

[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 200610111067.8

[51] Int. Cl.

*C07K 14/47 (2006.01)*  
*C07K 19/00 (2006.01)*  
*C12N 15/12 (2006.01)*  
*C12N 15/62 (2006.01)*  
*C12N 15/63 (2006.01)*  
*C12N 5/00 (2006.01)*

[43] 公开日 2007年2月14日

[11] 公开号 CN 1911962A

[51] Int. Cl. (续)

*A01K 67/027 (2006.01)*  
*C07K 16/18 (2006.01)*  
*G01N 33/68 (2006.01)*  
*G01N 33/53 (2006.01)*  
*C12Q 1/68 (2006.01)*  
*A61K 38/17 (2006.01)*  
*A61K 39/00 (2006.01)*  
*A61K 48/00 (2006.01)*  
*A61P 35/00 (2006.01)*

[22] 申请日 2003.9.23

[21] 申请号 200610111067.8

分案原申请号 03823129.8

[30] 优先权

[32] 2002.9.27 [33] EP [31] 02021696.6

[32] 2002.10.3 [33] US [31] 60/415913

[71] 申请人 奥里迪斯生物医学研究及开发有限责任公司

地址 奥地利格拉茨

[72] 发明人 C·居利 C·布克  
K·扎特罗卡尔

[74] 专利代理机构 中国专利代理(香港)有限公司  
代理人 李波 刘玥

权利要求书7页 说明书135页 附图8页

[54] 发明名称

多肽与编码这些多肽的核酸和它们用于预防、  
诊断或治疗肝脏失调和上皮癌的运用

[57] 摘要

本发明涉及多肽与编码这些多肽的核酸和它们用于预防、诊断和/或治疗肝脏失调和致瘤性失调，特别是肝脏和其他上皮组织的癌、良性肝脏肿瘤如腺瘤和其他增生性肝脏失调如病灶性结节性增生症(FNH)和肝硬变的用途。本发明还涉及诊断和治疗这些失调的方法。

1. 含有根据 SEQ ID 2 的序列的分离的多肽，或者其功能变体。
2. 融合蛋白，其含有根据权利要求 1 的多肽。
3. 编码根据权利要求 1 的多肽的分离的核酸，或者其变体。
4. 根据权利要求 3 的核酸，其中该核酸是单链或双链 RNA。
5. 根据权利要求 3 的核酸，其中该核酸含有根据 SEQ ID 11 的核酸。
6. 一种载体，其含有核酸，所述核酸选自根据权利要求 3 的核酸和编码根据 SEQ ID 1 到 9 或 SEQ ID 47 的多肽的核酸。
7. 根据权利要求 6 的载体，其中该载体选自敲除基因构建体、质粒、穿梭载体、噬菌粒、粘粒、病毒载体和表达载体。
8. 含有根据权利要求 3 的核酸的细胞。
9. 含有根据权利要求 6 的载体的细胞。
10. 根据权利要求 9 的细胞，其中该细胞是转基因胚胎非人干细胞。
11. 含有根据权利要求 3 的核酸的转基因非人哺乳动物。
12. 抗体或者其抗体片段，其中该抗体针对根据权利要求 1 的多肽或者针对根据权利要求 3 的核酸。
13. 一种核酸，其含有具有与根据权利要求 3 的核酸互补的序列的核酸或者根据权利要求 3 的核酸的非功能突变变体。
14. 根据权利要求 13 的核酸，其中具有互补序列的核酸是反义分子或者 RNA 干扰分子。
15. 含有根据权利要求 13 的核酸的载体。
16. 根据权利要求 15 的载体，其中该载体选自质粒、穿梭载体、噬菌粒、粘粒、病毒载体，和表达载体。
17. 含有根据权利要求 13 的核酸的细胞。
18. 含有根据权利要求 15 的载体的细胞。
19. 一种诊断试剂，其含有至少一种化合物，所述化合物选自根据权利要求 1 的多肽、根据 SEQ ID 1 到 9 或者 SEQ ID 47 的多肽、编码前述多肽之一的核酸、前述核酸之一的变体和针对前述多肽之一的抗体或者抗体片段，与适宜的添加剂或者助剂组合或在一起。
20. 根据权利要求 19 的诊断试剂，其中核酸是探针。

21. 根据权利要求 20 的诊断试剂，其中探针是 DNA 探针。

22. 药物组合物，其含有至少一种组分，所述组分选自根据权利要求 1 的多肽、根据 SEQ ID 1 到 9 或者 SEQ ID 47 的多肽、前述多肽之一的功能变体、编码前述多肽之一的核酸、前述核酸之一的变体、为前述核酸之一的非功能突变变体的核酸、具有与前述核酸之一互补的序列的核酸、含有前述核酸之一的载体、含有前述核酸之一的细胞、含有前述载体的细胞、针对前述多肽之一的抗体或该抗体的片段、针对前述多肽之一的功能变体的抗体或者该抗体的片段、含有编码前述抗体之一的核酸的载体、含有含有编码前述抗体之一的核酸的载体的细胞和含有含有编码前述抗体片段之一的核酸的载体的细胞，与适宜的添加剂或者助剂组合或在一起。

23. 根据权利要求 22 的药物组合物，其中具有互补序列的核酸是反义分子或者 RNA 干扰分子。

24. 肝脏失调或者上皮癌的诊断方法，其中选自根据 SEQ ID 1 到 9 或者 SEQ ID 47 的序列的多肽、前述多肽之一的功能变体、编码前述多肽之一的核酸、前述核酸之一的变体、为前述核酸之一的非功能突变变体的核酸、具有与前述核酸之一互补的序列的核酸、针对前述多肽之一的抗体或该抗体的片段和针对前述多肽之一的功能变体的抗体或者该抗体的片段中的至少一种化合物在患者的样品中被鉴定并与参比文库或参比样品的至少一种化合物比较。

25. 根据权利要求 24 的方法，其中肝脏失调是选自肝硬化、酒精肝疾病、慢性肝炎、威尔逊病、血色素沉着、肝细胞癌、良性肝肿瘤和病灶性结节性增生症的失调。

26. 根据权利要求 24 的方法，其中上皮癌是选自肺、胃、肾、结肠、前列腺、皮肤和乳腺的器官的腺癌。

27. 治疗患有肝脏失调或者上皮癌的患者方法，其中选自根据 SEQ ID 1 到 9 或者 SEQ ID 47 的多肽、前述多肽之一的功能变体、编码前述多肽之一的核酸、前述核酸之一的变体、为前述核酸之一的非功能突变变体的核酸、具有与前述核酸之一互补的序列的核酸、含有前述核酸之一的载体、含有前述核酸之一的细胞、含有前述载体的细胞、针对前述多肽之一的抗体或该抗体的片段、针对前述多肽之一的功能变体的抗体或者该抗体的片段、含有编码前述抗体的核酸的载

体、含有含有编码前述抗体的核酸的载体的细胞和含有含有编码前述抗体片段的核酸的载体的细胞中的至少一种组分，与适宜的添加剂或者助剂组合或在一起，以治疗有效量被施用于需要该治疗的患者。

28. 根据权利要求 27 的治疗方法，其中具有互补序列的核酸是反义分子或者 RNA 干扰分子。

29. 根据权利要求 28 的治疗方法，其中 RNA 干扰分子以双链 RNA 或者表达双链 RNA 的载体的形式施用。

30. 根据权利要求 29 的方法，其中 RNA 干扰分子具有选自 15 到 30 个核苷酸的大小范围。

31. 根据权利要求 27 到 30 之一的方法，其中肝脏失调是选自肝硬化、酒精肝疾病、慢性肝炎、威尔逊病、血色素沉着、肝细胞癌、良性肝肿瘤和病灶性结节性增生症的失调。

32. 根据权利要求 27 到 30 之一的方法，其中上皮癌是选自肺、胃、肾、结肠、前列腺、皮肤和乳腺的器官的腺癌。

33. 在患有肝脏失调或者上皮癌的患者中刺激针对根据 SEQ ID 1 到 SEQ ID 9 或者 SEQ ID 47 的序列的多肽或者其功能变体的免疫应答的方法，其中选自根据 SEQ ID 1 到 9 或者 SEQ ID 47 的多肽、其功能变体、编码前述多肽之一的核酸、前述核酸之一的变体、含有前述核酸之一的载体、含有前述核酸之一的细胞和含有前述载体的细胞中的至少一种组分以有效刺激患者中免疫应答的量施用于需要这种治疗的患者。

34. 鉴定相对于参比文库或者参比样品，在分离自患者的样品中差别表达的根据 SEQ ID 10 到 SEQ ID 19 的至少一种核酸或者其变体的方法，该方法包括下面的步骤：

- (a) 检测从患者分离的样品中至少一种根据 SEQ ID 10 到 SEQ ID 19 的核酸，或者其变体的表达，
- (b) 将步骤 (a) 中检测的所述核酸的表达与参比文库或参比样品中所述核酸的表达比较，
- (c) 鉴定与参比文库或参比样品相比，在从患者分离的样品中差别表达的所述核酸。

35. 诊断肝脏失调或者上皮癌的方法，该方法包括下面的步骤：

- (a) 检测从患者分离的样品中至少一种根据 SEQ ID 10 到 SEQ

ID 19 的核酸，或者其变体的表达，

- (b) 将步骤 (a) 中检测的所述核酸的表达与参比文库或参比样品中所述核酸的表达比较，
- (c) 鉴定与参比文库或参比样品相比，在从患者分离的样品中差别表达的所述核酸，
- (d) 将步骤 (c) 中鉴定的所述核酸与病理学参比样品或病理学参比文库中差别表达的所述核酸匹配，

其中匹配的核酸表明该患者患有肝脏失调或上皮癌。

36. 根据权利要求 35 的方法，其中在步骤 (a) 中鉴定了至少两种核酸。

37. 根据权利要求 35 的方法，其中在步骤 (a) 中所述核酸的检测是通过基于 PCR 的检测或者通过杂交测定法。

38. 根据权利要求 35 到 37 之一的方法，其中在步骤 (b) 中通过选自基于固相的筛选方法、杂交、扣除杂交、差别展示和 RNA 酶保护测定法的方法比较所述核酸的表达。

39. 根据权利要求 35 到 38 之一的方法，其中从患者分离的样品选自肝脏组织、肝脏细胞、来自受到癌性转化的另一器官的组织、来自该器官的细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液、和粪便。

40. 根据权利要求 35 到 39 之一的方法，其中参比样品分离自选自相同患者的未患病样品和来自另一受试者的未患病样品的来源。

41. 根据权利要求 35 到 40 之一的方法，其中参比样品选自肝脏组织、肝脏细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液、和粪便。

42. 根据权利要求 35 到 41 之一的方法，其中参比文库是含有关于步骤 (a) 中所述核酸的肝脏失调特异性表达的克隆或数据的表达文库或者数据库。

43. 根据权利要求 35 到 42 之一的方法，其中病理学参比样品分离自选自来自患有肝脏失调或者上皮癌的另一患者的患病样品的来源。

44. 根据权利要求 35 到 43 的方法，其中病理学参比文库是数据库，该数据库含有关于相对于参比样品或参比文库中的对照表达，从

患有肝脏失调或者上皮癌的另一患者分离的样品中步骤(a)中所述核酸的差别表达的数据。

45. 根据权利要求 35 到 44 的方法, 其中肝脏失调是选自肝细胞癌、良性肝肿瘤和肝硬变的失调。

46. 根据权利要求 35 到 44 的方法, 其中上皮癌是选自肺、胃、肾、结肠、前列腺、皮肤和乳腺的器官的腺癌。

47. 鉴定相对于参比文库或者参比样品, 在从患者分离的样品中差别表达的至少一种根据 SEQ ID 1 到 SEQ ID 9 或 SEQ ID 47 的多肽, 或者其功能变体的方法, 该方法包括下面的步骤:

- (a) 检测从患者分离的样品中至少一种根据 SEQ ID 1 到 SEQ ID 9 或 SEQ ID 47 的多肽, 或者其功能变体的表达,
- (b) 将步骤(a)中检测的所述多肽的表达与参比文库或参比样品中所述多肽的表达比较,
- (c) 鉴定与参比文库或参比样品相比, 在从患者分离的样品中差别表达的所述多肽。

48. 肝脏失调或上皮癌的诊断方法, 该方法包括下面的步骤:

- (a) 检测从患者分离的样品中至少一种根据 SEQ ID 1 到 SEQ ID 9 或 SEQ ID 47 的多肽, 或者其变体的表达,
- (b) 将步骤(a)中检测的所述多肽的表达与参比文库或参比样品中所述多肽的表达比较,
- (c) 鉴定与参比文库或参比样品相比, 在从患者分离的样品中差别表达的所述多肽, 和
- (d) 将步骤(c)中鉴定的所述多肽与病理学参比样品或病理学参比文库中差别表达的所述多肽匹配,

其中所匹配的多肽表明该患者患有肝脏失调, 或上皮癌。

49. 根据权利要求 48 的方法, 其中鉴定了至少 2 种多肽。

50. 根据权利要求 48 或 49 的方法, 其中通过选自凝胶电泳、层析技术、免疫印迹分析、免疫组织化学、基于酶的免疫测定法、表面胞质团共振、高压液相色谱、质谱、免疫组织化学和基于酶的免疫测定法的方法检测多肽。

51. 根据权利要求 48 到 50 之一的方法, 其中通过选自二维凝胶电泳、层析分离技术、免疫印迹分析、表面胞质团共振、免疫组织化

学和基于酶的免疫测定法的方法比较多肽。

52. 根据权利要求 48 到 51 之一的方法，其中从患者分离的样品选自肝脏组织、肝细胞、来自受到癌性转化的另一器官的组织、来自该器官的细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液和粪便。

53. 根据权利要求 48 到 52 之一的方法，其中参比样品分离自选自相同患者的未患病样品和另一受试者的未患病样品的来源。

54. 根据权利要求 48 到 53 之一的方法，其中参比样品选自肝脏组织、肝细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿液、精液和粪便。

55. 根据权利要求 48 到 54 之一的方法，其中参比文库是含有关于步骤 (a) 的所述多肽的肝脏失调特异的表达的克隆或者数据的表达文库或者数据库。

56. 根据权利要求 48 到 55 的方法，其中病理学参比样品分离自选自来自患有肝脏失调和上皮癌的另一患者的患病样品的来源。

57. 根据权利要求 48 到 56 的方法，其中病理学参比文库是数据库，该数据库含有关于相对于参比样品或参比文库中的对照表达，从患有肝脏失调或上皮癌的另一患者分离的样品中步骤 (a) 的所述多肽的差别表达的数据。

58. 根据权利要求 48 到 57 的方法，其中肝脏失调是选自肝细胞癌、良性肝肿瘤、和肝硬变的失调。

59. 根据权利要求 48 到 57 之一的方法，其中上皮癌是选自肺、胃、肾、结肠、前列腺、皮肤和乳腺的器官的腺癌。

60. 预防患者发生肝脏失调或上皮癌的方法，其中选自根据 SEQ ID 1 到 9 或者 SEQ ID 47 的多肽、其功能变体、编码前述多肽之一的核酸、前述核酸之一的变体、具有与前述核酸之一互补的序列的核酸、为前述核酸之一的非功能突变变体的核酸、含有前述核酸之一的载体，或其变体、含有前述核酸之一的细胞、或其变体、和含有前述载体的细胞中的至少一种组分，以治疗有效量施用于需要这种预防性治疗的患者。

