

[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 200610100220.7

[51] Int. Cl.

C07K 14/705 (2006.01)

C07K 16/28 (2006.01)

C07K 19/00 (2006.01)

C12N 15/63 (2006.01)

C12N 15/62 (2006.01)

[43] 公开日 2007 年 3 月 28 日

[11] 公开号 CN 1935840A

[22] 申请日 1998.5.7

[21] 申请号 200610100220.7

分案原申请号 200510092715.5

[30] 优先权

[32] 1997.5.7 [33] US [31] 60/044293

[32] 1998.1.22 [33] US [31] 60/072212

[32] 1998.3.5 [33] US [31] 60/076947

[71] 申请人 先灵公司

地址 美国新泽西州凯尼尔沃斯盖洛平希尔
路 2000 号

[72] 发明人 G·T·哈迪曼 F·L·罗克
J·F·巴赞 R·A·卡斯特莱恩

[74] 专利代理机构 中国专利代理(香港)有限公司
代理人 刘 玥

权利要求书 2 页 说明书 160 页 附图 5 页

[54] 发明名称

人 Toll 样受体蛋白、相关试剂和方法

[57] 摘要

编码与果蝇属 Toll 受体和人 IL-1 受体同源的 9 种人受体的核酸,所述受体命名为 DNAX Toll 样受体 2-10(DTLR2-10);纯化的 DTLR 蛋白及其片段;针对这些受体的单克隆抗体/多克隆抗体;和诊断和治疗使用的方法。

1. 大致纯的或重组的 DTLR2 蛋白或肽, 它们在至少大约 12 个氨基酸的长度内, 表现出与 SEQ ID NO: 4 的序列同一性至少为大约 85%。

2. 大致纯的或重组的 DTLR3 蛋白或肽, 它们在至少大约 12 个氨基酸的长度内, 表现出与 SEQ ID NO: 6 的序列同一性至少为大约 85%。

3. 大致纯的或重组的 DTLR4 蛋白或肽, 它们在至少大约 12 个氨基酸的长度内, 表现出与 SEQ ID NO: 26 的序列同一性至少为大约 85%。

4. 大致纯的或重组的 DTLR5 蛋白或肽, 它们在至少大约 12 个氨基酸的长度内, 表现出与 SEQ ID NO: 10 的序列同一性至少为大约 85%。

5. 大致纯的或重组的 DTLR6 蛋白或肽, 它们在至少大约 12 个氨基酸的长度内, 表现出与 SEQ ID NO: 12 的序列同一性至少为大约 85%。

6. 大致纯的或重组的 DTLR7 蛋白或肽, 它们在至少大约 12 个氨基酸的长度内, 表现出与 SEQ ID NO: 16 或 18 的序列同一性至少为大约 85%。

7. 大致纯的或重组的 DTLR8 蛋白或肽, 它们在至少大约 12 个氨基酸的长度内, 表现出与 SEQ ID NO: 32 的序列同一性至少为大约 85%。

8. 大致纯的或重组的 DTLR9 蛋白或肽, 它们在至少大约 12 个氨基酸的长度内, 表现出与 SEQ ID NO: 22 的序列同一性至少为大约 85%。

9. 大致纯的或重组的 DTLR10 蛋白或肽, 它们在至少大约 12 个氨基酸的长度内, 表现出与 SEQ ID NO: 34 的序列同一性至少为大约

85%。

10. 融合蛋白，它包含权利要求 1-9 中任一项的蛋白或肽。

11. 结合化合物，它与权利要求 1-9 中任一项的蛋白或肽特异性地结合。

12. 权利要求 11 的结合化合物，它是一种抗体或抗体片段。

13. 核酸，它编码权利要求 1-9 中任一项的蛋白或肽。

14. 表达载体，它包含权利要求 13 的核酸。

15. 宿主细胞，它包含权利要求 14 的载体。

16. 重组生产多肽的方法，所述方法包括在表达所述多肽的条件下，培养权利要求 15 的宿主细胞。

人 Toll 样受体蛋白、相关试剂和方法

本申请是申请日为 1998 年 5 月 7 日，申请号为 200510092715.5，发明名称为“人 Toll 样受体蛋白、相关试剂和方法”的发明专利申请的分案申请。

该申请要求 1997 年 5 月 7 日申请的美国专利申请 USSN 60/044,293、1998 年 1 月 22 日申请的 USSN 60/072,212 和 1998 年 3 月 5 日申请的 USSN 60/076,947 的优先权，这些专利申请的每个均通过引用结合到本文中。

发明领域

本发明涉及影响哺乳动物生理学的组合物和方法，所述哺乳动物生理学包括形态发生或免疫系统功能。具体地说，本发明提供调节发育和/或免疫系统的核酸、蛋白质和抗体。也公开了这些物质的诊断用途和治疗用途。

发明背景

重组 DNA 技术一般是指这些技术，所述技术通过将来自供体来源的遗传信息整合到用于后续加工的载体中，诸如通过引入到宿主中，由此使所转移的遗传信息在该新环境中复制和/或表达。通常，该遗传信息以互补 DNA (cDNA) 的形式存在，所述互补 DNA 衍生自编码所需的蛋白产物的信使 RNA (mRNA)。该载体通常是一种质粒，该质粒具有将用于随后复制的 cDNA 加入宿主中的能力，在某些情况下，实际控制该 cDNA 的表达，并由此指导所编码的产物在该宿主中合成。

一段时间以来，已经知道哺乳动物的免疫应答基于称为“免疫网

络”的一系列复杂的细胞相互作用。最近的研究已经提供了对该网络内部运作的新的了解。尽管已经清楚许多免疫应答实际上确实围绕淋巴细胞、巨噬细胞、粒细胞和其它细胞的网络样相互作用运转,但免疫学家现在一般坚持这样一种观点:称为淋巴因子、细胞因子或单核因子的可溶性蛋白,在控制这些细胞相互作用中起关键作用。因此,人们对细胞调节因子的分离、特征鉴定和作用机制相当感兴趣,对这些信息的了解将导致在许多医学异常(例如免疫系统失调)的诊断和治疗方面取得显著的进展。

淋巴因子看来以多种方式介导细胞活性。已经表明,它们支持多能造血干细胞的增殖、生长和/或分化为大量的祖细胞(progenitor),所述祖细胞包括构成复杂的免疫系统的多样化的细胞谱系。细胞组分之间适当和平衡的相互作用是健康的免疫应答所必需的。当结合其它因子给予淋巴因子时,不同的细胞谱系常常以不同方式应答。

对免疫应答尤其重要的细胞谱系包括两类淋巴细胞: B 细胞,它们可以产生和分泌免疫球蛋白(具有识别并结合外源物质以实现将其去除的能力的蛋白);和各种亚群的 T 细胞,它们分泌淋巴因子,并诱导或抑制 B 细胞和构成免疫网络的各种其它细胞(包括其它 T 细胞)。这些淋巴细胞与许多其它细胞类型相互作用。

另一种重要的细胞谱系是肥大细胞(在所有哺乳动物物种中尚未被明确鉴定),这是一种含颗粒的结缔组织细胞,邻近遍布机体的毛细血管定位。发现这些细胞在肺、皮肤以及胃肠道和生殖泌尿道中的浓度尤其高。肥大细胞如下在变态反应相关的病征(特别是过敏反应)中起中心作用:当选定的抗原与结合于肥大细胞表面上受体的一类免疫球蛋白交联时,该肥大细胞脱粒并释放引起变态反应(例如过敏反应)的介质,例如组胺、5-羟色胺、肝素和前列腺素。

一般不能在体外维持免疫系统的细胞,这已经阻碍了为更好地了解和治疗各种免疫失调而进行的研究。免疫学家已经发现,通过使用 T 细胞和其它细胞的上清液,可以完成对许多这些细胞的培养,所述

上清液含有各种生长因子, 包括许多淋巴因子。

蛋白的白介素-1 家族包括 IL-1 α 、IL-1 β 、IL-1RA 和最近的 IL-1 γ (也命名为干扰素- γ 诱导因子, IGIF)。在范围广泛的生物学功能中已经牵涉到该基因相关家族。参见 Dinarello (1994) EASEB J 8:1314-1325; Dinarello (1991) Blood 77:1627-1652; 和 Okamura 等 (1995) Nature 378:88-91。

此外, 存在调节形态发生发育的各种生长因子和调节因子。这包括例如 Toll 配体, 它们通过结合于享有 IL-1 受体特征性的结构和机制特征的受体而发送信号。参见例如 Lemaitre 等 (1996) Cell 86:973-983; Belvin 和 Anderson (1960) Ann. Rev. Cell & Devel. Biol. 12:393-416。

根据以上所述, 可明显看出, 新的可溶性蛋白及其受体(包括类似于淋巴因子的蛋白)的发现和研制, 应该有助于用于范围广泛的退行性病征或异常病征的新疗法, 所述病征直接或间接涉及例如免疫系统和/或造血细胞的发育、分化或功能。特别是, 对于增加或增强其它淋巴因子有益活性的淋巴因子样分子, 其新受体的发现和了解应该是非常有利。本发明提供新的受体和相关化合物及其使用方法, 其中所述受体为表现出与白介素-1 样组合物相似性的配体的受体。

附图简述

图 1 显示了果蝇属和人的 DTLR 蛋白结构的图示比较、以及它们与脊椎动物 IL-1 受体和植物抗病性蛋白的关系。三种果蝇属(Dm) DTLR (Toll、18w 和 Mst ORF 片段) (Morisato 和 Anderson (1995) Ann. Rev. Genet. 29:371-399; Chiang 和 Beachy (1994) Mech. Develop. 47:225-239; Mitcham 等(1996) J. Biol. Chem. 271:5777-5783; 和 Eldon 等 (1994) Develop. 120:885-899)排列在四个完整的(DTLR 1-4)和一个部分的(DTLR5)人(Hu)受体旁边。清楚地框注了通过 PRINTS (Attwood 等 (1997) Nucleic Acids Res. 25:212-217)标记的所述受体胞外域的各个 LRR; 邻接 LRR 阵列(LRR array) C 末端或 N 末端的“顶部”和“底部”

富含 Cys 的簇, 分别用汇合的半圆画出。当与 Toll 和 18w 的 784 aa 和 977 aa 的延伸物相比时, DTLR 1-5 中内部 Cys 富含区的丧失, 主要产生其较小的胞外域(分别为 558、570、690 和 652 aa)。DmMst 和 HuDTLR5 的不完整链(分别为 519 aa 和 153 aa 的胞外域)由虚线表示。与 DTLR、IL-1 型受体(IL-1R)、胞内蛋白 Myd88 和烟草抗病性基因 N 产物(DRgN)共同的胞内信号组件, 显示在该质膜下面。参见例如 Hardiman 等 (1996) Oncogene 13:2467-2475; 和 Rock 等 (1998) Proc. Nat'l Acad. Sci. USA 95:588-。其它域包括 IL-1R 中 Ig 样组件的 trio (二硫键连接的环); DRgN 蛋白的特征是一个 NTP 酶域(方框), 而 Myd88 具有一个死域(death domain)(黑色椭圆)。

图 2A-2B 显示 Toll 样和 IL-1 样细胞因子受体和两个分歧的组件蛋白信号域中的保守结构模式。图 2A 显示共同的 TH 域的序列对比。DTLR 按图 1 中进行标记; 人(Hu)或小鼠(Mo) IL-1 家族受体(IL-1R1-6)按早年提出的方法(Hardiman 等 (1996) Oncogene 13:2467-2475)顺序编号; Myd88 和来自烟草(To)和亚麻 *L. usitatissimum* (Lu)的序列, 分别代表较大的多域分子的 C 末端域和 N 末端域。框注了不间断的序列区组(编号为 1-10)。三角表示有害的突变, 而箭头 N 末端的截短物消除人 IL-1R1 中的生物活性(Heguy 等 (1992) J. Biol. Chem. 267:2605-2609)。标出了 α -螺旋(H)、 β -链(E)或螺旋(L)的 PHD (Rost 和 Sander (1994) Proteins 19:55-72)和 DSC (King 和 Sternberg (1996) Protein Sci. 5:2298-2310)二级结构预测。氨基酸遮蔽方案描述了化学上相似的残基: 疏水性、酸性、碱性、Cys、芳族、结构断裂性和微小的残基。通过 75% 严格性下的共有序列, 得出 IL-1R、DTLR 和全部序列对比(ALL)的诊断序列模式。氨基酸亚组的符号是(关于细节请参见互联网网站(internet site)): o, 醇; l, 脂族的; •, 任一氨基酸; a, 芳族的; c, 带电荷的; h, 疏水性; -, 阴性; p, 极性; +, 阳性; s, 小的; u, 微小的; t, 转角样的。图 2B 显示了建议的 TH β/α 域折叠的拓扑图。在其 C 末端观察到平行的 β -片(具有黄色三角的 β -链 A-E); α -螺旋(标记 1-5 的圆)

连接所述 β -链；链连接在前面(可见的)或后面(隐藏的)。以灰色(Asp)或作为单独的黑色(Arg)残基,标注 β -片 C 末端保守的带电荷的残基(参见正文)。

图 3 显示信号域超家族的进化。通过相邻连接法(Thompson 等, (1994) *Nucleic Acids Res.* 22:4673-4680), 用图 2A 的多重 TH 组件序列对比得出进化系统树。同序列对比中一样标记蛋白; 用 TreeView 提供该树。

图 4A-4D 显示人 DTLR 基因的 FISH 染色体作图。使来自人淋巴细胞同步培养物的变性染色体, 与用于定位的生物素化 DTLR cDNA 探针杂交。通过使 FISH 信号与 DAPI 显带的染色体重叠(中心图板), 得到具有染色体带的 FISH 作图数据(左边, 图 4A, DTLR2; 4B, DTLR3; 4C, DTLR4; 4D, DTLR5)的对比。Heng 和 Tsui (1994) *Meth. Molec. Biol.* 33:109-122. 以人染色体模式图(ideograms)(右板)的形式总结分析。

图 5A-5F 显示人 DTLR 的 mRNA 印迹分析。用放射标记的、编码上述 DTLR1 (图 5A-5 C)、DTLR2 (图 5D)、DTLR3 (图 5E)和 DTLR4 (图 5F)的 cDNA, 探测每泳道含有大约 2 μ g 多聚(A)⁺ RNA 的人多种组织的印迹(He, 心脏; Br, 脑; Pl, 胎盘; Lu, 肺; Li, 肝脏; Mu, 肌肉; Ki, 肾脏; Pn, 胰脏; Sp, 脾脏; Th, 胸腺; Pr, 前列腺; Te, 睾丸; Ov, 卵巢; SI, 小肠; Co, 结肠; PBL, 外周血淋巴细胞)和癌细胞系(早幼粒细胞性白血病, HL60; 宫颈癌, HELAS3; 慢性髓细胞性白血病, K562; 淋巴细胞性白血病, Molt4; 结肠直肠腺癌, SW480; 黑素瘤, G361; 伯基特淋巴瘤 Raji, Burkitt's; 结肠直肠腺癌, SW480; 肺癌, A549)。于-70℃用增感屏将印迹对 X 光片曝光 2 天(图 5A-5C)或一周(图 5D-5F)。在某些泳道中出现异常的 0.3 kB 种类; 杂交实验排除了编码 DTLR 胞质片段的信息。

发明概述

本发明涉及 9 种新的相关哺乳动物受体(例如人)-Toll 受体样分子的结构及其生物活性,所述受体命名为 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 和 DTLR10。本发明包括编码所述多肽本身的核酸以及它们的生产和使用方法。本发明核酸的特征,部分在于它们与本文公开的克隆的互补 DNA (cDNA)序列的同源性。

在某些实施方案中,本发明提供选自以下物质的组合物:大致纯的或重组的 DTLR2 蛋白或肽,它在至少大约 12 个氨基酸的长度内,表现出与 SEQ ID NO: 4 的序列同一性至少为 85%;SEQ ID NO: 4 的天然序列 DTLR2;包含 DTLR2 序列的融合蛋白;大致纯的或重组的 DTLR3 蛋白或肽,它在至少大约 12 个氨基酸的长度内,表现出与 SEQ ID NO: 6 的序列同一性至少为 85%;SEQ ID NO: 6 的天然序列 DTLR3;包含 DTLR3 序列的融合蛋白;大致纯的或重组的 DTLR4 蛋白或肽,它在至少大约 12 个氨基酸的长度内,表现出与 SEQ ID NO: 26 的序列同一性至少为 85%;SEQ ID NO: 26 的天然序列 DTLR4;包含 DTLR4 序列的融合蛋白;大致纯的或重组的 DTLR5 蛋白或肽,它在至少大约 12 个氨基酸的长度内,表现出与 SEQ ID NO: 10 的序列同一性至少为 85%;SEQ ID NO: 10 的天然序列 DTLR5;和包含 DTLR5 序列的融合蛋白。

在其它实施方案中,本发明提供选自以下物质的组合物:大致纯的或重组的 DTLR6 蛋白或肽,它在至少大约 12 个氨基酸的长度内,表现出与 SEQ ID NO: 12 的序列同一性至少为 85%;SEQ ID NO: 12 的天然序列 DTLR6;包含 DTLR6 序列的融合蛋白;大致纯的或重组的 DTLR7 蛋白或肽,它在至少大约 12 个氨基酸的长度内,表现出与 SEQ ID NO: 16 或 18 的序列同一性至少为 85%,或;SEQ ID NO: 16 或 18 的天然序列 DTLR7;包含 DTLR7 序列的融合蛋白;大致纯的或重组的 DTLR8 蛋白或肽,它在至少大约 12 个氨基酸的长度内,表现

出与 SEQ ID NO: 32 的序列同一性至少为 85%; SEQ ID NO: 32 的天然序列 DTLR8; 包含 DTLR8 序列的融合蛋白; 大致纯的或重组的 DTLR9 蛋白或肽, 它在至少大约 12 个氨基酸的长度内, 表现出与 SEQ ID NO: 22 的序列同一性至少为 85%; SEQ ID NO: 22 的天然序列 DTLR9; 和包含 DTLR9 序列的融合蛋白; 大致纯的或重组的 DTLR10 蛋白或肽, 它在至少大约 12 个氨基酸的长度内, 表现出与 SEQ ID NO: 34 的序列同一性至少为 85%; SEQ ID NO: 34 的天然序列 DTLR10; 和包含 DTLR10 序列的融合蛋白。

最好是, 所述大致纯的或分离的蛋白包含与以下序列的相应部分表现出序列同一性的区段: DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 或 DTLR10, 其中: 所述同源性至少为大约 90% 的同一性, 而所述部分为至少大约 9 个氨基酸; 所述同源性至少为大约 80% 的同一性, 而所述部分为至少大约 17 个氨基酸; 或所述同源性至少为大约 70% 的同一性, 而所述部分为至少大约 25 个氨基酸。在具体实施方案中, 所述物质的组合物: 为 DTLR2, 它包含 SEQ ID NO: 4 的成熟序列; 或表现出与天然 DTLR2 不同的翻译后 + 修饰模式; 为 DTLR3, 它: 包含 SEQ ID NO: 6 的成熟序列, 或表现出与天然 DTLR3 不同的翻译后修饰模式; 为 DTLR4, 它: 包含 SEQ ID NO: 26 的成熟序列, 或表现出与天然 DTLR4 不同的翻译后修饰模式; 或为 DTLR5, 它: 包含 SEQ ID NO: 10 的完整序列, 或表现出与天然 DTLR5 不同的翻译后修饰模式; 或为 DTLR6, 它: 包含 SEQ ID NO: 12 的成熟序列, 或表现出与天然 DTLR6 不同的翻译后修饰模式; 为 DTLR7, 它: 包含 SEQ ID NO: 16 或 18 的成熟序列, 或表现出与天然 DTLR7 不同的翻译后修饰模式; 为 DTLR8, 它: 包含 SEQ ID NO: 32 的成熟序列, 或表现出与天然 DTLR8 不同的翻译后修饰模式; 或为 DTLR9, 它: 包含 SEQ ID NO: 22 的完整序列, 或表现出与天然 DTLR9 不同的翻译后修饰模式; 或为 DTLR10, 它: 包含 SEQ ID NO: 34 的完整序列, 或表现出与天然 DTLR10 不同的翻译后修饰模式; 或所述物质的组合

物可以是一种蛋白或肽，该蛋白或肽：来自选自哺乳动物(包括诸如人的灵长类)的温血动物；包含 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 的至少一种肽区段；表现出多个部分显示所述同一性；为 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 或 DTLR10 的天然等位基因变异体；其长度至少为大约 30 个氨基酸；表现出至少两个非重叠表位是灵长类 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 或 DTLR10 特有的；在至少大约 20 个氨基酸的长度内，表现出与灵长类 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6 的序列同一性至少为大约 90%；表现出至少两个非重叠表位是灵长类 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 或 DTLR10 特有的；在至少大约 20 个氨基酸的长度内，表现出与灵长类 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 或 DTLR10 的序列同一性至少为大约 90%；为糖基化的；其分子量至少为 100 kD，并且被天然糖基化；为合成多肽；连接至固体基质；与另一化学部分缀合；为天然序列的 5 重或更低重取代的取代体；或为天然序列的缺失或插入变异体。

其它实施方案包括组合物，它包含：无菌 DTLR2 蛋白或肽；或所述 DTLR2 蛋白或肽和一种载体，其中所述载体为：水性化合物，包括水、盐水和/或缓冲液；和/或配制用于经口、直肠、鼻、局部或胃肠外给药；无菌 DTLR3 蛋白或肽；或所述 DTLR3 蛋白或肽和一种载体，其中所述载体为：水性化合物，包括水、盐水和/或缓冲液；和/或配制用于经口、直肠、鼻、局部或胃肠外给药；无菌 DTLR4 蛋白或肽；或所述 DTLR4 蛋白或肽和一种载体，其中所述载体为：水性化合物，包括水、盐水和/或缓冲液；和/或配制用于经口、直肠、鼻、局部或胃肠外给药；无菌 DTLR5 蛋白或肽；或所述 DTLR5 蛋白或肽和一种载体，其中所述载体为：水性化合物，包括水、盐水和/或缓冲液；和/或配制用于经口、直肠、鼻、局部或胃肠外给药；无菌 DTLR6 蛋白或肽；或

所述 DTLR6 蛋白或肽和一种载体, 其中所述载体为: 水性化合物, 包括水、盐水和/或缓冲液; 和/或配制用于经口、直肠、鼻、局部或胃肠外给药; 无菌 DTLR7 蛋白或肽; 或所述 DTLR7 蛋白或肽和一种载体, 其中所述载体为: 水性化合物, 包括水、盐水和/或缓冲液; 和/或配制用于经口、直肠、鼻、局部或胃肠外给药; 无菌 DTLR8 蛋白或肽; 或所述 DTLR8 蛋白或肽和一种载体, 其中所述载体为: 水性化合物, 包括水、盐水和/或缓冲液; 和/或配制用于经口、直肠、鼻、局部或胃肠外给药; 无菌 DTLR9 蛋白或肽; 或所述 DTLR9 蛋白或肽和一种载体, 其中所述载体为: 水性化合物, 包括水、盐水和/或缓冲液; 和/或配制用于经口、直肠、鼻、局部或胃肠外给药; 无菌 DTLR10 蛋白或肽; 或所述 DTLR10 蛋白或肽和一种载体, 其中所述载体为: 水性化合物, 包括水、盐水和/或缓冲液; 和/或配制用于经口、直肠、鼻、局部或胃肠外给药。

在某些融合蛋白实施方案中, 本发明提供一种融合蛋白, 它包含: SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 的成熟蛋白序列; 一种检测或纯化标记, 包括 FLAG、His6 或 Ig 序列; 或另一受体蛋白的序列。

各种试剂盒实施方案包括的试剂盒包含一种 DTLR 蛋白或多肽, 和: 包含所述蛋白或多肽的区室; 和/或使用或处理该试剂盒中试剂的说明。

结合化合物的实施方案包括那些化合物, 它们包含来自抗体的抗原结合位点, 该位点特异性地结合于天然的 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 或 DTLR10 蛋白, 其中所述蛋白为灵长类蛋白; 所述结合化合物为 Fv、Fab 或 Fab2 片段; 所述结合化合物与另一化学部分缀合; 或所述抗体: 针对 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 的成熟肽的肽序列而产生; 针对成熟的 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 或 DTLR10 而产生; 针对纯化的人 DTLR2、DTLR3、DTLR4、

DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 或 DTLR10 而产生；为免疫选择性的；为多克隆抗体；结合于变性的 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 或 DTLR10；表现出与抗原的 K_d 至少为 30 μM ；连接至固体基质，包括珠粒或塑料膜；在无菌组合物中；或进行了可检测标记，包括放射性标记或荧光标记。结合组合物的试剂盒通常包含所述结合化合物，和：包含所述结合化合物的区室；和/或使用或处理该试剂盒中试剂的说明。该试剂盒通常能够进行定性或定量分析。

其它组合物包括这样的组合物，它包含：无菌结合化合物，或所述结合化合物和一种载体，其中所述载体为：水性化合物，包括水、盐水和/或缓冲液；和/或配制用于经口、直肠、鼻、局部或胃肠外给药。

核酸的实施方案包括分离的或重组的核酸，它编码 DTLR2-10 的蛋白或肽或融合蛋白，其中：所述 DTLR 来自哺乳动物；或所述核酸：编码 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 的抗原肽序列；编码多个 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 的抗原肽序列；表现出与编码所述区段的天然 cDNA 的同一性至少为大约 80%；为一种表达载体；还包含一个复制起点；来自天然来源；包含一种可检测标记；包含合成的核苷酸序列；小于 6 kb，最好小于 3 kb；来自哺乳动物，包括灵长类；包含一种天然的全长编码序列；为编码所述 DTLR 的基因的杂交探针；或为一种 PCR 引物、PCR 产物或诱变引物。也提供包含这样一种重组核酸的细胞、组织或器官。所述细胞最好是：原核细胞；真核细胞；细菌细胞；酵母细胞；昆虫细胞；哺乳动物细胞；小鼠细胞；灵长类细胞；或人细胞。提供的试剂盒包含这类核酸，和：包含所述核酸的区室；还包含灵长类 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 或 DTLR10 蛋白或多肽的区室；和/或使用或处理该试剂盒中试剂的说明。通常，该试剂盒能够进行定性或定量分析。

其它实施方案包括一种核酸，它：在 30℃ 和低于 2M 盐的洗涤条

件下与 SEQ ID NO: 3 杂交; 在 30℃和低于 2M 盐的洗涤条件下与 SEQ ID NO: 5 杂交; 在 30℃和低于 2M 盐的洗涤条件下与 SEQ ID NO: 25 杂交; 在 30℃和低于 2M 盐的洗涤条件下与 SEQ ID NO: 9 杂交; 在 30℃和低于 2M 盐的洗涤条件下与 SEQ ID NO: 11 杂交; 在 30℃和低于 2M 盐的洗涤条件下与 SEQ ID NO: 15 或 17 杂交; 在 30℃和低于 2M 盐的洗涤条件下与 SEQ ID NO: 31 杂交; 在 30℃和低于 2M 盐的洗涤条件下与 SEQ ID NO: 21 杂交; 在 30℃和低于 2M 盐的洗涤条件下与 SEQ ID NO: 33 杂交; 在至少大约 30 个核苷酸的一段序列内, 表现出与灵长类 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 或 DTLR10 的同一性至少为大约 85%。

优选的是, 这种核酸将具有这类性质, 其中: 洗涤条件为 45℃和/或 500 mM 盐; 或所述同一性至少为 90%和/或所述一段序列至少为 55 个核苷酸。更优选的是, 所述洗涤条件为 55℃和/或 150 mM 盐; 或所述同一性至少为 95%和/或所述一段序列至少为 75 个核苷酸。

本发明也提供调节细胞或组织培养细胞生理学或发育的方法, 该方法包括使所述细胞与哺乳动物 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 或 DTLR10 的激动剂或拮抗剂接触。

优选实施方案的详述

I. 一般描述

本发明提供哺乳动物、本文的灵长类 DNAX Toll 样受体分子 (DTLR) 的氨基酸序列和 DNA 序列, 所述受体分子具有特别限定的结构和生物学特性。这些受体分子在本文中已经分别命名为 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 和 DTLR10, 它们将人 Toll 样受体家族的成员数由 1 增加至 10。由例如人的灵长类的 cDNA 序列文库, 获得编码这些分子的各种 cDNA。也可能需要其它灵长类或其它哺乳动物的相应物。

例如在 Maniatis 等 (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor Press; Sambrook 等 (1989) Molecular Cloning: A Laboratory Manual, (第二版), 第 1-3 卷, CHS Press, NY; Ausubel 等, Biology, Greene Publishing Associates, Brooklyn, NY; 或 Ausubel 等 (1987 和定期增刊) Current Protocols in Molecular Biology, Greene/Wiley, New York 中描述或引用了可应用的某些标准方法, 每个参考文献通过引用结合到本文中。

在 SEQ ID NO: 1 和 2 中显示了人 DTLR1 编码区段的完整的核苷酸序列和相应的氨基酸序列。也参见 Nomura 等 (1994) DNA Res 1:27-35。在 SEQ ID NO: 3 和 4 中显示了人 DTLR2 编码区段的完整的核苷酸序列和相应的氨基酸序列。在 SEQ ID NO: 5 和 6 中显示了人 DTLR3 编码区段的完整的核苷酸序列和相应的氨基酸序列。在 SEQ ID NO: 7 和 8 中显示了人 DTLR4 编码区段的完整的核苷酸序列和相应的氨基酸序列。在 SEQ ID NO: 25 和 26 中显示了人 DTLR4 编码区段的替代的核苷酸序列和相应的氨基酸序列。在 SEQ ID NO: 9 和 10 中显示了人 DTLR5 编码区段的部分核苷酸序列和相应的氨基酸序列。在 SEQ ID NO: 11 和 12 中显示了人 DTLR6 编码区段的完整的核苷酸序列和相应的氨基酸序列, 而在 SEQ ID NO: 13 和 14 中提供了小鼠 DTLR6 的部分序列。在 SEQ ID NO: 27 和 29 (核苷酸序列) 和 SEQ ID NO: 28 和 30 (氨基酸序列) 中提供了另一小鼠 DTLR6 的序列。也提供了人 DTLR7 编码区段的部分核苷酸序列 (SEQ ID NO: 15 和 17) 和相应的氨基酸序列 (SEQ ID NO: 16 和 18)。在 SEQ ID NO: 19 和 20 中显示了人 DTLR8 编码区段的部分核苷酸序列和相应的氨基酸序列。在 SEQ ID NO: 31 和 32 中显示了一种人 DTLR 编码区段的更完整的核苷酸序列和相应的氨基酸序列。在 SEQ ID NO: 21 和 22 中显示了人 DTLR9 编码区段的部分核苷酸序列和相应的氨基酸序列。在 SEQ ID NO: 23 和 24 中显示了人 DTLR10 编码区段的部分核苷酸序列和相应的氨基酸序列。在 SEQ ID NO: 33 和 34 中显示了人 DTLR10 编码区段的更完整

的核苷酸序列和相应的氨基酸序列。在 SEQ ID NO: 35 中显示了小鼠 DTLR10 编码区段的部分核苷酸序列。

表 1: 人 DTLR 胞内域的比较。DTLR1 为 SEQ ID NO: 2; DTLR2 为 SEQ ID NO: 4; DTLR3 为 SEQ ID NO: 6; DTLR4 为 SEQ ID NO: 8; DTLR5 为 SEQ ID NO: 10; 而 DTLR6 为 SEQ ID NO: 12. 在所述 DTLR 中, 特别重要和保守的(例如特征性的)残基对应于 SEQ ID NO: 18 的残基 try10-tyr13; trp26; cys46; trp52; pro54-gly55; ser69; lys71; trp134-pro135; 和 phe144-trp145.

```
DTLR1  QRNLOFHAFISYSGHD---SFWVKNELLPNLEKEG-----MQICLHERNF
DTLR9  KENLOFHAFISYSEHD---SAWVKSELVPYLEKED-----IQICLHERNF
DTLR8  -----NELIPNLEKEDGS---ILICLYESYF
DTLR2  SRNICYDAFVSYSERD---AYWVENLMVQELFNFPF---FKLCLHKRDF
DTLR6  SPDCCYDAFIVYDTKDPVTEWVLAELVAKLEDPREK--HFNLCLLEERDW
DTLR7  TSQTFYDAYISYDTKDASVTDWVINELRYHLEESRDK--NVLLCLEERDW
DTLR10 EDALPYDAFVVFDTKTXSAVADWVYNELRGQLEECRGRW-ALRLCLEERDW
DTLR4  RGENIYDAFVIYSSQD---EDWVRNELVKNLLEGVPP---FQLCLHYRDF
DTLR5  PDMKYDAYLCFSSKD---FTWVQNALLKHLDTQYSDQNRFNLCFEERDF
DTLR3  TEQFEYAAYIIHAYKD---KDWVWEHFSMEKEDQS---LKFCLEERDF
      :      .      .      : * :      :
```

```
DTLR1  VPGKSIVENIITC-IEKSYKSIFVLSPNFVQSEWCH-YELYFAHHNLFHE
DTLR9  VPGKSIVENIINC-IEKSYKSIFVLSPNFVQSEWCH-YELYFAHHNLFHE
DTLR8  DPGKSISENIVSF-IEKSYKSIFVLSPNFVQNEWCH-YEPYFAHHNLFHE
DTLR2  IPGKWIIDNIIDS-IEKSHKTVFVLSENFVKSEWCK-YELDFSHFRLFEE
DTLR6  LPGQPVLENLSQS-IQLSKKTVFVMTDKYAKTENFK-IAFYLSHORLMDE
DTLR7  DPGLAIIIDNMQS-INQSKKTVFVLTKKYAKSWNFK-TAFYLLXQRLMGE
DTLR10 LPGKTLFENLWAS-VYGSRKTLFVLAHTDRVSGLLR-AIFLLAQORLLE-
DTLR4  IPGVAIAANIIHEGFHKSRKVIVVVSQHFIQSRWCI-FEYELAQTWQFLS
DTLR5  VPGENRIANIQDA-IWNSRKIVCLVSRHFLRDGWCL-EAFSYAQGRCLSD
DTLR3  EAGVFELEAIVNS-IKRSRKIIFVITHLLKQPLCKRFKVHHAQQQAIEQ
      . *      :      .      * * : : : :      :
```

```
DTLR1  GSNLILILILLEPIPOYSIPSSYHKLKSLMARRTYLEWPKEKSKRGLFWAN
DTLR9  GSNLILILILLEPIQNSIPNKYHKLKALMTORTYLQWPKEKSKRGLFWA-
DTLR8  NSDHIILILLEPIPFYCIPTRYHKLKLEALLEKKAYLEWPKDRRKCGLFWAN
DTLR2  NNDAAIILILLEPIEKKAIPQRFCKLRKIMNTKTYLEWPMDEAOREGFVWN
DTLR6  KVDVIIILIFLEKPFQK---SKFLQLRKRLCGSSVLEWPTNPQAHFYFWQC
DTLR7  NMDVIFILILEPVLQH---SPYLRRLRQIRICKSSILQWPDNPKAERLFWQT
DTLR10 -----
DTLR4  SRAGIIFIVLQKVEKT-LLRQQVELYRLLSRNTYLEWEDSVLGRHIFWRR
DTLR5  LNSALIMVVVGSLSQY-QLMKHQSIQGFVQKQQLRWPEDLQDVGWFLHK
DTLR3  NLDSSIILVLEEIPDYKLNHALCLRRGMFKSHCIILNWPVQKERIGAFRHK
```

```
DTLR1  LRAAINIKLTEQAKK-----
DTLR9  -----
DTLR8  LRAAVNVNVLATREMYELQTFTELNEESRGSTISLMRTDCL
DTLR2  LRAAIKS-----
DTLR6  LKNALATDNHVAYSQVFKETV-----
DTLR7  LKNVVLTEENDSRYNMIVDSIKQY-----
DTLR10 -----
DTLR4  LRKALLDGKSWNPEGTVGTGCNWOEATSI-----
DTLR5  LSQQILKKEKEKKKDNNIPLQTVATIS-----
DTLR3  LQVALGSKNSVH-----
```

本文所用的术语 DNAX Toll 样受体 2 (DTLR2)将用来描述一种蛋白, 包括蛋白或肽区段, 所述蛋白或肽区段具有和享有 SEQ ID NO: 4 中所示的氨基酸序列或其大致的片段。DTLR3 和 SEQ ID NO: 6、DTLR4 和 SEQ ID NO: 26、DTLR5 和 SEQ ID NO: 10、DTLR6 和 SEQ ID NO: 12、DTLR7 和 SEQ ID NO: 16 和 18、DTLR8 和 SEQ ID NO: 32、DTLR9 和 SEQ ID NO: 22、以及 DTLR10 和 SEQ ID NO: 34 具有类似的对应关系。

本发明也包括已提供其序列的相应 DTLR 等位基因的蛋白变异体, 例如变异蛋白激动剂或拮抗剂。通常, 这类激动剂或拮抗剂会表现出低于大约 10% 的序列差异, 因此常常会具有 1-11 重取代, 例如 2 重、3 重、5 重、7 重和其它重取代。也包括所述蛋白的等位基因变异体和其它变异体, 例如天然的多态变异体。通常, 它会高亲和性地结合于其相应的生物学受体, 其亲和性例如为至少大约 100 nM, 通常优于大约 30 nM, 优选优于大约 10 nM, 更优选优于大约 3 nM。在本文中该术语也将用来指所述哺乳动物蛋白相关的天然存在的形式, 例如等位基因变异体、多态变异体和代谢变异体。

本发明也包括与 SEQ ID NO: 4 中的氨基酸序列具有大致氨基酸序列同一性的蛋白或肽。它将包括具有相对少取代的序列变异体, 例如最好少于 3-5 重取代。将相似的特征应用于 SEQ ID NO: 6、26、10、12、16、18、32、22 和 34 中提供的其它 DTLR 序列。

大致的多肽“片段”或“区段”是指一段氨基酸残基序列, 它至少大约为 8 个氨基酸, 一般至少为 10 个氨基酸, 更一般至少为 12 个氨基酸, 经常至少为 14 个氨基酸, 更经常至少为 16 个氨基酸, 典型至少为 18 个氨基酸, 更典型至少为 20 个氨基酸, 通常至少为 22 个氨基酸, 更通常至少为 24 个氨基酸, 优选至少为 26 个氨基酸, 更优选至少为 28 个氨基酸, 而在特别优选的实施方案中, 至少为大约 30 个或更多的氨基酸。可以在合适长度的数段序列内, 将不同蛋白区段的序列与另一蛋白区段的序列比较。

通过使残基匹配最优化,必要时,根据需要引入间隙,确定氨基酸序列的同源性或序列同一性。参见例如 Needleham 等 (1970) J. Mol. Biol. 48:443-453; Sankoff 等, (1983) Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison 的第 1 章, Addison-Wesley, Reading, MA; 和 IntelliGenetics (Mountain View, CA) 的软件包; 和 Wisconsin Genetics Computer Group (GCG) (Madison, WI) 的软件包; 以上每个参考文献均通过引用结合到本文中。当将保守取代考虑为匹配时,会有变化。保守取代通常包括以下组内的取代: 甘氨酸、丙氨酸; 缬氨酸、异亮氨酸、亮氨酸; 天冬氨酸、谷氨酸; 天冬酰胺、谷氨酰胺; 丝氨酸、苏氨酸; 赖氨酸、精氨酸; 和苯丙氨酸、酪氨酸。同源的氨基酸序列计划包括该细胞因子序列中的天然等位基因变异和种间变异。典型的同源蛋白或肽与 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 的氨基酸序列区段的同源性将从 50-100% (如果可以引入间隙) 至 60-100% (如果包括保守取代)。同源性测量结果将至少为大约 70%, 一般至少为 76%, 更一般至少为 81%, 常常至少为 85%, 更常常至少为 88%, 典型至少为 90%, 更典型至少为 92%, 通常至少为 94%, 更通常至少为大约 95%, 优选至少为大约 96%, 而更优选至少为大约 97%, 而在特别优选的实施方案中, 至少为 98% 或更高。同源性的程度将随所比较的区段长度而变化。诸如等位基因变异体的同源蛋白或肽, 将与 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 中描述的实施方案共享大多数生物活性。在氨基酸水平或核苷酸水平上, 特别感兴趣的比较区域对应于区组 1-10 中每个区组内的那些区域、或区组内的区域, 所述区组对应于图 2A 中指定的区组。

本文所用的术语“生物活性”用来描述(但不限于)相应配体对炎症反应、先天免疫和/或形态发生发育的作用。例如, 这些受体应该同 IL-1 受体一样, 介导磷酸酶或磷酸化酶活性, 所述活性容易用标准方法测定。参见例如 Hardie 等(编辑 1995) The Protein Kinase FactBook 第

I 和 II 卷, Academic Press, San Diego, CA; Hanks 等 (1991) Meth. Enzymol. 200:38-62; Hunter 等 (1992) Cell 70:375-388; Lewin (1990) Cell 61:743-752; Pines 等 (1991) Cold Spring Harbor Symp. Quant. Biol. 56:449-463; 和 Parker 等 (1993) Nature 363:736-738。所述受体表现出同可调节酶非常相似的生物活性, 该活性通过配体结合而被调节。然而, 该酶的转换数比受体复合物更接近酶。此外, 诱导这种酶活性所必需的被占据的受体数低于大多数受体系统, 与在每个细胞数以千计的数值下触发的大多数受体相反, 被占据的受体数可能更接近每个细胞数打。所述受体或其部分作为标记通用底物或特定底物的磷酸标记酶, 可能是有用的。

术语例如 DTLR 的配体、激动剂、拮抗剂和类似物, 包括调节对 Toll 配体样蛋白特征性细胞应答的分子、以及加工配体-受体相互作用的更标准的结构结合竞争特征的分子, 例如在所述受体是天然受体或抗体的情况下。细胞应答可能通过不同的 Toll 配体与细胞受体结合而介导, 其中所述细胞受体与 I 型或 II 型 IL-1 受体相关, 但形式可能不同。参见例如 Belvin 和 Anderson (1996) Ann. Rev. Cell Dev. Biol. 12:393-416; Marisato 和 Anderson (1995) Ann. Rev. Genetics 29:371-3991 和 Hultmark (1994) Nature 367:116-117。

此外, 配体是这样一种分子, 它或者作为所述受体或其类似物结合的天然配体起作用, 或者作为所述天然配体的功能类似物的分子而起作用。所述功能类似物可以是具有结构修饰的配体, 或可以是完全不相关的分子, 该分子具有与所述合适的配体结合决定基相互作用的分子形状。所述配体可以用作激动剂或拮抗剂, 参见例如 Goodman 等 (编辑) (1990) Goodman & Gilman's: The Pharmacological Bases of Therapeutics, Pergamon Press, New York。

合理的药物设计也可以基于受体或抗体以及其它效应物或配体的分子形状的结构研究。效应物可以是对配体结合应答而介导其它功能的其它蛋白、或正常与该受体相互作用的其它蛋白。确定哪些位点

与特定的其它蛋白相互作用的方法是物理结构测定,例如X射线晶体学或2维NMR技术。这些将提供对于哪些氨基酸残基形成分子接触区的指南。关于蛋白结构测定的详细描述,参见例如Blundell和Johnson(1976) Protein Crystallography, Academic Press, New York, 该文献通过引用结合到本文中。

II. 活性

Toll样受体蛋白将具有许多不同的生物活性,例如在磷酸代谢中,将磷酸加入特定底物(通常为蛋白)或从特定底物去除磷酸。这类活性通常导致对炎症功能、其它先天的免疫应答或形态学效应的调节。DTLR2、3、4、5、6、7、8、9或10蛋白与其它Toll样受体蛋白同源,但它们每种均具有结构差异。例如,人DTLR2基因的编码序列可能与小鼠DTLR2的核苷酸编码序列具有大约70%的同一性。在氨基酸水平上,也可能有相当的同一性。

所述DTLR的生物活性与将磷酸部分添加至底物或从底物去除相关,通常以特异性方式相关,但有时以非特异性方式相关。可以鉴定底物,或可以用标准方法分析酶活性的条件,例如如以下文献中所述:Hardie等(编辑1995) The Protein Kinase FactBook 第I和II卷, Academic Press, San Diego, CA; Hanks等(1991) Meth. Enzymol. 200: 38-62; Hunter等(1992) Cell 70:375-388; Lewin(1990) Cell 61:743-752; Pines等(1991) Cold Spring Harbor Symp. Quant. Biol. 56:449-463; Parker等(1993) Nature 363:736-738。

III. 核酸

本发明设想了分离的核酸或片段的用途,例如它们编码这些蛋白或密切相关的蛋白、或它们的片段,例如以编码相应的多肽,最好是具有生物活性的多肽。另外,本发明包括分离或重组的DNA,所述DNA编码的这类蛋白或多肽,具有所述相应DTLR(单独的或作为一

组)的特征性序列。通常,该核酸能够在合适的条件下,与 SEQ ID NO: 3、5、25、9、11、15、17、31、21 或 33 中所示的核酸序列区段杂交,但最好与 SEQ ID NO: 1 的相应区段不杂交。所述生物活性蛋白或多肽可以是全长蛋白或片段,并且通常会具有其氨基酸序列与 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 中所示的氨基酸序列同源的区段。此外,本发明包括分离的或重组的核酸或其片段的用途,所述核酸编码的蛋白具有等同于 DTLR2-10 蛋白的片段。所述分离的核酸在 5'和 3'侧翼可以具有相应的调节序列,例如启动子、增强子、聚腺苷酸添加信号和来自所述天然基因的其它调节序列。

“分离的”核酸是大致纯的核酸,例如 RNA、DNA 或混合聚合物,它与天然伴随天然序列的其它组分分离,所述其它组分诸如核糖体、聚合酶和来自起源物种的侧翼基因组序列。该术语包括已经从其天然存在的环境中取出的核酸序列,包括:重组的或克隆的 DNA 分离物,它们因此可与天然存在的组分区别开来;和化学合成的类似物或由异源系统生物合成的类似物。大致纯的分子包括该分子的分离形式,或者为完全纯的,或者为大致纯的。

