

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C07K 14/435



[12] 发明专利申请公开说明书

C07K 14/47 C07K 16/18

C12N 15/11 C12N 15/12

C12N 15/63 C12N 15/64

C07H 21/00 C12P 21/02

[21] 申请号 02137722.7

[43] 公开日 2004 年 5 月 5 日

[11] 公开号 CN 1493586A

[22] 申请日 2002. 10. 30 [21] 申请号 02137722.7

[71] 申请人 上海博德基因开发有限公司

地址 200092 上海市中山北二路 1111 号 3 号楼 5 层

[72] 发明人 毛裕民 谢毅

[74] 专利代理机构 上海市华诚律师事务所

代理人 徐申民

权利要求书 1 页 说明书 21 页 附图 2 页

[54] 发明名称 一种多肽 - 人转录调控蛋白 38 和编码这种多肽的多核苷酸

[57] 摘要

本发明公开了一种多肽——人转录调控蛋白 38，编码此多肽的多核苷酸和经 DNA 重组技术产生这种多肽的方法。本发明还公开了此多肽用于治疗多种疾病的方法，如恶性肿瘤，血液病，HIV 感染和免疫性疾病和各类炎症等。本发明还公开了抗此多肽的拮抗剂及其治疗作用。本发明还公开了编码这种新的人转录调控蛋白 38 的多核苷酸的用途。

I S S N 1 0 0 8 - 4 2 7 4

- 1、一种分离的多肽-人转录调控蛋白 38，其特征在于它包含有：<210> 2 所示的氨基酸序列的多肽、或其多肽的活性片段、类似物或衍生物。
- 2、如权利要求 1 所述的多肽，其特征在于所述多肽、类似物或衍生物的氨基酸序列具有与<210> 2 所示的氨基酸序列至少 95%的相同性。
- 3、如权利要求 2 所述的多肽，其特征在于它包含具有<210> 2 所示的氨基酸序列的多肽。
- 4、一种分离的多核苷酸，其特征在于所述多核苷酸包含选自下组中的一种：
 - (a) 编码具有<210> 2 所示氨基酸序列的多肽或其片段、类似物、衍生物的多核苷酸；
 - (b) 与多核苷酸 (a) 互补的多核苷酸；或
 - (c) 与 (a) 或 (b) 有至少 70%相同性的多核苷酸。
- 5、如权利要求 4 所述的多核苷酸，其特征在于所述多核苷酸包含编码具有<210> 2 所示氨基酸序列的多核苷酸。
- 6、如权利要求 4 所述的多核苷酸，其特征在于所述多核苷酸的序列包含有<210> 1 中 328-1365 位的序列或<210> 1 中 1-2278 位的序列。
- 7、一种含有外源多核苷酸的重组载体，其特征在于它是由权利要求 4-6 中的任一权利要求所述多核苷酸与质粒、病毒或运载体表达载体构建而成的重组载体。
- 8、一种含有外源多核苷酸的遗传工程化宿主细胞，其特征在于它是选自于下列一种宿主细胞：
 - (a) 用权利要求 7 所述的重组载体转化或转导的宿主细胞；或
 - (b) 用权利要求 4-6 中的任一权利要求所述多核苷酸转化或转导的宿主细胞。
- 9、一种具有人转录调控蛋白 38 活性的多肽的制备方法，其特征在于所述方法包括：
 - (a) 在表达人转录调控蛋白 38 条件下，培养权利要求 8 所述的工程化宿主细胞；
 - (b) 从培养物中分离出具有人转录调控蛋白 38 活性的多肽。
- 10、一种能与多肽结合的抗体，其特征在于所述抗体是能与人转录调控蛋白 38 特异性结合的抗体。

一种多肽——人转录调控蛋白 38 和编码这种多肽的多核苷酸

技术领域

- 5 本发明属于生物技术领域，具体地说，本发明描述了一种多肽——人转录调控蛋白 38，以及编码此多肽的多核苷酸序列。本发明还涉及此多核苷酸和多肽的制备方法和应用。

背景技术

- 10 真核基因的转录调控对于基因的正常表达及发挥生物学功能是十分重要的，通常由转录调控因子来完成这一过程。转录调控因子在生物体内参与决定基因在何种组织及何种发育阶段开始转录，编码这类蛋白的基因如发生突变，不但该基因自身不能正常表达，而且受其调节的许多基因也不能正常的进行转录与表达。转录因子对基因表达的调控主要通过转录因子与特定的 DNA 序列结合、转录因子间的相互作用及转录因子与常规转录机构的相互作用在完成。根据结构基序的不同，已知的 DNA 结合蛋白可主要分为两类：含有螺旋-转角-螺旋基序的蛋白及锌指蛋白 [Kamal Chowdhury, Heidi Rohdewohld et al., *Nucleic Acids Research*, 1988, 16: 9995-10011]。

- 20 锌指蛋白为编码锌离子介导的核苷酸结合蛋白多基因家族中的成员，锌指蛋白按其结构特征又可分为各种不同的家族。人们已从酵母、果蝇、鼠及人等多种生物体中分离得到了各种类型的锌指蛋白。果蝇 Kruppel 基因类似的锌指蛋白分布最为广泛，且在生物体内有着重要的生物学功能。这些基因均含有锌指蛋白的特征性连续重复的 C2-H2 锌指蛋白结构域。研究发现，这些蛋白与基因的转录活化及抑制有关，这些蛋白的表达异常将引发各种发育紊乱性疾病、各种肿瘤的发生、各种遗传性疾病及免疫系统疾病 [Kamal Chowdhury, Heidi Rohdewohld et al., *Nucleic Acids Research*, 1988, 16: 9995-10011]。

- 30 所有的锌指蛋白 Kruppel 家族的成员均含有 28-30 个氨基酸长的保守的指重复序列 (F/Y) XCXXCXXFXXXXXLXXHXXHTGKPK，其中一些特定的氨基酸残基位点为高度保守的。这一序列在很多不同的锌指蛋白中均含有多个拷贝，其拷贝数不同 (锌指个数不同) 则功能也不同。锌指蛋白与不同长度的 DNA 的结合依赖于指结构的数量，多指结构可能与复合物的结合稳定性有关，而复合物是 RNA 聚合酶转录的作用位点。研究发现，许多锌指蛋白的锌指结构域相互连接区域也是高度保

守的,这一区域通常含有下列序列:His-Thr-Gly-Gly-Lys-Pro-(Tyr,Phe)-X-Cys,其中组氨酸与半胱氨酸为金属离子的结合位点,而X为可变氨基酸残基。这一区域对于锌指结构的形成是必需的,指结构的数量将直接影响锌指蛋白与不同长度的DNA结合,且多指结构与复合物的结合稳定性有关[Jeremy M. Berg, Annu. Rev.