61. 鉴定药理学活性化合物的方法，该方法包括下面的步骤：

(a) 提供根据 SEQ ID 1 到 9 或 SEQ ID 47 的至少一种多肽，

或者其功能变体，

- (b) 将所述多肽与怀疑具有药理学活性的化合物接触，
- (c) 测定步骤(a)的所述多肽与怀疑具有药理学活性的所述化合物的相互作用，
- (d) 鉴定直接或间接与步骤(a)的所述多肽相互作用的怀疑具有药理学活性的所述化合物。

62. 根据权利要求61的方法，其中步骤(a)的所述多肽附着到柱子、所述多肽附着到阵列、包含在电泳凝胶中、附着到膜或者被细胞表达。

63. 根据权利要求61或62的方法，其中通过基于酶或荧光的细胞报道分子的方法测定相互作用。

64. 根据权利要求61或62的方法，其中通过表面胞质团共振、HPL、或质谱测定相互作用。

65. 根据权利要求61的方法，其中步骤(d)的直接或间接功能相互作用选自步骤(a)的所述多肽的表达的诱导、所述多肽的抑制、所述多肽的功能的激活和所述多肽的功能的抑制。



## 多肽与编码这些多肽的核酸 和它们用于预防、诊断或治疗肝脏失调和上皮癌的用途

本申请是申请日为2003年9月23日的中国专利申请98807702.7“多肽与编码这些多肽的核酸和它们用于预防、诊断或治疗肝脏失调和上皮癌的用途”的分案申请。

### 技术领域

本发明涉及多肽与编码这些多肽的核酸和它们用于预防、诊断和/或治疗肝脏失调和致瘤性失调，特别是肝脏和其他上皮组织的癌、良性肝脏肿瘤如腺瘤和其他增生性肝脏失调如病灶性结节性增生症(FNH)和肝硬变的用途。本发明还涉及诊断和治疗这些失调的方法。

### 背景技术

癌的发展的一般特征是遗传突变，该突变改变重要的细胞途径包括，例如，增殖、凋亡(细胞死亡)、对胁迫的应答和上皮/基质相互作用的活性。日益认识到在癌中去调节的核酸的鉴定可以提供对肿瘤转化机理的重要的新的认识。前癌阶段(如肝癌中巨大再生性瘤状体和“大”和“小”细胞改变)中去调节核酸的鉴定可提供对恶性转化中早期事件的理解。类似地，以组织增殖和重塑为特征的失调(如肝脏中FNH和肝硬变)中去调节基因表达的鉴定可以区分参与增殖和恶性转化的核酸。这些去调节的核酸和被编码的基因产物一起具有作为癌的新诊断标记的潜力。此外，这些去调节核酸的产物本身是人类患者中这些失调的预防和/或治疗中治疗性干预靶标。

肝脏在蛋白质、脂质、糖类、核酸和维生素代谢中起重要作用。有许多影响肝脏而不能被有效诊断、预防或治疗的失调，如肝细胞癌(HCC)。HCC的检查尤其好地适于癌中去调节的基因表达的鉴定。这是因为HCC的组织样品可以从手术切除的肿瘤得到并且该肿瘤是几乎没有基质组织的被良好局限的实体结构。此外，如上面指出的，可能比较分析良性和恶性肿瘤以及肝硬变——一种非瘤性状态。如果本领域中与肝脏失调相关的差别表达基因鉴定的局限性可以被克服，那么该比较方法可以使得能够鉴定与成熟器官中细胞增殖和组织重塑过程

(例如, 肝硬变)特别相关的去调节的核酸以及鉴定和分辨与增生(如 FNH)和良性和恶性瘤(例如, 腺瘤和 HCC)相关的基因表达改变。在 HCC 中对于新的和更好的诊断和治疗能力有迫切的需求。肝癌中去调节的基因可以与胃肠道的其他癌高度相关并且实际上与其他瘤(上皮起源的癌)相关, 因为这些组织具有共同的胚胎学起源。

在全球基础上, 肝细胞癌(HCC)属于最常见的恶性肿瘤, 导致每年约1百万例死亡(Ishak 等人, 1999. 肿瘤病理学图表册(Atlas of Tumor Pathology). 31 卷. Armed Forces Institute of Pathology, 华盛顿, DC)。

瘤性肝脏失调如 HCC 和许多其他肿瘤的明确诊断依赖于活检标本的组织病理学评估。一般不采取侵入性手术方法直到症状出现并且那时该疾病通常处于晚期, 从而限制了治疗性干预选择。从而, 需要改进诊断学和诊断方法。此外, 早诊断是关键的是受到晚发作或者甚至缺乏特定临床症状的妨碍。诊断时, 大多数 HCC 肿瘤不再顺从手术切除(除了被包裹的肿瘤或者羽层状变体)(Chen 和 Jeng, 1997, J. Gastroenterol. Hepatol. 12: 329-34); 而且, 它们对细胞抑制疗法具有高度抗性(Kawata 等人, 2001 Br. J. Cancer 84: 886-91)。总之, 诊断后一年内通常发生死亡。从而, 在该领域中高度希望 HCC 的早检测标记、预防指示剂, 和有效预防和/或治疗方案。

相反, 不像结直肠癌中充分研究的状况, 肝脏腺瘤不可能代表 HCC 的前体病变。类似地, 虽然肝硬变和肝炎病毒感染是 HCC 的明显危险因素, 但是这些病症不是发生 HCC 的前提。某些肝病如巨大再生性小结节性增生可能代表 HCC 前期阶段, 但是这还没有被证实(Shortell 和 Schwartz, 1991, Surg Gynecol Obstet. 173: 426-31; Anthony, P. in MacSween 等人, 编者 Pathology of the Liver. 2001, Churchill Livingstone, Edinburgh)。尽管通过肝脏切除术和肝脏活检的组织病理学研究可以诊断这些失调, 但是没有有效方法可以更早或者非侵入性检测这些状态。再一次迫切需要这些瘤的诊断和预后标记和这些失调的非侵入性检测。

在过去的十年中, 一些技术已经使得可能同时监视细胞内许多转录物的表达水平(见, 例如, Schena 等人, 1995, Science 270: 467-470; Lockhart 等人, 1996, Nature Biotechnology 14: 1675-1680; Blanchard

等人, 1996, *Nature Biotechnology* 14, 1649; 1996, US 5.569.588)。转录物阵列技术已经被用于鉴定在各种失调状态中被上调或下调的基因。一些最近的研究已经利用该技术检查 HCC 中基因表达的变化。这些研究已经不同地揭示了相对于对照 HCC 细胞系和 HCC 组织中编码肝脏特定蛋白质的基因的去调节(例如, 过表达和表达不足)。此外, 研究揭示了细胞周期控制、胁迫应答、凋亡、脂质代谢、细胞-细胞-相互作用、DNA 修复和细胞因子和生长因子产生所必需的基因(Graveel 等人, 2001, *Oncogene* 20: 2704-12; Kawai 等人, 2001, *Hepatology* 33: 676-91; Lau 等人, 2000, *Oncol. Res.* 12: 59-69; Nagai 等人, 1998, *Cancer* 82: 454-61; Okabe 等人, 2001, *Cancer Res* 61: 2129-37; Salwcci 等人, 1999, *Oncogene* 18: 181-187; Shirota 等人, 2001, *Hepatology* 33: 832-40; Tackels-Horne 等人, 2001, *Cancer* 92: 395-405; Wu 等人, 2001, *Oncogene* 20: 2674-3682; Xu 等人, 2001, *Cancer Res.* 61: 3176-81)。然而, 在这些研究中报导的基因表达模式中没有一致性, 这可能是由于实验设计中的不同和/或由于 HCC 组织本身的异质性。而且, 这些 HCC 的病原学是一个重要的因素: 慢性乙肝和丙肝病毒感染是 HCC 的主要原因但是也认识到酒精和慢性肝代谢失调导致 HCC 并且造成来自这些不同病原学的肿瘤发展的机理可能不同。合起来看, 直到现在还没有开发出满意的诊断学和诊断方法以便能够干预肝脏失调。

同样的情况应用于肝脏失调和上皮癌的疗法。例如, 对于 HCC, 除了切除和移植, 没有有效的治疗选择, 但是切除和移植方法仅仅可应用于 HCC 的早期, 受到得到供体肝脏的限制, 并且与患者的严重风险相关。此外, 这些方法极端昂贵。这些癌对化疗应答极弱, 最可能是由于解毒和输出有害化合物的正常肝功能。一些其他治疗选择, 如化疗栓塞(chemoembolization)、冷冻疗法和乙醇注射仍然处于实验阶段并且还没有确立这些方法的功效。手术干预仍然是最好的治疗选择但是不可能精确限定肿瘤的程度。因此, 该侵入性方法从治疗观点看不是最佳的。此外, 特定肝脏功能异常的早期诊断学的缺乏最常导致疾病的进一步发展, 其进一步混淆治疗选择并显著地增加这些疾病的患者死亡率(Jansen P. L., 1999, *Neth. J. Med.* 55: 287-292)。从而, 直到现在还没有开发出令人满意的疗法以便能够干预肝脏失

调，和其他上皮癌。此外，在本领域状态中，通过差别基因表达揭示的肝脏失调的不同亚型如 HCC 前体病变、良性肝脏瘤，和代谢性肝脏疾病如酒精肝脏疾病和肝硬变的识别还没有被公开。在表 1 中提供了本发明中评定的一些失调的关键疾病特征的概要。

表 1: 疾病特征

疾病	细胞增殖	组织重塑	克隆细胞扩张	瘤	转化/恶性潜力
肝硬变	+	+			
FNH	+	+	+/-		
腺瘤	+	+	+	+	
HCC	+	+	+	+	+

### 发明内容

本发明涉及多肽与编码这些多肽的核酸和它们用于预防、诊断和/或治疗肝脏失调，特别是肝细胞癌（HCC），和上皮癌、前癌性肝脏病变、良性瘤如腺瘤，和其他增生性肝脏失调如病灶性结节性增生症（FNH）和肝硬变的用途，这些预防、诊断和/或治疗克服了本领域中的局限性。本发明还涉及含有这些核酸的载体和细胞，还涉及针对所述多肽和核酸的抗体或抗体片段。

本发明还涉及诊断和治疗这些失调的方法。多种失调的评定具有重叠但是明显的形态学和临床特征，该评定为鉴定和辨别根据本发明的这些失调和最后关于这些失调的新的治疗策略提供了新的信息。

### 发明详述

本发明中使用的一种独特方法利用了不同的、病理学家证实的肝癌病理，其用于产生与非-癌性人肝脏库相比富含在 HCC 中被特异上调或下调的基因的疾病特异的 cDNA 文库。该文库在基因组范围内代表 HCC 中去调节的基因表达并因此包括所有潜在的 HCC 去调节基因。用来自许多额外的不同的、病理学家证实的肝癌样品（HCCs）和非恶性肝病变的被标记表达的核酸与这些文库克隆反复杂交指出在 HCC 中被高度去调节的核酸。令人惊奇的发现是该方法提供了以前没有被鉴定的去调节核酸以及以前不与 HCC 相关的许多去调节的核酸，这些核

酸上升的表达也可以与其他癌相关。这些 HCC 去调节的基因和蛋白质是本发明的主题。

筛选和检定策略由于参数的精细且限定的选择而自身具有创造性。根据本发明的差别表达的基因的鉴定依赖于组织病理学上可区分的肝脏疾病组织，其用于人肝脏疾病中基因表达变化的比较。也通过诊断确定用于实验的非-患病的参比肝脏样品。

通过肝脏失调、肝癌和/或上皮癌的一种诊断方法解决了本发明的目的，其中选自根据 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 (表 2) 的序列的多肽、其功能变体、编码前述多肽之一的核酸、前述核酸之一的变体、针对前述多肽或其变体之一的抗体或抗体片段的至少一种化合物在患者的样品中被鉴定并且与参比文库或参比样品的至少一种化合物相比较。

还通过治疗患有肝脏失调或上皮癌的患者的一种方法解决了本发明的目标，其中选自根据 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 的多肽、前述多肽之一的功能变体、编码前述多肽之一的核酸，或者其功能变体、前述核酸之一的变体、为前述核酸之一的非功能突变变体的核酸、具有与前述核酸之一互补的序列的核酸、含有前述核酸之一的载体、含有前述核酸之一的细胞、含有前述载体的细胞、针对前述多肽或其功能变体之一的抗体或者前述抗体片段、含有编码前述抗体之一的核酸的载体、含有编码前述抗体片段之一的核酸的载体、含有含有编码前述抗体之一的核酸的载体的细胞和含有含有编码前述抗体片段之一的核酸的载体的细胞中的至少一种组分被以治疗有效量施用于需要这种治疗的患者。

本发明的另一方面提供了药物组合物，其含有选自根据本发明的多肽、其功能变体、编码该多肽的核酸、前述核酸之一的变体、为前述核酸之一的非功能突变变体的核酸、具有与前述核酸之一互补的序列的核酸、含有前述核酸之一的载体、含有前述核酸之一的细胞、含有前述载体的细胞、针对前述多肽之一的抗体、针对前述多肽之一的功能变体的抗体、前述抗体之一的片段、含有编码前述抗体之一的核酸的载体、含有编码前述抗体片段之一的核酸的载体、含有含有编码前述抗体之一的核酸的载体的细胞和含有含有编码前述抗体片段之一的核酸的载体的细胞中的至少一种组分，该组合物被以治疗有效量施

用于需要这种治疗的患者。

根据本发明的多肽的记录号和它们的 cDNAs 在表 2 中显示。

表 2: 来自 GenBank 数据库的核酸和多肽与它们各自的 SEQ ID 号和记录号。

分子	多肽 (SEQ ID)	记录号	DNA (SEQ ID)	记录号
OBc11	1	NP_443111	10	AL833272
OBc15	2	新的	11	新的
IK2	3	NP_079436	12	NM_025160
IK5	4	NP_006398	13	NM_006407
DAP3	5	NP_387506	14	NM_033657
LOC5	6	NP_060917	15	NM_018447
SEC14L2	7	NP_036561	16	NM_012429
SSP29	8	NP_006392	17	NM_006401
HS16	9	NP_057223	18	NM_016139
IK3	47	NP_131462	19	AL049338

通过 RT-PCR 分析已经表明根据本发明的这些核酸和多肽的一个子集在上皮起源的其他癌中并优选不在相应的正常人组织中被特异表达或去调节。这些核酸优选包括 SEQ ID NO: 11 到 16 和 19 (在表 6 和图 3 中提供)。肝癌中的去调节核酸优选与胃肠道的其他癌高度相关, 因为这些组织具有共同的胚胎学起源。因此, 这些核酸和所编码的多肽可以优选被类似地用于诊断的诊断学方法、药物组合物和预防和/或治疗这些上皮癌的方法。

根据本发明的多肽和核酸的共同之处是与参比样品相比它们在从患有根据本发明的失调的患者中分离的样品中差别表达。根据本发明的多肽和核酸的调节是病理过程必需的并且从而与根据本发明的失调的诊断、预防和/或治疗有直接或间接的关系。根据本发明的多肽和核酸不属于直到现在所知道的靶标从而从本发明得到令人惊奇的并且全新的诊断和治疗方法。

通常, 组织中差别表达的基因的分析比细胞培养系统的分析较少

可能导致人为假阳性克隆形式的误差。除了现有的细胞培养系统不能充分模拟组织中的病理学过程的复杂性这一事实外，培养环境中细胞行为的变化也导致核酸和多肽表达模式与实际的病理状态可疑的联系。通过利用正常的和患病人组织中的基因表达，这些问题可能较不显著，但是多变量再次混淆了直接与疾病相关的差别基因表达的清楚鉴定。例如，差别表达的核酸可以来自个体间差异、代谢状态和/或临床治疗范例。此外，使用 cDNA 微阵列的大规模基因表达研究不能指出基因表达中变化的细胞来源。此外，包括所有或多数基因的差别基因表达研究产生非常大量的数据，它们使关键疾病-相关的基因表达变化的鉴定变得混淆不清。因此，包括来自仅仅一般定义（例如，肝脏肿瘤的）的肝脏失调的组织的基因表达的大规模描绘（profiling）的方法不可能阐明与该疾病过程相关的关键基因，并且就是这些关键基因代表诊断学和治疗干预的最好靶标。