分离的核酸一般为分子的均质组合物,但在某些实施方案中,含有异质性,最好具有较少的异质性。通常在该聚合物末端或对于所需的生物功能或活性不重要的部分发现这种异质性。

“重组的”核酸通常根据生产方法或其结构而定义。对于其生产方法,例如用使用重组核酸技术的方法生产的产物,所述重组技术例如涉及在所述核苷酸序列中人的干涉。通常,这种干涉涉及体外操作,尽管在某些情况下,它可以涉及更经典的动物育种技术。或者,它可以是用以下方法制备的核酸:产生包含天然相互不相邻的两种片段的融合体的序列,但是指排除天然产物,例如在其天然状态中发现的天然存在的突变体。因此,例如,包括用任何非天然存在的载体转化细胞而制备的产物,也包括包含用任何合成寡核苷酸方法所衍生序列的核酸。通常进行这样一种方法,以用编码同一种或保守的氨基酸的丰

余密码子取代一个密码子，同时通常引入或去除一个限制性酶的序列识别位点。或者，进行该方法，以将所需功能的核酸区段连接在一起，以产生包括所需功能组合的单一的遗传实体，例如编码融合蛋白的遗传实体，其中所述功能组合是未在通常可得到的天然形式中发现的组合。限制性酶的识别位点常常是这类人工修饰的靶，但可以通过设计，加入其它位点特异性靶，例如启动子、DNA 复制位点、调节序列、控制序列或其它有用的特征。相似的概念将用于重组(例如融合)多肽。这将包括二体重复序列。具体包括：合成的核酸，它由于遗传密码的冗余性而编码相同的 DTLR2-10 的多肽或片段；和来自各种不同的相关分子(例如其它 IL-1 受体家族成员)的序列的融合体。

在核酸方面，“片段”是一连续的区段，它至少为大约 17 个核苷酸，一般至少为 21 个核苷酸，更一般至少为 25 个核苷酸，通常至少 30 个核苷酸，更通常至少为 35 个核苷酸，常常至少为 39 个核苷酸，更常常至少为 45 个核苷酸，典型至少为 50 个核苷酸，更典型至少为 55 个核苷酸，经常至少为 60 个核苷酸，更经常至少为 66 个核苷酸，更优选至少为 72 个核苷酸，更优选至少为 79 个核苷酸，在特别优选的实施方案中，将至少为 85 个或更多个核苷酸。通常，可以在合适长度的数段序列内，将不同遗传序列的片段与另一序列的片段进行比较，特别限定的区段诸如下述的域。

编码 DTLR2-10 的核酸将特别有用，以鉴定编码蛋白本身或其相关蛋白的基因、mRNA 和 cDNA 种类，以及鉴定编码来自不同个体或相关物种的多态变异体、等位基因变异体或其它遗传变异体。用于这类筛选的优选探针是白介素的那些区域，它们在不同多态变异体之间是保守的，或者它们含有缺乏特异性的核苷酸，并且它们最好是全长的或接近全长的。在其它情况下，多态变异体特异性序列将是更有用的。

本发明还包括重组核酸分子和片段，它们具有与本文提出的分离的 DNA 相同或高度同源的核酸序列。特别是，所述序列将常常与控制

转录、翻译和DNA复制的DNA区段操作性地连接。这些额外的区段通常有助于表达所需的核酸区段。

当相互比较或与SEQ ID NO: 3、5、25、9、11、15、17、31、21或33中所示的序列比较时，同源的、或高度相同的核酸序列表现出显著的相似性。核酸中的同源性标准是以下两个测量结果之一：本领域一般使用的通过序列比较得到的同源性，或基于杂交条件的同源性。以下更详细地描述比较性的杂交条件。

在核酸序列比较方面，大致同一性是指当对比时，或者所述区段或者其互补链，在插入或缺失合适的核苷酸进行最佳对比时，至少大约60%的核苷酸相同，一般至少为66%，经常至少为71%，常常至少为76%，更常常至少为80%，通常至少为84%，更通常至少为大约88%，典型至少为91%，更典型至少为大约93%，优选至少为大约95%，更优选至少为大约96-98%或更多，而在特定实施方案中，高达大约99%或更高的核苷酸相同，所述区段包括例如编码结构域的区段，诸如下述的区段。另一方面，当所述区段在选择性杂交条件下与一条链或其互补链杂交时，存在着大致相似性，一般使用由SEQ ID NO: 3、5、25、9、11、15、17、31、21或33衍生的序列。通常，当在一段至少大约14个核苷酸的序列中有大约55%同源性时，将发生选择性杂交，同源性更通常至少为大约65%，优选至少为大约75%，而更优选至少为大约90%。参见Kanehisa (1984) Nuc. Acids Res. 12:203-213，该文献通过引用结合到本文中。所述的同源性比较的长度如上所述，可以为更长的序列，在某些实施方案中为一段至少为大约17个核苷酸的序列，一般至少为大约20个核苷酸，通常至少为大约24个核苷酸，经常至少为大约28个核苷酸，典型至少为大约32个核苷酸，更典型至少为大约40个核苷酸，优选至少为大约50个核苷酸，而更优选至少为大约75-100或更多个核苷酸。

关于杂交方面的同源性，严格条件将是通常在杂交反应中控制的盐、温度、有机溶剂和其它参数的严格性的组合条件。严格性温度条

件通常包括超过大约30℃的温度，更通常为超过大约37℃的温度，典型的为超过大约45℃的温度，更典型的为超过大约55℃的温度，优选为超过大约65℃的温度，而更优选为超过大约70℃的温度。严格性的盐条件一般低于大约500 mM，通常低于大约400 mM，更通常低于大约300 mM，典型的低于大约200 mM，优选低于大约100 mM，而更优选低于大约80 mM，甚至降至低于大约20 mM。然而，参数的组合比任何单个参数的测量值更为重要。参见例如，Wetmur和Davidson (1968) J. Mol. Biol. 31:349-370，该文献通过引用结合到本文中。

另一方面，对于序列比较，通常一个序列用作参比序列，将测试序列与其比较。当使用一种序列比较算法时，将测试序列和参比序列输入计算机，必要时指定亚序列坐标，并且指定序列运算规则程序的参数。然后，该序列比较算法根据指定的程序参数，计算测试序列相对于参比序列的序列同一性百分率。

可以用如下进行用于比较的序列的光学序列对比：例如用 Smith 和 Waterman (1981) Adv. Appl. Math. 2:482 的局部同源性算法、用 Needleman 和 Wunsch (1970) J. Mol. Biol. 48:443 的同源性序列对比算法、通过 Pearson 和 Lipman (1988) Proc. Nat'l Acad. Sci. USA 85:2444 的相似性方法的检索、通过计算机实现这些算法(Wisconsin Genetics 软件包中的 GAP、BESTFIT、FASTA 和 TFASTA, Genetics Computer Group, 575 Science Dr., Madison, WI)、或通过目测检查(一般参见 Ausubel 等, 见上述)。

有用的算法的一个实例是 PILEUP。PILEUP 采用渐进的成对序列对比，产生来自一组相关序列的一个多重序列对比，以显示关系和序列同一性的百分率。它也绘出一个树(tree)或枝叉图(dendogram)，显示用来产生所述序列对比的成簇关系。PILEUP 采用 Feng 和 Doolittle (1987) J. Mol. Evol. 35:351-360 的渐进序列对比法的简化方法。所用的方法类似于 Higgins 和 Sharp (1989) CABIOS 5: 151-153 描述的方法。该程序可以对多达 300 个序列进行序列对比，每个序列的最大长度为

5,000 个核苷酸或氨基酸。多重序列对比法以两个最相似的序列的成对序列对比开始，产生一组(a cluster of)两个对比的序列。然后将该组与下一个最相关的序列或下一组对比的序列进行序列对比。通过简单地延伸两个单独序列的成对序列对比，将两组序列进行序列对比。通过一系列渐进的成对序列对比，得到最后的序列对比。通过指定序列比较区的特定序列及其氨基酸或核苷酸坐标，并且指定该程序的参数，运行该程序。例如，可以将参比序列与其它测试序列比较，以采用以下参数测定序列同一性关系的百分率：默认间隙加权值(3.00)、默认间隙长度加权值(0.10)和加权的末端间隙。

适用于测定序列同一性和序列相似性百分率的算法的另一实例是 BLAST 算法，在 Altschul 等 (1990) J. Mol. Biol. 215:403-410 中有描述。公众可从国家生物技术信息中心(<http://www.ncbi.nlm.nih.gov/>)得到进行 BLAST 分析的软件。该算法包括首先通过在查询序列中鉴别长度为 W 的短字(short words)，而鉴别高计分序列对(HSP)，当与一数据库序列中相同长度的字进行序列对比时，它们或者匹配或者满足某些阳性值的阈值 T。T 称为相邻字计分阈值(Altschul 等，见上述)。这些起始的相邻字命中(word hit)用作种子，起始搜寻以寻找含所述字命中的较长的 HSP。然后，只要可以增加累积的序列对比计分，则所述字命中在两个方向上沿每个序列延伸。在以下情况下，每个方向上所述字命中的延伸停止：由于来自其最大得到的数值的数量 X，累积的序列对比计分下降；由于一个或多个负计分的残基序列对比的累积，累积计分转到零或零以下；或达到两个序列之一的末端。BLAST 算法的参数 W、T 和 X 决定该序列对比的灵敏度和速度。BLAST 程序使用时，默认字长(W)为 11，BLOSUM62 计分矩阵(参见 Henikoff 和 Henikoff (1989) Proc. Nat'l Acad. Sci. USA 89:10915)序列(B)为 50，期望值(E)为 10，M=5，并且比较两条链。

除计算序列同一性百分率外，BLAST 算法也进行两个序列之间相似性的统计学分析(参见例如 Karlin 和 Altschul (1993) Proc. Nat'l Acad.

Sci. USA 90:5873-5787)。BLAST 算法提供的相似性的一种测量是最小和概率($P(N)$),这提供偶然发生两个核苷酸序列或氨基酸序列之间匹配概率的指标。例如,如果在测试核酸与参比核酸的比较中,最小和概率低于大约 0.1,更优选低于大约 0.01,最优选低于大约 0.001 时,认为该核酸与参比序列相似。

两个多肽的核酸序列大致相同的再一指标是:第一种核酸编码的多肽与第二种核酸编码多肽免疫交叉反应,如以下所述。因此,例如在一种多肽与第二种多肽仅由于保守取代而不同时,这两种多肽通常大致相同。两个核酸序列大致相同的另一指标是:所述两种分子在严格条件下相互杂交,如下所述。

通过核苷酸置换、核苷酸缺失、核苷酸插入和核苷酸序列倒位,可以容易地修饰分离的 DNA。这些修饰产生新的 DNA 序列,它们编码该蛋白或其衍生物。这些修饰的序列可以用来产生突变的蛋白(突变蛋白),或增强变异体种类的表达。表达增强可以包括基因扩增、转录增加、翻译增加和其它机制。这类突变的 DTLR 样衍生物,包括该蛋白或其片段的预定突变或定点突变,包括利用遗传密码简并性的沉默突变。本文所用的“突变的 DTLR”包括这样的多肽,它在其它情况下属于上述 DTLR 同源性定义范围内,但其氨基酸序列无论由于缺失、置换,还是由于插入,均不同于自然界中发现的其它 DTLR 样蛋白的氨基酸序列。特别是,“定点突变 DTLR”包括与 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 的蛋白大致同源的蛋白,并通常共享大多数本文公开形式的生物活性或作用。

尽管定点突变位点是预定的,但突变体不必是定点的。通过在该基因中制造氨基酸插入或缺失,可以达到对哺乳动物 DTLR 诱变,并将其与表达偶联。可以产生置换、缺失、插入或任何组合,以得到最终的构成物。插入包括氨基末端融合或羧基末端融合。可以于靶密码子进行随机诱变,然后可以根据所需活性,筛选表达的哺乳动物 DTLR 突变体。在序列已知的 DNA 中于预定位点制造置换突变的方法是本领

域众所周知的, 例如通过 M13 引物诱变。也参见 Sambrook 等 (1989) 和 Ausubel 等 (1987 和定期增刊)。

所述 DNA 中的突变应该不将编码序列置于读框之外, 最好不产生可以杂交而产生诸如环或发夹的 mRNA 二级结构的互补区。

Beaucage 和 Carruthers (1981) Tetra. Letts 22:1859-1862 描述的亚磷酸胺法将产生合适的合成 DNA 片段。常常或者通过合成互补链, 并在合适的条件下使所述链一起退火, 或通过加入互补链, 用 DNA 聚合酶与合适的引物序列, 获得双链片段。

聚合酶链式反应(PCR)技术通常可以用于诱变。另一方面, 诱变引物是通常用来于预定位点产生限定的突变的方法。参见例如 Innis 等 (编辑 1990) PCR Protocols: A Guide to Methods and Applications Academic Press, San Diego, CA; 和 Dieffenbach 和 Dveksler (1995; 编辑) PCR Primer: A Laboratory Manual Cold Spring Harbor Press, CSH, NY。

IV. 蛋白质、肽

如上所述, 本发明包括灵长类 DTLR2-10, 例如其序列公开于 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 和上述的灵长类 DTLR2-10。也设想了等位基因变异体和其它变异体, 包括例如将这类序列与其它序列的部分组合的融合蛋白, 所述其它序列包括表位标记和功能域。

本发明也提供重组蛋白, 例如采用来自这些啮齿动物蛋白区段的异源融合蛋白。异源融合蛋白是一种蛋白或区段的融合体, 所述蛋白或区段在天然状态下通常不以同一方式融合。因此, DTLR 与 IL-1 受体的融合产物是一种连续的蛋白分子, 它具有以典型的肽键融合的序列, 通常作为单一的翻译产物制备, 并表现出例如得自每种源肽的序列或抗原性的特性。相似的概念适用于异源核酸序列。

另外, 通过组合来自其它相关蛋白的相似功能域或结构域, 可以制备新的构成物, 其中所述其它相关蛋白例如为 IL-1 受体或其它 DTLR, 包括种变异体。例如, 在不同的新融合多肽或片段之间, 可以

“交换”配体结合区段或其它区段。参见例如 Cunningham 等 (1989) Science 243:1330-1336; 和 O'Dowd 等 (1988) J. Biol. Chem. 263: 15985-15992, 每个文献通过引用结合到本文中。因此, 由于受体结合特异性的功能的结合, 将产生表现出新的特异性组合的新嵌合多肽。例如, 可以加入来自其它相关受体分子的配体结合域, 或用其取代该蛋白或相关蛋白的其它域。所产生的蛋白通常会具有杂种功能或特性。例如, 融合蛋白可以包含一个引导域, 它可以用以将该融合蛋白结合至特定的亚细胞细胞器。

可以从各种序列数据库中选择候选的融合伴侣和序列, 所述数据库例如 GenBank, c/o IntelliGenetics, Mountain View, CA; 和 BCG, University of Wisconsin Biotechnology Computing Group, Madison, WI, 它们通过引用结合到本文中。

本发明特别提供结合 Toll 配体和/或影响信号转导的突变蛋白。人 DTLR1-10 与 IL-1 家族其它成员的结构对比, 显示出保守的特征/残基。参见例如图 3A。人 DTLR 序列与 IL-1 家族其它成员的序列对比, 显示不同的结构和功能共享特征。也参见 Bazan 等(1996) Nature 379:591; Lodi 等(1994) Science 263:1762-1766; Sayle 和 Milner-White (1995) TIBS 20:374-376; 和 Gronenberg 等(1991) Protein Engineering 4:263-269。

IL-1 α 和 IL-1 β 配体结合作为第一受体的 IL-1 受体 I 型, 然后该复合物与 IL-1 受体 III 型形成一种高亲和性的受体复合物。这类受体亚基可能与所述新 IL-1 家族成员共享。

DTLR2-10 序列的其它物种相应物中的相似变异(例如在相应的区域中), 应该提供与配体或底物的相似相互作用。特别优选或者小鼠序列或人序列的取代。相反, 离开配体结合相互作用区的保守取代将可能保留大多数信号活性。

灵长类 DTLR2-10 的“衍生物”, 包括氨基酸序列的突变体、糖基化变异体、代谢衍生物以及与其它化学部分的共价缀合物或聚集缀

合物。通过例如采用本领域熟知的方法，将官能团(functionalities)与所述 DTLR 氨基酸侧链中或于 N 末端或 C 末端发现的基团连接，可以制备共价缀合物。这些衍生物可以包括但不限于：羧基末端的脂族酯或酰胺、或含羧基侧链的残基、含羟基残基的 O-酰基衍生物、和氨基末端氨基酸或含氨基残基(例如赖氨酸或精氨酸)的 N-酰基衍生物。酰基选自烷基部分，包括 C3-C18 的正烷基，由此形成烷酰基芳酰基种类。

特别是，包括糖基化改变，例如通过多肽合成和加工期间，或在进一步加工步骤中，修饰该多肽的糖基化模式而制备的糖基化改变。完成这一步的特别优选的方法是：将多肽暴露于糖基化酶，例如哺乳动物糖基化酶，所述酶得自正常提供这种加工的细胞。也设想了去糖基化酶。也包括一级氨基酸序列相同的形式，它们具有其它少量修饰，包括磷酸化的氨基酸残基，例如磷酸酪氨酸、磷酸丝氨酸或磷酸苏氨酸。

主要一组衍生物是所述受体或其片段与多肽的其它蛋白的共价缀合物。这些衍生物可以在重组培养物中合成，诸如为 N 末端融合体或 C 末端融合体，通过利用本领域已知的试剂，根据它们在通过反应性侧基使蛋白交联的有用性，合成这些衍生物。优选用交联剂衍生的位点是游离氨基、碳水化合物部分和半胱氨酸残基。

也提供所述受体和其它同源或异源蛋白之间的融合多肽。同源多肽可以是不同受体之间的融合体，产生例如表现出多种不同 Toll 配体的结合特异性的杂种蛋白，或产生可以扩大或减弱底物效应特异性的受体。同样，可以构建表现出所述衍生蛋白特性或活性组合的异源融合体。典型的实例是报道多肽(例如荧光素酶)与受体区段或域(例如配体结合区段)的融合体，使得可以容易地确定所需配体的存在或位置。参见例如 Dull 等的美国专利第 4,859,609 号，它通过引用结合到本文中。其它基因融合伴侣包括谷胱甘肽-S-转移酶(GST)、细菌 β -半乳糖苷酶、trpE、A 蛋白、 β -内酰胺酶、 α 淀粉酶、醇脱氢酶和酵母 α 交配因子。参见例如 Godowski 等(1988) *Science* 241:812-816。

Beaucage 和 Carruthers (1981) Tetra. Letts. 22:1859-1862 描述的亚磷酸胺法, 将产生合适的合成 DNA 片段。常常或者通过合成互补链, 并在合适的条件下使所述链一起退火, 或通过使用 DNA 聚合酶与合适的引物序列加入互补链, 获得双链片段。

这类多肽也可以具有已经进行化学修饰的氨基酸残基, 所述化学修饰通过磷酸化、磺化、生物素化、或添加或去除其它部分(特别是分子形状类似于磷酸基团的那些部分)而进行。在某些实施方案中, 所述修饰将是有用的标记试剂, 或用作纯化的靶, 例如亲和配体。

通常通过或者重组核酸方法, 或者通过合成肽方法, 制备融合蛋白。一般例如在 Sambrook 等 (1989) Molecular Cloning: A Laboratory Manual (第二版), 第 1-3 卷, Cold Spring Harbor Laboratory, 和 Ausubel 等(编辑 1987 和定期增刊) Current Protocols in Molecular Biology, Greene/Wiley, New York 中描述了核酸操作和表达的技术, 所述文献通过引用结合到本文中。例如在 Merrifield (1963) J. Amer. Chem. Soc. 85:2149-2156; Merrifield (1986) Science 232: 341-347; 和 Atherton 等 (1989) Solid Phase Peptide Synthesis: A Practical Approach, IRL Press, Oxford 中描述了多肽合成的技术, 每个文献通过引用结合到本文中。也参见 Dawson 等(1994) Science 266:776-779 关于制备较大多肽的方法。

本发明也设想了非氨基酸序列变异或糖基化变异的 DTLR2-10 衍生物的运用。这类衍生物可以包括与化学部分的共价结合或聚集结合。这些衍生物一般分为三类: (1)盐, (2)侧链和末端残基共价修饰, 和(3)吸附复合体, 例如与细胞膜的吸附复合体。这类共价或聚集衍生物可用作免疫原、免疫测定中的试剂, 或用于纯化方法中, 诸如用于受体或其它结合分子(例如抗体)的亲和纯化。例如, 可以采用本领域熟知的方法, 通过将 Toll 配体共价结合至固相支持体(诸如溴化氰活化的 Sepharose), 将 Toll 配体固定化, 或将其吸附至聚烯烃表面上, 并且进行或不进行戊二醛交联, 用于 DTLR 受体、抗体或其它相似分子

的测定或纯化。该配体也可以用可检测基团标记，例如用氯胺 T 法放射碘化、共价结合至稀土螯合物、或缀合至用于诊断测定的另一荧光部分。

本发明的 DTLR 可以用作免疫原，用于生产对所述 DTLR 或其各种片段特异性的抗血清或抗体，例如它们能够区别其它 IL-1 受体家族成员。所述纯化的 DTLR 可以用来筛选单克隆抗体或抗原结合片段，所述抗体或片段是通过用各种形式的含该蛋白的不纯制剂免疫而制备的。具体地说，术语“抗体”也包括天然抗体的抗原结合片段，例如 Fab、Fab2、Fv 等等。所述纯化的 DTLR 也可以用作试剂，来检测对存在水平提高的表达应答而产生的抗体，或检测导致针对该内源受体的抗体产生的免疫失调。另外，DTLR 片段也可以用作免疫原，以生产本发明的抗体，如下面紧接描述的。例如，本发明设想了这样的抗体或其片段或各种同源肽，所述抗体对 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 中所示的氨基酸序列具有结合亲和性，或针对其而产生的。特别是，本发明设想了对特定片段具有结合亲和性或针对其产生的抗体，所述特定片段预期或实际上暴露于所述天然 DTLR 的外部蛋白表面上。

由于可能通过竞争抑制，抑制该配体与所述受体的结合，可以阻断对所述受体配体的生理反应。因此，本发明的体外测定常常会采用抗体或这些抗体的抗原结合区段，或连接至固相基质的片段。这些测定也提供对或者配体结合区突变和修饰、或者其它突变和修饰(例如影响信号功能或酶功能的突变和修饰)的效应的诊断测定。

本发明也设想了竞争性药物筛选测定的用途，例如在针对所述受体或片段的中和抗体与测试化合物竞争配体或其它抗体结合的情况下。以该方式，可以用所述中和抗体或片段检测下述多肽的存在，所述多肽享有结合受体的一个或多个结合位点，并且也可以用来占据受体上的结合位点，其中所述结合位点在其它情况下可能结合配体。

V. 制备核酸和蛋白

可以通过化学合成、筛选 DNA 文库、或通过筛选由多种细胞系或组织样品制备的基因组文库，获得编码所述蛋白或其片段的 DNA。可以采用标准方法和本文提供的序列，分离天然序列。通过杂交技术，或通过各种 PCR 技术，结合在例如 GenBank 的序列数据库进行检索，或通过数据库检索，可以鉴定其它物种的相应物。

该 DNA 可以在种类繁多的宿主细胞中表达，用以合成全长受体或片段，所述全长受体或片段可以进而例如用来产生多克隆抗体或单克隆抗体；用于结合研究；用于构建和表达修饰的配体结合域或激酶/磷酸酶域；和用于结构/功能研究。可以在用合适的表达载体转化或转染的宿主细胞中，表达变异体或片段。这些分子可以大致不含蛋白质或细胞污染物，不是得自重组宿主中的那些分子，因此当结合药用可接受的载体和/或稀释剂时，在药用组合物中特别有用。所述蛋白或其部分可以作为与其它蛋白的融合体表达。

表达载体通常是自主复制的 DNA 或 RNA 构成物，它们含有所需的受体基因或其片段，通常与在合适宿主细胞中识别的合适的遗传控制元件操作性地连接。这些控制元件能够在合适的宿主中实现表达。实现表达所必需的特定类型的控制元件，将取决于最终使用的宿主细胞。一般而言，所述遗传控制元件可以包括原核启动子系统或真核启动子表达控制系统，通常包括一个转录启动子、一个控制转录启动的可选的操纵子、提高 mRNA 表达水平的转录增强子、一段编码合适的核糖体结合位点的序列、以及终止转录和翻译的序列。表达载体通常也含有一个复制起点，使得该载体独立于宿主细胞而复制。

本发明的载体包括这样的载体，它们含有编码所述蛋白的 DNA 或其编码生物活性相当的多肽的片段。所述 DNA 可以处于病毒启动子的控制之下，并且可以编码一个选择标记。本发明还设想了这类表达载体的用途，当所述载体与所述宿主匹配，并且将编码该受体的真核 cDNA 插入该载体中，使得含所述载体的宿主的生长表达所述 cDNA

时, 所述表达载体能够在原核或真核宿主中表达编码这种蛋白的真核 cDNA。通常, 设计表达载体, 用于在其宿主细胞中稳定复制, 或用于扩增, 以大大提高每个细胞所需基因的总拷贝数。不必总是需要表达载体在宿主细胞中复制, 例如, 可以采用不含有被所述宿主细胞识别的复制起点的载体, 实现该蛋白或其片段在各种宿主中瞬时表达。使用通过重组引起该蛋白编码部分或其片段整合入宿主 DNA 中的载体, 这也可能。

本文所用的载体包括质粒、病毒、噬菌体、可整合 DNA 片段和能够将 DNA 片段整合入该宿主基因组的其它载体。表达载体是特化的载体, 它们含有实现表达操作性连接的基因的遗传控制元件。质粒是最常用的载体形式, 但本领域已知的、或成为已知的、起作相当用的所有其它形式的载体均适用于本文。参见例如 Pouewls 等 (1985 和增刊) Cloing Vectors: A Laboratory Manual, Elsevier, N.Y. 和 Rodriguez 等 (编辑) Vectors: A Survey of Molecular Cloning Vectors and Their Uses, Butterworth, Boston, 1988, 所述文献通过引用结合到本文中。

转化的细胞是已经用采用重组 DNA 技术构建的受体载体转化或转染的细胞, 最好是哺乳动物细胞。转化的宿主细胞通常表达所需的蛋白或其片段, 但对于克隆、扩增和操作其 DNA 的目的, 不需要表达题述蛋白。本发明还设想了在营养培养基中培养转化的细胞, 由此允许该受体在细胞膜中累积。可以或者从该培养物中, 或者在某些情况下从培养基中, 回收该蛋白。

对于本发明的目的, 当核酸序列在功能上相互相关时, 将核酸序列操作性地连接。例如, 如果前序列或分泌前导序列的 DNA 作为蛋白原表达, 或参与将多肽导向至细胞膜或参与该多肽的分泌, 则将该 DNA 与所述多肽操作性地连接。如果启动子控制该多肽的分泌, 则将该 DNA 与多肽操作性地连接。假如一种启动子控制该多肽的转录, 则将该启动子与编码序列操作性地连接; 如果将一个核糖体结合位点定位以允许翻译, 则将其与编码序列操作性地连接。通常, 操作性连接

的是指连续的，并且在读框内，然而，诸如抑制基因的某些遗传元件是不连续连接的，但仍结合于操纵子序列，操纵子序列进而控制表达。

合适的宿主细胞包括原核细胞、低等真核细胞和高等真核细胞。原核细胞包括革兰氏阴性和革兰氏阳性生物，例如大肠杆菌和枯草杆菌(*B. subtilis*)。低等真核细胞包括酵母，例如酿酒酵母(*S. cerevisiae*)和毕赤酵母属(*Pichia*)；以及网柄菌属(*Dictyostelium*)物种。高等真核细胞包括由动物细胞建立的组织培养细胞系，所述动物细胞包括非哺乳动物来源，例如昆虫细胞和鸟类，也包括哺乳动物来源，例如人类、灵长类和啮齿动物。

原核宿主-载体系统包括种类繁多的用于许多不同物种的载体。本文所用的大肠杆菌及其载体在种属上用来包括用于其它原核细胞的相当的载体。用来扩增 DNA 的代表性载体是 pBR322 或许多其衍生物。可以用来表达该受体或其片段的载体，包括但不限于诸如这样的载体，它们含有 lac 启动子(pUC 系列)；trp 启动子(pBR322-trp)；Ipp 启动子(pIN 系列)； λ -pP 或 pR 启动子(pOTS)；或杂种启动子，诸如 ptac (pDR540)。参见 Brosius 等 (1988) “利用 λ -、trp-、lac-和 Ipp 衍生的启动子的表达载体”，载于 Vectors: A Survey of Molecular Cloning Vectors and Their Uses, (编辑, Rodriguez 和 Denhardt), Butterworth, Boston, 第 10 章, 第 205-236 页, 该文献通过引用结合到本文中。

低等真核细胞，例如酵母和网柄菌属，可以用含 DTLR 序列的载体转化。对于本发明的目的，最常用的低等真核宿主是面包酵母，即酿酒酵母。它将用来在种属上代表低等真核细胞，尽管可利用许多其它菌种和物种。酵母载体通常由一个复制起点(整合型除外)、一个选择基因、一个启动子、编码该受体或其片段的 DNA、以及用于翻译终止、聚腺苷酸化和转录终止的序列。合适的酵母表达载体包括组成型启动子，诸如 3-磷酸甘油激酶和各种其它糖酵解酶基因的启动子；或诱导型启动子，诸如醇脱氢酶 2 启动子或金属硫蛋白启动子。合适的载体包括以下类型的衍生物：自主复制低拷贝数型(诸如 YRp 系列)、自主复制

高拷贝数型(诸如YEp系列); 整合型(诸如YIp系列), 或微型染色体(诸如YCp系列)。

高等真核生物组织培养细胞一般是用于表达功能活性的白介素蛋白的优选宿主细胞。原则上, 可以使用任何高等真核生物组织培养细胞系, 例如昆虫杆状病毒表达系统, 不论其来自无脊椎动物源还是来自脊椎动物源。然而, 最好是哺乳动物细胞。这种细胞的转化或转染和繁殖已经成为常规技术。有用的细胞系的实例包括海拉细胞、中国仓鼠卵巢(CHO)细胞系、幼鼠肾(BRK)细胞系、昆虫细胞系、鸟细胞系和猴(COS)细胞系。用于这些细胞系的表达载体一般包括一个复制起点、一个启动子、一个翻译起始位点、RNA剪接位点(如果使用基因组DNA)、一个多腺苷酸化位点和一个转录终止位点。这些载体也通常含有一个选择基因或扩增基因。合适的表达载体可以是携带启动子的质粒、病毒或反转录病毒, 所述启动子例如得自诸如腺病毒、SV40、细小病毒、痘苗病毒或巨细胞病毒这类来源。合适的表达载体的代表性实例包括pCDNA1; pCD, 参见Okayama等(1985) Mol. Cell Biol. 5:1136-1142; pMC1neo PolyA, 参见Thomas等(1987) Cell 51:503-512; 和诸如pAC 373或pAC 610的杆状病毒载体。

对于分泌蛋白, 一个可读框通常编码一种多肽, 它由于其N末端共价连接于信号肽的成熟的或分泌的产物组成。在该成熟的、或有活性的多肽分泌之前, 该信号肽被切割。根据经验性的规律(例如 von-Heijne (1986) Nucleic Acids Research 14:4683-4690), 可以以高度的精确性预测该切割位点, 信号肽精确的氨基酸组成看来对其功能不重要, 例如 Randall 等(1989) Science 243:1156-1159; Kaiser 等(1987) Science 253:312-317。

通常需要在提供特定或限定的糖基化模式的系统中, 表达这些多肽。在这种情况下, 通常的模式是由该表达系统天然提供的。然而, 将通过将该多肽(例如非糖基化形式)暴露于引入异源表达系统中的合适的糖基化蛋白, 来修饰该模式。例如, 该受体基因可以与一种或多

种编码哺乳动物糖基化酶或其它糖基化酶的基因共转化。采用该方法,可在原核细胞或其它细胞中得到某些哺乳动物糖基化模式。

DTLR 源可以是表达重组 DTLR 的真核或原核宿主,诸如以上描述的。该源也可以是诸如小鼠 Swiss 3T3 成纤维细胞的细胞系,但本发明也设想了其它哺乳动物细胞系,优选的细胞系来自人的种类。

因为所述序列是已知的,所以可以通过合成肽的常规方法,制备灵长类 DTLR、其片段或衍生物。这些包括诸如在 Stewart 和 Young (1984) Solid Phase Peptide Synthesis, Pierce Chemical Co., Rockford, IL; Bodanszky 和 Bodanszky (1984) The Practice of Peptide Synthesis, Springer-Verlag, New York; 和 Bodanszky (1984) The Principles of Peptide Synthesis, Springer-Verlag, New York 中所述的方法,所有这些文献均通过引用结合到本文中。例如,可以使用叠氮化物法、盐酸法、酸酐法、混合酐法、活性酯法(例如对硝基苯酯、N-羧基琥珀酰亚胺酯或氰基甲酯)、碳化二咪唑法、氧化还原法、或二环己基碳二亚胺(DCCD)/加成法。固相和液相合成都适用于前述方法。可以使用相似的技术与部分 DTLR 序列。

按照通常用于肽合成的以上方法,适当地制备所述 DTLR 蛋白、片段或衍生物,一般或者用所谓的分段法,该方法包括按顺序一个接一个地将氨基酸缩合至末端氨基酸,或者通过将肽片段偶联至末端氨基酸。通常必须保护在偶联反应中不使用的氨基,以防止于不正确的位置进行偶联。

如果采用固相合成,则将 C 末端氨基酸通过其羧基结合至不溶性载体或支持体。不特别限制所述不溶性载体,只要它对反应性羧基具有结合能力即可。这类不溶性载体的实例包括:卤代甲基树脂(诸如氯代甲基树脂或溴代甲基树脂)、羟甲基树脂、酚树脂、叔烷氧羰基酰肼化(hydrazidated)树脂等等。

通过将氨基保护的氨基酸的活化羧基与先前形成的肽或链的反应性氨基缩合,将氨基保护的氨基酸按顺序结合,以逐步合成所述肽。

合成所述完整的序列后,将该肽切离所述不溶性载体,以产生该肽。
Merrifield 等(1963) *J. Am. Chem. Soc.* 85:2149-2156 中一般描述了这种固相方法,该文献通过引用结合到本文中。

通过肽分离,例如通过抽提、沉淀、电泳、各种形式的层析等等,可以从反应混合物中分离和纯化所制备的蛋白或其片段。根据所需的用途,本发明的受体可以以不同程度的纯度获得。利用本文公开的蛋白纯化技术(参见下文),或利用本文在免疫吸附亲和层析方法中描述的抗体,完成纯化。通过首先将所述抗体连接至固相支持体,然后使连接的抗体与合适细胞的溶解的裂解液、表达所述受体的其它细胞的裂解液或通过 DNA 技术产生该蛋白的细胞上清液接触,进行这种免疫吸附亲和层析,见下述。

一般而言,该纯化蛋白的纯度至少为大约 40%,一般至少为大约 50%,通常至少为大约 60%,典型的至少为大约 70%,更典型的至少为大约 80%,优选至少为大约 90%,更优选至少为大约 95%,而在特别优选的实施方案中,为 97%-99%或更高。纯度通常基于重量,但也可以基于摩尔。将适当地应用不同的测定。

VI. 抗体

可以针对不同哺乳动物(例如灵长类)的 DTLR 蛋白及其片段产生抗体,所述 DTLR 蛋白及其片段包括天然存在的天然形式和其重组形式,不同的是,针对该活性受体的抗体更可能识别仅存在于天然构象中的表位。变性抗原的检测也可以用于例如蛋白质分析。也设想了抗独特型抗体,它可以用作天然受体或抗体的激动剂或拮抗剂。

通过用蛋白预定片段与免疫原性蛋白的缀合物免疫动物,可以产生针对所述片段的抗体,包括结合片段和单链形式。由分泌所需抗体的细胞制备单克隆抗体。可以根据与正常蛋白或缺陷蛋白的结合,或根据激动剂或拮抗剂活性,筛选这些抗体。这些单克隆抗体结合的 K_D 值通常至少为大约 1 mM,更通常至少为大约 300 μ M,典型的至少为

大约 100 μM ，更典型的至少为大约 30 μM ，优选至少为 10 μM ，更优选至少为 3 μM 。

包括抗原结合片段在内的本发明的抗体，可以具有显著的诊断或治疗价值。它们可以是有效的拮抗剂，它们结合于所述受体，并抑制与配体的结合，或抑制所述受体引发生物学反应(例如作用于其底物)的能力。它们也可以用作非中和性抗体，并可以偶联于毒素或放射性核素，以结合生产细胞或定位为白介素源的细胞。此外，这些抗体可以与药物或其它治疗剂缀合，或直接缀合，或者通过接头缀合。

本发明的抗体也可以用于诊断应用。作为捕捉物或非中和性抗体，它们可以结合至所述受体，而不抑制配体或底物的结合。作为中和抗体，它们可以用于竞争性结合测定。它们也可以用来检测配体或对配体定量。它们可以用作蛋白质印迹分析的试剂，或用于相应蛋白的免疫沉淀或免疫纯化。

可以将蛋白片段连接至其它材料，特别是多肽，如待用作免疫原的融合的或共价连接的多肽。哺乳动物 DTLR 及其片段可以与多种免疫原融合或共价连接，所述免疫原诸如匙孔 血蓝蛋白、牛血清白蛋白、破伤风类毒素等。关于制备多克隆抗血清的方法的描述，参见 Microbiology, Hoeber Medical Division, Harper 和 Row, 1969; Landsteiner (1962) Specificity of Serological Reactions, Dover Publications, New York; 和 Williams 等(1967) Methods in Immunology and Immunochemistry, 第 1 卷, Academic Press, New York; 每个参考文献均通过引用结合到本文中。典型的方法包括用免抗原超免疫动物。然后在重复免疫后不久，收集该动物的血液，并分离 γ 球蛋白。

在某些情况下，最好从各种哺乳动物宿主制备单克隆抗体，所述哺乳动物宿主诸如小鼠、啮齿动物、灵长类、人类等等。在例如以下文献中可以发现制备这类单克隆抗体的技术的描述：Stites 等(编辑) Basic and Clinical Immunology (第 4 版), Lange Medical Publications, Los Altos, CA 和其中引用的参考文献；Harlow 和 Lane (1988) Antibodies: A

Laboratory Manual, CSH Press; Goding (1986) Monoclonal Antibodies: Principles and Practice (第2版) Academic Press, New York; 以及特别是 Kohler 和 Milstein (1975) Nature 256:495-497 (它讨论了产生单克隆抗体的一种方法)。这些参考文献中的每个均通过引用结合到本文中。总之, 该方法包括给动物注射一种免疫原。然后处死该动物, 取出其脾脏细胞, 然后将脾细胞与骨髓瘤细胞融合。结果是杂种细胞或“杂交瘤”, 它能够在体外繁殖。然后筛选杂交瘤群体, 以分离各个克隆, 每个克隆分泌针对该免疫原的单种抗体。以该方式, 获得的各个抗体种类是来自免疫动物的无限增殖化的和克隆的单种 B 细胞, 是对该免疫原性物质上识别的特定位点应答而产生的 B 细胞。

其它合适的技术包括, 使淋巴细胞体外暴露于该抗原多肽或者暴露于在噬菌体或相似载体中的抗体文库的选择物。参见 Huse 等(1989) “免疫球蛋白所有组成成分在噬菌体 λ 中的大组合文库的产生”, Science 246: 1275-1281; 和 Ward 等(1989) Nature 341:544-546, 每个参考文献均通过引用结合到本文中。本发明的多肽和抗体可以修饰或不修饰使用, 包括嵌合抗体和人源化抗体。通常, 通过或者共价或者非共价连接一种提供可检测信号的物质, 来标记所述多肽和抗体。种类繁多的标记和缀合技术是已知的, 并且在科学文献和专利文献中均广泛地进行了报道。合适的标记包括放射性核素、酶、底物、辅因子、抑制剂、荧光部分、化学发光部分、磁性颗粒等等。指出这类标记用法的专利, 包括美国专利第 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; 和 4,366,241 号。此外, 可以产生重组或嵌合的免疫球蛋白, 参见 Cabilly 的美国专利第 4,816,567 号; 或在转基因小鼠中制备, 参见 Mendez 等(1997) Nature Genetics 15:146-156。这些参考文献均通过引用结合到本文中。

本发明的抗体也可以用于分离所述 DTLR 的亲亲和层析。当所述抗体连接至固相支持体(例如颗粒, 诸如琼脂糖、Sephadex 等)时, 可以制备柱子, 在可以使细胞裂解液通过该柱的情况下, 洗涤该柱, 然后

提高温和变性剂的浓度，由此释放所纯化的蛋白。该蛋白可以用来纯化抗体。

所述抗体也可以用来筛选特定表达产物的表达文库。通常，用于这种方法中的抗体用一个部分标记，所述部分通过抗体结合，允许容易地检测抗原的存在。

针对 DTLR 产生的抗体也可以用来产生抗独特型抗体。这些抗体将可用于检测或诊断与该蛋白表达或表达该蛋白的细胞相关的各种免疫疾病。它们也可用作所述配体的激动剂或拮抗剂，它们可以是天然存在的配体的竞争性抑制剂或替代物。

通常在免疫测定中，确定与针对特定免疫原产生的抗体特异性地结合或特异性地免疫反应的 DTLR 蛋白，所述特定免疫原诸如由 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 的氨基酸序列组成的免疫原。所述免疫测定通常使用多克隆抗血清，它是针对例如 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 的蛋白产生的。选择对其它 IL-1R 家族成员(例如 DTLR1)具有低交叉反应性的抗血清，所述抗血清最好来自相同的物种，并且在用于该免疫测定之前，通过免疫吸附除去任何这种交叉反应性。

为了产生用于免疫测定的抗血清，按本文所述分离 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 的蛋白或其组合的蛋白。例如，可以在哺乳动物细胞系中产生重组蛋白。通常使用标准佐剂(诸如弗氏佐剂)和标准的小鼠免疫方案(参见 Harlow 和 Lane, 见上述)，用选定的蛋白免疫合适的宿主，例如小鼠的近交品系，诸如 balb/c。或者，由本文公开的序列衍生、且与载体蛋白缀合的合成肽，可以用作免疫原。收集多克隆抗血清，并且在免疫测定中测定针对该免疫原蛋白的效价，例如在固相免疫测定中，用固定至固相支持体的免疫原测定。选择效价为 10^4 或更高的多克隆抗血清，并用竞争性结合免疫测定，测试其对其它 IL-1R 家族成员(例如小鼠 DTLR 或人 DTLR1)的交叉反应性，所述竞争性免疫测定诸如 Harlow 和 Lane (见上述，第

570-573 页)所述的免疫测定。最好是,在该测定中,结合人 DTLR2-10 中任一种或其中某些 DTLR,使用至少两个 DTLR 家族成员。这些 IL-1R 家族成员可以作为重组蛋白产生,并用本文所述的标准分子生物学和蛋白化学技术进行分离。

竞争性结合形式的免疫测定,可以用来测定交叉反应性。例如,可以将 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 的蛋白或其各种片段固定于固相支持体。加入该测定中的蛋白,与所述抗血清竞争与所述固定化抗原的结合。将上述蛋白与所述抗血清竞争与所述固定化抗原结合的能力,与 SEQ ID NO: 4、6、26、10、12、16、18、32、22 和/或 34 的蛋白进行比较。采用标准计算方法,计算上述蛋白的交叉反应性的百分率。选择并合并与每种以上所列蛋白的交叉反应性低于 10% 的抗血清。然后,通过用以上列出的蛋白进行免疫吸附,从合并的抗血清中,除去交叉反应性抗体。

然后,将免疫吸附的合并的抗血清用于如上所述的竞争性结合免疫测定中,以将第二种蛋白与该免疫原蛋白(例如 SEQ ID NO: 4、6、26、10、12、16、18、32、22 和/或 34 的 IL-1R 样蛋白)进行比较。为了进行这种比较,每次在很宽的浓度范围内分析所述两种蛋白,并测定抑制该抗血清与该固定化抗原结合的 50% 所需的每种蛋白的量。如果第二种蛋白的所需量低于选定蛋白或所需蛋白的蛋白量的 2 倍,则认为第二种蛋白特异性地结合于针对该免疫原产生的抗体。

不用说,这些 DTLR 蛋白是一个同源蛋白家族的成员,该家族包括迄今鉴定的至少 10 个基因。关于特定基因的产物,诸如 DTLR2-10,该术语不仅是指本文公开的氨基酸序列,而且是指为等位基因变异体、非等位基因变异体或种变异体的其它蛋白。人们也理解,所述术语包括通过以下方法引入的非天然突变,所述方法为采用常规的重组技术(诸如单位点突变)的有意突变,或切下编码相应蛋白的 DNA 的短片段,或取代新的氨基酸或添加新的氨基酸。这类少量改变必须大体维持原始分子的免疫同一性和/或其生物活性。因此,这些改变包括与

指定的天然存在的 IL-1R 相关蛋白特异性免疫反应的蛋白, 所述相关蛋白例如为 SEQ ID NO: 4、6、26、10、12、16、18、32、22 或 34 中所示的 DTLR 蛋白。可以通过在合适的细胞系中表达所述改变的蛋白, 并测量对淋巴细胞的合适作用, 可以确定该蛋白的生物学特性。认为修饰较少的具体蛋白修饰, 包括化学性质相似的氨基酸的保守取代, 如以上关于 IL-1R 大体上所述的。通过将一种蛋白与 DTLR2-10 的蛋白进行最佳序列对比, 并采用本文所述的常规免疫测定测定免疫同一性, 人们可以确定本发明的蛋白组成。

VII. 试剂盒和定量测定

天然存在形式和重组形式的本发明 IL-1R 样分子在试剂盒和测定方法中均特别有用。例如, 这些方法也可以用来根据例如配体对这些蛋白的结合活性进行筛选。最近数年, 已经开发了几种自动化测定方法, 以便允许每年选数万种化合物。参见例如 BIOMEK 自动化工作站, Beckman Instruments, Palo Alto, California, 和 Fodor 等 (1991) *Science* 251:767-773, 所述文献通过引用结合到本文中。后一篇文献描述了用固相基质上合成的多种特定聚合物测试结合的方法。大量纯化的活性状态的可溶性 DTLR (诸如本发明提供的) 的可得性, 可以大大促进开发用来筛选配体或激动剂/拮抗剂同源蛋白的合适测定。

可以将纯化的 DTLR 直接包被于平板上, 用于在上述配体筛选技术中。然而, 针对这些蛋白的非中和抗体, 可以用作捕捉抗体, 以将相应受体固定至固相上, 例如可用于诊断用途。

本发明也设想了 DTLR2-10、其片段、肽和它们的融合产物的用途, 即用于检测该蛋白或其配体存在的多种诊断试剂盒和方法中。或者, 或另外, 可以将针对所述分子的抗体加入所述试剂盒和方法中。通常, 该试剂盒将具有一个区室, 该区室含有或者特定的 DTLR 肽或基因区段, 或者含有识别一种或另一种肽或基因的试剂。通常, 在肽的情况下, 识别试剂是一种受体或抗体, 或在基因区段的情况下, 识

别试剂通常是杂交探针。

用于测定例如 DTLR4 的一种样品浓度的优选试剂盒，通常包含一种标记化合物，例如对 DTLR4 具有已知结合亲和性的配体或抗体、作为阳性对照的一种 DTLR4 源(天然存在的或重组的)和一种分离结合的与游离的标记化合物的一种工具，例如用于固定该测试样品中的 DTLR4 的固相。通常提供含有试剂的区室和说明。

对哺乳动物 DTLR 或肽片段、或受体片段特异性的抗体，包括抗原结合区段，可用于诊断应用中，以检测水平提高的配体/或其片段的存在。诊断测定可以是均相的(在游离试剂和抗体-抗原复合物之间没有分离步骤)或多相的(有分离步骤)。存在各种商业化测定，诸如放射免疫测定(RIA)、酶联免疫吸附测定(ELISA)、酶免疫测定(EIA)、酶倍增(enzyme-multiplied)免疫测定技术(EMIT)、底物标记荧光免疫测定(SLFIA)等等。例如，通过采用被标记并识别抗 DTLR 或抗其特定片段的抗体的第二抗体，可以利用未标记的抗体。这些测定也已经在参考文献中进行了广泛地讨论。参见例如 Harlow 和 Lane (1988) Antibodies: A Laboratory Manual, CSH., 和 Coligan (编辑) (1991)和定期增刊, Current Protocols In Immunology Greene/Wiley, New York.

