

[19]中华人民共和国国家知识产权局

[51]Int. Cl⁷

C12N 15/52

C12N 9/00 C12N 15/63

C07K 16/40 C12Q 1/25

C12Q 1/68 A61K 38/43

[12] 发明专利申请公开说明书

[21] 申请号 00116594.1

[43]公开日 2002 年 1 月 9 日

[11]公开号 CN 1330144A

[22]申请日 2000.6.19 [21]申请号 00116594.1
[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 12 层
[72]发明人 毛裕民 谢 毅

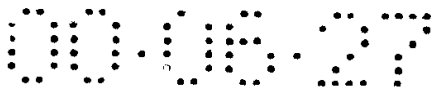
权利要求书 2 页 说明书 29 页 附图页数 1 页

[54]发明名称 一种新的多肽——人 I 类氨酰基 - tRNA 合成酶 10 和编码这种多肽的多核苷酸

[57]摘要

本发明公开了一种新的多肽——人 I 类氨酰基 - tRNA 合成酶 10, 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, 如各种肿瘤、蛋白质代谢紊乱疾病等。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 I 类氨酰基 - tRNA 合成酶 10 的多核苷酸的用途。

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



权 利 要 求 书

- 1、一种分离的多肽-人 I 类氨酰基-tRNA 合成酶 10, 其特征在于它包含有: SEQ ID NO: 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 SEQ ID NO: 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 SEQ ID NO: 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 SEQ ID NO: 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 SEQ ID NO: 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 SEQ ID NO: 1 中 214-501 位的序列或 SEQ ID NO: 1 中 1-2004 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 I 类氨酰基-tRNA 合成酶 10 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 I 类氨酰基-tRNA 合成酶 10 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 I 类氨酰基-tRNA 合成酶 10 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 I 类氨酰基-tRNA 合成酶 10 特异性结合的抗体。
- 11、一类模拟或调节多肽活性或表达的化合物, 其特征在于它们是模拟、促进、拮抗或抑制人 I 类氨酰基-tRNA 合成酶 10 的活性的化合物。
- 12、如权利要求 11 所述的化合物, 其特征在于它是 SEQ ID NO: 1 所示的多核苷酸序

列或其片段的反义序列。

13、一种权利要求 11 所述化合物的应用，其特征在于所述化合物用于调节人 I 类氨酰基-tRNA 合成酶 10 在体内、体外活性的方法。

14、一种检测与权利要求 1-3 中的任一权利要求所述多肽相关的疾病或疾病易感性的方法，其特征在于其包括检测所述多肽的表达量，或者检测所述多肽的活性，或者检测多核苷酸中引起所述多肽表达量或活性异常的核苷酸变异。

15、如权利要求 1-3 中的任一权利要求所述多肽的应用，其特征在于它应用于筛选人 I 类氨酰基-tRNA 合成酶 10 的模拟物、激动剂，拮抗剂或抑制剂；或者用于肽指纹图谱鉴定。

16、如权利要求 4-6 中的任一权利要求所述的核酸分子的应用，其特征在于它作为引物用于核酸扩增反应，或者作为探针用于杂交反应，或者用于制造基因芯片或微阵列。

17、如权利要求 1-6 及 11 中的任一权利要求所述的多肽、多核苷酸或化合物的应用，其特征在于用所述多肽、多核苷酸或其模拟物、激动剂、拮抗剂或抑制剂以安全有效剂量与药学上可接受的载体组成作为诊断或治疗与人 I 类氨酰基-tRNA 合成酶 10 异常相关的疾病的药物组合物。

18、权利要求 1-6 及 11 中的任一权利要求所述的多肽、多核苷酸或化合物的应用，其特征在于用所述多肽、多核苷酸或化合物制备用于治疗如恶性肿瘤，血液病，HIV 感染和免疫性疾病和各类炎症的药物。

说明书

一种新的多肽——人 I 类氨酰基-tRNA 合成酶 10 和编码这种多肽的多核苷酸

5 本发明属于生物技术领域，具体地说，本发明描述了一种新的多肽——人 I 类氨酰基-tRNA 合成酶 10，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。

10 在 mRNA 的翻译过程中，tRNA 起着将 mRNA 的密码子与相应的氨基酸对应起来的功能。tRNA 分子上的反密码子与其对应的氨基酸之间并无结构上的相关性，而且 tRNA 本身并不能催化相应氨基酸的荷载，将 tRNA 与相应的氨基酸结合起来的反应是由氨酰基-tRNA 合成酶催化的。

15 氨酰基 tRNA 合成酶合成氨酰基-tRNA 的反应分为两步：氨基酸首先活化成氨酰基腺苷酸，之后，氨酰基腺苷酸再与 tRNA 生成氨酰基-tRNA。氨酰基-tRNA 合成酶通过 tRNA 分子上的“副密码子”来识别不同的 tRNA，以便把正确的氨基酸与 tRNA 结合起来。

20 每一种氨基酸都对应着一种氨酰基-tRNA 合成酶，因此，20 种氨基酸就有 20 种不同的氨酰基-tRNA 合成酶。但这 20 种氨酰基-tRNA 合成酶彼此是有关系的。根据结构和序列的不同，20 种氨酰基-tRNA 合成酶可以分成两类，I 类和 II 类。I 类酶的催化部分含有 Rossmann 折叠构成的骨架，这是一种核酸结合区域；II 类酶的骨架则由一种新的反向平行折叠构成。

有些氨酰基-tRNA 合成酶的 N 端序列相当保守，特别是四肽序列 His-Ile-Gly-His ('HIGH')。此序列在精氨酸，色氨酸，谷氨酸，异亮氨酸，蛋氨酸，酪氨酸等的相应的氨酰基-tRNA 合成酶中都有发现。它们都属于 I 类合成酶。

25 其保守区域是：P-x(0,2)-[GSTAN]-[DENQGAPK]-x-[LIVMFPP]-[HT]-[LIVMYAC]-G-[HNTG]-[LIVMFYSTAGPC]

