg aac 1264
Asp Leu Pro Tyr Met Gly Ser Tyr His Ala His Pro Gln Lys Met Asn
300 305 310

tll gtg gcg ccc cac cct cca gcc ctc ccc gtg aca tct tcc agt ttt 1312
Phe Val Ala Pro His Pro Pro Ala Leu Pro Val Thr Ser Ser Ser Phe
315 320 325

ttt gct gcc cca aac cca tac tgg aat tca cca act ggg ggt ata tac 1360
Phe Ala Ala Pro Asn Pro Tyr Trp Asn Ser Pro Thr Gly Gly Ile Tyr
330 335 340 345

ccc aac act agg ctc ccc acc agc cat atg cct tct cat ctg ggc act 1408
Pro Asn Thr Arg Leu Pro Thr Ser His Met Pro Ser His Leu Gly Thr
350 355 360

tac tac taa agacctggcg gaggttttc ccatcagcgt gcattcacca 1457
Tyr Tyr *

```

```

gcccatcgcc acaaaactcta tcggagaaca tgaatcaaaa gtgcctcaag aggaatgaaa 1517
aaagcttttac tggggctggg gaagggaagcc ggggaagaga tccaaagact cttgggaggg 1577
agttactgaa gtcttactac agaaatgagg aggatgctaa aaatgtcacg aatatggaca 1637
tatcatctgt ggactgacct tgtaaaagac agtgtatgta gaagcatgaa gtcttaagga 1697
caaagtgcc aagaaagtgg tcttaagaaa tgtataaact ttagagtaga gtttggaatc 1757
ccactaatgc aaactgggat gaaactaaag caatagaaac aacacagtgt tgacctaaca 1817
taccgtttat aatgccattt taaggaaaac taccgtgtatt taaaaataga aacatatcaa 1877
aaacaagaga aaagacacga gagagactgt ggcccatcaa cagacgttga tatgcaactg 1937
calggcatgt gctgttttgg ttgaaatcaa atacattccg ttgatggac agctgicagc 1997
tttccaaac tgtgaagatg acccaaagt tccaactcct ttacagtatt accgggacta 2057
tgaactaaaa ggtgggactg aggatgtgta tagagtggc gtgtgattgt agacagaggg 2117
gtgaagaagg aggaggaaga ggcagagaag gaggagacca gggcigggaa agaaacttct 2177
caagcaatga agactggact caggacattt ggggactgtg tacaatgagt tatggagact 2237
cgagggttca tgcagtcatg gtatatacaa acccagtgtt aggagaaaagg acacagcgta 2297
alggagaaag ggggaagtgt agaattcaga aacaaaaatg cgcattctct tctttgtttg 2357
tcaaatgaaa attttaactg gaattgtctg atatttaaga gaaacattca ggacctcatt 2417
attatgtggg ggcgtttgtc tccacagggt caggtaagag atggccttct tggctgccac 2477
aatcagaaat cacgcaggca ttttgggtag gcggcctcca gttttccttt gactcgcgaa 2537
cgctgtgcgt ttgtcagaat gaagtataca agtcaatgtt tttcccccct tttatataat 2597
aattatataa cttatgcatt tatacaciac gatttgatct cggccagcca aagacacacg 2657
acaaaagaga caatcgatat aatgtggcct tgaattttta cctgtatgc ttaatgttta 2717
caatatgaag ttatlagttc ttagaatgca gaatgtatgt aataaaataa gcttggccta 2777
gcatggcaaa tcagatttat acaggagtct gcatttgcac tttttttagt gactaaagtt 2837
gctlaatgaa aacatgtgct gaattgtgtg gattttgtgt tataatttac tttgtccagg 2897
aacttgtgca agggagagcc aaggaaatag gatgtttggc acccaaatgg cgtcagcctc 2957
tccaggtect tcttgcctcc cctcctgtct tttatttcta gccctttttg gaacagaagg 3017
acccgggtt tcacattgga gcctccatat ttaigcctgg aatggaaaga ggcctatgaa 3077
gctggggttg tcatlgagaa attctagttc agcaccigtg cacaatcac ccttaattcc 3137
tgctatgatt aaaatacatt tgttgaacag tgaacaagct accactcgt agggcaactg 3197
tattattact ggcaataaaa gcgtcatgga tagctgcaat ttctcatttt aca 3250

```

<210> 16
 <211> 363
 <212> PRT
 <213> 人

<400> 16
 Met Val Gly Ser Pro Asp Thr Val Gly Met Asn Tyr Gly Ser Tyr Met
 1 5 10 15
 Glu Glu Lys His Met Pro Pro Pro Asn Met Thr Thr Asn Glu Arg Arg
 20 25 30
 Val Ile Val Pro Ala Asp Pro Thr Leu Trp Ser Thr Asp His Val Arg
 35 40 45
 Gln Trp Leu Glu Trp Ala Val Lys Glu Tyr Gly Leu Pro Asp Val Asn
 50 55 60

```

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

```

<210> 17

<211> 3304

<212> DNA

<213> 人

<220>

<221> CDS

<222> (104)...(1471)

<400> 17

```

aaatcctatcc gctctaaaca accatcatcaa aactactttc tggctcagaga gaagcaataa 60
ttattatttaa catttatttaa cgaatcaataa acttgatcgc att atg gcc agc act 115
Met Ala Ser Thr
1

```

```

att aag gaa gcc tta tca gtt gtg agt gag gac cag tcg ttg ttt gag 163
Ile Lys Glu Ala Leu Ser Val Val Ser Glu Asp Gln Ser Leu Phe Glu
5      10      15      20

```

```

tgt gcc tac gga acg cca cac ctg gct aag aca gag atg acc gcg tcc 211
Cys Ala Tyr Gly Thr Pro His Leu Ala Lys Thr Glu Met Thr Ala Ser
      25      30      35

```

```

tcc tcc agc gac tat gga cag act tcc aag atg agc cca cgc gtc cct 259
Ser Ser Ser Asp Tyr Gly Gln Thr Ser Lys Met Ser Pro Arg Val Pro
      40      45      50

```

```

cag cag gat tgg ctg tct caa ccc cca gcc agg gtc acc atc aaa atg 307
Gln Gln Asp Trp Leu Ser Gln Pro Pro Ala Arg Val Thr Ile Lys Met
      55      60      65

```

```

gaa tgt aac cct agc cag gtg aat ggc tca agg aac tct cct gat gaa 355
Glu Cys Asn Pro Ser Gln Val Asn Gly Ser Arg Asn Ser Pro Asp Glu
      70      75      80

```

```

tgc agt gtg gcc aaa ggc ggg aag atg gtg ggc agc cca gac acc gtt 403
Cys Ser Val Ala Lys Gly Gly Lys Met Val Gly Ser Pro Asp Thr Val

```

85	90	95	100	
ggg atg aac tac Gly Met Asn Tyr	ggc agc tac atg Gly Ser Tyr Met	gag gag aag cac Glu Glu Lys His	atg cca ccc cca Met Pro Pro Pro	451
	105	110	115	
aac atg acc acg Asn Met Thr Thr	aac gag cgc aga Asn Glu Arg Arg	ggt atc gtg cca Val Ile Val Pro	gca gat cct acg Ala Asp Pro Thr	499
	120	125	130	
cta tgg agt aca Leu Trp Ser Thr	gac cat gtg cgg Asp His Val Arg	cag tgg ctg gag Gln Trp Leu Glu	tgg gcg gtg aaa Trp Ala Val Lys	547
	135	140	145	
gaa tat ggc ctt Glu Tyr Gly Leu	cca gac gtc aac Pro Asp Val Asn	atc ttg tta ttc Ile Leu Leu Phe	cag aac atc gat Gln Asn Ile Asp	595
	150	155	160	
ggg aag gaa ctg Gly Lys Glu Leu	tgc aag atg acc Cys Lys Met Thr	aag gag gac gac Lys Asp Asp Phe	ttc cag agg ctc Gln Arg Leu Thr	643
	165	170	175	180
ccc agc tac aac Pro Ser Tyr Asn	gcc gac atc ctt Ala Asp Ile Leu	ctc tca cat ctc Leu Ser His Leu	cac tac ctc aga His Tyr Leu Arg	691
	185	190	195	
gag act cct ctt Glu Thr Pro Leu	cca cat ttg act Pro His Leu Thr	tca gat gat gtt Ser Asp Asp Val	gat aaa gcc tta Asp Lys Ala Leu	739
	200	205	210	
caa aac tct cca Gln Asn Ser Pro	cgg tta atg cat Arg Leu Met His	gct aga aac aca Ala Arg Asn Thr	gat tta cca tat Asp Leu Pro Tyr	787
	215	220	225	
gag ccc ccc agg Glu Pro Pro Arg	aga tca gcc tgg Arg Ser Ala Trp	acc ggt cac ggc Thr Thr Gly His	cac ccc acg ccc His Pro Thr Pro	835
	230	235	240	
cag tcg aaa gct Gln Ser Lys Ala	gct gct caa cca Ala Ala Gln Pro	tct cct tcc aca Ser Pro Ser Thr	gtg ccc aaa act Val Pro Lys Thr	883
	245	250	255	260
gac cag cgt cct Asp Gln Arg Pro	cag tta gat cct Gln Leu Asp Pro	tat cag att ctt Tyr Gln Ile Leu	gga cca aca agt Gly Pro Thr Ser	931
	265	270	275	
agc cgc ctt gca Ser Arg Leu Ala	aat cca ggc agt Asn Pro Gly Ser	ggc cag atc cag Gly Gln Ile Gln	ctt tgg cag ttc Leu Trp Gln Phe	979
	280	285	290	
ctc ctg gag ctc Leu Leu Glu Leu	ctg tcg gac agc Leu Ser Asp Ser	tcc aac tcc agc Ser Asn Ser Ser	tgc atc acc tgg Cys Ile Thr Trp	1027
	295	300	305	
gaa ggc acc aac Glu Gly Thr Asn	ggg gag ttc aag Gly Glu Phe Lys	atg acg gat ccc Met Thr Asp Pro	gac gag gtg gcc Asp Glu Val Ala	1075
	310	315	320	
cgg cgc tgg gga Arg Arg Trp Gly	gag cgg aag agc Glu Arg Lys Ser	aaa ccc aac atg Pro Asn Met Asn	aac tac gat aag Tyr Asp Lys Lys	1123
	325	330	335	340
ctc agc cgc gcc Leu Ser Arg Ala	ctc cgt tac tac Leu Arg Tyr Tyr	tat gac aag aac Tyr Asp Lys Asn	atc atg acc aag Ile Met Thr Lys	1171
	345	350	355	
gtc cat ggg aag Val His Gly Lys	cgc tac gcc tac Arg Tyr Ala Tyr	aag ttc gac ttc Lys Phe Asp Phe	cac ggg atc gcc His Gly Ile Ala	1219
	360	365	370	
cag gcc ctc cag Gln Ala Leu Gln	ccc cac ccc ccg Pro His Pro Glu	gag tca tct ctg Ser Ser Leu Tyr	tac aag tac ccc Lys Tyr Pro Pro	1267
	375	380	385	
tca gac ctc ccg Ser Asp Leu Pro	tac atg ggc tcc Tyr Met Gly Ser	tat cac gcc cac Tyr His Ala His	cca cag aag atg Pro Gln Lys Met	1315
	390	395	400	

aac ttt gtg gcg ccc cac cct cca gcc ctc ccc gtg aca tct tcc agt 1363
 Asn Phe Val Ala Pro His Pro Pro Ala Leu Pro Val Thr Ser Ser Ser
 405 410 415 420

ttt ttt gct gcc cca aac cca tac tgg aat tca cca act ggg ggt ata 1411
 Phe Phe Ala Ala Pro Asn Pro Tyr Trp Asn Ser Pro Thr Gly Gly Ile
 425 430 435

tac ccc aac act agg ctc ccc acc agc cat atg cct tct cat ctg ggc 1459
 Tyr Pro Asn Thr Arg Leu Pro Thr Ser His Met Pro Ser His Leu Gly
 440 445 450

act tac tac taa agaccctggcg gaggccttttc ccatcagcgt gcattcacca 1511
 Thr Tyr Tyr *
 455

gccccatgcc acaaaactcta tcggagaaca tgaatcaaaa gtcctcaag aggaatgaaa 1571
 aaagctttac tggggctggg gaaggaagcc ggggaagaga tccaaagact ctggggaggg 1631
 agttactgaa gtcttactac agaaatgagg aggatgctaa aaatgtcacg aatatggaca 1691
 tatcatcigt ggactgacct tgtaaaagac agtgtatgta gaagcatgaa gtcttaagga 1751
 caaagtgcc aagaaaatgg tcttaagaaa tglataaact ttagagtaga gtttggaatc 1811
 ccactaatgc aaactgggat gaaactaaag caatagaaac aacacagttt tgacctaaac 1871
 taccgtttat aatgccattt taaggaaaac tacctgtatt taaaataga aacatatcaa 1931
 aaacaagaga aaagacacga gagagactgt ggcctatcaa cagacgttga tatgcaactg 1991
 catggcatgt gctgttttgg ttgaaatcaa atacattccg ttgatggac agctgtcagc 2051
 tttctcaaac tgtgaagatg acccaaagt tccaactcct ttacagtatt accgggacta 2111
 tgaactaaaa ggtgggactg aggatgtgta tagagtgagc gtgtgattgt agacagaggg 2171
 gtgaagaagg agggaggaaga ggcagagaag gaggagacca gggctgggaa agaaacttct 2231
 caagcaatga agactggact caggacattt ggggactgtg tacaatgagt tatggagact 2291
 cgagggttca tgcagtcagt gttataccaa acccagtgtt aggagaaaagg acacagcgta 2351
 atggagaaag gggaagtagt agaattcaga aacaaaaatg cgcattctct tctttgtttg 2411
 tcaaatgaaa attttaactg gaattgtctg atatttaaga gaaacattca ggacctcacc 2471
 attatgtggg ggccttttgc tccacagggt caggtaagag atggccttct tggctgccac 2531
 aatcagaaat cagcgaggca ttttgggtag gcggcctcca gtttctctt gagtgcgcaa 2591
 cgctgtgcgt ttgtcagaat gaagtataca agtcaatgtt tttccccctt tttatataat 2651
 aattatataa cttatgcatt tatacactac gagttgatct cggccagcca aagacacacg 2711
 acaaaaagaga caatcgatat aatgtggcct tgaattttaa ctctgtatgc ttaatgttta 2771
 caatatgaag ttattagttc ttagaatgca gaatgtatgt aataaaataa gcttggccta 2831
 gcatggcaaa tcagatttat acaggagtc tgcatttgcac ttttttagt gactaaagtt 2891
 gcttaatgaa aacatgtgct gaattgtgtg galtttgtgt tataatttac ttgtccagg 2951
 aacttgtgca agggagagcc aaggaaaatag gatgtttggc acccaaatgg cgtcagctc 3011
 tcagggtcct tcttgctcc cctcctgtct tttatttcta gcccttttg gaacagaagg 3071
 accccgggtt tcacattgga gcctccatat ttatgcctgg aatggaaaga ggcctatgaa 3131
 tctggggttg tcattagaaa attctagttc agcacctggt cacaatacac ccttaattcc 3191
 tgciaatgatt aaaatacatt tgttgaacag tgaacaagct accactcgt aggcaaaactg 3251
 tattattact ggcaataaaa gcgtcatgga tagctgcaat ttctcacttt aca 3304

<210> 18
 <211> 455
 <212> PRT
 <213> 人

<400> 18

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

```

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

```

<210> 19
 <211> 662
 <212> DNA
 <213> 人

<220>
 <221> CDS
 <222> (49)... (456)