抗独特型抗体可以具有相似的用途，作为 DTLR4 的激动剂或拮抗剂起作用。这些抗体应该在合适的情况下用作治疗剂。

通常，该试剂盒中提供用于诊断测定的试剂，以优化该测定的灵敏度。对于本发明，根据该测定的性质，提供所述方案和所述标记、或者标记或者未标记的抗体、或标记的配体。这通常结合其它添加剂，诸如缓冲液、稳定剂、用于产生信号所需的物质，诸如酶的底物等等。最好是，该试剂盒也含有正确使用内容物和使用后处理内容物的说明。该试剂盒通常具有用于每种有用试剂的区室，并且将含有正确使用和处理试剂的说明。希望的是，当可以在水性介质中重建具有合适浓度的所述试剂，以进行该测定时，所述试剂作为干的冻干粉提供。

可以不加修饰地使用诊断测定的上述组分，或可以对其以多种方

式进行修饰。例如,通过共价或非共价连接直接或间接提供可检测信号的部分,可以实现标记。在这些测定的任何一种中,可以或者直接、或者间接地标记测试化合物、DTLR 或其抗体。可能用于直接标记的包括标记基团:诸如 ^{125}I 的放射标记、诸如过氧化物酶和碱性磷酸酶的酶(美国专利第 3,645,090 号)、和能够监测荧光强度变化、波长偏移或荧光偏振的荧光标记(美国专利第 3,940,475 号)。这两个专利均通过引用结合到本文中。可能用于间接标记的包括将一种组分生物素化,然后与偶联至一个上述标记基团的抗生物素蛋白结合。

也有许多从游离配体中分离结合配体、或者从游离的测试化合物中分离结合化合物的方法。可以在各种基质上固定所述 DTLR,然后洗涤。合适的基质包括诸如 ELISA 板的塑料、滤膜和珠粒。将所述受体固定化至基质上的方法包括但不限于,直接附着到塑料上、使用捕捉抗体、化学偶联和生物素-抗生物素蛋白。该方案中的最后一步包括通过使用几种方法中的任何一个,沉淀抗体/抗原复合物,所述方法包括那些例如使用诸如聚乙二醇的有机溶剂或诸如硫酸铵的盐的方法。其它合适的分离技术,包括但不限于在 Rattle 等(1984) *Clin. Chem.* 30(9):1457-1461 中描述的荧光素抗体可磁化微粒法,和在美国专利第 4,659,678 号中描述的双抗体磁性微粒分离,所述文献通过引用结合到本文中。

在所述文献中广泛报告了用于将蛋白或片段同各种标记连接的方法,本文并不需要详细讨论。所述技术中的许多技术涉及使用活化羧基,或者通过使用碳二亚胺或者通过使用活性酯形成肽键;涉及通过巯基与诸如氯乙酰的活化卤素或诸如马来酰亚胺的活化烯烃反应,形成硫醚来用于连接等等。在这些应用中也发现使用融合蛋白。

本发明另一诊断方面,包括使用取自 DTLR 序列的寡核苷酸或多核苷酸序列。可以使用这些序列作为探针,以检测怀疑患有免疫失调的患者中相应 DTLR 水平。RNA 和 DNA 二者核苷酸序列的制备、所述序列的标记和所述序列的优选大小,已经在所述文献中得到了充分描述

和讨论。一般寡核苷酸探针应当具有至少大约14个核苷酸，通常至少大约18个核苷酸，而所述多核苷酸探针最多可以达到几千个碱基。可以使用各种标记，最常用的是放射性核素，特别是 ^{32}P 。然而，也可以使用其它技术，诸如使用生物素修饰的用于引入多核苷酸中的核苷酸。然后，该生物素用作结合抗生物素蛋白或抗体的位点，所述抗生物素蛋白或抗体可以用各种各样的标记物标记，诸如放射性核素、荧光物质(fluorescer)、酶等等。另一方面，可以使用可识别特定双链体的抗体，所述双链体包括DNA双链体、RNA双链体、DNA-RNA杂种双链体或DNA-蛋白双链体。可以进而标记所述抗体，并进行所述测定，其中所述双链体结合于一表面，使得根据在所述表面上双链体的形成，可以检测结合于所述双链体的抗体的存在。使用任何常规技术，诸如核酸杂交、加减筛选、重组探针探测、杂交分子释放翻译(HRT)和杂交分子终止翻译(HART)，可以实现对新型反义RNA的探针的应用。这也包括诸如聚合酶链式反应(PCR)的扩增技术。

也考虑了诊断试剂盒，所述试剂盒也定性或定量测试其它标记的存在。诊断或预后可以依靠用作标记的多种指示的组合。因此，试剂盒可以测试标记物的组合。参见例如Viallet等(1989) Progress in Growth Factor Res. 1:89-97。

VIII. 治疗用途

本发明提供具有重大治疗价值的试剂。所述DTLR (天然存在的或重组的)、其片段、突变蛋白受体和抗体以及鉴定为对所述受体或抗体具有结合亲和性的化合物，应该可用于治疗表现出其配体的受体表达异常的病征。这类异常通常表现为免疫失调。另外，本发明应该在与该配体表达异常或对该配体应答的触发异常相关的各种疾病或病征方面，提供治疗价值。已经提示，所述Toll配体参与形态发育(例如背腹极性的确定)和免疫应答，特别是原始的先天应答。参见例如Sun等(1991) Eur. J. Biochem. 196:247-254; Hultmark (1994) Nature 367:116-

117.

可以纯化重组DTLR、突变蛋白、其激动剂或拮抗剂抗体、或抗体，然后将其给予患者。这些试剂可以例如在常规药用可接受的载体或稀释剂以及生理上无害的稳定剂和赋形剂中，与另外的活性组分结合用于治疗用途。可以例如无菌过滤这些组合物，并通过在剂量瓶中冻干将其以剂量形式放置，或以稳定化的水性制剂贮藏。本发明也考虑了使用非补体结合的抗体或其结合片段。

可以进行采用DTLR或其片段的配体筛选，以鉴别与所述受体具有结合亲和性的分子。然后，可以利用后续的生物测定，确定推定的配体是否可以提供可阻断内在刺激活性的竞争性结合。受体片段可以用作阻断剂或拮抗剂，因为它阻断配体的活性。同样，具有内在刺激活性的化合物可以激活所述受体，并因此是一种激动剂，因为它刺激配体的活性，例如诱导信号。本发明还考虑了抗DTLR抗体作为拮抗剂的治疗用途。

用于有效治疗所必需的试剂量取决于许多不同因素，包括给药方法、靶位点、患者的生理状态和给予的其它药物。因此应当确定治疗剂量，以最优化安全性和有效性。一般在体外使用的剂量可以为这些试剂的原位给药用量提供有用的指南。用于治疗特定疾病的有效剂量的动物测试，将对人的剂量提供进一步的预测指示。例如在Gilman等(编辑)(1990) Goodman and Gilman's: The Pharmacological Bases of Therapeutics (第8版) Pergamon Press; 和 Remington's Pharmaceutical Sciences (最新版) Mack Publishing Co., Easton, Penn. 中描述了各种考虑，所述每个文献通过引用结合到本文中。在其中和下文讨论了给药方法，例如经口、静脉内、腹膜内、或肌肉内给药、经皮扩散和其它方法。药用可接受的载体包括水、盐、缓冲液和例如在 Merck Index (Merck & Co., Rahway, New Jersey) 中所述的其它化合物。因为推定的配体与其受体之间可能高亲和性结合或转换数高，所以最初预期低剂量的这些试剂是有效的。并且信号途径提示，极低量的配体可能具有效应。

因此，一般预期具有合适载体的剂量范围低于1mM的浓度，通常低于大约10 μ M的浓度，一般低于大约100 nM，优选低于大约10 pM(皮摩尔浓度)，最优选低于大约1 fM(费摩尔浓度)。对于连续给药，经常使用缓释配方或缓释设备。

可以直接向待治疗的宿主给予 DTLR、其片段以及抗体或其片段、拮抗剂和激动剂，或者根据所述化合物的大小，可能最好在给药前将所述化合物同载体蛋白缀合，所述载体蛋白诸如卵清蛋白或血清白蛋白。可以以任何常规给药制剂给予治疗制剂。尽管可能独自给予所述活性成分，但最好是作为药用制剂将其提呈。制剂包含至少一种如上文定义的活性成分，连同一种或多种可接受的载体。每种载体在与其它成分相容、且不伤害患者的意义上，必须既是药用可接受的，也是生理上可接受的。制剂包括那些适于经口、直肠、鼻或肠胃外(包括皮下、肌内、静脉在内和皮内)给药的制剂。所述制剂通常可以以单位剂量形式存在，并可以通过药剂学领域熟知的任何方法制备。参见例如 Gilman 等(编辑)(1990) Goodman and Gilman's: The Pharmacological Bases of Therapeutics 第 8 版 Pergamon Press；和 Remington's Pharmaceutical Sciences (最新版) Mack Publishing Co., Easton, Penn.；Avis 等(编辑, 1993) Pharmaceutical Dosage Forms: Parenteral Medications Dekker, NY；Lieberman 等(编辑, 1990) Pharmaceutical Dosage Forms: Tablets Dekker, NY；和 Lieberman 等(编辑, 1990) Pharmaceutical Dosage Forms: Disperse Systems Dekker, NY。本发明的疗法可以和其它治疗剂组合，或与其它治疗剂结合使用，所述治疗剂特别是其它 IL-1 家族成员的激动剂或拮抗剂。

IX. 配体

本文对所述 Toll 受体的描述，提供了鉴定配体的方法，如上所述。这种配体应该以适当的高亲和性，特异性地结合于相应的受体。可得到各种构成物，它们允许对所述受体进行任何一种标记，以检测其配体。例如，直接标记 DTLR，将其融合至用于第二标记的标记物(例如

FLAG 或其它表位标记等), 将允许检测受体。这在组织学上可以作为生化纯化的亲和方法、或在表达克隆方法中的标记或选择。也可以应用双杂种选择系统, 制备具有可利用的 DTLR 序列的合适构成物。参见例如 Fields 和 Song (1989) Nature 340:245-246。

一般而言, 对 DTLR 的描述类似地适用于针对 DTLR2、DTLR3、DTLR4、DTLR5、DTLR6、DTLR7、DTLR8、DTLR9 和/或 DTLR10 试剂和组合物的各个具体实施方案。

参考以下实施例, 可最好地理解本发明的广泛范围, 所述实施例不会将本发明限于所述具体的实施方案。

实施例

I. 一般方法

例如在以下文献中描述和引用了某些标准方法: Maniatis等(1982) 分子克隆, 实验室指南, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, NY; Sambrook等(1989) 分子克隆: 实验室指南 (第2版) 第1-3卷, CSH Press, NY; Ausubel等, Biology, Greene Publishing Associates, Brooklin, NY; 或 Ausubel等(1987和增补本) Current Protocols in Molecular Biology, Wiley/Greene, NY。用于蛋白纯化的方法, 包括诸如硫酸铵沉淀、柱层析、电泳、离心、结晶的方法和其它方法。参见例如 Ausubel等(1987和定期增补本); Coligan等(编辑, 1996) 和定期增补本, Current Protocols In Protein Science, Wiley/Greene, NY; Deutscher (1990) Methods in Enzymology 第182卷和在这个系列中的其它卷中的“蛋白质纯化指南”; 和生产商关于蛋白纯化产品用法的文献, 例如 Pharmacia, Piscataway, NJ 或 Bio-Rad, Richmond, CA。与重组技术结合, 允许同合适的区段融合, 例如同 FLAG 序列或可以通过蛋白酶可去除序列融合的相当序列融合。参见例如 Hochuli (1989) Chemische Industrie 12:69-70; Hochuli (1990) “用金属螯合物吸收剂纯化重组蛋白”, 载于 Setlow (编辑) Genetic Engineering, Principle and

Methods 12:87-98, Plenum Press, NY; 和Crowe等(1992) OIAexpress: The High Level Expression & Protein Purification System QUIAGEN, Inc., Chatsworth, CA.

例如在Hertzenberg等 (编辑, 1996) Weir's Handbook of Experimental Immunology 第1-4卷, Blackwell Science; Coligan (1991) Current Protocols in Immunology Wiley/Greene, NY; 和Methods in Enzymology 第70、73、74、84、92、93、108、116、121、132、150、162和163卷中描述了标准免疫学技术和测定。

对血管生物活性的测定是本领域熟知的。它们将包括在肿瘤或其它组织中的血管生成活性和血管抑制活性, 例如动脉平滑肌增生(参见例如Koyoma等(1996) Cell 87:1069-1078)、单核细胞附着于血管上皮(参见McEvoy等 (1997) J. Exp. Med. 185:2069-2077)等。也参见Ross (1993) Nature 362:801-809; Rekhter和Gordon (1995) Am. J. Pathol. 147:668-677; Thyberg等(1990) Atherosclerosis 10:966-990; 和Gumbiner (1996) Cell 84:345-357。

例如在Wouterlood (1995年编辑) Neuroscience Protocols 第10单元, Elsevier; Methods in Neurosciences Academic Press; 和Neuromethods Humana Press, Totowa, NJ中描述了对神经细胞生物活性的测定。例如在Meisami (编辑) Handbook of Human Growth and Developmental Biology CRC Press; 和Chrispeels (编辑) Molecular Techniques and Approaches in Developmental Biology Interscience中描述了发育系统方法学。

例如采用可得到的软件程序, 包括GCG (U. Wisconsin)和GenBank来源的那些程序, 进行计算机序列分析。也使用例如来自GenBank、NCBI、EMBO和其它数据库的公开的序列数据库。

对于所有目的, 适用于IL-10受体的许多技术可以用于DTLR, 例如在USSN 08/110,683 (IL-10受体)中所述, 该文献通过引用结合到本文中。

II. 人受体的新家族

缩写: DTLR, Toll样受体; IL-1R, 白介素-1受体; TH, Toll同源性; LRR, 富含亮氨酸的重复序列; EST, 表达的序列标记; STS, 序列标志位点; FISH, 原位杂交中的荧光。

在果蝇属Toll和人白介素-1的胞质域之间发现序列同源性,已经传播了这两种分子触发依赖于Rel型转录因子核转位的相关信号途径的信念。在昆虫和脊椎动物中,这种保守的信号模式支配进化上古老的免疫应答。我们报道了一类新的推定的人受体的分子克隆,所述人受体在胞内和胞外区段中具有与果蝇属Toll非常相似的蛋白结构。命名为DTLR 1-5的5种人Toll样受体可能是蝇分子的直接同系物,同样可能组成人先天免疫的重要的和未识别的一种组分;有趣的是,在脊椎动物中DTLR在进化上保留,可能表明在果蝇属胚胎背腹化中类似于Toll的另一种作用,即作为早期形态发生图式发育的调节物。多种组织的mRNA印迹,显示人DTLR显著不同的表达模式。采用原位杂交中的荧光和序列标志位点数据库分析,我们也显示了位于第4 (DTLR 1、2和3)、9 (DTLR4)和1 (DTLR5)染色体上的同系DTLR基因。对来自以下的序列对比的Toll同源(TH)域的结构预测,识别出一个具有酸性活性位点的平行 β/α 折叠:不同的昆虫和人的DTLR、脊椎动物IL-1受体和MyD88因子以及植物抗病性蛋白;在细菌中广泛参与转导感觉信息的一类应答调节物中,明显重现相似的结构。

如此显著地区分蝇与人的形态发生鸿沟的种子植于熟悉的胚胎形状和模式之中,但产生非常不同的细胞复杂度。DeRobertis 和 Sasai (1996) Nature 380:37-40; 和 Arendt 和 Nulbler-Jung (1997) Mech. Develop. 61:7-21。显著相似的信号途径设计了昆虫和脊椎动物之间的这种发育计划的趋异,强调了来自不同基因所有组分成分的蛋白网络和生化机制更大的保守性。Miklos 和 Rubin (1996) Cell 86:521-529; 和 Chothia (1994) Develop. 1994 增刊, 27-33。绘出这些调节途径进化设计

的一个有功效的方法,是通过蛋白序列和结构的种间比较推断出它们可能的分子组分(和生物学功能)。Miklos 和 Rubin (1996) Cell 86:521-529; Chothia (1994) Develop. 1994 增刊, 27-33 (3-5); 和 Banfi 等(1996) Nature Genet. 13:167-174。

胚胎发育中的一个关键的通用步骤是体轴(body axes)的特化,所述特化或者来自先天的不对称,或者由外部信号触发。DeRobertis 和 Sasai (1996) Nature 380:37-40; 和 Arendt 和 Nulbler-Jung (1997) Mech. Develop. 61:7-21。作为一个模型系统,已经将特殊的关注集中于系统发育基础和背腹极化的细胞机制。DeRobertis 和 Sasai (1996) Nature 380:37-40; 和 Arendt 和 Nulbler-Jung (1997) Mech. Develop. 61: 7-21。由果蝇属胚胎形成了这种转化的一个原型分子策略,其中少数基因的顺序作用,产生转录因子 Dorsal 的腹侧梯度。St. Johnston 和 Nusslein Volhard (1992) Cell 68:201-209; 和 Morisato 和 Anderson (1995) Ann. Rev. Genet. 29:371-399。

该信号途径以跨膜受体 Toll 为中心, Toll 将母体分泌的腹侧因子 Spatzle 转导结合到辅助分子 Tube 的胞质接合(cytoplasmic engagement)中并转导 Ser/Thr 激酶 Pelle 的激活, Pelle 催化 Dorsal 与抑制剂 Cactus 的解离,并允许 Dorsal 转移至腹侧核(Morisato 和 Anderson (1995) Ann. Rev. Genet. 29:371-399; 和 Belvin 和 Anderson (1996) Ann. Rev. Cell Develop. Biol. 12:393-416。所述 Toll 途径也控制成体蝇中有效抗微生物因子的诱导(Lemaitre 等(1996) Cell 86:973-983); 在果蝇属免疫防御中的该作用加强了与 IL-1 途径平行的机械途径,它们在脊椎动物中支配免疫应答和炎症反应。Belvin 和 Anderson (1996) Ann. Rev. Cell Develop. Biol. 12:393-416; 和 Wasserman (1993) Molec. Biol. Cell 4:767-771。IL-1 受体中的 Toll 相关胞质域,指导 Pelle 样激酶 IRAK 的结合和反映出含有 Dorsal 和 Cactus 的潜伏 NF/ κ B/I- κ B 复合物的激活。Belvin 和 Anderson (1996) Ann. Rev. Cell Develop. Biol. 12:393-416; 和 Wasserman (1993) Molec. Biol. Cell 4:767-771。

我们描述了人类 4 种新 Toll 样分子的克隆和分子鉴定, 所述 4 种分子命名为 DTLR 2-5 (根据 Chiang 和 Beachy (1994) Mech. Develop. 47:225-239), 这揭示了一个与果蝇属 Toll 同系物的关系比与脊椎动物 IL-1 受体的关系更密切的受体家族。所述 DTLR 序列得自人 EST; 这些部分 cDNA 用来绘出人组织中 5 种 DTLR 的完整的表达分布型, 对同源基因的染色体位置进行作图, 并缩减用于全长 cDNA 挽回的 cDNA 文库的选择。我们被其它努力(Banfi 等(1996) Nature Genet. 13:167-174; 和 Wang 等(1996) J. Biol. Chem. 271:4468-4476)所激励, 我们根据结构的保守和分子简约性, 装配了一个人类的生物学系统, 它是果蝇属中令人瞩目的调节方案的对应方案。此外, 提出的由 DTLR、广泛的 IL-1 受体家族、哺乳动物 MyD88 因子和植物抗病性蛋白共享的一个核心组件 - Toll 同源 (TH) 域的三级折叠, 提示驱动 Toll 信号的生物化学机制。Mitcham 等(1996) J. Biol. Chem. 271:5777-5783; 和 Hardiman 等(1996) Oncogene 13:2467-2475。我们提出: 在昆虫、植物和动物中将形态发生和原始免疫偶联的信号途径(Belvin 和 Anderson (1996) Ann. Rev. Cell Develop. Biol. 12:393-416; 和 Wilson 等(1997) Curr. Biol. 7:175-178)可能来源于细菌的双组分途径。

计算机分析。

采用 BLAST 服务器(Altschul 等(1994) Nature Genet. 6:119-129), 从国家生物技术信息中心(NCBI)的 EST 数据库(dbEST), 鉴定出与昆虫 DTLR 相关的人类序列。使用更敏感的基于模式和分布型的方法(Bork 和 Gibson (1996) Meth. Enzymol. 266:162-184), 分离出所述 DTLR 家族中与非冗余数据库中存在的脊椎动物和植物蛋白共享的信号域。用 ClustalW (Thompson, 等(1994) Nucleic Acids Res. 22: 4673-4680), 进行 DTLR 胞内域和胞外域序列的渐进序列对比; 该程序也通过相邻连接法(5000 自引导程序的复制提供树归组的可信度值), 计算序列对比的序列的分支顺序。

用 Consensus 程序(internet URL <http://www.bork.embl-heidelberg.de/Alignment/consensus.html>), 绘出在几个严格性程度上分辨的保守的序列对比模式。用灵活地匹配趋异 LRR 的 N 末端和 C 末端特征的一个化合物基序(PRINTS 密码 Leurichrpt), 蛋白指纹的 RPINTS 文库 (<http://www.biochem.ucl.ac.uk/bsm/dbbrowser/PRINTS/PRINTS.html>) (Attwood 等(1997) Nucleic Acids Res. 25:212-217)可靠地鉴别出 DTLR 胞外区段中存在的无数富含亮氨酸的重复序列(LRR)。三态精度 (three-state accuracy)高于 72%的两种预测算法, 用来得出胞内域序列对比的共有二级结构, 作为折叠识别努力的桥梁(Fischer 等(1996) EASEB.L 10:126-136)。神经网络程序 PHD (Rost 和 Sander (1994) Proteins 19:55-72)和统计学预测方法 DSC (King 和 Sternberg (1996) Protein Sci. 5:2298-2310)有 internet 服务器(分别为 URLs http://www.embl-heidelberg.de/predictprotein/phd_pred.html 和 http://bonsai.lif.icnet.uk/bmm/dsc/dsc_read_align.html)。胞内区编码例如在 Hardiman 等(1996) Oncogene 13:2467-2475; 和 Rock 等(1998) Proc Nat'l Acad. Sci. USA 95:588-593 (每个文献通过引用结合到本文中)中讨论的 THD 区。该域在受体的信号机制中非常重要, 它将磷酸基团转移至底物。

全长人 DTLR cDNA 的克隆

用得自 Toll 样 Humrsc786 序列(GenBank 登记密码 D13637) (Normura 等(1994) DNA Res 1:27-35)的 PCR 引物作为探针, 探测人红白血病 TF-1 细胞系衍生的 cDNA 文库(Kitamura 等(1989) Blood 73:375-380), 产生 DTLR1 cDNA 序列。通过 Research Genetics (Huntsville, AL), 从 dbEST 和由 I.M.A.G.E.聚生体获得的相关的 EST 克隆(Lennon 等(1996) Genomics 33:151-152)中, 标记其余的 DTLR 序列: CloneID# 80633 和 117262 (DTLR2)、144675 (DTLR3)、202057

(DTLR4)和 227229 (DTLR5)。通过 DNA 杂交分别筛选 λ gt10 噬菌体、成人肺脏、胎盘和胎肝 5'-Stretch Plus cDNA 文库(Clontech), 克隆了人 DTLR 2-4 的全长 cDNA; 所述 DTLR5 序列得自人多发性硬化斑的 EST。对所有阳性克隆进行测序和序列对比, 以鉴定各个 DTLR ORF: DTLR1 (2366 bp 的克隆, 786 aa ORF)、DTLR2 (2600 bp, 784 aa)、DTLR3 (3029 bp, 904 aa)、DTLR4 (3811 bp, 879 aa)和 DTLR5 (1275 bp, 370 aa)。分别用人胎盘(Stratagene)和成人肝(Clontech)的 cDNA 文库作为模板, 通过 PCR 产生用于 DTLR3 和 DTLR4 杂交的探针; 引物对得自相应的 ESR 序列。用栖热水生菌(*T. aquaticus*)的 Taqplus DNA 聚合酶(Stratagene), 在以下条件下进行 PCR 反应: $1 \times (94^{\circ}\text{C}, 2 \text{ 分钟})$ 、 $30 \times (55^{\circ}\text{C}, 20 \text{ 秒}; 72^{\circ}\text{C} 30 \text{ 秒}; 94^{\circ}\text{C} 20 \text{ 秒})$ 、 $1 \times (72^{\circ}\text{C}, 8 \text{ 分钟})$ 。对于 DTLR2 全长 cDNA 的筛选, 通过 EcoRI/XbaI 消化第一个 EST 克隆(ID# 80633)产生的 900 bp 片段用作探针。

mRNA 印迹和染色体定位

每泳道含有大约 $2 \mu\text{g poly(A)}^+$ RNA 的人多种组织印迹(Cat#1、2)和癌细胞系印迹(Cat# 7757-1), 购自 Clontech (Palo Alto, CA)。对于 DTLR 1-4, 将分离的全长 cDNA 用作探针; 对于 DTLR5, 使用 EST 克隆(ID #277229)质粒插入片段。简而言之, 采用 Amersham Rediprime 随机标记试剂盒(RPN1633), 用 $[\alpha\text{-}^{32}\text{P}] \text{ dATP}$ 放射标记所述探针。于 65°C , 在 $0.5 \text{ M Na}_2\text{HPO}_4$ 、7% SDS、 $0.5 \text{ M EDTA (pH 8.0)}$ 中进行预杂交和杂交。于 65°C 下进行所有严格性洗涤, 最初 2 次洗涤在 $2 \times \text{SSC}$ 、0.1% SDS 中进行 40 分钟, 然后在 $0.1 \times \text{SSC}$ 、0.1% SDS 中洗涤 20 分钟。然后, 在增感屏存在下, 于 -70°C 将膜对 X 光片(Kodak)曝光。用选定的人 DTLR 克隆通过 cDNA 文库 Southern (14)进行的更详细的研究, 以检查它们在造血细胞亚群中的表达。

采用 Heng 和 Tsui (1994) Meth. Molec. Biol. 33:109-122 中描述的原位杂交中的荧光(FISH)方法, 用各种全长(DTLR 2-4)或部分(DTLR5)

cDNA 克隆作为探针, 进行人染色体作图。这些分析作为一种服务, 由 SeeDNA Biotech Inc. (Ontario, 加拿大) 进行。通过 internet 服务器 (http://www.hgmp.mrc.ac.uk/DHMHD/hum_chromel.html), 在畸形人-小鼠同源性数据库中, 进行人综合征(或同线基因座中的小鼠缺陷)的检索。

昆虫和人 DTLR 胞外域的保守结构。

果蝇属中的 Toll 家族至少包括 4 种不同的基因产物: 参与蝇胚胎背腹图式发育的原型受体 Toll (Morisato 和 Anderson (1995) Ann. Rev. Genet. 29:371-399) 以及第二种称为 “18 Wheeler” (18w) 的产物, 后者也可能参与早期的胚胎发育 (Chiang 和 Beachy (1994) Mech. Develop. 47:225-239; 和 Eldon 等 (1994) Develop. 120:885-899); 两种另外的受体根据雄性特异性转录物 (Mst) 基因座 (GenBank 密码 X67703) 下游的不完整的 Toll 样 ORF 来预测, 或由 “序列标志位点” (STS) Dm2245 (GenBank 密码 G01378) 编码 (Mitscham 等 (1996) J. Biol. Chem. 271:5777-5783)。Toll 和 18w 的胞外区段与众不同地由不完全的约 24 个氨基酸的 LRR 基元组成 (Chiang 和 Beachy (1994) Mech. Develop. 47:225-239; 和 Eldon 等 (1994) Develop. 120:885-899)。相似的 LRR 串联排列通常形成不同细胞表面分子的粘附天线, 推定它们的属三级结构, 以模拟一种核糖核酸酶抑制剂折叠的马蹄形支架, 其中 17 种 LRR 显示出重复的 β/α 发夹, 这是一种 28 个残基的基元 (Buchanan 和 Gay (1996) Prog. Biophys. Molec. Biol. 65:1-44)。Toll 对 Spatzle 的特异性地识别可能遵循一种模型, 该模型是用弯曲 β -片的凹面, 为蛇受体的多 LRR 胞外域对胱氨酸结折叠糖蛋白激素的结合而提出的 (Kajava 等 (1995) Structure 3:867-877); 有趣的是, 在 Spatzle 和果蝇属孤独配体 Trunk 中的半胱氨酸模式, 预测了一个相似的胱氨酸结三级结构 (Belvin 和 Anderson (1996) Ann. Rev. Cell & Devel. Biol. 12:393-416; 和 Casanova 等 (1995) Genes Develop. 9:2539-2544)。

根据序列和模式分析, Toll 和 18w 分别为 22 个和 31 个 LRR 的胞外域(Mst ORF 片段呈现出 16 个 LRR)与 DTLR 1-4 (目前不完整的 DTLR5 链包括 4 个膜近接的(membrane-proximal) LRR)相当的 18 个、19 个、24 个和 22 个 LRR 排列最密切相关(Altschul 等(1994) Nature Genet. 6:119-129; 和 Bork 和 Gibson (1996) Meth. Enzymol. 266:162-184) (图 1)。然而, 人 DTLR 链中显著的差异是通常丧失约 90 个残基的半胱氨酸富含区, 该区可变地包含于 Toll、18w 和 Mst ORF 中(分别距膜边界 4 个、6 个和 2 个 LRR)。这些半胱氨酸簇是等分的, 具有明显的“顶部”(结束一个 LRR)和“底部”(堆叠在一个 LRR 顶部)两半(Chiang 和 Beachy (1994) Mech. Develop. 47:225-239; Eldon 等(1994) Develop. 120:885-899; 以及 Buchanan 和 Gay (1996) Prog. Biophys. Molec. Biol. 65:1-44); 在果蝇属和人 DTLR 中均作为保守的近膜间隔区重现所述“顶部”组件(图 1)。我们提出, 当“顶部”与“底部”接合时, 果蝇属受体(和其它 LRR 蛋白)中的可变定位的半胱氨酸簇, 形成一个完整的组件, 它具有匹配末端, 可以插入任何 LRR 对之间, 而不改变 DTLR 胞外域的整体折叠; 类似的“挤出”域装饰其它蛋白的结构(Russell (1994) Protein Engin. 7:1407-1410)。

TH 信号域的分子设计

Toll 和 IL-1 I 型(IL-1R1)受体的序列比较, 已经揭示了一个约略相似的约 200 个氨基酸的胞质域, 推测它由相似的 Rel 型转录因子来介导信号(Belvin 和 Anderson (1996) Ann. Rev. Cell Develop. Biol. 12:393-416; 和 Wasserman (1993) Molec. Biol. Cell 4:767-771)。新近加入该功能性范例的包括来自烟草和亚麻的一对植物抗病性蛋白, 它们的特征为一个 N 末端 TH 组件, 后接核苷酸结合(NTP 酶)区段和 LRR 区段(Wilson 等(1997) Curr. Biol. 7:175-178); 相反, 一个“死域”在 MyD88 的 TH 链之前, MyD88 是一种胞内骨髓分化标记(Mitcham 等(1997) J. Biol. Chem. 271:5777-5783; 和 Hardiman 等(1996) Oncogene 13:2467-

2475)(图 1)。新的 IL-1 型受体包括一种辅助信号分子 IL-1R3、和孤独受体 IL-1R4 (也称为 ST2/Fit-1/T1)、IL-1R5 (IL-1R 相关蛋白)和 IL-1R6 (IL-1R 相关蛋白-2) (Mitcham 等(1996) J. Biol. Chem. 271:5777-5783; Hardiman 等 (1996) Oncogene 13:2467-2475)。用所述新的人 DTLR 序列,通过分析共同 TH 域的构象,我们已经找到该进化线索的一个结构定义:包含 128 个氨基酸的 10 个区组的保守序列,形成最小的 TH 域折叠;序列对比中的间隙,标记了序列和长度可变的环的可能位置(图 2a)。

利用多个序列对比序列中保守和变异模式的两个预测算法 PHD (Rost 和 Sander (1994) Proteins 19:55-72)和 DSC (King 和 Sternberg (1996) Protein Sci. 5:2298-2310),产生 TH 信号组件的强的一致结果(图 2a)。每个区组含有一个独立的二级结构元件:交替的 β -链(标记为 A-E)和 α -螺旋(编号为 1-5)的特征,是在平行 β -片两面上具有 α -螺旋的 β/α 类折叠的诊断性特征。预测疏水性 β -链 A、C 和 D 形成 β -片中的“内部”板条,而较短的两性 β -链 B 和 E 装配典型的“边缘”单位(图 2a)。这种布置与核心 β -片中 B-A-C-D-E 的链顺序一致(图 2b);折叠比较(“映射(mapping)”)和识别(“threading”)程序(Fischer 等 1996) FASEB J. 10:126-136)强烈地返回该双缠绕 β/α 拓扑学。TH 域的这种轮廓结构的一个惊人的功能预测是:将多重序列对比中的许多保守的带电荷的残基作图至该 β -片的 C 末端: β A 末端的残基 Asp16 (区组编码方案-图 2a)、 β B 之后的 Arg39 和 Asp40、 α 3 的第一个转角中的 Glu75 以及 β D- α 4 环中或 β E 之后更松散保守的 Glu/Asp 残基(图 2a)。其它 4 个保守残基(Asp7、Glu28 和 Arg57-Arg/Lys58 对)的位置与该 β -片相对的 N 末端的盐桥网络相匹配(图 2a)。

信号功能依赖于 TH 域的结构完整性。对于 IL-1R1 和 Toll,已经将所述组件边界内的失活突变或缺失(图 2a)按目录分类。Heguy 等 (1992) J. Biol. Chem. 267:2605-2609; Croston 等(1995) J. Biol. Chem. 270:16514-16517; Schneider 等(1991) Genes Develop. 5:797-807; Norris

和 Manley (1992) Genes Develop. 6:1654-1667; Norris 和 Manley (1995) Genes Develop. 9:358-369; 和 Norris 和 Manley (1996) Genes Develop. 10:862-872。延伸越过所述最小 TH 域(分别长 8、0、6、22 和 18 个残基)的人 DTLR1-5 链, 与 Mst ORF 的短粗的 4 aa “尾”最为相似。Toll 和 18w 表现出不相关的 102 个和 207 个残基的尾(图 2a), 它可以负调节所述融合 TH 域的信号。Norris 和 Manley (1995) Genes Develop. 9:358-369; 和 Norris 和 Manley (1996) Genes Develop. 10:862-872。

携带该 TH 域的不同蛋白之间的进化关系, 可能最好通过得自多重序列对比的系统发育树来鉴别(图 3)。4 个主要的分支分开了所述植物蛋白、MyD88 因子、IL-1 受体和 Toll 样分子; 最后一个分支聚集于果蝇属和人的 DTLR。

人 DTLR 基因的染色体散布。

为了研究该新生的人 DTLR 基因家族的遗传连锁, 我们用 FISH 对 5 个基因中的 4 个的染色体位点进行了作图(图 4)。人类基因组计划先前已经绘制了 DTLR1 基因: Humrc786 cDNA 存在一个 STS 数据库基因座(dbSTS 登记号 G06709, 对应于 STS WI-7804 或 SHGC-12827) (Nomura 等, (1994) DNA Res 1:27-35), 并且将该基因固定至第 4 染色体标记间隔 D4S1587-D42405 (50-56 cM) circa 4p14。最近已经通过 FISH 分析证实这种排列。Taguchi 等(1996) Genomics 32:486-488。在本发明工作中, 我们可靠地将其余的 DTLR 基因排列于染色体 4q32 (DTLR2)、4q35 (DTLR3)、9q32-33 (DTLR4)和 1q33.3 (DTLR5)上的基因座。在该工作期间, 已经产生了亲代 DTLR2 EST (CloneID # 80633) 的 STS (STS SHGC-33147 的 dbSTS 登记号为 T57791), 并且根据我们的发现, 将其作图至第 4 染色体 4q32 的标记间隔 D4S424-D4S1548 (143-153 cM)。在第 4 染色体长臂上的 DTLR2 和 DTLR3 基因之间有一个约 50 cM 的间隙。

差异表达 DTLR 基因。

在果蝇属中, Toll 和 18w 均具有复杂的空间和时间表达模式, 这可能指出胚胎模式以外的功能。St. Johnston 和 Nusslein-Volhard (1992) Cell 68:201-219; Morisato 和 Anderson (1995) Ann. Rev. Genet. 29:371-399; Belvin 和 Anderson (1996) Ann. Rev. Cell Develop. Biol. 12:393-416; Lemaitre 等(1996) Cell 86:973-983; Chiang 和 Beachy (1994) Mech. Develop. 47:225-239; 和 Eldon 等(1994) Develop. 120:885-899。我们采用放射标记的 DTLR cDNA, 通过用不同的人组织和癌细胞系的 mRNA 印迹分析, 检查了 DTLR 转录物的空间分布(图 5)。发现 DTLR1 是普遍表达的, 并且表达水平高于其它受体。“短的” 3.0 kB 和“长的” 8.0 kB 的 DTLR1 转录物形式分别存在于卵巢和脾脏中, 推定反映出替代剪接(图 5, A 板和 B 板)。癌细胞 mRNA 一板也显示出 DTLR1 在伯基特淋巴瘤 Raji 细胞系中占优势的过量表达(图 5, C 板)。DTLR2 mRNA 的表达没有 DTLR1 的范围广, 在肺脏中检测到一个 4.0 kB 的种类, 在心脏、脑和肌肉中明显检测到一个 4.4 kB 的转录物。DTLR3 的组织分布模式重复了 DTLR2 的组织分布模式(图 5, E 板)。DTLR3 也以大小约为 4.0 kB 和 6.0 kB 的两个主要转录物存在, 在胎盘和胰脏中观察到最高水平的表达。相反, DTLR4 和 DTLR5 信息看来是极度组织特异性的。DTLR4 仅在胎盘中作为约 7.0 kB 的单一转录物检测到。在卵巢和外周血单核细胞中观察到 DTLR5 的 4.0 kB 弱信号。

进化上古老的调节系统的组分。

通过比较基因组法, 可以重建信号途径的原始分子蓝图和趋异命运。Miklos 和 Rubin (1996) Cell 86:521-529; Chothia (1994) Develop. 1994 增刊, 27-33; Banfi 等(1996) Nature Genet. 13:167-174; 和 Wang 等(1996) J. Biol. Chem. 271:4468-4476。我们已经采用该逻辑, 鉴别出一个人类中出现的基因家族, 它们目前编码 5 个受体的平行进化同源物

DTLR 1-5, 它们是以 Toll 为首的果蝇属基因家族的直接进化的对应物(图 1-3)。人和蝇 DTLR 的保守结构即保守的 LRR 胞外域和胞内 TH 域(图 1), 暗示与果蝇属中 Toll 偶联的坚固的途径(6、7)在脊椎动物中继续存在。最佳证据来自重复途径: 多功能 IL-1 系统及其受体融合的 TH 域的所有组成成分 IRAK、NF- κ B 和 I- κ B 同系物(Belvin 和 Anderson (1996) Ann. Rev. Cell Develop. Biol. 12:393-416; 和 Wasserman (1993) Molec. Biol. Cell 4:767-771; Hardiman 等 (1996) Oncogene 13:2467-2475; 和 Cao 等(1996) Science 271:1128-1131); 也已经特征鉴定出一种 Tube 样因子。尚不知 DTLR 是否可以有效地偶联于 IL-1R 信号机器, 或者, 而是使用平行的一组蛋白。与 IL-1 受体不同, 预测人 DTLR 的 LRR 支架保留对 Spatzle-Trunk 相关胱氨酸结因子的亲和性; 已经分离出适合该模型的候选 DTLR 配体(所谓的 PEN)。

通过途径中互作蛋白折叠的保守性, 可以估计信号转导的生化机制。Miklos 和 Rubin (1996) Cell 86:521-529; Chothia (1994) Develop. 1994 增刊, 27-33。目前, Toll 信号范例包括其作用受其结构、作用或命运精密限定的某些分子: Pelle 是一种 Ser/Thr 激酶(磷酸化), Dorsal 是一种 NF- κ B 样转录因子(DNA-结合), 而 Cactus 是一种锚蛋白重复抑制剂(Dorsal 结合, 降解)。Belvin 和 Anderson (1996) Ann. Rev. Cell Develop. Biol. 12:393-416。相反, Toll TH 域和 Tube 的功能尚不清楚。与其它细胞因子受体一样(Heldin (1995) Cell 80:213-223), 配体介导的 Toll 二聚看来是触发事件: Toll 近膜区中的游离半胱氨酸产生组成型活性的受体对(Schneider 等(1991) Genes Develop. 5:797-807), 而嵌合 Torso-Toll 受体以二聚体作为信号(Galindo 等(1995) Develop. 121:2209-2218); 此外, Toll 胞外域的严重截短或全部丧失, 产生杂乱的胞内信号(Norris 和 Manley (1995) Genes Develop. 9:358-369; Winans 和 Hashimoto (1995) Molec. Biol. Cell 6:587-596), 这是具有催化域的癌基因受体的回忆(Heldin (1995) Cell 80:213-223)。Tube 是膜定位的, 它连接 Pelle 的 N 末端(死)域并且磷酸化, 但双杂种分析均未显示出 Toll-

Tube 和 Toll-Pelle 的相互作用(Galindo 等(1995) Develop. 121:2209-2218; 和 Großhans 等(1994) Nature 372:563-566); 这后一结果提示: Toll TH 域的构象“状态”略微影响因子的募集。Norris 和 Manley (1996) Genes Develop. 10:862-872; 和 Galindo 等(1995) Develop. 121:2209-2218。

这些令人苦恼的问题的本质是 Toll TH 域的结构性质。为了解决该问题, 我们已经利用了来自昆虫、植物和脊椎动物的 TH 序列的进化趋异性, 加入人 DTLR 链, 并提取用于结构预测和折叠识别的最小的保守蛋白核心(图 2)。强烈预测的 $(\beta/\alpha)_5$ TH 域折叠具有其酸性残基不对称簇, 在拓扑学上与细菌中双组分信号途径中的效应调节物的结构相同(Volz (1993) Biochemistry 32:11741-11753; 和 Parkinson (1993) Cell 73:857-871)(图 2)。原型趋化性调节物 CheY 瞬时结合核心 β -片 C 末端“天冬氨酸袋”中的二价阳离子; 该阳离子提供静电稳定性, 并促进不变的 Asp 的活化磷酸化。Volz (1993) Biochemistry 32:11741-11753。同样, 该 TH 域可以捕捉其酸性巢中的阳离子, 但激活和下游信号依赖于带负电部分的特异性结合: 阴离子配体可以通过将其锁入精确的氢键网络中, 而克服强负电结合位点的电势。Ledvina 等(1996) Proc. Nat'l Acad. Sci. USA 93:6786-6791。有趣的是, 该 TH 域可以不仅仅用作装配 Toll 或植物和脊椎动物中同源系统的 Tube/Pelle 复合物的被动支架, 而是作为信号转导机器中真正的构象触发剂而主动参与。Toll 的二聚可以促进由调节受体尾引起的解掩蔽, 也许解释了 Tube/Pelle 复合物的条件结合(Norris 和 Manley (1995) Genes Develop. 9:358-369; 和 Norris 和 Manley (1996) Genes Develop. 10:862-872), 或促进 TH 袋的小分子激活物的结合。然而, 该细胞内的“游离”TH 域(Norris 和 Manley (1995) Genes Develop. 9:358-369; Winans 和 Hashimoto (1995) Molec. Biol. Cell 6:587-596)可能通过激活并与游走的 Tube/Pelle 复合物对接, 作为催化性的 CheY 样触发剂起作用。

形态发生受体和免疫防御

昆虫和脊椎动物免疫系统之间的进化联系印在 DNA 中: 昆虫中编码抗微生物因子的基因显示出的上游基序与哺乳动物中急性期效应元件相似, 已知所述急性期效应元件结合 NF- κ B 转录因子。Hultmark (1993) Trends Genet. 9:178-183。Dorsal 和两种 Dorsal 相关因子 Dif 和 Relish, 帮助在细菌攻击后诱导这些防御蛋白(Reichhart 等(1993) C. R. Acad. Sci. Paris 316:1218-1224; Ip 等(1993) Cell 75:753-763; 和 Dushay 等(1996) Proc Nat'l Acad. Sci. USA 93:10343-10347); Toll 或其它 DTLR 可能在果蝇属成体中介导这些快速免疫应答(Lamaitre 等(1996) Cell 86:973-983; 和 Rosetto 等(1995) Biochem. Biophys. Res. Commun. 209:111-116)。在脊椎动物中, 与 IL-1 炎症反应机制平行的反应明显具有 Toll 信号途径的功能多面性, 并提示在胚胎图式发育和先天免疫之间古老的协同作用(Belvin 和 Anderson (1996) Ann. Rev. Cell Develop. Biol. 12:393-416; Lamaitre 等(1996) Cell 86:973-983; Wasserman (1993) Molec. Biol. Cell 4:767-771; Wilson 等(1997) Curr. Biol. 7:175-178; Hultmark (1993) Trends Genet. 9:178-183; Reichhart 等(1993) C. R. Acad. Sci. Paris 316:1218-1224; Ip 等(1993) Cell 75:753-763; 和 Dushay 等(1996) Proc Nat'l Acad. Sci. USA 93:10343-10347; Rosetto 等(1995) Biochem. Biophys. Res. Commun. 209:111-116; Medzhitov 和 Janeway (1997) Curr. Opin. Immunol. 9:4-9; 和 Medzhitov 和 Janeway (1997) Curr. Opin. Immunol. 9:4-9)。昆虫和人 DTLR 蛋白的更密切的同系物, 引起甚至更强的生物学功能的重叠, 这些功能替代与 IL-1 系统平行的纯免疫系统, 并提供对脊椎动物胚胎背腹转化和其它转化的潜在的分子调节物。DeRobertis 和 Sasai (1996) Nature 380:37-40; 和 Arendt 和 Nubler-Jung (1997) Mech. Develop. 61:7-21。

本文对人类中一个出现的强大的(robust)受体家族的描述, 反映出最近对 Wnt 图式发育因子的脊椎动物 Frizzled 受体的发现。Wang 等(1996) J. Biol. Chem. 271:4468-4476。因为许多其它细胞因子受体系统在早期发育中有作用(Lemaire 和 Kodjabachian (1996) Trends Genet.