30 氨酰基-tRNA 合成酶提供合成蛋白质的可用的氨酰基-tRNA。如果某种氨酰基-tRNA 合成酶缺失，会造成对应的氨酰基-tRNA 不能正确形成，导致合成蛋白质过程的过早终止，形成没有功能的蛋白。如果某种氨酰基-tRNA 合成酶发生突变，识别另一种氨基酸，则形成的蛋白质包含不正确的氨基酸序列，如果这种变异发生在对该蛋白质的功能来说关键的氨基酸位点上，那么形成的蛋白质将不能正确折叠而没有功能，或者功能减弱或增强，或者改变了作用的底物而使其功能发生变化。这种功能上的变化常常导致疾病。

例如，在 25%—30%的人肿瘤中，发现异常的 Ras 蛋白，在 Ras 的 12 位获 61 位残基的点突变，又甘氨酸代替了缬氨酸，抑制 Ras 本身以及 GAP 刺激的 GTP 酶的活性，因此不能是 Ras-GTP 转变成 Ras-GDP，导致 Ras 长期处于活性状态，不停顿的向下游发放信号，引起细胞的异常增殖，从而导致癌症的发生。

5 氨酰基-tRNA 合成酶的突变还可能蛋白质的分拣和分泌紊乱以及膜受体的功能失常等，引起各种内分泌疾病，免疫系统疾病，神经肌肉系统疾病等。

本发明的蛋白质含有 I 类氨酰基-tRNA 合成酶的特征性保守区域，并且，通过对表达谱的分析发现，正常肝组织中的 II 类氨酰基-tRNA 合成酶表达量比肝癌组织中的表达量高，在正常喉组织和喉癌组织中也有相同的发现。综上所述，认为本发
10 明的多肽是一种新的 I 类氨酰基-tRNA 合成酶，具有相似的生物学功能，并命名为人 I 类氨酰基-tRNA 合成酶 10。

由于如上所述人 I 类氨酰基-tRNA 合成酶 10 蛋白在机体内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 I 类氨酰基-tRNA 合成酶 10 蛋白，特别是鉴定这种蛋白的氨基酸序列。
15 新人 I 类氨酰基-tRNA 合成酶 10 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/或治疗药的基础，因此分离其编码 DNA 是非常重要的。

20 本发明的一个目的是提供分离的新的多肽——人 I 类氨酰基-tRNA 合成酶 10 以及其片段、类似物和衍生物。

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 I 类氨酰基-tRNA 合成酶 10 的多核苷酸的重组载体。

25 本发明的另一个目的是提供含有编码人 I 类氨酰基-tRNA 合成酶 10 的多核苷酸的基因工程化宿主细胞。

本发明的另一个目的是提供生产人 I 类氨酰基-tRNA 合成酶 10 的方法。

本发明的另一个目的是提供针对本发明的多肽——人 I 类氨酰基-tRNA 合成酶 10 的抗体。

30 本发明的另一个目的是提供了针对本发明多肽——人 I 类氨酰基-tRNA 合成酶 10 的模拟化合物、拮抗剂、激动剂、抑制剂。

本发明的另一个目的是提供诊断治疗与人 I 类氨酰基-tRNA 合成酶 10 异常相关

的疾病的方法。

本发明涉及一种分离的多肽，该多肽是人源的，它包含：具有 SEQ ID No. 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 SEQ ID NO: 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 SEQ ID No. 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 SEQ ID NO: 1 中 214-501 位的序列；和 (b) 具有 SEQ ID NO: 1 中 1-2004 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。

15 本发明还涉及一种能与本发明多肽特异性结合的抗体。

本发明还涉及一种筛选的模拟、激活、拮抗或抑制人 I 类氨酰基-tRNA 合成酶 10 蛋白活性的化合物的方法，其包括利用本发明的多肽。本发明还涉及用该方法获得的化合物。

20 本发明还涉及一种体外检测与人 I 类氨酰基-tRNA 合成酶 10 蛋白异常表达相关的疾病或疾病易感性的方法，包括检测生物样品中所述多肽或其编码多核苷酸序列中的突变，或者检测生物样品中本发明多肽的量或生物活性。

本发明也涉及一种药物组合物，它含有本发明多肽或其模拟物、激活剂、拮抗剂或抑制剂以及药学上可接受的载体。

25 本发明还涉及本发明的多肽和/或多核苷酸在制备用于治疗癌症、发育性疾病或免疫性疾病或其它由于人 I 类氨酰基-tRNA 合成酶 10 表达异常所引起疾病的药物的用途。

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。

30 本说明书和权利要求书中使用的下列术语除非特别说明具有如下的含义：

“核酸序列”是指寡核苷酸、核苷酸或多核苷酸及其片段或部分，也可以指基因组或合成的 DNA 或 RNA，它们可以是单链或双链的，代表有义链或反义链。类似地，术

语“氨基酸序列”是指寡肽、肽、多肽或蛋白质序列及其片段或部分。当本发明中的“氨基酸序列”涉及一种天然存在的蛋白质分子的氨基酸序列时，这种“多肽”或“蛋白质”不意味着将氨基酸序列限制为与所述蛋白质分子相关的完整的天然氨基酸。

5 蛋白质或多核苷酸“变体”是指一种具有一个或多个氨基酸或核苷酸改变的氨基酸序列或编码它的多核苷酸序列。所述改变可包括氨基酸序列或核苷酸序列中氨基酸或核苷酸的缺失、插入或替换。变体可具有“保守性”改变，其中替换的氨基酸具有与原氨基酸相类似的结构或化学性质，如用亮氨酸替换异亮氨酸。变体也可具有非保守性改变，如用色氨酸替换甘氨酸。

10 “缺失”是指在氨基酸序列或核苷酸序列中一个或多个氨基酸或核苷酸的缺失。

“插入”或“添加”是指在氨基酸序列或核苷酸序列中的改变导致与天然存在的分子相比，一个或多个氨基酸或核苷酸的增加。“替换”是指由不同的氨基酸或核苷酸替换一个或多个氨基酸或核苷酸。