```

<400> 19
accgccgacg cagacccctc tcigcacgcc agcccgcccg caccaccc atg gcc aca 57
                                     Met Ala Thr
                                     1

gtt cag cag ctg gaa gga aga tgg cgc ctg gtg gac agc aaa ggc ttt 105
Val Gln Gln Leu Glu Gly Arg Trp Arg Leu Val Asp Ser Lys Gly Phe
5      10      15

gat gaa tac atg aag gag cta gga gtg gga ata gct ttg cga aaa atg 153
Asp Glu Tyr Met Lys Glu Leu Gly Val Gly Ile Ala Leu Arg Lys Met
20      25      30      35

ggc gca atg gcc aag cca gat tgt atc atc act tgt gat ggt aaa aac 201
Gly Ala Met Ala Lys Pro Asp Cys Ile Ile Thr Cys Asp Gly Lys Asn
40      45      50

ctc acc ata aaa act gag agc act ttg aaa aca aca cag ttt tct tgt 249
Leu Thr Ile Lys Thr Glu Ser Thr Leu Lys Thr Thr Gln Phe Ser Cys
55      60      65

acc ctg gga gag aag ttt gaa gaa acc aca gct gat ggc aga aaa act 297
Thr Leu Gly Glu Lys Phe Glu Glu Thr Thr Ala Asp Gly Arg Lys Thr
70      75      80

cag act gtc tgc aac ttt aca gat ggt gca ttg gtt cag cat cag gag 345
Gln Thr Val Cys Asn Phe Thr Asp Gly Ala Leu Val Gln His Gln Glu
85      90      95

tgg gat ggg aag gaa agc aca ata aca aga aaa ttg aaa gat ggg aaa 393
Trp Asp Gly Lys Glu Ser Thr Ile Thr Arg Lys Leu Lys Asp Gly Lys

```

100	105	110	115	
tta gtg gtg gag tgt gtc atg aac aat gtc acc tgt act cgg atc tat				441
Leu Val Val Glu Cys Val Met Asn Asn Val Thr Cys Thr Arg Ile Tyr				
120		125	130	
gaa aaa gta gaa taa aaattccatc atcactttgg acaggagtta attaagagaa				496
Glu Lys Val Glu *				
135				
tgaccaagct cagttcaatg agcaaatctc catactgttt ctttcttttt tttttcatta				556
cigtgttcaa ttatctttat cataaacatt ttacatgcag ctatttcaaa gtgtgttgga				616
itaattagga tcatcccttt ggtaataaaa taaatgtgtt tgtgct				662
<210> 20				
<211> 135				
<212> PRT				
<213> 人				
<400> 20				
Met Ala Thr Val Gln Gln Leu Glu Gly Arg Trp Arg Leu Val Asp Ser				
1 5 10 15				
Lys Gly Phe Asp Glu Tyr Met Lys Glu Leu Gly Val Gly Ile Ala Leu				
20 25 30				
Arg Lys Met Gly Ala Met Ala Lys Pro Asp Cys Ile Ile Thr Cys Asp				
35 40 45				
Gly Lys Asn Leu Thr Ile Lys Thr Glu Ser Thr Leu Lys Thr Thr Gln				
50 55 60				
Phe Ser Cys Thr Leu Gly Glu Lys Phe Glu Glu Thr Thr Ala Asp Gly				
65 70 75 80				
Arg Lys Thr Gln Thr Val Cys Asn Phe Thr Asp Gly Ala Leu Val Gln				
85 90 95				
His Gln Glu Trp Asp Gly Lys Glu Ser Thr Ile Thr Arg Lys Leu Lys				
100 105 110				
Asp Gly Lys Leu Val Val Glu Cys Val Met Asn Asn Val Thr Cys Thr				
115 120 125				
Arg Ile Tyr Glu Lys Val Glu				
130 135				
<210> 21				
<211> 4466				
<212> DNA				
<213> 人				
<220>				
<221> CDS				
<222> (103)... (1482)				
<400> 21				
gctgagctgc aggcgcggcg aaacttccct ctacccgccc ggcccgcggc gcgcaccgtt				60
ggcgctggac gcttcctcct tggaaagcgc tctccctcag tt atg gag aaa act				114
Met Glu Lys Thr				
1				
tgt ata gat gca ctt cct ctt act atg aat tct tca gaa aag caa gag				162
Cys Ile Asp Ala Leu Pro Leu Thr Met Asn Ser Ser Glu Lys Gln Glu				
5 10 15 20				
act gta tgt att ttt gga act ggt gat ttt gga aga tca ctg gga ttg				210
Thr Val Cys Ile Phe Gly Thr Gly Asp Phe Gly Arg Ser Leu Gly Leu				
25 30 35				
aaa atg ctc cag tgt ggt tat tct gtt gtt ttt gga agt cga aac ccc				258
Lys Met Leu Gln Cys Gly Tyr Ser Val Val Phe Gly Ser Arg Asn Pro				
40 45 50				
cag aag acc acc cta ctg ccc agt ggt gca gaa gtc ttg agc tat tca				306
Gln Lys Thr Thr Leu Leu Pro Ser Gly Ala Glu Val Leu Ser Tyr Ser				
55 60 65				
gaa gca gcc aag aag tct gac atc ata atc ata gca atc cac aga gag				354
Glu Ala Ala Lys Lys Ser Asp Ile Ile Ile Ile Ala Ile His Arg Glu				
70 75 80				
cat tat gat ttt ctc aca gaa tta act gag gtt ctc aat gga aaa ata				402
His Tyr Asp Phe Leu Thr Glu Leu Thr Glu Val Leu Asn Gly Lys Ile				

85	90	95	100	
ttg gta gac atc agc aac aac ctc aaa atc aat caa tat cca gaa tct Leu Val Asp Ile Ser Asn Asn Leu Lys Ile Asn Gln Tyr Pro Glu Ser	105	110	115	450
aat gca gag tac ctt gct cat ttg gtg cca gga gcc cac gtg gta aaa Asn Ala Glu Tyr Leu Ala His Leu Val Pro Gly Ala His Val Val Lys	120	125	130	498
gca ttt aac acc atc tca gcc tgg gct ctc cag tca gga gca ctg gat Ala Phe Asn Thr Ile Ser Ala Trp Ala Leu Gln Ser Gly Ala Leu Asp	135	140	145	546
gca agt cgg cag gtg ttt gtg tgt gga aat gac agc aaa gcc aag caa Ala Ser Arg Gln Val Phe Val Cys Gly Asn Asp Ser Lys Ala Lys Gln	150	155	160	594
aga gtg atg gat att gtt cgt aat ctt gga ctt act cca atg gat caa Arg Val Met Asp Ile Val Arg Asn Leu Gly Leu Thr Pro Met Asp Gln	165	170	175	642
gga tca ctc atg gca gcc aaa gaa att gaa aag tac ccc ctg cag cta Gly Ser Leu Met Ala Ala Lys Glu Ile Glu Lys Tyr Pro Leu Gln Leu	185	190	195	690
ttt cca atg tgg agg ttc ccc ttc tat ttg tct gct gtg ctg tgt gtc Phe Pro Met Trp Arg Phe Pro Phe Tyr Leu Ser Ala Val Leu Cys Val	200	205	210	738
ttc ttg ttt ttc tat tgt gtt ata aga gac gta atc tac cct tat gtt Phe Leu Phe Phe Tyr Cys Val Ile Arg Asp Val Ile Tyr Pro Tyr Val	215	220	225	786
tat gaa aag aaa gat aat aca ttt cgt atg gct att tcc att cca aat Tyr Glu Lys Lys Asp Asn Thr Phe Arg Met Ala Ile Ser Ile Pro Asn	230	235	240	834
cgt atc ttt cca ata aca gca ctt aca ctg ctt gct ttg gtt tac ctc Arg Ile Phe Pro Ile Thr Ala Leu Thr Leu Leu Ala Leu Val Tyr Leu	245	250	255	882
cct ggt gtt att gct gcc att cta caa ctg tac cga ggc aca aaa tac Pro Gly Val Ile Ala Ala Ile Leu Gln Leu Tyr Arg Gly Thr Lys Tyr	265	270	275	930
cgt cga ttc cca gac tgg ctt gac cac tgg atg ctt tgc cga aag cag Arg Arg Phe Pro Asp Trp Leu Asp His Trp Met Leu Cys Arg Lys Gln	280	285	290	978
ctt ggc ttg gta gct ctg gga ttt gcc ttc ctt cat gtc ctc tac aca Leu Gly Leu Val Ala Leu Gly Phe Ala Phe Leu His Val Leu Tyr Thr	295	300	305	1026
ctt gtg att cct att cga tat tat gta cga tgg aga ttg gga aac tta Leu Val Ile Pro Ile Arg Tyr Tyr Val Arg Trp Arg Leu Gly Asn Leu	310	315	320	1074
acc gtt acc cag gca ata ctc aag aag gag aat cca ttt agc acc tcc Thr Val Thr Gln Ala Ile Leu Lys Lys Glu Asn Pro Phe Ser Thr Ser	325	330	335	1122
tca gcc tgg ctc agt gat tca tai gtg gct ttg gga ata ctt ggg ttt Ser Ala Trp Leu Ser Asp Ser Tyr Val Ala Leu Gly Ile Leu Gly Phe	345	350	355	1170
ttt ctg ttt gta ctc ttg gga atc act tct ttg cca tct gtt agc aat Phe Leu Phe Val Leu Leu Gly Ile Thr Ser Leu Pro Ser Val Ser Asn	360	365	370	1218
gca gtc aac tgg aga gag ttc cga ttt gtc cag tcc aaa ctg ggt tat Ala Val Asn Trp Arg Glu Phe Arg Phe Val Gln Ser Lys Leu Gly Tyr	375	380	385	1266
ttg acc ctg atc ttg tgt aca gcc cac acc ctg gtg tac ggt ggg aag Leu Thr Leu Ile Leu Cys Thr Ala His Thr Leu Val Tyr Gly Gly Lys	390	395	400	1314

aga ttc ctc agc cct tca aat ctc aga tgg tat ctt cct gca gcc tac 1362
 Arg Phe Leu Ser Pro Ser Asn Leu Arg Trp Tyr Leu Pro Ala Ala Tyr
 405 410 415 420

gtg tta ggg ctt atc att cct tgc act gtg ctg gtg atc aag ttt gtc 1410
 Val Leu Gly Leu Ile Ile Pro Cys Thr Val Leu Val Ile Lys Phe Val
 425 430 435

cta atc atg cca tgt gta gac aac acc ctt aca agg atc cgc cag ggc 1458
 Leu Ile Met Pro Cys Val Asp Asn Thr Leu Thr Arg Ile Arg Gln Gly
 440 445 450

tgg gaa agg aac tca aaa cac tag aaaaagcatt gaatggaaaa tcaatatatta 1512
 Trp Glu Arg Asn Ser Lys His *
 455

aaacaaagt caatttagct ggattttctga actatggttt tgaatgttta aagaagaatg 1572
 atgggtacag ttaggaaagt ttttttctta caccgtgact gagggaaaaca ttgcttgtct 1632
 ttgagaaatt gactgacata ctggaagaga acaccatttt atctcaggtt agtgaagaat 1692
 cagtgacagt ccctgactct tattttccca gaggccatgg agctgagatt gagactagcc 1752
 ttgtggtttc acactaaaga gtttcctgt tatgggcaac atgcatgacc taatgtcttg 1812
 caaaatccaa tagaagtatt gcagcttctt tctctggctc aagggtgag ttaagtgaag 1872
 ggaaaaacag cacaatgggt accactgata aaggctttat taggtatatc tgaggaaagt 1932
 ggtcacatga aatgtaaaaa gggaatgagg tttttgttgt tttttggaag taaaggcaaa 1992
 cataaatatt accatgaatt attctagtga aatgaccct tgactttgct tttcttaata 2052
 cagatatatta ctgagaggaa ctatttttat aacacaagaa aaatttaca ttgattaaaa 2112
 gtatccatgt cttggataca tacgtatcta tagagctggc atgtaattct tctctataa 2172
 agaataggta taggaagac tgaataaaaa tggagggata tcccttggg tttcacttgc 2232
 attgtcaaat aagcaaaaga gggttgataa aagttcttga tcaaaaagt caaagaaacc 2292
 agaattttag acagcaagct aaataaatat tgtaaaattg cactatatta ggttaagtat 2352
 tatttaggta ttataatatg ctttgtaaat ttatatctcc aaatattgct caatattttt 2412
 catctattaa attaatctt agtgaataa agtagcttct atactgtct tagtctatta 2472
 taattgtaag gaggtaaat aaatgaatag tctgcaggta taaatttgaa caatgcatag 2532
 atgatcgaat attacgaaa atcatagggc agagaggtgt gaagattcat cattatgtga 2592
 aatttggatc tttctcaaat ccttgcgtgaa atttaggatg gttctcactg ttttctgtg 2652
 ctgatagtac cctttccaag gtgaccttca gggggattaa ccttccctagc tcaagcaatg 2712
 agctaaaagg agccttatgc atgatcttcc cacatatcaa aataactaaa aggcactgag 2772
 ttggcattt tctgcttgc tctgctaaga cctttttttt ttttacttt cattataaca 2832
 taattatcat gacattatca aaaaatgatt aaaaatatatt aaacaacat caacaatcca 2892
 ggaatatttt ctataaaat ttttaaaaat aattgtatct atatatcaa tttacatcc 2952
 tttttcaag ctttgtttt tctaaaggct ttgttttct ttttattatt ttttctttt 3012
 ttatttttt gagacagtct tgcctgtctg ctgaggttgg agtgcagtgg cacgatctca 3072
 gctcactgca acctcttct cccaggttca agtgattctt gtccatcagc ctcccagta 3132
 gctgggacta caggcatgtg ccactatgcc cagctaattt ttgactttt agtagagaca 3192
 gggtttcacc acattgtgtca ggcgtgtctt gaaatgcctg cgtcaagtga tctgcttgc 3252
 tccgcctcac aaagcactgg gat tacaggc atgaatctgg ccttacgtaa tatattttct 3312
 taatggctgc ataataatca atcaaatagg catttttcaa acctcttcc ttattaaaca 3372
 ttagactat atccatttt tactaaaaa aataacattt cagataatat cttgactg 3432
 ataattgtgc caagccattt ctaaagtgc cttatcaatt taattaccat tggatgagg 3492
 tgttgccttc atcgacccat ttagattgt cttttttatt tcaatttgcg tttatttata 3552
 actggttgca aaggtacaca gaacacacgc tcttcaact tatctttgat aaaccaagc 3612
 aaggatcaaa aaagtggac gacattgagt agagtcatgg tatacgtgc tgacctaca 3672
 gtatcagtgg aaaagalaag gaaaatgtca ctactacct atgttatgca aaacagttag 3732
 gtgtgtctgg gctggatact gctcttttac ttgagcattg gttgattaaa gtttaggtac 3792
 catccaggct ggtctagaga agtctttgga gtttaacctg cttttttgt taaagaagag 3852
 agtaattgtt ttatcctggc tcatagtcg tcaccgaaaa tagaaaatgc catccatagg 3912
 taaaatgcig acctatagaa aaaaatgaac tctactttta tagcctagta aaaatgctct 3972
 acctgagtag ttaaaagcaa ttcatgaagc ctgaagctaa agagcactct gatgttttg 4032
 gcataatagc tgcatttcca gacctgacct ttggccccc aaacaagtgc tccaagcccc 4092
 accagctgac caaagaaagc ccaagttctc cttctgtcct tcccacaacc tccctgctcc 4152
 caaaactatg aaattaattt gaccatatta acacagctga ctctccagt ttacttaagg 4212
 tagaaagaat gagtttaca cagatgaaaa taagtgtctt gggcgaactg tattcctttt 4272
 aacagatcca aactatttta catltaaaaa aaaagttaaa ctaaaactct ttactgtcga 4332
 tatgtttcct gtattctaga aaaaatttta cactttcaca ttatttttgt acactttccc 4392
 catgttaagg gatgatggct ttataaatg tgtattcatt aaatgttact ttaaaaaata 4452
 aaaaaaaaa aaaa 4466

<210> 22
 <211> 459
 <212> PRT
 <213> 人

<400> 22
 Met Glu Lys Thr Cys Ile Asp Ala Leu Pro Leu Thr Met Asn Ser Ser
 1 5 10 15
 Glu Lys Gln Glu Thr Val Cys Ile Phe Gly Thr Gly Asp Phe Gly Arg


```

      20      25      30
Ser Leu Gly Leu Lys Met Leu Gln Cys Gly Tyr Ser Val Val Phe Gly
      35      40      45
Ser Arg Asn Pro Gln Lys Thr Thr Leu Leu Pro Ser Gly Ala Glu Val
      50      55      60
Leu Ser Tyr Ser Glu Ala Lys Lys Ser Asp Ile Ile Ile Ala
      65      70      75      80
Ile His Arg Glu His Tyr Asp Phe Leu Thr Glu Leu Thr Glu Val Leu
      85      90      95
Asn Gly Lys Ile Leu Val Asp Ile Ser Asn Asn Leu Lys Ile Asn Gln
      100      105      110
Tyr Pro Glu Ser Asn Ala Glu Tyr Leu Ala His Leu Val Pro Gly Ala
      115      120      125
His Val Val Lys Ala Phe Asn Thr Ile Ser Ala Trp Ala Leu Gln Ser
      130      135      140
Gly Ala Leu Asp Ala Ser Arg Gln Val Phe Val Cys Gly Asn Asp Ser
      145      150      155      160
Lys Ala Lys Gln Arg Val Met Asp Ile Val Arg Asn Leu Gly Leu Thr
      165      170      175
Pro Met Asp Gln Gly Ser Leu Met Ala Ala Lys Glu Ile Glu Lys Tyr
      180      185      190
Pro Leu Gln Leu Phe Pro Met Trp Arg Phe Pro Phe Tyr Leu Ser Ala
      195      200      205
Val Leu Cys Val Phe Leu Phe Phe Tyr Cys Val Ile Arg Asp Val Ile
      210      215      220
Tyr Pro Tyr Val Tyr Glu Lys Lys Asp Asn Thr Phe Arg Met Ala Ile
      225      230      235      240
Ser Ile Pro Asn Arg Ile Phe Pro Ile Thr Ala Leu Thr Leu Leu Ala
      245      250      255
Leu Val Tyr Leu Pro Gly Val Ile Ala Ala Ile Leu Gln Leu Tyr Arg
      260      265      270
Gly Thr Lys Tyr Arg Arg Phe Pro Asp Trp Leu Asp His Trp Met Leu
      275      280      285
Cys Arg Lys Gln Leu Gly Leu Val Ala Leu Gly Phe Ala Phe Leu His
      290      295      300
Val Leu Tyr Thr Leu Val Ile Pro Ile Arg Tyr Tyr Val Arg Trp Arg
      305      310      315      320
Leu Gly Asn Leu Thr Val Thr Gln Ala Ile Leu Lys Lys Glu Asn Pro
      325      330      335
Phe Ser Thr Ser Ser Ala Trp Leu Ser Asp Ser Tyr Val Ala Leu Gly
      340      345      350
Ile Leu Gly Phe Phe Leu Phe Val Leu Leu Gly Ile Thr Ser Leu Pro
      355      360      365
Ser Val Ser Asn Ala Val Asn Trp Arg Glu Phe Arg Phe Val Gln Ser
      370      375      380
Lys Leu Gly Tyr Leu Thr Leu Ile Leu Cys Thr Ala His Thr Leu Val
      385      390      395      400
Tyr Gly Gly Lys Arg Phe Leu Ser Pro Ser Asn Leu Arg Trp Tyr Leu
      405      410      415
Pro Ala Ala Tyr Val Leu Gly Leu Ile Ile Pro Cys Thr Val Leu Val
      420      425      430
Ile Lys Phe Val Leu Ile Met Pro Cys Val Asp Asn Thr Leu Thr Arg
      435      440      445
Ile Arg Gln Gly Trp Glu Arg Asn Ser Lys His
      450      455

```

<210> 23
 <211> 3080
 <212> DNA
 <213> 人

<220>
 <221> CDS
 <222> (199)... (1404)