12:525-531), 也许致密的胚胎和组合的(gangly)成体的不同细胞环境 (cellular context)仅仅产生熟悉的信号途径和在不同时间具有不同生物学结果的其可扩散触发剂, 所述不同的生物学结果例如 DTLR 的形态发生相对于免疫防御。对于昆虫、植物和人的 Toll 相关系统(Hardiman 等 (1996) *Oncogene* 13:2467-2475; Wilson 等(1997) *Curr. Biol.* 7:175-178), 这些信号通过调节性 TH 域运行, 所述 TH 域有趣地类似细菌转导的发动机(Parkinson (1993) *Cell* 73:857-871)。

特别是, DTLR6 表现出确立在该家族中其成员资格的结构特征。此外, 该家族成员已经涉及许多明显的发育疾病病征, 并与先天免疫系统有关。具体地说, 已经将 DTLR6 作图至 X 染色体的一个位置, 该位置是主要的发育异常的热点。参见例如 The Sanger Center: 人 X 染色体网站 <http://sanger.ac.uk/HGP/ChrX/index.shtml>; 和 the Baylor College of Medicine 人类基因组测序网站 <http://gc.bcm.tmc.edu:8088/cgi-bin/seq/home>。

登记的 PAC 的登记号是 AC003046。该登记号带有来自两种 PAC 的序列: RPC-164K3 和 RPC-263P4。这两种 PAC 序列作图至人染色体 Xp22, 于 Baylor 网站标记 DX704 和 DXS7166 之间。该区是严重的发育异常的“热点”。

III. DTLR 片段的 PCR 扩增

选择两种合适的引物序列(参见表 1-10)。在根据产生部分或全长 cDNA 的信息的存在而选择的合适 mRNA 样品上, 例如在表达该基因的样品上, 使用 RT-PCR。参见例如 Innis 等(编辑, 1990) *PCR Protocols: A Guide to Methods and Applications* Academic Press, San Diego, CA; 和 Dieffenbach 和 Dveksler(1995; 编辑) *PCR Primer: A Laboratory Manual* Cold Spring Harbor Press, CSH, NY。这样将允许确定有用的序列, 以探测 cDNA 文库中的全长基因。TLR6 是该基因组中的一个连续序列,

这可能提示其它 TLR 也是连续的序列。因此,对基因组 DNA 进行 PCR 可能产生全长的连续序列,然后可应用染色体步查方法。或者,序列数据库会含有对应于所需实施方案的部分的序列,或密切相关的形式,例如替代剪接等。表达克隆技术也可以应用于 cDNA 文库。

IV. DTLR 的组织分布

已经检测到编码这些 DTLR 的每个基因的信息。参见图 5A-5F。将用合适的技术分析其它细胞和组织,所述合适的技术例如为 PCR、免疫测定、杂交或其它方法。组织和器官 cDNA 制备物可得自 Clontech, Mountain View, CA。天然表达源的鉴定是有用的,如所述的。

DNA 分析: 来自首次扩增的 cDNA 文库的 DNA (5 μ g)用合适的限制性酶消化,以释放插入片段,在 1%琼脂糖凝胶上电泳,并转移至尼龙膜(Schleicher 和 Schuell, Keene, NH)。

用于人 mRNA 分离的样品通常包括例如: 外周血单核细胞(单核细胞、T 细胞、NK 细胞、粒细胞、B 细胞),静息的(T100); 外周血单核细胞,用抗 CD3 激活 2、6、12 小时的混合物(pooled)(T101); T 细胞, TH0 克隆 Mot 72, 静息的(T102); T 细胞, TH0 克隆 Mot 72, 用抗 CD28 和抗 CD3 激活 3、6、12 小时的混合物(T103); T 细胞, TH0 克隆 Mot 72, 用特定肽无反应性处理 2、7、12 小时的混合物(T104); T 细胞, TH1 克隆 HY06, 静息的(T107); T 细胞, TH1 克隆 HY06, 用抗 CD28 和抗 CD3 激活 3、6、12 小时的混合物(T108); T 细胞, TH1 克隆 HY06, 用特定肽无反应性处理 2、6、12 小时的混合物(T109); T 细胞, TH2 克隆 HY935, 静息的(T110); T 细胞, TH2 克隆 HY935, 用抗 CD28 和抗 CD3 激活 2、7、12 小时的混合物(T111); T 细胞 CD4+CD45RO-, 在抗 CD28、IL-4 和抗 IFN- γ 中极化的 T 细胞, 极化的 TH2, 用抗 CD3 和抗 CD28 激活 4 小时(T116); T 细胞肿瘤系 Jurkat 和 Hut78, 静息的(T117); T 细胞克隆, 混合的 AD130.2、Tc783.12、Tc783.13、Tc783.58、Tc783.69, 静息的(T118); T 细胞随机 $\gamma\delta$ T 细胞克隆, 静息的(T119);

脾细胞, 静息的(B100); 脾细胞, 用抗 CD40 和 IL-4 激活的(B101); B 细胞 EBV 系, 混合的 WT49、RSB、JY、CVIR、721.221、RM3、HSY, 静息的(B102); B 细胞系 JY, 用 PMA 和离子霉素激活 1、6 小时的混合物(B103); 混合的 K 20 克隆, 静息的(K100); 混合的 NK20 克隆, 用 PMA 和离子霉素激活 6 小时(K101); NKL 克隆, 得自 LGL 白血病患者的外周血, IL-2 处理的(K106); NK 细胞毒性克隆 640-A30-1, 静息的(K107); 造血前体系 TF1, 用 PMA 和离子霉素激活 1、6 小时的混合物(C100); U937 前单核细胞系, 静息的(M100); U937 前单核细胞系, 用 PMA 和离子霉素激活 1、6 小时的混合物(M101); 淘析的单核细胞, 用 LPS、IFN γ 、抗 IL-10 激活 1、2、6、12、24 小时的混合物(M102); 淘析的单核细胞, 用 LPS、IFN γ 、IL-10 激活 1、2、6、12、24 小时的混合物(M103); 淘析的单核细胞, 用 LPS、IFN γ 、抗 IL-10 激活 4、16 小时的混合物(M106); 淘析的单核细胞, 用 LPS、IFN γ 、IL-10 激活 4、16 小时的混合物(M107); 淘析的单核细胞, 用 LPS 激活 1 小时(M108); 淘析的单核细胞, 用 LPS 激活 6 小时(M109); DC 70% CD1a+, 来自 CD34+ GM-CSF、TNF α 12 天, 静息的(D101); DC 70% CD1a+, 来自 CD34+ GM-CSF、TNF α 12 天, 用 PMA 和离子霉素激活 1 小时(D102); DC 70% CD1a+, 来自 CD34+ GM-CSF、TNF α 12 天, 用 PMA 和离子霉素激活 6 小时(D103); DC 95% CD1a+, 来自 CD34+ GM-CSF、TNF α 12 天, 通过 FACS 分选, 用 PMA 和离子霉素激活 1、6 小时的混合物(D104); DC 95% CD14+, 来自 CD34+ GM-CSF、TNF α 12 天, 通过 FACS 分选, 用 PMA 和离子霉素激活 1、6 小时的混合物(D105); DC CD1a+ CD86+, 来自 CD34+ GM-CSF、TNF α 12 天, 通过 FACS 分选, 用 PMA 和离子霉素激活 1、6 小时的混合物(D106); DC, 来自单核细胞 GM-CSF、IL-4 5 天, 静息的(D107); DC, 来自单核细胞 GM-CSF、IL-4 5 天, 静息的(D108); DC, 来自单核细胞 GM-CSF、IL-4 5 天, 用 LPS 激活 4、16 小时的混合物(D109); DC, 来自单核细胞 GM-CSF、IL-4 5 天, 用 TNF α 、单核细胞上清液(supe)

激活 4、16 小时的混合物(D110); 平滑肌瘤 L11 良性肿瘤(X101); 正常子宫肌层 M5 (O115); 恶性平滑肌肉瘤 GS1 (X103); 肺成纤维细胞肉瘤系 MRC5, 用 PMA 和离子霉素激活 1、6 小时的混合物(C101); 肾上皮癌细胞系 CHA, 用 PMA 和离子霉素激活 1、6 小时的混合物(C102); 28 周的雄性胎肾(O100); 28 周的雄性胎肺(O101); 28 周的雄性胎肝(O102); 28 周的雄性胎心(O103); 28 周的雄性胎脑(O104); 28 周的雄性胎儿膀胱(O106); 28 周的雄性胎儿小肠(O107); 28 周的雄性胎儿脂肪组织(O108); 25 周的雌性胎儿卵巢(O109); 25 周的雌性胎儿子宫(O110); 28 周的雄性胎儿睾丸(O111); 28 周的雄性胎脾(O112); 28 周的成人胎盘(O113); 和来自 12 岁的发炎扁桃体(X100)。

用于小鼠 mRNA 分离的样品可以包括例如: 静息的小鼠成纤维细胞的 L 细胞系(C200); Braf:ER (与雌激素受体融合的 Braf 融合体)转染的细胞, 对照(C201); T 细胞, TH1, 极化的(Me114 bright, 来自脾脏的 CD4+细胞, 用 IFN- γ 和抗 IL-4 极化 7 天; T200); T 细胞, TH2, 极化的(Me114 bright, 来自脾脏的 CD4+细胞, 用 IL-4 和抗 IFN- γ 极化 7 天; T201); T 细胞, 高度极化的 TH1 (参见 Openshaw 等(1995) L. Exp. Med. 182:1357-1367; 用抗 CD3 激活 2、6、16 小时的混合物; T202); T 细胞, 高度极化的 TH2 (参见 Openshaw 等(1995) L. Exp. Med. 182:1357-1367; 用抗 CD3 激活 2、6、16 小时的混合物; T203); CD44-CD25+前 T 细胞, 从胸腺分选的(T204); TH1 T 细胞克隆 D1.1, 用抗原最后一次刺激后静息 3 周(T205); TH1 T 细胞克隆 D1.1, 10 μ g/ml ConA 激活 15 小时(T206); TH2 T 细胞克隆 CDC35, 用抗原最后一次刺激后静息 3 周(T207); TH2 T 细胞克隆 CDC35, 10 μ g/ml ConA 激活 15 小时(T208); Me114+原初 T 细胞, 来自脾脏, 静息的(T209); Me114+T 细胞, 用 IFN- γ /IL-12/抗 IL-4 极化 6、12、24 小时的混合物, 极化成 Th1 (T210); Me114+ T 细胞, 用 IL-4/抗 IFN- γ 极化 6、13、24 小时的混合物, 极化成 Th2 (T211); 未刺激的成熟 B 细胞白血病细胞系 A20 (B200); 未刺激的 B 细胞系 CH12 (B201); 来自脾脏的未刺激的大 B

细胞(B202); 来自全脾的 B 细胞, LPS 激活(B203); 来自脾脏的甲泛葡胺富集的树突细胞, 静息的(D200); 来自骨髓的树突细胞, 静息的(D201); 单核细胞细胞系 RW 264.7, 用 LPS 激活 4 小时(M200); 骨髓巨噬细胞, 用 GM 和 M-CSF 衍生的(M201); 巨噬细胞系 J774, 静息的(M202); 于 0.5、1、3、6、12 小时混合的巨噬细胞系 J774 + LPS + 抗 IL-10 (M203); 于 0.5、1、3、6、12 小时混合的巨噬细胞系 J774 + LPS + IL-10 (M204); 气溶胶攻击的小鼠肺组织, Th2 引物(primer)、气溶胶 OVA 攻击 7、14、23 小时的混合物(参见 Garlisi 等(1995) Clinical Immunology and Immunopathology 75:75-83; X206); 日圆线虫属 (*Nippostrongylus*) 感染的肺组织 (参见 Coffman 等(1989) Science 245:308-310; X200); 全成人肺脏, 正常的(O200); 全肺脏, rag-1 (参见 Schwarz 等(1993) Immunodeficiency 4:249-252; O205); IL-10 K.O. 脾脏(参见 Kuhn 等(1991) Cell 75:263-274; X201); 全成人脾脏, 正常的(O201); 全脾脏, rag-1 (O207); IL-10 K.O. 肠系膜淋巴结(X202); 全肠系膜淋巴结, 正常的(O210); IL-10 K.O. 肠系膜淋巴结(X203); 全肠系膜淋巴结, 正常的(O211); IL-10 K.O. 结肠(X203); 全结肠, 正常的(O212); NOD 小鼠胰脏(参见 Makino 等(1980) Jikken Dobutsu 29:1-13; X205); 全胸腺, rag-1 (O208); 全肾脏, rag-1 (O209); 全心脏, rag-1 (O202); 全脑, rag-1 (O203); 全睾丸, rag-1 (O204); 全肝脏, rag-1 (O206); 大鼠正常关节组织(O300); 和大鼠关节炎关节组织(X300)。

V. DTLR 的种对应物的克隆

采用各种策略, 以获得这些 DTLR 的种对应物, 最好是来自其它灵长类的种对应物。一种方法是采用密切相关的物种的 DNA 探针进行交叉杂交。调查进化上相似的物种作为中间步骤, 可能是有用的。另一方法是采用特异性 PCR 引物, 所述引物基于特定物种(例如人)基因(例如多肽或核苷酸序列高度保守或非保守的区)之间相似性或差异的区组鉴定。或者, 可以将抗体用于表达克隆。

VI. 哺乳动物 DTLR 蛋白的生产

工程改造合适的(例如 GST)融合构成物, 用于在例如大肠杆菌中表达。例如, 构建小鼠 IGIF pGex 质粒, 并将其转化入大肠杆菌中。在含有 50 $\mu\text{g/ml}$ 氨苄青霉素的 LB 培养基中培养新转化的细胞, 并用 IPTG (Sigma, St. Louis, MO) 诱导。过夜诱导后, 收获细菌, 并分离含有 DTLR 蛋白的沉淀。将所述沉淀在 2 升 TE 缓冲液(50 mM Tris 碱 pH 8.0、10 mM EDTA 和 2 mM pefabloc)中匀化。使该材料通过微流化装置(microfluidizer) (Microfluidics, Newton, MA)三次。在 Sorvall GS-3 转子上, 将流化的上清液以 13,000 rpm 离心 1 小时。所产生的上清液含所述 DTLR 蛋白, 将其过滤, 并使其通过谷胱甘肽-SEPHAROSE 柱, 该柱在 50 mM Tris 碱 pH 8.0 中平衡。合并含有所述 DTLR-GST 融合蛋白的流分, 并用凝血酶(Enzyme Research Laboratories, Inc., South Bend, IN)酶切。然后, 使酶切的合并物通过 Q-SEPHAROSE 柱, 该柱在 50 mM Tris 碱中平衡。合并含有 DTLR 的流分, 并在冷蒸馏水中稀释, 以降低电导率, 并再次仅通过新的 Q-Sepharose 柱, 或再次接连通过新的 Q-Sepharose 和免疫亲和抗体柱。合并含有所述 DTLR 蛋白的流分, 将其分成等分样品, 并于 -70°C 冰箱中贮存。

将 CD 谱与 DTLR1 蛋白比较, 可以提示该蛋白是正确地折叠的。参见 Hazuda 等(1969) J. Biol. Chem. 264:1689-1693。

VII. 用 DTLR 进行生物测定

生物测定一般涉及该蛋白的配体结合特征, 或涉及该受体的激酶/磷酸酶活性。该活性通常是可逆的, 许多其它酶的作用(介导磷酸酶或磷酸化酶活性)也是可逆的, 采用标准方法可容易地测量所述活性。参见例如 Hardie 等(编辑, 1995) The Protein Kinase FactBook 第 I 和 II 卷, Academic Press, San Diego, CA; Hanks 等(1991) Meth. Enzymol. 200:38-62; Hunter 等(1992) Cell 70:375-388; Lewin (1990) Cell 61:743-752; Pines

等(1991) Cold Spring Harbor Symp. Quant. Biol. 56:449-463; 和 Parker 等(1993) Nature 363:736-738.

白介素-1 家族含有这样的分子, 它们每种均是重要的炎症疾病的介质。关于综合性综述, 参见 Dinarello (1996) “疾病中白介素-1 的生物学基础” Blood 87:2095-2147。提示不同的 Toll 配体可能在疾病起始、特别是炎症反应中起重要作用。与 IL-1 家族相关的新蛋白的发现, 促进这种分子的鉴定, 所述分子提供疾病起始的分子基础, 并允许开发范围和效力增加的治疗策略。

VIII. 对例如 DTLR4 特异性的抗体的制备

近交 Balb/c 小鼠用重组形式的所述蛋白(例如纯化的 DTLR4)腹膜内免疫, 或用稳定转染的 NIH-3T3 细胞腹膜内免疫。于合适的时间点, 用加入或不加入额外佐剂的蛋白加强免疫动物, 以进一步刺激抗体产生。收集血清, 或用收获的脾脏产生杂交瘤。

或者, Balb/c 小鼠用以该基因或其片段转化的细胞(或者内源细胞或者异源细胞)免疫, 或用分离的富集膜免疫, 以表达该抗原。于合适的时间点收集血清, 通常在多次进一步给予免疫后收集。在原位生产蛋白中, 各种基因治疗技术可能可用以产生免疫应答。

可以制备单克隆抗体。例如, 将脾细胞与合适的融合伴侣融合, 通过标准方法, 在生长培养基中选择杂交瘤。例如通过 ELISA 或其它测定, 根据与所需 DTLR 结合的抗体的存在, 筛选杂交瘤上清液。也可以选择或制备特异性识别特定 DTLR 实施方案的抗体。

在另一方法中, 可以将合成的肽或纯化的蛋白提呈给免疫系统, 以产生单克隆抗体或多克隆抗体。参见例如 Coligan (1991) Current Protocols in Immunology Wiley/Greene; 和 Harlow 和 Lane (1989) Antibodies: A Laboratory Manual Cold Spring Harbor Press。在合适的情况下, 结合试剂或者如上所述进行标记, 例如荧光标记或其它标记, 或者将其固定至一种基质, 以用于淘选方法。也可以将核酸引入动物

细胞中,以产生用来诱出免疫应答的抗原。参见例如 Wang 等(1993) Proc. Nat'l Acad. Sci. 90:4156-4160; Barry 等(1994) BioTechniques 16:616-619; 和 Xiang 等(1995) Immunity 2:129-135。

IX. 与例如 DTLR5 融合的融合蛋白的产生

用 DTLR5 制备不同的融合构成物。将该基因的部分融合于一种表位标记,例如 FLAG 标记,或融合于双杂种系统构成物。参见例如 Fields 和 Song (1989) Nature 340:245-246。

所述表位标记可以用于表达克隆方法中,用抗 FLAG 抗体检测,以检测结合伴侣,例如相应 DTLR5 的配体。所述双杂种系统也可以用来分离特异性结合于 DTLR5 的蛋白。

X. DTLR 的染色体作图

制备染色体线(chromosome spreads)。在染色体制备物上进行原位杂交,用以植物凝集素刺激、培养 72 小时的淋巴细胞获得所述染色体制备物。加入 5-溴脱氧尿苷培养最后 7 小时(60 µg/ml 培养基),以确保良好质量的杂交后染色体显带。

将借助引物在总 B 细胞 cDNA 模板上扩增的合适片段(例如 PCR 片段),克隆入合适的载体中。通过切口平移,用 ³H 标记该载体。使放射标记的探针杂交至中期线上,如 Mattei 等(1985) Hum. Genet. 69: 327-331 中所述。

玻片用核示踪乳液(KODAK NTB₂)包被后,于 4℃曝光例如 18 天。为了避免在显带步骤期间有任何银颗粒滑落,染色体线首先用缓冲的吉姆萨溶液染色,并拍摄中期照片。然后,通过荧光染料-质子传递作用-吉姆萨(FPG)法,进行 R-带显带,并且在分析之前,再次拍摄中期照片。

或者,可以如上所述进行 FISH。将所述 DTLR 基因定位于不同的染色体。DTLR2 和 DTLR3 定位至人类第 4 染色体; DTLR4 定位于人

类第9染色体, 而 DTLR5 定位于人类第1染色体。参见图 4A-4D。

XI. 结构活性关系

采用标准方法和分析, 测定特定残基关键性的信息。例如通过于确定的位置(例如于以上鉴定的位置)产生许多不同的变异体, 并且评价所述变异体的生物活性, 进行标准诱变分析。在某种程度上进行该分析, 以确定修饰活性的位置, 或集中于特定的几个位置, 以确定可以取代的残基, 以或者保留、阻断或调节生物活性。

或者, 天然变异体的分析可以指明什么位置耐受天然突变。这可以得自个体、品系或物种间变异的群体分析。例如通过 PCR 分析并测序, 分析选定个体的样品。这允许评价群体多态性。

XI. DTLR 配体的分离

利用 DTLR 的结合特异性, 可以将 DTLR 用作特异性的结合试剂, 以鉴定其结合伴侣, 这非常象使用抗体。结合试剂或者如上所述标记, 例如荧光标记或其它标记, 或者固定至一种基质, 以用于淘选方法。

所述结合组合物用来筛选表达文库, 所述表达文库由表达结合伴侣(例如配体, 最好是膜结合的)的细胞系制备。标准染色方法用来检测或分选表面表达的配体, 或通过淘选, 筛选表面表达的转化细胞。采用各种染色方法或免疫荧光法, 进行胞内表达的筛选。也参见 McMahan 等(1991) EMBO J 10:2821-2832。

例如, 于第0天, 用每室1 ml 纤连蛋白(10 ng/ml)的 PBS 于室温下预包被2室的 permanox 玻片30分钟。用 PBS 冲洗1次。然后, 将 COS 细胞以 $2-3 \times 10^5$ 细胞/室, 在1.5 ml 生长培养基中平板接种。于37℃培养过夜。

对于每种样品, 第一天, 制备0.5 ml 无血清 DME 溶液, 该溶液含有 66 µg/ml DEAE-葡聚糖、66 µM 氯喹和 4 µg DNA。对于每组, 制

备阳性对照, 即 1 和 1/200 稀释度的 DTLR-FLAG cDNA, 以及制备阴性模拟物。用无血清 DME 冲洗细胞。加入该 DNA 溶液, 于 37℃ 温育 5 小时。去除培养基, 加入 0.5 ml 10% DMSO 的 DME 2.5 分钟。去除 DME, 并用 DME 洗涤 1 次。加入 1.5 ml 生长培养基, 并且培养过夜。

第 2 天, 更换培养基。第 3 或 4 天, 将所述细胞固定并染色。用 Hank 缓冲盐溶液(HBSS)冲洗细胞 2 次, 并于 4%低聚甲醛(PFA)/葡萄糖中固定 5 分钟。用 HBSS 洗涤 3 次。去除所有液体后, 将所述玻片于 -80℃ 贮存。对于每个室, 如下进行 0.5 ml 温育。加入具有 32 $\mu\text{l/ml}$ 1 M NaN_3 的 HBSS/皂苷(0.1%) 20 分钟。然后细胞用 HBSS/皂苷洗涤 1 次。将合适的 DTLR 或 DTLR/抗体复合物加入细胞中, 并温育 30 分钟。用 HBSS/皂苷洗涤细胞 2 次。如果合适, 加入第一抗体 30 分钟。以 1/200 的稀释度加入第二抗体, 例如 Vector 抗小鼠抗体, 并温育 30 分钟。制备 ELISA 溶液, 例如 Vector Elite ABC 辣根过氧化物酶溶液, 并预温育 30 分钟。使用例如每 2.5 ml HBSS/皂苷 1 滴 A 溶液(抗生物素蛋白)和 1 滴 B 溶液(生物素)。用 HBSS/皂苷洗涤细胞 2 次。加入 ABC HRP 溶液, 并温育 30 分钟。用 HBSS 洗涤细胞 2 次, 第二次洗涤 2 分钟, 封闭细胞。然后, 加入 Vector 的二氨基苯甲酸(DAB) 5-10 分钟。使用每 5 ml 玻璃蒸馏水 2 滴缓冲液加 4 滴 DAB 加 2 滴 H_2O_2 。小心地除去室, 并在水中冲洗玻片。风干数分钟, 然后加入 1 滴 Crystal Mount, 加上盖玻片。于 85-90℃ 烘 5 分钟。

评价库和渐进亚克隆的阳性染色, 以分离负责结合的单基因。

或者, 用 DTLR 试剂来亲和纯化或分选出表达推定配体的细胞。参见例如 Sambrook 等或 Ausubel 等。

另一策略是通过淘选, 筛选膜结合受体。如上所述构建该受体的 cDNA。可以将该配体固定化, 并用以固定表达的细胞。利用识别例如 DTLR 融合构成物的 FLAG 序列的合适抗体, 或利用针对第一抗体产生的抗体, 可以达到固定化。筛选和扩增的递归(recursive)循环导致合适克隆的富集, 并且最终分离受体表达克隆。

用哺乳动物 DTLR 可以筛选噬菌体表达文库。例如抗 FLAG 抗体的合适的标记技术，将提供合适克隆的特异性标记。

本文引用的所有参考文献均通过引用结合到本文中，其程度与具体和单独地表明每个单独的出版物或专利申请通过引用结合到本文中的程度一样。

不偏离本发明的精神和范围，可以进行本发明的许多修改和变化，这对本领域技术人员是显而易见的。提供本文所述的具体实施方案仅仅是作为实例，本发明由所附的权利要求书以及赋予这类权利要求的相当权利要求的全部范围所限制；本发明不受作为实例提出的具体实施方案的限制。

序列表

(1) 一般资料

(i) 申请人:

- (A) 姓名: Schering Corporation
- (B) 街道: 2000 Galloping Hill Road
- (C) 城市: Kenilworth
- (D) 州: New Jersey
- (E) 国家: 美国
- (F) 邮政编码: 07033
- (G) 电话: (908) 298-4000
- (H) 传真: (908) 298-5388

(ii) 发明名称: 人受体蛋白; 相关的药剂和方法

(iii) 序列数: 35

(v) 计算机可读形式

- (A) 媒体类型: 软盘
- (B) 计算机: Macintosh Power PC
- (C) 操作系统: 8.0
- (D) 软件: Microsoft Word 6.0

(vi) 当前申请数据

- (A) 申请号:
- (B) 提交日期:
- (C) 分类:

(vii) 在先申请数据

- (A) 申请号: USSN 60/044, 293
- (B) 提交日期: 1997 年 5 月 7 日

- (A) 申请号: USSN 60/072, 212
- (B) 提交日期: 1998 年 1 月 22 日

- (A) 申请号: USSN 60/076, 947
- (B) 提交日期: 1998 年 3 月 5 日

(2) SEQ ID NO: 1 的信息:

(i) 顺序特征:

- (A) 长度: 2367 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

(A) 名称/关键词: CDS

(B) 位置: 1...2358

(ix) 特征:

(A) 名称/关键词: 成熟肽

(B) 位置: 67...2358

(xi) 顺序描述: SEQ ID NO: 1:

ATG ACT AGC ATC TTC CAT TTT GCC ATT ATC TTC ATG TTA ATA CTT CAG	48
Met Thr Ser Ile Phe His Phe Ala Ile Ile Phe Met Leu Ile Leu Gln	
-22 -20 -15 -10	
ATC AGA ATA CAA TTA TCT GAA GAA AGT GAA TTT TTA GTT GAT AGG TCA	96
Ile Arg Ile Gln Leu Ser Glu Glu Ser Glu Phe Leu Val Asp Arg Ser	
-5 1 5 10	
AAA AAC GGT CTC ATC CAC GTT CCT AAA GAC CTA TCC CAG AAA ACA ACA	144
Lys Asn Gly Leu Ile His Val Pro Lys Asp Leu Ser Gln Lys Thr Thr	
15 20 25	
ATC TTA AAT ATA TCG CAA AAT TAT ATA TCT GAG CTT TGG ACT TCT GAC	192
Ile Leu Asn Ile Ser Gln Asn Tyr Ile Ser Glu Leu Trp Thr Ser Asp	
30 35 40	
ATC TTA TCA CTG TCA AAA CTG AGG ATT TTG ATA ATT TCT CAT AAT AGA	240
Ile Leu Ser Leu Ser Lys Leu Arg Ile Leu Ile Ile Ser His Asn Arg	
45 50 55	
ATC CAG TAT CTT GAT ATC AGT GTT TTC AAA TTC AAC CAG GAA TTG GAA	288
Ile Gln Tyr Leu Asp Ile Ser Val Phe Lys Phe Asn Gln Glu Leu Glu	
60 65 70	
TAC TTG GAT TTG TCC CAC AAC AAG TTG GTG AAG ATT TCT TGC CAC CCT	336
Tyr Leu Asp Leu Ser His Asn Lys Leu Val Lys Ile Ser Cys His Pro	
75 80 85 90	
ACT GTG AAC CTC AAG CAC TTG GAC CTG TCA TTT AAT GCA TTT GAT GCC	384
Thr Val Asn Leu Lys His Leu Asp Leu Ser Phe Asn Ala Phe Asp Ala	
95 100 105	
CTG CCT ATA TGC AAA GAG TTT GGC AAT ATG TCT CAA CTA AAA TTT CTG	432
Leu Pro Ile Cys Lys Glu Phe Gly Asn Met Ser Gln Leu Lys Phe Leu	
110 115 120	
GGG TTG AGC ACC ACA CAC TTA GAA AAA TCT AGT GTG CTG CCA ATT GCT	480
Gly Leu Ser Thr Thr His Leu Glu Lys Ser Ser Val Leu Pro Ile Ala	
125 130 135	

CAT TTG AAT ATC AGC AAG GTC TTG CTG GTC TTA GGA GAG ACT TAT GGG His Leu Asn Ile Ser Lys Val Leu Leu Val Leu Gly Glu Thr Tyr Gly 140 145 150	528
GAA AAA GAA GAC CCT GAG GGC CTT CAA GAC TTT AAC ACT GAG AGT CTG Glu Lys Glu Asp Pro Glu Gly Leu Gln Asp Phe Asn Thr Glu Ser Leu 155 160 165 170	576
CAC ATT GTG TTC CCC ACA AAC AAA GAA TTC CAT TTT ATT TTG GAT GTG His Ile Val Phe Pro Thr Asn Lys Glu Phe His Phe Ile Leu Asp Val 175 180 185	624
TCA GTC AAG ACT GTA GCA AAT CTG GAA CTA TCT AAT ATC AAA TGT GTG Ser Val Lys Thr Val Ala Asn Leu Glu Leu Ser Asn Ile Lys Cys Val 190 195 200	672
CTA GAA GAT AAC AAA TGT TCT TAC TTC CTA AGT ATT CTG GCG AAA CTT Leu Glu Asp Asn Lys Cys Ser Tyr Phe Leu Ser Ile Leu Ala Lys Leu 205 210 215	720
CAA ACA AAT CCA AAG TTA TCA AGT CTT ACC TTA AAC AAC ATT GAA ACA Gln Thr Asn Pro Lys Leu Ser Ser Leu Thr Leu Asn Asn Ile Glu Thr 220 225 230	768
ACT TGG AAT TCT TTC ATT AGG ATC CTC CAA CTA GTT TGG CAT ACA ACT Thr Trp Asn Ser Phe Ile Arg Ile Leu Gln Leu Val Trp His Thr Thr 235 240 245 250	816
GTA TGG TAT TTC TCA ATT TCA AAC GTG AAG CTA CAG GGT CAG CTG GAC Val Trp Tyr Phe Ser Ile Ser Asn Val Lys Leu Gln Gly Gln Leu Asp 255 260 265	864
TTC AGA GAT TTT GAT TAT TCT GGC ACT TCC TTG AAG GCC TTG TCT ATA Phe Arg Asp Phe Asp Tyr Ser Gly Thr Ser Leu Lys Ala Leu Ser Ile 270 275 280	912
CAC CAA GTT GTC AGC GAT GTG TTC GGT TTT CCG CAA AGT TAT ATC TAT His Gln Val Val Ser Asp Val Phe Gly Phe Pro Gln Ser Tyr Ile Tyr 285 290 295	960
GAA ATC TTT TCG AAT ATG AAC ATC AAA AAT TTC ACA GTG TCT GGT ACA Glu Ile Phe Ser Asn Met Asn Ile Lys Asn Phe Thr Val Ser Gly Thr 300 305 310	1008
CGC ATG GTC CAC ATG CTT TGC CCA TCC AAA ATT AGC CCG TTC CTG CAT Arg Met Val His Met Leu Cys Pro Ser Lys Ile Ser Pro Phe Leu His 315 320 325 330	1056

TTG GAT TTT TCC AAT AAT CTC TTA ACA GAC ACG GTT TTT GAA AAT TGT Leu Asp Phe Ser Asn Asn Leu Leu Thr Asp Thr Val Phe Glu Asn Cys 335 340 345	1104
GGG CAC CTT ACT GAG TTG GAG ACA CTT ATT TTA CAA ATG AAT CAA TTA Gly His Leu Thr Glu Leu Glu Thr Leu Ile Leu Gln Met Asn Gln Leu 350 355 360	1152
AAA GAA CTT TCA AAA ATA GCT GAA ATG ACT ACA CAG ATG AAG TCT CTG Lys Glu Leu Ser Lys Ile Ala Glu Met Thr Thr Gln Met Lys Ser Leu 365 370 375	1200
CAA CAA TTG GAT ATT AGC CAG AAT TCT GTA AGC TAT GAT GAA AAG AAA Gln Gln Leu Asp Ile Ser Gln Asn Ser Val Ser Tyr Asp Glu Lys Lys 380 385 390	1248
GGA GAC TGT TCT TGG ACT AAA AGT TTA TTA AGT TTA AAT ATG TCT TCA Gly Asp Cys Ser Trp Thr Lys Ser Leu Leu Ser Leu Asn Met Ser Ser 395 400 405 410	1296
AAT ATA CTT ACT GAC ACT ATT TTC AGA TGT TTA CCT CCC AGG ATC AAG Asn Ile Leu Thr Asp Thr Ile Phe Arg Cys Leu Pro Pro Arg Ile Lys 415 420 425	1344
GTA CTT GAT CTT CAC AGC AAT AAA ATA AAG AGC ATT CCT AAA CAA GTC Val Leu Asp Leu His Ser Asn Lys Ile Lys Ser Ile Pro Lys Gln Val 430 435 440	1392
GTA AAA CTG GAA GCT TTG CAA GAA CTC AAT GTT GCT TTC AAT TCT TTA Val Lys Leu Glu Ala Leu Gln Glu Leu Asn Val Ala Phe Asn Ser Leu 445 450 455	1440
ACT GAC CTT CCT GGA TGT GGC AGC TTT AGC AGC CTT TCT GTA TTG ATC Thr Asp Leu Pro Gly Cys Gly Ser Phe Ser Ser Leu Ser Val Leu Ile 460 465 470	1488
ATT GAT CAC AAT TCA GTT TCC CAC CCA TCA GCT GAT TTC TTC CAG AGC Ile Asp His Asn Ser Val Ser His Pro Ser Ala Asp Phe Phe Gln Ser 475 480 485 490	1536
TGC CAG AAG ATG AGG TCA ATA AAA GCA GGG GAC AAT CCA TTC CAA TGT Cys Gln Lys Met Arg Ser Ile Lys Ala Gly Asp Asn Pro Phe Gln Cys 495 500 505	1584
ACC TGT GAG CTC GGA GAA TTT GTC AAA AAT ATA GAC CAA GTA TCA AGT Thr Cys Glu Leu Gly Glu Phe Val Lys Asn Ile Asp Gln Val Ser Ser 510 515 520	1632

GAA GTG TTA GAG GGC TGG CCT GAT TCT TAT AAG TGT GAC TAC CCG GAA Glu Val Leu Glu Gly Trp Pro Asp Ser Tyr Lys Cys Asp Tyr Pro Glu 525 530 535	1680
AGT TAT AGA GGA ACC CTA CTA AAG GAC TTT CAC ATG TCT GAA TTA TCC Ser Tyr Arg Gly Thr Leu Leu Lys Asp Phe His Met Ser Glu Leu Ser 540 545 550	1728
TGC AAC ATA ACT CTG CTG ATC GTC ACC ATC GTT GCC ACC ATG CTG GTG Cys Asn Ile Thr Leu Leu Ile Val Thr Ile Val Ala Thr Met Leu Val 555 560 565 570	1776
TTG GCT GTG ACT GTG ACC TCC CTC TGC ATC TAC TTG GAT CTG CCC TGG Leu Ala Val Thr Val Thr Ser Leu Cys Ile Tyr Leu Asp Leu Pro Trp 575 580 585	1824
TAT CTC AGG ATG GTG TGC CAG TGG ACC CAG ACC CGG CGC AGG GCC AGG Tyr Leu Arg Met Val Cys Gln Trp Thr Gln Thr Arg Arg Arg Ala Arg 590 595 600	1872
AAC ATA CCC TTA GAA GAA CTC CAA AGA AAT CTC CAG TTT CAT GCA TTT Asn Ile Pro Leu Glu Glu Leu Gln Arg Asn Leu Gln Phe His Ala Phe 605 610 615	1920
ATT TCA TAT AGT GGG CAC GAT TCT TTC TGG GTG AAG AAT GAA TTA TTG Ile Ser Tyr Ser Gly His Asp Ser Phe Trp Val Lys Asn Glu Leu Leu 620 625 630	1968
CCA AAC CTA GAG AAA GAA GGT ATG CAG ATT TGC CTT CAT GAG AGA AAC Pro Asn Leu Glu Lys Glu Gly Met Gln Ile Cys Leu His Glu Arg Asn 635 640 645 650	2016
TTT GTT CCT GGC AAG AGC ATT GTG GAA AAT ATC ATC ACC TGC ATT GAG Phe Val Pro Gly Lys Ser Ile Val Glu Asn Ile Ile Thr Cys Ile Glu 655 660 665	2064
AAG AGT TAC AAG TCC ATC TTT GTT TTG TCT CCC AAC TTT GTC CAG AGT Lys Ser Tyr Lys Ser Ile Phe Val Leu Ser Pro Asn Phe Val Gln Ser 670 675 680	2112
GAA TGG TGC CAT TAT GAA CTC TAC TTT GCC CAT CAC AAT CTC TTT CAT Glu Trp Cys His Tyr Glu Leu Tyr Phe Ala His His Asn Leu Phe His 685 690 695	2160
GAA GGA TCT AAT AGC TTA ATC CTG ATC TTG CTG GAA CCC ATT CCG CAG Glu Gly Ser Asn Ser Leu Ile Leu Ile Leu Leu Glu Pro Ile Pro Gln 700 705 710	2208

```

TAC TCC ATT CCT AGC AGT TAT CAC AAG CTC AAA AGT CTC ATG GCC AGG      2256
Tyr Ser Ile Pro Ser Ser Tyr His Lys Leu Lys Ser Leu Met Ala Arg
715                      720                      725                      730

AGG ACT TAT TTG GAA TGG CCC AAG GAA AAG AGC AAA CGT GGC CTT TTT      2304
Arg Thr Tyr Leu Glu Trp Pro Lys Glu Lys Ser Lys Arg Gly Leu Phe
                      735                      740                      745

TGG GCT AAC TTA AGG GCA GCC ATT AAT ATT AAG CTG ACA GAG CAA GCA      2352
Trp Ala Asn Leu Arg Ala Ala Ile Asn Ile Lys Leu Thr Glu Gln Ala
                      750                      755                      760

AAG AAA TAGTCTAGA      2367
Lys Lys

```

(2) SEQ ID NO: 2 的信息:

- (i) 顺序特征:
- (A) 长度: 786 个氨基酸
 - (B) 类型: 氨基酸
 - (D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 顺序描述: SEQ ID NO: 2:

```

Met Thr Ser Ile Phe His Phe Ala Ile Ile Phe Met Leu Ile Leu Gln
-22      -20                      -15                      -10

Ile Arg Ile Gln Leu Ser Glu Glu Ser Glu Phe Leu Val Asp Arg Ser
      -5                      1                      5                      10

Lys Asn Gly Leu Ile His Val Pro Lys Asp Leu Ser Gln Lys Thr Thr
                      15                      20                      25

Ile Leu Asn Ile Ser Gln Asn Tyr Ile Ser Glu Leu Trp Thr Ser Asp
      30                      35                      40

Ile Leu Ser Leu Ser Lys Leu Arg Ile Leu Ile Ile Ser His Asn Arg
      45                      50                      55

Ile Gln Tyr Leu Asp Ile Ser Val Phe Lys Phe Asn Gln Glu Leu Glu
      60                      65                      70

Tyr Leu Asp Leu Ser His Asn Lys Leu Val Lys Ile Ser Cys His Pro
      75                      80                      85                      90

```

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

Gly His Leu Thr Glu Leu Glu Thr Leu Ile Leu Gln Met Asn Gln Leu
 350 355 360
 Lys Glu Leu Ser Lys Ile Ala Glu Met Thr Thr Gln Met Lys Ser Leu
 365 370 375
 Gln Gln Leu Asp Ile Ser Gln Asn Ser Val Ser Tyr Asp Glu Lys Lys
 380 385 390
 Gly Asp Cys Ser Trp Thr Lys Ser Leu Leu Ser Leu Asn Met Ser Ser
 395 400 405 410
 Asn Ile Leu Thr Asp Thr Ile Phe Arg Cys Leu Pro Pro Arg Ile Lys
 415 420 425
 Val Leu Asp Leu His Ser Asn Lys Ile Lys Ser Ile Pro Lys Gln Val
 430 435 440
 Val Lys Leu Glu Ala Leu Gln Glu Leu Asn Val Ala Phe Asn Ser Leu
 445 450 455
 Thr Asp Leu Pro Gly Cys Gly Ser Phe Ser Ser Leu Ser Val Leu Ile
 460 465 470
 Ile Asp His Asn Ser Val Ser His Pro Ser Ala Asp Phe Phe Gln Ser
 475 480 485 490
 Cys Gln Lys Met Arg Ser Ile Lys Ala Gly Asp Asn Pro Phe Gln Cys
 495 500 505
 Thr Cys Glu Leu Gly Glu Phe Val Lys Asn Ile Asp Gln Val Ser Ser
 510 515 520
 Glu Val Leu Glu Gly Trp Pro Asp Ser Tyr Lys Cys Asp Tyr Pro Glu
 525 530 535
 Ser Tyr Arg Gly Thr Leu Leu Lys Asp Phe His Met Ser Glu Leu Ser
 540 545 550
 Cys Asn Ile Thr Leu Leu Ile Val Thr Ile Val Ala Thr Met Leu Val
 555 560 565 570
 Leu Ala Val Thr Val Thr Ser Leu Cys Ile Tyr Leu Asp Leu Pro Trp
 575 580 585
 Tyr Leu Arg Met Val Cys Gln Trp Thr Gln Thr Arg Arg Arg Ala Arg
 590 595 600

```

Asn Ile Pro Leu Glu Glu Leu Gln Arg Asn Leu Gln Phe His Ala Phe
    605                      610                      615

Ile Ser Tyr Ser Gly His Asp Ser Phe Trp Val Lys Asn Glu Leu Leu
    620                      625                      630

Pro Asn Leu Glu Lys Glu Gly Met Gln Ile Cys Leu His Glu Arg Asn
    635                      640                      645                      650

Phe Val Pro Gly Lys Ser Ile Val Glu Asn Ile Ile Thr Cys Ile Glu
    655                      660                      665

Lys Ser Tyr Lys Ser Ile Phe Val Leu Ser Pro Asn Phe Val Gln Ser
    670                      675                      680

Glu Trp Cys His Tyr Glu Leu Tyr Phe Ala His His Asn Leu Phe His
    685                      690                      695

Glu Gly Ser Asn Ser Leu Ile Leu Ile Leu Leu Glu Pro Ile Pro Gln
    700                      705                      710

Tyr Ser Ile Pro Ser Ser Tyr His Lys Leu Lys Ser Leu Met Ala Arg
    715                      720                      725                      730

Arg Thr Tyr Leu Glu Trp Pro Lys Glu Lys Ser Lys Arg Gly Leu Phe
    735                      740                      745

Trp Ala Asn Leu Arg Ala Ala Ile Asn Ile Lys Leu Thr Glu Gln Ala
    750                      755                      760

Lys Lys