15 “生物活性”是指具有天然分子的结构、调控或生物化学功能的蛋白质。类似地，术语“免疫学活性”是指天然的、重组的或合成蛋白质及其片段在合适的动物或细胞中诱导特定免疫反应以及与特异性抗体结合的能力。

“激动剂”是指当与人 I 类氨酰基-tRNA 合成酶 10 结合时，一种可引起该蛋白质改变从而调节该蛋白质活性的分子。激动剂可以包括蛋白质、核酸、碳水化合物或任何其它可结合人 I 类氨酰基-tRNA 合成酶 10 的分子。

20 “拮抗剂”或“抑制物”是指当与人 I 类氨酰基-tRNA 合成酶 10 结合时，一种可封闭或调节人 I 类氨酰基-tRNA 合成酶 10 的生物学活性或免疫学活性的分子。拮抗剂和抑制物可以包括蛋白质、核酸、碳水化合物或任何其它可结合人 I 类氨酰基-tRNA 合成酶 10 的分子。

25 “调节”是指人 I 类氨酰基-tRNA 合成酶 10 的功能发生改变，包括蛋白质活性的升高或降低、结合特性的改变及人 I 类氨酰基-tRNA 合成酶 10 的任何其它生物学性质、功能或免疫性质的改变。

30 “基本上纯”是指基本上不含天然与其相关的其它蛋白、脂类、糖类或其它物质。本领域的技术人员能用标准的蛋白质纯化技术纯化人 I 类氨酰基-tRNA 合成酶 10。基本上纯的人 I 类氨酰基-tRNA 合成酶 10 在非还原性聚丙烯酰胺凝胶上能产生单一的主带。人 I 类氨酰基-tRNA 合成酶 10 多肽的纯度可用氨基酸序列分析。

“互补的”或“互补”是指在允许的盐浓度和温度条件下通过碱基配对的多核苷酸天然结合。例如，序列“C-T-G-A”可与互补的序列“G-A-C-T”结合。两个单链

分子之间的互补可以是部分的或全部的。核酸链之间的互补程度对于核酸链之间杂交的效率及强度有明显影响。

“同源性”是指互补的程度，可以是部分同源或完全同源。“部分同源”是指一种部分互补的序列，其至少可部分抑制完全互补的序列与靶核酸的杂交。这种杂交的抑制可通过在严格性程度降低的条件下进行杂交（Southern印迹或Northern印迹等）来检测。基本上同源的序列或杂交探针可竞争和抑制完全同源的序列与靶序列在的严格性程度降低的条件下的结合。这并不意味严格性程度降低的条件允许非特异性结合，因为严格性程度降低的条件要求两条序列相互的结合为特异性或选择性相互作用。

“相同性百分率”是指在两种或多种氨基酸或核酸序列比较中序列相同或相似的百分率。可用电子方法测定相同性百分率，如通过MEGALIGN程序(Lasergene software package, DNASTAR, Inc., Madison Wis.). MEGALIGN程序可根据不同的方法如Cluster法比较两种或多种序列(Higgins, D. G. 和 P.M. Sharp (1988) Gene 73: 237-244)。Cluster法通过检查所有配对之间的距离将各组序列排列成簇。然后将各簇以成对或成组分配。两个氨基酸序列如序列A和序列B之间的相同性百分率通过下式计算：

序列A与序列B之间匹配的残基个数

$100 \text{ (序列A的残基数 - 序列A中间隔残基数 - 序列B中间隔残基数)}$

也可以通过Cluster法或用本领域周知的方法如Jotun Hein 测定核酸序列之间的相同性百分率(Hein J., (1990) Methods in enzymology 183: 625-645)。

“相似性”是指氨基酸序列之间排列对比时相应位置氨基酸残基的相同或保守性取代的程度。用于保守性取代的氨基酸例如，带负电荷的氨基酸可包括天冬氨酸和谷氨酸；带正电荷的氨基酸可包括赖氨酸和精氨酸；具有不带电荷的头部基团有相似亲水性的氨基酸可包括亮氨酸、异亮氨酸和缬氨酸；甘氨酸和丙氨酸；天冬酰胺和谷氨酰胺；丝氨酸和苏氨酸；苯丙氨酸和酪氨酸。

“反义”是指与特定的DNA或RNA序列互补的核苷酸序列。“反义链”是指与“有义链”互补的核酸链。

“衍生物”是指HFP或编码其的核酸的化学修饰物。这种化学修饰物可以用烷基、酰基或氨基替换氢原子。核酸衍生物可编码保留天然分子的主要生物学特性的多肽。

“抗体”是指完整的抗体分子及其片段，如 Fa、F(ab')₂ 及Fv，其能特异性结合人I类氨酰基-tRNA合成酶10的抗原决定簇。

“分离的”一词指将物质从它原来的环境（例如，若是自然产生的就指其天然环境）之中移出。比如说，一个自然产生的多核苷酸或多肽存在于活动物中就是没有被分离出来，但同样的多核苷酸或多肽同一些或全部在自然系统中与之共存的物质分开就是分离的。这样的多核苷酸可能是某一载体的一部分，也可能这样的多核苷酸或多肽是某一组合物的一部分。既然载体或组合物不是它天然环境的成分，它们仍然是分离的。

如本发明所用,“分离的”是指物质从其原始环境中分离出来(如果是天然物质,原始环境即是天然环境)。如活体细胞内的天然状态下的多聚核苷酸和多肽是没有分离纯化的,但同样的多聚核苷酸或多肽如从天然状态中同存在的其他物质中分开,则为分离纯化的。

如本文所用,“分离的人 I 类氨酰基-tRNA 合成酶 10”是指人 I 类氨酰基-tRNA 合成酶 10 基本上不含天然与其相关的其它蛋白、脂类、糖类或其它物质。本领域的技术人员能用标准的蛋白质纯化技术纯化人 I 类氨酰基-tRNA 合成酶 10。基本上纯的多肽在非还原聚丙烯酰胺凝胶上能产生单一的主带。人 I 类氨酰基-tRNA 合成酶 10 多肽的纯度能用氨基酸序列分析。

本发明提供了一种新的多肽——人I类氨酰基-tRNA合成酶10,其基本上是由SEQ ID NO: 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。