5 Biophys. Chem, 1990, 19: 405-421]。

本发明的新的人转录调控蛋白与已知的人转录调控蛋白 Kruppel 家族中的成员 ZNF135 在蛋白水平上有 61% 的同一性及 75% 的相似性。且两者的蛋白序列均含有人转录调控蛋白 Kruppel 家族的特征性连续指重复序列及指结构相连区域。因而,本发明的新人转录调控蛋白 38 与人转录调控蛋白 Kruppel 家族的成员 ZNF135
10 相似,同为人转录调控蛋白 Kruppel 家族的成员,并具有相似的生物学功能。其在生物体内与各种固体肿瘤的发生有关[Tommerup N., Vissing H, 1995, Genomics, 27: 259-264],可用于诊断及治疗各种相关的恶性肿瘤、癌症及发育、代谢紊乱症等。

15 由于如上所述人转录调控蛋白 38 蛋白在调节细胞分裂和胚胎发育等机体重要功能中起重要作用,而且相信这些调节过程中涉及大量的蛋白,因而本领域中一直需要鉴定更多参与这些过程的人转录调控蛋白 38 蛋白,特别是鉴定这种蛋白的氨基酸序列。新人转录调控蛋白 38 蛋白编码基因的分离也为研究确定该蛋白在健康和疾病状态下的作用提供了基础。这种蛋白可能构成开发疾病诊断和/或治疗
20 药的基础,因此分离其编码 DNA 是非常重要的。

发明内容

本发明的一个目的是提供分离的新的多肽——人转录调控蛋白 38 以及其片段、类似物和衍生物。

25 本发明的另一个目的是提供编码该多肽的多核苷酸。

本发明的另一个目的是提供含有编码人转录调控蛋白 38 的多核苷酸的重组载体。

本发明的另一个目的是提供含有编码人转录调控蛋白 38 的多核苷酸的基因工程化宿主细胞。

30 本发明的另一个目的是提供生产人转录调控蛋白 38 的方法。

本发明的另一个目的是提供针对本发明的多肽——人转录调控蛋白 38 的抗体。

本发明涉及一种分离的多肽，该多肽是人源的，它包含：具有<210> 2 氨基酸序列的多肽、或其保守性变体、生物活性片段或衍生物。较佳地，该多肽是具有<210> 2 氨基酸序列的多肽。

5 本发明还涉及一种分离的多核苷酸，它包含选自下组的一种核苷酸序列或其变体：

(a) 编码具有<210> 2 氨基酸序列的多肽的多核苷酸；

(b) 与多核苷酸(a)互补的多核苷酸；

(c) 与(a)或(b)的多核苷酸序列具有至少70%相同性的多核苷酸。

10 更佳地，该多核苷酸的序列是选自下组的一种：(a) 具有<210> 1 中 328-1365 位的序列；和(b) 具有<210> 1 中 1-2278 位的序列。

本发明另外涉及一种含有本发明多核苷酸的载体，特别是表达载体；一种用该载体遗传工程化的宿主细胞，包括转化、转导或转染的宿主细胞；一种包括培养所述宿主细胞和回收表达产物的制备本发明多肽的方法。

本发明还涉及一种能与本发明多肽特异性结合的抗体。

15 本发明的其它方面由于本文的技术的公开，对本领域的技术人员而言是显而易见的。

20 本发明提供了一种多肽——人转录调控蛋白 38，其基本上是由<210> 2 所示的氨基酸序列组成的。本发明的多肽可以是重组多肽、天然多肽、合成多肽，优选重组多肽。本发明的多肽可以是天然纯化的产物，或是化学合成的产物，或使用重组技术从原核或真核宿主(例如，细菌、酵母、高等植物、昆虫和哺乳动物细胞)中产生。根据重组生产方案所用的宿主，本发明的多肽可以是糖基化的，或可以是非糖基化的。本发明的多肽还可包括或不包括起始的甲硫氨酸残基。

25 本发明还包括人转录调控蛋白 38 的片段、衍生物和类似物。如本发明所用，术语“片段”、“衍生物”和“类似物”是指基本上保持本发明的人转录调控蛋白 38 相同的生物学功能或活性的多肽。本发明多肽的片段、衍生物或类似物可以是：(I) 这样一种，其中一个或多个氨基酸残基被保守或非保守氨基酸残基(优选的是保守氨基酸残基)取代，并且取代的氨基酸可以是也可以不是由遗传密码子编码的；或者(II) 这样一种，其中一个或多个氨基酸残基上的某个基团被其它基团取代包含取代基；或者(III) 这样一种，其中成熟多肽与另一种化合物(比如延长多肽半衰期的化合物，例如聚乙二醇)融合；或者(IV) 这样一种，其中附加的氨基酸序列融合进成熟多肽而形成的多肽序列(如前导序列或分泌序列或用来纯化此多肽的序列或蛋白原序列)通过本文的阐述，这样的片段、衍生物和

类似物被认为在本领域技术人员的知识范围之内。

本发明提供了分离的核酸（多核苷酸），基本由编码具有<210> 2 氨基酸序列的多肽的多核苷酸组成。本发明的多核苷酸序列包括<210> 1 的核苷酸序列。本发明的多核苷酸是从人胎脑组织的 cDNA 文库中发现的。它包含的多核苷酸序列全长为 2278 个碱基，其开放读框（328-1365）编码了 345 个氨基酸。根据氨基酸序列同源比较发现，此多肽与 ZNF135 有 61% 的同源性，可推断出该人转录调控蛋白 38 具有 ZNF135 相似的结构和功能。

本发明的多核苷酸可以是 DNA 形式或是 RNA 形式。DNA 形式包括 cDNA、基因组 DNA 或人工合成的 DNA。DNA 可以是单链的或是双链的。DNA 可以是编码链或非编码链。编码成熟多肽的编码区序列可以与<210> 1 所示的编码区序列相同或者是简并的变异体。如本发明所用，“简并的变异体”在本发明中是指编码具有<210> 2 的蛋白质或多肽，但与<210> 1 所示的编码区序列有差别的核酸序列。

编码<210> 2 的成熟多肽的多核苷酸包括：只有成熟多肽的编码序列；成熟多肽的编码序列和各种附加编码序列；成熟多肽的编码序列（和任选的附加编码序列）以及非编码序列。

本发明还涉及上述描述多核苷酸的变异体，其编码与本发明有相同的氨基酸序列的多肽或多肽的片断、类似物和衍生物。此多核苷酸的变异体可以是天然发生的等位变异体或非天然发生的变异体。这些核苷酸变异体包括取代变异体、缺失变异体和插入变异体。如本领域所知的，等位变异体是一个多核苷酸的替换形式，它可能是一个或多个核苷酸的取代、缺失或插入，但不会从实质上改变其编码的多肽的功能。

本发明还涉及与以上所描述的序列杂交的多核苷酸（两个序列之间具有至少 50%，优选具有 70% 的相同性）。本发明特别涉及在严格条件下与本发明所述多核苷酸可杂交的多核苷酸。在本发明中，“严格条件”是指：(1) 在较低离子强度和较高温度下的杂交和洗脱，如 0.2×SSC, 0.1%SDS, 60℃；或 (2) 杂交时加用变性剂，如 50% (v/v) 甲酰胺，0.1%小牛血清/0.1%Ficoll, 42℃等；或 (3) 仅在两条序列之间的相同性至少在 95% 以上，更好是 97% 以上时才发生杂交。并且，可杂交的多核苷酸编码的多肽与<210> 2 所示的成熟多肽有相同的生物学功能和活性。

本发明还涉及与以上所描述的序列杂交的核酸片段。如本发明所用，“核酸片段”的长度至少含 10 个核苷酸，较好是至少 20-30 个核苷酸，更好是至少 50-60 个核苷酸，最好是至少 100 个核苷酸以上。核酸片段也可用于核酸的扩增技术（如 PCR）以确定和/或分离编码人转录调控蛋白 38 的多核苷酸。