由于这些困难，所以筛选的成功严重依赖于实验参数的选择。尽管所用的方法基于成熟的步骤，但是筛选和证实策略由于参数的精细和限定的选择而本身已经具有创造性。本发明中使用的一种独特方法利用了不同的、病理学家证实的肝癌病理，其用于产生与非癌性人肝脏库相比富含在 HCC 中被特异上调或下调的基因的疾病特异的 cDNA 文库。用于实验的非患病参比肝样品也通过诊断确定并合并三个独立的样品以减少个体间变化导致的假阳性的检测。在参比肝脏库和患病肝脏（即，HCC）中通常以类似水平表达的核酸通过产生扣除抑制杂交（subtractive suppressive hybridization）（SSH）cDNA 文库（Diatchenko 等人，1996，Proc. Natl. Acad. Sci. USA 93: 6025-6030）而被除去。这些 cDNA 高度富含在 HCC 中被上调和下调的核酸但是不代表未差别表达的那些核酸。数千 SSH 克隆的每一种都通过聚合酶链式反应（PCR）扩增并被固定到定制的 cDNA 微阵列中的载玻片上。来自额外的病理学家证实的肝脏失调的 RNA 被转化成荧光标记的 cDNA，其用于与微阵列上合并的未患病肝脏 RNA 竞争性杂交。所得杂交强度的比例揭示了肝脏失调中被特异下调的核酸。除了提供高度富含差别表达的基因的候选 cDNA 库之外，SSH 文库还在基因组范围规模内代表大多数（如果不是全部）差别表达的基因，这些基因的克隆比标准 cDNA 文库中的要少得多。该特征因此集中于在疾病中被特异下

调的核酸。本发明中产生的 SSH 文库包括来自基本不在正常肝脏中表达并因此不在常规 cDNA 文库中或者基因组规模的 cDNA 微阵列上提供的核酸的 cDNA 克隆。

通过序列特异的定量 RT-PCR (Q-PCR) 对 RNA 水平的独立分析, 证实与正常肝脏相比肝脏失调组织中根据本发明的序列的过表达 (图 2)。在这些验证实验中, 相应于根据本发明的序列的细胞 RNA 水平的 PCR 产物通过特定 PCR 产物的荧光检测来监视。通过 TaqMan 方法中的序列特异的水解探针寡核苷酸 (引物) 或者通过荧光双链 DNA 结合染料如 SYBR 绿提供荧光信号。通过与“管家”基因包括甘油醛脱氢酶 (GAPDH) 和  $\beta$ -肌动蛋白的水平比较将相应于本发明序列的 PCR 产物的水平关于实验变异性标准化, “管家”基因被认为在疾病或下面的实验操作中是相对不变的。这些 Q-PCR 步骤也用于测量实验条件中基因表达的水平例如当根据本发明的一种序列的水平在实验上随着小干扰 RNA 寡核苷酸降低的情况 (图 6, 表 10)。用于 TaqMan Q-PCR 分析的参比基因引物是 GAPDH-p1, SEQ ID 56; GAPDH-p2, SEQ ID 57; GAPDH-p3, SEQ ID 58; bActin-p1, SEQ ID 59; bActin-p2, SEQ ID 60; 和 bActin-p3, SEQ ID 61。用于 SYBR 绿分析的参比基因引物是 bActin-p4, SEQ ID 62; 和 bActin-p5, SEQ ID 63。Q-PCR 实验中相对于这些管家基因的 RNA 水平的确定, 根据 Pffaf1 的方法 (Nucleic Acids Research (2001) 5 月 1 日, 29 (9): e45) 进行。这些技术是本领域中技术人员熟知的。

此外, 根据本发明的 HCC 去调基因的表达与静止细胞的 8 小时和 12 小时血清刺激后肝细胞瘤细胞 (Hep3B) 的增殖有关 (图 8)。该发现支持如下提议: 根据本发明的序列的过表达对于增殖性肝脏失调如肝癌在功能上是重要的。

与本领域状态相比, 这些多肽和核酸令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断。根据本发明的核酸和多肽可用于肝脏失调和上皮癌的诊断、预防和治疗。

本发明涉及含有根据 SEQ ID 2 的序列的多肽, 或者其功能变体。本发明还涉及编码该多肽, 或者其功能变体的核酸, 尤其涉及根据 SEQ ID 11 的核酸和其变体。



在优选的实施方案中，该多肽由根据 SEQ ID 2 的序列组成。在另一优选的实施方案中，该核酸由根据 SEQ ID 11 的序列组成。

与本领域状态相比，这些多肽和核酸令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性诊断。

在本发明的另一方面，本发明涉及根据 SEQ ID 1 到 9 和/或 SEQ ID 47 的多肽、该多肽的功能变体、编码前述多肽之一的核酸、编码功能变体的核酸、前述核酸之一的变体、为前述核酸之一的非功能突变变体的核酸、具有与前述核酸之一互补的序列的核酸、含有前述核酸之一的载体、含有前述核酸之一的细胞、含有前述载体的细胞、针对前述多肽之一的抗体、针对前述多肽之一的功能变体的抗体、前述抗体之一的片段、含有编码前述抗体之一的核酸的载体、含有编码前述抗体片段之一的核酸的载体、含有含有编码前述抗体之一的核酸的载体的细胞、和/或含有含有编码前述抗体片段之一的核酸的载体的细胞的至少一种用于根据本发明的失调的诊断、预防和/或治疗。本发明的其他实施方案在下面详述。

当与肝脏失调，和/或上皮癌的治疗的本领域状况相比时，这些组分的使用令人惊奇地提供了对根据本发明的失调的改良的、持续的和/或更有效的诊断、预防和/或治疗。

术语“多肽”指根据本发明的多肽的全长。在优选的实施方案中，术语“多肽”还包括分离的多肽和通过重组方法制备的多肽，例如，通过从样品分离和纯化，通过筛选文库和通过常规方法进行蛋白质合成，所有这些方法对于本领域技术人员是公知的。优选地，可以合成全长多肽或其部分，例如，通过常规合成如 Merrifield 技术合成。在另一优选的实施方案中，根据本发明的多肽的部分可用于得到抗血清或特异单克隆抗体，它们可用于筛选合适的基因文库，所制备的基因文库用于表达编码的蛋白质序列以鉴定根据本发明的多肽的其他功能变体。

术语“根据本发明的多肽”指根据 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 的多肽（表 2）。

术语多肽的“功能变体”在本发明的意义内指与具有根据 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 之一的氨基酸序列的多肽具有约 70%、优选约 80%，尤其约 90%，特别约 95%，最优选 98% 的序列同源性，尤

其是序列同一性的多肽。这些功能变体为，例如，与根据本发明的多肽同源的多肽，它们来自非人生物，优选来自非人哺乳动物如，例如，小鼠、大鼠、猴子和猪。功能变体的其他实例为不同个体中、生物的不同器官中或者不同发育阶段中不同等位基因编码的多肽。功能变体，例如，还包括被从非肝脏组织，例如，胚胎组织，分离的核酸编码，但是在与肝脏失调相关的细胞中表达后具有指定的功能的多肽。功能变体优选还包括天然发生的或合成的突变，尤其是定量地改变这些序列编码的肽的活性的突变。此外，这些变体可以优选来自编码基因的差别剪接。

“功能变体”指与根据本发明的相应多肽具有基本相同的生物学功能的多肽。这种生物学功能可以在功能测定法中测定。

为了检验候选多肽是否是根据本发明的多肽的功能变体，候选多肽可以在本领域技术人员公知的功能测定法中分析，该测定法适于测定根据本发明的相应多肽的生物学功能。这种功能测定法包括例如细胞培养系统；其中基因被缺失（“敲除”）的小鼠或对于编码候选多肽的基因是转基因的小鼠；酶测定法等。如果候选多肽展示或者直接干扰与根据本发明的相应多肽基本相同的生物学功能，那么该候选多肽是相应多肽的功能变体，条件是候选多肽满足上面提到的%序列同一性水平的要求。

此外，术语“功能变体”包括相对于参比样品或者参比文库，在患有肝脏失调，或者其他上皮癌的患者中优选差别表达的多肽，包括从突变基因或者从差别剪接的基因表达的多肽，条件是候选功能变异多肽在%序列同一性水平上满足功能变体的标准。这种表达分析可以通过本领域中技术人员公知的方法实施。

多肽的“功能变体”还可以是根据本发明的多肽的部分，其长度至少为约7到约1000个氨基酸，优选至少10个氨基酸，更优选至少20个，最优选至少50个，例如至少100个，例如至少200个，例如至少300个，例如至少400个，例如至少500个，例如至少600个氨基酸，条件是它们具有与根据本发明的相应多肽基本相同的生物学功能。还包括根据本发明的多肽的缺失，缺失的范围为约1-30，优选约1-15，尤其约1-5个氨基酸，条件是它们具有与根据本发明的相应多肽基本相同的生物学功能。例如，可以缺少第一个氨基酸甲硫氨酸而

不显著改变该多肽的功能。而且，在变体中可以存在翻译后修饰，例如，脂质锚定或者磷酸基的存在或缺失。

“序列同一性”指两条序列的同一性程度(%同一性)，对于多肽其可以通过例如 BLASTP 2.0.1 确定，对于核酸可以通过例如 BLASTN 2.014 确定，其中 Filter 设为关并且 BLOSUM 为 62 (Altschul 等人, 1997, *Nucleic Acids Res.*, 25: 3389-3402)。

“序列同源性”指通过例如 BLASTP 2.0.1 确定的两条多肽序列的相似性(%阳性)，BLASTP 2.0.1 中 Filter 设为关并且 BLOSUM 为 62 (Altschul 等人, 1997, *Nucleic Acids Res.*, 25: 3389-3402)。

术语“肝脏失调”表示并包括优选影响肝脏的解剖学、生理学、代谢的和/或遗传活动，优选影响新肝脏细胞的产生，和/或肝脏的再生的所有类型的失调，这些失调作为整体或者其部分优选短暂地、暂时地、长期地或永久地处于病理途径中。优选还包括遗传的肝脏失调和癌性肝脏失调。肝脏失调被进一步理解为优选包括损伤、中毒、尤其是酒精、药物或食物中毒、辐射、感染、胆汁郁积、免疫反应，和遗传的代谢肝脏疾病导致的肝脏失调。肝脏失调的优选实例包括肝硬化、酒精肝疾病、慢性肝炎、威尔逊病和血色素沉着。优选还包括自身免疫失调，其中自身免疫应答针对至少一种根据本发明的多肽。在本发明的意思内，术语“肝脏失调”优选还包括肝癌，例如，肝细胞癌(HCC)、良性肝肿瘤如腺瘤和/或 FNH。HCC 优选还包括所提到的失调的亚型，包括以细胞内蛋白质内含体为特征的肝癌、以肝细胞脂肪变性为特征的 HCC 和纤维层的 HCC。例如，癌前病变优选还包括如以肝细胞增大(“大细胞”改变)为特征的病变，和以肝细胞减小(“小细胞”改变)以及巨大再生(增生的)的小结为特征的病变 (Anthony, P. in MacSween 等人, 编辑 *Pathology of the Liver*. 2001, Churchill Livingstone, Edinburgh)。

术语“上皮癌”在本发明意思内包括除了肝脏之外的任何器官，优选肺、胃、肾、结肠、前列腺、皮肤和乳腺的腺癌，并且指这些器官的失调，在这些器官中组织的上皮细胞组分被转化，导致根据本领域技术人员公知的标准诊断方法鉴定的恶性肿瘤。

在本发明的意思内，术语“根据本发明的失调”包括如上定义的上皮癌和肝脏失调。

对于多肽，术语“多肽的差别表达”指与参比样品或者参比文库中多肽的表达相比，来自患者的分离样品中多肽的表达的相对水平。该表达可以通过本领域技术人员公知的方法确定。这些方法的实例包括用对多肽特异的抗体对该多肽进行的免疫组织化学或免疫印迹或ELISA检测。优选还包括通过遗传操作以标记多肽并在模型体系中检测进行该多肽的检测，如通过在用于在模型系统中表达的转基因中标记多肽。

术语“样品”指生物材料，其包括肝脏组织或肝细胞，优选来自受到恶性转化的另一器官的组织或者来自该器官的细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液或粪便。

通过包括侵入性或非侵入性方法的方法可以从患者或者另一受试者分离样品。侵入性方法是技术人员公知的并且包括例如通过穿刺分离样品、从打开的身体或者通过内窥镜仪器手术去除样品。最小的侵入性和非侵入性方法也是本领域技术人员公知的并且包括例如，收集体液如血液、血清、血浆、腹水、脑膜液和脑脊液、唾液、尿、精液和粪便。非侵入性方法优选不需要穿刺或者通过不同于天然存在的身体开口如口、耳、鼻、直肠、尿道和创口的开口打开患者或者受试者身体。

术语“最小侵入性”方法指为了得到患者样品材料本领域中特别是技术人员公知的方法，该方法不需要麻醉，可以在医务室或者诊所中常规地完成并且或者不疼痛或者仅仅微痛。最小侵入性方法的最常见实例是静脉穿刺(venupuncture)。

术语“参比样品”指用以评价从患者中分离的给定样品中根据本发明的核酸和/或多肽的差别表达作为适宜对照的样品；这种适宜参比样品的选择是本领域中技术人员公知的。参比样品的实例包括从相同患者或者另一受试者的非患病器官或组织或细胞分离的样品，其中非患病器官或组织或细胞选自肝脏组织或肝脏细胞、血液，或者上面描述的样品。为了与从肝脏失调患者分离的样品中的表达比较，参比样品还可以包括从不同患者的非患病器官或者组织或者细胞分离的样品，其中肝脏失调的组织或细胞选自上面列出的样品组。此外，参比可以包括来自健康供体，优选与患者的年龄和性别匹配的健康供体的样品。

术语“参比文库”指展示所表达基因的克隆的文库，该文库优选从非患病肝组织或细胞制备。参比还可以来自非患病肝脏组织或细胞的 mRNA 并且可以包括含有关于核酸的非患病组织表达数据的数据库。为了比较从具有失调的肝脏患者分离的样品中根据本发明的核酸或多肽的表达，参比文库可以含有从肝脏失调的患病肝脏组织或者细胞制备的表达文库和含有关于核酸的肝脏失调的特异表达数据的数据库。

术语“患者”在本发明意思内包括死亡的或活的动物，优选哺乳动物，和人。患者或者患有肝脏失调，和/或其他上皮癌，或者受肝脏失调和/或其他上皮癌的分析、预防措施、治疗和/或诊断的影响。

术语“受试者”在本发明范围内包括，死亡的或活的动物，优选哺乳动物，和人，它们不患有肝脏失调和/或其他上皮癌并从而代表用于确定患者中根据本发明的核酸和/或多肽的差别表达的优选的合适对照。