-6-

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 SEQ ID NO: 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 SEQ ID NO: 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 2004 个碱基,其开放读框(214-501)编码了 87 个氨基酸。此多肽具有人 I 类氨酰基-tRNA 合成酶的特征序列,可推断出该人 I 类氨酰基-tRNA 合成酶 10 具有人 I 类氨酰基-tRNA 合成酶所代表的结构和功能。

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 SEQ ID NO: 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 SEQ ID NO: 2 的蛋白质或多肽,但与 SEQ ID NO: 1 所示的编码区序列有差别的核酸序列。

编码 SEQ ID NO: 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。

术语“编码多肽的多核苷酸”是指包括编码此多肽的多核苷酸和包括附加编码和/或非编码序列的多核苷酸。

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 SEQ ID NO: 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 I 类氨酰基-tRNA 合成酶 10 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。

5 本发明的编码人 I 类氨酰基-tRNA 合成酶 10 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。

10 本发明的 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
15 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。

20 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 I 类氨酰基-tRNA 合成酶 10 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。

25 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 探针的标记可用放射性同位素, 荧光素或酶(如碱性磷酸酶)等。

30 在第(4)种方法中, 检测人 I 类氨酰基-tRNA 合成酶 10 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。

5 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。

10 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 I 类氨酰基-tRNA 合成酶 10 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 I 类氨酰基-tRNA 合成酶 10 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
15 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
20 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 I 类氨酰基-tRNA 合成酶
10 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,
25 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
30 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,

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

5 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。

10 本发明中，编码人 I 类氨酰基-tRNA 合成酶 10 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。

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 I 类氨酰基-tRNA 合成酶 10 (Science, 1984; 224: 1431)。一般来说有以下步骤:

- 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。

- 10 -

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(HPLC)和其它各种液相层析技术及这些方法的结合。

本发明的多肽以及该多肽的拮抗剂、激动剂和抑制剂可直接用于疾病治疗，例如，可治疗恶性肿瘤、肾上腺缺乏症、皮肤病、各类炎症、HIV 感染和免疫性疾病等。

在 mRNA 的翻译过程中，tRNA 起着将 mRNA 的密码子与相应的氨基酸对应起来的功能。tRNA 分子上的反密码子与其对应的氨基酸之间并无结构上的相关性，而且 tRNA 本身并不能催化相应氨基酸的荷载，将 tRNA 与相应的氨基酸结合起来的反应是由氨酰基-tRNA 合成酶催化的。每一种氨基酸都对应着一种氨酰基-tRNA 合成酶，因此，20 种氨基酸就有 20 种不同的氨酰基-tRNA 合成酶。研究发现，在多数氨酰基-tRNA 合成酶中存在着一段特征性的保守序列 His-Ile-Gly-His ('HIGH')。该特异的保守序列是形成氨酰基-tRNA 合成酶其活性 motif 所必需。该特征序列的表达异常将导致氨酰基-tRNA 合成酶的功能丧失，导致蛋白质的翻译障碍，产生异常或/和无功能的蛋白质，进而引发各种肿瘤、蛋白质代谢紊乱疾病等的发生。

本发明新的多肽在结构与功能上与人氨酰基-tRNA 合成酶蛋白有高度的同源性及相似性，并且氨基酸序列中含有上述保守的特征序列模板。上述特异性的保守序列的表达异常将导致本发明的含有此 motif 的多肽的功能异常，从而导致蛋白质的翻译障碍，产生异常或/和无功能的蛋白质，进而引发各种肿瘤、蛋白质代谢紊乱疾病等的发生，这些疾病包括但不限于：

25 各种组织常见肿瘤：

一. 上皮组织：

乳头状瘤，鳞状细胞癌【皮肤、鼻咽、喉、宫颈】，腺瘤(癌)【乳腺、甲状腺】，粘/浆液性囊腺瘤(癌)【卵巢】，基底细胞癌【头面部皮肤】，(恶性)多型性腺瘤【涎腺】，乳头状瘤、移行上皮癌【膀胱、肾盂】等；

30 二. 间叶组织：

纤维(肉)瘤【四肢】，(恶性)纤维组织细胞瘤【四肢】，脂肪(肉)瘤【皮下组织、下肢、腹膜后】，平滑肌(肉)瘤【子宫和胃肠】，横纹肌(肉)瘤【头颈、生殖泌尿

道、四肢】，血管(肉)瘤、淋巴管(肉)瘤【皮肤、皮下组织、舌、唇】，骨(肉)瘤【颅骨、长骨】，(恶)性巨细胞瘤【股/胫/肱骨上端】，软骨(肉)瘤【手足短骨、盆/肋/股/肱/肩胛骨】，滑膜(肉)瘤【膝/踝/腕/肩/肘关节附近】，(恶)性间皮瘤【胸/腹膜】等；

5 三. 淋巴造血组织：

恶性淋巴瘤【颈部、纵隔、肠系膜和腹膜后淋巴结】，各种白血病【淋巴造血组织】，多发性骨髓瘤【椎/胸/肋/颅骨和长骨】等；

四. 神经组织：

10 神经纤维(肉)瘤【全身皮神经/深部神经及内脏】，(恶性)神经鞘瘤【头、颈、四肢等处神经】，(恶性)胶质细胞瘤【大脑】，髓母细胞瘤【小脑】，(恶性)脑膜瘤【脑膜】，节神经细胞瘤/神经母细胞瘤【纵隔和腹膜后/肾上腺髓质】等；

五. 其他肿瘤：

15 黑痣，恶性黑色素瘤【皮肤、粘膜】，(恶性)葡萄胎，绒毛膜上皮癌【子宫】，(恶性)支持细胞、间质细胞瘤，(恶性)颗粒细胞瘤【卵巢、睾丸】，精原细胞瘤【睾丸】，无性细胞瘤【卵巢】，胚胎性癌【睾丸、卵巢】，(恶性)畸胎瘤【卵巢、睾丸、纵隔和骶尾部】等；