```

<400> 23
gggggcgggg ccttgcgcgt gcggaggccg gagacgcggc ggcgctggac gcggaggcgc 60
tgggcgcacg gcgcggagcc ggccggagct cgaggccggc ggcggcggga gagcgaccg 120
ggcggcctcg tagcggggcc cggatcccc gagtggcggc cggagcctcg aaaagagatt 180
ctcagcgctg attttgag atg atg ggc ttg gga aac ggg cgt cgc agc atg 231
      Met Met Gly Leu Gly Asn Gly Arg Arg Ser Met
      1          5          10

aag tcg ccg ccc ctc gtg ctg gcc gcc ctg gtg gcc tgc atc atc gtc 279
Lys Ser Pro Pro Leu Val Leu Ala Ala Leu Val Ala Cys Ile Ile Val
      15          20          25

```

tig ggc ttc aac tac tgg att gcg agc tcc cgg agc gtg gac ctc cag Leu Gly Phe Asn Tyr Trp Ile Ala Ser Ser Arg Ser Val Asp Leu Gln 30 35 40	327
aca cgg atc atg gag ctg gaa ggc agg gtc cgc agg gcg gct gca gag Thr Arg Ile Met Glu Leu Glu Gly Arg Val Arg Arg Ala Ala Ala Glu 45 50 55	375
aga ggc gcc gtg gag ctg aag aag aac gag ttc cag gga gag ctg gag Arg Gly Ala Val Glu Leu Lys Lys Asn Glu Phe Gln Gly Glu Leu Glu 60 65 70 75	423
aag cag cgg gag cag ctt gac aaa atc cag tcc agc cac aac ttc cag Lys Gln Arg Glu Gln Leu Asp Lys Ile Gln Ser Ser His Asn Phe Gln 80 85 90	471
ctg gag agc gtc aac aag ctg tac cag gac gaa aag gcg gtt ttg gtg Leu Glu Ser Val Asn Lys Leu Tyr Gln Asp Glu Lys Ala Val Leu Val 95 100 105	519
aat aac atc acc aca ggt gag agg ctc atc cga gtg ctg caa gac cag Asn Asn Ile Thr Thr Gly Glu Arg Leu Ile Arg Val Leu Gln Asp Gln 110 115 120	567
tta aag acc ctg cag agg aat tac ggc agg ctg cag cag gat gtc ctc Leu Lys Thr Leu Gln Arg Asn Tyr Gly Arg Leu Gln Gln Asp Val Leu 125 130 135	615
cag ttt cag aag aac cag acc aac ctg gag agg aag ttc tcc tac gac Gln Phe Gln Lys Asn Gln Thr Asn Leu Glu Arg Lys Phe Ser Tyr Asp 140 145 150 155	663
ctg agc cag tgc atc aat cag atg aag gag gtg aag gaa cag tgt gag Leu Ser Gln Cys Ile Asn Gln Met Lys Glu Val Lys Glu Gln Cys Glu 160 165 170	711
gag cga ata gaa gag gtc acc aaa aag ggg aat gaa gct gta gct tcc Glu Arg Ile Glu Glu Val Thr Lys Lys Gly Asn Glu Ala Val Ala Ser 175 180 185	759
aga gac ctg agt gaa aac aac gac cag aga cag cag ctc caa gcc ctc Arg Asp Leu Ser Glu Asn Asn Asp Gln Arg Gln Gln Leu Gln Ala Leu 190 195 200	807
agt gag cct cag ccc agg ctg cag gca gca ggc ctg cca cac aca gag Ser Glu Pro Gln Pro Arg Leu Gln Ala Ala Gly Leu Pro His Thr Glu 205 210 215	855
gtg cca caa ggg aag gga aac gtg ctt ggt aac agc aag tcc cag aca Val Pro Gln Gly Lys Gly Asn Val Leu Gly Asn Ser Lys Ser Gln Thr 220 225 230 235	903
cca gcc ccc agt tcc gaa gtg gtt ttg gat tca aag aga caa gtt gag Pro Ala Pro Ser Ser Glu Val Val Leu Asp Ser Lys Arg Gln Val Glu 240 245 250	951
aaa gag gaa acc aat gag atc cag gtg gtg aat gag gag cct cag agg Lys Glu Glu Thr Asn Glu Ile Gln Val Val Asn Glu Glu Pro Gln Arg 255 260 265	999
gac agg ctg ccg cag gag cca ggc cgg gag cag gtg gtg gaa gac aga Asp Arg Leu Pro Gln Glu Pro Gly Arg Glu Gln Val Val Glu Asp Arg 270 275 280	1047
cct gta ggt gga aga ggc ttc ggg gga gcc gga gaa ctg ggc cag acc Pro Val Gly Gly Arg Gly Phe Gly Gly Ala Gly Glu Leu Gly Gln Thr 285 290 295	1095
cca cag gtg cag gct gcc ctg tca gtg agc cag gaa aat cca gag atg Pro Gln Val Gln Ala Ala Leu Ser Val Ser Gln Glu Asn Pro Glu Met 300 305 310 315	1143
gag ggc cct gag cga gac cag ctt gtc atc ccc gac gga cag gag gag Glu Gly Pro Glu Arg Asp Gln Leu Val Ile Pro Asp Gly Gln Glu Glu 320 325 330	1191

gag cag gaa gct gcc ggg gaa ggg aga aac cag cag aaa ctg aga gga 1239
 Glu Gln Glu Ala Ala Gly Glu Gly Arg Asn Gln Gln Lys Leu Arg Gly
 335 340 345

gaa gat gac tac aac atg gat gaa aat gaa gca gaa tct gag aca gac 1287
 Glu Asp Tyr Asn Met Asp Glu Asn Glu Ala Glu Ser Glu Thr Asp
 350 355 360

aag caa gca gcc ctg gca ggg aat gac aga aac ata gat gtt ttt aat 1335
 Lys Gln Ala Ala Leu Ala Gly Asn Asp Arg Asn Ile Asp Val Phe Asn
 365 370 375

gtt gaa gat cag aaa aga gac acc ata aat tta ctt gat cag cgt gaa 1383
 Val Glu Asp Gln Lys Arg Asp Thr Ile Asn Leu Leu Asp Gln Arg Glu
 380 385 390 395

aag cgg aat cat aca ctc tga attgaactgg aatcacatat ttcacaacag 1434
 Lys Arg Asn His Thr Leu *
 400

ggccgaagag atgactataa aatgttcatg agggactgaa tactgaaaac tgtgaaatgt 1494
 actaaataaa atgtacatct gaagatgatt attgtgaaat tttagtatgc actttgtgtg 1554
 ggaaaaaatg gaattgttctt ttaaacagct tttggggggt actttggaag tgtctaataa 1614
 ggtgtcacaa tttttgtag taggtatttc gtgagaagtt caacaccaa actggaacat 1674
 agttctctt caagtgttgg cgacagcggg gcttcctgat tctggaatat aactttgtgt 1734
 aaattaacag ccacctatag aagagtcctat ctgctgtgaa ggagagacag agaactctgg 1794
 gticcgtcgt cctgtccacg tgcgtacca agtgcgtggg ccagcctgtt acctgttctc 1854
 actgaaaagt ctggctaagt ctcttgtgtg gtcacttctg attctgacaa tcaatcaatc 1914
 aatggcctag agcactgact gttaacacaa acgtcactag caaagtagca acagctttaa 1974
 gtctaataac aaagctgttc tgtgtgagaa ttttttaaaa ggctacttgt ataataaccc 2034
 ttgtcatttt taatgtacaa aacgctatta agtggcttag aatttgaaca tttgtgtgtc 2094
 ttattttactt tgccttctgt gtgggcaaag caacatcttc cctaaatata tattaccaag 2154
 aaaagcaaga agcagattag gtttttgaca aaacaaacag gccaaaagg ggctgacctg 2214
 gagcagagca tggtagagg caaggcatga gagggcaagt ttgtgttgga cagatctgtg 2274
 cctactttat tactggagta aaagaaaaca aagttcattg atgtcgaagg atatatacag 2334
 tgttagaaat taggactgtt tagaaaaaca ggaatacaat ggttgttttt atcatagtgt 2394
 acacatttag ctgtgtgtaa atgactcaca aaactgattt taaaatcaag ttaatgtgaa 2454
 ttttgaatat tactacttaa tcctaattca caataacaat ggcatttaagg ttgacttga 2514
 gttgttctt agtattattt atggtaaata ggctcttacc acttgcaaat aactggccac 2574
 atcattaatg actgacttcc cagtaaggct ctctaagggg taagtaggag gatccacagg 2634
 atttagatg ctaaggcccc agagatcgtt tgatecaacc ctcttatttt cagaggggaa 2694
 aatggggcct agaagttaac gagcatctag ctgggtcgtt ggcacccctg gctcacaca 2754
 gactcccag tagctgggac tacaggcaca cagtcactga agcaggccct gtttgcaatt 2814
 caggttgcca cctccaactt aaacattctt catatgtgat gtccttagtc actaaggtaa 2874
 aactttccca cccagaaaag gcaacttaga taaaatctta gactacttcc atactcttct 2934
 aagtcctctt ccagcctcac tttagtcct cttgggggtt gataggaatt ttctcttgc 2994
 ttctcaataa agtctctatt catctcatgt ttaatttgta cgcatagaat tgcagagaaa 3054
 taaaatgttc tgttcaactt ataaaa 3080

<210> 24

<211> 401

<212> PRT

<213> 人

<400> 24

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

Val Thr Lys Lys Gly Asn Glu Ala Val Ala Ser Arg Asp Leu Ser Glu
 180 185 190
 Asn Asn Asp Gln Arg Gln Gln Leu Gln Ala Leu Ser Glu Pro Gln Pro
 195 200 205
 Arg Leu Gln Ala Ala Gly Leu Pro His Thr Glu Val Pro Gln Gly Lys
 210 215 220
 Gly Asn Val Leu Gly Asn Ser Lys Ser Gln Thr Pro Ala Pro Ser Ser
 225 230 235 240
 Glu Val Val Leu Asp Ser Lys Arg Gln Val Glu Lys Glu Glu Thr Asn
 245 250 255
 Glu Ile Gln Val Val Asn Glu Glu Pro Gln Arg Asp Arg Leu Pro Gln
 260 265 270
 Glu Pro Gly Arg Glu Gln Val Val Glu Asp Arg Pro Val Gly Gly Arg
 275 280 285
 Gly Phe Gly Gly Ala Gly Glu Leu Gly Gln Thr Pro Gln Val Gln Ala
 290 295 300
 Ala Leu Ser Val Ser Gln Glu Asn Pro Glu Met Glu Gly Pro Glu Arg
 305 310 315 320
 Asp Gln Leu Val Ile Pro Asp Gly Gln Glu Glu Glu Gln Glu Ala Ala
 325 330 335
 Gly Glu Gly Arg Asn Gln Gln Lys Leu Arg Gly Glu Asp Asp Tyr Asn
 340 345 350
 Met Asp Glu Asn Glu Ala Glu Ser Glu Thr Asp Lys Gln Ala Ala Leu
 355 360 365
 Ala Gly Asn Asp Arg Asn Ile Asp Val Phe Asn Val Glu Asp Gln Lys
 370 375 380
 Arg Asp Thr Ile Asn Leu Leu Asp Gln Arg Glu Lys Arg Asn His Thr
 385 390 395 400
 Leu

<210> 25

<211> 4412

<212> DNA

<213> 人

<220>

<221> CDS

<222> (331)... (1974)

<400> 25

gcgaagtgagg ccicgctctc ccgggttagc agccgiagcc agatccgttg aaaggagtgc 60
 agagaggtct cattgcgctc ccgaacagac ctgacgtaga tccgaagtgg cccgcgccat 120
 ctcaactatg aggggacacc cgtaggcggc gggagaggga cgccgcgagg agccaataaa 180
 gctccgcaac cggaagtgtc ttctgggagg ggicgtaccc ggaagtgtgg caccctcccg 240
 gccgcaccgg gaagtgtgat gccaccgccc ctacggggaa gtaatggtat ccggccaatt 300
 gagattcgga gttaaacaag ggaatgtgcag atg gag gtc gga gga gac act gct 354
 Met Glu Val Gly Gly Asp Thr Ala
 1 5

gcc ccg gcc ccc ggg ggc gcg gag gac ttg gag gac acg cag ttc ccc 402
 Ala Pro Ala Pro Gly Gly Ala Glu Asp Leu Glu Asp Thr Gln Phe Pro
 10 15 20

agt gag gaa gct aga gaa ggt gga ggg gtt cac gcg gtc ccg ccg gat 450
 Ser Glu Glu Ala Arg Glu Gly Gly Val His Ala Val Pro Pro Asp
 25 30 35 40

ccc gaa gac gag ggc ctg gag gaa aca gga tcc aag gac aag gac cag 498
 Pro Glu Asp Glu Gly Leu Glu Glu Thr Gly Ser Lys Asp Lys Asp Gln
 45 50 55

cca ccc agc cca tca cca ccg ccc cag tca gag gcc ctg tca agc acc 546
 Pro Pro Ser Pro Ser Pro Pro Pro Gln Ser Glu Ala Leu Ser Ser Thr
 60 65 70

tct cgg ctc tgg agt cct gca gcc cct gag aat agt ccc aca tgt agc 594
 Ser Arg Leu Trp Ser Pro Ala Ala Pro Glu Asn Ser Pro Thr Cys Ser
 75 80 85

cct gag agt agc tct gga ggc cag ggc ggg gac ccc agt gat gag gag 642
 Pro Glu Ser Ser Ser Gly Gly Gln Gly Gly Asp Pro Ser Asp Glu Glu
 90 95 100

tgg cgc agc cag cgg aag cat gtg ttt gtg ctg agt gag gct ggc aag 690
 Trp Arg Ser Gln Arg Lys His Val Phe Val Leu Ser Glu Ala Gly Lys

105	110	115	120	
ccc atc tac tcg cgg tat ggt agt gtg gag gcg ctg tcg gct acc atg Pro Ile Tyr Ser Arg Tyr Gly Ser Val Glu Ala Leu Ser Ala Thr Met	125	130	135	738
ggt gta atg acc gcc ctg gtg tcc ttt gtg cag agt gcg gga gat gcc Gly Val Met Thr Ala Leu Val Ser Phe Val Gln Ser Ala Gly Asp Ala	140	145	150	786
atc cgt gcc atc tac gct gag gac cac aag ctg gtg ttc cta caa cag Ile Arg Ala Ile Tyr Ala Glu Asp His Lys Leu Val Phe Leu Gln Gln	155	160	165	834
ggc cca ctg ttg ctc gtg gcc atg tca cgg act tct cag tca gca gcc Gly Pro Leu Leu Leu Val Ala Met Ser Arg Thr Ser Gln Ser Ala Ala	170	175	180	882
cag ctg cgg ggg gag ctg cta gct gtg cac gca cag atc gtg agc aca Gln Leu Arg Gly Glu Leu Leu Ala Val His Ala Gln Ile Val Ser Thr	185	190	195	930
ctt aca cgt gca agt gtc gcc cgc atc ttc gca cac aag cag aac tat Leu Thr Arg Ala Ser Val Ala Arg Ile Phe Ala His Lys Gln Asn Tyr	205	210	215	978
gac ctc cgc cgc ctg ctg gct ggt tca gag cgc aca ctg gac cga ctt Asp Leu Arg Arg Leu Leu Ala Gly Ser Glu Arg Thr Leu Asp Arg Leu	220	225	230	1026
ctg gac agt atg gag cag gac cca gga gcc ctg ctc ctg ggt gcc gtg Leu Asp Ser Met Glu Gln Asp Pro Gly Ala Leu Leu Leu Gly Ala Val	235	240	245	1074
cgc tgt gtg ccc ctt gcc cgc ccg ctg cga gac gca cta ggt gcg ctc Arg Cys Val Pro Leu Ala Arg Pro Leu Arg Asp Ala Leu Gly Ala Leu	250	255	260	1122
ctc cga cgt tgc aca gcg cct ggc ctg gcg ctg tca gtg ctg gca gta Leu Arg Arg Cys Thr Ala Pro Gly Leu Ala Leu Ser Val Leu Ala Val	265	270	275	1170
ggc ggt cga ctt ata aca gca gcc cag gag cga aat gtg ctg gcc gag Gly Gly Arg Leu Ile Thr Ala Ala Gln Glu Arg Asn Val Leu Ala Glu	285	290	295	1218
tgc cgg ctg gac cca gct gac ctg cag ttg ctg ctc gac tgg gtg ggt Cys Arg Leu Asp Pro Ala Asp Leu Gln Leu Leu Leu Asp Trp Val Gly	300	305	310	1266
gca cca gcc ttt gcg gcg ggt gag gct tgg gca cct gtg tgc ctg ccc Ala Pro Ala Phe Ala Ala Gly Glu Ala Trp Ala Pro Val Cys Leu Pro	315	320	325	1314
cgc ttc aac cct gat ggt ttt ttc tac gcc tac gtg gcc cgc ctg gat Arg Phe Asn Pro Asp Gly Phe Phe Tyr Ala Tyr Val Ala Arg Leu Asp	330	335	340	1362
gct atg cct gtc tgc ctg ctg ctg ctt ggc acc caa cgt gaa gcc ttc Ala Met Pro Val Cys Leu Leu Leu Leu Gly Thr Gln Arg Glu Ala Phe	345	350	355	1410
cat gcc atg gcc gcc tgc cgg cgc ctg gtt gaa gat ggg atg cat gcc His Ala Met Ala Ala Cys Arg Arg Leu Val Glu Asp Gly Met His Ala	365	370	375	1458
ctt ggt gcc atg cgt gcc ctt ggg gag gct gcc agc ttc tct aat gcc Leu Gly Ala Met Arg Ala Leu Gly Glu Ala Ala Ser Phe Ser Asn Ala	380	385	390	1506
tca tca gcc agt gct cct gcc tac agc gtg cag gct gtc ggg gcg ccg Ser Ser Ala Ser Ala Pro Ala Tyr Ser Val Gln Ala Val Gly Ala Pro	395	400	405	1554
ggc ctc cgg cac ttc ctg tat aag ccg ctg gac atc cct gac cac cac Gly Leu Arg His Phe Leu Tyr Lys Pro Leu Asp Ile Pro Asp His His	410	415	420	1602