```

(2) SEQ ID NO: 3 的信息:

- (i) 顺序特征:
 - (A) 长度: 2355 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线性
- (ii) 分子类型: cDNA
- (ix) 特征:
 - (A) 名称/关键词: CDS
 - (B) 位置: 1...2352
- (ix) 特征:

GAG GAA CTT GAG ATT GAT GCT TCA GAT CTA CAG AGC TAT GAG CCA AAA Glu Glu Leu Glu Ile Asp Ala Ser Asp Leu Gln Ser Tyr Glu Pro Lys 155 160 165 170	576
AGT TTG AAG TCA ATT CAG AAC GTA AGT CAT CTG ATC CTT CAT ATG AAG Ser Leu Lys Ser Ile Gln Asn Val Ser His Leu Ile Leu His Met Lys 175 180 185	624
CAG CAT ATT TTA CTG CTG GAG ATT TTT GTA GAT GTT ACA AGT TCC GTG Gln His Ile Leu Leu Leu Glu Ile Phe Val Asp Val Thr Ser Ser Val 190 195 200	672
GAA TGT TTG GAA CTG CGA GAT ACT GAT TTG GAC ACT TTC CAT TTT TCA Glu Cys Leu Glu Leu Arg Asp Thr Asp Leu Asp Thr Phe His Phe Ser 205 210 215	720
GAA CTA TCC ACT GGT GAA ACA AAT TCA TTG ATT AAA AAG TTT ACA TTT Glu Leu Ser Thr Gly Glu Thr Asn Ser Leu Ile Lys Lys Phe Thr Phe 220 225 230	768
AGA AAT GTG AAA ATC ACC GAT GAA AGT TTG TTT CAG GTT ATG AAA CTT Arg Asn Val Lys Ile Thr Asp Glu Ser Leu Phe Gln Val Met Lys Leu 235 240 245 250	816
TTG AAT CAG ATT TCT GGA TTG TTA GAA TTA GAG TTT GAT GAC TGT ACC Leu Asn Gln Ile Ser Gly Leu Leu Glu Leu Glu Phe Asp Asp Cys Thr 255 260 265	864
CTT AAT GGA GTT GGT AAT TTT AGA GCA TCT GAT AAT GAC AGA GTT ATA Leu Asn Gly Val Gly Asn Phe Arg Ala Ser Asp Asn Asp Arg Val Ile 270 275 280	912
GAT CCA GGT AAA GTG GAA ACG TTA ACA ATC CGG AGG CTG CAT ATT CCA Asp Pro Gly Lys Val Glu Thr Leu Thr Ile Arg Arg Leu His Ile Pro 285 290 295	960
AGG TTT TAC TTA TTT TAT GAT CTG AGC ACT TTA TAT TCA CTT ACA GAA Arg Phe Tyr Leu Phe Tyr Asp Leu Ser Thr Leu Tyr Ser Leu Thr Glu 300 305 310	1008
AGA GTT AAA AGA ATC ACA GTA GAA AAC AGT AAA GTT TTT CTG GTT CCT Arg Val Lys Arg Ile Thr Val Glu Asn Ser Lys Val Phe Leu Val Pro 315 320 325 330	1056
TGT TTA CTT TCA CAA CAT TTA AAA TCA TTA GAA TAC TTG GAT CTC AGT Cys Leu Leu Ser Gln His Leu Lys Ser Leu Glu Tyr Leu Asp Leu Ser 335 340 345	1104

GAA AAT TTG ATG GTT GAA GAA TAC TTG AAA AAT TCA GCC TGT GAG GAT Glu Asn Leu Met Val Glu Glu Tyr Leu Lys Asn Ser Ala Cys Glu Asp 350 355 360	1152
GCC TGG CCC TCT CTA CAA ACT TTA ATT TTA AGG CAA AAT CAT TTG GCA Ala Trp Pro Ser Leu Gln Thr Leu Ile Leu Arg Gln Asn His Leu Ala 365 370 375	1200
TCA TTG GAA AAA ACC GGA GAG ACT TTG CTC ACT CTG AAA AAC TTG ACT Ser Leu Glu Lys Thr Gly Glu Thr Leu Leu Thr Leu Lys Asn Leu Thr 380 385 390	1248
AAC ATT GAT ATC AGT AAG AAT AGT TTT CAT TCT ATG CCT GAA ACT TGT Asn Ile Asp Ile Ser Lys Asn Ser Phe His Ser Met Pro Glu Thr Cys 395 400 405 410	1296
CAG TGG CCA GAA AAG ATG AAA TAT TTG AAC TTA TCC AGC ACA CGA ATA Gln Trp Pro Glu Lys Met Lys Tyr Leu Asn Leu Ser Ser Thr Arg Ile 415 420 425	1344
CAC AGT GTA ACA GGC TGC ATT CCC AAG ACA CTG GAA ATT TTA GAT GTT His Ser Val Thr Gly Cys Ile Pro Lys Thr Leu Glu Ile Leu Asp Val 430 435 440	1392
AGC AAC AAC AAT CTC AAT TTA TTT TCT TTG AAT TTG CCG CAA CTC AAA Ser Asn Asn Asn Leu Asn Leu Phe Ser Leu Asn Leu Pro Gln Leu Lys 445 450 455	1440
GAA CTT TAT ATT TCC AGA AAT AAG TTG ATG ACT CTA CCA GAT GCC TCC Glu Leu Tyr Ile Ser Arg Asn Lys Leu Met Thr Leu Pro Asp Ala Ser 460 465 470	1488
CTC TTA CCC ATG TTA CTA GTA TTG AAA ATC AGT AGG AAT GCA ATA ACT Leu Leu Pro Met Leu Leu Val Leu Lys Ile Ser Arg Asn Ala Ile Thr 475 480 485 490	1536
ACG TTT TCT AAG GAG CAA CTT GAC TCA TTT CAC ACA CTG AAG ACT TTG Thr Phe Ser Lys Glu Gln Leu Asp Ser Phe His Thr Leu Lys Thr Leu 495 500 505	1584
GAA GCT GGT GGC AAT AAC TTC ATT TGC TCC TGT GAA TTC CTC TCC TTC Glu Ala Gly Gly Asn Asn Phe Ile Cys Ser Cys Glu Phe Leu Ser Phe 510 515 520	1632
ACT CAG GAG CAG CAA GCA CTG GCC AAA GTC TTG ATT GAT TGG CCA GCA Thr Gln Glu Gln Gln Ala Leu Ala Lys Val Leu Ile Asp Trp Pro Ala 525 530 535	1680

AAT TAC CTG TGT GAC TCT CCA TCC CAT GTG CGT GGC CAG CAG GTT CAG Asn Tyr Leu Cys Asp Ser Pro Ser His Val Arg Gly Gln Gln Val Gln 540 545 550	1728
GAT GTC CGC CTC TCG GTG TCG GAA TGT CAC AGG ACA GCA CTG GTG TCT Asp Val Arg Leu Ser Val Ser Glu Cys His Arg Thr Ala Leu Val Ser 555 560 565 570	1776
GGC ATG TGC TGT GCT CTG TTC CTG CTG ATC CTG CTC ACG GGG GTC CTG Gly Met Cys Cys Ala Leu Phe Leu Leu Ile Leu Leu Thr Gly Val Leu 575 580 585	1824
TGC CAC CGT TTC CAT GGC CTG TGG TAT ATG AAA ATG ATG TGG GCC TGG Cys His Arg Phe His Gly Leu Trp Tyr Met Lys Met Met Trp Ala Trp 590 595 600	1872
CTC CAG GCC AAA AGG AAG CCC AGG AAA GCT CCC AGC AGG AAC ATC TGC Leu Gln Ala Lys Arg Lys Pro Arg Lys Ala Pro Ser Arg Asn Ile Cys 605 610 615	1920
TAT GAT GCA TTT GTT TCT TAC AGT GAG CGG GAT GCC TAC TGG GTG GAG Tyr Asp Ala Phe Val Ser Tyr Ser Glu Arg Asp Ala Tyr Trp Val Glu 620 625 630	1968
AAC CTT ATG GTC CAG GAG CTG GAG AAC TTC AAT CCC CCC TTC AAG TTG Asn Leu Met Val Gln Glu Leu Glu Asn Phe Asn Pro Pro Phe Lys Leu 635 640 645 650	2016
TGT CTT CAT AAG CGG GAC TTC ATT CCT GGC AAG TGG ATC ATT GAC AAT Cys Leu His Lys Arg Asp Phe Ile Pro Gly Lys Trp Ile Ile Asp Asn 655 660 665	2064
ATC ATT GAC TCC ATT GAA AAG AGC CAC AAA ACT GTC TTT GTG CTT TCT Ile Ile Asp Ser Ile Glu Lys Ser His Lys Thr Val Phe Val Leu Ser 670 675 680	2112
GAA AAC TTT GTG AAG AGT GAG TGG TGC AAG TAT GAA CTG GAC TTC TCC Glu Asn Phe Val Lys Ser Glu Trp Cys Lys Tyr Glu Leu Asp Phe Ser 685 690 695	2160
CAT TTC CGT CTT TTT GAA GAG AAC AAT GAT GCT GCC ATT CTC ATT CTT His Phe Arg Leu Phe Glu Glu Asn Asn Asp Ala Ala Ile Leu Ile Leu 700 705 710	2208
CTG GAG CCC ATT GAG AAA AAA GCC ATT CCC CAG CGC TTC TGC AAG CTG Leu Glu Pro Ile Glu Lys Lys Ala Ile Pro Gln Arg Phe Cys Lys Leu 715 720 725 730	2256

```

CGG AAG ATA ATG AAC ACC AAG ACC TAC CTG GAG TGG CCC ATG GAC GAG      2304
Arg Lys Ile Met Asn Thr Lys Thr Tyr Leu Glu Trp Pro Met Asp Glu
          735                740                745

GCT CAG CGG GAA GGA TTT TGG GTA AAT CTG AGA GCT GCG ATA AAG TCC      2352
Ala Gln Arg Glu Gly Phe Trp Val Asn Leu Arg Ala Ala Ile Lys Ser
          750                755                760

TAG                                                                    2355

```

(2) SEQ ID NO: 4 的信息:

(i) 顺序特征:

(A) 长度: 784 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 顺序描述: SEQ ID NO: 4:

```

Met Pro His Thr Leu Trp Met Val Trp Val Leu Gly Val Ile Ile Ser
-22   -20               -15               -10

```

```

Leu Ser Lys Glu Glu Ser Ser Asn Gln Ala Ser Leu Ser Cys Asp Arg
  -5               1               5               10

```

```

Asn Gly Ile Cys Lys Gly Ser Ser Gly Ser Leu Asn Ser Ile Pro Ser
      15               20               25

```

```

Gly Leu Thr Glu Ala Val Lys Ser Leu Asp Leu Ser Asn Asn Arg Ile
      30               35               40

```

```

Thr Tyr Ile Ser Asn Ser Asp Leu Gln Arg Cys Val Asn Leu Gln Ala
      45               50               55

```

```

Leu Val Leu Thr Ser Asn Gly Ile Asn Thr Ile Glu Glu Asp Ser Phe
      60               65               70

```

```

Ser Ser Leu Gly Ser Leu Glu His Leu Asp Leu Ser Tyr Asn Tyr Leu
      75               80               85               90

```

```

Ser Asn Leu Ser Ser Ser Trp Phe Lys Pro Leu Ser Ser Leu Thr Phe
      95               100              105

```

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

Ala Trp Pro Ser Leu Gln Thr Leu Ile Leu Arg Gln Asn His Leu Ala
 365 370 375
 Ser Leu Glu Lys Thr Gly Glu Thr Leu Leu Thr Leu Lys Asn Leu Thr
 380 385 390
 Asn Ile Asp Ile Ser Lys Asn Ser Phe His Ser Met Pro Glu Thr Cys
 395 400 405 410
 Gln Trp Pro Glu Lys Met Lys Tyr Leu Asn Leu Ser Ser Thr Arg Ile
 415 420 425
 His Ser Val Thr Gly Cys Ile Pro Lys Thr Leu Glu Ile Leu Asp Val
 430 435 440
 Ser Asn Asn Asn Leu Asn Leu Phe Ser Leu Asn Leu Pro Gln Leu Lys
 445 450 455
 Glu Leu Tyr Ile Ser Arg Asn Lys Leu Met Thr Leu Pro Asp Ala Ser
 460 465 470
 Leu Leu Pro Met Leu Leu Val Leu Lys Ile Ser Arg Asn Ala Ile Thr
 475 480 485 490
 Thr Phe Ser Lys Glu Gln Leu Asp Ser Phe His Thr Leu Lys Thr Leu
 495 500 505
 Glu Ala Gly Gly Asn Asn Phe Ile Cys Ser Cys Glu Phe Leu Ser Phe
 510 515 520
 Thr Gln Glu Gln Gln Ala Leu Ala Lys Val Leu Ile Asp Trp Pro Ala
 525 530 535
 Asn Tyr Leu Cys Asp Ser Pro Ser His Val Arg Gly Gln Gln Val Gln
 540 545 550
 Asp Val Arg Leu Ser Val Ser Glu Cys His Arg Thr Ala Leu Val Ser
 555 560 565 570
 Gly Met Cys Cys Ala Leu Phe Leu Leu Ile Leu Leu Thr Gly Val Leu
 575 580 585
 Cys His Arg Phe His Gly Leu Trp Tyr Met Lys Met Met Trp Ala Trp
 590 595 600
 Leu Gln Ala Lys Arg Lys Pro Arg Lys Ala Pro Ser Arg Asn Ile Cys
 605 610 615

Tyr Asp Ala Phe Val Ser Tyr Ser Glu Arg Asp Ala Tyr Trp Val Glu
620 625 630

Asn Leu Met Val Gln Glu Leu Glu Asn Phe Asn Pro Pro Phe Lys Leu
635 640 645 650

Cys Leu His Lys Arg Asp Phe Ile Pro Gly Lys Trp Ile Ile Asp Asn
655 660 665

Ile Ile Asp Ser Ile Glu Lys Ser His Lys Thr Val Phe Val Leu Ser
670 675 680

Glu Asn Phe Val Lys Ser Glu Trp Cys Lys Tyr Glu Leu Asp Phe Ser
685 690 695

His Phe Arg Leu Phe Glu Glu Asn Asn Asp Ala Ala Ile Leu Ile Leu
700 705 710

Leu Glu Pro Ile Glu Lys Lys Ala Ile Pro Gln Arg Phe Cys Lys Leu
715 720 725 730

Arg Lys Ile Met Asn Thr Lys Thr Tyr Leu Glu Trp Pro Met Asp Glu
735 740 745

Ala Gln Arg Glu Gly Phe Trp Val Asn Leu Arg Ala Ala Ile Lys Ser
750 755 760

(2) SEQ ID NO: 5 的信息:

- (i) 顺序特征:
- (A) 长度: 2715 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线性

- (ii) 分子类型: cDNA

- (ix) 特征:
- (A) 名称/关键词: CDS
 - (B) 位置: 1...2712

- (ix) 特征:
- (A) 名称/关键词: 成熟肽
 - (B) 位置: 64...2712

- (xi) 顺序描述: SEQ ID NO: 5:

ATG AGA CAG ACT TTG CCT TGT ATC TAC TTT TGG GGG GGC CTT TTG CCC Met Arg Gln Thr Leu Pro Cys Ile Tyr Phe Trp Gly Gly Leu Leu Pro -21 -20 -15 -10	48
TTT GGG ATG CTG TGT GCA TCC TCC ACC ACC AAG TGC ACT GTT AGC CAT Phe Gly Met Leu Cys Ala Ser Ser Thr Thr Lys Cys Thr Val Ser His -5 1 5 10	96
GAA GTT GCT GAC TGC AGC CAC CTG AAG TTG ACT CAG GTA CCC GAT GAT Glu Val Ala Asp Cys Ser His Leu Lys Leu Thr Gln Val Pro Asp Asp 15 20 25	144
CTA CCC ACA AAC ATA ACA GTG TTG AAC CTT ACC CAT AAT CAA CTC AGA Leu Pro Thr Asn Ile Thr Val Leu Asn Leu Thr His Asn Gln Leu Arg 30 35 40	192
AGA TTA CCA GCC GCC AAC TTC ACA AGG TAT AGC CAG CTA ACT AGC TTG Arg Leu Pro Ala Ala Asn Phe Thr Arg Tyr Ser Gln Leu Thr Ser Leu 45 50 55	240
GAT GTA GGA TTT AAC ACC ATC TCA AAA CTG GAG CCA GAA TTG TGC CAG Asp Val Gly Phe Asn Thr Ile Ser Lys Leu Glu Pro Glu Leu Cys Gln 60 65 70 75	288
AAA CTT CCC ATG TTA AAA GTT TTG AAC CTC CAG CAC AAT GAG CTA TCT Lys Leu Pro Met Leu Lys Val Leu Asn Leu Gln His Asn Glu Leu Ser 80 85 90	336
CAA CTT TCT GAT AAA ACC TTT GCC TTC TGC ACG AAT TTG ACT GAA CTC Gln Leu Ser Asp Lys Thr Phe Ala Phe Cys Thr Asn Leu Thr Glu Leu 95 100 105	384
CAT CTC ATG TCC AAC TCA ATC CAG AAA ATT AAA AAT AAT CCC TTT GTC His Leu Met Ser Asn Ser Ile Gln Lys Ile Lys Asn Asn Pro Phe Val 110 115 120	432
AAG CAG AAG AAT TTA ATC ACA TTA GAT CTG TCT CAT AAT GGC TTG TCA Lys Gln Lys Asn Leu Ile Thr Leu Asp Leu Ser His Asn Gly Leu Ser 125 130 135	480
TCT ACA AAA TTA GGA ACT CAG GTT CAG CTG GAA AAT CTC CAA GAG CTT Ser Thr Lys Leu Gly Thr Gln Val Gln Leu Glu Asn Leu Gln Glu Leu 140 145 150 155	528
CTA TTA TCA AAC AAT AAA ATT CAA GCG CTA AAA AGT GAA GAA CTG GAT Leu Leu Ser Asn Asn Lys Ile Gln Ala Leu Lys Ser Glu Glu Leu Asp 160 165 170	576

ATC TTT GCC AAT TCA TCT TTA AAA AAA TTA GAG TTG TCA TCG AAT CAA Ile Phe Ala Asn Ser Ser Leu Lys Lys Leu Glu Leu Ser Ser Asn Gln 175 180 185	624
ATT AAA GAG TTT TCT CCA GGG TGT TTT CAC GCA ATT GGA AGA TTA TTT Ile Lys Glu Phe Ser Pro Gly Cys Phe His Ala Ile Gly Arg Leu Phe 190 195 200	672
GGC CTC TTT CTG AAC AAT GTC CAG CTG GGT CCC AGC CTT ACA GAG AAG Gly Leu Phe Leu Asn Asn Val Gln Leu Gly Pro Ser Leu Thr Glu Lys 205 210 215	720
CTA TGT TTG GAA TTA GCA AAC ACA AGC ATT CGG AAT CTG TCT CTG AGT Leu Cys Leu Glu Leu Ala Asn Thr Ser Ile Arg Asn Leu Ser Leu Ser 220 225 230 235	768
AAC AGC CAG CTG TCC ACC ACC AGC AAT ACA ACT TTC TTG GGA CTA AAG Asn Ser Gln Leu Ser Thr Thr Ser Asn Thr Thr Phe Leu Gly Leu Lys 240 245 250	816
TGG ACA AAT CTC ACT ATG CTC GAT CTT TCC TAC AAC AAC TTA AAT GTG Trp Thr Asn Leu Thr Met Leu Asp Leu Ser Tyr Asn Asn Leu Asn Val 255 260 265	864
GTT GGT AAC GAT TCC TTT GCT TGG CTT CCA CAA CTA GAA TAT TTC TTC Val Gly Asn Asp Ser Phe Ala Trp Leu Pro Gln Leu Glu Tyr Phe Phe 270 275 280	912
CTA GAG TAT AAT AAT ATA CAG CAT TTG TTT TCT CAC TCT TTG CAC GGG Leu Glu Tyr Asn Asn Ile Gln His Leu Phe Ser His Ser Leu His Gly 285 290 295	960
CTT TTC AAT GTG AGG TAC CTG AAT TTG AAA CGG TCT TTT ACT AAA CAA Leu Phe Asn Val Arg Tyr Leu Asn Leu Lys Arg Ser Phe Thr Lys Gln 300 305 310 315	1008
AGT ATT TCC CTT GCC TCA CTC CCC AAG ATT GAT GAT TTT TCT TTT CAG Ser Ile Ser Leu Ala Ser Leu Pro Lys Ile Asp Asp Phe Ser Phe Gln 320 325 330	1056
TGG CTA AAA TGT TTG GAG CAC CTT AAC ATG GAA GAT AAT GAT ATT CCA Trp Leu Lys Cys Leu Glu His Leu Asn Met Glu Asp Asn Asp Ile Pro 335 340 345	1104
GGC ATA AAA AGC AAT ATG TTC ACA GGA TTG ATA AAC CTG AAA TAC TTA Gly Ile Lys Ser Asn Met Phe Thr Gly Leu Ile Asn Leu Lys Tyr Leu 350 355 360	1152

AGT CTA TCC AAC TCC TTT ACA AGT TTG CGA ACT TTG ACA AAT GAA ACA Ser Leu Ser Asn Ser Phe Thr Ser Leu Arg Thr Leu Thr Asn Glu Thr 365 370 375	1200
TTT GTA TCA CTT GCT CAT TCT CCC TTA CAC ATA CTC AAC CTA ACC AAG Phe Val Ser Leu Ala His Ser Pro Leu His Ile Leu Asn Leu Thr Lys 380 385 390 395	1248
AAT AAA ATC TCA AAA ATA GAG AGT GAT GCT TTC TCT TGG TTG GGC CAC Asn Lys Ile Ser Lys Ile Glu Ser Asp Ala Phe Ser Trp Leu Gly His 400 405 410	1296
CTA GAA GTA CTT GAC CTG GGC CTT AAT GAA ATT GGG CAA GAA CTC ACA Leu Glu Val Leu Asp Leu Gly Leu Asn Glu Ile Gly Gln Glu Leu Thr 415 420 425	1344
GGC CAG GAA TGG AGA GGT CTA GAA AAT ATT TTC GAA ATC TAT CTT TCC Gly Gln Glu Trp Arg Gly Leu Glu Asn Ile Phe Glu Ile Tyr Leu Ser 430 435 440	1392
TAC AAC AAG TAC CTG CAG CTG ACT AGG AAC TCC TTT GCC TTG GTC CCA Tyr Asn Lys Tyr Leu Gln Leu Thr Arg Asn Ser Phe Ala Leu Val Pro 445 450 455	1440
AGC CTT CAA CGA CTG ATG CTC CGA AGG GTG GCC CTT AAA AAT GTG GAT Ser Leu Gln Arg Leu Met Leu Arg Arg Val Ala Leu Lys Asn Val Asp 460 465 470 475	1488
AGC TCT CCT TCA CCA TTC CAG CCT CTT CGT AAC TTG ACC ATT CTG GAT Ser Ser Pro Ser Pro Phe Gln Pro Leu Arg Asn Leu Thr Ile Leu Asp 480 485 490	1536
CTA AGC AAC AAC AAC ATA GCC AAC ATA AAT GAT GAC ATG TTG GAG GGT Leu Ser Asn Asn Asn Ile Ala Asn Ile Asn Asp Asp Met Leu Glu Gly 495 500 505	1584
CTT GAG AAA CTA GAA ATT CTC GAT TTG CAG CAT AAC AAC TTA GCA CGG Leu Glu Lys Leu Glu Ile Leu Asp Leu Gln His Asn Asn Leu Ala Arg 510 515 520	1632
CTC TGG AAA CAC GCA AAC CCT GGT GGT CCC ATT TAT TTC CTA AAG GGT Leu Trp Lys His Ala Asn Pro Gly Gly Pro Ile Tyr Phe Leu Lys Gly 525 530 535	1680
CTG TCT CAC CTC CAC ATC CTT AAC TTG GAG TCC AAC GGC TTT GAC GAG Leu Ser His Leu His Ile Leu Asn Leu Glu Ser Asn Gly Phe Asp Glu 540 545 550 555	1728

ATC CCA GTT GAG GTC TTC AAG GAT TTA TTT GAA CTA AAG ATC ATC GAT Ile Pro Val Glu Val Phe Lys Asp Leu Phe Glu Leu Lys Ile Ile Asp 560 565 570	1776
TTA GGA TTG AAT AAT TTA AAC ACA CTT CCA GCA TCT GTC TTT AAT AAT Leu Gly Leu Asn Asn Leu Asn Thr Leu Pro Ala Ser Val Phe Asn Asn 575 580 585	1824
CAG GTG TCT CTA AAG TCA TTG AAC CTT CAG AAG AAT CTC ATA ACA TCC Gln Val Ser Leu Lys Ser Leu Asn Leu Gln Lys Asn Leu Ile Thr Ser 590 595 600	1872
GTT GAG AAG AAG GTT TTC GGG CCA GCT TTC AGG AAC CTG ACT GAG TTA Val Glu Lys Lys Val Phe Gly Pro Ala Phe Arg Asn Leu Thr Glu Leu 605 610 615	1920
GAT ATG CGC TTT AAT CCC TTT GAT TGC ACG TGT GAA AGT ATT GCC TGG Asp Met Arg Phe Asn Pro Phe Asp Cys Thr Cys Glu Ser Ile Ala Trp 620 625 630 635	1968
TTT GTT AAT TGG ATT AAC GAG ACC CAT ACC AAC ATC CCT GAG CTG TCA Phe Val Asn Trp Ile Asn Glu Thr His Thr Asn Ile Pro Glu Leu Ser 640 645 650	2016
AGC CAC TAC CTT TGC AAC ACT CCA CCT CAC TAT CAT GGG TTC CCA GTG Ser His Tyr Leu Cys Asn Thr Pro Pro His Tyr His Gly Phe Pro Val 655 660 665	2064
AGA CTT TTT GAT ACA TCA TCT TGC AAA GAC AGT GCC CCC TTT GAA CTC Arg Leu Phe Asp Thr Ser Ser Cys Lys Asp Ser Ala Pro Phe Glu Leu 670 675 680	2112
TTT TTC ATG ATC AAT ACC AGT ATC CTG TTG ATT TTT ATC TTT ATT GTA Phe Phe Met Ile Asn Thr Ser Ile Leu Leu Ile Phe Ile Phe Ile Val 685 690 695	2160
CTT CTC ATC CAC TTT GAG GGC TGG AGG ATA TCT TTT TAT TGG AAT GTT Leu Leu Ile His Phe Glu Gly Trp Arg Ile Ser Phe Tyr Trp Asn Val 700 705 710 715	2208
TCA GTA CAT CGA GTT CTT GGT TTC AAA GAA ATA GAC AGA CAG ACA GAA Ser Val His Arg Val Leu Gly Phe Lys Glu Ile Asp Arg Gln Thr Glu 720 725 730	2256
CAG TTT GAA TAT GCA GCA TAT ATA ATT CAT GCC TAT AAA GAT AAG GAT Gln Phe Glu Tyr Ala Ala Tyr Ile Ile His Ala Tyr Lys Asp Lys Asp 735 740 745	2304

TGG GTC TGG GAA CAT TTC TCT TCA ATG GAA AAG GAA GAC CAA TCT CTC Trp Val Trp Glu His Phe Ser Ser Met Glu Lys Glu Asp Gln Ser Leu 750 755 760	2352
AAA TTT TGT CTG GAA GAA AGG GAC TTT GAG GCG GGT GTT TTT GAA CTA Lys Phe Cys Leu Glu Glu Arg Asp Phe Glu Ala Gly Val Phe Glu Leu 765 770 775	2400
GAA GCA ATT GTT AAC AGC ATC AAA AGA AGC AGA AAA ATT ATT TTT GTT Glu Ala Ile Val Asn Ser Ile Lys Arg Ser Arg Lys Ile Ile Phe Val 780 785 790 795	2448
ATA ACA CAC CAT CTA TTA AAA GAC CCA TTA TGC AAA AGA TTC AAG GTA Ile Thr His His Leu Leu Lys Asp Pro Leu Cys Lys Arg Phe Lys Val 800 805 810	2496
CAT CAT GCA GTT CAA CAA GCT ATT GAA CAA AAT CTG GAT TCC ATT ATA His His Ala Val Gln Gln Ala Ile Glu Gln Asn Leu Asp Ser Ile Ile 815 820 825	2544
TTG GTT TTC CTT GAG GAG ATT CCA GAT TAT AAA CTG AAC CAT GCA CTC Leu Val Phe Leu Glu Glu Ile Pro Asp Tyr Lys Leu Asn His Ala Leu 830 835 840	2592
TGT TTG CGA AGA GGA ATG TTT AAA TCT CAC TGC ATC TTG AAC TGG CCA Cys Leu Arg Arg Gly Met Phe Lys Ser His Cys Ile Leu Asn Trp Pro 845 850 855	2640
GTT CAG AAA GAA CGG ATA GGT GCC TTT CGT CAT AAA TTG CAA GTA GCA Val Gln Lys Glu Arg Ile Gly Ala Phe Arg His Lys Leu Gln Val Ala 860 865 870 875	2688
CTT GGA TCC AAA AAC TCT GTA CAT TAA Leu Gly Ser Lys Asn Ser Val His 880	2715

(2) SEQ ID NO: 6 的信息:

- (i) 顺序特征:
 - (A) 长度: 904 个氨基酸
 - (B) 类型: 氨基酸
 - (D) 拓扑学: 线性
- (ii) 分子类型: 蛋白质
- (xi) 顺序描述: SEQ ID NO: 6:

Met Arg Gln Thr Leu Pro Cys Ile Tyr Phe Trp Gly Gly Leu Leu Pro
 -21 -20 -15 -10
 Phe Gly Met Leu Cys Ala Ser Ser Thr Thr Lys Cys Thr Val Ser His
 -5 1 5 10
 Glu Val Ala Asp Cys Ser His Leu Lys Leu Thr Gln Val Pro Asp Asp
 15 20 25
 Leu Pro Thr Asn Ile Thr Val Leu Asn Leu Thr His Asn Gln Leu Arg
 30 35 40
 Arg Leu Pro Ala Ala Asn Phe Thr Arg Tyr Ser Gln Leu Thr Ser Leu
 45 50 55
 Asp Val Gly Phe Asn Thr Ile Ser Lys Leu Glu Pro Glu Leu Cys Gln
 60 65 70 75
 Lys Leu Pro Met Leu Lys Val Leu Asn Leu Gln His Asn Glu Leu Ser
 80 85 90
 Gln Leu Ser Asp Lys Thr Phe Ala Phe Cys Thr Asn Leu Thr Glu Leu
 95 100 105
 His Leu Met Ser Asn Ser Ile Gln Lys Ile Lys Asn Asn Pro Phe Val
 110 115 120
 Lys Gln Lys Asn Leu Ile Thr Leu Asp Leu Ser His Asn Gly Leu Ser
 125 130 135
 Ser Thr Lys Leu Gly Thr Gln Val Gln Leu Glu Asn Leu Gln Glu Leu
 140 145 150 155
 Leu Leu Ser Asn Asn Lys Ile Gln Ala Leu Lys Ser Glu Glu Leu Asp
 160 165 170
 Ile Phe Ala Asn Ser Ser Leu Lys Lys Leu Glu Leu Ser Ser Asn Gln
 175 180 185
 Ile Lys Glu Phe Ser Pro Gly Cys Phe His Ala Ile Gly Arg Leu Phe
 190 195 200
 Gly Leu Phe Leu Asn Asn Val Gln Leu Gly Pro Ser Leu Thr Glu Lys
 205 210 215
 Leu Cys Leu Glu Leu Ala Asn Thr Ser Ile Arg Asn Leu Ser Leu Ser
 220 225 230 235

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

Leu Ser Asn Asn Asn Ile Ala Asn Ile Asn Asp Asp Met Leu Glu Gly
495 500 505

Leu Glu Lys Leu Glu Ile Leu Asp Leu Gln His Asn Asn Leu Ala Arg
510 515 520

Leu Trp Lys His Ala Asn Pro Gly Gly Pro Ile Tyr Phe Leu Lys Gly
525 530 535

Leu Ser His Leu His Ile Leu Asn Leu Glu Ser Asn Gly Phe Asp Glu
540 545 550 555

Ile Pro Val Glu Val Phe Lys Asp Leu Phe Glu Leu Lys Ile Ile Asp
560 565 570

Leu Gly Leu Asn Asn Leu Asn Thr Leu Pro Ala Ser Val Phe Asn Asn
575 580 585

Gln Val Ser Leu Lys Ser Leu Asn Leu Gln Lys Asn Leu Ile Thr Ser
590 595 600

Val Glu Lys Lys Val Phe Gly Pro Ala Phe Arg Asn Leu Thr Glu Leu
605 610 615

Asp Met Arg Phe Asn Pro Phe Asp Cys Thr Cys Glu Ser Ile Ala Trp
620 625 630 635

Phe Val Asn Trp Ile Asn Glu Thr His Thr Asn Ile Pro Glu Leu Ser
640 645 650

Ser His Tyr Leu Cys Asn Thr Pro Pro His Tyr His Gly Phe Pro Val
655 660 665

Arg Leu Phe Asp Thr Ser Ser Cys Lys Asp Ser Ala Pro Phe Glu Leu
670 675 680

Phe Phe Met Ile Asn Thr Ser Ile Leu Leu Ile Phe Ile Phe Ile Val
685 690 695

Leu Leu Ile His Phe Glu Gly Trp Arg Ile Ser Phe Tyr Trp Asn Val
700 705 710 715

Ser Val His Arg Val Leu Gly Phe Lys Glu Ile Asp Arg Gln Thr Glu
720 725 730

Gln Phe Glu Tyr Ala Ala Tyr Ile Ile His Ala Tyr Lys Asp Lys Asp
735 740 745

Trp Val Trp Glu His Phe Ser Ser Met Glu Lys Glu Asp Gln Ser Leu
750 755 760

Lys Phe Cys Leu Glu Glu Arg Asp Phe Glu Ala Gly Val Phe Glu Leu
765 770 775

Glu Ala Ile Val Asn Ser Ile Lys Arg Ser Arg Lys Ile Ile Phe Val
780 785 790 795

Ile Thr His His Leu Leu Lys Asp Pro Leu Cys Lys Arg Phe Lys Val
800 805 810

His His Ala Val Gln Gln Ala Ile Glu Gln Asn Leu Asp Ser Ile Ile
815 820 825

Leu Val Phe Leu Glu Glu Ile Pro Asp Tyr Lys Leu Asn His Ala Leu
830 835 840

Cys Leu Arg Arg Gly Met Phe Lys Ser His Cys Ile Leu Asn Trp Pro
845 850 855

Val Gln Lys Glu Arg Ile Gly Ala Phe Arg His Lys Leu Gln Val Ala
860 865 870 875

Leu Gly Ser Lys Asn Ser Val His
880

(2) SEQ ID NO: 7 的信息:

- (i) 顺序特征:
- (A) 长度: 2400 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线性

(ii) 分子类型: cDNA

- (ix) 特征:
- (A) 名称/关键词: CDS
 - (B) 位置: 1...2397

(xi) 顺序描述: SEQ ID NO: 7:

ATG GAG CTG AAT TTC TAC AAA ATC CCC GAC AAC CTC CCC TTC TCA ACC
Met Glu Leu Asn Phe Tyr Lys Ile Pro Asp Asn Leu Pro Phe Ser Thr
1 5 10 15

48

AAG AAC CTG GAC CTG AGC TTT AAT CCC CTG AGG CAT TTA GGC AGC TAT Lys Asn Leu Asp Leu Ser Phe Asn Pro Leu Arg His Leu Gly Ser Tyr 20 25 30	96
AGC TTC TTC AGT TTC CCA GAA CTG CAG GTG CTG GAT TTA TCC AGG TGT Ser Phe Phe Ser Phe Pro Glu Leu Gln Val Leu Asp Leu Ser Arg Cys 35 40 45	144
GAA ATC CAG ACA ATT GAA GAT GGG GCA TAT CAG AGC CTA AGC CAC CTC Glu Ile Gln Thr Ile Glu Asp Gly Ala Tyr Gln Ser Leu Ser His Leu 50 55 60	192
TCT ACC TTA ATA TTG ACA GGA AAC CCC ATC CAG AGT TTA GCC CTG GGA Ser Thr Leu Ile Leu Thr Gly Asn Pro Ile Gln Ser Leu Ala Leu Gly 65 70 75 80	240
GCC TTT TCT GGA CTA TCA AGT TTA CAG AAG CTG GTG GCT GTG GAG ACA Ala Phe Ser Gly Leu Ser Ser Leu Gln Lys Leu Val Ala Val Glu Thr 85 90 95	288
AAT CTA GCA TCT CTA GAG AAC TTC CCC ATT GGA CAT CTC AAA ACT TTG Asn Leu Ala Ser Leu Glu Asn Phe Pro Ile Gly His Leu Lys Thr Leu 100 105 110	336
AAA GAA CTT AAT GTG GCT CAC AAT CTT ATC CAA TCT TTC AAA TTA CCT Lys Glu Leu Asn Val Ala His Asn Leu Ile Gln Ser Phe Lys Leu Pro 115 120 125	384
GAG TAT TTT TCT AAT CTG ACC AAT CTA GAG CAC TTG GAC CTT TCC AGC Glu Tyr Phe Ser Asn Leu Thr Asn Leu Glu His Leu Asp Leu Ser Ser 130 135 140	432
AAC AAG ATT CAA AGT ATT TAT TGC ACA GAC TTG CGG GTT CTA CAT CAA Asn Lys Ile Gln Ser Ile Tyr Cys Thr Asp Leu Arg Val Leu His Gln 145 150 155 160	480
ATG CCC CTA CTC AAT CTC TCT TTA GAC CTG TCC CTG AAC CCT ATG AAC Met Pro Leu Leu Asn Leu Ser Leu Asp Leu Ser Leu Asn Pro Met Asn 165 170 175	528
TTT ATC CAA CCA GGT GCA TTT AAA GAA ATT AGG CTT CAT AAG CTG ACT Phe Ile Gln Pro Gly Ala Phe Lys Glu Ile Arg Leu His Lys Leu Thr 180 185 190	576
TTA AGA AAT AAT TTT GAT AGT TTA AAT GTA ATG AAA ACT TGT ATT CAA Leu Arg Asn Asn Phe Asp Ser Leu Asn Val Met Lys Thr Cys Ile Gln 195 200 205	624

GGT CTG GCT GGT TTA GAA GTC CAT CGT TTG GTT CTG GGA GAA TTT AGA Gly Leu Ala Gly Leu Glu Val His Arg Leu Val Leu Gly Glu Phe Arg 210 215 220	672
AAT GAA GGA AAC TTG GAA AAG TTT GAC AAA TCT GCT CTA GAG GGC CTG Asn Glu Gly Asn Leu Glu Lys Phe Asp Lys Ser Ala Leu Glu Gly Leu 225 230 235 240	720
TGC AAT TTG ACC ATT GAA GAA TTC CGA TTA GCA TAC TTA GAC TAC TAC Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr 245 250 255	768
CTC GAT GAT ATT ATT GAC TTA TTT AAT TGT TTG ACA AAT GTT TCT TCA Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Thr Asn Val Ser Ser 260 265 270	816
TTT TCC CTG GTG AGT GTG ACT ATT GAA AGG GTA AAA GAC TTT TCT TAT Phe Ser Leu Val Ser Val Thr Ile Glu Arg Val Lys Asp Phe Ser Tyr 275 280 285	864
AAT TTC GGA TGG CAA CAT TTA GAA TTA GTT AAC TGT AAA TTT GGA CAG Asn Phe Gly Trp Gln His Leu Glu Leu Val Asn Cys Lys Phe Gly Gln 290 295 300	912
TTT CCC ACA TTG AAA CTC AAA TCT CTC AAA AGG CTT ACT TTC ACT TCC Phe Pro Thr Leu Lys Leu Lys Ser Leu Lys Arg Leu Thr Phe Thr Ser 305 310 315 320	960
AAC AAA GGT GGG AAT GCT TTT TCA GAA GTT GAT CTA CCA AGC CTT GAG Asn Lys Gly Gly Asn Ala Phe Ser Glu Val Asp Leu Pro Ser Leu Glu 325 330 335	1008
TTT CTA GAT CTC AGT AGA AAT GGC TTG AGT TTC AAA GGT TGC TGT TCT Phe Leu Asp Leu Ser Arg Asn Gly Leu Ser Phe Lys Gly Cys Cys Ser 340 345 350	1056
CAA AGT GAT TTT GGG ACA ACC AGC CTA AAG TAT TTA GAT CTG AGC TTC Gln Ser Asp Phe Gly Thr Thr Ser Leu Lys Tyr Leu Asp Leu Ser Phe 355 360 365	1104
AAT GGT GTT ATT ACC ATG AGT TCA AAC TTC TTG GGC TTA GAA CAA CTA Asn Gly Val Ile Thr Met Ser Ser Asn Phe Leu Gly Leu Glu Gln Leu 370 375 380	1152
GAA CAT CTG GAT TTC CAG CAT TCC AAT TTG AAA CAA ATG AGT GAG TTT Glu His Leu Asp Phe Gln His Ser Asn Leu Lys Gln Met Ser Glu Phe 385 390 395 400	1200

TCA GTA TTC CTA TCA CTC AGA AAC CTC ATT TAC CTT GAC ATT TCT CAT Ser Val Phe Leu Ser Leu Arg Asn Leu Ile Tyr Leu Asp Ile Ser His 405 410 415	1248
ACT CAC ACC AGA GTT GCT TTC AAT GGC ATC TTC AAT GGC TTG TCC AGT Thr His Thr Arg Val Ala Phe Asn Gly Ile Phe Asn Gly Leu Ser Ser 420 425 430	1296
CTC GAA GTC TTG AAA ATG GCT GGC AAT TCT TTC CAG GAA AAC TTC CTT Leu Glu Val Leu Lys Met Ala Gly Asn Ser Phe Gln Glu Asn Phe Leu 435 440 445	1344
CCA GAT ATC TTC ACA GAG CTG AGA AAC TTG ACC TTC CTG GAC CTC TCT Pro Asp Ile Phe Thr Glu Leu Arg Asn Leu Thr Phe Leu Asp Leu Ser 450 455 460	1392
CAG TGT CAA CTG GAG CAG TTG TCT CCA ACA GCA TTT AAC TCA CTC TCC Gln Cys Gln Leu Glu Gln Leu Ser Pro Thr Ala Phe Asn Ser Leu Ser 465 470 475 480	1440
AGT CTT CAG GTA CTA AAT ATG AGC CAC AAC AAC TTC TTT TCA TTG GAT Ser Leu Gln Val Leu Asn Met Ser His Asn Asn Phe Phe Ser Leu Asp 485 490 495	1488
ACG TTT CCT TAT AAG TGT CTG AAC TCC CTC CAG GTT CTT GAT TAC AGT Thr Phe Pro Tyr Lys Cys Leu Asn Ser Leu Gln Val Leu Asp Tyr Ser 500 505 510	1536
CTC AAT CAC ATA ATG ACT TCC AAA AAA CAG GAA CTA CAG CAT TTT CCA Leu Asn His Ile Met Thr Ser Lys Lys Gln Glu Leu Gln His Phe Pro 515 520 525	1584
AGT AGT CTA GCT TTC TTA AAT CTT ACT CAG AAT GAC TTT GCT TGT ACT Ser Ser Leu Ala Phe Leu Asn Leu Thr Gln Asn Asp Phe Ala Cys Thr 530 535 540	1632
TGT GAA CAC CAG AGT TTC CTG CAA TGG ATC AAG GAC CAG AGG CAG CTC Cys Glu His Gln Ser Phe Leu Gln Trp Ile Lys Asp Gln Arg Gln Leu 545 550 555 560	1680
TTG GTG GAA GTT GAA CGA ATG GAA TGT GCA ACA CCT TCA GAT AAG CAG Leu Val Glu Val Glu Arg Met Glu Cys Ala Thr Pro Ser Asp Lys Gln 565 570 575	1728
GGC ATG CCT GTG CTG AGT TTG AAT ATC ACC TGT CAG ATG AAT AAG ACC Gly Met Pro Val Leu Ser Leu Asn Ile Thr Cys Gln Met Asn Lys Thr 580 585 590	1776

ATC ATT GGT GTG TCG GTC CTC AGT GTG CTT GTA GTA TCT GTT GTA GCA Ile Ile Gly Val Ser Val Leu Ser Val Leu Val Val Ser Val Val Ala 595 600 605	1824
GTT CTG GTC TAT AAG TTC TAT TTT CAC CTG ATG CTT CTT GCT GGC TGC Val Leu Val Tyr Lys Phe Tyr Phe His Leu Met Leu Leu Ala Gly Cys 610 615 620	1872
ATA AAG TAT GGT AGA GGT GAA AAC ATC TAT GAT GCC TTT GTT ATC TAC Ile Lys Tyr Gly Arg Gly Glu Asn Ile Tyr Asp Ala Phe Val Ile Tyr 625 630 635 640	1920
TCA AGC CAG GAT GAG GAC TGG GTA AGG AAT GAG CTA GTA AAG AAT TTA Ser Ser Gln Asp Glu Asp Trp Val Arg Asn Glu Leu Val Lys Asn Leu 645 650 655	1968
GAA GAA GGG GTG CCT CCA TTT CAG CTC TGC CTT CAC TAC AGA GAC TTT Glu Glu Gly Val Pro Pro Phe Gln Leu Cys Leu His Tyr Arg Asp Phe 660 665 670	2016
ATT CCC GGT GTG GCC ATT GCT GCC AAC ATC ATC CAT GAA GGT TTC CAT Ile Pro Gly Val Ala Ile Ala Ala Asn Ile Ile His Glu Gly Phe His 675 680 685	2064
AAA AGC CGA AAG GTG ATT GTT GTG GTG TCC CAG CAC TTC ATC CAG AGC Lys Ser Arg Lys Val Ile Val Val Val Ser Gln His Phe Ile Gln Ser 690 695 700	2112
CGC TGG TGT ATC TTT GAA TAT GAG ATT GCT CAG ACC TGG CAG TTT CTG Arg Trp Cys Ile Phe Glu Tyr Glu Ile Ala Gln Thr Trp Gln Phe Leu 705 710 715 720	2160
AGC AGT CGT GCT GGT ATC ATC TTC ATT GTC CTG CAG AAG GTG GAG AAG Ser Ser Arg Ala Gly Ile Ile Phe Ile Val Leu Gln Lys Val Glu Lys 725 730 735	2208
ACC CTG CTC AGG CAG CAG GTG GAG CTG TAC CGC CTT CTC AGC AGG AAC Thr Leu Leu Arg Gln Gln Val Glu Leu Tyr Arg Leu Leu Ser Arg Asn 740 745 750	2256
ACT TAC CTG GAG TGG GAG GAC AGT GTC CTG GGG CGG CAC ATC TTC TGG Thr Tyr Leu Glu Trp Glu Asp Ser Val Leu Gly Arg His Ile Phe Trp 755 760 765	2304
AGA CGA CTC AGA AAA GCC CTG CTG GAT GGT AAA TCA TGG AAT CCA GAA Arg Arg Leu Arg Lys Ala Leu Leu Asp Gly Lys Ser Trp Asn Pro Glu 770 775 780	2352