蛋白质代谢紊乱可影响蛋白质下列主要生理功能，进而导致相关疾病的发生：

一. 提供机体能量，维持组织生长、更新和修补：

20 肌肉萎缩，四肢软弱，身体消瘦，严重者可呈“恶液质”表现；

二. 产生一些生理活性物质，如激素、抗体、胺类等：

1. 蛋白质肽类激素功能紊乱可导致如下疾病发生：

1) 胰岛素和胰高血糖素：糖尿病，低血糖症等；

25 2) 下丘脑和垂体激素：巨人症，侏儒症，肢端肥大症，皮质醇增多症(柯兴综合征)，原发性性醛固酮增多症，继发慢性肾上腺皮质功能减退，甲状腺机能亢进症，甲状腺机能减退症(呆小病，幼年型甲减，成年型甲减)，男/女不育症，月经失调(功能性子宫出血、闭经、多囊卵巢综合症、经前期紧张综合症、更年期综合症)，性发育障碍，尿崩症、抗利尿激素分泌不当综合症，泌乳异常等；

3) 甲状旁腺素：甲状旁腺机能亢进症，甲状旁腺机能减退症等；

30 4) 胃肠道激素：消化性溃疡，慢性消化不良，慢性胃炎等；

2. 胺类物质代谢紊乱可导致如下疾病发生：

心律失常，休克，精神错乱，癫痫，舞蹈病，肝性脑病(去甲肾上腺素、 γ -氨

3. 抗体缺陷易发生各种感染:

三. 某些蛋白质的特殊生理作用:

10 综合上述, 本发明的多肽以及该多肽的拮抗剂, 激动剂和抑制剂可直接用于多种疾病的治疗, 例如各种肿瘤、蛋白质代谢紊乱疾病等。

人 I 类氨酰基-tRNA 合成酶 10 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 I 类氨酰基-tRNA 合成酶 10 的拮抗剂可以与人 I 类氨酰基-tRNA 合成酶 10 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。

30 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 I 类氨酰基-tRNA 合成酶 10 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。

多克隆抗体的生产可用人 I 类氨酰基-tRNA 合成酶 10 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 I 类氨酰基-tRNA 合成酶 10 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 I 类氨酰基-tRNA 合成酶 10 的单链抗体。

抗人 I 类氨酰基-tRNA 合成酶 10 的抗体可用于免疫组织化学技术中，检测活检标本中的人 I 类氨酰基-tRNA 合成酶 10。

与人 I 类氨酰基-tRNA 合成酶 10 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。

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 I 类氨酰基-tRNA 合成酶 10 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 I 类氨酰基-tRNA 合成酶 10 阳性的细胞。

本发明中的抗体可用于治疗或预防与人 I 类氨酰基-tRNA 合成酶 10 相关的疾病。给予适当剂量的抗体可以刺激或阻断人 I 类氨酰基-tRNA 合成酶 10 的产生或活性。

本发明还涉及定量和定位检测人 I 类氨酰基-tRNA 合成酶 10 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 I 类氨酰基-tRNA 合成酶 10 水平，可以用作解释人 I 类氨酰基-tRNA 合成酶 10 在各种疾病中的重要性和用于诊断人 I 类氨酰基-tRNA 合成酶 10 起作用的疾病。

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。

编码人 I 类氨酰基-tRNA 合成酶 10 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 I 类氨酰基-tRNA 合成酶 10 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 I 类氨酰基-tRNA 合成酶 10，以抑制内源性的人 I 类氨酰基

5

10

15

25

30

tRNA 合成酶 10 相关的疾病。人 I 类氨酰基-tRNA 合成酶 10 突变的形式包括与正常野生型人 I 类氨酰基-tRNA 合成酶 10 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。

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)。

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。

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。

药物组合物可以以方便的方式给药，如通过局部、静脉内、腹膜内、肌内、皮下、鼻内或皮内的给药途径。人 I 类氨酰基-tRNA 合成酶 10 以有效地治疗和/或预防具体的适应症的量来给药。施用于患者的人 I 类氨酰基-tRNA 合成酶 10 的量和剂量范围将取决于许多因素，如给药方式、待治疗者的健康条件和诊断医生的判断。

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

图1是本发明人 I 类氨酰基-tRNA 合成酶 10 在 3-49 共 47 个氨基酸和结构域人 I 类氨酰基-tRNA 合成酶的氨基酸序列同源性比较图。上方序列是人 I 类氨酰基-tRNA 合成酶 10，下方序列是结构域人 I 类氨酰基-tRNA 合成酶。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 I 类氨酰基-tRNA 合成酶 10 的聚丙烯酰胺凝胶电泳图 (SDS-PAGE)。10kDa 为蛋白质的分子量。箭头所指为分离出的蛋白条带。

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限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 I 类氨酰基-tRNA 合成酶 10 的克隆

用异硫氰酸胍/酚/氯仿一步法提取人胎脑总 RNA。用 Quik mRNA Isolation Kit

10

实施例2: cDNA 克隆的结构域分析

15

20

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Qiagene的试剂盒纯化后,用下列引物进行PCR扩增:

Primer1: 5'-CATCCTGAGAACTGAAATTGATCGC-3' (SEQ ID NO: 3)

Primer2: 5'-ATAAAATTTTGAATTTATGTTCAA-3' (SEQ ID NO: 4)

25

Primer2为SEQ ID NO: 1的中的3'端反向序列。

30

DNA序列与SEQ ID NO: 1所示的1-2004bp完全相同。

实施例4: Northern 印迹法分析人I类氨酰基-tRNA合成酶10基因的表达:

用一步法提取总RNA[Anal. Biochem 1987, 162,156-159]。该法包括酸性硫氰
5 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-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
10 移至硝酸纤维素膜上。用 α - ^{32} P dATP通过随机引物法制备 ^{32} P-标记的DNA探针。所用的DNA探针为图1所示的PCR扩增的人I类氨酰基-tRNA合成酶10编码区序列 (214bp至501bp)。将 ^{32} P-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
15 中于42°C杂交过夜, 该溶液包含50%甲酰胺-25mM KH_2PO_4 (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I类氨酰基-tRNA合成酶10的体外表达、分离和纯化