术语“有效治疗”在本发明意思内指优选治愈患者的至少一种根据本发明的失调和/或关于与该失调相关的至少一种症状，优选 3 种症状、更优选 5 种症状、最优选与该失调相关的 10 种症状；优选基于短暂的、短期的（数小时到数天）、长期的（数周、数月或数年）或者永久的基础，改善患者的病理状况的治疗，其中病理状况的改善情况可以优选为不变的、渐增的、渐减的、不断改变的或者摆动的，只要总效果与对照患者相比是显著改善的。治疗效能和毒性，例如， $ED_{50}$  和  $LD_{50}$  可以通过细胞培养或实验动物中标准的药理学方法确定。治疗和毒性效果之间的剂量比为治疗指数并且可通过比例  $LD_{50}/ED_{50}$  表达。优选显示出大治疗指数的药物组合物。必须根据所治疗的单个患者的年龄、体重和状况，以及施用途径、剂量形式和方案、所希望的结果调整剂量，并且精确的剂量当然应该由执业医师决定。

实际的剂量依赖于所治疗的失调的性质和严重性，在医生的判断力之内，并且随着为产生所希望的治疗效果本发明的特定环境中剂量的滴度而变。然而，现今预期每剂含有约 0.1 到 500mg，优选约 1 到 100mg，最优选约 1 到 10mg 活性成分的药物组合物适于治疗性治疗。

活性成分可以每天以一个或几个剂量施用。满意的结果可以，在一些情况中，在静脉内（i.v.）0.1  $\mu\text{g}/\text{kg}$  和经口地（p.o.）1  $\mu\text{g}$  的低剂量下得到。优选的范围为 0.1  $\mu\text{g}/\text{kg}/\text{天}$  到约 10  $\text{mg}/\text{kg}/\text{天}$  i.v. 和

1  $\mu\text{g}/\text{kg}/\text{天}$ 到约 100  $\text{mg}/\text{kg}/\text{天 p.o.}$ 。

本发明的另一方面涉及含有根据 SEQ ID 1 到 9 和/或 SEQ ID 47 的多肽的融合蛋白质，或者其功能变体。

“融合蛋白质”指多肽，其含有至少一种根据 SEQ ID 1 到 9 和/或 SEQ ID 47 的多肽、其功能变体或部分和选自通过本领域中技术人员公知的共价或非共价结合如，例如，氢键结合到根据本发明的多肽的多肽、肽和/或肽类似物的至少一种组分 A。融合蛋白质的组分 A 的优选实例为多肽、肽和/或肽类似物，其使该融合蛋白质更易检测；这些实例为，例如，“绿色荧光蛋白质”，或者其变体。还包括方便了重组蛋白质的纯化的融合蛋白质如“his-标签”，或者增加蛋白质的免疫原性的融合。

根据本发明的融合蛋白质可以通过本领域技术人员公知的方法产生。根据本发明的融合蛋白质可用于肝脏失调和/或上皮癌的诊断、预防和/或治疗。

与本领域的状况相比，这些融合蛋白质令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

根据本发明的优选核酸具有根据 SEQ ID 10 到 SEQ ID No.19 之一的序列，或者其变体。本发明尤其涉及已经分离的根据本发明的核酸。

与本领域状况相比，这些核酸和多肽令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

术语“根据本发明的核酸”指相应于 SEQ ID 10 到 SEQ ID 19 的核酸和/或其变体。

术语“编码核酸”涉及编码根据本发明的可分离的生物活性多肽或者其前体的 DNA 序列。该多肽可被全长序列或者该编码序列的部分所编码，只要生物功能，如例如受体-活性，被基本上保留（参考功能变体的定义）。

公知上述核酸的序列中可以存在小的改变，例如，由于遗传密码的简并性，或者未翻译的序列可以粘附到核酸的 5'和/或 3'末端而不显著影响所编码多肽的活性。本发明因此还包括上述核酸的所谓天然

发生的和人工产生的“变体”。

优选地，根据本发明使用的核酸为 DNA 或 RNA，优选 DNA，尤其是双链 DNA。根据本发明的核酸尤其可以是 RNA 分子，优选单链或双链 RNA 分子。核酸的序列可以还含有至少一个内含子和/或一个 polyA 序列。

根据本发明的核酸可以通过技术人员公知的方法产生并且也已经下面详述。

在本发明的意思内“变体”指与 DNA 序列互补的所有 DNA 序列，其在严格条件下与参比序列杂交并且具有相应于根据本发明的多肽的类似活性。根据本发明的核酸也可以以它们的反义序列的形式使用。

核酸的“变体”也可以与来自其他物种的核酸同源，序列同一性优选为 80%，尤其是 90%，最优选 95%。

核酸的“变体”也可以是根据本发明的核酸的部分，其长为至少约 8 个核苷酸，优选至少约 16 个核苷酸，尤其至少约 21 个核苷酸，更优选至少约 30 个核苷酸，甚至更优选至少约 40 个核苷酸，最优选至少约 50 个核苷酸，只要这些部分具有与根据本发明的相应多肽类似的活性。这些活性可以使用上面描述的功能测定法测定。

在本发明的优选实施方案中，核酸含有具有与根据本发明的核酸互补的序列的核酸，或者其变体。该核酸优选含有根据本发明的核酸的非功能突变变体，或者其变体。

本发明尤其涉及具有互补序列的核酸，其中该核酸是反义分子或者 RNA 干扰分子。

术语“核酸的非功能突变变体”指从根据本发明的核酸得到的核酸或者其变体，该核酸已经被突变从而该核酸的非功能突变变体编码的多肽显示出与未突变的多肽相比显著降低或被消除的生物学活性。非功能突变的变体核酸编码的多肽的这种活性可以通过如上述用于评价功能变体的功能测定法确定。从根据本发明的核酸得到的这种非功能突变的变体的构建和筛选是本领域技术人员公知的。这种根据本发明的“核酸的非功能突变变体”可以在患者中表达并且将优选通过与用以被核糖体翻译成多肽的天然 mRNA 分子竞争而消除或减小目标核酸的表达水平。

“严格杂交条件”指那些条件，其中杂交在 60°C 2.5 × SSC 缓冲

液中发生并且在 37℃ 低盐浓度缓冲液中许多洗涤步骤后保持稳定。

术语“核酸的差别表达”指与参比样品或参比文库中核酸的表达相比，从患者分离的核酸的相对表达水平。参比样品和参比文库的定义已经在上面详述。该表达可以通过本领域技术人员公知的方法确定。这些方法的实例包括 RNA 印迹 (northern) 分析、核酸酶保护、原位杂交、反转录酶 PCR (RT-PCR; 包括定量动态 RT-PCR)。这些方法还包括 cDNA 和寡核苷酸微阵列。

在优选的实施方案中，根据本发明的核酸为 Obcl1 cDNA (SEQ ID 10)，其通过鉴定来自非丰余的和人 EST GenBank 序列数据库的重叠序列装配。相应于装配序列 SEQ ID 的 RNA 的 HCC 中表达通过实验证实。相对于被鉴定为 SSH cDNA 克隆的非患病肝脏，HCC 中被上调的最初序列相应于 GenBank 序列 AL050205。该序列的 5' 末端与 AF131755 重叠；该序列以 XM113703、AK055521 和 AY004310 的 5' 不断延伸。后三个序列包括编码 Obcl1. pr 的可读框 (SEQ ID 1)。为支持该 mRNA 构建体，所有重叠的 cDNA 序列可以定位于相同的染色体。此外，使用来自该克隆的 SSH 序列作为杂交探针通过 HCC 但不是正常肝脏 RNA 的 RNA 印迹分析鉴定了约 6 kb 的 mRNA (图 4)。相应于该克隆的序列的表达以前还没有在肝脏或 HCC 中报导。

在优选的实施方案中，根据本发明的多肽为 Obcl1. pr 多肽 (SEQ ID 1)，其从相对于未患病肝脏在 HCC 中被鉴定上调平均 2.9 倍的 mRNA 令人惊奇地鉴定出来 (Obcl1, SEQ ID 10) (见表 3A/3B)。用反转录 PCR 分析鉴定编码该多肽并与该 mRNA 重叠的 cDNA 序列并且这些核酸在 HCC 中类似地升高。关于该多肽序列在 HCC 中上升的水平以前未被识别。从 Obcl1. pr 多肽的序列，以用 BLAST 序列分析工具可以得到的保守结构域预测 CDD 算法可以鉴定两个保守序列结构域 (Altschul 等人, 1997, Nucleic Acids Res., 25: 3389-3402); 狼疮 La 多肽型 RNA 结合结构域 (SEQ ID 1, 氨基酸 47 到 125)，和功能未知的 GTPase 酶结构域 (SEQ ID 1, 氨基酸 90 到 203)。Obcl1. pr 序列在 GenBank 序列数据库中已经被指定为细胞骨髓增殖性白血病受体 (c-Mpl) 结合多肽。虽然是骨髓增殖性白血病病毒受体 (也称作血小板生成素受体) 的潜在调节剂，但是该多肽的功能角色还没有被在任何系统中描述。类似地，该多肽的表达模式还没有被公开。在受到表达描绘的 21



个肝脏肿瘤中的 11 个肿瘤 (52%) 中编码该多肽的 mRNA 相对于未患病肝脏上升 2 倍以上。在被描绘的 4 个病灶性结节性增生症 (FNHs) 的四个 (100%) 中编码该多肽的 mRNA 类似地上升至少 2 倍 (表 3A/3B)。对于根据本发明的该核酸和其他核酸, 该表达的值包括来自置于 cDNA 微阵列表达描绘实验的所有 21 个 HCC 样品的表达值比例数据, 包括不上升 2 倍或更高的来自样品的值。

该 mRNA 的表达对于肝脏失调显著特异, 因为在所分析的其他癌或者未患病组织包括肝脏、肾、胃、肺、皮肤和其他组织中都没有检测到表达 (见表 6)。Obc11 mRNA 的表达水平的独立的 RT-PCR 分析用基因特异的寡核苷酸引物包括 SEQ ID 22 和 SEQ ID 23 来确定。因此, 令人惊奇地发现分别在 OBc11. pr 多肽 (SEQ ID 1) 和编码该多肽的核酸 (SEQ ID 10) 与根据本发明的失调之间有强而且特异的相关性。因此, 该多肽和编码核酸可用于诊断根据本发明的失调, 例如用于诊断增生性 (包括癌性) 肝脏疾病和肝硬化之间的区别。

此外, 该 HCC-去调节的基因的表达与肝癌细胞的增殖相关, 表现出用 8 小时和 12 小时血清刺激静置细胞时, Hep3B 细胞系中 Obc11 mRNA 分别增加 3.4 倍和 6.3 倍 (见图 8)。

这些结果证明 OBc11. pr 多肽 (SEQ ID 1) 和编码该多肽的核酸 (SEQ ID 10) 可用于防止和治疗根据本发明的失调, 尤其用于治疗增生性 (包括癌性) 肝脏疾病。对于治疗, 优选实施该治疗从而 OBc11. pr 多肽或者编码该多肽的核酸的表达被减小和/或抑制, 例如, 通过施用与编码 OBc11. pr 多肽的核酸特异相互作用的反义寡核苷酸或 RNA 干扰分子实现治疗。备选地, 实施治疗从而 OBc11. pr 多肽的活性被减小和/或抑制, 例如, 通过将封闭 OBc11. pr 多肽活性的针对 OBc11. pr 多肽的抗体或者其抗体片段施用于需要这种治疗的患者实现该治疗。与本领域的状况相比, 该 OBc11. pr 多肽和/或 Obc11 核酸令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

在另一实施方案中, 根据本发明的核酸是 OBc15 核酸 (SEQ ID 11), 其是编码 OBc15. pr 多肽 (SEQ ID 2) 的编码序列。

完整序列从许多 GenBank 表达序列标签 (EST) 数据库序列和 GenBank 基因组数据库序列建立, 并且验证每个片段在 HCC 中的过表

达。例如，相对于未患病肝脏参比物，cDNA 微阵列上该核酸的序列平均上升 24.7 倍（表 3A/3B）。

相应于该克隆的部分序列的表达已经在一些组织和一些肿瘤（包括胎儿肝脏、结肠腺癌和局限于肝脏中的肿瘤转移）中报导，但是根据本发明的完整序列以前还没有描述。因此 OBc15 的上升的表达对于肝脏失调十分特异。OBc15 核酸或者推导的多肽的编码序列在根据本发明的失调，优选 HCC 中的上升的水平还没有被认识到。

检索公共结构域数据库（如 PubMed 和 SOURCE）得到关于该序列和根据本发明的所有序列的信息。期刊文章还没有公开根据本发明的大多数序列。来自自动测序的 cDNA 文库的 cDNA 克隆的相对丰度因此提供了关于根据本发明的该序列和其他序列此处所引用的证据。通过数据库如“SOURCE”（由斯坦福大学的遗传学系提供）可以得到该信息，该数据库包括来自（curate from）UniGene、Swiss-Prot、GeneMap99、RHdb、dbEST、GeneCards 和 Locus-Link 数据库的数据。

在另一优选的实施方案中根据本发明的多肽是 OBc15. pr 多肽（SEQ ID 2），其代表来自该去调节的 mRNA 序列的最大的可读框。该多肽序列不含有与特征多肽或者公知的结构基序同源的已识别序列。还没有描述该多肽的表达模式。潜在编码该多肽的 RNA 的表达与非-患病肝脏相比在 100% 的 HCC 病例中上升 2 倍以上，并且在所描绘的 21 个病例的 17 个病例中（81%）大于 8 倍。相对于非-患病肝脏，编码 mRNA 的上升的表达在肝脏腺瘤、FNH 和肝硬变肝脏中也是明显的，但是转录在肝硬变中上调的显著性稍小。编码该多肽的 mRNA 在非-患病人肺、脑（皮质）、结肠、睾丸组织中可以检测到但是在多数其他评价的癌中检测不到（表 6）。Obc15 RNA 表达水平的独立 RT-PCR 分析用包括 SEQ ID 24 和 SEQ ID 25 的基因特异的寡核苷酸引物确定。OBc15 cDNA 的高表达特异性通过与正常组织中表达模式相比 HCC、FNH 和如图 2 阐明的其他类型的癌中的定量评定（Q-PCR）证实。TaqMan 方法证实了与非癌性肝脏相比 HCC 和 FNH 中 OBc15 RNA（SEQ ID 11）的大量表达（图 2），TaqMan 方法利用了 GAPDH 和  $\beta$ -肌动蛋白的平行检查。相对于这些管家基因，Q-PCR 揭示与正常肝相比，肝癌和 FNH 中 OBc15 RNA 水平上升并且其他组织和其他癌中 OBc15 RNA 水平比正常肝脏低的多。对于 TaqMan 分析，用包括 SEQ ID 66、SEQ ID 67 和 SEQ ID 68

(“水解”探针)的基因特异性寡核苷酸引物确定了 OBc15 表达。

此外，原位杂交分析使用通过与 OBc15 RNA 特异杂交的放射性同位素标记的 OBc15 RNA 反义探针清楚地指出与正常肝组织切片中边缘信号相对比 HCC 中 OBc15 RNA 的位置(图 5)。

因此该多肽和/或编码 RNA 的过表达，可用于诊断肝脏失调。这些结果清楚地阐明 OBc15. pr 多肽和编码该多肽的核酸 (SEQ ID 11) 和其功能变体可用于诊断、预防和治疗根据本发明的失调，尤其是 HCC、肝腺瘤、FNH 和肝硬变。

关于治疗，优选实施治疗使得 OBc15. pr 多肽和/或其功能变体；或者编码该多肽的核酸和/或其功能变体的表达被减弱和/或抑制，例如，通过施用反义寡核苷酸或小干扰 RNA 分子实现治疗，这些分子与潜在编码 OBc15. pr 多肽和/或其功能变体的 SEQ ID 11 中定义的核酸特异相互作用。

备选地可以实施治疗使得 OBc15. pr 多肽和/或其功能变体的活性被减小和/或抑制，例如，通过将可以封闭 OBc15. pr 多肽和/或其功能变体的活性的针对 OBc15. pr 多肽和/或其功能变体的抗体，或者抗体片段施用于需要这种治疗的患者实现治疗。与本领域的状况相比，OBc15. pr 多肽和/或其功能变体；和/或 OBc15 核酸令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