GGA ACA GTG GGT ACA GGA TGC AAT TGG CAG GAA GCA ACA TCT ATC 2397
 Gly Thr Val Gly Thr Gly Cys Asn Trp Gln Glu Ala Thr Ser Ile
 785 790 795

TGA 2400

(2) SEQ ID NO: 8 的信息:

(i) 顺序特征:

(A) 长度: 799 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 顺序描述: SEQ ID NO: 8:

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

Met Pro Leu Leu Asn Leu Ser Leu Asp Leu Ser Leu Asn Pro Met Asn
 165 170 175

Phe Ile Gln Pro Gly Ala Phe Lys Glu Ile Arg Leu His Lys Leu Thr
 180 185 190

Leu Arg Asn Asn Phe Asp Ser Leu Asn Val Met Lys Thr Cys Ile Gln
 195 200 205

Gly Leu Ala Gly Leu Glu Val His Arg Leu Val Leu Gly Glu Phe Arg
 210 215 220

Asn Glu Gly Asn Leu Glu Lys Phe Asp Lys Ser Ala Leu Glu Gly Leu
 225 230 235 240

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr
 245 250 255

Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Thr Asn Val Ser Ser
 260 265 270

Phe Ser Leu Val Ser Val Thr Ile Glu Arg Val Lys Asp Phe Ser Tyr
 275 280 285

Asn Phe Gly Trp Gln His Leu Glu Leu Val Asn Cys Lys Phe Gly Gln
 290 295 300

Phe Pro Thr Leu Lys Leu Lys Ser Leu Lys Arg Leu Thr Phe Thr Ser
 305 310 315 320

Asn Lys Gly Gly Asn Ala Phe Ser Glu Val Asp Leu Pro Ser Leu Glu
 325 330 335

Phe Leu Asp Leu Ser Arg Asn Gly Leu Ser Phe Lys Gly Cys Cys Ser
 340 345 350

Gln Ser Asp Phe Gly Thr Thr Ser Leu Lys Tyr Leu Asp Leu Ser Phe
 355 360 365

Asn Gly Val Ile Thr Met Ser Ser Asn Phe Leu Gly Leu Glu Gln Leu
 370 375 380

Glu His Leu Asp Phe Gln His Ser Asn Leu Lys Gln Met Ser Glu Phe
 385 390 395 400

Ser Val Phe Leu Ser Leu Arg Asn Leu Ile Tyr Leu Asp Ile Ser His
 405 410 415

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

Ile Pro Gly Val Ala Ile Ala Ala Asn Ile Ile His Glu Gly Phe His
675 680 685

Lys Ser Arg Lys Val Ile Val Val Val Ser Gln His Phe Ile Gln Ser
690 695 700

Arg Trp Cys Ile Phe Glu Tyr Glu Ile Ala Gln Thr Trp Gln Phe Leu
705 710 715 720

Ser Ser Arg Ala Gly Ile Ile Phe Ile Val Leu Gln Lys Val Glu Lys
725 730 735

Thr Leu Leu Arg Gln Gln Val Glu Leu Tyr Arg Leu Leu Ser Arg Asn
740 745 750

Thr Tyr Leu Glu Trp Glu Asp Ser Val Leu Gly Arg His Ile Phe Trp
755 760 765

Arg Arg Leu Arg Lys Ala Leu Leu Asp Gly Lys Ser Trp Asn Pro Glu
770 775 780

Gly Thr Val Gly Thr Gly Cys Asn Trp Gln Glu Ala Thr Ser Ile
785 790 795

(2) SEQ ID NO: 9 的信息:

- (i) 顺序特征:
 (A) 长度: 1275 个碱基对
 (B) 类型: 核酸
 (C) 链型: 单链
 (D) 拓扑学: 线性

- (ii) 分子类型: cDNA

- (ix) 特征:
 (A) 名称/关键词: CDS
 (B) 位置: 1...1095

- (xi) 顺序描述: SEQ ID NO: 9:

TGT TGC GAT GTT TTT GAG GGA CTT TCT CAT CTT CAA GTT CTG TAT TTG 48
 Cys Trp Asp Val Phe Glu Gly Leu Ser His Leu Gln Val Leu Tyr Leu
 1 5 10 15

AAT CAT AAC TAT CTT AAT TCC CTT CCA CCA GGA GTA TTT AGC CAT CTG 96
 Asn His Asn Tyr Leu Asn Ser Leu Pro Pro Gly Val Phe Ser His Leu
 20 25 30

ACT GCA TTA AGG GGA CTA AGC CTC AAC TCC AAC AGG CTG ACA GTT CTT Thr Ala Leu Arg Gly Leu Ser Leu Asn Ser Asn Arg Leu Thr Val Leu 35 40 45	144
TCT CAC AAT GAT TTA CCT GCT AAT TTA GAG ATC CTG GAC ATA TCC AGG Ser His Asn Asp Leu Pro Ala Asn Leu Glu Ile Leu Asp Ile Ser Arg 50 55 60	192
AAC CAG CTC CTA GCT CCT AAT CCT GAT GTA TTT GTA TCA CTT AGT GTC Asn Gln Leu Leu Ala Pro Asn Pro Asp Val Phe Val Ser Leu Ser Val 65 70 75 80	240
TTG GAT ATA ACT CAT AAC AAG TTC ATT TGT GAA TGT GAA CTT AGC ACT Leu Asp Ile Thr His Asn Lys Phe Ile Cys Glu Cys Glu Leu Ser Thr 85 90 95	288
TTT ATC AAT TGG CTT AAT CAC ACC AAT GTC ACT ATA GCT GGG CCT CCT Phe Ile Asn Trp Leu Asn His Thr Asn Val Thr Ile Ala Gly Pro Pro 100 105 110	336
GCA GAC ATA TAT TGT GTG TAC CCT GAC TCG TTC TCT GGG GTT TCC CTC Ala Asp Ile Tyr Cys Val Tyr Pro Asp Ser Phe Ser Gly Val Ser Leu 115 120 125	384
TTC TCT CTT TCC ACG GAA GGT TGT GAT GAA GAG GAA GTC TTA AAG TCC Phe Ser Leu Ser Thr Glu Gly Cys Asp Glu Glu Glu Val Leu Lys Ser 130 135 140	432
CTA AAG TTC TCC CTT TTC ATT GTA TGC ACT GTC ACT CTG ACT CTG TTC Leu Lys Phe Ser Leu Phe Ile Val Cys Thr Val Thr Leu Thr Leu Phe 145 150 155 160	480
CTC ATG ACC ATC CTC ACA GTC ACA AAG TTC CGG GGC TTC TGT TTT ATC Leu Met Thr Ile Leu Thr Val Thr Lys Phe Arg Gly Phe Cys Phe Ile 165 170 175	528
TGT TAT AAG ACA GCC CAG AGA CTG GTG TTC AAG GAC CAT CCC CAG GGC Cys Tyr Lys Thr Ala Gln Arg Leu Val Phe Lys Asp His Pro Gln Gly 180 185 190	576
ACA GAA CCT GAT ATG TAC AAA TAT GAT GCC TAT TTG TGC TTC AGC AGC Thr Glu Pro Asp Met Tyr Lys Tyr Asp Ala Tyr Leu Cys Phe Ser Ser 195 200 205	624
AAA GAC TTC ACA TGG GTG CAG AAT GCT TTG CTC AAA CAC CTG GAC ACT Lys Asp Phe Thr Trp Val Gln Asn Ala Leu Leu Lys His Leu Asp Thr 210 215 220	672

CAA TAC AGT GAC CAA AAC AGA TTC AAC CTG TGC TTT GAA GAA AGA GAC Gln Tyr Ser Asp Gln Asn Arg Phe Asn Leu Cys Phe Glu Glu Arg Asp 225 230 235 240	720
TTT GTC CCA GGA GAA AAC CGC ATT GCC AAT ATC CAG GAT GCC ATC TGG Phe Val Pro Gly Glu Asn Arg Ile Ala Asn Ile Gln Asp Ala Ile Trp 245 250 255	768
AAC AGT AGA AAG ATC GTT TGT CTT GTG AGC AGA CAC TTC CTT AGA GAT Asn Ser Arg Lys Ile Val Cys Leu Val Ser Arg His Phe Leu Arg Asp 260 265 270	816
GGC TGG TGC CTT GAA GCC TTC AGT TAT GCC CAG GGC AGG TGC TTA TCT Gly Trp Cys Leu Glu Ala Phe Ser Tyr Ala Gln Gly Arg Cys Leu Ser 275 280 285	864
GAC CTT AAC AGT GCT CTC ATC ATG GTG GTG GTT GGG TCC TTG TCC CAG Asp Leu Asn Ser Ala Leu Ile Met Val Val Val Gly Ser Leu Ser Gln 290 295 300	912
TAC CAG TTG ATG AAA CAT CAA TCC ATC AGA GGC TTT GTA CAG AAA CAG Tyr Gln Leu Met Lys His Gln Ser Ile Arg Gly Phe Val Gln Lys Gln 305 310 315 320	960
CAG TAT TTG AGG TGG CCT GAG GAT CTC CAG GAT GTT GGC TGG TTT CTT Gln Tyr Leu Arg Trp Pro Glu Asp Leu Gln Asp Val Gly Trp Phe Leu 325 330 335	1008
CAT AAA CTC TCT CAA CAG ATA CTA AAG AAA GAA AAG GAA AAG AAG AAA His Lys Leu Ser Gln Gln Ile Leu Lys Lys Glu Lys Glu Lys Lys Lys 340 345 350	1056
GAC AAT AAC ATT CCG TTG CAA ACT GTA GCA ACC ATC TCC TAATCAAAGG Asp Asn Asn Ile Pro Leu Gln Thr Val Ala Thr Ile Ser 355 360 365	1105
AGCAATTTCC AACTTATCTC AAGCCACAAA TAACTCTTCA CTTTGTATTT GCACCAAGTT	1165
ATCATTTTGG GGTCTCTCTT GGAGGTTTTT TTTTCTTTT TGCTACTATG AAAACAACAT	1225
AAATCTCTCA ATTTTCGTAT CAAAAAAAAA AAAAAAAAAA TGGCGGCCGC	1275

(2) SEQ ID NO: 10 的信息:

- (i) 顺序特征:
- (A) 长度: 365 个氨基酸
 - (B) 类型: 氨基酸
 - (D) 拓扑学: 线性

(ii) 分子类型：蛋白质

(xi) 顺序描述：SEQ ID NO: 10:

```

Cys Trp Asp Val Phe Glu Gly Leu Ser His Leu Gln Val Leu Tyr Leu
 1             5             10             15

Asn His Asn Tyr Leu Asn Ser Leu Pro Pro Gly Val Phe Ser His Leu
      20             25             30

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

Ser His Asn Asp Leu Pro Ala Asn Leu Glu Ile Leu Asp Ile Ser Arg
      50             55             60

Asn Gln Leu Leu Ala Pro Asn Pro Asp Val Phe Val Ser Leu Ser Val
      65             70             75             80

Leu Asp Ile Thr His Asn Lys Phe Ile Cys Glu Cys Glu Leu Ser Thr
      85             90             95

Phe Ile Asn Trp Leu Asn His Thr Asn Val Thr Ile Ala Gly Pro Pro
      100            105            110

Ala Asp Ile Tyr Cys Val Tyr Pro Asp Ser Phe Ser Gly Val Ser Leu
      115            120            125

Phe Ser Leu Ser Thr Glu Gly Cys Asp Glu Glu Glu Val Leu Lys Ser
      130            135            140

Leu Lys Phe Ser Leu Phe Ile Val Cys Thr Val Thr Leu Thr Leu Phe
      145            150            155            160

Leu Met Thr Ile Leu Thr Val Thr Lys Phe Arg Gly Phe Cys Phe Ile
      165            170            175

Cys Tyr Lys Thr Ala Gln Arg Leu Val Phe Lys Asp His Pro Gln Gly
      180            185            190

Thr Glu Pro Asp Met Tyr Lys Tyr Asp Ala Tyr Leu Cys Phe Ser Ser
      195            200            205

Lys Asp Phe Thr Trp Val Gln Asn Ala Leu Leu Lys His Leu Asp Thr
      210            215            220

```

```

Gln Tyr Ser Asp Gln Asn Arg Phe Asn Leu Cys Phe Glu Glu Arg Asp
225                      230                      235                      240

Phe Val Pro Gly Glu Asn Arg Ile Ala Asn Ile Gln Asp Ala Ile Trp
                      245                      250                      255

Asn Ser Arg Lys Ile Val Cys Leu Val Ser Arg His Phe Leu Arg Asp
                      260                      265                      270

Gly Trp Cys Leu Glu Ala Phe Ser Tyr Ala Gln Gly Arg Cys Leu Ser
                      275                      280                      285

Asp Leu Asn Ser Ala Leu Ile Met Val Val Val Gly Ser Leu Ser Gln
290                      295                      300

Tyr Gln Leu Met Lys His Gln Ser Ile Arg Gly Phe Val Gln Lys Gln
305                      310                      315                      320

Gln Tyr Leu Arg Trp Pro Glu Asp Leu Gln Asp Val Gly Trp Phe Leu
                      325                      330                      335

His Lys Leu Ser Gln Gln Ile Leu Lys Lys Glu Lys Glu Lys Lys Lys
340                      345                      350

Asp Asn Asn Ile Pro Leu Gln Thr Val Ala Thr Ile Ser
355                      360                      365

```

(2) SEQ ID NO: 11 的信息:

- (i) 顺序特征:
 - (A) 长度: 3138 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线性
- (ii) 分子类型: cDNA
- (ix) 特征:
 - (A) 名称/关键词: CDS
 - (B) 位置: 1...3135
- (ix) 特征:
 - (A) 名称/关键词: 成熟肽
 - (B) 位置: 67...3135
- (xi) 顺序描述: SEQ ID NO: 11:

ATC TGG ACA CTG AAG AGA CTA ATT CTT ATC CTT TTT AAC ATA ATC CTA Met Trp Thr Leu Lys Arg Leu Ile Leu Ile Leu Phe Asn Ile Ile Leu -22 -20 -15 -10	48
ATT TCC AAA CTC CTT GGG GCT AGA TGG TTT CCT AAA ACT CTG CCC TGT Ile Ser Lys Leu Leu Gly Ala Arg Trp Phe Pro Lys Thr Leu Pro Cys -5 1 5 10	96
GAT GTC ACT CTG GAT GTT CCA AAG AAC CAT GTG ATC GTG GAC TGC ACA Asp Val Thr Leu Asp Val Pro Lys Asn His Val Ile Val Asp Cys Thr 15 20 25	144
GAC AAG CAT TTG ACA GAA ATT CCT GGA GGT ATT CCC ACG AAC ACC ACG Asp Lys His Leu Thr Glu Ile Pro Gly Gly Ile Pro Thr Asn Thr Thr 30 35 40	192
AAC CTC ACC CTC ACC ATT AAC CAC ATA CCA GAC ATC TCC CCA GCG TCC Asn Leu Thr Leu Thr Ile Asn His Ile Pro Asp Ile Ser Pro Ala Ser 45 50 55	240
TTT CAC AGA CTG GAC CAT CTG GTA GAG ATC GAT TTC AGA TGC AAC TGT Phe His Arg Leu Asp His Leu Val Glu Ile Asp Phe Arg Cys Asn Cys 60 65 70	288
GTA CCT ATT CCA CTG GGG TCA AAA AAC AAC ATG TGC ATC AAG AGG CTG Val Pro Ile Pro Leu Gly Ser Lys Asn Asn Met Cys Ile Lys Arg Leu 75 80 85 90	336
CAG ATT AAA CCC AGA AGC TTT AGT GGA CTC ACT TAT TTA AAA TCC CTT Gln Ile Lys Pro Arg Ser Phe Ser Gly Leu Thr Tyr Leu Lys Ser Leu 95 100 105	384
TAC CTG GAT GGA AAC CAG CTA CTA GAG ATA CCG CAG GGC CTC CCG CCT Tyr Leu Asp Gly Asn Gln Leu Leu Glu Ile Pro Gln Gly Leu Pro Pro 110 115 120	432
AGC TTA CAG CTT CTC AGC CTT GAG GCC AAC AAC ATC TTT TCC ATC AGA Ser Leu Gln Leu Leu Ser Leu Glu Ala Asn Asn Ile Phe Ser Ile Arg 125 130 135	480
AAA GAG AAT CTA ACA GAA CTG GCC AAC ATA GAA ATA CTC TAC CTG GGC Lys Glu Asn Leu Thr Glu Leu Ala Asn Ile Glu Ile Leu Tyr Leu Gly 140 145 150	528
CAA AAC TGT TAT TAT CGA AAT CCT TGT TAT GTT TCA TAT TCA ATA GAG Gln Asn Cys Tyr Tyr Arg Asn Pro Cys Tyr Val Ser Tyr Ser Ile Glu 155 160 165 170	576

AAA GAT GCC TTC CTA AAC TTG ACA AAG TTA AAA GTG CTC TCC CTG AAA Lys Asp Ala Phe Leu Asn Leu Thr Lys Leu Lys Val Leu Ser Leu Lys 175 180 185	624
GAT AAC AAT GTC ACA GCC GTC CCT ACT GTT TTG CCA TCT ACT TTA ACA Asp Asn Asn Val Thr Ala Val Pro Thr Val Leu Pro Ser Thr Leu Thr 190 195 200	672
GAA CTA TAT CTC TAC AAC AAC ATG ATT GCA AAA ATC CAA GAA GAT GAT Glu Leu Tyr Leu Tyr Asn Asn Met Ile Ala Lys Ile Gln Glu Asp Asp 205 210 215	720
TTT AAT AAC CTC AAC CAA TTA CAA ATT CTT GAC CTA AGT GGA AAT TGC Phe Asn Asn Leu Asn Gln Leu Gln Ile Leu Asp Leu Ser Gly Asn Cys 220 225 230	768
CCT CGT TGT TAT AAT GCC CCA TTT CCT TGT GCG CCG TGT AAA AAT AAT Pro Arg Cys Tyr Asn Ala Pro Phe Pro Cys Ala Pro Cys Lys Asn Asn 235 240 245 250	816
TCT CCC CTA CAG ATC CCT GTA AAT GCT TTT GAT GCG CTG ACA GAA TTA Ser Pro Leu Gln Ile Pro Val Asn Ala Phe Asp Ala Leu Thr Glu Leu 255 260 265	864
AAA GTT TTA CGT CTA CAC AGT AAC TCT CTT CAG CAT GTG CCC CCA AGA Lys Val Leu Arg Leu His Ser Asn Ser Leu Gln His Val Pro Pro Arg 270 275 280	912
TGG TTT AAG AAC ATC AAC AAA CTC CAG GAA CTG GAT CTG TCC CAA AAC Trp Phe Lys Asn Ile Asn Lys Leu Gln Glu Leu Asp Leu Ser Gln Asn 285 290 295	960
TTC TTG GCC AAA GAA ATT GGG GAT GCT AAA TTT CTG CAT TTT CTC CCC Phe Leu Ala Lys Glu Ile Gly Asp Ala Lys Phe Leu His Phe Leu Pro 300 305 310	1008
AGC CTC ATC CAA TTG GAT CTG TCT TTC AAT TTT GAA CTT CAG GTC TAT Ser Leu Ile Gln Leu Asp Leu Ser Phe Asn Phe Glu Leu Gln Val Tyr 315 320 325 330	1056
CGT GCA TCT ATG AAT CTA TCA CAA GCA TTT TCT TCA CTG AAA AGC CTG Arg Ala Ser Met Asn Leu Ser Gln Ala Phe Ser Ser Leu Lys Ser Leu 335 340 345	1104
AAA ATT CTG CGG ATC AGA GGA TAT GTC TTT AAA GAG TTG AAA AGC TTT Lys Ile Leu Arg Ile Arg Gly Tyr Val Phe Lys Glu Leu Lys Ser Phe 350 355 360	1152

AAC CTC TCG CCA TTA CAT AAT CTT CAA AAT CTT GAA GTT CTT GAT CTT Asn Leu Ser Pro Leu His Asn Leu Gln Asn Leu Glu Val Leu Asp Leu 365 370 375	1200
GGC ACT AAC TTT ATA AAA ATT GCT AAC CTC AGC ATG TTT AAA CAA TTT Gly Thr Asn Phe Ile Lys Ile Ala Asn Leu Ser Met Phe Lys Gln Phe 380 385 390	1248
AAA AGA CTG AAA GTC ATA GAT CTT TCA GTG AAT AAA ATA TCA CCT TCA Lys Arg Leu Lys Val Ile Asp Leu Ser Val Asn Lys Ile Ser Pro Ser 395 400 405 410	1296
GGA GAT TCA AGT GAA GTT GGC TTC TGC TCA AAT GCC AGA ACT TCT GTA Gly Asp Ser Ser Glu Val Gly Phe Cys Ser Asn Ala Arg Thr Ser Val 415 420 425	1344
GAA AGT TAT GAA CCC CAG GTC CTG GAA CAA TTA CAT TAT TTC AGA TAT Glu Ser Tyr Glu Pro Gln Val Leu Glu Gln Leu His Tyr Phe Arg Tyr 430 435 440	1392
GAT AAG TAT GCA AGG AGT TGC AGA TTC AAA AAC AAA GAG GCT TCT TTC Asp Lys Tyr Ala Arg Ser Cys Arg Phe Lys Asn Lys Glu Ala Ser Phe 445 450 455	1440
ATG TCT GTT AAT GAA AGC TGC TAC AAG TAT GGG CAG ACC TTG GAT CTA Met Ser Val Asn Glu Ser Cys Tyr Lys Tyr Gly Gln Thr Leu Asp Leu 460 465 470	1488
AGT AAA AAT AGT ATA TTT TTT GTC AAG TCC TCT GAT TTT CAG CAT CTT Ser Lys Asn Ser Ile Phe Phe Val Lys Ser Ser Asp Phe Gln His Leu 475 480 485 490	1536
TCT TTC CTC AAA TGC CTG AAT CTG TCA GGA AAT CTC ATT AGC CAA ACT Ser Phe Leu Lys Cys Leu Asn Leu Ser Gly Asn Leu Ile Ser Gln Thr 495 500 505	1584
CTT AAT GGC AGT GAA TTC CAA CCT TTA GCA GAG CTG AGA TAT TTG GAC Leu Asn Gly Ser Glu Phe Gln Pro Leu Ala Glu Leu Arg Tyr Leu Asp 510 515 520	1632
TTC TCC AAC AAC CGG CTT GAT TTA CTC CAT TCA ACA GCA TTT GAA GAG Phe Ser Asn Asn Arg Leu Asp Leu Leu His Ser Thr Ala Phe Glu Glu 525 530 535	1680
CTT CAC AAA CTG GAA GTT CTG GAT ATA AGC AGT AAT AGC CAT TAT TTT Leu His Lys Leu Glu Val Leu Asp Ile Ser Ser Asn Ser His Tyr Phe 540 545 550	1728

CAA TCA GAA GGA ATT ACT CAT ATG CTA AAC TTT ACC AAG AAC CTA AAG Gln Ser Glu Gly Ile Thr His Met Leu Asn Phe Thr Lys Asn Leu Lys 555 560 565 570	1776
GTT CTG CAG AAA CTG ATG ATG AAC GAC AAT GAC ATC TCT TCC TCC ACC Val Leu Gln Lys Leu Met Met Asn Asp Asn Asp Ile Ser Ser Ser Thr 575 580 585	1824
AGC AGG ACC ATG GAG AGT GAG TCT CTT AGA ACT CTG GAA TTC AGA GGA Ser Arg Thr Met Glu Ser Glu Ser Leu Arg Thr Leu Glu Phe Arg Gly 590 595 600	1872
AAT CAC TTA GAT GTT TTA TGG AGA GAA GGT GAT AAC AGA TAC TTA CAA Asn His Leu Asp Val Leu Trp Arg Glu Gly Asp Asn Arg Tyr Leu Gln 605 610 615	1920
TTA TTC AAG AAT CTG CTA AAA TTA GAG GAA TTA GAC ATC TCT AAA AAT Leu Phe Lys Asn Leu Leu Lys Leu Glu Glu Leu Asp Ile Ser Lys Asn 620 625 630	1968
TCC CTA AGT TTC TTG CCT TCT GGA GTT TTT GAT GGT ATG CCT CCA AAT Ser Leu Ser Phe Leu Pro Ser Gly Val Phe Asp Gly Met Pro Pro Asn 635 640 645 650	2016
CTA AAG AAT CTC TCT TTG GCC AAA AAT GGG CTC AAA TCT TTC AGT TGG Leu Lys Asn Leu Ser Leu Ala Lys Asn Gly Leu Lys Ser Phe Ser Trp 655 660 665	2064
AAG AAA CTC CAG TGT CTA AAG AAC CTG GAA ACT TTG GAC CTC AGC CAC Lys Lys Leu Gln Cys Leu Lys Asn Leu Glu Thr Leu Asp Leu Ser His 670 675 680	2112
AAC CAA CTG ACC ACT GTC CCT GAG AGA TTA TCC AAC TGT TCC AGA AGC Asn Gln Leu Thr Thr Val Pro Glu Arg Leu Ser Asn Cys Ser Arg Ser 685 690 695	2160
CTC AAG AAT CTG ATT CTT AAG AAT AAT CAA ATC AGG AGT CTG ACG AAG Leu Lys Asn Leu Ile Leu Lys Asn Asn Gln Ile Arg Ser Leu Thr Lys 700 705 710	2208
TAT TTT CTA CAA GAT GCC TTC CAG TTG CGA TAT CTG GAT CTC AGC TCA Tyr Phe Leu Gln Asp Ala Phe Gln Leu Arg Tyr Leu Asp Leu Ser Ser 715 720 725 730	2256
AAT AAA ATC CAG ATG ATC CAA AAG ACC AGC TTC CCA GAA AAT GTC CTC Asn Lys Ile Gln Met Ile Gln Lys Thr Ser Phe Pro Glu Asn Val Leu 735 740 745	2304

AAC AAT CTG AAG ATG TTG CTT TTG CAT CAT AAT CGG TTT CTG TGC ACC Asn Asn Leu Lys Met Leu Leu Leu His His Asn Arg Phe Leu Cys Thr 750 755 760	2352
TGT GAT GCT GTG TGG TTT GTC TGG TGG GTT AAC CAT ACG GAG GTG ACT Cys Asp Ala Val Trp Phe Val Trp Trp Val Asn His Thr Glu Val Thr 765 770 775	2400
ATT CCT TAC CTG GCC ACA GAT GTG ACT TGT GTG GGG CCA GGA GCA CAC Ile Pro Tyr Leu Ala Thr Asp Val Thr Cys Val Gly Pro Gly Ala His 780 785 790	2448
AAG GGC CAA AGT GTG ATC TCC CTG GAT CTG TAC ACC TGT GAG TTA GAT Lys Gly Gln Ser Val Ile Ser Leu Asp Leu Tyr Thr Cys Glu Leu Asp 795 800 805 810	2496
CTG ACT AAC CTG ATT CTG TTC TCA CTT TCC ATA TCT GTA TCT CTC TTT Leu Thr Asn Leu Ile Leu Phe Ser Leu Ser Ile Ser Val Ser Leu Phe 815 820 825	2544
CTC ATG GTG ATG ATG ACA GCA AGT CAC CTC TAT TTC TGG GAT GTG TGG Leu Met Val Met Met Thr Ala Ser His Leu Tyr Phe Trp Asp Val Trp 830 835 840	2592
TAT ATT TAC CAT TTC TGT AAG GCC AAG ATA AAG GGG TAT CAG CGT CTA Tyr Ile Tyr His Phe Cys Lys Ala Lys Ile Lys Gly Tyr Gln Arg Leu 845 850 855	2640
ATA TCA CCA GAC TGT TGC TAT GAT GCT TTT ATT GTG TAT GAC ACT AAA Ile Ser Pro Asp Cys Cys Tyr Asp Ala Phe Ile Val Tyr Asp Thr Lys 860 865 870	2688
GAC CCA GCT GTG ACC GAG TGG GTT TTG GCT GAG CTG GTG GCC AAA CTG Asp Pro Ala Val Thr Glu Trp Val Leu Ala Glu Leu Val Ala Lys Leu 875 880 885 890	2736
GAA GAC CCA AGA GAG AAA CAT TTT AAT TTA TGT CTC GAG GAA AGG GAC Glu Asp Pro Arg Glu Lys His Phe Asn Leu Cys Leu Glu Glu Arg Asp 895 900 905	2784
TGG TTA CCA GGG CAG CCA GTT CTG GAA AAC CTT TCC CAG AGC ATA CAG Trp Leu Pro Gly Gln Pro Val Leu Glu Asn Leu Ser Gln Ser Ile Gln 910 915 920	2832
CTT AGC AAA AAG ACA GTG TTT GTG ATG ACA GAC AAG TAT GCA AAG ACT Leu Ser Lys Lys Thr Val Phe Val Met Thr Asp Lys Tyr Ala Lys Thr 925 930 935	2880

GAA AAT TTT AAG ATA GCA TTT TAC TTG TCC CAT CAG AGG CTC ATG GAT	2928
Glu Asn Phe Lys Ile Ala Phe Tyr Leu Ser His Gln Arg Leu Met Asp	
940 945 950	
GAA AAA GTT GAT GTG ATT ATC TTG ATA TTT CTT GAG AAG CCC TTT CAG	2976
Glu Lys Val Asp Val Ile Ile Leu Ile Phe Leu Glu Lys Pro Phe Gln	
955 960 965 970	
AAG TCC AAG TTC CTC CAG CTC CGG AAA AGG CTC TGT GGG AGT TCT GTC	3024
Lys Ser Lys Phe Leu Gln Leu Arg Lys Arg Leu Cys Gly Ser Ser Val	
975 980 985	
CTT GAG TGG CCA ACA AAC CCG CAA GCT CAC CCA TAC TTC TGG CAG TGT	3072
Leu Glu Trp Pro Thr Asn Pro Gln Ala His Pro Tyr Phe Trp Gln Cys	
990 995 1000	
CTA AAG AAC GCC CTG GCC ACA GAC AAT CAT GTG GCC TAT AGT CAG GTG	3120
Leu Lys Asn Ala Leu Ala Thr Asp Asn His Val Ala Tyr Ser Gln Val	
1005 1010 1015	
TTC AAG GAA ACG GTC TAG	3138
Phe Lys Glu Thr Val	
1020	

(2) SEQ ID NO: 12 的信息:

- (i) 顺序特征:
- (A) 长度: 1045 个氨基酸
 - (B) 类型: 氨基酸
 - (D) 拓扑学: 线性
- (ii) 分子类型: 蛋白质
- (xi) 顺序描述: SEQ ID NO: 12:

Met Trp Thr Leu Lys Arg Leu Ile Leu Ile Leu Phe Asn Ile Ile Leu	
-22 -20 -15 -10	
Ile Ser Lys Leu Leu Gly Ala Arg Trp Phe Pro Lys Thr Leu Pro Cys	
-5 1 5 10	
Asp Val Thr Leu Asp Val Pro Lys Asn His Val Ile Val Asp Cys Thr	
15 20 25	
Asp Lys His Leu Thr Glu Ile Pro Gly Gly Ile Pro Thr Asn Thr Thr	
30 35 40	

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

Phe Leu Ala Lys Glu Ile Gly Asp Ala Lys Phe Leu His Phe Leu Pro
 300 305 310
 Ser Leu Ile Gln Leu Asp Leu Ser Phe Asn Phe Glu Leu Gln Val Tyr
 315 320 325 330
 Arg Ala Ser Met Asn Leu Ser Gln Ala Phe Ser Ser Leu Lys Ser Leu
 335 340 345
 Lys Ile Leu Arg Ile Arg Gly Tyr Val Phe Lys Glu Leu Lys Ser Phe
 350 355 360
 Asn Leu Ser Pro Leu His Asn Leu Gln Asn Leu Glu Val Leu Asp Leu
 365 370 375
 Gly Thr Asn Phe Ile Lys Ile Ala Asn Leu Ser Met Phe Lys Gln Phe
 380 385 390
 Lys Arg Leu Lys Val Ile Asp Leu Ser Val Asn Lys Ile Ser Pro Ser
 395 400 405 410
 Gly Asp Ser Ser Glu Val Gly Phe Cys Ser Asn Ala Arg Thr Ser Val
 415 420 425
 Glu Ser Tyr Glu Pro Gln Val Leu Glu Gln Leu His Tyr Phe Arg Tyr
 430 435 440
 Asp Lys Tyr Ala Arg Ser Cys Arg Phe Lys Asn Lys Glu Ala Ser Phe
 445 450 455
 Met Ser Val Asn Glu Ser Cys Tyr Lys Tyr Gly Gln Thr Leu Asp Leu
 460 465 470
 Ser Lys Asn Ser Ile Phe Phe Val Lys Ser Ser Asp Phe Gln His Leu
 475 480 485 490
 Ser Phe Leu Lys Cys Leu Asn Leu Ser Gly Asn Leu Ile Ser Gln Thr
 495 500 505
 Leu Asn Gly Ser Glu Phe Gln Pro Leu Ala Glu Leu Arg Tyr Leu Asp
 510 515 520
 Phe Ser Asn Asn Arg Leu Asp Leu Leu His Ser Thr Ala Phe Glu Glu
 525 530 535
 Leu His Lys Leu Glu Val Leu Asp Ile Ser Ser Asn Ser His Tyr Phe
 540 545 550

Gln Ser Glu Gly Ile Thr His Met Leu Asn Phe Thr Lys Asn Leu Lys
 555 560 565 570
 Val Leu Gln Lys Leu Met Met Asn Asp Asn Asp Ile Ser Ser Ser Thr
 575 580 585
 Ser Arg Thr Met Glu Ser Glu Ser Leu Arg Thr Leu Glu Phe Arg Gly
 590 595 600
 Asn His Leu Asp Val Leu Trp Arg Glu Gly Asp Asn Arg Tyr Leu Gln
 605 610 615
 Leu Phe Lys Asn Leu Leu Lys Leu Glu Glu Leu Asp Ile Ser Lys Asn
 620 625 630
 Ser Leu Ser Phe Leu Pro Ser Gly Val Phe Asp Gly Met Pro Pro Asn
 635 640 645 650
 Leu Lys Asn Leu Ser Leu Ala Lys Asn Gly Leu Lys Ser Phe Ser Trp
 655 660 665
 Lys Lys Leu Gln Cys Leu Lys Asn Leu Glu Thr Leu Asp Leu Ser His
 670 675 680
 Asn Gln Leu Thr Thr Val Pro Glu Arg Leu Ser Asn Cys Ser Arg Ser
 685 690 695
 Leu Lys Asn Leu Ile Leu Lys Asn Asn Gln Ile Arg Ser Leu Thr Lys
 700 705 710
 Tyr Phe Leu Gln Asp Ala Phe Gln Leu Arg Tyr Leu Asp Leu Ser Ser
 715 720 725 730
 Asn Lys Ile Gln Met Ile Gln Lys Thr Ser Phe Pro Glu Asn Val Leu
 735 740 745
 Asn Asn Leu Lys Met Leu Leu Leu His His Asn Arg Phe Leu Cys Thr
 750 755 760
 Cys Asp Ala Val Trp Phe Val Trp Trp Val Asn His Thr Glu Val Thr
 765 770 775
 Ile Pro Tyr Leu Ala Thr Asp Val Thr Cys Val Gly Pro Gly Ala His
 780 785 790
 Lys Gly Gln Ser Val Ile Ser Leu Asp Leu Tyr Thr Cys Glu Leu Asp
 795 800 805 810

```

Leu Thr Asn Leu Ile Leu Phe Ser Leu Ser Ile Ser Val Ser Leu Phe
      815                      820                      825

Leu Met Val Met Met Thr Ala Ser His Leu Tyr Phe Trp Asp Val Trp
      830                      835                      840

Tyr Ile Tyr His Phe Cys Lys Ala Lys Ile Lys Gly Tyr Gln Arg Leu
      845                      850                      855

Ile Ser Pro Asp Cys Cys Tyr Asp Ala Phe Ile Val Tyr Asp Thr Lys
      860                      865                      870

Asp Pro Ala Val Thr Glu Trp Val Leu Ala Glu Leu Val Ala Lys Leu
      875                      880                      885                      890

Glu Asp Pro Arg Glu Lys His Phe Asn Leu Cys Leu Glu Glu Arg Asp
      895                      900                      905

Trp Leu Pro Gly Gln Pro Val Leu Glu Asn Leu Ser Gln Ser Ile Gln
      910                      915                      920

Leu Ser Lys Lys Thr Val Phe Val Met Thr Asp Lys Tyr Ala Lys Thr
      925                      930                      935

Glu Asn Phe Lys Ile Ala Phe Tyr Leu Ser His Gln Arg Leu Met Asp
      940                      945                      950

Glu Lys Val Asp Val Ile Ile Leu Ile Phe Leu Glu Lys Pro Phe Gln
      955                      960                      965                      970

Lys Ser Lys Phe Leu Gln Leu Arg Lys Arg Leu Cys Gly Ser Ser Val
      975                      980                      985

Leu Glu Trp Pro Thr Asn Pro Gln Ala His Pro Tyr Phe Trp Gln Cys
      990                      995                      1000

Leu Lys Asn Ala Leu Ala Thr Asp Asn His Val Ala Tyr Ser Gln Val
      1005                      1010                      1015

Phe Lys Glu Thr Val
      1020

```

(2) SEQ ID NO: 13 的信息:

- (i) 顺序特征:
 - (A) 长度: 180 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链

(D) 拓扑学：线性

(ii) 分子类型：cDNA

(ix) 特征：

(A) 名称/关键词：CDS

(B) 位置：1...177

(xi) 顺序描述：SEQ ID NO: 13:

CTT GGA AAA CCT CTT CAG AAG TCT AAG TTT CTT CAG CTC AGG AAG AGA	48
Leu Gly Lys Pro Leu Gln Lys Ser Lys Phe Leu Gln Leu Arg Lys Arg	
1 5 10 15	
CTC TGC AGG AGC TCT GTC CTT GAG TGG CCT GCA AAT CCA CAG GCT CAC	96
Leu Cys Arg Ser Ser Val Leu Glu Trp Pro Ala Asn Pro Gln Ala His	
20 25 30	
CCA TAC TTC TGG CAG TGC CTG AAA AAT GCC CTG ACC ACA GAC AAT CAT	144
Pro Tyr Phe Trp Gln Cys Leu Lys Asn Ala Leu Thr Thr Asp Asn His	
35 40 45	
GTG GCT TAT AGT CAA ATG TTC AAG GAA ACA GTC TAG	180
Val Ala Tyr Ser Gln Met Phe Lys Glu Thr Val	
50 55	

(2) SEQ ID NO: 14 的信息：

(i) 顺序特征：

(A) 长度：59 个氨基酸

(B) 类型：氨基酸

(D) 拓扑学：线性

(ii) 分子类型：蛋白质

(xi) 顺序描述：SEQ ID NO: 14:

Leu Gly Lys Pro Leu Gln Lys Ser Lys Phe Leu Gln Leu Arg Lys Arg	
1 5 10 15	
Leu Cys Arg Ser Ser Val Leu Glu Trp Pro Ala Asn Pro Gln Ala His	
20 25 30	
Pro Tyr Phe Trp Gln Cys Leu Lys Asn Ala Leu Thr Thr Asp Asn His	
35 40 45	
Val Ala Tyr Ser Gln Met Phe Lys Glu Thr Val	
50 55	

(2) SEQ ID NO: 15 的信息:

(i) 顺序特征：

- (A) 长度: 990 个碱基对
(B) 类型: 核酸
(C) 链型: 单链
(D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

- (A) 名称/关键词: CDS
(B) 位置: 2...988

(xi) 顺序描述: SEQ ID NO: 15:

G AAT TCC AGA CTT ATA AAC TTG AAA AAT CTC TAT TTG GCC TGG AAC	46
Asn Ser Arg Leu Ile Asn Leu Lys Asn Leu Tyr Leu Ala Trp Asn	
1 5 10 15	
TGC TAT TTT AAC AAA GTT TGC GAG AAA ACT AAC ATA GAA GAT GGA GTA	94
Cys Tyr Phe Asn Lys Val Cys Glu Lys Thr Asn Ile Glu Asp Gly Val	
20 25 30	
TTT GAA ACG CTG ACA AAT TTG GAG TTG CTA TCA CTA TCT TTC AAT TCT	142
Phe Glu Thr Leu Thr Asn Leu Glu Leu Leu Ser Leu Ser Phe Asn Ser	
35 40 45	
CTT TCA CAT GTG CCA CCC AAA CTG CCA AGC TCC CTA CGC AAA CTT TTT	190
Leu Ser His Val Pro Pro Lys Leu Pro Ser Ser Leu Arg Lys Leu Phe	
50 55 60	
CTG AGC AAC ACC CAG ATC AAA TAC ATT AGT GAA GAA GAT TTC AAG GGA	238
Leu Ser Asn Thr Gln Ile Lys Tyr Ile Ser Glu Glu Asp Phe Lys Gly	
65 70 75	
TTG ATA AAT TTA ACA TTA CTA GAT TTA AGC GGG AAC TGT CCG AGG TGC	286
Leu Ile Asn Leu Thr Leu Leu Asp Leu Ser Gly Asn Cys Pro Arg Cys	
80 85 90 95	
TTC AAT GCC CCA TTT CCA TGC GTG CCT TGT GAT GGT GGT GCT TCA ATT	334
Phe Asn Ala Pro Phe Pro Cys Val Pro Cys Asp Gly Gly Ala Ser Ile	
100 105 110	
AAT ATA GAT CGT TTT GCT TTT CAA AAC TTG ACC CAA CTT CGA TAC CTA	382
Asn Ile Asp Arg Phe Ala Phe Gln Asn Leu Thr Gln Leu Arg Tyr Leu	
115 120 125	

AAC CTC TCT AGC ACT TCC CTC AGG AAG ATT AAT GCT GCC TGG TTT AAA Asn Leu Ser Ser Thr Ser Leu Arg Lys Ile Asn Ala Ala Trp Phe Lys 130 135 140	430
AAT ATG CCT CAT CTG AAG GTG CTG GAT CTT GAA TTC AAC TAT TTA GTG Asn Met Pro His Leu Lys Val Leu Asp Leu Glu Phe Asn Tyr Leu Val 145 150 155	478
GGA GAA ATA GCC TCT GGG GCA TTT TTA ACG ATG CTG CCC CGC TTA GAA Gly Glu Ile Ala Ser Gly Ala Phe Leu Thr Met Leu Pro Arg Leu Glu 160 165 170 175	526
ATA CTT GAC TTG TCT TTT AAC TAT ATA AAG GGG AGT TAT CCA CAG CAT Ile Leu Asp Leu Ser Phe Asn Tyr Ile Lys Gly Ser Tyr Pro Gln His 180 185 190	574
ATT AAT ATT TCC AGA AAC TTC TCT AAA CTT TTG TCT CTA CGG GCA TTG Ile Asn Ile Ser Arg Asn Phe Ser Lys Leu Leu Ser Leu Arg Ala Leu 195 200 205	622
CAT TTA AGA GGT TAT GTG TTC CAG GAA CTC AGA GAA GAT GAT TTC CAG His Leu Arg Gly Tyr Val Phe Gln Glu Leu Arg Glu Asp Asp Phe Gln 210 215 220	670
CCC CTG ATG CAG CTT CCA AAC TTA TCG ACT ATC AAC TTG GGT ATT AAT Pro Leu Met Gln Leu Pro Asn Leu Ser Thr Ile Asn Leu Gly Ile Asn 225 230 235	718
TTT ATT AAG CAA ATC GAT TTC AAA CTT TTC CAA AAT TTC TCC AAT CTG Phe Ile Lys Gln Ile Asp Phe Lys Leu Phe Gln Asn Phe Ser Asn Leu 240 245 250 255	766
GAA ATT ATT TAC TTG TCA GAA AAC AGA ATA TCA CCG TTG GTA AAA GAT Glu Ile Ile Tyr Leu Ser Glu Asn Arg Ile Ser Pro Leu Val Lys Asp 260 265 270	814
ACC CGG CAG AGT TAT GCA AAT AGT TCC TCT TTT CAA CGT CAT ATC CGG Thr Arg Gln Ser Tyr Ala Asn Ser Ser Ser Phe Gln Arg His Ile Arg 275 280 285	862
AAA CGA CGC TCA ACA GAT TTT GAG TTT GAC CCA CAT TCG AAC TTT TAT Lys Arg Arg Ser Thr Asp Phe Glu Phe Asp Pro His Ser Asn Phe Tyr 290 295 300	910
CAT TTC ACC CGT CCT TTA ATA AAG CCA CAA TGT GCT GCT TAT GGA AAA His Phe Thr Arg Pro Leu Ile Lys Pro Gln Cys Ala Ala Tyr Gly Lys 305 310 315	958

GCC TTA GAT TTA AGC CTC AAC AGT ATT TTC TT
 Ala Leu Asp Leu Ser Leu Asn Ser Ile Phe
 320 325

990

(2) SEQ ID NO: 16 的信息:

- (i) 顺序特征:
 (A) 长度: 329 个氨基酸
 (B) 类型: 氨基酸
 (D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 顺序描述: SEQ ID NO: 16:

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

```

Leu Asp Leu Ser Phe Asn Tyr Ile Lys Gly Ser Tyr Pro Gln His Ile
    180                      185                      190

Asn Ile Ser Arg Asn Phe Ser Lys Leu Leu Ser Leu Arg Ala Leu His
    195                      200                      205

Leu Arg Gly Tyr Val Phe Gln Glu Leu Arg Glu Asp Asp Phe Gln Pro
    210                      215                      220

Leu Met Gln Leu Pro Asn Leu Ser Thr Ile Asn Leu Gly Ile Asn Phe
    225                      230                      235                      240

Ile Lys Gln Ile Asp Phe Lys Leu Phe Gln Asn Phe Ser Asn Leu Glu
    245                      250                      255

Ile Ile Tyr Leu Ser Glu Asn Arg Ile Ser Pro Leu Val Lys Asp Thr
    260                      265                      270

Arg Gln Ser Tyr Ala Asn Ser Ser Ser Phe Gln Arg His Ile Arg Lys
    275                      280                      285

Arg Arg Ser Thr Asp Phe Glu Phe Asp Pro His Ser Asn Phe Tyr His
    290                      295                      300

Phe Thr Arg Pro Leu Ile Lys Pro Gln Cys Ala Ala Tyr Gly Lys Ala
    305                      310                      315                      320

Leu Asp Leu Ser Leu Asn Ser Ile Phe
    325

```

(2) SEQ ID NO: 17 的信息:

- (i) 顺序特征:
 - (A) 长度: 1557 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线性
- (ii) 分子类型: cDNA
- (ix) 特征:
 - (A) 名称/关键词: CDS
 - (B) 位置: 1...513
- (ix) 特征:
 - (A) 名称/关键词: misc 特征
 - (B) 位置: 278

(D) 其它信息：/注释 = “核苷酸 278 标示为 G，可以为 G 或 C”

(ix) 特征：

(A) 名称/关键词：misc 特征

(B) 位置：445

(D) 其它信息：/注释 = “核苷酸 445 标示为 A，可以为 A 或 T”

(ix) 特征：

(A) 名称/关键词：misc 特征

(B) 位置：572

(D) 其它信息：/注释 = “核苷酸 572、593、600、607、617、622、631、640、646、653、719、755 和 861 标示为 C；每个可以为 A、C、G 或 T”

(xi) 顺序描述：SEQ ID NO: 17:

CAG TCT CTT TCC ACA TCC CAA ACT TTC TAT GAT GCT TAC ATT TCT TAT	48
Gln Ser Leu Ser Thr Ser Gln Thr Phe Tyr Asp Ala Tyr Ile Ser Tyr	
1 5 10 15	
GAC ACC AAA GAT GCC TCT GTT ACT GAC TGG GTG ATA AAT GAG CTG CGC	96
Asp Thr Lys Asp Ala Ser Val Thr Asp Trp Val Ile Asn Glu Leu Arg	
20 25 30	
TAC CAC CTT GAA GAG AGC CGA GAC AAA AAC GTT CTC CTT TGT CTA GAG	144
Tyr His Leu Glu Glu Ser Arg Asp Lys Asn Val Leu Leu Cys Leu Glu	
35 40 45	
GAG AGG GAT TGG GAC CCG GGA TTG GCC ATC ATC GAC AAC CTC ATG CAG	192
Glu Arg Asp Trp Asp Pro Gly Leu Ala Ile Ile Asp Asn Leu Met Gln	
50 55 60	
AGC ATC AAC CAA AGC AAG AAA ACA GTA TTT GTT TTA ACC AAA AAA TAT	240
Ser Ile Asn Gln Ser Lys Lys Thr Val Phe Val Leu Thr Lys Lys Tyr	
65 70 75 80	
GCA AAA AGC TGG AAC TTT AAA ACA GCT TTT TAC TTG GGC TTG CAG AGG	288
Ala Lys Ser Trp Asn Phe Lys Thr Ala Phe Tyr Leu Gly Leu Gln Arg	
85 90 95	
CTA ATG GGT GAG AAC ATG GAT GTG ATT ATA TTT ATC CTG CTG GAG CCA	336
Leu Met Gly Glu Asn Met Asp Val Ile Ile Phe Ile Leu Leu Glu Pro	
100 105 110	
GTG TTA CAG CAT TCT CCG TAT TTG AGG CTA CGG CAG CGG ATC TGT AAG	384
Val Leu Gln His Ser Pro Tyr Leu Arg Leu Arg Gln Arg Ile Cys Lys	
115 120 125	

AGC TCC ATC CTC CAG TGG CCT GAC AAC CCG AAG GCA GAA AGG TTG TTT Ser Ser Ile Leu Gln Trp Pro Asp Asn Pro Lys Ala Glu Arg Leu Phe 130 135 140	432
TGG CAA ACT CTG AGA AAT GTG GTC TTG ACT GAA AAT GAT TCA CGG TAT Trp Gln Thr Leu Arg Asn Val Val Leu Thr Glu Asn Asp Ser Arg Tyr 145 150 155 160	480
AAC AAT ATG TAT GTC GAT TCC ATT AAG CAA TAC TAACTGACGT TAAGTCATGA Asn Asn Met Tyr Val Asp Ser Ile Lys Gln Tyr 165 170	533
TTTCGCGCCA TAATAAAGAT GCAAAGGAAT GACATTTCCG TATTAGTTAT CTATTGCTAC	593
GGTAACCAAA TTA CTCCCAA AAACCTTACG TCGGTTTCAA AACAACCACA TTCTGCTGGC	653
CCCACAGTTT TTGAGGGTCA GGAGTCCAGG CCCAGCATAA CTGGGTCTTC TGCTTCAGGG	713
TGTCTCCAGA GGCTGCAATG TAGGTGTTCA CCAGAGACAT AGGCATCACT GGGGTCACAC	773
TCCATGTGGT TGTTTTCTGG ATTCAATTCC TCCTGGGCTA TTGGCCAAAG GCTATACTCA	833
TGTAAGCCAT GCGAGCCTAT CCCACAACGG CAGCTTGCTT CATCAGAGCT AGCAAAAAAG	893
AGAGGTTGCT AGCAAGATGA AGTCACAATC TTTTGTAATC GAATCAAAAA AGTGATATCT	953
CATCACTTTG GCCATATTCT ATTTGTTAGA AGTAAACCAC AGGTCCCACC AGCTCCATGG	1013
GAGTGACCAC CTCAGTCCAG GGAAAACAGC TGAAGACCAA GATGGTGAGC TCTGATTGCT	1073
TCAGTTGGTC ATCAACTATT TTCCCTTGAC TGCTGTCTTG GGATGGCCGG CTATCTTGAT	1133
GGATAGATTG TGAATATCAG GAGGCCAGGG ATCACTGTGG ACCATCTTAG CAGTTGACCT	1193
AACACATCTT CTTTTCATA TCTAAGAACT TTTGCCACTG TGAATAATGG TCCTAATATT	1253
AAGCTGTTGT TTATATTTAT CATATATCTA TGGCTACATG GTTATATTAT GCTGTGGTTG	1313
CGTTCGGTTT TATTTACAGT TGCTTTTACA AATATTTGCT GTAACATTTG ACTTCTAAGG	1373
TTTAGATGCC ATTTAAGAAC TGAGATGGAT AGCTTTTAAA GCATCTTTTA CTTCTTACCA	1433
TTTTTTAAAA GTATGCAGCT AAATTCGAAG CTTTGTGCT ATATTGTAA TTGCCATTGC	1493
TGTAAATCTT AAAATGAATG AATAAAAATG TTTCAATTTA AAAAAAAAAA AAAAAAAAAA	1553
AAAA	1557

(2) SEQ ID NO: 18 的信息:

(i) 顺序特征:

(A) 长度: 171 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 顺序描述: SEQ ID NO: 18:

```

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

```

(2) SEQ ID NO: 19 的信息:

(i) 顺序特征:

(A) 长度: 629 个碱基对

(B) 类型: 核酸
(C) 链型: 单链
(D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

(A) 名称/关键词: CDS

(B) 位置: 1...486

(ix) 特征:

(A) 名称/关键词: misc 特征

(B) 位置: 144

(D) 其它信息: / 注释 = “核苷酸 144 和 225 标示为 C; 可以为 C 或 T”

(xi) 顺序描述: SEQ ID NO: 19:

AAT GAA TTG ATC CCC AAT CTA GAG AAG GAA GAT GGT TCT ATC TTG ATT 48
Asn Glu Leu Ile Pro Asn Leu Glu Lys Glu Asp Gly Ser Ile Leu Ile
1 5 10 15

TGC CTT TAT GAA AGC TAC TTT GAC CCT GGC AAA AGC ATT AGT GAA AAT 96
Cys Leu Tyr Glu Ser Tyr Phe Asp Pro Gly Lys Ser Ile Ser Glu Asn
20 25 30

ATT GTA AGC TTC ATT GAG AAA AGC TAT AAG TCC ATC TTT GTT TTG TCC 144
Ile Val Ser Phe Ile Glu Lys Ser Tyr Lys Ser Ile Phe Val Leu Ser
35 40 45

CCC AAC TTT GTC CAG AAT GAG TGG TGC CAT TAT GAA TTC TAC TTT GCC 192
Pro Asn Phe Val Gln Asn Glu Trp Cys His Tyr Glu Phe Tyr Phe Ala
50 55 60

CAC CAC AAT CTC TTC CAT GAA AAT TCT GAT CAC ATA ATT CTT ATC TTA 240
His His Asn Leu Phe His Glu Asn Ser Asp His Ile Ile Leu Ile Leu
65 70 75 80

CTG GAA CCC ATT CCA TTC TAT TGC ATT CCC ACC AGG TAT CAT AAA CTG 288
Leu Glu Pro Ile Pro Phe Tyr Cys Ile Pro Thr Arg Tyr His Lys Leu
85 90 95

GAA GCT CTC CTG GAA AAA AAA GCA TAC TTG GAA TGG CCC AAG GAT AGG 336
Glu Ala Leu Leu Glu Lys Lys Ala Tyr Leu Glu Trp Pro Lys Asp Arg
100 105 110

CGT AAA TGT GGG CTT TTC TGG GCA AAC CTT CGA GCT GCT GTT AAT GTT 384
Arg Lys Cys Gly Leu Phe Trp Ala Asn Leu Arg Ala Ala Val Asn Val
115 120 125

AAT GTA TTA GCC ACC AGA GAA ATG TAT GAA CTG CAG ACA TTC ACA GAG 432
 Asn Val Leu Ala Thr Arg Glu Met Tyr Glu Leu Gln Thr Phe Thr Glu
 130 135 140

TTA AAT GAA GAG TCT CGA GGT TCT ACA ATC TCT CTG ATG AGA ACA GAC 480
 Leu Asn Glu Glu Ser Arg Gly Ser Thr Ile Ser Leu Met Arg Thr Asp
 145 150 155 160

TGT CTA TAAATCCCA CAGTCCTTGG GAAGTTGGGG ACCACATACA CTGTTGGGAT 536
 Cys Leu

GTACATTGAT ACAACCTTTA TGATGGCAAT TTGACAATAT TTATTAAAAT AAAAAATGGT 596

TATCCCTTC AAAAAAAAAA AAAAAAAAAA AAA 629

(2) SEQ ID NO: 20 的信息:

- (i) 顺序特征:
 (A) 长度: 162 个氨基酸
 (B) 类型: 氨基酸
 (D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 顺序描述: SEQ ID NO: 20:

Asn Glu Leu Ile Pro Asn Leu Glu Lys Glu Asp Gly Ser Ile Leu Ile
 1 5 10 15

Cys Leu Tyr Glu Ser Tyr Phe Asp Pro Gly Lys Ser Ile Ser Glu Asn
 20 25 30

Ile Val Ser Phe Ile Glu Lys Ser Tyr Lys Ser Ile Phe Val Leu Ser
 35 40 45

Pro Asn Phe Val Gln Asn Glu Trp Cys His Tyr Glu Phe Tyr Phe Ala
 50 55 60

His His Asn Leu Phe His Glu Asn Ser Asp His Ile Ile Leu Ile Leu
 65 70 75 80

Leu Glu Pro Ile Pro Phe Tyr Cys Ile Pro Thr Arg Tyr His Lys Leu
 85 90 95

Glu Ala Leu Leu Glu Lys Lys Ala Tyr Leu Glu Trp Pro Lys Asp Arg
 100 105 110

Arg Lys Cys Gly Leu Phe Trp Ala Asn Leu Arg Ala Ala Val Asn Val
115 120 125

Asn Val Leu Ala Thr Arg Glu Met Tyr Glu Leu Gln Thr Phe Thr Glu
130 135 140

Leu Asn Glu Glu Ser Arg Gly Ser Thr Ile Ser Leu Met Arg Thr Asp
145 150 155 160

Cys Leu

(2) SEQ ID NO: 21 的信息:

(i) 顺序特征:

- (A) 长度: 427 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

- (A) 名称/关键词: CDS
- (B) 位置: 1...426

(xi) 顺序描述: SEQ ID NO: 21:

AAG AAC TCC AAA GAA AAC CTC CAG TTT CAT GCT TTT ATT TCA TAT AGT	48
Lys Asn Ser Lys Glu Asn Leu Gln Phe His Ala Phe Ile Ser Tyr Ser	
1 5 10 15	
GAA CAT GAT TCT GCC TGG GTG AAA AGT GAA TTG GTA CCT TAC CTA GAA	96
Glu His Asp Ser Ala Trp Val Lys Ser Glu Leu Val Pro Tyr Leu Glu	
20 25 30	
AAA GAA GAT ATA CAG ATT TGT CTT CAT GAG AGA AAC TTT GTC CCT GGC	144
Lys Glu Asp Ile Gln Ile Cys Leu His Glu Arg Asn Phe Val Pro Gly	
35 40 45	
AAG AGC ATT GTG GAA AAT ATC ATC AAC TGC ATT GAG AAG AGT TAC AAG	192
Lys Ser Ile Val Glu Asn Ile Ile Asn Cys Ile Glu Lys Ser Tyr Lys	
50 55 60	
TCC ATC TTT GTT TTG TCT CCC AAC TTT GTC CAG AGT GAG TGG TGC CAT	240
Ser Ile Phe Val Leu Ser Pro Asn Phe Val Gln Ser Glu Trp Cys His	
65 70 75 80	

TAC GAA CTC TAT TTT GCC CAT CAC AAT CTC TTT CAT GAA GGA TCT AAT	288
Tyr Glu Leu Tyr Phe Ala His His Asn Leu Phe His Glu Gly Ser Asn	
85 90 95	
AAC TTA ATC CTC ATC TTA CTG GAA CCC ATT CCA CAG AAC AGC ATT CCC	336
Asn Leu Ile Leu Ile Leu Leu Glu Pro Ile Pro Gln Asn Ser Ile Pro	
100 105 110	
AAC AAG TAC CAC AAG CTG AAG GCT CTC ATC ACG CAG CGG ACT TAT TTG	384
Asn Lys Tyr His Lys Leu Lys Ala Leu Met Thr Gln Arg Thr Tyr Leu	
115 120 125	
CAG TGG CCC AAG GAG AAA AGC AAA CGT GGG CTC TTT TGG GCT	426
Gln Trp Pro Lys Glu Lys Ser Lys Arg Gly Leu Phe Trp Ala	
130 135 140	
A	427

(2) SEQ ID NO: 22 的信息:

(i) 顺序特征:

(A) 长度: 142 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 顺序描述: SEQ ID NO: 22:

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

Asn Leu Ile Leu Ile Leu Leu Glu Pro Ile Pro Gln Asn Ser Ile Pro
 100 105 110

Asn Lys Tyr His Lys Leu Lys Ala Leu Met Thr Gln Arg Thr Tyr Leu
 115 120 125

Gln Trp Pro Lys Glu Lys Ser Lys Arg Gly Leu Phe Trp Ala
 130 135 140

(2) SEQ ID NO: 23 的信息:

(i) 顺序特征:

- (A) 长度: 662 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

- (A) 名称/关键词: CDS
- (B) 位置: 1...627

(ix) 特征:

- (A) 名称/关键词: misc 特征
- (B) 位置: 54
- (D) 其它信息: /注释 = "核苷酸 54、103 和 345 标示为 A; 每个可以为 A 或 G"

(ix) 特征:

- (A) 名称/关键词: misc 特征
- (B) 位置: 313
- (D) 其它信息: /注释 = "核苷酸 313 标示为 G; 可以为 G 或 T"

(ix) 特征:

- (A) 名称/关键词: misc 特征
- (B) 位置: 316
- (D) 其它信息: /注释 = "核苷酸 316、380、407 和 408 标示为 C; 可以为 A、C、G 或 T"

(xi) 顺序描述: SEQ ID NO: 23:

GCT TCC ACC TGT GCC TGG CCT GGC TTC CCT GGC GGG GGC GGC AAA GTG 48
 Ala Ser Thr Cys Ala Trp Pro Gly Phe Pro Gly Gly Gly Lys Val
 1 5 10 15

GGC GAA ATG AGG ATG CCC TGC CCT ACG ATG CCT TCG TGG TCT TCG ACA 96
 Gly Glu Met Arg Met Pro Cys Pro Thr Met Pro Ser Trp Ser Ser Thr
 20 25 30

AAA CGC AGA GCG CAG TGG CAG ACT GGG TGT ACA ACG AGC TTC GGG GGC Lys Arg Arg Ala Gln Trp Gln Thr Gly Cys Thr Thr Ser Phe Gly Gly 35 40 45	144
AGC TGG AGG AGT GCC GTG GGC GCT GGG CAC TCC GCC TGT GCC TGG AGG Ser Trp Arg Ser Ala Val Gly Ala Gly His Ser Ala Cys Ala Trp Arg 50 55 60	192
AAC GCG ACT GGC TGC CTG GCA AAA CCC TCT TTG AGA ACC TGT GGG CCT Asn Ala Thr Gly Cys Leu Ala Lys Pro Ser Leu Arg Thr Cys Gly Pro 65 70 75 80	240
CGG TCT ATG GCA GCC GCA AGA CGC TGT TTG TGC TGG CCC ACA CGG ACC Arg Ser Met Ala Ala Ala Arg Arg Cys Leu Cys Trp Pro Thr Arg Thr 85 90 95	288
GGG TCA GTG GTC TCT TGC GCG CCA GTT CTC CTG CTG GCC CAG CAG CGC Gly Ser Val Val Ser Cys Ala Pro Val Leu Leu Leu Ala Gln Gln Arg 100 105 110	336
CTG CTG GAA GAC CGC AAG GAC GTC GTG GTG CTG GTG ATC CTA ACG CCT Leu Leu Glu Asp Arg Lys Asp Val Val Val Leu Val Ile Leu Thr Pro 115 120 125	384
GAC GGC CAA GCC TCC CGA CTA CCC GAT GCG CTG ACC AGC GCC TCT GCC Asp Gly Gln Ala Ser Arg Leu Pro Asp Ala Leu Thr Ser Ala Ser Ala 130 135 140	432
GCC AGA GTG TCC TCC TCT GGC CCC ACC AGC CCA GTG GTC GCG CAG CTT Ala Arg Val Ser Ser Ser Gly Pro Thr Ser Pro Val Val Ala Gln Leu 145 150 155 160	480
CTG AGG CCA GCA TGC ATG GCC CTG ACC AGG GAC AAC CAC CAC TTC TAT Leu Arg Pro Ala Cys Met Ala Leu Thr Arg Asp Asn His His Phe Tyr 165 170 175	528
AAC CGG AAC TTC TGC CAG GGA ACC CAC GGC CGA ATA GCC GTG AGC CGG Asn Arg Asn Phe Cys Gln Gly Thr His Gly Arg Ile Ala Val Ser Arg 180 185 190	576
AAT CCT GCA CGG TGC CAC CTC CAC ACA CAC CTA ACA TAT GCC TGC CTG Asn Pro Ala Arg Cys His Leu His Thr His Leu Thr Tyr Ala Cys Leu 195 200 205	624
ATC TGACCAACAC ATGCTCGCCA CCCTCACCAC ACACC Ile	662

(2) SEQ ID NO: 24 的信息:

(i) 顺序特征:

(A) 长度: 209 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 顺序描述: SEQ ID NO: 24:

Ala Ser Thr Cys Ala Trp Pro Gly Phe Pro Gly Gly Gly Gly Lys Val
 1 5 10 15

Gly Glu Met Arg Met Pro Cys Pro Thr Met Pro Ser Trp Ser Ser Thr
 20 25 30

Lys Arg Arg Ala Gln Trp Gln Thr Gly Cys Thr Thr Ser Phe Gly Gly
 35 40 45

Ser Trp Arg Ser Ala Val Gly Ala Gly His Ser Ala Cys Ala Trp Arg
 50 55 60

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

Arg Ser Met Ala Ala Ala Arg Arg Cys Leu Cys Trp Pro Thr Arg Thr
 85 90 95

Gly Ser Val Val Ser Cys Ala Pro Val Leu Leu Leu Ala Gln Gln Arg
 100 105 110

Leu Leu Glu Asp Arg Lys Asp Val Val Val Leu Val Ile Leu Thr Pro
 115 120 125

Asp Gly Gln Ala Ser Arg Leu Pro Asp Ala Leu Thr Ser Ala Ser Ala
 130 135 140

Ala Arg Val Ser Ser Ser Gly Pro Thr Ser Pro Val Val Ala Gln Leu
 145 150 155 160

Leu Arg Pro Ala Cys Met Ala Leu Thr Arg Asp Asn His His Phe Tyr
 165 170 175

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

Asn Pro Ala Arg Cys His Leu His Thr His Leu Thr Tyr Ala Cys Leu
 195 200 205

Ile

(2) SEQ ID NO: 25 的信息:

- (i) 顺序特征:
 - (A) 长度: 4865 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线性
- (ii) 分子类型: cDNA
- (ix) 特征:
 - (A) 名称/关键词: CDS
 - (B) 位置: 107...2617
- (ix) 特征:
 - (A) 名称/关键词: 成熟肽
 - (B) 位置: 173...2617
- (ix) 特征:
 - (A) 名称/关键词: misc 特征
 - (B) 位置: 81
 - (D) 其它信息: /注释 = "核苷酸 81、3144、3205 和 3563 标示为 A, 每个可以为 A、C、G 或 T"
- (ix) 特征:
 - (A) 名称/关键词: misc 特征
 - (B) 位置: 84
 - (D) 其它信息: /注释 = "核苷酸 84 标示为 C, 可以为 C 或 G"
- (ix) 特征:
 - (A) 名称/关键词: misc 特征
 - (B) 位置: 739
 - (D) 其它信息: /注释 = "核苷酸 739 标示为 C, 可以为 C 或 T"
- (ix) 特征:
 - (A) 名称/关键词: misc 特征
 - (B) 位置: 3132
 - (D) 其它信息: /注释 = "核苷酸 3132、3532、3538 和 3553 标示为 G, 每个可以为 G 或 T"
- (ix) 特征:
 - (A) 名称/关键词: misc 特征
 - (B) 位置: 3638
 - (D) 其它信息: /注释 = "核苷酸 A, 可以为 A 或 T"

(ix) 特征:

(A) 名称/关键词: misc 特征

(B) 位置: 3677

(D) 其它信息: /注释 = "核苷酸 3677、3685 和 3736 标示为 C, 每个可以为 A 或 C"

(xi) 顺序描述: SEQ ID NO: 25:

```

AAAATACTCC CTTGCCTCAA AAAGTGTCTG GTCAAACGGT GATAGCAAAC CACGCATTCA      60
CAGGGCCACT GCTGCTCACA AAACCACTGA GGATGATGCC AGGATG ATG TCT GCC      115
                               Met Ser Ala
                               -22    -20

TCG CGC CTG GCT GGG ACT CTG ATC CCA GCC ATG GCC TTC CTC TCC TGC      163
Ser Arg Leu Ala Gly Thr Leu Ile Pro Ala Met Ala Phe Leu Ser Cys
                -15                -10                -5

GTG AGA CCA GAA AGC TGG GAG CCC TGC GTG GAG GTT CCT AAT ATT ACT      211
Val Arg Pro Glu Ser Trp Glu Pro Cys Val Glu Val Pro Asn Ile Thr
                1                5                10

TAT CAA TGC ATG GAG CTG AAT TTC TAC AAA ATC CCC GAC AAC CTC CCC      259
Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys Ile Pro Asp Asn Leu Pro
                15                20                25

TTC TCA ACC AAG AAC CTG GAC CTG AGC TTT AAT CCC CTG AGG CAT TTA      307
Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe Asn Pro Leu Arg His Leu
                30                35                40                45

GGC AGC TAT AGC TTC TTC AGT TTC CCA GAA CTG CAG GTG CTG GAT TTA      355
Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu Leu Gln Val Leu Asp Leu
                50                55                60

TCC AGG TGT GAA ATC CAG ACA ATT GAA GAT GGG GCA TAT CAG AGC CTA      403
Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp Gly Ala Tyr Gln Ser Leu
                65                70                75

AGC CAC CTC TCT ACC TTA ATA TTG ACA GGA AAC CCC ATC CAG AGT TTA      451
Ser His Leu Ser Thr Leu Ile Leu Thr Gly Asn Pro Ile Gln Ser Leu
                80                85                90

GCC CTG GGA GCC TTT TCT GGA CTA TCA AGT TTA CAG AAG CTG GTG GCT      499
Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser Leu Gln Lys Leu Val Ala
                95                100                105

```

GTG GAG ACA AAT CTA GCA TCT CTA GAG AAC TTC CCC ATT GGA CAT CTC Val Glu Thr Asn Leu Ala Ser Leu Glu Asn Phe Pro Ile Gly His Leu 110 115 120 125	547
AAA ACT TTG AAA GAA CTT AAT GTG GCT CAC AAT CTT ATC CAA TCT TTC Lys Thr Leu Lys Glu Leu Asn Val Ala His Asn Leu Ile Gln Ser Phe 130 135 140	595
AAA TTA CCT GAG TAT TTT TCT AAT CTG ACC AAT CTA GAG CAC TTG GAC Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr Asn Leu Glu His Leu Asp 145 150 155	643
CTT TCC AGC AAC AAG ATT CAA AGT ATT TAT TGC ACA GAC TTG CGG GTT Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr Cys Thr Asp Leu Arg Val 160 165 170	691
CTA CAT CAA ATG CCC CTA CTC AAT CTC TCT TTA GAC CTG TCC CTG AAC Leu His Gln Met Pro Leu Leu Asn Leu Ser Leu Asp Leu Ser Leu Asn 175 180 185	739
CCT ATG AAC TTT ATC CAA CCA GGT GCA TTT AAA GAA ATT AGG CTT CAT Pro Met Asn Phe Ile Gln Pro Gly Ala Phe Lys Glu Ile Arg Leu His 190 195 200 205	787
AAG CTG ACT TTA AGA AAT AAT TTT GAT AGT TTA AAT GTA ATG AAA ACT Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser Leu Asn Val Met Lys Thr 210 215 220	835
TGT ATT CAA GGT CTG GCT GGT TTA GAA GTC CAT CGT TTG GTT CTG GGA Cys Ile Gln Gly Leu Ala Gly Leu Glu Val His Arg Leu Val Leu Gly 225 230 235	883
GAA TTT AGA AAT GAA GGA AAC TTG GAA AAG TTT GAC AAA TCT GCT CTA Glu Phe Arg Asn Glu Gly Asn Leu Glu Lys Phe Asp Lys Ser Ala Leu 240 245 250	931
GAG GGC CTG TGC AAT TTG ACC ATT GAA GAA TTC CGA TTA GCA TAC TTA Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Ala Tyr Leu 255 260 265	979
GAC TAC TAC CTC GAT GAT ATT ATT GAC TTA TTT AAT TGT TTG ACA AAT Asp Tyr Tyr Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Thr Asn 270 275 280 285	1027
GTT TCT TCA TTT TCC CTG GTG AGT GTG ACT ATT GAA AGG GTA AAA GAC Val Ser Ser Phe Ser Leu Val Ser Val Thr Ile Glu Arg Val Lys Asp 290 295 300	1075

TTT TCT TAT AAT TTC GGA TGG CAA CAT TTA GAA TTA GTT AAC TGT AAA Phe Ser Tyr Asn Phe Gly Trp Gln His Leu Glu Leu Val Asn Cys Lys 305 310 315	1123
TTT GGA CAG TTT CCC ACA TTG AAA CTC AAA TCT CTC AAA AGG CTT ACT Phe Gly Gln Phe Pro Thr Leu Lys Leu Lys Ser Leu Lys Arg Leu Thr 320 325 330	1171
TTC ACT TCC AAC AAA GGT GGG AAT GCT TTT TCA GAA GTT GAT CTA CCA Phe Thr Ser Asn Lys Gly Gly Asn Ala Phe Ser Glu Val Asp Leu Pro 335 340 345	1219
AGC CTT GAG TTT CTA GAT CTC AGT AGA AAT GGC TTG AGT TTC AAA GGT Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn Gly Leu Ser Phe Lys Gly 350 355 360 365	1267
TGC TGT TCT CAA AGT GAT TTT GGG ACA ACC AGC CTA AAG TAT TTA GAT Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr Ser Leu Lys Tyr Leu Asp 370 375 380	1315
CTG AGC TTC AAT GGT GTT ATT ACC ATG AGT TCA AAC TTC TTG GGC TTA Leu Ser Phe Asn Gly Val Ile Thr Met Ser Ser Asn Phe Leu Gly Leu 385 390 395	1363
GAA CAA CTA GAA CAT CTG GAT TTC CAG CAT TCC AAT TTG AAA CAA ATG Glu Gln Leu Glu His Leu Asp Phe Gln His Ser Asn Leu Lys Gln Met 400 405 410	1411
AGT GAG TTT TCA GTA TTC CTA TCA CTC AGA AAC CTC ATT TAC CTT GAC Ser Glu Phe Ser Val Phe Leu Ser Leu Arg Asn Leu Ile Tyr Leu Asp 415 420 425	1459
ATT TCT CAT ACT CAC ACC AGA GTT GCT TTC AAT GGC ATC TTC AAT GGC Ile Ser His Thr His Thr Arg Val Ala Phe Asn Gly Ile Phe Asn Gly 430 435 440 445	1507
TTG TCC AGT CTC GAA GTC TTG AAA ATG GCT GGC AAT TCT TTC CAG GAA Leu Ser Ser Leu Glu Val Leu Lys Met Ala Gly Asn Ser Phe Gln Glu 450 455 460	1555
AAC TTC CTT CCA GAT ATC TTC ACA GAG CTG AGA AAC TTG ACC TTC CTG Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu Arg Asn Leu Thr Phe Leu 465 470 475	1603
GAC CTC TCT CAG TGT CAA CTG GAG CAG TTG TCT CCA ACA GCA TTT AAC Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu Ser Pro Thr Ala Phe Asn 480 485 490	1651

TCA CTC TCC AGT CTT CAG GTA CTA AAT ATG AGC CAC AAC AAC TTC TTT Ser Leu Ser Ser Leu Gln Val Leu Asn Met Ser His Asn Asn Phe Phe 495 500 505	1699
TCA TTG GAT ACG TTT CCT TAT AAG TGT CTG AAC TCC CTC CAG GTT CTT Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu Asn Ser Leu Gln Val Leu 510 515 520 525	1747
GAT TAC AGT CTC AAT CAC ATA ATG ACT TCC AAA AAA CAG GAA CTA CAG Asp Tyr Ser Leu Asn His Ile Met Thr Ser Lys Lys Gln Glu Leu Gln 530 535 540	1795
CAT TTT CCA AGT AGT CTA GCT TTC TTA AAT CTT ACT CAG AAT GAC TTT His Phe Pro Ser Ser Leu Ala Phe Leu Asn Leu Thr Gln Asn Asp Phe 545 550 555	1843
GCT TGT ACT TGT GAA CAC CAG AGT TTC CTG CAA TGG ATC AAG GAC CAG Ala Cys Thr Cys Glu His Gln Ser Phe Leu Gln Trp Ile Lys Asp Gln 560 565 570	1891
AGG CAG CTC TTG GTG GAA GTT GAA CGA ATG GAA TGT GCA ACA CCT TCA Arg Gln Leu Leu Val Glu Val Glu Arg Met Glu Cys Ala Thr Pro Ser 575 580 585	1939
GAT AAG CAG GGC ATG CCT GTG CTG AGT TTG AAT ATC ACC TGT CAG ATG Asp Lys Gln Gly Met Pro Val Leu Ser Leu Asn Ile Thr Cys Gln Met 590 595 600 605	1987
AAT AAG ACC ATC ATT GGT GTG TCG GTC CTC AGT GTG CTT GTA GTA TCT Asn Lys Thr Ile Ile Gly Val Ser Val Leu Ser Val Leu Val Val Ser 610 615 620	2035
GTT GTA GCA GTT CTG GTC TAT AAG TTC TAT TTT CAC CTG ATG CTT CTT Val Val Ala Val Leu Val Tyr Lys Phe Tyr Phe His Leu Met Leu Leu 625 630 635	2083
GCT GGC TGC ATA AAG TAT GGT AGA GGT GAA AAC ATC TAT GAT GCC TTT Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu Asn Ile Tyr Asp Ala Phe 640 645 650	2131
GTT ATC TAC TCA AGC CAG GAT GAG GAC TGG GTA AGG AAT GAG CTA GTA Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp Val Arg Asn Glu Leu Val 655 660 665	2179
AAG AAT TTA GAA GAA GGG GTG CCT CCA TTT CAG CTC TGC CTT CAC TAC Lys Asn Leu Glu Glu Gly Val Pro Pro Phe Gln Leu Cys Leu His Tyr 670 675 680 685	2227

AGA GAC TTT ATT CCC GGT GTG GCC ATT GCT GCC AAC ATC ATC CAT GAA Arg Asp Phe Ile Pro Gly Val Ala Ile Ala Ala Asn Ile Ile His Glu 690 695 700	2275
GGT TTC CAT AAA AGC CGA AAG GTG ATT GTT GTG GTG TCC CAG CAC TTC Gly Phe His Lys Ser Arg Lys Val Ile Val Val Val Ser Gln His Phe 705 710 715	2323
ATC CAG AGC CGC TGG TGT ATC TTT GAA TAT GAG ATT GCT CAG ACC TGG Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr Glu Ile Ala Gln Thr Trp 720 725 730	2371
CAG TTT CTG AGC AGT CGT GCT GGT ATC ATC TTC ATT GTC CTG CAG AAG Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile Phe Ile Val Leu Gln Lys 735 740 745	2419
GTG GAG AAG ACC CTG CTC AGG CAG CAG GTG GAG CTG TAC CGC CTT CTC Val Glu Lys Thr Leu Leu Arg Gln Gln Val Glu Leu Tyr Arg Leu Leu 750 755 760 765	2467
AGC AGG AAC ACT TAC CTG GAG TGG GAG GAC AGT GTC CTG GGG CGG CAC Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp Ser Val Leu Gly Arg His 770 775 780	2515
ATC TTC TGG AGA CGA CTC AGA AAA GCC CTG CTG GAT GGT AAA TCA TGG Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu Leu Asp Gly Lys Ser Trp 785 790 795	2563
AAT CCA GAA GGA ACA GTG GGT ACA GGA TGC AAT TGG CAG GAA GCA ACA Asn Pro Glu Gly Thr Val Gly Thr Gly Cys Asn Trp Gln Glu Ala Thr 800 805 810	2611
TCT ATC TGAAGAGGAA AAATAAAAC CTCCTGAGGC ATTTCTTGCC CAGCTGGGTC Ser Ile 815	2667
CAACACTTGT TCAGTTAATA AGTATTAAAT GCTGCCACAT GTCAGGCCTT ATGCTAAGGG	2727
TGAGTAATTC CATGGTGCAC TAGATATGCA GGGCTGCTAA TCTCAAGGAG CTTCCAGTGC	2787
AGAGGGAATA AATGCTAGAC TAAAATACAG AGTCTTCCAG GTGGGCATTT CAACCAACTC	2847
AGTCAAGGAA CCCATGACAA AGAAAGTCAT TTCAACTCTT ACCTCATCAA GTTGAATAAA	2907
GACAGAGAAA ACAGAAAGAG ACATTGTTCT TTTCCTGAGT CTTTGAATG GAAATTGTAT	2967
TATGTTATAG CCATCATAAA ACCATTTTGG TAGTTTTGAC TGAAGTGGGT GTTCACTTTT	3027

TCCTTTTGA TTGAATACAA TTAAATTCT ACTTGATGAC TGCAGTCGTC AAGGGGCTCC 3087
TGATGCAAGA TCCCCCTTCC ATTTTAAGTC TGTCTCCTTA CAGAGGTAA AGTCTAATGG 3147
CTAATTCCTA AGGAAACCTG ATTAACACAT GCTCACAACC ATCCTGGTCA TTCTCGAACA 3207
TGTTCTATTT TTAACTAAT CACCCTGAT ATATTTTAT TTTTATATAT CCAGTTTCA 3267
TTTTTTTAGG TCTTGCCTAT AAGCTAATAT CATAAATAAG GTTGTTAAG ACGTGCTTCA 3327
AATATCCATA TTAACCACTA TTTTCAAGG AAGTATGGAA AAGTACACTC TGTCACCTTG 3387
TCACTCGATG TCATTCCAAA GTTATTGCCT ACTAAGTAAT GACTGTCATG AAAGCAGCAT 3447
TGAAATAATT TGTTTAAAGG GGGCACTCTT TTAAACGGGA AGAAAATTTC CGCTTCCTGG 3507
TCTTATCATG GACAATTTGG GCTAGAGGCA GGAAGGAAGT GGGATGACCT CAGGAAGTCA 3567
CCTTTTCTTG ATTCCAGAAA CATATGGGCT GATAAACCCG GGGTGACCTC ATGAAATGAG 3627
TTGCAGCAGA AGTTTATTTT TTTCAGAACA AGTGATGTTT GATGGACCTC TGAATCTCTT 3687
TAGGGAGACA CAGATGGCTG GGATCCCTCC CCTGTACCCT TCTCACTGCC AGGAGAACTA 3747
CGTGTGAAGG TATTCAAGGC AGGGAGTATA CATTGCTGTT TCCTGTTGGG CAATGCTCCT 3807
TGACCACATT TTGGGAAGAG TGGATGTTAT CATTGAGAAA ACAATGTGTC TGGAATTAAT 3867
GGGGTTCTTA TAAAGAAGGT TCCCAGAAAA GAATGTTCAT TCCAGCTTCT TCAGGAAACA 3927
GGAACATTCA AGGAAAAGGA CAATCAGGAT GTCATCAGG AAATGAAAAT AAAAACCACA 3987
ATGAGATATC ACCTTATACC AGGTAGATGG CTAATAAAA AAAATGAAGT GTCATCAAGG 4047
ATATAGAGAA ATTGGAACCC TTCTTCACTG CTGGAGGGAA TGAAAAATGG TGAGCCGTT 4107
ATGAAAAACA GTACGGAGGT TTCTCAAAAA TAAAAATAG AACTGCTATA TGATCCAGCA 4167
ATCTCACTTC TGTATATATA CCCAAAATAA TTGAAATCAG AATTTCAAGA AAATATTTAC 4227
ACTCCCATGT TCATTGTGGC ACTCTTCACA ATCACTGTTT CCAAAGTTAT GGAAACAACC 4287
CAAATTTCCA TTGGAAAATA AATGGACAAA GGAAATGTGC ATATAACGTA CAATGGGGAT 4347
ATTATTCAGC CTAAAAAAAG GGGGGATCCT GTTATTTATG ACAACATGAA TAAACCCGGA 4407
GGCCATTATG CTATGTAAAA TGAGCAAGTA ACAGAAAGAC AAATACTGCC TGATTTCAAT 4467

TATATGAGGT TCTAAAATAG TCAAATCAT AGAAGCAGAG AATAGAACAG TGGTTCCTAG 4527
 GGAAAAGGAG GAAGGGAGAA ATGAGGAAAT AGGGAGTTGT CTAATTGGTA TAAAATTATA 4587
 GTATGCAAGA TGAATTAGCT CTAAAGATCA GCTGTATAGC AGAGTTCGTA TAATGAACAA 4647
 TACTGTATTA TGCACTTAAC ATTTTGTAA GAGGGTACCT CTCATGTAA GTGTTCTTAC 4707
 CATATACATA TACACAAGGA AGCTTTTGA GGTGATGGAT ATATTATTA CCTTGATTGT 4767
 GGTGATGGTT TGACAGGTAT GTGACTATGT TCAAATCAT CAAATTGTAT ACATTAAATA 4827
 TATGCAGTTT TATAATATCA AAAAAAAAAA AAAAAAAAAA 4865

(2) SEQ ID NO: 26 的信息:

(i) 顺序特征:

(A) 长度: 837 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 顺序描述: SEQ ID NO: 26:

Met Ser Ala Ser Arg Leu Ala Gly Thr Leu Ile Pro Ala Met Ala Phe
 -22 -20 -15 -10

Leu Ser Cys Val Arg Pro Glu Ser Trp Glu Pro Cys Val Glu Val Pro
 -5 1 5 10

Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys Ile Pro Asp
 15 20 25

Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe Asn Pro Leu
 30 35 40

Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu Leu Gln Val
 45 50 55

Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp Gly Ala Tyr
 60 65 70

Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly Asn Pro Ile
 75 80 85 90

Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser Leu Gln Lys
 95 100 105

Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn Phe Pro Ile
110 115 120

Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His Asn Leu Ile
125 130 135

Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr Asn Leu Glu
140 145 150

His Leu Asp Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr Cys Thr Asp
155 160 165 170

Leu Arg Val Leu His Gln Met Pro Leu Leu Asn Leu Ser Leu Asp Leu
175 180 185

Ser Leu Asn Pro Met Asn Phe Ile Gln Pro Gly Ala Phe Lys Glu Ile
190 195 200

Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser Leu Asn Val
205 210 215

Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val His Arg Leu
220 225 230

Val Leu Gly Glu Phe Arg Asn Glu Gly Asn Leu Glu Lys Phe Asp Lys
235 240 245 250

Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu
255 260 265

Ala Tyr Leu Asp Tyr Tyr Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys
270 275 280

Leu Thr Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr Ile Glu Arg
285 290 295

Val Lys Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu Glu Leu Val
300 305 310

Asn Cys Lys Phe Gly Gln Phe Pro Thr Leu Lys Leu Lys Ser Leu Lys
315 320 325 330

Arg Leu Thr Phe Thr Ser Asn Lys Gly Gly Asn Ala Phe Ser Glu Val
335 340 345

Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn Gly Leu Ser
350 355 360

Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr Ser Leu Lys
365 370 375

Tyr Leu Asp Leu Ser Phe Asn Gly Val Ile Thr Met Ser Ser Asn Phe
380 385 390

Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His Ser Asn Leu
395 400 405 410

Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg Asn Leu Ile
415 420 425

Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe Asn Gly Ile
430 435 440

Phe Asn Gly Leu Ser Ser Leu Glu Val Leu Lys Met Ala Gly Asn Ser
445 450 455

Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu Arg Asn Leu
460 465 470

Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu Ser Pro Thr
475 480 485 490

Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met Ser His Asn
495 500 505

Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu Asn Ser Leu
510 515 520

Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser Lys Lys Gln
525 530 535

Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn Leu Thr Gln
540 545 550

Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu Gln Trp Ile
555 560 565 570

Lys Asp Gln Arg Gln Leu Leu Val Glu Val Glu Arg Met Glu Cys Ala
575 580 585

Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu Asn Ile Thr
590 595 600

Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Leu Ser Val Leu
605 610 615

```

Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr Phe His Leu
  620                625                630

Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu Asn Ile Tyr
  635                640                645                650

Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp Val Arg Asn
      655                660                665

Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe Gln Leu Cys
      670                675                680

Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala Ala Asn Ile
      685                690                695

Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val Val Val Ser
      700                705                710

Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr Glu Ile Ala
      715                720                725                730

Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile Phe Ile Val
      735                740                745

Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val Glu Leu Tyr
      750                755                760

Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp Ser Val Leu
      765                770                775

Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu Leu Asp Gly
      780                785                790

Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys Asn Trp Gln
      795                800                805                810

Glu Ala Thr Ser Ile
      815

```

(2) SEQ ID NO: 27 的信息:

- (i) 顺序特征:
 - (A) 长度: 300 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线性

(ii) 分子类型: cDNA

(ix) 特征:

(A) 名称/关键词: CDS

(B) 位置: 1...300

(ix) 特征:

(A) 名称/关键词: misc 特征

(B) 位置: 186

(D) 其它信息:/注释 = "核苷酸 186、196、217、276 和 300 标示为 C, 每个可以为 A、C、G 或 T"

(xi) 顺序描述: SEQ ID NO: 27:

TCC TAT TCT ATG GAA AAA GAT GCT TTC CTA TTT ATG AGA AAT TTG AAG	48
Ser Tyr Ser Met Glu Lys Asp Ala Phe Leu Phe Met Arg Asn Leu Lys	
1 5 10 15	
GTT CTC TCA CTA AAA GAT AAC AAT GTC ACA GCT GTC CCC ACC ACT TTG	96
Val Leu Ser Leu Lys Asp Asn Asn Val Thr Ala Val Pro Thr Thr Leu	
20 25 30	
CCA CCT AAT TTA CTA GAG CTC TAT CTT TAT AAC AAT ATC ATT AAG AAA	144
Pro Pro Asn Leu Leu Glu Leu Tyr Leu Tyr Asn Asn Ile Ile Lys Lys	
35 40 45	
ATC CAA GAA AAT GAT TTC AAT AAC CTC AAT GAG TTG CAA GTC CTT GAC	192
Ile Gln Glu Asn Asp Phe Asn Asn Leu Asn Glu Leu Gln Val Leu Asp	
50 55 60	
CTA CGT GGA AAT TGC CCT CGA TGT CAT AAT GTC CCA TAT CCG TGT ACA	240
Leu Arg Gly Asn Cys Pro Arg Cys His Asn Val Pro Tyr Pro Cys Thr	
65 70 75 80	
CCG TGT GAA AAT AAT TCC CCC TTA CAG ATC CAT GAC AAT GCT TTC AAT	288
Pro Cys Glu Asn Asn Ser Pro Leu Gln Ile His Asp Asn Ala Phe Asn	
85 90 95	
TCA TCG ACA GAC	300
Ser Ser Thr Asp	
100	