```

cgc caa ctc ccc cag tti acc agc cct gag cta gag gcc ccc tac agc 1650
Arg Gln Leu Pro Gln Phe Thr Ser Pro Glu Leu Glu Ala Pro Tyr Ser
425 430 435 440

aga gag gag gag cgg cag cgg ctg tgc gac ctg tac cac cgc ctg cat 1698
Arg Glu Glu Glu Arg Gln Arg Leu Ser Asp Leu Tyr His Arg Leu His
445 450 455

gct cgt ctc cac agc acc tcc cga ccc ctg cgc ctc att tac cac gtg 1746
Ala Arg Leu His Ser Thr Ser Arg Pro Leu Arg Leu Ile Tyr His Val
460 465 470

gct gag aag gag aca cta ctg gcc tgg gtg acc tcc aaa ttc gag ctc 1794
Ala Glu Lys Glu Thr Leu Leu Ala Trp Val Thr Ser Lys Phe Glu Leu
475 480 485

tat acc tgc ctc agc cct ctg gtg acc aag gca ggt gca atc ttg gta 1842
Tyr Thr Cys Leu Ser Pro Leu Val Thr Lys Ala Gly Ala Ile Leu Val
490 495 500

gtg acc aaa ctc ctg cgc tgg gtg aag aaa gag gag gac cgg ctc ttc 1890
Val Thr Lys Leu Leu Arg Trp Val Lys Lys Glu Glu Asp Arg Leu Phe
505 510 515 520

att cgt tac cca ccc aag tac tcc aca cca cca gcc acc tct acg gac 1938
Ile Arg Tyr Pro Pro Lys Tyr Ser Thr Pro Pro Ala Thr Ser Thr Asp
525 530 535

caa gct gcc cat aat ggc ttg ttc act gga ctc tga tagttggagc 1984
Gln Ala Ala His Asn Gly Leu Phe Thr Gly Leu *
540 545

1cccagacca ggcagtgctg ggagcaacca cctttgtttt ttaccttctg tctaccctgg 2044
aaatgtgtgt ggggggtgtg ctgtggccag tcatgtcttc cctaagcaat ggggcaaggt 2104
ctgagggccc accgatgaga gagatgggtg cagccgccag gcgagcaggc tgccttccct 2164
gccagtcgat gcacctcccc ctctggggaa atcctttagc ctccctctcc ctccctctg 2224
tctcatctcc tccacttggg tgaigtctca gccctgttca gggactgtcc cctccaaact 2284
tgcttccgtg gtctgcctcc tagtgaatc tcagccctga gtgtccagat ctggccaagg 2344
tgcttagggt ggcccacggg ggtgctggaa ttggcacttc agggccaggc tatgcttggg 2404
actggcctga ggttatttta aagaaaaaaa ctacataaaa ggcctaaaaa taagaccac 2464
aaggatatcc ctttgccttt cttgtacttt ttcatctttt accctgccag aaatgaccgc 2524
ccctcaatgc tggctgctgc taacattaat gagaagggtg ccttcagtgt ccacctgtgg 2584
aaccaggac acagcacctg actgcacaca gtggctgaaa tccagcattt ttacatagga 2644
gatgcactta ggttatcttc ctgtttttac tcatctgtga aacagagata agtaaccctc 2704
tctcatgaac tctttgatga ggatttgtaa acgaaaacag actcgaacta ttgtgtacca 2764
ccacatagca catgcacgtc tgtcccagac ttgacaacc tgcacaagac aagcagccta 2824
aagcaggaga gacctcccta gggttttgtg tgtgtgcaca ctaccctcac tcccaactg 2884
gccattaccc tagttctgcc cttgtttgtg gatttacagc ctcaagggtg tagcatgtgt 2944
gctggcaatc agggcccgag tgtgttctgc gcctgccagc agctgactcc tgatttaacc 3004
gctggcgtaa ccgcggttg caccgatcgc tgcgaaaag cctttcaccc tcacgtggtt 3064
tcttttttaa ccagtcacatc agcagggtc gcgcgcaggc cccgcgttgg aaaatggcgg 3124
ggaagctgaa accctgaat gtggaggcgc cagaagctac tgaggaggct gaaggtagtg 3184
agggcaagtg ggcctgcactc ctttctctcc aaccagggca gaaaggaggg aggatctgtc 3244
ccattacaat aatgaaataa tgatattcta atttttttta ataaaaatgt aagccttttg 3304
tlattgaaga aaaacaattt tttaacglt cagcacagtg gagataaatt aacaggcata 3364
ttcttatcac cgagattaac tttgtcaac tgtagtgtat acgttgttga aagtaacaa 3424
catgtagttc agtacgaaag tcttcaaaca aaagtggggc ggtggggatg ttctcataac 3484
aatcaccgtg ctgcagccac ccttatata cagtaattcc cagtgtcgcc atttcacacc 3544
taacacatat gacactttga tggactctta aacctcctaa tcgaacatgg cctccccac 3604
agcctctttt ctccagggta ttgacttgaa tgacaggcaa aggagggtgg aggggtgtcg 3664
gcctagacac tattggaaat tgtattccat ccaaaagaaa aaaaaatctc aggcggggca 3724
cggaggctca cgcctgtaat ccagcactt tgggaggcgg aggcgggcgg atcacagggt 3784
taggagttcg aggcagcct gaccaacatg gtgaaacacc atctctacta aaaaactata 3844
tatacatata catatgtgtg tatgtgtatt tatatgtatg tatgtatgtg tgtgtatata 3904
tatatatatg tatatgtgta tatatatata tatgtgtata tataatatata tatacacaca 3964
ctaatacagcc gggcgcggtg gtgtggccct glaatcgag ctactcagga agctgaggca 4024
ggagaatcac tlgacaccga gaggcggagg ttgcagttag ccgagatcat gtcactgcac 4084
tccagccitg gtgacagggt gagactcgt ctcaaaagaa aaaaaattct ccttggccg 4144
ggcttgggtg ctatgcctg taatcccagc actttggaag gccgaggcgg gtgggtcacc 4204
tgaggctcagg agtttgagac cagcctggcc agcatggcga aacctcgtct ctactaaaaa 4264
tacaataaatt gtccgggtgt ggtggcgggc gcctglagtc ccagctactg gggaggctga 4324
ggcaggggaa ttgcttgaac ccggaggcga gaggttatcg tgagctgaca ttgcgccact 4384
gcactccagc gcggaagaca gagttagc 4412

```

<210> 26

<211> 547

<212> PRT

<213> 人

<400> 26

```

Met Glu Val Gly Gly Asp Thr Ala Ala Pro Ala Pro Gly Gly Ala Glu
 1      5      10      15
Asp Leu Glu Asp Thr Gln Phe Pro Ser Glu Glu Ala Arg Glu Gly Gly
 20      25      30
Gly Val His Ala Val Pro Pro Asp Pro Glu Asp Glu Gly Leu Glu Glu
 35      40      45
Thr Gly Ser Lys Asp Lys Asp Gln Pro Pro Ser Pro Ser Pro Pro Pro
 50      55      60
Gln Ser Glu Ala Leu Ser Thr Ser Arg Leu Trp Ser Pro Ala Ala
 65      70      75      80
Pro Glu Asn Ser Pro Thr Cys Ser Pro Glu Ser Ser Ser Gly Gly Gln
 85      90      95
Gly Gly Asp Pro Ser Asp Glu Glu Trp Arg Ser Gln Arg Lys His Val
100      105      110
Phe Val Leu Ser Glu Ala Gly Lys Pro Ile Tyr Ser Arg Tyr Gly Ser
115      120      125
Val Glu Ala Leu Ser Ala Thr Met Gly Val Met Thr Ala Leu Val Ser
130      135      140
Phe Val Gln Ser Ala Gly Asp Ala Ile Arg Ala Ile Tyr Ala Glu Asp
145      150      155      160
His Lys Leu Val Phe Leu Gln Gln Gly Pro Leu Leu Leu Val Ala Met
165      170      175
Ser Arg Thr Ser Gln Ser Ala Ala Gln Leu Arg Gly Glu Leu Leu Ala
180      185      190
Val His Ala Gln Ile Val Ser Thr Leu Thr Arg Ala Ser Val Ala Arg
195      200      205
Ile Phe Ala His Lys Gln Asn Tyr Asp Leu Arg Arg Leu Leu Ala Gly
210      215      220
Ser Glu Arg Thr Leu Asp Arg Leu Leu Asp Ser Met Glu Gln Asp Pro
225      230      235      240
Gly Ala Leu Leu Leu Gly Ala Val Arg Cys Val Pro Leu Ala Arg Pro
245      250      255
Leu Arg Asp Ala Leu Gly Ala Leu Leu Arg Arg Cys Thr Ala Pro Gly
260      265      270
Leu Ala Leu Ser Val Leu Ala Val Gly Gly Arg Leu Ile Thr Ala Ala
275      280      285
Gln Glu Arg Asn Val Leu Ala Glu Cys Arg Leu Asp Pro Ala Asp Leu
290      295      300
Gln Leu Leu Leu Asp Trp Val Gly Ala Pro Ala Phe Ala Ala Gly Glu
305      310      315      320
Ala Trp Ala Pro Val Cys Leu Pro Arg Phe Asn Pro Asp Gly Phe Phe
325      330      335
Tyr Ala Tyr Val Ala Arg Leu Asp Ala Met Pro Val Cys Leu Leu Leu
340      345      350
Leu Gly Thr Gln Arg Glu Ala Phe His Ala Met Ala Ala Cys Arg Arg
355      360      365
Leu Val Glu Asp Gly Met His Ala Leu Gly Ala Met Arg Ala Leu Gly
370      375      380
Glu Ala Ala Ser Phe Ser Asn Ala Ser Ser Ala Ser Ala Pro Ala Tyr
385      390      395      400
Ser Val Gln Ala Val Gly Ala Pro Gly Leu Arg His Phe Leu Tyr Lys
405      410      415
Pro Leu Asp Ile Pro Asp His His Arg Gln Leu Pro Gln Phe Thr Ser
420      425      430
Pro Glu Leu Glu Ala Pro Tyr Ser Arg Glu Glu Glu Arg Gln Arg Leu
435      440      445
Ser Asp Leu Tyr His Arg Leu His Ala Arg Leu His Ser Thr Ser Arg
450      455      460
Pro Leu Arg Leu Ile Tyr His Val Ala Glu Lys Glu Thr Leu Leu Ala
465      470      475      480
Trp Val Thr Ser Lys Phe Glu Leu Tyr Thr Cys Leu Ser Pro Leu Val
485      490      495
Thr Lys Ala Gly Ala Ile Leu Val Val Thr Lys Leu Leu Arg Trp Val
500      505      510
Lys Lys Glu Glu Asp Arg Leu Phe Ile Arg Tyr Pro Pro Lys Tyr Ser
515      520      525
Thr Pro Pro Ala Thr Ser Thr Asp Gln Ala Ala His Asn Gly Leu Phe
530      535      540
Thr Gly Leu
545

```

<210> 27

<211> 6079

<212> DNA

<213> 人

<220>

<221> CDS

<222> (121)... (1911)

<400> 27

```

ggggacggcg cccacaccgc gagtcacttg tcagccctgt ctgaggcgga ggcagccccg 60
cgccgcgcgc gacccgagca tatttcattt tctgtcattg gactttgagc cattagaacc 120
atg agc aac tac agt gtg tca ctg gtt ggc cca gct cct tgg ggt ttc 168
Met Ser Asn Tyr Ser Val Ser Leu Val Gly Pro Ala Pro Trp Gly Phe
1 5 10 15

cgg ctg cag ggc ggt aag gat ttc aac atg cct ctg aca atc tct agt 216
Arg Leu Gln Gly Lys Asp Phe Asn Met Pro Leu Thr Ile Ser Ser
20 25 30

cta aaa gat ggc ggc aag gca gcc cag gca aat gta aga ata ggc gat 264
Leu Lys Asp Gly Gly Lys Ala Ala Gln Ala Asn Val Arg Ile Gly Asp
35 40 45

gtg gtt ctg agc att gat gga ata aat gca caa gga atg act cat ctt 312
Val Val Leu Ser Ile Asp Gly Ile Asn Ala Gln Met Thr His Leu
50 55 60

gaa gcc cag aat aag att aag ggt tgt aca ggc tct ttg aat atg act 360
Glu Ala Gln Asn Lys Ile Lys Gly Cys Thr Gly Ser Leu Asn Met Thr
65 70 75 80

ctg caa aga gca tct gct gca ccc aag cct gag ccg gtt cct gtt caa 408
Leu Gln Arg Ala Ser Ala Ala Pro Lys Pro Glu Pro Val Pro Val Gln
85 90 95

aag gga gaa cct aaa gaa gta gtt aaa cct gtg ccc att aca tct cct 456
Lys Gly Glu Pro Lys Glu Val Val Lys Pro Val Pro Ile Thr Ser Pro
100 105 110

gct gtg tcc aaa gtc act tcc aca aac aac atg gcc tac aat aag gca 504
Ala Val Ser Lys Val Thr Ser Thr Asn Asn Met Ala Tyr Asn Lys Ala
115 120 125

cca cgg cct ttt ggt tct gtg tct tca cca aaa gtc aca tcc atc cca 552
Pro Arg Pro Phe Gly Ser Val Ser Ser Pro Lys Val Thr Ser Ile Pro
130 135 140

tca cca tcg tct gcc ttc acc cca gcc cat gcg acc acc tca tca cat 600
Ser Pro Ser Ser Ala Phe Thr Pro Ala His Ala Thr Thr Ser Ser His
145 150 155 160

gct tcc cct tca ccc gtg gct gcc gtc act cct ccc ctg ttc gct gca 648
Ala Ser Pro Ser Pro Val Ala Ala Val Thr Pro Pro Leu Phe Ala Ala
165 170 175

tct gga ctg cat gct aat gcc aat ctt agt gct gac cag tct cca tct 696
Ser Gly Leu His Ala Asn Ala Asn Leu Ser Ala Asp Gln Ser Pro Ser
180 185 190

gca ctg agc gct ggt aaa act gca gtt aat gtc cca cgg cag ccc aca 744
Ala Leu Ser Ala Gly Lys Thr Ala Val Asn Val Pro Arg Gln Pro Thr
195 200 205

gtc acc agc gtg tgt tcc gag act tct cag gag cta gca gag gga cag 792
Val Thr Ser Val Cys Ser Glu Thr Ser Gln Glu Leu Ala Glu Gly Gln
210 215 220

aga aga gga tcc cag ggt gac agt aaa cag caa aat ggc cca cca aga 840
Arg Arg Gly Ser Gln Gly Asp Ser Lys Gln Gln Asn Gly Pro Pro Arg
225 230 235 240

aaa cac att gtg gag cgc tat aca gag ttt tat cat gta ccc act cac 888
Lys His Ile Val Glu Arg Tyr Thr Glu Phe Tyr His Val Pro Thr His
245 250 255