详细的序列分析揭示 OBc15 mRNA 与其他真核生物的非编码 RNA 之间的序列相似性。此外，旨在检测该 RNA 的蛋白质产物的使用不同方法学的多种尝试还没有揭示这种产物。因此，该 RNA 可能不被翻译成多肽但是可能自身具有功能(例如，调节的)性质。非编码调节性 RNA 的疾病相关性现在正变得明显，如例如通过与真核生物果蝇中细胞增殖相关的非编码 RNA “bantam” (Brennecke J, Hipfner DR, Stark A, Russell RB, Cohen SM. Cell (2003) 4月4日; 113(1): 25-36), 和与转录因子 HES-1 相互作用而阻碍神经元分化的微 RNA-23 和通过微小 RNA-23 的角色所证明的 (Kawasaki, H. 和 Tiara, K. Nature (2003) 423: 838-842)。

用小干扰 RNA (siRNA) 寡核苷酸降低正增殖的人肝癌细胞中 OBc15 RNA 的水平(击倒)，这证明了 OBc15 RNA 升高的表达在肝脏失调，特

别是肝癌中的功能上的重要角色。在该实验中，当 OBC15 RNA 的水平降低时，编码肿瘤抑制基因成视网膜细胞瘤蛋白 1 (RB1) 的 mRNA 的水平被上调若干倍，OBC15 RNA 的水平使用如上述的 TaqMan Q-PCR 确定。RB1 mRNA 的水平使用 SYBR 绿定量 PCR 分析用引物 RB 1-p1 (SEQ ID 64) 和 RB 1-p2 (SEQ ID 65) 确定。通过 RB1 的负调节，HCC 中 OBC15 RNA 的上升的表达可因此促进肿瘤细胞生长 (图 6)。

在另一优选的实施方案中，根据本发明的核酸是 Gene Bank 序列 NM025160 代表的 IK2 核酸 (SEQ ID 12)，序列 NM025160 包括编码 IK2. pr 多肽 (SEQ ID 3) 的开放读框。IK2. pr 多肽是本发明的另一实施方案。已经报导在来自包括肝脏的一些组织的 cDNA 文库和腺瘤中相应于该克隆的 EST 序列，但是该序列以前还没有被包含在 HCC 中。该多肽的表达还没有在任何细胞或组织中描述。该多肽序列没有公知的功能，尽管该序列在演变中相当保守 (预测的多肽在一些哺乳动物、果蝇 (*Drosophila*) 和植物 (拟南芥) 中发现)。CDD 算法预测了在根据本发明的多肽序列中 WD40 型多肽-多肽相互作用域。在来自 HCC 患者的肝脏样品中，编码该多肽的 mRNA 的表达相对于非患病肝脏令人惊奇地在描绘的 21 例的 15 例 (71%) 中上升了平均 4.67 倍。相对于非-患病肝脏编码 mRNA 的上升表达在肝硬变肝脏中也是明显的 (表 3A/3B)。相对于非患病肝脏，在 FNH 中观察到编码该肽的 mRNA 的最高差别表达水平；在所描述的 4 例的 4 例中上调 8 倍。编码该多肽的 mRNA 还在其他人类癌中表达，这些癌包括乳腺、肺和肾的癌，并且在所检查的 17 种非患病组织中的 2 种 (乳腺和肾) 中也有表达。用包括 SEQ ID 26 和 SEQ ID 27 的基因特异性寡核苷酸引物进行了 IK2 mRNA 表达水平的独立 RT-PCR 分析。

这些结果表明该多肽和/或编码 mRNA 的过表达可用于根据本发明的失调的诊断、预防和治疗，尤其用于 HCC、FNH、肝硬变和上皮来源的癌的诊断。关于治疗，优选实施治疗使得 IK2. pr 多肽或者编码该多肽的核酸的表达被减弱和/或抑制，例如，通过施用反义寡核苷酸或 RNA 干扰分子实现治疗，这些分子与编码 IK2. pr 多肽的核酸特异相互作用。备选地可以实施治疗使得 IK2. pr 多肽的活性被减小和/或抑制，例如，通过将封闭 IK2. pr 多肽活性的针对 IK2. pr 多肽的抗体或者其抗体片段施用于需要这种治疗的患者实现治疗。与本领域

的状况相比, IK2. pr 多肽和/或 IK2 核酸令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

在另一优选的实施方案中, 根据本发明的核酸是代表 HCC 去调节的 cDNA 克隆的序列的 IK5 核酸 (SEQ ID 13)。已经报导在一些组织 (包括肝脏) 和一些肿瘤 (包括垂体和前列腺) 中相应于该克隆的序列的表达, 但是该序列以前还没有描述在 HCC 中被上调。在优选的实施方案中, 根据本发明的多肽是 IK5. pr 多肽 (SEQ ID 4), 其被 IK5 cDNA (SEQ ID 13) 编码。从 GenBank 数据库 (记录号 NM006407) 推论出该多肽序列为 JWA——一种维生素 A 应答多肽。尽管从用维生素 A 刺激培养的细胞描述了编码该假定的多肽的基因, 但是还没有描述该多肽在任何细胞或组织中存在并且功能是未知的。在 GenBank 数据库中 JWA 被进一步描述为细胞骨架相关的多肽。该多肽还和与 EAAC1 谷氨酸转运蛋白相互作用并降低其活性的啮齿动物多肽具有同源性。该序列的保守结构域检索指出可能存在异戊二烯基化的 rab 接纳体 1 结构域 (PRA1), 其可能调节与 G 蛋白信号分子的相互作用。在 100% 的所描绘的 HCC 病例中编码该多肽的 mRNA 的表达相对于未-患病肝脏平均升高 9.14 倍。类似地, 在腺瘤和 FNH 中编码 mRNA 的上升的表达也是明显的。编码 mRNA 的表达在肝硬变肝脏中也差异表达但是比在其他肝脏失调中的程度低。编码该多肽的 mRNA 在人的肺、肾和结肠癌中表达但是在所检查的 17 例未患病组织仅 1 例中有表达。用包括 SEQ ID 28 和 SEQ ID 29 的基因特异性寡核苷酸引物确定了 IK5 mRNA 的表达的独立 RT-PCR 分析。该多肽和/或编码 mRNA 的过表达可能标志特异的上皮来源的癌, 包括肝癌。这些结果表明 IK5 cDNA 序列的差异的上调表达对于根据本发明的失调是高度特异的。

此外, 该 HCC-去调节的基因的表达与肝癌细胞的增殖相关, 当静止细胞的 8 小时和 12 小时血清刺激时显示出 Hep3B 细胞系中 IK5 mRNA 分别增加 10.9 倍和 4.3 倍。

因此, IK5. pr 多肽和/或编码核酸可用于根据本发明的失调的诊断、预防和治疗, 尤其用以 HCC、腺瘤、FNH、肝硬变和上皮来源的癌的诊断。关于治疗, 优选实施治疗使得 IK5. pr 多肽或者编码该多肽的核酸的表达被减弱和/或抑制, 例如, 通过施用反义寡核苷酸或 RNA

干扰分子实现治疗，这些分子与编码 IK5. pr 多肽的核酸特异相互作用。备选地可以实施治疗使得 IK5. pr 多肽的活性被减小和/或抑制，例如，通过将可以封闭 IK5. pr 多肽活性的针对 IK5. pr 多肽的抗体或者其抗体片段施用于需要这种治疗的患者实现治疗。与本领域的状况相比，该 IK5. pr 多肽和/或 IK5 核酸令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

在另一优选的实施方法中，根据本发明的核酸是 DAP3 核酸 (SEQ ID 14)，其以前被公开 (记录号 X83544) 编码 DAP3. pr 多肽 (SEQ ID 5)。本发明还涉及死亡相关的多肽 3 (DAP3, SEQ ID 5)，其在培养的细胞中过表达时促进凋亡细胞死亡 (Kissil 等人, 1995, J. Biol. Chem., 270: 27932-6)。

该多肽促进线粒体 28S 核糖体复合物。同样，该多肽可能在许多 (如果不是全部) 组织和细胞中普遍表达，尽管显然以较低水平表达。没有描述内源 DAP3 的特定功能 (Cadvar Koc 等人, 2001, FEBS Lett., 492: 166-170)。在肝脏的结肠腺癌转移中描述了 DAP3 mRNA 的下调 (PCT/US01/30589)，但是还没有认识到根据本发明的失调，优选 HCC 中 DAP3 核酸或 DAP3 多肽的上升的水平。

纯化的基因组 DNA 的定量 RT-PCR (Q-PCR) 扩增分析表明在 10 例 HCC 的 8 例中 DAP3 基因在肝癌中以约 4-6 拷贝扩增，在 13 个非-癌性肝脏样品 (包括接近和远离肿瘤的肝硬变组织) 中都没有扩增。这些分析用 TaqMan 方法进行以精确定量 DAP3 基因组 DNA 的相对量，使用引物 DAP3-p5 (SEQ ID 71)、DAP3-p6 (SEQ ID 72) 和水解探针 DAP3 p-7 (SEQ ID 73)。实际上，DAP3 基因位于染色体 1q 上，通常发现该区域在 HCC 中扩增 (Buendia MA., Med Pediatr Oncol. (2002) 11 月; 39 (5): 530-5)。该发现表明表现为基因扩增的正选择压力可能驱动 HCC 中 DAP3 RNA 的过表达，证明 DAP3 在 HCC 中的功能上重要的角色。

在所描绘的 21 例 HCC 的 18 例 (86%) 中编码该多肽的 mRNA 的表达相对于未患病肝脏平均升高 5.5 倍。编码 mRNA 的上升的表达在其他肝脏失调中也是明显的但是比 HCC 中的程度低。用包括 SEQ ID 30 和 SEQ ID 31 的基因特异的寡核苷酸引物确定了 DAP3 mRNA 的表达水

平的独立 RT-PCR 分析。通过 Q-PCR 分析, 以 SYBR 绿技术, 使用  $\beta$ -肌动蛋白作为参比基因, 进一步证实了与正常肝脏相比 HCC 中上升的 DAP3 mRNA。在从所检查的 5 例 HCC 的每一种分离的 RNA 中, DAP3 mRNA 与  $\beta$ -肌动蛋白 mRNA 水平的比例与从 2 个正常肝脏样品分离的 RNA 中这些比例相比升高了 (DAP3 mRNA 与  $\beta$ -肌动蛋白 mRNA 的平均 HCC 比例=12.8, DAP3 mRNA 与  $\beta$ -肌动蛋白 mRNA 的平均正常肝脏比例=1.03)。使用 SYBR 绿分析以包括 SEQ ID 69 和 SEQ ID 70 的基因特异的寡核苷酸引物确定了 DAP3 mRNA 的 Q-PCR 分析。

DAP3 蛋白质的表达在 HCC 中被显著特异上调, 因为在所分析的其他癌中和包括肝脏、肾、胃、肺、皮肤和其他组织的未患病组织中表达非常低或者没有检测到。与正常肝脏和其他正常和患病组织相比 HCC 中 DAP3 蛋白质表达水平的特异增加进一步证明了 DAP3 在 HCC 中的功能相关性 (见表 6 和图 7)。用小干扰 RNA 分子 (siRNA; SEQ ID 54 和 SEQ ID 55) 实验减少肝癌细胞中 DAP3 mRNA 导致肝癌细胞中显著的形态变化和明显的生物化学改变从而细胞变大并且用标准方法提取 RNA 和蛋白质对于处理的细胞是不可能的。这些发现进一步支持了 HCC 中增加的 DAP3 的功能重要性。

这些结果表明 DAP3 cDNA 序列和 DAP3<sub>pr</sub> 多肽的强烈的上调表达对于根据本发明的失调, 特别是 HCC 是高度特异的。因此, DAP3 多肽和/或编码核酸可用于根据本发明的失调的诊断、预防和治疗, 尤其用于 HCC 的诊断。关于治疗, 优选实施治疗使得 DAP3 多肽或者编码该多肽的核酸的表达被减弱和/或抑制, 例如, 通过施用反义寡核苷酸或 RNA 干扰分子实现治疗, 这些分子与编码 DAP3 多肽的核酸特异相互作用。备选地可以实施治疗使得 DAP3 多肽的活性被减小和/或抑制, 例如, 通过将封闭 DAP3 多肽活性的针对 DAP3 多肽的抗体或者其抗体片段施用于需要这种治疗的患者实现治疗。与本领域的状况相比, DAP3 多肽和/或 DAP3 核酸令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

在另一优选的实施领域中本发明涉及 HCC 上调的 LOC5<sub>pr</sub> 假定多肽 (SEQ ID 6) 和编码该多肽的核酸 LOC5 (SEQ ID 15)。已经从包括肝脏的一些人组织的 cDNA 文库中鉴定了相应于该 mRNA 的 cDNA (信息

来自如上述的 SOURCE 数据库), 但是以前没有报导该序列在根据本发明的失调, 尤其是 HCC 中被上调。在 71% 的所描绘的 HCC 病例中该 mRNA 的表达相对于未患病肝脏上升 5 倍 (表 3B)。类似的分析揭示在置于该 cDNA 微阵列表达描绘步骤的 FNH 和大多数肝硬变肝脏中该 mRNA 的上升的表达。该 mRNA 在其他胃肠道癌中表达但是仅在所检查的 17 例未患病组织的脑和骨髓中表达。用包括 SEQ ID 32 和 SEQ ID 33 的基因特异的寡核苷酸引物确定 LOC5 mRNA 的表达水平的独立 RT-PCR 分析。LOC5. pr (SEQ ID 6) 是一种预测的 30 kDa 多肽 (GenBank 数据库中记录号 NP\_60917.1)。还没有描述任何细胞或组织中存在该多肽。还没有描述该预测的多肽的功能, CDD 结构域算法检索也没有揭示保守结构域。这些结果表明 LOC5 cDNA 序列的强烈的上调表达对于本发明的失调, 特别是在 HCC、FNH 和大多数肝硬变肝脏中是高度特异的。此外, 该 HCC-去调节基因的表达与肝癌细胞的增殖相关, 当对静止细胞用血清刺激 8 小时和 12 小时时 Hep3B 细胞系中 LOC5 mRNA 分别增加 3.7 倍和 8.8 倍 (见图 8)。

因此, LOC5. pr 多肽和/或其功能变体和/或编码核酸和/或其变体可用于根据本发明的失调的诊断、预防和治疗, 尤其用于 HCC、FNH 和大多数肝硬变肝脏中的诊断。关于治疗, 优选实施治疗使得 LOC5. pr 多肽或者编码该多肽的核酸的表达被减弱和/或抑制, 例如, 通过施用反义寡核苷酸或 RNA 干扰分子实现治疗, 这些分子与编码 LOC5. pr 多肽的核酸特异相互作用。备选地可以实施治疗使得 LOC5. pr 多肽的活性被减小和/或抑制, 例如, 通过将封闭 LOC5. pr 多肽活性的针对 LOC5. pr 多肽的抗体或者其抗体片段施用于需要这种治疗的患者实现治疗。与本领域的状况相比, LOC5. pr 多肽和/或 LOC5 核酸令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

在另一优选的实施方案中本发明涉及编码根据本发明的 SEC14L2. pr 多肽 (SEQ ID 7) 的 SEC14L2 核酸 cDNA (SEQ ID 16)。SEC14L2 mRNA 的表达已经在许多组织中描述, 但是该 mRNA 或者所编码的多肽的上升以前还没有在根据本发明的失调, 尤其是肝脏失调或癌中报导。SEC14L2. pr (SEQ ID 7) 是酵母 sec 多肽 14 的人同系物。尽管该多肽与酵母分泌途径有关, 但是该多肽或者其同系物的确切功能还没有