本发明中的多肽和多核苷酸优选以分离的形式提供，更佳地被纯化至均质。

本发 5 明的编码人转录调控蛋白 38 的特异的多核苷酸序列能用多种方法获得。例如，用本领域熟知的杂交技术分离多核苷酸。这些技术包括但不限于：1) 用探针与基因组或 cDNA 文库杂交以检出同源的多核苷酸序列，和 2) 表达文库的抗体筛选以检出具有共同结构特征的克隆的多核苷酸片段。

本发 10 明的 DNA 片段序列也能用下列方法获得：1) 从基因组 DNA 分离双链 DNA 序列；2) 化学合成 DNA 序列以获得所述多肽的双链 DNA。

上述提到的方法中，分离基因组 DNA 最不常用。DNA 序列的直接化学合成是经常选用的方法。更经常选用的方法是 cDNA 序列的分离。分离感兴趣的 cDNA 的标准方法是从高表达该基因的供体细胞分离 mRNA 并进行逆转录，形成质粒或噬菌体 cDNA 文库。提取 mRNA 的方法已有多种成熟的技术，试剂盒也可从商业途径获得 (Qiagene)。而构建 cDNA 文库也是通常的方法 (Sambrook, et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, New York, 1989)。还可得到商业供应的 cDNA 文库，如 Clontech 公司的不同 cDNA 文库。当 15 结合使用聚合酶反应技术时，即使极少的表达产物也能克隆。

可用常规方法从这些 cDNA 文库中筛选本发明的基因。这些方法包括(但不限于)：(1) DNA-DNA 或 DNA-RNA 杂交；(2) 标志基因功能的出现或丧失；(3) 测定人转录调控蛋白 38 的转录本的水平；(4) 通过免疫学技术或测定生物学活性，来检测基因表达的蛋白产物。上述方法可单用，也可多种方法联合应用。

20 在第(1)种方法中，杂交所用的探针是与本发明的多核苷酸的任何一部分同源，其长度至少 10 个核苷酸，较好是至少 30 个核苷酸，更好是至少 50 个核苷酸，最好是至少 100 个核苷酸。此外，探针的长度通常在 2000 个核苷酸之内，较佳的为 1000 个核苷酸之内。此处所用的探针通常是在本发明的基因序列信息的基础上化学合成的 DNA 序列。本发明的基因本身或者片段当然可以用作探针。DNA 探针 25 的标记可用放射性同位素，荧光素或酶(如碱性磷酸酶)等。

在第(4)种方法中，检测人转录调控蛋白 38 基因表达的蛋白产物可用免疫学技术如 Western 印迹法，放射免疫沉淀法，酶联免疫吸附法 (ELISA) 等。

应用 PCR 技术扩增 DNA/RNA 的方法 (Saiki, et al. Science 1985; 230: 1350-1354) 被优选用于获得本发明的基因。特别是很难从文库中得到全 30 长的 cDNA 时，可优选使用 RACE 法 (RACE - cDNA 末端快速扩增法)，用于 PCR 的引物可根据本文所公开的本发明的多核苷酸序列信息适当地选择，并可用常规方法合成。可用常规方法如通过凝胶电泳分离和纯化扩增的 DNA/RNA 片段。

如上所述得到的本发明的基因，或者各种 DNA 片段等的多核苷酸序列可用常规方法如双脱氧链终止法 (Sanger et al. PNAS, 1977, 74: 5463-5467) 测定。这类多核苷酸序列测定也可用商业测序试剂盒等。为了获得全长的 cDNA 序列，测序需反复进行。有时需要测定多个克隆的 cDNA 序列，才能拼接成全长的 cDNA 序列。

5 本发明也涉及包含本发明的多核苷酸的载体，以及用本发明的载体或直接用人转录调控蛋白 38 编码序列经基因工程产生的宿主细胞，以及经重组技术产生本发明所述多肽的方法。

本发明中，编码人转录调控蛋白 38 的多核苷酸序列可插入到载体中，以构成含有本发明所述多核苷酸的重组载体。术语“载体”指本领域熟知的细菌质粒、噬菌体、酵母质粒、植物细胞病毒、哺乳动物细胞病毒如腺病毒、逆转录病毒或其它载体。在本发明中适用的载体包括但不限于：在细菌中表达的基于 T7 启动子的表达载体 (Rosenberg, et al. Gene, 1987, 56: 125)；在哺乳动物细胞中表达的 pMSXND 表达载体 (Lee and Nathans, J Bio Chem. 263: 3521, 1988) 和在昆虫细胞中表达的来源于杆状病毒的载体。总之，只要能在宿主体内复制和稳定，任何质粒和载体都可以用于构建重组表达载体。表达载体的一个重要特征是通常含有复制起始点、启动子、标记基因和翻译调控元件。

本领域的技术人员熟知的方法能用于构建含编码人转录调控蛋白 38 的 DNA 序列和合适的转录/翻译调控元件的表达载体。这些方法包括体外重组 DNA 技术、DNA 合成技术、体内重组技术等 (Sambrook, et al. Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory. New York, 1989)。所述的 DNA 序列可有效连接到表达载体中的适当启动子上，以指导 mRNA 合成。这些启动子的代表性例子有：大肠杆菌的 lac 或 trp 启动子；λ 噬菌体的 PL 启动子；真核启动子包括 CMV 立即早期启动子、HSV 胸苷激酶启动子、早期和晚期 SV40 启动子、反转录病毒的 LTRs 和其它一些已知的可控制基因在原核细胞或真核细胞或其病毒中表达的启动子。表达载体还包括翻译起始用的核糖体结合位点和转录终止子等。在载体中插入增强子序列将会使其在高等真核细胞中的转录得到增强。增强子是 DNA 表达的顺式作用因子，通常大约有 10 到 300 个碱基对，作用于启动子以增强基因的转录。可举的例子包括在复制起始点晚期一侧的 100 到 270 个碱基对的 SV40 增强子、在复制起始点晚期一侧的多瘤增强子以及腺病毒增强子等。

30 此外，表达载体优选地包含一个或多个选择性标记基因，以提供用于选择转化的宿主细胞的表型性状，如真核细胞培养用的二氢叶酸还原酶、新霉素抗性以及绿色荧光蛋白 (GFP)，或用于大肠杆菌的四环素或氨苄青霉素抗性等。

本领域一般技术人员都清楚如何选择适当的载体/转录调控元件(如启动子、增强子等)和选择性标记基因。

5 本发明中,编码人转录调控蛋白 38 的多核苷酸或含有该多核苷酸的重组载体可转化或转导入宿主细胞,以构成含有该多核苷酸或重组载体的基因工程化宿主细胞。术语“宿主细胞”指原核细胞,如细菌细胞;或是低等真核细胞,如酵母细胞;或是高等真核细胞,如哺乳动物细胞。代表性例子有:大肠杆菌,链霉菌属;细菌细胞如鼠伤寒沙门氏菌;真菌细胞如酵母;植物细胞;昆虫细胞如果蝇 S2 或 Sf9;动物细胞如 CHO、COS 或 Bowes 黑素瘤细胞等。