根据SEQ ID NO: 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:

20 Primer3: 5'-CCCCATATGATGCTCTGTACCTTCAAAGGATGG-3' (Seq ID No: 5)

Primer4: 5'-CCCAAGCTTCTTCAACATGCCGCTTCTGTTCTTC-3' (Seq ID No: 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-
25 1325b06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-1325b06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
30 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-1325b06) 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

5

10

15

20

25

30

5

从本发明的多核苷酸 SEQ ID NO: 1 中选择寡核苷酸片段用作杂交探针, 应遵循以下原则和需要考虑的几个方面:

- 15

20

5'-TGGATGAGCAAAGAAAAATATTTTCTTGAGATGGAATATAAT-3' (SEQ ID NO: 8)

25

与以下具体实验步骤有关的其它未列出的常用试剂及其配制方法请参考文献:

30

1, 从新鲜或冰冻组织中提取 DNA

步骤：1) 将新鲜或新鲜解冻的正常肝组织放入浸在冰上并盛有磷酸盐缓冲液 (PBS) 的平皿中。用剪刀或手术刀将组织切成小块。操作中应保持组织湿润。2) 以 1000g 离心切碎组织 10 分钟。3) 用冷匀浆缓冲液 (0.25mol/L 蔗糖; 25mmol/L Tris-HCl, pH7.5; 25mmol/L NaCl; 25mmol/L MgCl₂) 悬浮沉淀 (大约 10ml/g)。4) 在 4°C 用电动匀浆器以全速匀浆组织悬液, 直至组织被完全破碎。5) 1000g 离心 10 分钟。6) 用重悬细胞沉淀 (每 0.1g 最初组织样品加 1-5ml), 再以 1000g 离心 10 分钟。7) 用裂解缓冲液重悬沉淀 (每 0.1g 最初组织样品加 1ml), 然后接以下的苯酚抽提法。

2, DNA 的苯酚抽提法

步骤：1) 用 1-10ml 冷 PBS 洗细胞, 1000g 离心 10 分钟。2) 用冷细胞裂解液重悬浮沉淀的细胞 (1×10^8 细胞/ml) 最少应用 100ul 裂解缓冲液。3) 加 SDS 至终浓度为 1%, 如果在重悬细胞之前将 SDS 直接加入到细胞沉淀中, 细胞可能会形成大的团块而难以破碎, 并降低的总产率。这一点在抽提 $>10^7$ 细胞时特别严重。4) 加蛋白酶 K 至终浓度 200ug/ml。5) 50°C 保温反应 1 小时或在 37°C 轻轻振摇过夜。6) 用等体积苯酚: 氯仿: 异戊醇 (25: 24: 1) 抽提, 在小离心机管中离心 10 分钟。两相应清楚分离, 否则重新进行离心。7) 将水相转移至新管。8) 用等体积氯仿: 异戊醇 (24: 1) 抽提, 离心 10 分钟。9) 将含 DNA 的水相转移至新管。然后进行 DNA 的纯化和乙醇沉淀。

3, DNA 的纯化和乙醇沉淀

步骤：1) 将 1/10 体积 2mol/L 醋酸钠和 2 倍体积冷 100%乙醇加到 DNA 溶液中, 混匀。在 -20°C 放置 1 小时或至过夜。2) 离心 10 分钟。3) 小心吸出或倒出乙醇。4) 用 70%冷乙醇 500ul 洗涤沉淀, 离心 5 分钟。5) 小心吸出或倒出乙醇。用 500ul 冷乙醇洗涤沉淀, 离心 5 分钟。6) 小心吸出或倒出乙醇, 然后在吸水纸上倒置使残余乙醇流尽。空气干燥 10-15 分钟, 以使表面乙醇挥发。注意不要使沉淀完全干燥, 否则较难重新溶解。7) 以小体积 TE 或水重悬 DNA 沉淀。低速涡旋振荡或用滴管吹吸, 同时逐渐增加 TE, 混合至 DNA 充分溶解, 每 $1-5 \times 10^6$ 细胞所提取的大约加 1ul。

以下第 8-13 步骤仅用于必须除去污染时, 否则可直接进行第 14 步骤。

8) 将 RNA 酶 A 加到 DNA 溶液中, 终浓度为 100ug/ml, 37°C 保温 30 分钟。9) 加入 SDS 和蛋白酶 K, 终浓度分别为 0.5%和 100ug/ml。37°C 保温 30 分钟。10) 用等体积的苯酚: 氯仿: 异戊醇 (25: 24: 1) 抽提反应液, 离心 10 分钟。11) 小心移出水相, 用等体积的氯仿: 异戊醇 (24: 1) 重新抽提, 离心 10 分钟。12) 小心移出

水相，加 1/10 体积 2mol/L 醋酸钠和 2.5 体积冷乙醇，混匀置-20℃ 1 小时。13) 用 70%乙醇及 100%乙醇洗涤沉淀，空气干燥，重悬核酸，过程同第 3-6 步骤。14) 测定 A_{260} 和 A_{280} 以检测 DNA 的纯度及产率。15) 分装后存放于-20℃。

样膜的制备:

5 1) 取 4×2 张适当大小的硝酸纤维素膜 (NC 膜)，用铅笔在其上轻轻标出点样位置及样号，每一探针需两张 NC 膜，以便在后面的实验步骤中分别用高强度条件和强度条件洗膜。

2) 吸取及对照各 15 微升，点于样膜上，在室温中晾干。

3) 置于浸润有 0.1mol/LNaOH, 1.5mol/LNaCl 的滤纸上 5 分钟 (两次)，晾干置于浸润有 0.5mol/L Tris-HCl (pH7.0), 3mol/LNaCl 的滤纸上 5 分钟 (两次)，晾干。

4) 夹于干净滤纸中，以铝箔包好，60-80℃ 真空干燥 2 小时。

探针的标记

15 1) 3 μ l Probe (0.10D/10 μ l)，加入 2 μ l Kinase 缓冲液，8-10 uCi γ - ^{32} P-dATP+2U Kinase，以补加至终体积 20 μ l。