在任何物种中报导。还提议该人序列结合生育酚并且预测该多肽参与鲨烯转移、胆固醇生物合成或者更一般地细胞内运输 (Zimmer 等人, 2000, J. Biol. Chem. 275: 25672-25680)。还没有报导人细胞或组织中该多肽序列的表达。该多肽序列包括可能的 G-多肽结合和磷脂酰肌醇转移结构域和共有 CRAL-TRIO 结构域。CRAL-TRIO 结构域通过顺式-视黄醛 CRAL 基序参与维生素结合。在 71% 的 HCC 样品、所有描绘的 FNH 疾病样品、但是不在胰腺中、在仅一半肝硬变样品中编码该多肽的 mRNA 相对于未患病肝脏平均上升 5.14 倍或更高 (表 3A/3B)。在肾和结肠癌和正常胰腺中但是未在所检查的其他正常组织中检测到编码该多肽的 mRNA 的表达 (表 6)。用包括 SEQ ID 34 和 SEQ ID 35 的特异寡核苷酸引物确定了 SEC14L2 mRNA 的表达水平的独立 RT-PCR 分析。此外, 该 HCC-去调节的基因的表达与肝癌细胞的增殖相关, 表现出静止细胞的 8 小时和 12 小时血清刺激时 Hep3B 细胞系中 SEC14L2 mRNA 的分别为 10.6 倍和 1.9 倍的增加 (见图 8)。

这些结果表明 SEC14L2 cDNA 序列的强烈上调的表达对于根据本发明的失调, 特别是 HCC 和 FNH 是高度特异的。因此, SEC14L2.pr 多肽和/或编码核酸可用于根据本发明的失调的诊断、预防和治疗, 尤其用于 HCC、FNH 和还优选肝硬变中的诊断。关于治疗, 优选实施治疗使得 SEC14L2.pr 多肽或者编码该多肽的核酸的表达被减弱和/或抑制, 例如, 通过施用反义寡核苷酸或 RNA 干扰分子实现治疗, 这些分子与编码 SEC14L2.pr 多肽的核酸特异相互作用。备选地可以实施治疗使得 SEC14L2.pr 多肽的活性被减小和/或抑制, 例如, 通过将封闭 SEC14L2.pr 多肽活性的针对 SEC14L2.pr 多肽的抗体或者其抗体片段施用于需要这种治疗的患者实现治疗。与本领域的状况相比, SEC14L2.pr 多肽和/或 SEC14L2 核酸令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

在另一优选的实施案中, 本发明涉及编码 SSP29.pr 或 APRIL 多肽的核酸 (SEQ ID 17), 其已经在许多组织和肿瘤中描述。以前还没有报导编码该假定的肿瘤坏死家族成员的基因在根据本发明的失调, 尤其是 HCC 中以升高的水平表达。此外, 本发明涉及可银染的 29kDa 多肽 (SSP29.pr; SEQ ID 8), 其被根据本发明的核酸 (SEQ ID 17)

编码。该多肽已经被鉴定为富亮氨酸的分泌多肽，可能属于 TNF 细胞因子家族。其还被称为 APRIL（富含亮氨酸的酸性多肽）并且在 N-末端附近含有富亮氨酸重复（LRRs），其可能参与抗原-介导的细胞应答（Zhu 等人，1997, *Biochem. Mol. Biol. Int.* 42: 927-935; Mencinger 等人，1998, *Biochim. Biophys. Acta* 1395: 176-180）。还没有报导人细胞或组织中有 SSP29. pr 多肽的表达。在描绘的 21 例 HCC 的 17 例中编码该多肽的 mRNA 相对于未患病肝脏平均上升 3.77 倍。令人惊奇地，在铜中毒导致的肝硬变中编码该多肽的 mRNA 水平比未患病肝脏库中高 30 倍（表 3A/3B）。相对于未患病肝脏在描绘的其他肝脏失调中 mRNA 水平或多或少地上升了并且否则在受到表达描绘的正常和患病组织中仅仅偶尔检测到该 mRNA。用包括 SEQ ID 36 和 SEQ ID 37 的基因特异的寡核苷酸引物确定了 SSP29 mRNA 表达的独立 RT-PCR 分析。此外，该 HCC-去调节的基因的表达与肝癌细胞的增殖有关，显示出当静止细胞的 8 小时和 12 小时血清刺激时在 HeP3B 细胞系中 SSP29 增加分别为 2.4 倍和 4.3 倍（见图 8）。这些结果表明 SSP29 cDNA 的强烈上调的表达对于根据本发明的失调，特别是 HCC 中和某些类型的肝硬变疾病中是高度特异的。

因此，SSP29. pr 多肽和/或编码核酸可用于根据本发明的失调的诊断、预防和治疗，尤其用于 HCC、FNH 和还优选肝硬变中的诊断。关于治疗，优选实施治疗使得 SSP29. pr 多肽或者编码该多肽的核酸的表达被减弱和/或抑制，例如，通过施用反义寡核苷酸或 RNA 干扰分子实现治疗，这些分子与编码 SSP29. pr 多肽的核酸特异相互作用。备选地可以实施治疗使得 SSP29. pr 多肽的活性被减小和/或抑制，例如，通过将封闭 SSP29. pr 多肽活性的针对 SSP29. pr 多肽的抗体或者其抗体片段施用于需要这种治疗的患者实现治疗。与本领域的状况相比，SSP29. pr 多肽和/或 SSP29 核酸令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

在另一优选的实施方式中，本发明涉及 HS16 核酸（SEQ ID 18），相应于 HS16 mRNA 的 cDNA 克隆已经在包括结肠的腺癌的一些组织中鉴定，但是该 mRNA 和编码的多肽（HS16. pr, SEQ ID 9）以前都没有与根据本发明的失调，尤其是肝脏失调或 HCC 相关。本发明还涉及编

码 HS16 的多肽, 其是 16.7kDa 的预测的多肽 (SEQ ID 9; GenBank 中记录号 NP\_057223)。还没有描述任何细胞或组织中存在该多肽并且其功能还没被鉴定, 用 CDD 算法也没有鉴定其功能结构域。编码该多肽的 mRNA 在检查的 8 例 HCC 中升高至少 2.8 倍或更高, 在所检查的额外 4 个 HCC 样品中升高近 2 倍, 所有都相对于未患病的肝脏 (表 3A/3B)。用包括 SEQ ID 38 和 SEQ ID 39 的基因特异的寡核苷酸引物确定了 HS16 mRNA 表达的独立 RT-PCR 分析。这些结果表明 HS16 cDNA 的强烈上调的表达对于根据本发明的失调, 特别是 HCC 中是高度特异的。

因此, HS16.pr 多肽和/或编码核酸可用于根据本发明的失调的诊断、预防和治疗, 尤其用于 HCC 的诊断。关于治疗, 优选实施治疗使得 HS16.pr 多肽或者编码该多肽的核酸的表达被减弱和/或抑制, 例如, 通过施用反义寡核苷酸或 RNA 干扰分子实现治疗, 这些分子与编码 HS16.pr 多肽的核酸特异相互作用。备选地可以实施治疗使得 HS16.pr 多肽的活性被减小和/或抑制, 例如, 通过将封闭 HS16.pr 多肽活性的针对 HS16.pr 多肽的抗体或者其抗体片段施用于需要这种治疗的患者实现治疗。与本领域的状况相比, HS16.pr 多肽和/或 HS16 核酸令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

在一个优选的实施案中, 根据本发明的核酸是 IK3 cDNA (SEQ ID 19), 其通过从非丰余的 GenBank 序列数据库中鉴定重叠序列而得到。用 cDNA 微阵列分析相对于未患病肝脏在 HCC 中上调的最初序列相应于 GenBank 数据库中胎儿脑 cDNA (AL049338)。该序列与一种编码酪氨酸磷酸酶 D 型受体 (PTPRD) 的小鼠 cDNA XM 131462 (SEQ ID. No. 47) 重叠。尽管该小鼠 PTPRD 与人 PTPRD 转录单位高度同源, 但是在该人 PTPRD 转录单位序列中没有发现与该肝癌去调节的 RNA 同源的区。因此, 可能该 HCC-调节的序列编码还没有描述的人 PTPRD。备选地, 所提供的数据库序列可能含有错误, 其是缺乏开放读框的原因。另一备选原因是所编码的多肽可能来自该序列中小开放读框之一。甚至, 该 RNA 可能不被翻译成多肽但是可能自身具有功能 (例如, 调节的) 性质。

令人惊奇地，来自该 mRNA 的序列在 HCC 中比在正常人肝脏中以更高水平表现。另外，该 RNA 在正常的脑、骨骼肌、前列腺和肝脏中仅以低水平表达。在描绘的 21 个 HCC 样品中的 12 个样品 (57%) 中该 mRNA 相对于未患病肝脏平均上升 3.81 倍或更高。在所检查的 4 例 FNH 的 3 例中，相对于未患病肝脏，腺癌中和所检查的 6 个肝硬变样品中的 5 个中 IK3 也上升了 2 倍或更高 (表 3A/3B)。用包括 SEQ ID 40 和 SEQ ID 41 的基因特异的寡核苷酸引物确定了 IK3 mRNA 表达的独立 RT-PCR 分析。这些结果表明 IK3 cDNA 的强烈上调的表达对于根据本发明的失调，特别是 HCC、FNH、腺癌和肝硬变中是高度特异的。

因此，IK3 多肽和/或其功能变体，和/或编码核酸和/或其变体可用于根据本发明的失调的诊断、预防和治疗，尤其用于 HCC、FNH、腺癌和肝硬变的诊断。关于治疗，优选实施治疗使得 IK3 编码的多肽或者 IK3 核酸的表达被减弱和/或抑制，例如，通过施用反义寡核苷酸或 RNA 干扰分子实现治疗，这些分子与编码 IK3 核酸特异相互作用。备选地可以实施治疗使得 IK3 多肽的活性被减小和/或抑制，例如，通过将封闭 IK3 多肽活性的针对 IK3 多肽的抗体或者其抗体片段施用于需要这种治疗的患者实现治疗。与本领域的状况相比，该 IK3 核酸令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

评价了相对于未患病肝脏参比样品来自包括 HCC 的人肝脏失调的组织中根据本发明的序列的 cDNA 表达水平，这些 cDNA 表达水平在代表两组独立的实验的表 3A/3B 中显示。表 3B 中的值代表表达水平的 log2 比例而表 3A 是从对定制的 cDNA 微阵列的竞争性杂交得到的患病和未患病样品之间的未转化的数据。HCC=肝细胞癌样品；HCC (IHB) =含有 HCC 样品的 intrahyaline 体；FNH=病灶结节性增生样品；Cirrh=肝硬变样品。提供了每组 (HCC、FNH 和 Cirrh) 每个序列 (SEQ ID 10 到 19) 值的平均值、中位数 (值的第 50 个百分位数) 和标准误。

表 3A: cDNA 微阵列表达水平比 (未转化的值)

疾病样品 ↓	OBe1 SEQ ID 10	OBe15 SEQ ID 11	IK2 SEQ ID 12	IK5 SEQ ID 13	DAP3 (A) SEQ ID 14	DAP3 (B) SEQ ID 14	LOC5 SEQ ID 15	SEC14L2 SEQ ID 16	SSP29 SEQ ID 17	HS16 SEQ ID 18	IK3 SEQ ID 19
HCC11	4.0	27.7	2.4	8.8	2.2	4.7	1.0	15.6	2.0	1.1	1.5
HCC12	1.5	38.2	1.1	11.5	3.1	6.5	4.3	14.0	3.8	1.9	2.1
HCC13	1.4	44.6	7.8	6.7	7.3	9.3	7.1	3.3	2.1	1.8	2.8
HCC15	2.6	30.5	1.4	3.7	23.9	3.3	5.8	3.9	8.6	1.9	1.7
HCC1	2.4	40.5	7.6	9.6	1.8	2.4	7.0	9.4	5.6	1.6	12.8
HCC27	11.6	11.8	4.2	2.5	6.2	2.5	8.2	4.6	5.0	9.2	7.1
HCC29	10.9	22.9	13.5	3.9	6.7	7.6	5.4	1.9	7.0	4.7	3.4
HCC2	2.2	41.4	8.3	5.4	2.5	9.3	8.9	2.3	1.9	1.8	12.5
HCC30n	1.9	23.8	0.9	21.0	2.5	3.4	3.6	5.5	10.4	1.7	0.8
HCC31	1.3	9.7	0.6	13.8	2.9	3.4	3.0	0.9	3.3	3.6	0.9
HCC32	2.8	7.8	4.4	6.0	3.4	3.1	3.5	4.4	3.0	4.7	3.2
HCC33	0.9	7.1	2.0	4.1	1.9	3.4	8.3	7.2	1.9	0.9	3.4
HCC34	2.9	48.3	9.4	21.7	3.8	8.5	12.9	16.3	1.1	1.4	7.6
HCC35	4.0	3.1	4.2	7.2	5.4	2.8	4.3	3.3	5.0	4.9	3.0
HCC36	1.8	36.3	5.0	8.4	5.3	4.6	6.1	3.3	2.4	1.4	1.3
HCC4	1.7	21.4	8.3	15.4	10.6	19.0	2.0	0.9	2.5	3.3	1.8
HCC6	0.7	15.6	0.5	1.9	1.4	2.3	1.4	1.6	2.6	4.3	1.6
HCC9	1.2	52.7	3.6	15.6	2.7	2.3	1.1	4.4	3.8	1.4	0.9
HCC (IHB)	0.6	20.4	8.4	14.0	19.2	10.0	9.8	1.2	2.5	5.1	4.8
HCC22	3.2	10.5	2.2	5.0	2.4	1.7	0.8	2.4	2.7	1.2	5.1
HCC28	0.6	5.3	2.3	5.6	1.3	1.4	0.7	1.4	1.8	1.1	1.9
HCC平均值	2.9	24.7	4.7	5.1	5.6	5.3	5.0	5.1	3.8	2.8	3.8
HCC 中位数	1.9	22.9	4.2	7.2	3.1	3.4	4.3	3.3	2.7	1.8	2.8
HCC 标准误	3.0	15.4	3.6	5.8	5.9	4.2	3.4	4.8	2.4	2.1	3.5

疾病样品 ↓	OBe11 SEQ ID 10	OBe15 SEQ ID 11	IK2 SEQ ID 12	IK5 SEQ ID 13	DAP3 (A) SEQ ID 14	DAP3 (B) SEQ ID 14	LOC5 SEQ ID 15	SEC14L2 SEQ ID 16	SSP29 SEQ ID 17	HS16 SEQ ID 18	IK3 SEQ ID 19
<b>FNH1</b>	2.5	7.0	8.0	10.1	4.6	1.9	10.2	7.1	2.3	4.9	0.9
<b>FNH2</b>	4.7	7.1	10.9	16.2	2.2	4.4	7.1	4.2	2.2	2.1	16.6
<b>FNH3</b>	3.0	4.2	9.5	11.5	1.5	2.6	9.6	6.0	1.0	2.1	9.9
<b>FNH9</b>	3.4	15.1	7.7	9.9	1.7	3.2	2.4	3.8	0.9	1.3	7.5
<b>FNH 平均值</b>	<b>3.4</b>	<b>8.3</b>	<b>9.1</b>	<b>11.9</b>	<b>2.5</b>	<b>3.0</b>	<b>7.3</b>	<b>5.3</b>	<b>1.6</b>	<b>2.6</b>	<b>8.7</b>
<b>FNH 中位数</b>	<b>3.2</b>	<b>7.1</b>	<b>8.8</b>	<b>10.8</b>	<b>2.0</b>	<b>2.9</b>	<b>8.4</b>	<b>5.1</b>	<b>1.6</b>	<b>2.1</b>	<b>8.7</b>
<b>FNH 标准误</b>	<b>0.9</b>	<b>4.7</b>	<b>1.5</b>	<b>2.9</b>	<b>1.4</b>	<b>1.1</b>	<b>3.5</b>	<b>1.5</b>	<b>0.8</b>	<b>1.6</b>	<b>6.5</b>