10 用本发明所述的 DNA 序列或含有所述 DNA 序列的重组载体转化宿主细胞可用本领域技术人员熟知的常规技术进行。当宿主为原核生物如大肠杆菌时,能吸收 DNA 的感受态细胞可在指数生长期后收获,用 CaCl_2 法处理,所用的步骤在本领域众所周知。可供选择的是用 MgCl_2 。如果需要,转化也可用电穿孔的方法进行。当宿主是真核生物,可选用如下的 DNA 转染方法:磷酸钙共沉淀法,或者常规机械方法如显微注射、电穿孔、脂质体包装等。

15 通过常规的重组 DNA 技术,利用本发明的多核苷酸序列可用来表达或生产重组的人转录调控蛋白 38 (Science, 1984; 224: 1431)。一般来说有以下步骤:

(1). 用本发明的编码人 人转录调控蛋白 38 的多核苷酸(或变体),或用含有该多核苷酸的重组表达载体转化或转导合适的宿主细胞;

(2). 在合适的培养基中培养宿主细胞;

20 (3). 从培养基或细胞中分离、纯化蛋白质。

在步骤(2)中,根据所用的宿主细胞,培养中所用的培养基可选自各种常规培养基。在适于宿主细胞生长的条件下进行培养。当宿主细胞生长到适当的细胞密度后,用合适的方法(如温度转换或化学诱导)诱导选择的启动子,将细胞再培养一段时间。

25 在步骤(3)中,重组多肽可包被于细胞内、或在细胞膜上表达、或分泌到细胞外。如果需要,可利用其物理的、化学的和其它特性通过各种分离方法分离和纯化重组的蛋白。这些方法是本领域技术人员所熟知的。这些方法包括但不限于:常规的复性处理、蛋白沉淀剂处理(盐析方法)、离心、渗透破菌、超声波处理、超离心、分子筛层析(凝胶过滤)、吸附层析、离子交换层析、高效液相层析
30 (HPLC)和其它各种液相层析技术及这些方法的结合。

本发明也提供了筛选化合物以鉴定提高(激动剂)或阻遏(拮抗剂)人转录调控蛋白 38 的药剂的方法。激动剂提高人转录调控蛋白 38 刺激细胞增殖等生物功能,

而拮抗剂阻止和治疗与细胞过度增殖有关的紊乱如各种癌症。例如，能在药物的存在下，将哺乳动物细胞或表达人转录调控蛋白 38 的膜制剂与标记的人转录调控蛋白 38 一起培养。然后测定药物提高或阻遏此相互作用的能力。

5 人转录调控蛋白 38 的拮抗剂包括筛选出的抗体、化合物、受体缺失物和类似物等。人转录调控蛋白 38 的拮抗剂可以与人转录调控蛋白 38 结合并消除其功能，或是抑制该多肽的产生，或是与该多肽的活性位点结合使该多肽不能发挥生物学功能。

10 在筛选作为拮抗剂的化合物时，可以将人转录调控蛋白 38 加入生物分析测定中，通过测定化合物对人转录调控蛋白 38 和其受体之间相互作用的影响来确定化合物是否是拮抗剂。用上述筛选化合物的同样方法，可以筛选出起拮抗剂作用的受体缺失物和类似物。能与人转录调控蛋白 38 结合的多肽分子可通过筛选由各种可能组合的氨基酸结合于固相物组成的随机多肽库而获得。筛选时，一般应对人转录调控蛋白 38 分子进行标记。

15 本发明提供了用多肽，及其片段、衍生物、类似物或它们的细胞作为抗原以生产抗体的方法。这些抗体可以是多克隆抗体或单克隆抗体。本发明还提供了针对人转录调控蛋白 38 抗原决定簇的抗体。这些抗体包括(但不限于):多克隆抗体、单克隆抗体、嵌合抗体、单链抗体、Fab 片段和 Fab 表达文库产生的片段。

20 多克隆抗体的生产可用人转录调控蛋白 38 直接注射免疫动物(如家兔,小鼠,大鼠等)的方法得到,多种佐剂可用于增强免疫反应,包括但不限于弗氏佐剂等。制备人转录调控蛋白 38 的单克隆抗体的技术包括但不限于杂交瘤技术(Kohler and Milstein. Nature, 1975, 256: 495-497), 三瘤技术, 人 B-细胞杂交瘤技术, EBV-杂交瘤技术等。将人恒定区和非人源的可变区结合的嵌合抗体可用已有的技术生产(Morrison et al, PNAS, 1985, 81: 6851)。而已有的生产单链抗体的技术(U. S. Pat No. 4946778)也可用于生产抗人转录调控蛋白 38 的单链抗体。

25 抗人转录调控蛋白 38 的抗体可用于免疫组织化学技术中,检测活检标本中的人转录调控蛋白 38。

与人转录调控蛋白 38 结合的单克隆抗体也可用放射性同位素标记,注入体内可跟踪其位置和分布。这种放射性标记的抗体可作为一种非创伤性诊断方法用于肿瘤细胞的定位和判断是否有转移。

30 抗体还可用于设计针对体内某一特殊部位的免疫毒素。如人转录调控蛋白 38 高亲和性的单克隆抗体可与细菌或植物毒素(如白喉毒素,蓖麻蛋白,红豆碱等)共价结合。一种通常的方法是用巯基交联剂如 SPDP, 攻击抗体的氨基,通过二硫

键的交换，将毒素结合于抗体上，这种杂交抗体可用于杀灭人转录调控蛋白 38 阳性的细胞。

本发明中的抗体可用于治疗或预防与人转录调控蛋白 38 相关的疾病。给予适当剂量的抗体可以刺激或阻断人转录调控蛋白 38 的产生或活性。

- 5 本发明还涉及定量和定位检测人转录调控蛋白 38 水平的诊断试验方法。这些试验是本领域所熟知的，且包括 FISH 测定和放射免疫测定。试验中所检测的人转录调控蛋白 38 水平，可以用作解释人转录调控蛋白 38 在各种疾病中的重要性和用于诊断人转录调控蛋白 38 起作用的疾病。

- 10 本发明的多肽还可用作肽谱分析，例如，多肽可用物理的、化学或酶进行特异性切割，并进行一维或二维或三维的凝胶电泳分析，更好的是进行质谱分析。

- 15 编码人转录调控蛋白 38 的多核苷酸也可用于多种治疗目的。基因治疗技术可用于治疗由于人转录调控蛋白 38 的无表达或异常/无活性表达所致的细胞增殖、发育或代谢异常。重组的基因治疗载体(如病毒载体)可设计用于表达变异的人转录调控蛋白 38，以抑制内源性的人转录调控蛋白 38 活性。例如，一种变异的人
20 转录调控蛋白 38 可以是缩短的、缺失了信号传导功能域的人转录调控蛋白 38，虽可与下游的底物结合，但缺乏信号传导活性。因此重组的基因治疗载体可用于治疗人转录调控蛋白 38 表达或活性异常所致的疾病。来源于病毒的表达载体如逆转录病毒、腺病毒、腺病毒相关病毒、单纯疱疹病毒、细小病毒等可用于将编码人转录调控蛋白 38 的多核苷酸转移至细胞内。构建携带编码人转录调控蛋白 38
25 的多核苷酸的重组病毒载体的方法可见于已有文献(Sambrook, et al.)。另外重组编码人转录调控蛋白 38 的多核苷酸可包装到脂质体中转移至细胞内。