agt gat gcc agc aag aag aga ctg att gag gat act gaa gac tgg cgt 936
Ser Asp Ala Ser Lys Lys Arg Leu Ile Glu Asp Thr Glu Asp Trp Arg

```


260	265	270	
cca agg act gga aca act cag tct cgc tct ttc cga atc ctt gcc cag Pro Arg Thr Gly Thr Thr Gln Ser Arg Ser Phe Arg Ile Leu Ala Gln	275	280	285 984
atc act ggg act gaa cat ttg aaa gaa tct gaa gcc gat aat aca aag Ile Thr Gly Thr Glu His Leu Lys Glu Ser Glu Ala Asp Asn Thr Lys	290	295	300 1032
aag gca aat aac tct cag gag cct tct ccg cag ttg gct tcc tcg gta Lys Ala Asn Asn Ser Gln Glu Pro Ser Pro Gln Leu Ala Ser Ser Val	305	310	315 320 1080
gct tcc aca cgg agc atg ccc gag agc ctg gac agc cca acc tct ggc Ala Ser Thr Arg Ser Met Pro Glu Ser Leu Asp Ser Pro Thr Ser Gly	325	330	335 1128
aga cca ggg gtt acc agc ctc aca act gca gct gcc ttc aag cct gta Arg Pro Gly Val Thr Ser Leu Thr Thr Ala Ala Ala Phe Lys Pro Val	340	345	350 1176
gga tcc act ggc gtc atc aag tca cca agc tgg caa cgg cca aac caa Gly Ser Thr Gly Val Ile Lys Ser Pro Ser Trp Gln Arg Pro Asn Gln	355	360	365 1224
gga gta cct tcc act gga aga atc tca aac agc gct act tac tca gga Gly Val Pro Ser Thr Gly Arg Ile Ser Asn Ser Ala Thr Tyr Ser Gly	370	375	380 1272
tca gtg gca cca gcc aac tca gct ttg gga caa acc cag cca agt gac Ser Val Ala Pro Ala Asn Ser Ala Leu Gly Gln Thr Gln Pro Ser Asp	385	390	395 400 1320
cag gac act tta gtg caa aga gct gag cac att cca gca ggg aaa cga Gln Asp Thr Leu Val Gln Arg Ala Glu His Ile Pro Ala Gly Lys Arg	405	410	415 1368
act ccg atg tgc gcc cat tgt aac cag gtc atc aga gga cca ttc tta Thr Pro Met Cys Ala His Cys Asn Gln Val Ile Arg Gly Pro Phe Leu	420	425	430 1416
gtg gca ctg ggg aaa tct tgg cac cca gaa gaa ttc aac tgc gct cac Val Ala Leu Gly Lys Ser Trp His Pro Glu Glu Phe Asn Cys Ala His	435	440	445 1464
tgc aaa aat aca atg gcc tac att gga ttt gta gag gag aaa gga gcc Cys Lys Asn Thr Met Ala Tyr Ile Gly Phe Val Glu Glu Lys Gly Ala	450	455	460 1512
ctg tat tgt gag ctg tgc tat gag aaa ttc ttt gcc cct gaa tgt ggt Leu Tyr Cys Glu Leu Cys Tyr Glu Lys Phe Phe Ala Pro Glu Cys Gly	465	470	475 480 1560
cga tgc caa agg aag atc ctt gga gaa gtc atc aat gcg ttg aaa caa Arg Cys Gln Arg Ile Leu Gly Glu Val Ile Asn Ala Leu Lys Gln	485	490	495 1608
act tgg cat gtt tcc tgt ttt gtg tgt gta gcc tgt gga aag ccc att Thr Trp His Val Ser Cys Phe Val Cys Val Ala Cys Gly Lys Pro Ile	500	505	510 1656
cgg aac aat gtt ttt cac ttg gag gat ggt gaa ccc tac tgt gag act Arg Asn Asn Val Phe His Leu Glu Asp Gly Glu Pro Tyr Cys Glu Thr	515	520	525 1704
gat tat tat gcc ctc ttt ggt act ata tgc cat gga tgt gaa ttt ccc Asp Tyr Tyr Ala Leu Phe Gly Thr Ile Cys His Gly Cys Glu Phe Pro	530	535	540 1752
ata gaa gct ggt gac atg ttc ctg gaa gct ctg ggc tac acc tgg cat Ile Glu Ala Gly Asp Met Phe Leu Glu Ala Leu Gly Tyr Thr Trp His	545	550	555 560 1800
gac act tgc ttt gta tgc tca gtg tgt tgt gaa agt ttg gaa ggt cag Asp Thr Cys Phe Val Cys Ser Val Cys Cys Glu Ser Leu Glu Gly Gln	565	570	575 1848

acc ttt ttc tcc aag aag gac aag ccc ctg tgt aag aaa cat gct cat 1896
 Thr Phe Phe Ser Lys Lys Asp Lys Pro Leu Cys Lys Lys His Ala His
 580 585 590

tct gtg aat ttt tga aagtcaacag ttcaggagaa gagaaggaat ttgaagagaa 1951
 Ser Val Asn Phe *
 595

aaaggaaaat taaaattact aattaatttt tagattcaat atttatatgg agttttgaaa 2011
 aataatagtg gccctgaagg aataaattcc agcittaaaa accaagtcig aggaaatatt 2071
 iggcittcata aaglaaagag acggtttggc atttattatt actttttcct gtattttatg 2131
 cccataaaa aagcittata aaaaccaatt tccigtatga ctattaaatt catcttagaa 2191
 taaattagtg aagaatttaa ttttagaata aataatccaa tctgaaataa ttataccttc 2251
 tttccctgtt aggtagtatt gagtaaatct gcaaaaggca atgaaatgc cttaaatttt 2311
 atraataaca gaattattgt atttaaaaaa aaactaatat ttatctttaa aatagtaaat 2371
 aggattttaa acagagaall ttatcagtaa taggtgtcag tttttaaaaa attgcttgta 2431
 ggctgagcgc ggtggctcac gccgtgaatc ccagcacttt gggaggccaa ggtgggtgga 2491
 ccacatgagg tcaggagitt gagatcagcc tggccaacat ggtgaaaccc catctctact 2551
 aaaaatacaa aaattagccg gacgcagtgg cacgcgcctg taatcccagc tactcaagag 2611
 gctgaggcac gagaatcact tgaacccggg agggagaggt tgcagtgagc caagatcgta 2671
 ccactgcact ccagcctggg tgcacagagt agactctgtc tccaaaaaaa aactttgctt 2731
 gtataattatt ttgcccitac agtggatcat tctagtagga aaggacaata agatttttta 2791
 tcaaatgtg tcatgccagt aagagatgtt atattctitt ctcatttctt cccaccccaa 2851
 aaataagcta ccatatagct tataagtctc aaatttttgc cttttactaa aatgtgattg 2911
 tttctgttca ttgtgtatgc ttatcacct atattaggca aattccattt tttcccttgg 2971
 tgctaaggta aagattlaat laaataattt tggcctctca tagttttctc tctctttaaa 3031
 gagaataaat agagggccag gtgtgtgtgc tcacgcctgt gatcccagca ctttgggagg 3091
 ccaagacggg cggatcatga ggtcaagaga tcaagatcat cctggccaac atgggtgaaac 3151
 cctgtctcta ctaaaaatac aaaaatgagc tgggcatggt ggggcgtgccc tgtagtccca 3211
 tglacttggg aggcitgaggc aggaaaaatc ttgaaccag gagacggaag ttgcagttag 3271
 ctgagatcac accactgcac tccagcctgg tgacagagca agactccggc tcttaaaaaa 3331
 aaaaaaaaaa aaaaaaaaaa agagagagag aataaatalaga aaagaatgtg gctgggaatt 3391
 gtgaatcaga agattatacc ccccaattgt ttttcaatcc ccttttctca aataataaat 3451
 lagttaaatc agtttctgag ttatgccact ggctgatgaa gagttagagag gtctctttgc 3511
 agaatgatct tttgttttgc ttigttttct tcttctgcat ttaaaaaatta aaagatttgt 3571
 ttgaggatgt gatgaatttg agactttttg tggttttctc tcaataataa gtgaaccaat 3631
 ttcaaatgtg atcacaagct ttggaagact ttatctcaca gagggttgggt agtgttggga 3691
 ggggagttta attactcaga ttggcctgtt alttgatctc ctcccttggg aaaagaatta 3751
 ttagatatac acatggagac agggaaacaa ttgtggtaaa actgtggatc ctgttcttat 3811
 ttgcccagtg agaaaacaga ttcctggtatt tgatttgggt tttctctttg ttccagaat 3871
 ggaatgaagt ccatgaacct cctlaagttat aatttaaatt tgtttggggc aagggtattt 3931
 tatagtcgag acagagccct aggtccttcc tgcctcatca ctactttacg acatcacttc 3991
 caltgtgtgc atgtttgtta tagaggaggt tttaggctac aatatttgtt taacctccct 4051
 aagaactttc aaggcatctg tccgtaaaagc tgttaattta tggcttagca gatttatatt 4111
 atalgcagat aataattaac tggggataaa agaattggcaa ggggtgacac aaagtagcaa 4171
 actgaatact tctccaatag caaccccaag ctacctctc accctgcac ttggaggagg 4231
 gcaggaatt tcttttgaat aaagtgtctg gagctgaatt ctgcattatt talcgttgt 4291
 gctgaacaca cctataaaag acttgcctgc taatgtgcat tgtcatataa tttacactgt 4351
 cacatcttta cagtcttgta tttatagaa tacaataataa gttgatgtgt ttgtttgggt 4411
 tgagcttttt gttttgtttag ttttgccttc ataggttata tgccaagata gtatttgata 4471
 agtcaatgac atttggatgt ttctttcaaa gaattttatt tgaccagat ttcttataaa 4531
 gtlactctac attaaggatg tcaltttcat cagaccttct tttacatat tattcatgaa 4591
 gcataatgtt gcatttctcc aaattttatg cctgaaaggg tagtgttctc tccataaggta 4651
 tcatgttgtc ttgtgtcttt gtccatctct tccgtggcga agctttatai ctgttcttaa 4711
 aacagttaat cctgtgaaat aaataattgaa cataatccag aagaatctct ctatttccct 4771
 tggggaaatgc catatttaac tcaccagcag taatccttta ataactggca gagcacttta 4831
 ttcttctggt gagctccctg aatatttatt ttcttgatta taaattttct atattagtag 4891
 cattttttaa ttattacttc ttactatag agcatttact tttagtctct agatgtatat 4951
 ttlggaatgc tgtacttggc ataacataga tlaaaatcat aatgatgac taaaaactcc 5011
 ttggatttat ttccattttt aaaaatttta gcggtgaagt cagatttata atcttctct 5071
 agacttccat ggtctgaatg ttgcctgctg aagtagcaac ctaaaaagta tccccgtctt 5131
 atgcttctcc agttggccct ccatgtccat aggccttcgca tctgtgatcc agcccactgt 5191
 ggggtcaaaa latltgggga aaaaattgga ggttgcgcct ttgctgaaca tgtacaaact 5251
 tgtttttgtc attaaacaat atagtataac aactatttca aaagcattta cattgtatta 5311
 gctattalag gtaacttaga gatgatttaa agtgtatggt aggatgtgca caggttatat 5371
 gcaaatacta caccattttc tataagggac tlgaaatca tggacttag tatcctaggg 5431
 ggttcttga acccatcacc cataggggca ccataggaca actatagtac cgtgtttatt 5491
 tctattaat ttcaggttccg tttagagctc aaaaactaaaa cctaactatt tagtcacagt 5551
 gtaaaaacaa atggaaataa cagctcaaat cticaaaata ttaccatagc attatgttta 5611
 aataaatcta caacaaaaat gtaccatttt caagcagtag tacatttagga gcccttttat 5671
 agaaaaaat tttctcttta cccccgttcc agtgtgaatc tagtattctg ttaacatttg 5731
 tgtggcattt ggagtttctc attccccatt aaggagagac cttctcagac atgaagcaag 5791
 ggaaacatac tgaatagtgt tacacaaait tgatctggct tccatttctc cccctcattt 5851
 cccaaatgtt taaatgtatt ggatttgggt tctcaatgta taagtgtcct tatctgttaa 5911
 tgtctatctt ctgtctcttt aattttgtat atctgctgtt ttgcttttgg atacattttc 5971
 taattagaag tcaatgata aataataatc gtatagtaat aataccataa tgtgcacata 6031

ctcaataaat aaatgacigc attgttgtaa aaaaaaaaaa aaaaaaaaa

6079

<210> 28

<211> 596

<212> PRT

<213> 人

<400> 28

```

Met Ser Asn Tyr Ser Val Ser Leu Val Gly Pro Ala Pro Trp Gly Phe
 1      5      10      15
Arg Leu Gln Gly Gly Lys Asp Phe Asn Met Pro Leu Thr Ile Ser Ser
      20      25      30
Leu Lys Asp Gly Gly Lys Ala Ala Gln Ala Asn Val Arg Ile Gly Asp
      35      40      45
Val Val Leu Ser Ile Asp Gly Ile Asn Ala Gln Gly Met Thr His Leu
      50      55      60
Glu Ala Gln Asn Lys Ile Lys Gly Cys Thr Gly Ser Leu Asn Met Thr
65      70      75      80
Leu Gln Arg Ala Ser Ala Ala Pro Lys Pro Glu Pro Val Pro Val Gln
      85      90      95
Lys Gly Glu Pro Lys Glu Val Val Lys Pro Val Pro Ile Thr Ser Pro
      100      105      110
Ala Val Ser Lys Val Thr Ser Thr Asn Asn Met Ala Tyr Asn Lys Ala
      115      120      125
Pro Arg Pro Phe Gly Ser Val Ser Ser Pro Lys Val Thr Ser Ile Pro
      130      135      140
Ser Pro Ser Ser Ala Phe Thr Pro Ala His Ala Thr Thr Ser Ser His
145      150      155      160
Ala Ser Pro Ser Pro Val Ala Ala Val Thr Pro Pro Leu Phe Ala Ala
      165      170      175
Ser Gly Leu His Ala Asn Ala Asn Leu Ser Ala Asp Gln Ser Pro Ser
      180      185      190
Ala Leu Ser Ala Gly Lys Thr Ala Val Asn Val Pro Arg Gln Pro Thr
      195      200      205
Val Thr Ser Val Cys Ser Glu Thr Ser Gln Glu Leu Ala Glu Gly Gln
      210      215      220
Arg Arg Gly Ser Gln Gly Asp Ser Lys Gln Gln Asn Gly Pro Pro Arg
225      230      235      240
Lys His Ile Val Glu Arg Tyr Thr Glu Phe Tyr His Val Pro Thr His
      245      250      255
Ser Asp Ala Ser Lys Lys Arg Leu Ile Glu Asp Thr Glu Asp Trp Arg
      260      265      270
Pro Arg Thr Gly Thr Thr Gln Ser Arg Ser Phe Arg Ile Leu Ala Gln
      275      280      285
Ile Thr Gly Thr Glu His Leu Lys Glu Ser Glu Ala Asp Asn Thr Lys
      290      295      300
Lys Ala Asn Asn Ser Gln Glu Pro Ser Pro Gln Leu Ala Ser Ser Val
305      310      315      320
Ala Ser Thr Arg Ser Met Pro Glu Ser Leu Asp Ser Pro Thr Ser Gly
      325      330      335
Arg Pro Gly Val Thr Ser Leu Thr Thr Ala Ala Ala Phe Lys Pro Val
      340      345      350
Gly Ser Thr Gly Val Ile Lys Ser Pro Ser Trp Gln Arg Pro Asn Gln
      355      360      365
Gly Val Pro Ser Thr Gly Arg Ile Ser Asn Ser Ala Thr Tyr Ser Gly
      370      375      380
Ser Val Ala Pro Ala Asn Ser Ala Leu Gly Gln Thr Gln Pro Ser Asp
385      390      395      400
Gln Asp Thr Leu Val Gln Arg Ala Glu His Ile Pro Ala Gly Lys Arg
      405      410      415
Thr Pro Met Cys Ala His Cys Asn Gln Val Ile Arg Gly Pro Phe Leu
      420      425      430
Val Ala Leu Gly Lys Ser Trp His Pro Glu Glu Phe Asn Cys Ala His
      435      440      445
Cys Lys Asn Thr Met Ala Tyr Ile Gly Phe Val Glu Glu Lys Gly Ala
      450      455      460
Leu Tyr Cys Glu Leu Cys Tyr Glu Lys Phe Phe Ala Pro Glu Cys Gly
465      470      475      480
Arg Cys Gln Arg Lys Ile Leu Gly Glu Val Ile Asn Ala Leu Lys Gln
      485      490      495
Thr Trp His Val Ser Cys Phe Val Cys Val Ala Cys Gly Lys Pro Ile
      500      505      510
Arg Asn Asn Val Phe His Leu Glu Asp Gly Glu Pro Tyr Cys Glu Thr
      515      520      525
Asp Tyr Tyr Ala Leu Phe Gly Thr Ile Cys His Gly Cys Glu Phe Pro
530      535      540
Ile Glu Ala Gly Asp Met Phe Leu Glu Ala Leu Gly Tyr Thr Trp His

```

```

545      550      555      560
Asp Thr Cys Phe Val Cys Ser Val Cys Cys Glu Ser Leu Glu Gly Gln
      565      570      575
Thr Phe Phe Ser Lys Lys Asp Lys Pro Leu Cys Lys Lys His Ala His
      580      585      590
Ser Val Asn Phe
      595

<210> 29
<211> 4996
<212> DNA
<213> 人

<220>
<221> CDS
<222> (394)... (828)