2) 37℃ 保温 2 小时。

3) 加 1/5 体积的溴酚蓝指示剂 (BPB)。

4) 过 Sephadex G-50 柱。

5) 至有 ^{32} P-Probe 洗出前开始收集第一峰 (可用 Monitor 监测)。

20 6) 5 滴/管，收集 10-15 管。

7) 用液体闪烁仪监测同位素量

8) 合并第一峰的收集液后即所需制备的 ^{32} P-Probe (第二峰为游离 γ - ^{32} P-dATP)。

预杂交

25 将样膜置于塑料袋中，加入 3-10mg 预杂交液 (10×Denhardt's; 6×SSC, 0.1mg/ml CT DNA (小牛胸腺 DNA)。)，封好袋口后，68℃ 水浴摇 2 小时。

杂交

将塑料袋剪去一角，加入制备好的探针，封好袋口后，42℃ 水浴摇过夜。

洗膜:

高强度洗膜:

30 1) 取出已杂交好的样膜。

2) 2×SSC, 0.1%SDS 中，40℃ 洗 15 分钟 (2 次)。

3) 0.1×SSC, 0.1%SDS 中，40℃ 洗 15 分钟 (2 次)。

4) $0.1\times\text{SSC}$, $0.1\%\text{SDS}$ 中, 55°C 洗 30 分钟 (2 次), 室温晾干。

低强度洗膜:

1) 取出已杂交好的样膜。

2) $2\times\text{SSC}$, $0.1\%\text{SDS}$ 中, 37°C 洗 15 分钟 (2 次)。

5 3) $0.1\times\text{SSC}$, $0.1\%\text{SDS}$ 中, 37°C 洗 15 分钟 (2 次)。

4) $0.1\times\text{SSC}$, $0.1\%\text{SDS}$ 中, 40°C 洗 15 分钟 (2 次), 室温晾干。

X-光自显影:

-70°C , X-光自显影 (压片时间根据杂交斑放射性强弱而定)。

实验结果:

10 采用低强度洗膜条件所进行的杂交实验, 以上两个探针杂交斑放射性强弱没有明显区别; 而采用高强度洗膜条件所进行的杂交实验, 探针 1 的杂交斑放射性强程度明显强于另一个探针杂交斑的放射性强程度。因而可用探针 1 定性和定量地分析本发明的多核苷酸在不同组织中的存在和差异表达。

15

序 列 表

(1) 一般信息:

5

(ii) 发明名称: 人I类氨酰基-tRNA合成酶10及其编码序列

(iii) 序列数目: 9

(2) SEQ ID NO: 1的信息:

10

(i) 序列特征:

(A) 长度: 2004bp

(B) 类型: 核酸

(C) 链性: 双链

(D) 拓扑结构: 线性

15

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 1:

20

25

30

```
1  GGAGGAAGATGAAATATTAACATTAACAGGAGTTTGGGAAGAAGTTGATTCCAGCCCTCAT
61  GGATGACTTTGAGGGCTCAGGATGTCAGTGGAGGAAGTCCCTGCAGATGTGGTAGAAATC
121 ACAAGACAACCTAGAATTAGAATTAGGGCCTTAAGGTGAGATGAAATTGCTGTAAACTCAT
181 GATGAAACTTGAACAAATAGGGAGTTGCTTCTTATGGATGAGCAAAGAAAAATATTTTCTT
241 GAGATGGAATATAATCCAGGTGGAGATGCTATCAACATTGGTGAAATGACCACAAAGGAT
301 TTAGAATATTCCATAAACCTAGTTGATAAAGCAGCAGCAGGGTTTGAGAGGGTTTACTCC
361 AATTTTGAAGAAGTTCTACTGTTGATAAAATGCTGTCAAACAGCATCACATGCTACAGG
421 GAAATCTTTTGTGAAAGGAACAACTTCATTGTTTTAAAAAATTGACACAGGCTTCCAAC
481 CTTCAGCAGCCCCCACACTGATTAGCCAGCAGTCATCAATATAGAGGCAAGACTCTCACC
541 AGCAAAAAGATTATGACTTGGTAAGGCTCAGATATTCATTAGCATTTTTTAGCACTGACA
601 TATTTTAACCTAAGGTATGTACACAGTTTTTTAGACATGCTATTAATTACTAATTAATAA
661 TTATTAAATACTCATTAGACTACAATATAGTTTAAGCATAACTTTTATAAGCACTGGGAA
721 ATAAAATGTTTATGCAACTAACTTGATTGTAACATTTGCCTTATTGAAGTGGCCTGGAAC
781 AAAACCCACACTATCTCTGAGGTGTGCTTGTAGGTTTTTGTGTTGTAACATTTGCCTTACT
841 GCAGTTGTCTGGAACCAACCCACAATATCTCTGAGTATGCTTCTAGGTTTTTGGTTGTT
901 CTTTGTGTTTGAGATGGGGTTTCGCTCTGTGCCTAGGCCGGAGTGCGGTGGCGTGATCAT
```