多核苷酸导入组织或细胞内的方法包括：将多核苷酸直接注入到体内组织中；或在体外通过载体(如病毒、噬菌体或质粒等)先将多核苷酸导入细胞中，再将细胞移植到体内等。

- 25 抑制人转录调控蛋白 38 mRNA 的寡核苷酸(包括反义 RNA 和 DNA)以及核酶也在本发明的范围之内。核酶是一种能特异性分解特定 RNA 的酶样 RNA 分子，其作用机制是核酶分子与互补的靶 RNA 特异性杂交后进行核酸内切作用。反义的 RNA 和 DNA 及核酶可用已有的任何 RNA 或 DNA 合成技术获得，如固相磷酸酰胺化学合成法合成寡核苷酸的技术已广泛应用。反义 RNA 分子可通过编码该 RNA 的 DNA 序
30 列在体外或体内转录获得。这种 DNA 序列已整合到载体的 RNA 聚合酶启动子的下游。为了增加核酸分子的稳定性，可用多种方法对其进行修饰，如增加两侧的序列长度，核糖核苷之间的连接应用磷酸硫酸酯键或肽键而非磷酸二酯键。

编码人转录调控蛋白 38 的多核苷酸可用于与人转录调控蛋白 38 的相关疾病的诊断。编码人转录调控蛋白 38 的多核苷酸可用于检测人转录调控蛋白 38 的表达与否或在疾病状态下人转录调控蛋白 38 的异常表达。如编码人转录调控蛋白 38 的 DNA 序列可用于对活检标本进行杂交以判断人转录调控蛋白 38 的表达状况。杂交技术包括 Southern 印迹法, Northern 印迹法、原位杂交等。这些技术方法都是公开的成熟技术, 相关的试剂盒都可从商业途径得到。本发明的多核苷酸的一部分或全部可作为探针固定在微阵列 (Microarray) 或 DNA 芯片 (又称为“基因芯片”) 上, 用于分析组织中基因的差异表达分析和基因诊断。用人转录调控蛋白 38 特异的引物进行 RNA-聚合酶链反应 (RT-PCR) 体外扩增也可检测人转录调控蛋白 38 的转录产物。

检测人转录调控蛋白 38 基因的突变也可用于诊断人转录调控蛋白 38 相关的疾病。人转录调控蛋白 38 突变的形式包括与正常野生型人转录调控蛋白 38 DNA 序列相比的点突变、易位、缺失、重组和其它任何异常等。可用已有的技术如 Southern 印迹法、DNA 序列分析、PCR 和原位杂交检测突变。另外, 突变有可能影响蛋白的表达, 因此用 Northern 印迹法、Western 印迹法可间接判断基因有无突变。

本发明的序列对染色体鉴定也是有价值的。该序列会特异性地针对某条人染色体具体位置且并可以与其杂交。目前, 需要鉴定染色体上的各基因的具体位点。现在, 只有很少的基于实际序列数据 (重复多态性) 的染色体标记物可用于标记染色体位置。根据本发明, 为了将这些序列与疾病相关基因相关联, 其重要的第一步就是将这些 DNA 序列定位于染色体上。

简而言之, 根据 cDNA 制备 PCR 引物 (优选 15-35bp), 可以将序列定位于染色体上。然后, 将这些引物用于 PCR 筛选含各条人染色体的体细胞杂合细胞。只有那些含有相应于引物的人基因的杂合细胞会产生扩增的片段。

体细胞杂合细胞的 PCR 定位法, 是将 DNA 定位到具体染色体的快捷方法。使用本发明的寡核苷酸引物, 通过类似方法, 可利用一组来自特定染色体的片段或大量基因组克隆而实现亚定位。可用于染色体定位的其它类似策略包括原位杂交、用标记的流式分选的染色体预筛选和杂交预选, 从而构建染色体特异的 cDNA 库。

将 cDNA 克隆与中期染色体进行荧光原位杂交 (FISH), 可以在一个步骤中精确地进行染色体定位。此技术的综述, 参见 Verma 等, *Human Chromosomes: a Manual of Basic Techniques*, Pergamon Press, New York (1988)。

一旦序列被定位到准确的染色体位置, 此序列在染色体上的物理位置就可以

与基因图数据相关联。这些数据可见于例如, V. McKusick, Mendelian Inheritance in Man(可通过与Johns Hopkins University Welch Medical Library联机获得)。然后可通过连锁分析, 确定基因与业已定位到染色体区域上的疾病之间的关系。

- 5 接着, 需要测定患病和未患病个体间的cDNA或基因组序列差异。如果在一些或所有的患病个体中观察到某突变, 而该突变在任何正常个体中未观察到, 则该突变可能是疾病的病因。比较患病和未患病个体, 通常涉及首先寻找染色体中结构的变化, 如从染色体水平可见的或用基于cDNA序列的PCR可检测的缺失或易位。根据目前的物理作图和基因定位技术的分辨能力, 被精确定位至与疾病有关的染色体区域的cDNA, 可以是50至500个潜在致病基因间之一种(假定1兆碱基作图分辨能力和每20kb对应于一个基因)。

- 10 可以将本发明的多肽、多核苷酸及其模拟物、激动剂、拮抗剂和抑制剂与合适的药物载体组合后使用。这些载体可以是水、葡萄糖、乙醇、盐类、缓冲液、甘油以及它们的组合。组合物包含安全有效量的多肽或拮抗剂以及不影响药物效果的载体和赋形剂。这些组合物可以作为药物用于疾病治疗。

- 15 本发明还提供含有一种或多种容器的药盒或试剂盒, 容器中装有一种或多种本发明的药用组合物成分。与这些容器一起, 可以有由制造、使用或销售药品或生物制品的政府管理机构所给出的指示性提示, 该提示反映出生产、使用或销售的政府管理机构许可其在人体上施用。此外, 本发明的多肽可以与其它的治疗化合物结合使用。

附图说明

下列附图用于说明本发明的具体实施方案, 而不用于限定由权利要求书所界定的本发明范围。

- 25 图1是本发明人转录调控蛋白38和ZNF135的氨基酸序列同源性比较图。上方序列是人转录调控蛋白38, 下方序列是ZNF135。相同氨基酸在两个序列间用单字符氨基酸表示, 相似氨基酸用“+”表示。

图2为分离的人转录调控蛋白38的聚丙烯酰胺凝胶电泳图(SDS-PAGE)。38kDa为蛋白质的分子量。箭头所指为分离出的蛋白条带。

30

具体实施方式

下面结合具体实施例, 进一步阐述本发明。应理解, 这些实施例仅用于说明本

发明而不用于限制本发明的范围。下列实施例中未注明具体条件的实验方法，通常按照常规条件如 Sambrook 等人，分子克隆：实验室手册 (New York: Cold Spring Harbor Laboratory Press, 1989) 中所述的条件，或按照制造厂商所建议的条件。