<400> 29
attctcaatc ctacccatga tgcagacaga ttacctcatt tcagctcaat aaaacatcac 60
caaattccct taggcagaaa aagggtgggg agtgcccttt ctigaaggaa tcccttaagt 120
gcacgtctcc cgttttcttt tcacatcaat gcagagtccc tctgtgaatt tacaacaccc 180
aacacccaag agttcagtc ttaagaagga aaattgagca agtatitaa tggttggaca 240
ttgcataaaa gaggtccata tccagcctaa ctctagaaat ttattctca aacttagttg 300
gcctcttttt ctctgtcag aggaccattc ttagtggcac tggggaaatc ttggcaccac 360
gaagaattca actgcgtca ctgcaaaaat aca atg gcc tac att gga ttt gta 414
                                   Met Ala Tyr Ile Gly Phe Val
                                   1                               5

gag gag aaa gga gcc ctg tat tgt gag ctg tgc tat gag aaa ttc ttt 462
Glu Glu Lys Gly Ala Leu Tyr Cys Glu Leu Cys Tyr Glu Lys Phe Phe
      10                               15                               20

gcc cct gaa tgt ggt cga tgc caa agg aag atc ctt gga gaa gtc atc 510
Ala Pro Glu Cys Gly Arg Cys Gln Arg Lys Ile Leu Gly Glu Val Ile
      25                               30                               35

aat gcg ttg aaa caa act tgg cat gtt tcc tgt ttt gtg tgt gta gcc 558
Asn Ala Leu Lys Gln Thr Trp His Val Ser Cys Phe Val Cys Val Ala
      40                               45                               50                               55

tgt gga aag ccc att cgg aac aat gtt ttt cac ttg gag gat ggt gaa 606
Cys Gly Lys Pro Ile Arg Asn Asn Val Phe His Leu Glu Asp Gly Glu
      60                               65                               70

ccc tac tgt gag act gat tat tat gcc ctg ttt ggt act ata tgc cat 654
Pro Tyr Cys Glu Thr Asp Tyr Tyr Ala Leu Phe Gly Thr Ile Cys His
      75                               80                               85

gga tgt gaa ttt ccc ata gaa gct ggt gac atg ttc ctg gaa gct ctg 702
Gly Cys Glu Phe Pro Ile Glu Ala Gly Asp Met Phe Leu Glu Ala Leu
      90                               95                               100

ggc tac acc tgg cat gac act tgc ttt gta tgc tca gtg tgt tgt gaa 750
Gly Tyr Thr Trp His Asp Thr Cys Phe Val Cys Ser Val Cys Cys Glu
      105                               110                               115

agt ttg gaa ggt cag acc ttt ttc tcc aag aag gac aag ccc ctg tgt 798
Ser Leu Glu Gly Gln Thr Phe Phe Ser Lys Lys Asp Lys Pro Leu Cys
      120                               125                               130                               135

aag aaa cat gct cat tct gtg aat ttt tga aagtcaacag ttcaggagaa 848
Lys Lys His Ala His Ser Val Asn Phe *
      140

gagaaggaat tgaagagaa aaaggaat taaaattact aattaatttt tagattcaat 908
atttatatgg agttttgaaa aataatagtg gccctgaagg aataaattcc agctttaaaa 968
accaagtcgt aggaatatt tggcttcata aagtaaagag acggtttggc atttattatt 1028
acttttccct gtattttatg ccataaaaat aagctttata aaaaccaatt tcctgatgga 1088
ctattaaatt catcttagaa taaattagtg aagaatttaa ttttagaata aataatccaa 1148
tctgaaataa ttataacctc tticcttgtt aggtagttat gagtaaatct gcaaaaggca 1208
atgaaaatgc cttaaatttt atcaataaca gaattattgt atttaaaaaa aaactaatac 1268
ttatctttaa aatagtaaat aggatittaa acagagaatt ttatcagtaa taggtgtcag 1328
tttttaaaaa attgcttgta ggctgagcgc ggtggctcac gcctgtaate ccagcacitt 1388
gggaggccaa ggtgggtgga ccacatgagg tcaggagttt gagatcagcc tggccaacat 1448
ggtgaaaccc catcttact aaaaatacaa aaattagccg gacgcagtgg cacgcgctgt 1508
taatccagc tactcaagag gctgaggcac gagaatcac tgaacccggg agggagaggt 1568

```

```

tgcagtgcgc caagatcgta ccactgcact ccagcctggg tgacagagtg agactctgtc 1628
tccaaaaaaa aactttgctt gtattattat ttgaccttac agtggatcat tctagtagga 1688
aaggacaata agatttlltta tcaaaaatgtg tcatgccagt aagagatgtt atattctttt 1748
ctcatttctt cccaccacca aaataagcta ccatatagct tataagtctc aaatttttgc 1808
cttttactaa aatgtgattg tttctgttca ttgtgtatgc ttcacacact atattaggca 1868
aatccatttt ttttcccttg tgcctaaggta aagatttaaa taaataattt tggcctctca 1928
tagttttctc tctcttllaaa gagaataaat agagggccag gtgtgggtggc tcacgcctgt 1988
gatcccgaca ctttgggagg ccaagacggg cggatcatga ggtcaagaga tcaagatcat 2048
cctggccaac atgggtgaaa cctgtctcta ctaaaaatac aaaaatgagc tgggcatggt 2108
ggggcgtgcc tgtagtccca tctacttggg aggcctgaggc aggaaaattc ttgaaccag 2168
gagacggagag ttgcagttag ctgagatcac accactgcac tccagcctgg tgacagagca 2228
agactccggc tcttaaaaaa aaaaaaaaaa aaaaaaaaaa agagagagag aataaataga 2288
aaagaatgtg gctgggaatt gtgaatcaga agattatacc ccccaattgt tttcaatcc 2348
ccttttctca aataataaat tagttaaatc agtttctgag ttatgccact ggctgatgaa 2408
gagttgagag gtctctttgc agaatagtct tttgtttcgc ttttgtttct tcttctgcat 2468
tlaaaaatta aaagatttgt ttgaggatgt gatgaaattg agactttttg tggttttctc 2528
tcaataataa gtgaaccaat ttcaaatgtg atcacaaagt ttggaaagct tttattcaca 2588
gaggttgggt agtgttggga ggggagttaa tttactcaga ttggcctgtt atttgatttc 2648
ctcctttggg aaaaagaatta tglagatacc acatggagac agggaaacaa ttgttgtaaa 2708
actgtggatc ctgttgcctt ttgcccagtg agaaaacaga ttctgggtat tgatttgggt 2768
tttctctttg tttccagaat ggatgaaagt ccatgaacct cctaagttat aatttaaat 2828
tgtttggggc aaggttggca ttagtctgag acagagccct aggtccttcc tgcccatca 2888
ctcacctacg acatcacttc cattgtgtgc atgtttgtta tagaggaggt ttaggctac 2948
aatatttgtt taacctccct aagaacttcc aaggcatctg tccctgaaagc tgttaattta 3008
tggtctagca gatttatatt atatgcagat aataattaac tggggataaa agaattggca 3068
ggggtagcac aaagttagca actgaatact tctccaatag caaccccaag ctacctctc 3128
accctgcac tgggaggag gcaggaaatt tcttttgaag taaagtgtg gagctgaatt 3188
ctgcattatt tatcgttgc tctgaaacca cctataaaag acttgcctggc taatgtgcat 3248
tgcatataaa tgtacactgt cacatcttta cagtcttgta tgttatagaa taaaaataa 3308
gttgatgggt ttgtttgggt ttagcttttt gtltgtttag ttttgccttc atagggtata 3368
tgccaagata gtatttgata agtcaatgac atttgatgt tttcttcaa gaattttatt 3428
tgacccagat ttcttataaa gtattcttac attaaggatg tcatittcat cagaccttct 3488
ttctacatat tattcaigaa gcataatgtt gcatttctcc aaattttatg cctgaaaggg 3548
tagtgtgtct tccctaggta tcatgttgc ttgtgtcttt gtccatctct tccgtggcga 3608
agctttatat ctgttcttaa aacagttaat cctgtgaaat aaatatgaa cataatccag 3668
aagaatctct ctatttccct tggggaatgc cataittaat tcaccagcag taatccttta 3728
ataactggca gagcacitaa tttctctggt gagctccctg aatatattat tttctgatta 3788
taaaatttct atattagtag cattttttaa ttattacttc ttcactatag agcatttact 3848
tttagtctct agatgtatat ttlggaatgc tgtacttggc ataacataga ttaaatcat 3908
aatgcagtac taaaaactcc ttggatttat ttcccatttt aaaaatttta gcggtaagtt 3968
cagatttata atctttctct agacttccat ggcttgaatg ttgctgtctg aagtagcaac 4028
ctaaaaagta tcccctgctt atgcttctcc agttggccct ccatgtccat aggcttcgca 4088
tctgtgattc agcccactgt gggtcaaaaa tatttgggga aaaaatggat ggttgcgcct 4148
ttgctgaaca tgtacaaact tgtttttgtc attaaacaat atagtataac aactatttac 4208
aaagcatttta cattgtatta gctattatag gtaacttaga gatgatttaa agtgtatggt 4268
aggatgtgca caggttatat gcaaaactaa caccatttcc tataaggagc ttgaacatca 4328
tggaatttag tatcctaggg gtgtcttggg acccatcacc cataggggca ccataggaca 4388
actatagtag cgtgtttatt tcttattaat tcagggtccg tttagagtct aaaactaaaa 4448
cctaatacatt tagtcacagt gtaaaaacaa atggaaataa cagctcaaat cttcaaaata 4508
ttaccatagc attatgttta aaataatcta caacaaaaat gtaccatttt caagcagtag 4568
tacattagga gcccttttat agaaaataat ttcttcttta ccccggttcc agtgtgaatc 4628
tagtattctg ttaacatttg tgtggcatit ggagtttgtc atccccattg aaggagagc 4688
cttctcagac atgaagcaag gaaacatac tgaatagttt tacacaaatt tgaatcggct 4748
tccatttgtc cccctcattt cccaaatgtt taaatgtatt ggatttgggt tctcaatgta 4808
taagttgcct tatctgttaa tgtctatctt ctgtctcttt aattttgtat atctgtgtt 4868
ttgcttttgg atacatttcc taatlagaa tcacatgata aatataatca gtatagtaat 4928
aataccataa tgtgcacata ctcaataaat aatgactgac attgttgtta aaaaaaaaaa 4988
aaaaaaaaa 4996

```

<210> 30

<211> 144

<212> PRT

<213> 人

<400> 30

```

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

```

```

Asp Met Phe Leu Glu Ala Leu Gly Tyr Thr Trp His Asp Thr Cys Phe
100 105 110
Val Cys Ser Val Cys Cys Glu Ser Leu Glu Gly Gln Thr Phe Phe Ser
115 120 125
Lys Lys Asp Lys Pro Leu Cys Lys Lys His Ala His Ser Val Asn Phe
130 135 140

<210> 31
<211> 1875
<212> DNA
<213> 人

<220>
<221> CDS
<222> (207)... (821)

<400> 31
agctgcggcg gccgcaggtt ccaaagcggg tccgagccgc cgccgcgcgc gcgcccgcga 60
ctgcagcccc agggccccgc cccccacca cgtctgcgtt gctgccccgc ctgggccagg 120
ccccaaaggc aaggacaaag cagctgtcag ggaaccctcg ccggagtcca atttacgtgc 180
agctgccggc aaccacaggt tccaag atg gtt tgc ggg ggc ttc gcg tgt tcc 233
Met Val Cys Gly Gly Phe Ala Cys Ser
1 5

aag aac tgc ctg tgc gcc ctc aac ctg ctt tac acc ttg gtt agt ctg 281
Lys Asn Cys Leu Cys Ala Leu Asn Leu Leu Tyr Thr Leu Val Ser Leu
10 15 20 25

ctg cta att gga att gct gcg tgg ggc att ggc ttc ggg ctg att tcc 329
Leu Leu Ile Gly Ile Ala Ala Trp Gly Ile Gly Phe Gly Leu Ile Ser
30 35 40

agt ctc cga gtg gtc gcc gtg gtc att gca gtg ggc atc ttc ttg ttc 377
Ser Leu Arg Val Val Gly Val Val Ile Ala Val Gly Ile Phe Leu Phe
45 50 55

ctg att gct tta gtg ggt ctg att gga gct gta aaa cat cat cag gtg 425
Leu Ile Ala Leu Val Gly Leu Ile Gly Ala Val Lys His His Gln Val
60 65 70

ttg cta ttt ttt tat atg att att ctg tta ctt gta ttt att gtt cag 473
Leu Leu Phe Phe Tyr Met Ile Ile Leu Leu Leu Val Phe Ile Val Gln
75 80 85

ttt tct gta tct tgc gct tgt tta gcc ctg aac cag gag caa cag ggt 521
Phe Ser Val Ser Cys Ala Cys Leu Ala Leu Asn Gln Glu Gln Gln Gly
90 95 100 105

cag ctt ctg gag gtt ggt tgg aac aat acg gca agt gct cga aat gac 569
Gln Leu Leu Glu Val Gly Trp Asn Asn Thr Ala Ser Ala Arg Asn Asp
110 115 120

atc cag aga aat cta aac tgc tgt ggg ttc cga agt gtt aac cca aat 617
Ile Gln Arg Asn Leu Asn Cys Cys Gly Phe Arg Ser Val Asn Pro Asn
125 130 135

gac acc tgt ctg gct agc tgt gtt aaa agt gac cac tcg tgc tcg cca 665
Asp Thr Cys Leu Ala Ser Cys Val Lys Ser Asp His Ser Cys Ser Pro
140 145 150

tgt gct cca atc ata gga gaa tat gct gga gag gtt ttg aga ttt gtt 713
Cys Ala Pro Ile Ile Gly Glu Tyr Ala Gly Glu Val Leu Arg Phe Val
155 160 165

ggg ggc att ggc ctg ttc ttc agt ttt aca gag atc ctg ggt gtt tgg 761
Gly Gly Ile Gly Leu Phe Phe Ser Phe Thr Glu Ile Leu Gly Val Trp
170 175 180 185

ctg acc tac aga tac agg aac cag aaa gac ccc cgc gcg aat cct agt 809
Leu Thr Tyr Arg Tyr Arg Asn Gln Lys Asp Pro Arg Ala Asn Pro Ser
190 195 200

gca ttc ctt tga tgagaaaaca aggaagattt cctttcgtat tatgatcttg 861
Ala Phe Leu *

```

```

itcactttct gtaattttct gttaagctcc atttgccagt ttaaggaagg aaacactatc 921
tggaaaagta ccttattgat agtggaaatta tatattttta ctctatgttt ctctacatgt 981
ttttttcttt cegtgtctga aaaatatttg aaacttgttg tctctgaagc tcggtggcac 1041
ciggaattta ctgtattcat tgtcgggcac tgtccactgt ggcctttctt agcattttta 1101
cctgcagaaa aactttgtat ggtaccactg tgttggttat atgggtgaatc tgaacgtaca 1161
tctcactggg ataattatat gtacactgti gctgtgtaga tagttcctac tggaaaaaga 1221
gtggaaattt attaaaatca gaaagtatga gatcctgtta tgttaaggga aatccaaatt 1281
cccaattttt ttgtgtcttt tiaggaaaga tgtgttgttg taaaaagtgt tagtataaaa 1341
atgataattt acttgtagtc ttttatgatt acaccaatgt attctagaaa tagttatgtc 1401
ttaggaaatt gtggtttaat ttttgacttt tacaggtaag tgcaggagag aagtgggttc 1461
atgaaatgtt ctaatgtata ataacattta ccttcagcct ccatcagaat ggaacgagtt 1521
ttgagtaatc aggaagtata tctatatgat cttgatattg ttttataata atttgaagtc 1581
taaaagactg catttttaaa caagtttagta tlaatgcgtt ggcccacgta gcaaaaagat 1641
atttgattat cttaaaaaat gttaaatacc gttttcatga aagttctcag tattgtaaca 1701
gcaacttgtc aaacctaacg atatttgaat atgatctccc ataatttgaa attgaaatcg 1761
tattgtgtgg ctctgtatat tctgttaaaa aattaaagga cagaaacctt tctttgtgta 1821
tgcattgttg aattaaaaga aagtaatgga agaattgatc gatgaaaaaa aaaa 1875

```

<210> 32
 <211> 204
 <212> PRT
 <213> 人

```

<400> 32
Met Val Cys Gly Gly Phe Ala Cys Ser Lys Asn Cys Leu Cys Ala Leu
1      5      10      15
Asn Leu Leu Tyr Thr Leu Val Ser Leu Leu Leu Ile Gly Ile Ala Ala
20     25     30
Trp Gly Ile Gly Phe Gly Leu Ile Ser Ser Leu Arg Val Val Gly Val
35     40     45
Val Ile Ala Val Gly Ile Phe Leu Phe Leu Ile Ala Leu Val Gly Leu
50     55     60
Ile Gly Ala Val Lys His His Gln Val Leu Leu Phe Phe Tyr Met Ile
65     70     75     80
Ile Leu Leu Leu Val Phe Ile Val Gln Phe Ser Val Ser Cys Ala Cys
85     90     95
Leu Ala Leu Asn Gln Glu Gln Gln Gly Gln Leu Leu Glu Val Gly Trp
100    105    110
Asn Asn Thr Ala Ser Ala Arg Asn Asp Ile Gln Arg Asn Leu Asn Cys
115    120    125
Cys Gly Phe Arg Ser Val Asn Pro Asn Asp Thr Cys Leu Ala Ser Cys
130    135    140
Val Lys Ser Asp His Ser Cys Ser Pro Cys Ala Pro Ile Ile Gly Glu
145    150    155    160
Tyr Ala Gly Glu Val Leu Arg Phe Val Gly Gly Ile Gly Leu Phe Phe
165    170    175
Ser Phe Thr Glu Ile Leu Gly Val Trp Leu Thr Tyr Arg Tyr Arg Asn
180    185    190
Gln Lys Asp Pro Arg Ala Asn Pro Ser Ala Phe Leu
195    200

```

<210> 33
 <211> 551
 <212> DNA
 <213> 人

<220>
 <221> CDS
 <222> (87)... (380)

```

<400> 33
acccatccg ctggctctca cccctcggag acgctcgccc gacagcatag tacttgccgc 60
ccagccacgc ccgcgcgcca gccacc atg cta ggt aac aag cga ctg ggg ctg 113
Met Leu Gly Asn Lys Arg Leu Gly Leu
1      5

tcc gga ctg acc ctc gcc ctg tcc ctg ctc gtc tgc ctg ggt gcg ctg 161
Ser Gly Leu Thr Leu Ala Leu Ser Leu Leu Val Cys Leu Gly Ala Leu
10     15     20     25

gcc gag gcg tac ccc tcc aag ccg gac aac ccg ggc gag gac gca cca 209
Ala Glu Ala Tyr Pro Ser Lys Pro Asp Asn Pro Gly Glu Asp Ala Pro
30     35     40

gcg gag gac atg gcc aga tac tac tcg gcg ctg cga cac tac atc aac 257
Ala Glu Asp Met Ala Arg Tyr Tyr Ser Ala Leu Arg His Tyr Ile Asn

```

```

      45              50              55
ctc atc acc agg cag aga tat gga aaa cga tcc agc cca gag aca ctg 305
Leu Ile Thr Arg Gln Arg Tyr Gly Lys Arg Ser Ser Pro Glu Thr Leu
      60              65              70

att tca gac ctc ttg atg aga gaa agc aca gaa aat gtt ccc aga act 353
Ile Ser Asp Leu Leu Met Arg Glu Ser Thr Glu Asn Val Pro Arg Thr
      75              80              85

cgg ctt gaa gac cct gca atg tgg tga tgggaaatga gacttgctct 400
Arg Leu Glu Asp Pro Ala Met Trp *
      90              95

ctggcctttt cctattttca gcccatattt catcgtgtaa aacgagaatc caccatcct 460
accaatgcat gcagccactg tgctgaattc tgcaatgttt tccittgtca tcattgtata 520
tatgtgtgtt taaataaagt atcatgcatt c 551