<b>Cirrh34b</b>	7.6	17.7	6.0	6.0	13.7	3.2	9.3	2.3	19.6	8.6	4.2
<b>Cirrh5</b>	0.5	2.7	12.9	2.7	1.2	3.0	10.3	4.0	16.0	2.0	3.9
<b>Cirrh1</b>	1.0	1.8	2.2	2.8	7.5	3.0	1.9	2.3	9.3	12.2	10.1
<b>Cirrh2</b>	0.4	2.6	2.9	2.9	13.9	0.9	2.4	3.3	1.8	1.3	2.7
<b>Cirrh3</b>	0.4	4.0	15.2	22.1	1.3	2.8	1.4	0.8	2.4	3.6	1.7
<b>Cirrh4</b>	0.8	10.8	24.7	9.0	2.4	3.9	2.7	1.7	1.0	3.8	4.6
<b>Cirrh平均值</b>	<b>1.8</b>	<b>6.6</b>	<b>10.7</b>	<b>7.6</b>	<b>6.7</b>	<b>2.8</b>	<b>4.7</b>	<b>2.4</b>	<b>8.3</b>	<b>5.3</b>	<b>4.5</b>
<b>Cirrh. 中位数</b>	<b>0.7</b>	<b>3.4</b>	<b>9.5</b>	<b>4.5</b>	<b>5.0</b>	<b>3.0</b>	<b>2.6</b>	<b>2.3</b>	<b>5.9</b>	<b>3.7</b>	<b>4.1</b>
<b>Cirrh. 标准误</b>	<b>2.9</b>	<b>6.4</b>	<b>8.7</b>	<b>7.5</b>	<b>6.0</b>	<b>1.0</b>	<b>4.0</b>	<b>1.1</b>	<b>8.0</b>	<b>4.3</b>	<b>2.9</b>

<b>腺瘤</b>	1.9	10.0	1.7	6.9	1.6	3.6	1.8	1.1	2.2	1.5	3.7
-----------	-----	------	-----	-----	-----	-----	-----	-----	-----	-----	-----

<b>铜中毒</b>	2.3	18.7	3.5	7.2	7.0	8.4	13.0	7.3	35.5	22.4	9.5
------------	-----	------	-----	-----	-----	-----	------	-----	------	------	-----

<b>未患病肝脏</b>	0.7	0.6	n,d,	2.6	1.4	1.5	1.7	1.6	1.1	2.0	1.2
--------------	-----	-----	------	-----	-----	-----	-----	-----	-----	-----	-----

表 3B: cDNA 微阵列核酸表达水平比 (log<sub>2</sub> 值)

疾病样品 ↓	OBe11 SEQ ID 10	OBe15 SEQ ID 11	IK2 SEQ ID 12	IK5 SEQ ID 13	DAP3 SEQ ID 14	LOC5 SEQ ID 15	SEC14L2 SEQ ID 16	SSP29 SEQ ID 17	HS16 SEQ ID 18	IK3 SEQ ID 19
HCC11	1.99	4.79	1.24	3.13	1.12	0.05	3.96	1.02	0.14	0.57
HCC12	0.55	5.26	0.19	3.52	1.64	2.10	3.80	1.93	0.90	1.08
HCC13	0.46	5.48	2.97	2.74	2.87	2.82	1.73	1.05	0.81	1.50
HCC15	1.36	4.93	0.50	1.89	4.58	2.53	1.95	3.11	0.91	0.76
HCC1	1.24	5.34	2.92	3.26	0.83	2.82	3.23	2.47	0.64	3.68
HCC27	3.53	3.56	2.06	1.34	2.64	3.04	2.21	2.32	3.20	2.82
HCC29	3.45	4.52	3.75	1.96	2.75	2.43	0.91	2.81	2.22	1.75
HCC2	1.15	5.37	3.05	2.43	1.29	3.16	1.21	0.94	0.88	3.65
HCC30n	0.96	4.57	-0.19	4.40	1.31	1.84	2.46	3.38	0.77	-0.38
HCC31	0.42	3.27	-0.63	3.79	1.55	1.57	-0.12	1.74	1.85	-0.15
HCC32	1.48	2.96	2.15	2.59	1.77	1.81	2.13	1.59	2.22	1.68
HCC33	-0.17	2.83	0.97	2.05	0.96	3.06	2.85	0.94	-0.07	1.76
HCC34	1.54	5.59	3.23	4.44	1.93	3.69	4.03	0.16	0.44	2.92
HCC35	1.99	1.63	2.08	2.85	2.42	2.09	1.74	2.31	2.29	1.61
HCC36	0.86	5.18	2.32	3.08	2.40	2.61	1.74	1.28	0.52	0.35
HCC4	0.77	4.42	3.05	3.94	3.41	0.97	-0.13	1.35	1.73	0.87
HCC6	-0.60	3.96	-0.87	0.94	0.45	0.44	0.68	1.36	2.09	0.70
HCC9	0.29	5.72	1.84	3.97	1.44	0.17	2.15	1.92	0.47	-0.13
IHB-HCC	-0.85	4.35	3.06	3.81	4.27	3.29	0.32	1.34	2.35	2.26
HCC22	1.66	3.39	1.16	2.31	1.28	-0.31	1.26	1.45	0.32	2.35
HCC28	-0.75	2.41	1.22	2.49	0.42	-0.62	0.45	0.86	0.12	0.90
HCC平均值	1.02	4.26	1.72	2.90	1.97	1.88	1.84	1.68	1.18	1.45
HCC 中位数	0.96	4.52	2.06	2.85	1.64	2.10	1.74	1.45	0.88	1.50
HCC 标准误	1.17	1.16	1.35	0.97	1.14	1.29	1.26	0.81	0.93	1.18

疾病样品 ↓	OBe11 SEQ ID 10	OBe15 SEQ ID 11	IK2 SEQ ID 12	IK5 SEQ ID 13	DAP3 SEQ ID 14	LOC5 SEQ ID 15	SEC14L2 SEQ ID 16	SSP29 SEQ ID 17	HS16 SEQ ID 18	IK3 SEQ ID 19
<b>FNH1</b>	1.34	2.81	3.01	3.33	2.19	3.34	2.83	1.21	2.29	-0.18
<b>FNH2</b>	2.24	2.84	3.45	4.01	1.14	2.83	2.08	1.11	1.08	4.05
<b>FNH3</b>	1.58	2.07	3.25	3.53	0.59	3.27	2.58	0.06	1.06	3.30
<b>FNH9</b>	1.76	3.91	2.95	3.31	0.75	1.28	1.93	-0.12	0.35	2.91
<b>FNH平均值</b>	<b>1.56</b>	<b>2.99</b>	<b>2.68</b>	<b>3.39</b>	<b>1.07</b>	<b>2.32</b>	<b>1.91</b>	<b>0.67</b>	<b>1.07</b>	<b>2.40</b>
<b>FNH中位数</b>	<b>1.58</b>	<b>2.84</b>	<b>3.01</b>	<b>3.33</b>	<b>0.75</b>	<b>2.83</b>	<b>2.08</b>	<b>1.11</b>	<b>1.06</b>	<b>2.91</b>
<b>FNH标准误</b>	<b>0.50</b>	<b>0.68</b>	<b>1.10</b>	<b>0.44</b>	<b>0.66</b>	<b>1.16</b>	<b>1.05</b>	<b>0.65</b>	<b>0.75</b>	<b>1.64</b>

<b>Cirrh34b</b>	2.92	4.14	2.58	2.59	3.78	3.22	1.17	4.29	3.11	2.08
<b>Cirrh5</b>	-0.97	1.42	3.69	1.42	0.24	3.36	2.01	4.00	1.01	1.97
<b>Cirrh1</b>	0.02	0.86	1.16	1.48	2.91	0.92	1.19	3.22	3.61	3.34
<b>Cirrh2</b>	-1.43	1.37	1.55	1.55	3.80	1.29	1.70	0.81	0.40	1.44
<b>Cirrh3</b>	-1.28	1.99	3.93	4.47	0.40	0.53	-0.40	1.28	1.85	0.74
<b>Cirrh4</b>	-0.37	3.43	4.62	3.17	1.27	1.44	0.75	-0.05	1.92	2.20
<b>Cirrh平均值</b>	<b>-0.18</b>	<b>2.20</b>	<b>2.92</b>	<b>2.45</b>	<b>2.07</b>	<b>1.79</b>	<b>1.07</b>	<b>2.26</b>	<b>1.98</b>	<b>1.96</b>
<b>Cirrh中位数</b>	<b>-0.67</b>	<b>1.70</b>	<b>3.13</b>	<b>2.07</b>	<b>2.09</b>	<b>1.37</b>	<b>1.18</b>	<b>2.25</b>	<b>1.89</b>	<b>2.02</b>
<b>Cirrh标准误</b>	<b>1.62</b>	<b>1.30</b>	<b>1.39</b>	<b>1.22</b>	<b>1.63</b>	<b>1.20</b>	<b>0.84</b>	<b>1.82</b>	<b>1.22</b>	<b>0.86</b>

<b>腺瘤</b>	0.89	3.32	0.75	2.78	0.70	0.87	0.15	1.11	0.56	1.89
-----------	------	------	------	------	------	------	------	------	------	------

<b>铜中毒</b>	1.21	4.23	1.82	2.85	2.82	3.70	2.87	5.15	4.48	3.25
------------	------	------	------	------	------	------	------	------	------	------

<b>未患病肝脏</b>	-0.53	-0.84	n.d.	1.39	0.50	0.80	0.69	0.19	0.97	0.23
--------------	-------	-------	------	------	------	------	------	------	------	------



在表 4 中显示了 cDNA 微阵列核酸表达值的概述。应用 Mann-Whitney-U 检验统计地评估 RNA 表达水平：该检验等于成对旗标=“假”的 Wilcoxon Rank Sum 双边检验 (Hollander & Wolfe, 1973, 非参数统计学推论 (Nonparametric statistical inference) 纽约: John Wiley & Sons, 27- 33 页, 68-75; Bauer, D. F., 1972, J. Amer. Statistical Assoc. 67: 687-690)。表达值通常不符合正态分布从而将值平均可能是误导的。然而, 中位值的分析阐明了大多数情况下实验和参比值之间的显著差异, 在大数据组中尤其如此。Expt. 中位值=实验 (患病) 组织的中位值; Expt. iqr=实验值四分位间范围 (中位值的+/-25 百分位数); Contr. 中位值=对照 (未患病) 组织样品的中位值; Contr. iqr=对照值四分位间范围 (中位值的+/-25 百分位数); p 值=对实验值和对照值显著不同这一可能性的统计学评价得到的值。

表 4: cDNA 微阵列核酸 (SEQ ID 10 到 19) 表达值概述

HCC					
	实验 中位数	实验 iqr	对照 中位数	对照 iqr	P 值
<b>OBcl1</b>	6482	4915	3235	1050	0.0001
<b>OBC15</b>	995.5	1549.1	832.2	195	0.0156
<b>IK2</b>	582.7	348.9	874.3	344.1	0.0397
<b>IK5</b>	600.1	330.4	760.9	261.5	0.0056
<b>DAP3</b>	1202	1271.7	927	391.3	0.0499
<b>LOC5</b>	673.7	256.2	965	255.4	0.0255
<b>SEC14L2</b>	457.39	351.17	869.7	306.1	0.0003
<b>SSP29</b>	949.9	475.1	976.2	327.9	0.6792
<b>HS16</b>	1269	483	1083	494.4	0.2293
<b>IK3</b>	651.7	305.2	842.2	297.3	0.0080
FNH					
	实验 中位数	实验 iqr	对照 中位数	对照 iqr	P 值
<b>OBcl1</b>	8279.2	3205	3550.1	684	0.0286
<b>OBC15</b>	806.4	1563.4	737.6	106.5	0.4857
<b>IK2</b>	1165.1	222	887.2	137	0.6857
<b>IK5</b>	1358.9	383	882.1	196.6	0.4857
<b>DAP3</b>	1555.6	569	1046.2	136	0.3429
<b>LOC5</b>	971.3	459.3	890.7	131	0.6857
<b>SEC14L2</b>	807.3	262.9	806	176.6	0.6857
<b>SSP29</b>	1484.4	462	1139.9	101	0.2000
<b>HS16</b>	1556.2	644	1156.5	113	0.4857
<b>IK3</b>	1298.9	131	800.7	360.4	0.3429

肝硬化					
	实验 中位数	实验 iqr	对照 中位数	对照 iqr	P 值
<b>OBcl1</b>	2518	1923	4108	869	0.2403
<b>OBCI5</b>	318.4	187	1318	321	0.0087
<b>IK2</b>	408.3	235	1195	194	0.0022
<b>IK5</b>	244	251.7	1238	995	0.0022
<b>DAP3</b>	576.1	568.1	1417	446	0.0022
<b>LOC5</b>	355.6	360	1377	293	0.0022
<b>SEC14L2</b>	192.3	112.8	1287	243	0.0022
<b>SSP29</b>	361.3	140.4	1547	501	0.0087
<b>HS16</b>	246.7	250.5	1392	300	0.0022
<b>IK3</b>	378.6	446.6	1217	423	0.0043

在表 5 中显示了非癌性肝脏疾病和肝癌中核酸表达值的比较。对于根据本发明的每一种核酸，为 FNH、Cirrh. 和 HCC 样品之间比较的中位实验表达值提供了 P 值。对于每种核酸和比较，小于或等于 0.05 的 P 值表明患病组之间表达值的显著差异。用 Wilcoxon rank sum 检验评估重要性。在患病组之间表达中的统计学显著性差异是明显的。例如，在所有三个比较中 IK2 的表达值是显著差异的 (P 值小于 0.05)。FNH 样品组是小的并且显示出值的大分布。这可能是由于与该组的比较中较少的显著差异的原因。

表 5A: HCC 对 Cirrh、HCC 对 FNH、Cirrh 对 FNH 中核酸 (SEQ ID 10 到 19) 的表达特异性。

	HCC 对 Cirrh.	HCC 对 FNH	Cirrh. 对 FNH
<b>OBcl1</b>	0.0013	0.2718	0.0095
<b>OBcl5</b>	0.0010	0.7672	0.0667
<b>IK2</b>	0.0042	0.0081	0.0095
<b>IK5</b>	0.0078	0.0031	0.0095
<b>DAP3</b>	0.0078	0.4885	0.0667
<b>LOC5</b>	0.0042	0.1109	0.0095
<b>SEC14L2</b>	0.0004	0.0817	0.0095
<b>SSP29</b>	0.0052	0.0336	0.0095
<b>HS16</b>	0.0168	0.4085	0.0095
<b>IK3</b>	0.1273	0.0014	0.0095

表 5B 中的 Mann-Whitney U 指出当值以升序排列时第一组 (HCC) 中的值超过第二组 (分别为 FNH 和 Cirrh) 的次数。Wilcoxon W 是 Mann-Whitney Wilcoxon Rank Sum 检验中两组的较大者的等级的总和。Asymptotic 显著性 (Asymp. Sig.) (双尾) 提供了双侧检验的 P 值。该统计分析用于确定 OBcl5 (HCC 对 FNH、HCC 对 Cirrh) 的表达模式的总趋势, 其通过表 7 中提供的定量 RT-PCR (Q-PCR) 数据的统计学验证并在图 2 中显示。