实施例 1: 人转录调控蛋白 38 的克隆

- 5 用异硫氰酸胍/酚/氯仿一步法提取人胎脑总 RNA。用 Quik mRNA Isolation Kit (Qiagen 公司产品) 从总 RNA 中分离 poly(A) mRNA。2ug poly(A) mRNA 经逆转录形成 cDNA。用 Smart cDNA 克隆试剂盒 (购自 Clontech) 将 cDNA 片段定向插入到 pBSK (+) 载体 (Clontech 公司产品) 的多克隆位点上，转化 DH5 α ，细菌形成 cDNA 文库。用 Dye terminate cycle reaction sequencing kit (Perkin-Elmer 公司产品)
- 10 和 ABI 377 自动测序仪 (Perkin-Elmer 公司) 测定所有克隆的 5' 和 3' 末端的序列。将测定的 cDNA 序列与已有的公共 DNA 序列数据库 (Genebank) 进行比较，结果发现其中一个克隆 0385B11 的 cDNA 序列为新的 DNA。通过合成一系列引物对该克隆所含的插入 cDNA 片段进行双向测定。结果表明，0385B11 克隆所含的全长 cDNA 为 2278bp (如 <210> 1 所示)，从第 328bp 至 1365bp 有一个 1038bp 的开放阅读框架 (ORF)，
- 15 编码一个新的蛋白质 (如 <210> 2 所示)。我们将此克隆命名为 pBS-0385B11，编码的蛋白质命名为人转录调控蛋白 38。

实施例 2: cDNA 克隆的同源检索

- 20 将本发明的人转录调控蛋白 38 的序列及其编码的蛋白序列，用 Blast 程序 (Basiclocal Alignment search tool) [Altschul, SF et al. J. Mol. Biol. 1990; 215: 403-10]，在 Genbank、Swissport 等数据库进行同源检索。与本发明的人转录调控蛋白 38 同源性最高的基因是一种已知的 ZNF135，其编码的蛋白在 Genbank 的准入号为 U09413。蛋白质同源结果示于图 1，两者高度同源，其相同性为 61%；相似性为 75%。

25

实施例 3: 用 RT-PCR 方法克隆编码人转录调控蛋白 38 的基因

用胎脑细胞总 RNA 为模板，以 oligo-dT 为引物进行逆转录反应合成 cDNA，用 Qiagen 的试剂盒纯化后，用下列引物进行 PCR 扩增：

- Primer1: 5'- CTTTCTGATGAACATGGGAAATG -3' (<210> 3)
- 30 Primer2: 5'- AATAACAACTTCTCCACATTTAT -3' (<210> 4)
- Primer1 为位于 <210> 1 的 5' 端的第 1bp 开始的正向序列；
- Primer2 为 <210> 1 的中的 3' 端反向序列。

扩增反应的条件：在50 μ l的反应体积中含有50mmol/L KCl, 10mmol/L Tris-Cl, (pH8.5), 1.5mmol/L MgCl₂, 200 μ mol/L dNTP, 10pmol引物, 1U的Taq DNA聚合酶(Clontech公司产品)。在PE9600型DNA热循环仪(Perkin-Elmer公司)上按下列条件反应25个周期：94°C 30sec; 55°C 30sec; 72°C 2min。在RT-PCR时同时设

5 β -actin为阳性对照和模板空白为阴性对照。扩增产物用QIAGEN公司的试剂盒纯化, 用TA克隆试剂盒连接到pCR载体上(Invitrogen公司产品)。DNA序列分析结果表明PCR产物的DNA序列与<210> 1所示的1-2278bp完全相同。

实施例4: Northern 印迹法分析人转录调控蛋白38基因的表达:

10 用一步法提取总RNA[Anal. Biochem 1987, 162, 156-159]。该法包括酸性硫氰酸胍苯酚-氯仿抽提。即用4M异硫氰酸胍-25mM柠檬酸钠, 0.2M乙酸钠(pH4.0)对组织进行匀浆, 加入1倍体积的苯酚和1/5体积的氯仿-异戊醇(49:1), 混合后离心。吸出水相层, 加入异丙醇(0.8体积)并将混合物离心得到RNA沉淀。将得到的RNA沉淀用70%乙醇洗涤, 干燥并溶于水中。用20 μ g RNA, 在含20mM 3-(N-

15 吗啉代)丙磺酸(pH7.0)-5mM乙酸钠-1mM EDTA-2.2M甲醛的1.2%琼脂糖凝胶上进行电泳。然后转移至硝酸纤维素膜上。用 α -³²P dATP通过随机引物法制备³²P-标记的DNA探针。所用的DNA探针为图1所示的PCR扩增的人转录调控蛋白38编码区序列(328bp至1365bp)。将³²P-标记的探针(约 2×10^6 cpm/ml)与转移了RNA的硝酸纤维素膜在一溶液

20 中于42°C杂交过夜, 该溶液包含50%甲酰胺-25mM KH₂PO₄(pH7.4)-5 \times SSC-5 \times Denhardt's溶液和200 μ g/ml鲑精DNA。杂交之后, 将滤膜在1 \times SSC-0.1%SDS中于55°C洗30min。然后, 用Phosphor Imager进行分析和定量。

实施例5: 重组人转录调控蛋白38的体外表达、分离和纯化

25 根据<210> 1和图1所示的编码区序列, 设计出一对特异性扩增引物, 序列如下:

Primer3: 5'- CATGCTAGCATGGAAGCTCTGTGGTGGCAGTGAAT -3' (<210> 5)

Primer4: 5'- CATGGATCCTCAGGACTTCCCAAATCCATTACATG -3' (<210> 6)

此两段引物的5'端分别含有NheI和BamHI酶切位点, 其后分别为目的基因5'端和3'端的编码序列, NheI和BamHI酶切位点相应于表达载体质粒

30 pET-28b(+) (Novagen公司产品, Cat. No. 69865.3)上的选择性内切酶位点。以含有全长目的基因的pBS-0385B11质粒为模板, 进行PCR反应。PCR反应条件为: 总体积50 μ l中含pBS-0385B11质粒10pg、引物Primer-3和Primer-4分别为10pmol、

Advantage polymerase Mix (Clontech公司产品) 1 μ l。循环参数: 94°C 20s, 60°C 30s, 68°C 2 min, 共25个循环。用NheI和BamHI分别对扩增产物和质粒pET-28(+)进行双酶切, 分别回收大片段, 并用T4连接酶连接。连接产物转化用氯化钙法大肠杆菌DH5 α , 在含卡那霉素(终浓度30 μ g/ml)的LB平板培养过夜后, 用菌落PCR方法筛选阳性克隆, 并进行测序。挑选序列正确的阳性克隆(pET-0385B11)用氯化钙法将重组质粒转化大肠杆菌BL21(DE3)plySs (Novagen公司产品)。在含卡那霉素(终浓度30 μ g/ml)的LB液体培养基中, 宿主菌BL21(pET-0385B11)在37°C培养至对数生长期, 加入IPTG至终浓度1mmol/L, 继续培养5小时。离心收集菌体, 经超声波破菌, 离心收集上清, 用能与6个组氨酸(6His-Tag)结合的亲和层析柱His. Bind Quick Cartridge (Novagen公司产品)进行层析, 得到了纯化的目的蛋白人转录调控蛋白38。经SDS-PAGE电泳, 在38kDa处得到一单一的条带(图2)。将该条带转移至PVDF膜上用Edams水解法进行N-端氨基酸序列分析, 结果N-端15个氨基酸与<210> 2所示的N-端15个氨基酸残基完全相同。