<210> 34
<211> 97
<212> PRT
<213> 人

<400> 34
Met Leu Gly Asn Lys Arg Leu Gly Leu Ser Gly Leu Thr Leu Ala Leu
 1              5              10              15
Ser Leu Leu Val Cys Leu Gly Ala Leu Ala Glu Ala Tyr Pro Ser Lys
      20              25              30
Pro Asp Asn Pro Gly Glu Asp Ala Pro Ala Glu Asp Met Ala Arg Tyr
      35              40              45
Tyr Ser Ala Leu Arg His Tyr Ile Asn Leu Ile Thr Arg Gln Arg Tyr
      50              55              60
Gly Lys Arg Ser Ser Pro Glu Thr Leu Ile Ser Asp Leu Leu Met Arg
      65              70              75              80
Glu Ser Thr Glu Asn Val Pro Arg Thr Arg Leu Glu Asp Pro Ala Met
      85              90              95
Trp

<210> 35
<211> 997
<212> DNA
<213> 人

<220>
<221> CDS
<222> (273)... (707)

<400> 35
gaaggaaaaa gagcaacaga tccaggaggc attcacctgc cctgtctcca aacagccttg 60
tgccacacct accccaacc tccagagggg agcagctatt taaggggagc aggagtgcag 120
aacaacaag acggcctggg gatacaactc tggagtcctc tgagagagcc accaaggagg 180
agcaggggag cgacggccgg ggcagaagtt gagaccaccc agcagaggag ctaggccagt 240
ccatctgcat ttgtaccca agaactctta cc atg aag acc ctc cta ctg ttg 293
                                Met Lys Thr Leu Leu Leu Leu
                                1              5

gca gtg atc atg atc ttt ggc cta ctg cag gcc cat ggg aat ttg gtg 341
Ala Val Ile Met Ile Phe Gly Leu Leu Gln Ala His Gly Asn Leu Val
      10              15              20

aat ttc cac aga atg atc aag ttg acg aca gga aag gaa gcc gca ctc 389
Asn Phe His Arg Met Ile Lys Leu Thr Thr Gly Lys Glu Ala Ala Leu
      25              30              35

agt tat ggc ttc tac ggc tgc cac tgt ggc gtg ggt ggc aga gga tcc 437
Ser Tyr Gly Phe Tyr Gln Cys His Cys Gly Val Gly Gly Arg Gly Ser
      40              45              50              55

ccc aag gat gca acg gat cgc tgc tgt gtc act cat gac tgt tgc tac 485
Pro Lys Asp Ala Thr Asp Arg Cys Cys Val Thr His Asp Cys Cys Tyr
      60              65              70

aaa cgt ctg gag aaa cgt gga tgt ggc acc aaa ttt ctg agc tac aag 533
Lys Arg Leu Glu Lys Arg Gly Cys Gly Thr Lys Phe Leu Ser Tyr Lys
      75              80              85

```


ttt agc aac tcg ggg agc aga atc acc tgt gca aaa cag gac tcc tgc 581
Phe Ser Asn Ser Gly Ser Arg Ile Thr Cys Ala Lys Gln Asp Ser Cys
90 95 100

aga agt caa ctg tgt gag tgt gat aag gct gct gcc acc tgt ttt gct 629
Arg Ser Gln Leu Cys Glu Cys Asp Lys Ala Ala Ala Thr Cys Phe Ala
105 110 115

aga aac aag acg acc tac aat aaa aag tac cag tac tat tcc aat aaa 677
Arg Asn Lys Thr Thr Tyr Asn Lys Lys Tyr Gln Tyr Tyr Ser Asn Lys
120 125 130 135

cac tgc aga ggg agc acc cct cgt tgc tga gtccccctctt ccctggaaac 727
His Cys Arg Gly Ser Thr Pro Arg Cys *
140

cttcaccca gigtgaatt tccctctctc ataccctccc tccctaccct aaccaagttc 787
cttggccatg cagaaagcat ccctcaccca tcttagaggc caggcaggag ccctctata 847
cccaccaga atgagacatc cagcagattt ccagccttct actgctctcc tccacctcaa 907
ctccgtgctt aaccaaagaa gctgtactcc ggggggtctc ttcigaataa agcaattagc 967
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa 997

<210> 36
<211> 144
<212> PRT
<213> 人

<400> 36
Met Lys Thr Leu Leu Leu Leu Ala Val Ile Met Ile Phe Gly Leu Leu
1 5 10 15
Gln Ala His Gly Asn Leu Val Asn Phe His Arg Met Ile Lys Leu Thr
20 25 30
Thr Gly Lys Gly Ala Ala Leu Ser Tyr Gly Phe Tyr Gly Cys His Cys
35 40 45
Gly Val Gly Gly Arg Gly Ser Pro Lys Asp Ala Thr Asp Arg Cys Cys
50 55 60
Val Thr His Asp Cys Cys Tyr Lys Arg Leu Glu Lys Arg Gly Cys Gly
65 70 75 80
Thr Lys Phe Leu Ser Tyr Lys Phe Ser Asn Ser Gly Ser Arg Ile Thr
85 90 95
Cys Ala Lys Gln Asp Ser Cys Arg Ser Gln Leu Cys Glu Cys Asp Lys
100 105 110
Ala Ala Ala Thr Cys Phe Ala Arg Asn Lys Thr Thr Tyr Asn Lys Lys
115 120 125
Tyr Gln Tyr Tyr Ser Asn Lys His Cys Arg Gly Ser Thr Pro Arg Cys
130 135 140

<210> 37
<211> 1717
<212> DNA
<213> 人

<220>
<221> CDS
<222> (102)... (1202)

<400> 37
gggcgcagca cactccacgc cgccgcagc ctgacacgcc gcgcggcccc ccagtctccc 60
gcggctgctc cccagggcat ggcacagggc ctgcctcac t atg gca gca gca cgg 116
Met Ala Ala Ala Arg
1 5

cac agc acg ctg gac ttc atg ctg ggc gcc aaa gct gat ggt gag acc 164
His Ser Thr Leu Asp Phe Met Leu Gly Ala Lys Ala Asp Gly Glu Thr
10 15 20

att cta aaa ggc ctg cag tcc att ttc cag gag cag ggg atg gcg gag 212
Ile Leu Lys Gly Leu Gln Ser Ile Phe Gln Glu Gln Gly Met Ala Glu
25 30 35

tgc gtg cac acc tgg cag gac cat ggc tat tta gca acc tac aca aac 260
Ser Val His Thr Trp Gln Asp His Gly Tyr Leu Ala Thr Tyr Thr Asn
40 45 50

aag aac ggc agc ttt gcc aat ttg aga att tac cca cat gga ttg gtg 308

Lys	Asn	Gly	Ser	Phe	Ala	Asn	Leu	Arg	Ile	Tyr	Pro	His	Gly	Leu	Val	
55						60					65					
ttg	ctg	gac	ctt	cag	agt	tat	gat	ggt	gat	gcg	caa	ggc	aaa	gaa	gag	356
Leu	Leu	Asp	Leu	Gln	Ser	Tyr	Asp	Gly	Asp	Ala	Gln	Gly	Lys	Glu	Glu	
70					75					80					85	
atc	gac	agt	att	ttg	aac	aaa	gta	gag	gaa	aga	atg	aaa	gaa	ttg	agt	404
Ile	Asp	Ser	Ile	Leu	Asn	Lys	Val	Glu	Glu	Arg	Met	Lys	Glu	Leu	Ser	
				90					95					100		
cag	gac	agt	act	ggg	cgg	gtg	aaa	cga	tta	cca	ccc	ata	gtg	cga	gga	452
Gln	Asp	Ser	Thr	Gly	Arg	Val	Lys	Arg	Leu	Pro	Pro	Ile	Val	Arg	Gly	
			105					110					115			
gga	gcc	atc	gac	aga	tac	tgg	ccc	acc	gcc	gac	ggg	cgc	ctg	gtt	gaa	500
Gly	Ala	Ile	Asp	Arg	Tyr	Trp	Pro	Thr	Ala	Asp	Gly	Arg	Leu	Val	Glu	
		120					125					130				
tat	gac	ata	gat	gaa	gtg	gta	tat	gac	gaa	gat	tca	cct	tat	caa	aat	548
Tyr	Asp	Ile	Asp	Glu	Val	Val	Tyr	Asp	Glu	Asp	Ser	Pro	Tyr	Gln	Asn	
	135					140					145					
ata	aaa	att	cta	cac	tgc	aag	cag	ttt	gga	aat	att	ctc	atc	ctt	agt	596
Ile	Lys	Ile	Leu	His	Ser	Lys	Gln	Phe	Gly	Asn	Ile	Leu	Ile	Leu	Ser	
	150				155				160						165	
ggg	gat	gtt	aat	ttg	gca	gag	agt	gat	tig	gca	tat	acc	cgg	gcc	atc	644
Gly	Asp	Val	Asn	Leu	Ala	Glu	Ser	Asp	Leu	Ala	Tyr	Thr	Arg	Ala	Ile	
				170					175					180		
atg	ggc	agt	ggc	aaa	gaa	gat	tac	act	ggc	aaa	gat	gta	ctc	att	ctg	692
Met	Gly	Ser	Gly	Lys	Glu	Asp	Tyr	Thr	Gly	Lys	Asp	Val	Leu	Ile	Leu	
			185					190					195			
gga	ggt	gga	gac	gga	ggc	ata	tig	tgt	gaa	ata	gtc	aaa	cta	aaa	cca	740
Gly	Gly	Gly	Asp	Gly	Gly	Ile	Leu	Cys	Glu	Ile	Val	Lys	Leu	Lys	Pro	
		200					205					210				
aag	atg	gtc	act	atg	gta	gag	att	gac	caa	atg	gtg	att	gat	ggg	tgt	788
Lys	Met	Val	Thr	Met	Val	Glu	Ile	Asp	Gln	Met	Val	Ile	Asp	Gly	Cys	
	215					220					225					
aag	aaa	tac	atg	cga	aaa	acg	tgt	ggc	gat	gtc	tta	gac	aat	ctt	aaa	836
Lys	Lys	Tyr	Met	Arg	Lys	Thr	Cys	Gly	Asp	Val	Leu	Asp	Asn	Leu	Lys	
	230				235					240					245	
gga	gac	tgc	tat	cag	gtt	cta	ala	gaa	gac	tgt	atc	ccg	gta	ctg	aag	884
Gly	Asp	Cys	Tyr	Gln	Val	Leu	Ile	Glu	Asp	Cys	Ile	Pro	Val	Leu	Lys	
				250					255					260		
agg	tac	gcc	aaa	gaa	ggg	aga	gaa	ttt	gat	tat	gtg	att	aat	gat	tig	932
Arg	Tyr	Ala	Lys	Glu	Gly	Arg	Glu	Phe	Asp	Tyr	Val	Ile	Asn	Asp	Leu	
			265					270					275			
aca	gct	gtt	cca	atc	tcc	acg	tct	cca	gaa	gaa	gat	tcc	aca	tgg	gag	980
Thr	Ala	Val	Pro	Ile	Ser	Thr	Ser	Pro	Glu	Glu	Asp	Ser	Thr	Trp	Glu	
		280				285						290				
ttt	ctc	aga	ctg	att	ctt	gac	ctc	tca	atg	aaa	gtg	ttg	aaa	cag	gat	1028
Phe	Leu	Arg	Leu	Ile	Leu	Asp	Leu	Ser	Met	Lys	Val	Leu	Lys	Gln	Asp	
	295					300					305					
ggg	aaa	tat	ttt	aca	cag	ggg	aac	tgt	gtc	aat	ctg	aca	gaa	gca	ctg	1076
Gly	Lys	Tyr	Phe	Thr	Gln	Gly	Asn	Cys	Val	Asn	Leu	Thr	Glu	Ala	Leu	
	310				315					320				325		
tgc	cic	tat	gaa	gaa	cag	ctg	ggg	cgc	ctg	tat	tgt	cct	gtg	gaa	ttt	1124
Ser	Leu	Tyr	Glu	Glu	Gln	Leu	Gly	Arg	Leu	Tyr	Cys	Pro	Val	Glu	Phe	
				330				335						340		
tca	aag	gag	atc	gtc	tgt	gtc	cct	tca	tac	ttg	gaa	ttg	tgg	gta	ttt	1172
Ser	Lys	Glu	Ile	Val	Cys	Val	Pro	Ser	Tyr	Leu	Glu	Leu	Trp	Val	Phe	
			345					350					355			
tac	act	gtt	tgg	aag	aaa	gct	aaa	ccc	tga	agatcagtag	ccccctaatca					1222
Tyr	Thr	Val	Trp	Lys	Lys	Ala	Lys	Pro	*							

```

360      catgtgctgc aaatagcctt cctgacctcc atatgctgta catgacatca aaatgagtca 1282
      ggcaattgat tgtgaattcc ttaaagtittt ccttttttta ataattatatt ttaattttaa 1342
      aaagcaaatg gaaaatgtat attttgatga gcttaggggtg tttttttttt gaaagtcagc 1402
      tgaaggatgg ttagacagca cagcgaagac tgctaaatgc actgaccccc cccattagaa 1462
      tgtgattttt gttccttttt atttctctgt gggcttttgt ttttgttttt gttttggtag 1522
      atcttcaatt tggatatttg gaggagtga catcggttgtt ttgctggagg gaagatcttg 1582
      atgggtgttc ttccccaaa aattgactta gatattaaaa tttggtgctt ataagagaga 1642
      gttaaaaaaa aataggattg cttcaattaa aattacaaaa gagacaaaaa aaaaaaaaaa 1702
      aaaaaaaaaa aaaaaa 1717

```

<210> 38
 <211> 366
 <212> PRT
 <213> 人

```

<400> 38
Met Ala Ala Ala Arg His Ser Thr Leu Asp Phe Met Leu Gly Ala Lys
 1      5      10      15
Ala Asp Gly Glu Thr Ile Leu Lys Gly Leu Gln Ser Ile Phe Gln Glu
 20     25     30
Gln Gly Met Ala Glu Ser Val His Thr Trp Gln Asp His Gly Tyr Leu
 35     40     45
Ala Thr Tyr Thr Asn Lys Asn Gly Ser Phe Ala Asn Leu Arg Ile Tyr
 50     55     60
Pro His Gly Leu Val Leu Leu Asp Leu Gln Ser Tyr Asp Gly Asp Ala
 65     70     75     80
Gln Gly Lys Glu Glu Ile Asp Ser Ile Leu Asn Lys Val Glu Glu Arg
 85     90     95
Met Lys Glu Leu Ser Gln Asp Ser Thr Gly Arg Val Lys Arg Leu Pro
100    105    110
Pro Ile Val Arg Gly Gly Ala Ile Asp Arg Tyr Trp Pro Thr Ala Asp
115    120    125
Gly Arg Leu Val Glu Tyr Asp Ile Asp Glu Val Val Tyr Asp Glu Asp
130    135    140
Ser Pro Tyr Gln Asn Ile Lys Ile Leu His Ser Lys Gln Phe Gly Asn
145    150    155    160
Ile Leu Ile Leu Ser Gly Asp Val Asn Leu Ala Glu Ser Asp Leu Ala
165    170    175
Tyr Thr Arg Ala Ile Met Gly Ser Gly Lys Glu Asp Tyr Thr Gly Lys
180    185    190
Asp Val Leu Ile Leu Gly Gly Gly Asp Gly Gly Ile Leu Cys Glu Ile
195    200    205
Val Lys Leu Lys Pro Lys Met Val Thr Met Val Glu Ile Asp Gln Met
210    215    220
Val Ile Asp Gly Cys Lys Lys Tyr Met Arg Lys Thr Cys Gly Asp Val
225    230    235    240
Leu Asp Asn Leu Lys Gly Asp Cys Tyr Gln Val Leu Ile Glu Asp Cys
245    250    255
Ile Pro Val Leu Lys Arg Tyr Ala Lys Glu Gly Arg Glu Phe Asp Tyr
260    265    270
Val Ile Asn Asp Leu Thr Ala Val Pro Ile Ser Thr Ser Pro Glu Glu
275    280    285
Asp Ser Thr Trp Glu Phe Leu Arg Leu Ile Leu Asp Leu Ser Met Lys
290    295    300
Val Leu Lys Gln Asp Gly Lys Tyr Phe Thr Gln Gly Asn Cys Val Asn
305    310    315    320
Leu Thr Glu Ala Leu Ser Leu Tyr Glu Glu Gln Leu Gly Arg Leu Tyr
325    330    335
Cys Pro Val Glu Phe Ser Lys Glu Ile Val Cys Val Pro Ser Tyr Leu
340    345    350
Glu Leu Trp Val Phe Tyr Thr Val Trp Lys Lys Ala Lys Pro
355    360    365

```

<210> 39
 <211> 1807
 <212> DNA
 <213> 人

<220>
 <221> CDS
 <222> (276)... (1271)

```

<400> 39
gcacgaggga agagggtgat ccgaccgggg gaaggtcgct gggcaggggc agttgggaaa 60

```