961 AGCTCACTGCAGCCTCAAAGTGGTCAAGTGATTTTTCTGCCACGGCCTCCTGAGTA
 1021 GCTGGGACTGCAGGCATGCAGCACTATGCCTGGCTGTGTTTTTTAATACTTTTTTTGTA
 1081 GAGACGGGATCTCACTATGTTGCCACCTGGTCTTGAAGTCTGGGGCTCAGGCAATCCAG
 1141 CCACTTCAGCCTCCCAAAATGCAGCCATTACAGGCTTGAGCAACTGCACCCAGCTCGTCT
 5 1201 GCAGTTCTTAAAAACAAACACGGCCAGGTGCAGTGTCTGACACCTGTAATCCCAGCACTTT
 1261 GGGAGGCCAAGGCAGGTGAATCACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGG
 1321 TGAACGCAGTCTCTACTAAAAATACAAAAATTAGCCAGGCGCAGTGGCGGGCACCTGTAG
 1381 TCCCAGCTACTCGGGAGGTGAGGCAGGAGACTCGCTTGAACCCGGGAGGCGGAGGTTGC
 1441 AGTAAGCCGAGATCGTGCCACTGCACTCCTGCCTTGGCAGAGCAAGACTCCATCAAAA
 10 1501 AAAAAATGTAATTTAAAAAGATTAAAAATTAAGATAGATTTTTTACAAATGTATATATCTAC
 1561 CATTATAGTATCATACAGAATGGTTTTACTGCCGTGAAGATTCTCTGTGCTCTGTGTATC
 1621 CATCTCTTCCTTCCCCTTGGTCCATGGCAACCACTGATCTTTCTACTGATCCCATAGTTT
 1681 CACCGTTTCCAGAACGTCATATAATTGGAATCGAACAGTACATAGACTTTTTCGGATTGGC
 1741 TTCTTCTTTTAGCAATATGCATTTAAGTTTCCTCCATGTATTTTTTTGTGGCTTGATAGT
 15 1801 TCATCTCCTGTTAGCAGTGAATAATATCCCGTTGCCATTATACTTTTGTCAAGACTCAGA
 1861 ACGTACAACGTGAAGAGTAAACGTGAATGTGCACTGTAAGCTTTGAGTGTGTGGGAAAG
 1921 GGACATATGAGAAGTCTGTACTCTCTGCTCCGTTTAGCTGTGAACCTGCACTTCTCTAAA
 1981 AAATAAAATTAGTTAAAAATTTAAA

20

(3) SEQ ID NO: 2的信息:

(i) 序列特征:

(A) 长度: 95个氨基酸

(B) 类型: 氨基酸

25

(D) 拓扑结构: 线性

(ii) 分子类型: 多肽

(xi) 序列描述: SEQ ID NO: 2:

1 Met Asp Glu Gln Arg Lys Tyr Phe Leu Glu Met Glu Tyr Asn
 Pro
 30 16 Gly Gly Asp Ala Ile Asn Ile Gly Glu Met Thr Thr Lys Asp Leu
 31 Glu Tyr Ser Ile Asn Leu Val Asp Lys Ala Ala Ala Gly Phe Glu
 46 Arg Val Tyr Ser Asn Phe Glu Arg Ser Ser Thr Val Asp Lys Met

61 Leu Ser Asn Ser Ile Thr Cys Tyr Arg Glu Ile Phe Cys Glu Arg
 76 Asn Lys Leu His Cys Phe Lys Lys Leu Thr Gln Ala Ser Asn Leu
 91 Gln Gln Pro Pro His

5

(4) SEQ ID NO: 3的信息

(i) 序列特征

(A) 长度: 24碱基

(B) 类型: 核酸

10

(C) 链性: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 寡核苷酸

(xi) 序列描述: SEQ ID NO: 3:

GGAGGAAGATGAAATATTAACATT

24

15

(5) SEQ ID NO: 4的信息

(i) 序列特征

(A) 长度: 24碱基

(B) 类型: 核酸

20

(C) 链性: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 寡核苷酸

(xi) 序列描述: SEQ ID NO: 4:

TTTAAATTTTAACTAATTTTATTT

24

25

(6) SEQ ID NO: 5的信息

(i) 序列特征

(A) 长度: 33碱基

(B) 类型: 核酸

30

(C) 链性: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 寡核苷酸

(xi) 序列描述: SEQ ID NO : 5:

CATGCTAGCATGGATGAGCAAAGAAAATATTTT

33

(7) SEQ ID NO: 6的信息

5 (i) 序列特征

(A) 长度: 33碱基

(B) 类型: 核酸

(C) 链性: 单链

(D) 拓扑结构: 线性

10 (ii) 分子类型: 寡核苷酸

(xi) 序列描述: SEQ ID NO : 6:

CATGGATCCTCAGTGTGGGGCTGCTGAAGGTT

33

(8) SEQ ID NO: 7的信息:

15 (i) 序列特征:

(A) 长度: 15个氨基酸

(B) 类型: 氨基酸

(D) 拓扑结构: 线性

(ii) 分子类型: 多肽

20 (xi) 序列描述: SEQ ID NO: 7:

Met-Asp-Glu-Gln-Arg-Lys-Tyr-Phe-Leu-Glu-Met-Glu-Tyr-Asn-Pro 15

(9) SEQ ID NO: 8的信息

25 (i) 序列特征

(A) 长度: 41碱基

(B) 类型: 核酸

(C) 链性: 单链

(D) 拓扑结构: 线性

30 (ii) 分子类型: 寡核苷酸

(xi) 序列描述: SEQ ID NO : 8:

TGGATGAGCAAAGAAAATATTTTCTTGAGATGGAATATAAT

41

(10) SEQ ID NO: 9的信息

(i) 序列特征

(A) 长度: 41碱基

5 (B) 类型: 核酸

(C) 链性: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 寡核苷酸

(xi) 序列描述: SEQ ID NO : 9:

10 TGGATGAGCAAAGAAAATATCTTCTTGAGATGGAATATAAT

41

说明书附图

S 3 EQRKYFLEMEYNPGGDAINIGEMTTKDLEYSINLVDKAAAGFERVYS 49
 :::| . .: ||. |||: | : | . | :.. ||:
 5 P 11 NKKLIILSGIPCPNGSLHIGHLWALIILSDLLARYQRMQGFNVFFC 57

图 1

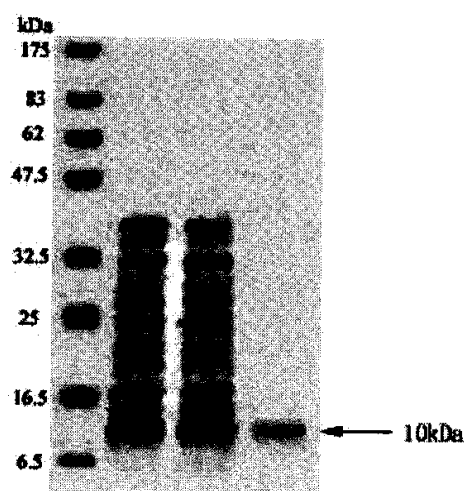


图 2