15 实施例6 抗人转录调控蛋白38抗体的产生

用多肽合成仪(PE公司产品)合成下述人转录调控蛋白38特异性的多肽:
NH₂-Met-Glu-Leu-Cys-Gly-Gly-Ser-Glu-Tyr-Gly-Lys-Thr-Ser-His-Leu-COOH。

将该多肽分别与血蓝蛋白和牛血清白蛋白耦合形成复合, 方法参见:
Avrameas, et al. *Immunochemistry*, 1969; 6: 43。用4mg上述血蓝蛋白多肽复合物加上完全弗氏佐剂免疫家兔, 15天后再用血蓝蛋白多肽复合物加不完全弗氏佐剂加强免疫一次。采用经15 μ g/ml牛血清白蛋白多肽复合物包被的滴定板做ELISA测定兔血清中抗体的滴度。用蛋白A-Sepharose从抗体阳性的家兔血清中分离总IgG。将多肽结合于溴化氰活化的Sepharose4B柱上, 用亲和层析法从总IgG中分离抗多肽抗体。免疫沉淀法证明纯化的抗体可特异性地与人转录调控蛋白38结合。

序列表

<110> 上海博德基因开发有限公司

5 <120> 一种多肽--人转录调控蛋白38和编码这种多肽的多核苷酸

<130> 0385b11

<160> 6

10 <170> PatentIn version 3.1

<210> 1

<211> 1380

15 <212> DNA

<213> Homo sapiens

<220>

<221> CDS

20 <222> (328)..(1365)

<223>

<400> 1

ctttctgatg aacatgggaa atgcagaaaa tccttttacc ggaaagcaca cctcattcag 60

25 catcagaggc cccactcagg agagaaaact taccaatatg aggaatgtgc aaaatccttt 120

tgttcaagtt cacatcctat tcagcatcct ggaacttatg tgggattcaa actttatgaa 180

30 tgtaatgaat gtgggaaagc tttctgtcag aattcaaacc tcagtaaaca tctgagaatt 240

cacacaaaag agaaacctg tgataacaat ggctgtggga gatcttaca gtcaccctc 300

```

ataggacacc agaaaacaga tgcagag atg gaa ctc tgt ggt ggc agt gaa tat      354
                                Met Glu Leu Cys Gly Gly Ser Glu Tyr
                                1                               5
5  ggg aag aca tca cat ctc aaa gga cat cag aga att ctc atg ggg gag      402
   Gly Lys Thr Ser His Leu Lys Gly His Gln Arg Ile Leu Met Gly Glu
   10                               15                               20                               25
   aaa ccc tat gaa tgt att gaa tgt ggg aaa act ttc tcc aag aca tca      450
   Lys Pro Tyr Glu Cys Ile Glu Cys Gly Lys Thr Phe Ser Lys Thr Ser
10                               30                               35                               40
   cat ctc aga gca cat cag aga att cac aca ggt gaa aaa ccc tat gaa      498
   His Leu Arg Ala His Gln Arg Ile His Thr Gly Glu Lys Pro Tyr Glu
                               45                               50                               55
   tgt gtt gaa tgt gag aaa act ttc tct cac aag aca cac ctc agt gta      546
15  Cys Val Glu Cys Glu Lys Thr Phe Ser His Lys Thr His Leu Ser Val
                               60                               65                               70
   cat cag aga gtt cac aca ggg gag aaa ccc tat gaa tgt aat gac tgt      594
   His Gln Arg Val His Thr Gly Glu Lys Pro Tyr Glu Cys Asn Asp Cys
                               75                               80                               85
20  ggg aaa tct ttt acc tat aac tca gcc ctg aga gca cat caa aga att      642
   Gly Lys Ser Phe Thr Tyr Asn Ser Ala Leu Arg Ala His Gln Arg Ile
   90                               95                               100                               105
   cac aca ggt gag aag ccc tat gaa tgc agt gac tgt gag aaa act ttt      690
   His Thr Gly Glu Lys Pro Tyr Glu Cys Ser Asp Cys Glu Lys Thr Phe
25                               110                               115                               120
   gcc cat aat tca gcc ctc aga gca cat cat aga att cac acg ggg gag      738
   Ala His Asn Ser Ala Leu Arg Ala His His Arg Ile His Thr Gly Glu
                               125                               130                               135
   aaa cct tat gaa tgt aat gaa tgt gga agg tct ttt gcc cat att tct      786
30  Lys Pro Tyr Glu Cys Asn Glu Cys Gly Arg Ser Phe Ala His Ile Ser
                               140                               145                               150
   gtt ctc aaa gca cat caa aga att cac aca ggg gag aaa ccc tat gaa      834

```

Val Leu Lys Ala His Gln Arg Ile His Thr Gly Glu Lys Pro Tyr Glu
 155 160 165
 tgt aat gac tgt ggg aag act ttc tcc aag aca tca cat ctc aga gca 882
 Cys Asn Asp Cys Gly Lys Thr Phe Ser Lys Thr Ser His Leu Arg Ala
 5 170 175 180 185
 cat ctt aga act cgc tca ggg gag aaa ccc tat gaa tgc agt gaa tgt 930
 His Leu Arg Thr Arg Ser Gly Glu Lys Pro Tyr Glu Cys Ser Glu Cys
 190 195 200
 ggg aaa acc ttc tct gag aag tca tat gtt agt gca cat cag aga gtt 978
 10 Gly Lys Thr Phe Ser Glu Lys Ser Tyr Val Ser Ala His Gln Arg Val
 205 210 215
 cat acg ggg gag aaa ccc tac gaa tgt aat gta tgt ggg aag cca ttt 1026
 His Thr Gly Glu Lys Pro Tyr Glu Cys Asn Val Cys Gly Lys Pro Phe
 220 225 230
 15 gcc cat aat tca acc ctc aga gta cat caa aga att cac aca ggg gag 1074
 Ala His Asn Ser Thr Leu Arg Val His Gln Arg Ile His Thr Gly Glu
 235 240 245
 aaa tcc tat gaa tgt aat gat tgt ggg aaa acg ttc tcc cag aaa tca 1122
 Lys Ser Tyr Glu Cys Asn Asp Cys Gly Lys Thr Phe Ser Gln Lys Ser
 20 250 255 260 265
 cac ctt agt gca cac cag aga att cac aca ggg gag aaa ccc tat gag 1170
 His Leu Ser Ala His Gln Arg Ile His Thr Gly Glu Lys Pro Tyr Glu
 270 275 280
 tgt aat gaa tgc gga aaa gct ttt gcc caa aat tca act ctc aga gta 1218
 25 Cys Asn Glu Cys Gly Lys Ala Phe Ala Gln Asn Ser Thr Leu Arg Val
 285 290 295
 cac cag aga att cac aca ggg gag aaa ccc tat gaa tgt gat gaa tgt 1266
 His Gln Arg Ile His Thr Gly Glu Lys Pro Tyr Glu Cys Asp Glu Cys
 300 305 310
 30 ggg aaa act ttt gtc cgt aag gca gct ctt agg gta cat cac acc aga 1314
 Gly Lys Thr Phe Val Arg Lys Ala Ala Leu Arg Val His His Thr Arg
 315 320 325