```

gcggcagccc ccgccgcccc cgcagccctt tctcctcctt tctcccacgt cctatctgcc 120
tctcgttgga ggccaggccg tgcagatcg aagacaggag gaactggagc ctcataggcc 180
ggccccgggc gccggcctcg ggcttaata ggagctccgg gctctggctg ggacccgacc 240
gctgccggcc gcgctcccg tgctcctgcc gggtg atg gaa aac ccc agc ccg 293
Met Glu Asn Pro Ser Pro
1 5

gcc gcc gcc ctg ggc aag gcc ctc tgc gct ctc ctc ctg gcc act ctc 341
Ala Ala Ala Leu Gly Lys Ala Leu Cys Ala Leu Leu Ala Thr Leu
10 15 20

ggc gcc gcc ggc cag cct ctt ggg gga gag tcc atc tgt tcc gcc aga 389
Gly Ala Ala Gly Gln Pro Leu Gly Gly Glu Ser Ile Cys Ser Ala Arg
25 30 35

gcc ccg gcc aaa tac agc atc acc ttc acg ggc aag tgg agc cag acg 437
Ala Pro Ala Lys Tyr Ser Ile Thr Phe Thr Gly Lys Trp Ser Gln Thr
40 45 50

gcc ttc ccc aag cag tac ccc ctg ttc cgc ccc cct gcg cag tgg tct 485
Ala Phe Pro Lys Gln Tyr Pro Leu Phe Arg Pro Pro Ala Gln Trp Ser
55 60 65

tcg ctg ctg ggg gcc gcg cat agc tcc gac tac agc atg tgg agg aag 533
Ser Leu Leu Gly Ala Ala His Ser Ser Asp Tyr Ser Met Trp Arg Lys
75 80 85

aac cag tac gtc agt aac ggg ctg cgc gac ttt gcg gag cgc ggc gag 581
Asn Gln Tyr Val Ser Asn Gly Leu Arg Asp Phe Ala Glu Arg Gly Glu
90 95 100

gcc tgg gcg ctg atg aag gag atc gag gcg gcg ggg gag gcg ctg cag 629
Ala Trp Ala Leu Met Lys Glu Ile Glu Ala Ala Gly Glu Ala Leu Gln
105 110 115

agc gtg cac gcg gtg ttt tgc gcg ccc gcc gtc ccc agc ggc acc ggg 677
Ser Val His Ala Val Phe Ser Ala Pro Ala Val Pro Ser Gly Thr Gly
120 125 130

cag acg tcg gcg gag ctg gag gtg cag cgc agg cac tcg ctg gtc tcg 725
Gln Thr Ser Ala Glu Leu Glu Val Gln Arg Arg His Ser Leu Val Ser
135 140 145

ttt gtg gtg cgc atc gtg ccc agc ccc gac tgg ttc gtg ggc gtg gac 773
Phe Val Val Arg Ile Val Pro Ser Pro Asp Trp Phe Val Gly Val Asp
155 160 165

agc ctg gac ctg tgc gac ggg gac cgt tgg cgg gaa cag gcg gcg ctg 821
Ser Leu Asp Leu Cys Asp Gly Asp Arg Trp Arg Glu Gln Ala Ala Leu
170 175 180

gac ctg tac ccc tac gac gcc ggg acg gac agc ggc ttc acc ttc tcc 869
Asp Leu Tyr Pro Tyr Asp Ala Gly Thr Asp Ser Gly Phe Thr Phe Ser
185 190 195

tcc ccc aac ttc gcc acc atc ccg cag gac acg gtg acc gag ata acg 917
Ser Pro Asn Phe Ala Thr Ile Pro Gln Asp Thr Val Thr Glu Ile Thr
200 205 210

tcc tcc tct ccc agc cac ccg gcc aac tcc ttc tac tac ccg cgg ctg 965
Ser Ser Ser Pro Ser His Pro Ala Asn Ser Phe Tyr Tyr Pro Arg Leu
215 220 225

aag gcc ctg cct ccc atc gcc agg gtg aca ctg gtg cgg ctg cga cag 1013
Lys Ala Leu Pro Pro Ile Ala Arg Val Thr Leu Val Arg Leu Arg Gln
235 240 245

agc ccc agg gcc ttc atc cct ccc gcc cca gtc ctg ccc agc agg gac 1061
Ser Pro Arg Ala Phe Ile Pro Pro Ala Pro Val Leu Pro Ser Arg Asp
250 255 260

aat gag att gta gac agc gcc tca gtt cca gaa acg ccg ctg gac tgc 1109
Asn Glu Ile Val Asp Ser Ala Ser Val Pro Glu Thr Pro Leu Asp Cys
265 270 275

gag gtc tcc ctg tgg tcg tcc tgg gga ctg tgc gga ggc cac tgt ggg 1157
Glu Val Ser Leu Trp Ser Ser Trp Gly Leu Cys Gly Gly His Cys Gly

```

280	285	290	
agg ctc ggg acc aag agc agg act cgc tac gtc cgg gtc cag ccc gcc			1205
Arg Leu Gly Thr Lys Ser Arg Thr Arg Tyr Val Arg Val Gln Pro Ala			
295	300	305	310
aac aac ggg agc ccc tgc ccc gag ctc gaa gaa gag gct gag tgc gtc			1253
Asn Asn Gly Ser Pro Cys Pro Glu Leu Glu Glu Glu Ala Glu Cys Val			
	315	320	325
cct gat aac tgc gtc taa gaccagagcc ccgcagccccc tggggccccc			1301
Pro Asp Asn Cys Val *			
	330		
ggagccatgg ggtgtcgggg gctcctgtgc aggtctatgc tgcaggcggc cgaggcacag			1361
ggggtttcgc gctgctcctg accgcgggtga ggccgcgccc accatctctg cactgaaggg			1421
ccctctgggt gccggcacgg gcatitgggaa acagcctcct cctttcccaa ccttgcttct			1481
taggggcccc cgtgtccgt ctgctctcag cctcctcctc ctgcaggata aagtcacccc			1541
caaggctcca gctactctaa attatggtct ccttataagt tattgctgct ccaggagatt			1601
gtccttcac gtccaggggc ctggctccca cgtgggttga gatacctcag acctgggtgt			1661
ctaggctgtg ctgagccac tctcccgagg gcgcatccaa gcgggggcca cttgagaagt			1721
gaataaatgg ggccggttgc gaagcgtcag tgtttccatg ttatggatct ctctgcgttt			1781
gaataaagac taictctgtt gctcac			1807
<210> 40			
<211> 331			
<212> PRT			
<213> 人			
<400> 40			
Met Glu Asn Pro Ser Pro Ala Ala Ala Leu Gly Lys Ala Leu Cys Ala			
1 5 10 15			
Leu Leu Leu Ala Thr Leu Gly Ala Ala Gly Gln Pro Leu Gly Glu			
20 25 30			
Ser Ile Cys Ser Ala Arg Ala Pro Ala Lys Tyr Ser Ile Thr Phe Thr			
35 40 45			
Gly Lys Trp Ser Gln Thr Ala Phe Pro Lys Gln Tyr Pro Leu Phe Arg			
50 55 60			
Pro Pro Ala Gln Trp Ser Ser Leu Leu Gly Ala Ala His Ser Ser Asp			
65 70 75 80			
Tyr Ser Met Trp Arg Lys Asn Gln Tyr Val Ser Asn Gly Leu Arg Asp			
85 90 95			
Phe Ala Glu Arg Gly Glu Ala Trp Ala Leu Met Lys Glu Ile Glu Ala			
100 105 110			
Ala Gly Glu Ala Leu Gln Ser Val His Ala Val Phe Ser Ala Pro Ala			
115 120 125			
Val Pro Ser Gly Thr Gly Gln Thr Ser Ala Glu Leu Glu Val Gln Arg			
130 135 140			
Arg His Ser Leu Val Ser Phe Val Val Arg Ile Val Pro Ser Pro Asp			
145 150 155 160			
Trp Phe Val Gly Val Asp Ser Leu Asp Leu Cys Asp Gly Asp Arg Trp			
165 170 175			
Arg Glu Gln Ala Ala Leu Asp Leu Tyr Pro Tyr Asp Ala Gly Thr Asp			
180 185 190			
Ser Gly Phe Thr Phe Ser Ser Pro Asn Phe Ala Thr Ile Pro Gln Asp			
195 200 205			
Thr Val Thr Glu Ile Thr Ser Ser Ser Pro Ser His Pro Ala Asn Ser			
210 215 220			
Phe Tyr Tyr Pro Arg Leu Lys Ala Leu Pro Pro Ile Ala Arg Val Thr			
225 230 235 240			
Leu Val Arg Leu Arg Gln Ser Pro Arg Ala Phe Ile Pro Pro Ala Pro			
245 250 255			
Val Leu Pro Ser Arg Asp Asn Glu Ile Val Asp Ser Ala Ser Val Pro			
260 265 270			
Glu Thr Pro Leu Asp Cys Glu Val Ser Leu Trp Ser Ser Trp Gly Leu			
275 280 285			
Cys Gly Gly His Cys Gly Arg Leu Gly Thr Lys Ser Arg Thr Arg Tyr			
290 295 300			
Val Arg Val Gln Pro Ala Asn Asn Gly Ser Pro Cys Pro Glu Leu Glu			
305 310 315 320			
Glu Glu Ala Glu Cys Val Pro Asp Asn Cys Val			
	325	330	

<210> 41
 <211> 1528
 <212> DNA

<213> 人

<220>

<221> CDS

<222> (179)... (1123)

<400> 41

```

cggcgagcga gcaccttcga cgcgggtccgg ggacccccctc gtcgctgtcc tcccgcgcgc 60
gacccgcgtg ccccaggcct cgcgctgccg ggccggctcc tcgtgtccca ctcccggcgc 120
acgccctccc gcgagtcgcc ggccccctcc gcgccctct tctcggcgcg cgcgcagc 178
atg gcg ccc ccg cag gtc ctc gcg ttc ggg ctt ctg ctt gcc gcg gcg 226
Met Ala Pro Pro Gln Val Leu Ala Phe Gly Leu Leu Leu Ala Ala Ala
  1           5           10           15

acg gcg act ttt gcc gca gct cag gaa gaa tgt gtc tgt gaa aac tac 274
Thr Ala Thr Phe Ala Ala Ala Gln Glu Glu Cys Val Cys Glu Asn Tyr
          20           25           30

aag ctg gcc gta aac tgc ttt gtg aat aat aat cgt caa tgc cag tgt 322
Lys Leu Ala Val Asn Cys Phe Val Asn Asn Asn Arg Gln Cys Gln Cys
          35           40           45

act tca gtt ggt gca caa aat act gtc att tgc tca aag ctg gct gcc 370
Thr Ser Val Gly Ala Gln Asn Thr Val Ile Cys Ser Lys Leu Ala Ala
          50           55           60

aaa tgt ttg gtg atg aag gca gaa atg aat ggc tca aaa ctt ggg aga 418
Lys Cys Leu Val Met Lys Ala Glu Met Asn Gly Ser Lys Leu Gly Arg
          65           70           75           80

aga gca aaa cct gaa ggg gcc ctc cag aac aat gat ggg ctt tat gat 466
Arg Ala Lys Pro Glu Gly Ala Leu Gln Asn Asn Asp Gly Leu Tyr Asp
          85           90           95

cct gac tgc gat gag agc ggg ctc ttt aag gcc aag cag tgc aac ggc 514
Pro Asp Cys Asp Glu Ser Gly Leu Phe Lys Ala Lys Gln Cys Asn Gly
          100          105          110

acc tcc acg tgc tgg tgt gtg aac act gct ggg gtc aga aga aca gac 562
Thr Ser Thr Cys Trp Cys Val Asn Thr Ala Gly Val Arg Arg Thr Asp
          115          120          125

aag gac act gaa ata acc tgc tct gag cga gtg aga acc tac tgg atc 610
Lys Asp Thr Glu Ile Thr Cys Ser Glu Arg Val Arg Thr Tyr Trp Ile
          130          135          140

atc att gaa cta aaa cac aaa gca aga gaa aaa cct tat gat agt aaa 658
Ile Ile Glu Leu Lys His Lys Ala Arg Glu Lys Pro Tyr Asp Ser Lys
          145          150          155          160

agt ttg cgg act gca ctt cag aag gag atc aca acg cgt tat caa ctg 706
Ser Leu Arg Thr Ala Leu Gln Lys Glu Ile Thr Thr Arg Tyr Gln Leu
          165          170          175

gat cca aaa ttt atc acg agt att ttg tat gag aat aat gtt atc act 754
Asp Pro Lys Phe Ile Thr Ser Ile Leu Tyr Glu Asn Asn Val Ile Thr
          180          185          190

att gat ctg gtt caa aat tct tct caa aaa act cag aat gat gtg gac 802
Ile Asp Leu Val Gln Asn Ser Ser Gln Lys Thr Gln Asn Asp Val Asp
          195          200          205

ata gct gat gtg gct tat tat ttt gaa aaa gat gtt aaa ggt gaa tcc 850
Ile Ala Asp Val Ala Tyr Tyr Phe Glu Lys Asp Val Lys Gly Glu Ser
          210          215          220

ttg ttt cat tct aag aaa atg gac ctg aca gta aat ggg gaa caa ctg 898
Leu Phe His Ser Lys Lys Met Asp Leu Thr Val Asn Gly Glu Gln Leu
          225          230          235          240

gat ctg gat cct ggt caa act tta att tat tat gtt gat gaa aaa gca 946
Asp Leu Asp Pro Gly Gln Thr Leu Ile Tyr Tyr Val Asp Glu Lys Ala
          245          250          255

cct gaa ttc tca atg cag ggt cta aaa gct ggt gtt att gct gtt att 994
Pro Glu Phe Ser Met Gln Gly Leu Lys Ala Gly Val Ile Ala Val Ile
          260          265          270

```

```

gtg gtt gtg gtg ata gca gtt gtt gct gga att gtt gtg ctg gtt att 1042
Val Val Val Val Ile Ala Val Val Ala Gly Ile Val Val Leu Val Ile
      275                280                285

tcc aga aag aag aga atg gca aag tat gag aag gct gag ata aag gag 1090
Ser Arg Lys Lys Arg Met Ala Lys Tyr Glu Lys Ala Glu Ile Lys Glu
      290                295                300

atg ggt gag atg cat agg gaa ctc aat gca taa ctatataatt tgaagattat 1143
Met Gly Glu Met His Arg Glu Leu Asn Ala *
      305                310

agaagaaggg aaatagcaaa tggacacaaa ttacaaatgt gtgtgctgg gacgaagaca 1203
tctttgaagg tcatgagttt gttagtttaa catcatatat ttgtaatagt gaaacctgta 1263
ctcaaaatat aagcagcttg aaactggctt taccaatctt gaaatttgac cacaagtgtc 1323
ttatatatgc agatctaagc taaaatccag aacttggact ccacgtttaa aattatttat 1383
gtgtaacatt caaatgtgtg cattaatat gctccacag taaaatctga aaaactgatt 1443
tgtattgaa agctgccttt ctatttactt gactcttgta catacactact tttttatgag 1503
ctatgaaata aaacatttta aactg 1528

```

<210> 42
 <211> 314
 <212> PRT
 <213> 人

```

<400> 42
Met Ala Pro Pro Gln Val Leu Ala Phe Gly Leu Leu Leu Ala Ala Ala
 1          5          10          15
Thr Ala Thr Phe Ala Ala Ala Gln Glu Glu Cys Val Cys Glu Asn Tyr
      20          25          30
Lys Leu Ala Val Asn Cys Phe Val Asn Asn Asn Arg Gln Cys Gln Cys
      35          40          45
Thr Ser Val Gly Ala Gln Asn Thr Val Ile Cys Ser Lys Leu Ala Ala
      50          55          60
Lys Cys Leu Val Met Lys Ala Glu Met Asn Gly Ser Lys Leu Gly Arg
      65          70          75          80
Arg Ala Lys Pro Glu Gly Ala Leu Gln Asn Asn Asp Gly Leu Tyr Asp
      85          90          95
Pro Asp Cys Asp Glu Ser Gly Leu Phe Lys Ala Lys Gln Cys Asn Gly
      100         105         110
Thr Ser Thr Cys Trp Cys Val Asn Thr Ala Gly Val Arg Arg Thr Asp
      115         120         125
Lys Asp Thr Glu Ile Thr Cys Ser Glu Arg Val Arg Thr Tyr Trp Ile
      130         135         140
Ile Ile Glu Leu Lys His Lys Ala Arg Glu Lys Pro Tyr Asp Ser Lys
      145         150         155         160
Ser Leu Arg Thr Ala Leu Gln Lys Glu Ile Thr Thr Arg Tyr Gln Leu
      165         170         175
Asp Pro Lys Phe Ile Thr Ser Ile Leu Tyr Glu Asn Asn Val Ile Thr
      180         185         190
Ile Asp Leu Val Gln Asn Ser Ser Gln Lys Thr Gln Asn Asp Val Asp
      195         200         205
Ile Ala Asp Val Ala Tyr Tyr Phe Glu Lys Asp Val Lys Gly Glu Ser
      210         215         220
Leu Phe His Ser Lys Lys Met Asp Leu Thr Val Asn Gly Glu Gln Leu
      225         230         235         240
Asp Leu Asp Pro Gly Gln Thr Leu Ile Tyr Tyr Val Asp Glu Lys Ala
      245         250         255
Pro Glu Phe Ser Met Gln Gly Leu Lys Ala Gly Val Ile Ala Val Ile
      260         265         270
Val Val Val Val Ile Ala Val Val Ala Gly Ile Val Val Leu Val Ile
      275         280         285
Ser Arg Lys Lys Arg Met Ala Lys Tyr Glu Lys Ala Glu Ile Lys Glu
      290         295         300
Met Gly Glu Met His Arg Glu Leu Asn Ala
      305         310

```

<210> 43
 <211> 1051
 <212> DNA
 <213> 人

<220>
 <221> CDS
 <222> (98)... (274)