(2) SEQ ID NO: 28 的信息:

(i) 顺序特征:

(A) 长度: 100 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii) 分子类型：蛋白质

(xi) 顺序描述：SEQ ID NO: 28:

```

Ser Tyr Ser Met Glu Lys Asp Ala Phe Leu Phe Met Arg Asn Leu Lys
 1              5              10              15

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

Pro Pro Asn Leu Leu Glu Leu Tyr Leu Tyr Asn Asn Ile Ile Lys Lys
      35              40              45

Ile Gln Glu Asn Asp Phe Asn Asn Leu Asn Glu Leu Gln Val Leu Asp
      50              55              60

Leu Arg Gly Asn Cys Pro Arg Cys His Asn Val Pro Tyr Pro Cys Thr
      65              70              75              80

Pro Cys Glu Asn Asn Ser Pro Leu Gln Ile His Asp Asn Ala Phe Asn
      85              90              95

Ser Ser Thr Asp
      100

```

(2) SEQ ID NO: 29 的信息:

(i) 顺序特征:

- (A) 长度：1756 个碱基对
- (B) 类型：核酸
- (C) 链型：单链
- (D) 拓扑学：线性

(ii) 分子类型：cDNA

(ix) 特征:

- (A) 名称/关键词：CDS
- (B) 位置：1...1182

(ix) 特征:

- (A) 名称/关键词：misc 特征
- (B) 位置：1643
- (D) 其它信息：/注释 = “核苷酸 1643 标示为 A，可以为 A 或 G”

(ix) 特征:

- (A) 名称/关键词：misc 特征
- (B) 位置：1664
- (D) 其它信息：/注释 = “核苷酸 1664 标示为 C，可以为 A、C、G 或 T”

GAA AAT GTC CTC AAC AAT CTG GAG ATG TTG GTT TTA CAT CAC AAT CGC Glu Asn Val Leu Asn Asn Leu Glu Met Leu Val Leu His His Asn Arg 115 120 125	384
TTT CTT TGC AAC TGT GAT GCT GTG TGG TTT GTC TGG TGG GTT AAC CAT Phe Leu Cys Asn Cys Asp Ala Val Trp Phe Val Trp Trp Val Asn His 130 135 140	432
ACA GAT GTT ACT ATT CCA TAC CTG GCC ACT GAT GTG ACT TGT GTA GGT Thr Asp Val Thr Ile Pro Tyr Leu Ala Thr Asp Val Thr Cys Val Gly 145 150 155 160	480
CCA GGA GCA CAC AAA GGT CAA AGT GTC ATA TCC CTT GAT CTG TAT ACG Pro Gly Ala His Lys Gly Gln Ser Val Ile Ser Leu Asp Leu Tyr Thr 165 170 175	528
TGT GAG TTA GAT CTC ACA AAC CTG ATT CTG TTC TCA GTT TCC ATA TCA Cys Glu Leu Asp Leu Thr Asn Leu Ile Leu Phe Ser Val Ser Ile Ser 180 185 190	576
TCA GTC CTC TTT CTT ATG GTA GTT ATG ACA ACA AGT CAC CTC TTT TTC Ser Val Leu Phe Leu Met Val Val Met Thr Thr Ser His Leu Phe Phe 195 200 205	624
TGG GAT ATG TGG TAC ATT TAT TAT TTT TGG AAA GCA AAG ATA AAG GGG Trp Asp Met Trp Tyr Ile Tyr Tyr Phe Trp Lys Ala Lys Ile Lys Gly 210 215 220	672
TAT CCA GCA TCT GCA ATC CCA TGG AGT CCT TGT TAT GAT GCT TTT ATT Tyr Pro Ala Ser Ala Ile Pro Trp Ser Pro Cys Tyr Asp Ala Phe Ile 225 230 235 240	720
GTG TAT GAC ACT AAA AAC TCA GCT GTG ACA GAA TGG GTT TTG CAG GAG Val Tyr Asp Thr Lys Asn Ser Ala Val Thr Glu Trp Val Leu Gln Glu 245 250 255	768
CTG GTG GCA AAA TTG GAA GAT CCA AGA GAA AAA CAC TTC AAT TTG TGT Leu Val Ala Lys Leu Glu Asp Pro Arg Glu Lys His Phe Asn Leu Cys 260 265 270	816
CTA GAA GAA AGA GAC TGG CTA CCA GGA CAG CCA GTT CTA GAA AAC CTT Leu Glu Glu Arg Asp Trp Leu Pro Gly Gln Pro Val Leu Glu Asn Leu 275 280 285	864
TCC CAG AGC ATA CAG CTC AGC AAA AAG ACA GTG TTT GTG ATG ACA CAG Ser Gln Ser Ile Gln Leu Ser Lys Lys Thr Val Phe Val Met Thr Gln 290 295 300	912

AAA TAT GCT AAG ACT GAG AGT TTT AAG ATG GCA TTT TAT TTG TCT CAT Lys Tyr Ala Lys Thr Glu Ser Phe Lys Met Ala Phe Tyr Leu Ser His 305 310 315 320	960
CAG AGG CTC CTG GAT GAA AAA GTG GAT GTG ATT ATC TTG ATA TTC TTG Gln Arg Leu Leu Asp Glu Lys Val Asp Val Ile Ile Leu Ile Phe Leu 325 330 335	1008
GAA AGA CCT CTT CAG AAG TCT AAG TTT CTT CAG CTC AGG AAG AGA CTC Glu Arg Pro Leu Gln Lys Ser Lys Phe Leu Gln Leu Arg Lys Arg Leu 340 345 350	1056
TGC AGG AGC TCT GTC CTT GAG TGG CCT GCA AAT CCA CAG GCT CAC CCA Cys Arg Ser Ser Val Leu Glu Trp Pro Ala Asn Pro Gln Ala His Pro 355 360 365	1104
TAC TTC TGG CAG TGC CTG AAA AAT GCC CTG ACC ACA GAC AAT CAT GTG Tyr Phe Trp Gln Cys Leu Lys Asn Ala Leu Thr Thr Asp Asn His Val 370 375 380	1152
GCT TAT AGT CAA ATG TTC AAG GAA ACA GTC TAGCTCTCTG AAGAATGTCA Ala Tyr Ser Gln Met Phe Lys Glu Thr Val 385 390	1202
CCACCTAGGA CATGCCTTGG TACCTGAAGT TTTCATAAAG GTTCCATAA ATGAAGGTCT	1262
GAATTTTTC TAACAGTTGT CATGGCTCAG ATTGGTGGGA AATCATCAAT ATATGGCTAA	1322
GAAATTAAGA AGGGGAGACT GATAGAAGAT AATTTCTTTC TTCATGTGCC ATGCTCAGTT	1382
AAATATTTCC CCTAGCTCAA ATCTGAAAAA CTGTGCCTAG GAGACAACAC AAGGCTTTGA	1442
TTTATCTGCA TACAATTGAT AAGAGCCACA CATCTGCCCT GAAGAAGTAC TAGTAGTTTT	1502
AGTAGTAGGG TAAAAATTAC ACAAGCTTTC TCTCTCTCTG ATACTGAACT GTACCAGAGT	1562
TCAATGAAAT AAAAGCCCAG AGAACTTCTC AGTAAATGGT TTCATTATCA TGTAGTATCC	1622
ACCATGCAAT ATGCCACAAA ACCGCTACTG GTACAGGACA GCTGGTAGCT GCTTCAAGGC	1682
CTCTTATCAT TTTCTTGGGG CCCATGGAGG GGTCTCTCTG GAAAAAGGGA AGGTTTTTTT	1742
TGGCCATCCA TGAA	1756

(2) SEQ ID NO: 30 的信息:

(i) 顺序特征:

(A) 长度: 394 个氨基酸

(B) 类型：氨基酸

(D) 拓扑学：线性

(ii) 分子类型：蛋白质

(xi) 顺序描述：SEQ ID NO: 30:

Ser Pro Glu Ile Pro Trp Asn Ser Leu Pro Pro Glu Val Phe Glu Gly
 1 5 10 15

Met Pro Pro Asn Leu Lys Asn Leu Ser Leu Ala Lys Asn Gly Leu Lys
 20 25 30

Ser Phe Phe Trp Asp Arg Leu Gln Leu Leu Lys His Leu Glu Ile Leu
 35 40 45

Asp Leu Ser His Asn Gln Leu Thr Lys Val Pro Glu Arg Leu Ala Asn
 50 55 60

Cys Ser Lys Ser Leu Thr Thr Leu Ile Leu Lys His Asn Gln Ile Arg
 65 70 75 80

Gln Leu Thr Lys Tyr Phe Leu Glu Asp Ala Leu Gln Leu Arg Tyr Leu
 85 90 95

Asp Ile Ser Ser Asn Lys Ile Gln Val Ile Gln Lys Thr Ser Phe Pro
 100 105 110

Glu Asn Val Leu Asn Asn Leu Glu Met Leu Val Leu His His Asn Arg
 115 120 125

Phe Leu Cys Asn Cys Asp Ala Val Trp Phe Val Trp Trp Val Asn His
 130 135 140

Thr Asp Val Thr Ile Pro Tyr Leu Ala Thr Asp Val Thr Cys Val Gly
 145 150 155 160

Pro Gly Ala His Lys Gly Gln Ser Val Ile Ser Leu Asp Leu Tyr Thr
 165 170 175

Cys Glu Leu Asp Leu Thr Asn Leu Ile Leu Phe Ser Val Ser Ile Ser
 180 185 190

Ser Val Leu Phe Leu Met Val Val Met Thr Thr Ser His Leu Phe Phe
 195 200 205

Trp Asp Met Trp Tyr Ile Tyr Tyr Phe Trp Lys Ala Lys Ile Lys Gly
 210 215 220

Tyr Pro Ala Ser Ala Ile Pro Trp Ser Pro Cys Tyr Asp Ala Phe Ile
225 230 235 240

Val Tyr Asp Thr Lys Asn Ser Ala Val Thr Glu Trp Val Leu Gln Glu
245 250 255

Leu Val Ala Lys Leu Glu Asp Pro Arg Glu Lys His Phe Asn Leu Cys
260 265 270

Leu Glu Glu Arg Asp Trp Leu Pro Gly Gln Pro Val Leu Glu Asn Leu
275 280 285

Ser Gln Ser Ile Gln Leu Ser Lys Lys Thr Val Phe Val Met Thr Gln
290 295 300

Lys Tyr Ala Lys Thr Glu Ser Phe Lys Met Ala Phe Tyr Leu Ser His
305 310 315 320

Gln Arg Leu Leu Asp Glu Lys Val Asp Val Ile Ile Leu Ile Phe Leu
325 330 335

Glu Arg Pro Leu Gln Lys Ser Lys Phe Leu Gln Leu Arg Lys Arg Leu
340 345 350

Cys Arg Ser Ser Val Leu Glu Trp Pro Ala Asn Pro Gln Ala His Pro
355 360 365

Tyr Phe Trp Gln Cys Leu Lys Asn Ala Leu Thr Thr Asp Asn His Val
370 375 380

Ala Tyr Ser Gln Met Phe Lys Glu Thr Val
385 390

(2) SEQ ID NO: 31 的信息:

- (i) 顺序特征:
 - (A) 长度: 999 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线性
- (ii) 分子类型: cDNA
- (ix) 特征:
 - (A) 名称/关键词: CDS
 - (B) 位置: 2...847

(ix) 特征:

(A) 名称/关键词: misc 特征

(B) 位置: 4

(D) 其它信息: /注释 = "核苷酸 4 和 23 标示为 C, 每个可以为 A、C、G"

或 T "

(ix) 特征:

(A) 名称/关键词: misc 特征

(B) 位置: 650

(D) 其它信息: /注释 = “核苷酸 650 标示为 G, 可以为 A 或 G”

(ix) 特征:

(A) 名称/关键词: misc 特征

(B) 位置: 715

(D) 其它信息：/注释 = “核苷酸 715、825 和 845 标示为 C，每个可以为 C

或 T ”

(xi) 顺序描述: SEQ ID NO: 31:

C TCC GAT GCC AAG ATT CGG CAC CAG GCA TAT TCA GAG GTC ATG ATG 46
Ser Asp Ala Lys Ile Arg His Gln Ala Tyr Ser Glu Val Met Met
1 5 10 15

GTT GGA TGG TCA GAT TCA TAC ACC TGT GAA TAC CCT TTA AAC CTA AGG 94
Val Gly Trp Ser Asp Ser Tyr Thr Cys Glu Tyr Pro Leu Asn Leu Arg
20 25 30

GGA ACT AGG TTA AAA GAC GTT CAT CTC CAC GAA TTA TCT TGC AAC ACA 142
Gly Thr Arg Leu Lys Asp Val His Leu His Glu Leu Ser Cys Asn Thr
35 40 45

GCT CTG TTG ATT GTC ACC ATT GTG GTT ATT ATG CTA GTT CTG GGG TTG 190
Ala Leu Leu Ile Val Thr Ile Val Val Ile Met Leu Val Leu Gly Leu
50 55 60

GCT GTG GCC TTC TGC TGT CTC CAC TTT GAT CTG CCC TGG TAT CTC AGG 238
Ala Val Ala Phe Cys Cys Leu His Phe Asp Leu Pro Trp Tyr Leu Arg
65 70 75

ATG CTA GGT CAA TGC ACA CAA ACA TGG CAC AGG GTT AGG AAA ACA ACC 286
Met Leu Gly Gln Cys Thr Gln Thr Trp His Arg Val Arg Lys Thr Thr
80 85 90 95

CAA GAA CAA CTC AAG AGA AAT GTC CGA TTC CAC GCA TTT ATT TCA TAC 334
Gln Glu Gln Leu Lys Arg Asn Val Arg Phe His Ala Phe Ile Ser Tyr
100 105 110

AGT GAA CAT GAT TCT CTG TGG GTG AAG AAT GAA TTG ATC CCC AAT CTA Ser Glu His Asp Ser Leu Trp Val Lys Asn Glu Leu Ile Pro Asn Leu 115 120 125	382
GAG AAG GAA GAT GGT TCT ATC TTG ATT TGC CTT TAT GAA AGC TAC TTT Glu Lys Glu Asp Gly Ser Ile Leu Ile Cys Leu Tyr Glu Ser Tyr Phe 130 135 140	430
GAC CCT GGC AAA AGC ATT AGT GAA AAT ATT GTA AGC TTC ATT GAG AAA Asp Pro Gly Lys Ser Ile Ser Glu Asn Ile Val Ser Phe Ile Glu Lys 145 150 155	478
AGC TAT AAG TCC ATC TTT GTT TTG TCT CCC AAC TTT GTC CAG AAT GAG Ser Tyr Lys Ser Ile Phe Val Leu Ser Pro Asn Phe Val Gln Asn Glu 160 165 170 175	526
TGG TGC CAT TAT GAA TTC TAC TTT GCC CAC CAC AAT CTC TTC CAT GAA Trp Cys His Tyr Glu Phe Tyr Phe Ala His His Asn Leu Phe His Glu 180 185 190	574
AAT TCT GAT CAC ATA ATT CTT ATC TTA CTG GAA CCC ATT CCA TTC TAT Asn Ser Asp His Ile Ile Leu Ile Leu Leu Glu Pro Ile Pro Phe Tyr 195 200 205	622
TGC ATT CCC ACC AGG TAT CAT AAA CTG GAA GCT CTC CTG GAA AAA AAA Cys Ile Pro Thr Arg Tyr His Lys Leu Glu Ala Leu Leu Glu Lys Lys 210 215 220	670
GCA TAC TTG GAA TGG CCC AAG GAT AGG CGT AAA TGT GGG CTT TTC TGG Ala Tyr Leu Glu Trp Pro Lys Asp Arg Arg Lys Cys Gly Leu Phe Trp 225 230 235	718
GCA AAC CTT CGA GCT GCT GTT AAT GTT AAT GTA TTA GCC ACC AGA GAA Ala Asn Leu Arg Ala Ala Val Asn Val Asn Val Leu Ala Thr Arg Glu 240 245 250 255	766
ATG TAT GAA CTG CAG ACA TTC ACA GAG TTA AAT GAA GAG TCT CGA GGT Met Tyr Glu Leu Gln Thr Phe Thr Glu Leu Asn Glu Glu Ser Arg Gly 260 265 270	814
TCT ACA ATC TCT CTG ATG ACA ACA GAC TGT CTA TAAATCCCA CAGTCCTTGG Ser Thr Ile Ser Leu Met Arg Thr Asp Cys Leu 275 280	867
GAAGTTGGGG ACCACATACA CTGTTGGGAT GTACATTGAT ACAACCTTTA TGATGGCAAT	927
TTGACAATAT TTATTAAAT AAAAAATGGT TATTCCTTC AAAAAAAAAA AAAAAAAAAA	987

AAAAAAAAAA AA

999

(2) SEQ ID NO: 32 的信息:

(i) 顺序特征:

(A) 长度: 282 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 顺序描述: SEQ ID NO: 32:

Ser Asp Ala Lys Ile Arg His Gln Ala Tyr Ser Glu Val Met Met Val
 1 5 10 15

Gly Trp Ser Asp Ser Tyr Thr Cys Glu Tyr Pro Leu Asn Leu Arg Gly
 20 25 30

Thr Arg Leu Lys Asp Val His Leu His Glu Leu Ser Cys Asn Thr Ala
 35 40 45

Leu Leu Ile Val Thr Ile Val Val Ile Met Leu Val Leu Gly Leu Ala
 50 55 60

Val Ala Phe Cys Cys Leu His Phe Asp Leu Pro Trp Tyr Leu Arg Met
 65 70 75 80

Leu Gly Gln Cys Thr Gln Thr Trp His Arg Val Arg Lys Thr Thr Gln
 85 90 95

Glu Gln Leu Lys Arg Asn Val Arg Phe His Ala Phe Ile Ser Tyr Ser
 100 105 110

Glu His Asp Ser Leu Trp Val Lys Asn Glu Leu Ile Pro Asn Leu Glu
 115 120 125

Lys Glu Asp Gly Ser Ile Leu Ile Cys Leu Tyr Glu Ser Tyr Phe Asp
 130 135 140

Pro Gly Lys Ser Ile Ser Glu Asn Ile Val Ser Phe Ile Glu Lys Ser
 145 150 155 160

Tyr Lys Ser Ile Phe Val Leu Ser Pro Asn Phe Val Gln Asn Glu Trp
 165 170 175

Cys His Tyr Glu Phe Tyr Phe Ala His His Asn Leu Phe His Glu Asn
 180 185 190

Ser Asp His Ile Ile Leu Ile Leu Leu Glu Pro Ile Pro Phe Tyr Cys
 195 200 205
 Ile Pro Thr Arg Tyr His Lys Leu Glu Ala Leu Leu Glu Lys Lys Ala
 210 215 220
 Tyr Leu Glu Trp Pro Lys Asp Arg Arg Lys Cys Gly Leu Phe Trp Ala
 225 230 235 240
 Asn Leu Arg Ala Ala Val Asn Val Asn Val Leu Ala Thr Arg Glu Met
 245 250 255
 Tyr Glu Leu Gln Thr Phe Thr Glu Leu Asn Glu Glu Ser Arg Gly Ser
 260 265 270
 Thr Ile Ser Leu Met Arg Thr Asp Cys Leu
 275 280

(2) SEQ ID NO: 33 的信息:

- (i) 顺序特征:
 - (A) 长度: 1173 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线性
- (ii) 分子类型: cDNA
- (ix) 特征:
 - (A) 名称/关键词: CDS
 - (B) 位置: 1...1008
- (ix) 特征:
 - (A) 名称/关键词: misc 特征
 - (B) 位置: 854
 - (D) 其它信息: /注释 = "核苷酸 854 标示为 A, 可以为 A 或 T"
- (ix) 特征:
 - (A) 名称/关键词: misc 特征
 - (B) 位置: 1171
 - (D) 其它信息: /注释 = "核苷酸 1171 和 1172 标示为 C, 每个可以为 A、C、G 或 T"
- (xi) 顺序描述: SEQ ID NO: 33:

CTG CCT GCT GGC ACC CGG CTC CGG AGG CTG GAT GTC AGC TGC AAC AGC
 Leu Pro Ala Gly Thr Arg Leu Arg Arg Leu Asp Val Ser Cys Asn Ser

48

1	5	10	15	
ATC AGC TTC GTG GCC CCC GGC TTC TTT TCC AAG GCC AAG GAG CTG CGA				96
Ile Ser Phe Val Ala Pro Gly Phe Phe Ser Lys Ala Lys Glu Leu Arg				
20	25	30		
GAG CTC AAC CTT AGC GCC AAC GCC CTC AAG ACA GTG GAC CAC TCC TGG				144
Glu Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp				
35	40	45		
TTT GGG CCC CTG GCG AGT GCC CTG CAA ATA CTA GAT GTA AGC GCC AAC				192
Phe Gly Pro Leu Ala Ser Ala Leu Gln Ile Leu Asp Val Ser Ala Asn				
50	55	60		
CCT CTG CAC TGC GCC TGT GGG GCG GCC TTT ATG GAC TTC CTG CTG GAG				240
Pro Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Glu				
65	70	75	80	
GTG CAG GCT GCC GTG CCC GGT CTG CCC AGC CGG GTG AAG TGT GGC AGT				288
Val Gln Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser				
85	90	95		
CCG GGC CAG CTC CAG GGC CTC AGC ATC TTT GCA CAG GAC CTG CGC CTC				336
Pro Gly Gln Leu Gln Gly Leu Ser Ile Phe Ala Gln Asp Leu Arg Leu				
100	105	110		
TGC CTG GAT GAG GCC CTC TCC TGG GAC TGT TTC GCC CTC TCG CTG CTG				384
Cys Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu				
115	120	125		
GCT GTG GCT CTG GGC CTG GGT GTG CCC ATG CTG CAT CAC CTC TGT GGC				432
Ala Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly				
130	135	140		
TGG GAC CTC TGG TAC TGC TTC CAC CTG TGC CTG GCC TGG CTT CCC TGG				480
Trp Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp				
145	150	155	160	
CGG GGG CGG CAA AGT GGG CGA GAT GAG GAT GCC CTG CCC TAC GAT GCC				528
Arg Gly Arg Gln Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala				
165	170	175		
TTC GTG GTC TTC GAC AAA ACG CAG AGC GCA GTG GCA GAC TGG GTG TAC				576
Phe Val Val Phe Asp Lys Thr Gln Ser Ala Val Ala Asp Trp Val Tyr				
180	185	190		
AAC GAG CTT CGG GGG CAG CTG GAG GAG TGC CGT GGG CGC TGG GCA CTC				624
Asn Glu Leu Arg Gly Gln Leu Glu Glu Cys Arg Gly Arg Trp Ala Leu				

195	200	205	
CGC CTG TGC CTG GAG GAA CGC GAC TGG CTG CCT GGC AAA ACC CTC TTT			672
Arg Leu Cys Leu Glu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe			
210	215	220	
GAG AAC CTG TGG GCC TCG GTC TAT GGC AGC CGC AAG ACG CTG TTT GTG			720
Glu Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val			
225	230	235	240
CTG GCC CAC ACG GAC CGG GTC AGT GGT CTC TTG CGC GCC AGC TTC CTG			768
Leu Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu			
245	250	255	
CTG GCC CAG CAG CGC CTG CTG GAG GAC CGC AAG GAC GTC GTG GTG CTG			816
Leu Ala Gln Gln Arg Leu Leu Glu Asp Arg Lys Asp Val Val Val Leu			
260	265	270	
GTG ATC CTG AGC CCT GAC GGC CGC CGC TCC CGC TAC GAG CGG CTG CGC			864
Val Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Glu Arg Leu Arg			
275	280	285	
CAG CGC CTC TGC CGC CAG AGT GTC CTC CTC TGG CCC CAC CAG CCC AGT			912
Gln Arg Leu Cys Arg Gln Ser Val Leu Leu Trp Pro His Gln Pro Ser			
290	295	300	
GGT CAG CGC AGC TTC TGG GCC CAG CTG GGC ATG GCC CTG ACC AGG GAC			960
Gly Gln Arg Ser Phe Trp Ala Gln Leu Gly Met Ala Leu Thr Arg Asp			
305	310	315	320
AAC CAC GAC TTC TAT AAC CGG AAC TTC TGC CAG GGA CCC ACG GCC GAA			1008
Asn His His Phe Tyr Asn Arg Asn Phe Cys Gln Gly Pro Thr Ala Glu			
325	330	335	
TAGCCGTGAG CCGGAATCCT GCACGGTGCC ACCTCCACAC TCACCTCACC TCTGCCTGCC			1068
TGGTCTGACC CTCCCCTGCT CGCCTCCCTC ACCCCACACC TGACACAGAG CAGGCACTCA			1128
ATAAATGCTA CCGAAGGCTA AAAAAAAAAA AAAAAAAAAA AACCA			1173

(2) SEQ ID NO: 34 的信息:

- (i) 顺序特征:
 - (A) 长度: 336 个氨基酸
 - (B) 类型: 氨基酸
 - (D) 拓扑学: 线性
- (ii) 分子类型: 蛋白质

(xi) 顺序描述: SEQ ID NO: 34:

```

Leu Pro Ala Gly Thr Arg Leu Arg Arg Leu Asp Val Ser Cys Asn Ser
 1           5           10           15

Ile Ser Phe Val Ala Pro Gly Phe Phe Ser Lys Ala Lys Glu Leu Arg
      20           25           30

Glu Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp
 35           40           45

Phe Gly Pro Leu Ala Ser Ala Leu Gln Ile Leu Asp Val Ser Ala Asn
 50           55           60

Pro Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Glu
 65           70           75           80

Val Gln Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser
      85           90           95

Pro Gly Gln Leu Gln Gly Leu Ser Ile Phe Ala Gln Asp Leu Arg Leu
 100           105           110

Cys Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu
 115           120           125

Ala Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly
 130           135           140

Trp Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp
 145           150           155           160

Arg Gly Arg Gln Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala
      165           170           175

Phe Val Val Phe Asp Lys Thr Gln Ser Ala Val Ala Asp Trp Val Tyr
 180           185           190

Asn Glu Leu Arg Gly Gln Leu Glu Glu Cys Arg Gly Arg Trp Ala Leu
 195           200           205

Arg Leu Cys Leu Glu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe
 210           215           220

Glu Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val
 225           230           235           240

```

Leu Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu
 245 250 255
 Leu Ala Gln Gln Arg Leu Leu Glu Asp Arg Lys Asp Val Val Val Leu
 260 265 270
 Val Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Glu Arg Leu Arg
 275 280 285
 Gln Arg Leu Cys Arg Gln Ser Val Leu Leu Trp Pro His Gln Pro Ser
 290 295 300
 Gly Gln Arg Ser Phe Trp Ala Gln Leu Gly Met Ala Leu Thr Arg Asp
 305 310 315 320
 Asn His His Phe Tyr Asn Arg Asn Phe Cys Gln Gly Pro Thr Ala Glu
 325 330 335

(2) SEQ ID NO: 35 的信息:

- (i) 顺序特征:
- (A) 长度: 497 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线性

(ii) 分子类型: cDNA

(xi) 顺序描述: SEQ ID NO: 35:

TGGCCACAC GGACCGGTC AGTGGCCTCC TGCGCACCAG CTCCTGCTG GCTCAGCAGC 60
 GCCTGTTGGA AGACCGCAAG GACGTGGTGG TGTGGTGAT CCTGCGTCCG GATGCCCCAC 120
 CGTCCCGCTA TGTGCGACTG CGCCAGCGTC TCTGCCGCCA GAGTGTGCTC TTCTGGCCCC 180
 AGCGACCCAA CGGGCAGGGG GGCTTCTGGG CCCAGCTGAG TACAGCCCTG ACTAGGGACA 240
 ACCGCCACTT CTATAACCAG AACTTCTGCC GGGGACCTAC AGCAGAATAG CTCAGAGCAA 300
 CAGCTGGAAG CAGCTGCATC TTCATGTCTG GTTCCCGAGT TGCTCTGCCT GCCTTGCTCT 360
 GTCTTACTAC ACCGCTATTT GGCAAGTGCG CAATATATGC TACCAAGCCA CCAGGCCAC 420
 GGAGCAAAGG TTGGCTGTAA AGGGTAGTTT TCTTCCCATG CATCTTTCAG GAGAGTGAAG 480
 ATAGACACCA AACCAC 497

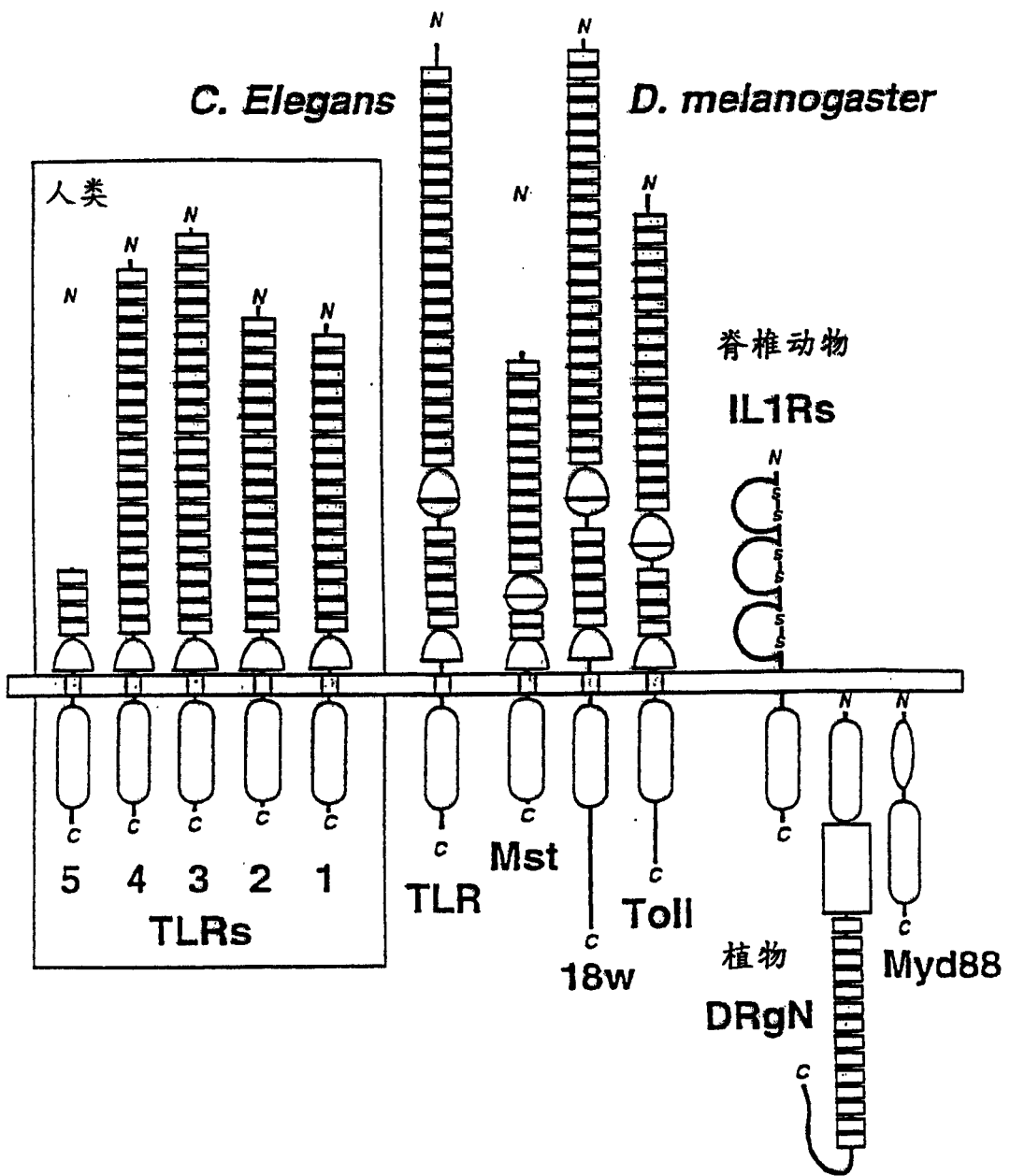


图 1

图 2A

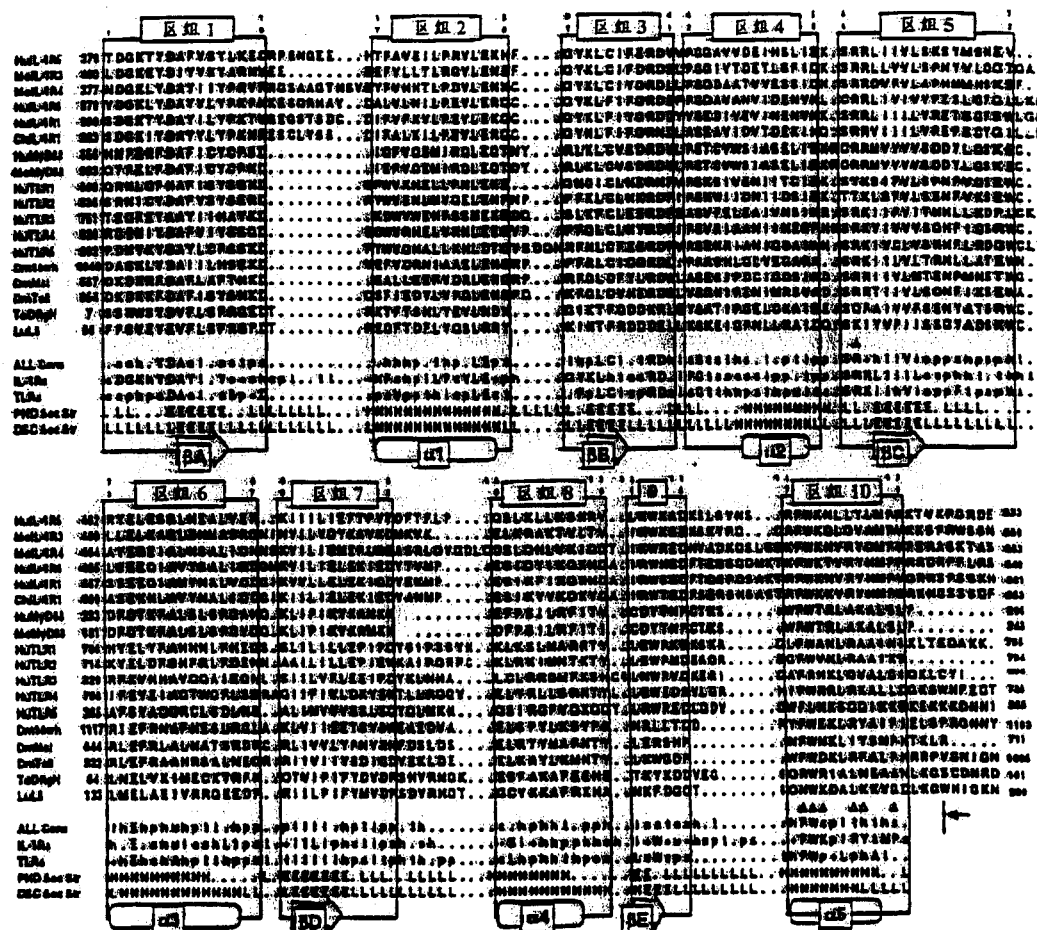
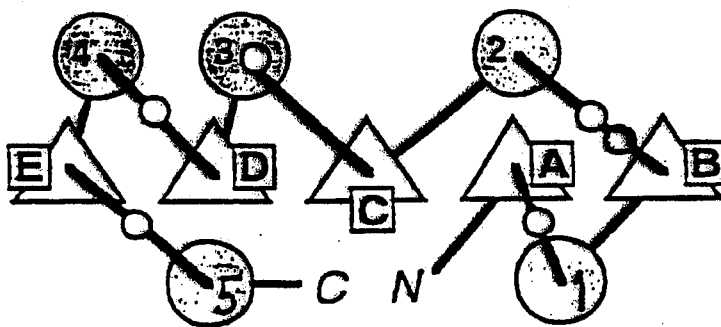


图 2B



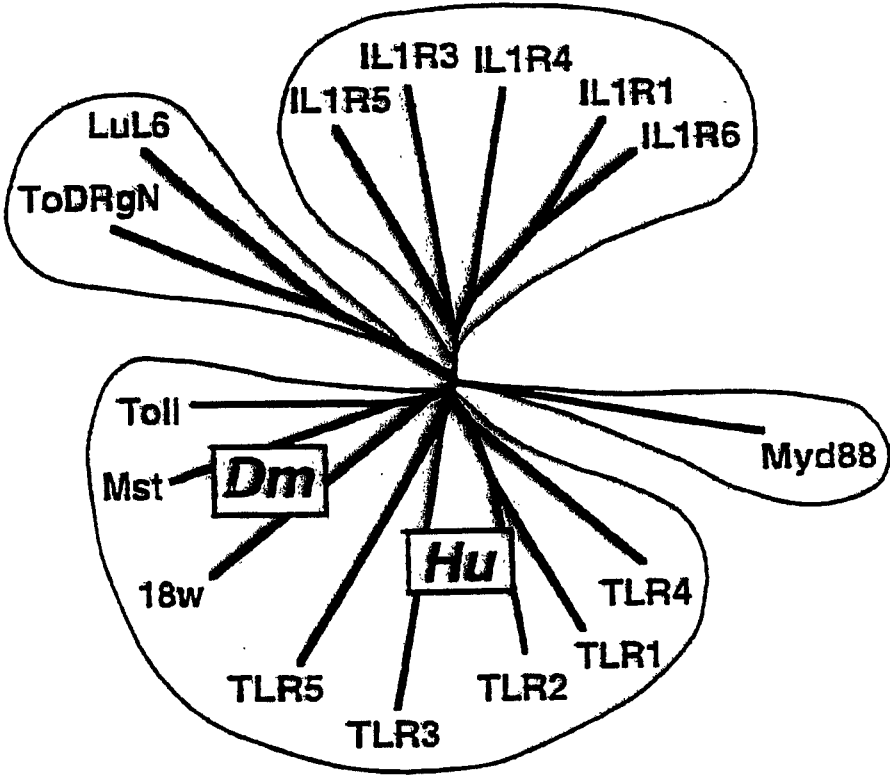


图 3

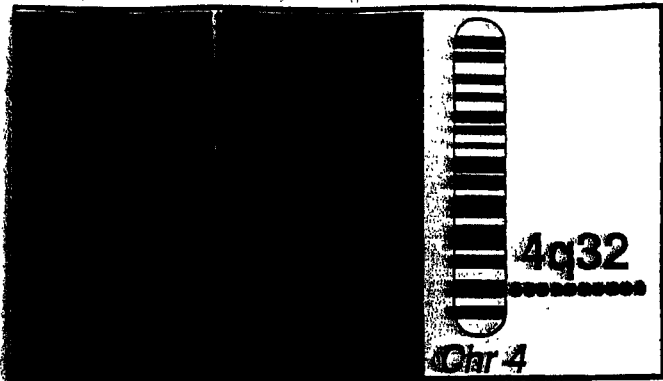


图 4A

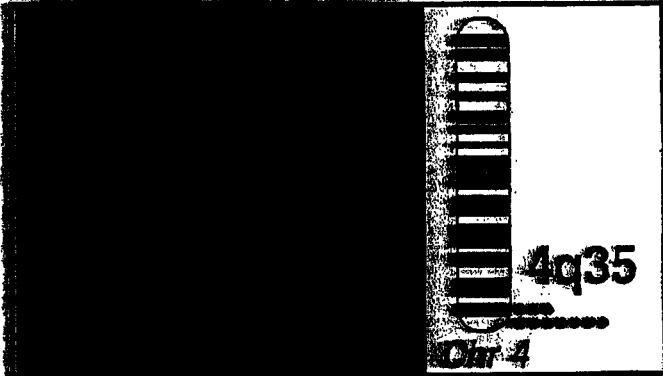


图 4B

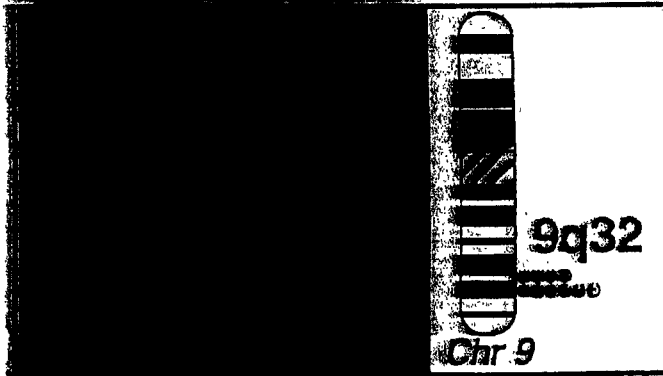


图 4C



图 4D

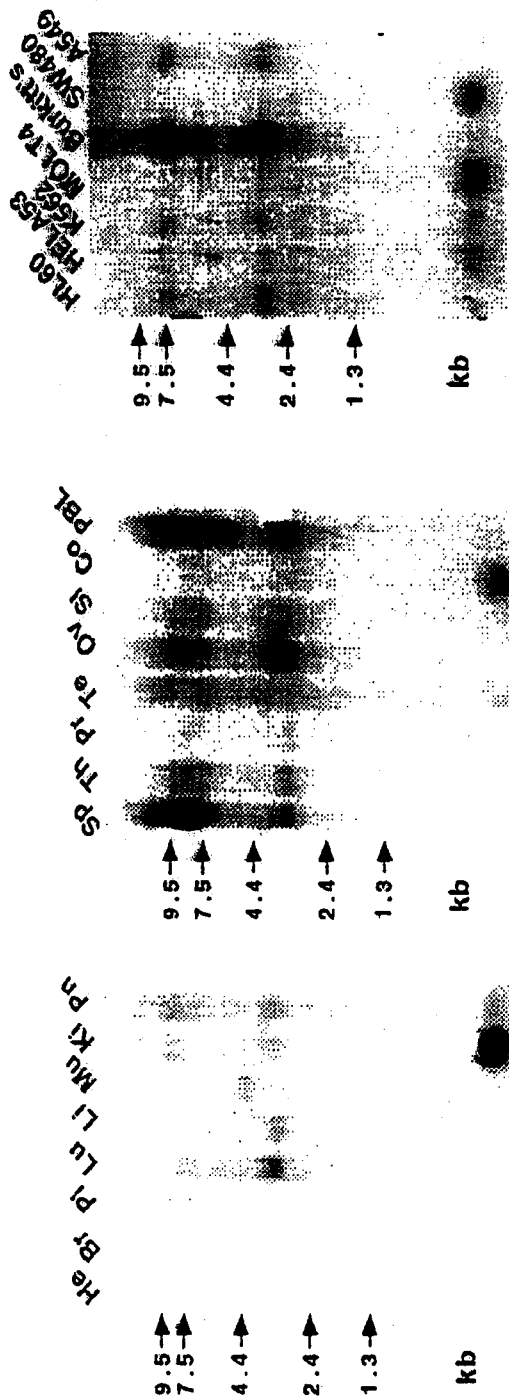


图 5A

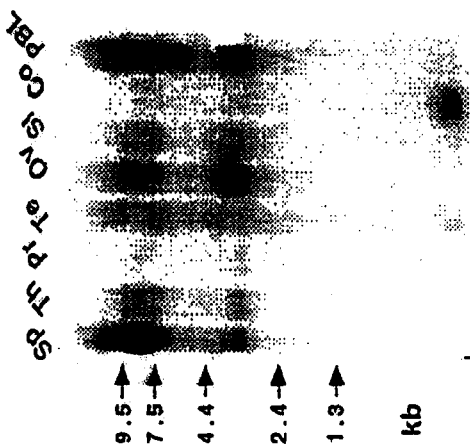


图 5B

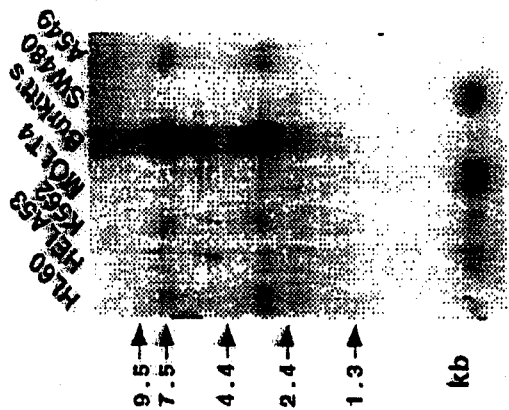


图 5C

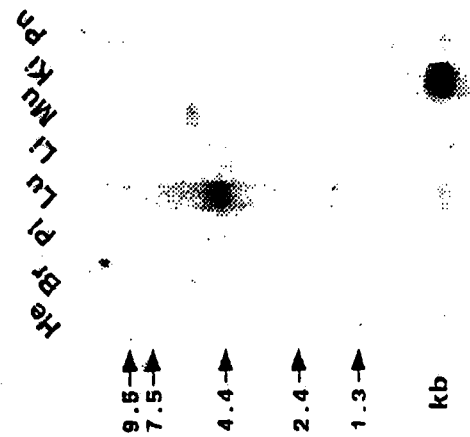


图 5D

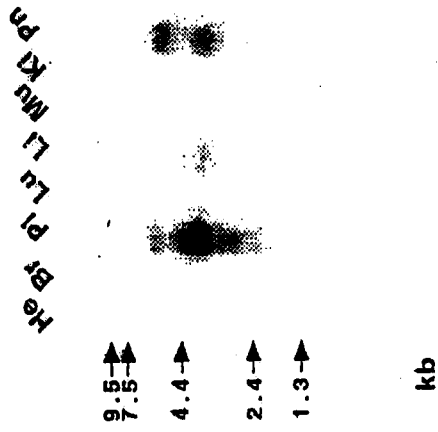


图 5E

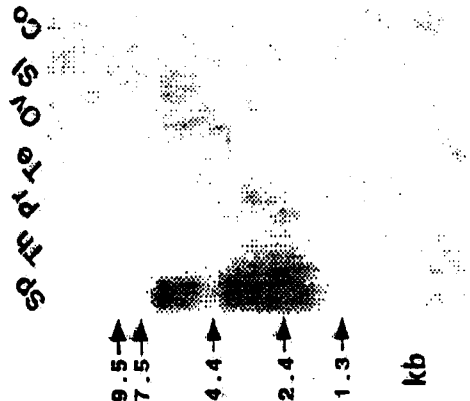


图 5F