atg cat acc aga gag aaa acc cta gca tgt aat gga ttt ggg aag tcc 1362

Met His Thr Arg Glu Lys Thr Leu Ala Cys Asn Gly Phe Gly Lys Ser

330 335 340 345

tga gggaatgcat acctt

5 1380

<210> 2

<211> 345

<212> PRT

10 <213> Homo sapiens

<400> 2

Met Glu Leu Cys Gly Gly Ser Glu Tyr Gly Lys Thr Ser His Leu Lys

15 1 5 10 15

Gly His Gln Arg Ile Leu Met Gly Glu Lys Pro Tyr Glu Cys Ile Glu

20 25 30

20 Cys Gly Lys Thr Phe Ser Lys Thr Ser His Leu Arg Ala His Gln Arg

35 40 45

Ile His Thr Gly Glu Lys Pro Tyr Glu Cys Val Glu Cys Glu Lys Thr

50 55 60

25

Phe Ser His Lys Thr His Leu Ser Val His Gln Arg Val His Thr Gly

65 70 75 80

Glu Lys Pro Tyr Glu Cys Asn Asp Cys Gly Lys Ser Phe Thr Tyr Asn

30 85 90 95

Ser Ala Leu Arg Ala His Gln Arg Ile His Thr Gly Glu Lys Pro Tyr

		100		105		110	
	Glu Cys Ser Asp Cys Glu Lys Thr Phe Ala His Asn Ser Ala Leu Arg						
		115		120		125	
5	Ala His His Arg Ile His Thr Gly Glu Lys Pro Tyr Glu Cys Asn Glu						
		130		135		140	
	Cys Gly Arg Ser Phe Ala His Ile Ser Val Leu Lys Ala His Gln Arg						
10	145		150		155		160
	Ile His Thr Gly Glu Lys Pro Tyr Glu Cys Asn Asp Cys Gly Lys Thr						
			165		170		175
15	Phe Ser Lys Thr Ser His Leu Arg Ala His Leu Arg Thr Arg Ser Gly						
		180		185		190	
	Glu Lys Pro Tyr Glu Cys Ser Glu Cys Gly Lys Thr Phe Ser Glu Lys						
		195		200		205	
20	Ser Tyr Val Ser Ala His Gln Arg Val His Thr Gly Glu Lys Pro Tyr						
		210		215		220	
	Glu Cys Asn Val Cys Gly Lys Pro Phe Ala His Asn Ser Thr Leu Arg						
25	225		230		235		240
	Val His Gln Arg Ile His Thr Gly Glu Lys Ser Tyr Glu Cys Asn Asp						
			245		250		255
30	Cys Gly Lys Thr Phe Ser Gln Lys Ser His Leu Ser Ala His Gln Arg						
		260		265		270	

Ile His Thr Gly Glu Lys Pro Tyr Glu Cys Asn Glu Cys Gly Lys Ala
275 280 285

Phe Ala Gln Asn Ser Thr Leu Arg Val His Gln Arg Ile His Thr Gly
5 290 295 300

Glu Lys Pro Tyr Glu Cys Asp Glu Cys Gly Lys Thr Phe Val Arg Lys
305 310 315 320

10 Ala Ala Leu Arg Val His His Thr Arg Met His Thr Arg Glu Lys Thr
325 330 335

Leu Ala Cys Asn Gly Phe Gly Lys Ser
340 345

15
<210> 3
<211> 23
<212> DNA
<213> Homo sapiens

20
<400> 3
ctttctgatg aacatgggaa atg
23

25 <210> 4
<211> 24
<212> DNA
<213> Homo sapiens

30 <400> 4
aataacaaac ttctccacat ttat
24

<210> 5
<211> 34
<212> DNA
<213> Homo sapiens

5

<400> 5
catgctagca tgggaactctg tgggtggcagt gaat
34

10 <210> 6
<211> 35
<212> DNA
<213> Homo sapiens

15 <400> 6
catggatcct caggacttcc caaatccatt acatg
35

相同性 = 207/337 (61%), 相似性 = 254/337 (75%)

5 Query: 9 YGKTSHLKGHRILMGEKPYECIECGKTFSTSHLRAHQRIHTGEKPYECVECEKTFSHK 68
 + +S L HQRI GEKPY+C +CG+TF++ + L HQR HTGEKPYEC EC K+FS +
 Sbjct: 62 FRNSSALTKHQRIHTGEKPYKCTQCGRITFNQIAPLIHQHRTHTGEKPYECSECGKSFSFR 121

10 Query: 69 THLSVHQRVHTGEKPYECNDCGKSFTYNSALRAHQRIHTGEKPYECSDCEKTFAHNSALR 128
 + S H+R HTGEKPYEC++CGK+F + L H RIHTGEKPY+C +C K F+H+S+L
 Sbjct: 122 SSFSQHERHTHTGEKPYECSECGKAFRQSIHLTQHLRIHTGEKPYQCGECGKAFSHSSSLT 181

15 Query: 129 AHHRIHTGEKPYECNECGRSFAHISVLKAHQRIHTGEKPYECNDCGKTFSTSHLRAHLR 188
 H RIHTGEKPYEC+ECG++F I+ L HQR HTGEKPYEC +CGK FS+++ L H R
 Sbjct: 182 KHQRIHTGEKPYECHECGKAFTQITPLIQHRTHTGEKPYECGECGKAQFSQSTLLTEHRR 241

20 Query: 189 TRSGEKPYECSECGKTFSEKSYVSAHQRVHTGEKPYECNVCGKPFANSTLRVHQRIHTG 248
 +GEKPY C+ECGKTFS S +S H+R HTGEKPYEC+ CGK F ++ L HQRIHTG
 Sbjct: 242 IHTGEKPYGCNECGKTFSSSSLSQHERHTHTGEKPYECSCGKAFRQSTHLTQHRIHTG 301

20 Query: 249 EKSYECNDCGKTFSSKSHLSAQRIHTGEKPYECNECGKAFAQNSTLRVHQRIHTGEKPY 308
 EK YECNDCGK FS S L+ HQRIHTGEKPYECN+CG+AF+Q + L HQRIHTGEKPY
 Sbjct: 302 EKPYECNDCGKAFSSSSSLTKHQRIHTGEKPYECNQCGRAFSQLAPLIHQRIHTGEKPY 361

25 Query: 309 ECDECGKTFVRKAALRVHHTRMHTREKTLACNGFGKS 345
 EC++CG+ R A L + H R+HT+EK CN GKS
 Sbjct: 362 ECNQCGRASAR-ATLLIEHQRIHTKEKPYGCNECGKS 397

图 1

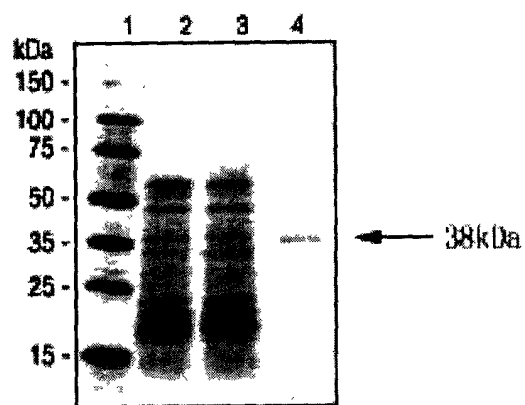


图 2