图 5B: HCC 对 FNH 和 HCC 对 Cirrh 中 OBcl5 的表达特异性

#### HCC 对 FNH

	Man-Whitney U	Wilcoxon W	Asymp. Sig. (双尾)
<b>OBC15</b>	18.0	33.0	0.025

## HCC 对 Cirrh

	Man-Whitney U	Wilcoxon W	Asymp. Sig. (双尾)
OBC15	15.0	36.0	0.005

用对在所列的每种组织中每种去调节的核酸特异的引物实施了反转录酶聚合酶链式反应 (RT-PCR) 以确定是否该序列存在于从每种组织制备的 RNA 中。所用的所有组织在用于 RNA (和 cDNA) 制备前通过诊断证实。在表 6 中符号 “+” 表示基因在组织中表达, “-” 表示来自该 RNA 样品的 cDNA 中没有检测到该基因; 空格表示没有对该基因和组织组合进行分析。提供了患者的年龄和性别。其他的样品信息包括肿瘤分期值 (T=肿瘤大小), 以及肿瘤分级得分 (G=肿瘤细胞分化); 大的数字分别指出更大的和分化较差的肿瘤。组织 cDNA 的阳性对照是来自甘油醛磷酸脱氢酶 mRNA (GAPDH) 的扩增。

表 6: 人未患病组织和患病组织中核酸表达的 RT-PCR 分析

样品	患者性别	患者年龄	诊断	T	G	GAPDH	OBcl1	OBcl5	IK2	IK5	DAP3	LOC5	SEC14L2	SSP29	HS16	IK3
肝脏	m	45	未患病组织			+				+						
肝脏	m	27	未患病组织			+	-	-	-	-	-	-	-	-	-	-
肝脏			未患病组织			+				+	-					
HCC1	m	66	小梁/管状HCC	3	1	+	-	+	+							+
HCC2	m	81	小梁/管状HCC	3	2	+	-	+	+	+						
HCC3	m	63	小梁/管状HCC	3	2	+	-	+	-	-						-
HCC4	m	72	小梁/管状HCC	3	2	+	-	+/-								
腺瘤	f	22	良性肝肿瘤			+	-	+	+	+		+		+/-		+
HCC (来自 HCV)	m	63	小梁/管状HCC	2		+	-	+								
HCC cDNA 文库							+	+	+	-		+	+	-	+	-
结肠	m	52	未患病组织			+	-	+	-	-	+	-	-	-	-	-
结肠肿瘤	m	69	管状腺瘤	4	2	+	-	-	-	-	+/-	-	-	+	-	-
结肠肿瘤	m	64	管状腺瘤	3	2	+	-	-	-	+	+	+	+	-	-	-
结肠肿瘤	m	52	管状腺瘤	3	2	+	-	-	-							

样品	患者性别	患者年龄	诊断	T	G	GAPDH	OBcl1	OBcl5	IK2	IK5	DAP3	LOC5	SEC14L	SSP29	HS16	IK3
胃	f	57	未患病组织			+	-	+			+/-					
胃	m	70	未患病组织				-	-	-	-						
胃肿瘤	f	61	腺癌	2		+	-	+			+					
胃肿瘤	f	78	腺癌	3	3	+	+/-	-			-					
胃肿瘤	f	70	管状腺癌	X	3	+	-	-			-					
胃肿瘤	m	69	腺癌	3	3	+	-	-								
胰腺	m	55	未患病组织			+	-	-	-	+	+/-					
胰腺肿瘤	m	69	腺癌	3	3	+	-	+			-					
胰腺肿瘤	m	69	腺癌	3	3	+	-	-								
皮肤	f	60	未患病组织			+	-	-			-					
皮肤肿瘤	m	50	鳞状细胞癌		2	+	-	-								
皮肤肿瘤	f	92	鳞状细胞癌	2	3	+	-	-			-					
皮肤肿瘤	m	73	鳞状细胞癌	2	1	+	-	-			+					
睾丸	m	48	未患病组织			+	-	+	-	-	-					
睾丸肿瘤	m	35	精原细胞瘤和卵黄囊肿瘤	3		+	-	-	-	-	+					
睾丸肿瘤	m	43	精原细胞瘤	2		+	-	-			-					
睾丸肿瘤	m	31	精原细胞瘤	1		+	-	-			-					

样品	患者性别	患者年龄	诊断	T	G	GAPDH	OBcl1	OBcl5	IK2	IK5	DAP3	LOC5	SEC14L2	SSP29	HS16	IK3
甲状腺肿瘤	f	60	乳头状癌	3 a		+	-	-	-	-	+/-	-	-	-	-	-
甲状腺肿瘤	f	57	乳头状癌	4 a		+	-	+	-	-		-	-	-	-	-
甲状腺肿瘤	f	17	乳头状癌	2 b		+	-	-	-	-	+	-	-	-	-	-
肾	f	33	未患病组织			+	-	-	+	-	-	-	-	-	-	-
肾肿瘤	f	33	透明细胞癌	1	1	+	-	-			+					
肾肿瘤	f	62	透明细胞癌	1	1	-	-	+	-	-	-	-	-	-	-	-
肾肿瘤	m	54	透明细胞癌	1	2	+	-	-	+	+	+/-	-	+	-	-	+
肺	f	64	未患病组织			+	-	+	-	-	-	-	-	-	-	-
肺	f	57	未患病组织			+	-	+	-	-	-	-	-	-	-	-
肺肿瘤	m	58	鳞状细胞癌	2	3	+	-	-	+	+	+/-	-	-	-	-	+
肺肿瘤	m	54	鳞状细胞癌	2	2- 3	+	-	+	+/-	-	-	-	-	-	-	+
肺肿瘤	f	57	鳞状细胞癌	2	2- 3	+	-	-	-	-	-	-	-	-	-	-
乳腺	f	38	未患病组织			+	-	-	+/-	-	+/-	-	+	-	-	-
乳腺肿瘤	f	55	侵袭性导管癌 (IDC)	2	2	+	-	-	+	-		-	-	-	-	-



样品	患者性别	患者年龄	诊断	T	G	GAPDH	OBcl1	OBcl5	IK2	IK5	DAP3	LOC5	SECTAL 2	SSP29	HS16	IK3
乳腺肿瘤	f	66	Mucinous 癌	2	1	+	-	+	-	-	-	-	-	-	-	-
脾	f	58	未患病组织			+	-	-	-	-	-	-	-	-	-	-
肌肉	m	65	未患病组织			+	-	-	-	-	-	-	-	-	-	-
脑(皮质)	m	27	未患病组织			+	-	+	-	-	-	-	-	-	-	-
脑髓质	m	27	未患病组织			+	-	-	-	-	-	+	-	+	-	-
心脏			未患病组织								+					+
骨髓			未患病组织													+
胎盘 cDNA 文库			未患病组织			+	+	-	+	+		+	+	+		+

在本发明的另一优选的实施方案中, 根据本发明的核酸可用于构建反义寡核苷酸 (Zheng 和 Kemeny, 1995, Clin. Exp. Immunol. 100: 380-2; Nellen 和 Lichtenstein, 1993, Trends Biochem. Sci. 18: 419-23; Stein, 1992, Leukemia 6: 967-74) 和/或核酶 (Amarzguioui, 等人 1998, Cell. Mol. Life Sci. 54: 1175-202; Vaish 等人, 1998, Nucleic Acids Res. 26: 5237-42; Persidis, 1997, Nat. Biotechnol. 15: 921-2; Couture 和 Stinchcomb, 1996, Trends Genet. 12: 510-5) 和/或小干涉双链 RNA (Elbashir 等人, 2001, Nature 411: 494-98; Brummelkamp 等人, 2002, Science 296: 550-553). 在进一步优选的实施方案中, 通过使用 RNA 干扰分子 (寡核苷酸) 可以降低根据本发明的核酸的稳定性和/或抑制根据本发明的核酸的翻译。从而, 例如, 细胞中相应基因的表达可以体内和体外都降低。因此寡核苷酸可适于作为治疗剂。该策略也适于例如, 肝脏细胞, 尤其如果反义寡核苷酸与脂质体复合的话。优选单链 DNA 或 RNA 用作探针或“反义”寡核苷酸。小干涉 RNA (siRNA) 双链寡核苷酸也适于用作治疗剂。使用该方法, 包括与所要治疗靶定的序列互补的序列的短序列或者 15 到 22 个核苷酸的序列暴露于疾病组织并用于显著降低或者“击倒”治疗靶标 RNA 序列的表达水平。其他疾病中的 siRNA 治疗方法最近也有报导并且也可应用于肝脏失调, 肝癌和其他上皮癌 (Filleur S, Courtin A, Ait-SiAli S, Guglielmi J, Merle C, Harel-Bellan A, Clezardin P, Cabon F. Cancer Res. 2003, 7 月 15 日; 63 (14): 39-22.)。

在优选的实施方案中, 通过筛选文库或者从来自患者或者受试者的样品分离, 用重组方法制备了根据本发明的核酸。在本发明的另一优选的实施方案中, 合成制备了根据本发明的核酸。从而, 例如借助于 SEQ ID 10 到 SEQ ID 19 中描述的 DNA 序列和/或借助于 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 中描述的蛋白质序列参考遗传密码, 例如根据磷酸三酯方法 (see, for example, Uhlmann and Peyman, 1990, Chemical Reviews 90: 543-584) 可以化学合成根据本发明的核酸。

在另一优选的实施方案中, 本发明涉及根据本发明的核酸或者为该核酸的非功能突变变体的核酸或者具有与前述核酸之一互补的序列的核酸, 该核酸通过粘附化学部分而被修饰, 该修饰用以稳定该核酸

以抵抗降解，从而细胞中长期保持高浓度的该核酸 (Beigelman 等人, 1995, *Nucleic Acids Res.* 23: 3989-94; Dudycz, 1995, WO 95/11910; Macadam 等人, 1998, WO 98/37240; Reese 等人, 1997, WO 97/29116)。典型地，可以通过导入一个或多个核苷酸间 (internucleotide) 磷基团或者通过导入一个或多个核苷酸间非磷基团得到这种稳定。

优选的适宜的被修饰的核苷酸间在 Uhlmann 和 Peymann (1990 *Chem. Rev.* 90, 544; 还见 Beigelman 等人, 1995 *Nucleic Acids Res.* 23: 3989-94; Dudycz, 1995, WO 95/11910; Macadam 等人, 1998, WO98/37240; Reese 等人, 1997, WO97/29116) 中概述。

在另一实施方案中，本发明涉及含有根据本发明的核酸的载体和/或其变体，或者为该核酸的非功能突变变体的核酸，或者具有与一种前述核酸互补序列的核酸。载体优选为可应用于基因治疗中的敲除基因构建体、质粒、穿梭载体、噬菌粒、粘粒、病毒载体、表达载体和/或可应用于基因治疗中的载体。这种构建体的制备是本领域技术人员公知的。

本发明意义内的“表达载体”优选包含用于在真核生物中表达的至少一个启动子或增强子，即至少一个含有一种翻译启动信号的调节元件、至少一种根据本发明的核酸或为该核酸的非功能突变变体的核酸、或者具有与一种前述核酸互补序列的核酸、一种翻译终止信号、转录终止信号、和聚腺苷酸化信号。

对于有关基因的表达，通常优选双链 DNA，尤其优选编码多肽的 DNA 区域。对于真核生物的情况，该区域以位于 Kozak 序列 (Kozak, 1987, *Nucleic. Acids Res.* 15: 8125-48) 中的第一个起始密码子 (ATG) 开始直到下一个终止密码子 (TAG、TGA 或 TAA)，该终止密码子位于与 ATG 相同的读框中。对于原核生物的情况，该区域以 SD 序列后的第一个 AUG (或者 GUG) 开始并以下一个终止密码子 (TAA、TAG 或 TGA) 结束，该终止密码子位于与 ATG 相同的读框中。

HCC 中差别表达的基因可以含有肝脏或肝癌基因特异调节序列。这些在组织特异性基因或疾病特异性基因中发现的未转录的序列可用于驱动所包括的治疗和/或肿瘤细胞毒性基因的组织或疾病特异的表达。这些调节序列可用于根据本发明的核酸或者为该核酸的非功能突变变体的核酸，或者具有与前述核酸之一互补序列的核酸的肝癌特异

的表达。这种调节序列的筛选和构建是本领域技术人员公知的。

适宜的表达载体可以是原核或真核表达载体。用于在大肠杆菌 (*E. coli*) 中表达的原核表达载体的实例为, 例如载体 pGEM 或 pUC 衍生载体, 真核表达载体的实例为, 用于在酿酒酵母 (*Saccharomyces cerevisiae*) 中表达的载体例如, 载体 p426Met25 或 p426GAL1 (Mumberg 等人 (1994) *Nucl. Acids Res.*, 22, 5767-5768), 用于在昆虫细胞中表达的为, 例如, 如在 EP-B1-0 127839 或 EP-B1-0 549721 中公开的杆状病毒载体, 和用于在哺乳动物细胞中表达的为例如, 载体 Rc/CMV 和 Rc/RSV 或 SV40 载体, 它们通常都是可得到的。还包括用于在转染后产生 RNA 干扰的特异载体, 如 pSUPER 载体 (Brummelkamp 等人, 2002, *Science* 296: 550-553)。

通常, 表达载体还包括适于各自细胞的启动子, 如, 例如, 用于在大肠杆菌中表达的 *trp* 启动子 (见, 例如, EP-B1-0 154 133)、用于在酵母中表达的 MET 25、GAL1 或 ADH2 启动子 (Rusell 等人 (1983), *J. Biol. Chem.* 258, 2674-2682; Mumberg, 同前)、用于在昆虫细胞中表达的杆状病毒多角体蛋白启动子 (见, 例如, EP-B1-0 127 839)。对于哺乳动物细胞中的表达, 例如, 适宜的启动子为允许真核细胞中组成性的、可调节的、组织特异的、细胞周期特异性的或代谢特异的表达的那些启动子。根据本发明的调节元件为启动子、激活子序列、增强子、沉默子和/或阻遏子序列。

使得在真核生物中组成性表达成为可能的适宜的调控元件的实例优选为被 RNA 聚合酶 III 识别的启动子或者病毒启动子、CMV 增强子、CMV 启动子、SV40 启动子或 LTR 启动子, 它们来自例如 MMTV (小鼠乳腺肿瘤病毒 Lee 等人 (1981) *Nature* 214, 228-232) 和其他病毒启动子和活化序列, 它们来自例如, 腺病毒或类腺病毒、HBV、HCV、HSV、HPV、EBV、HTLV 或 HIV。

使得真核生物中受调节的表达可能的调节元件的实例是四环素操作基因与相应阻遏基因的组合 (Gossen 等人, 1994, *Curr. Opin. Biotechnol.* 5: 516-20)。

翻译起始信号、翻译终止信号、转录终止信号、和聚腺苷酸化信号是本领域技术人员公知的并且可容易地从商业的实验室供应商得到。

优选地，肝脏失调和/或上皮癌相关基因的表达处于组织特异性启动子的控制下，例如，处于肝脏特异性启动子如白蛋白、 $\alpha$ 胎蛋白、载脂蛋白 AI、 $\alpha$ -1 抗胰蛋白酶和补体 C5 和 C8A 基因的控制下 (Schrem 等人, 2002, *Pharmacol. Rev.* 54 129-58; Pontoglio 等人, 2001, *J. Expt. Med.* 194: 1683-1689)。与如此处描述的 HCC 中高度去调节的基因相关的调节序列也提供了在这些失调中特异基因表达的优选方法。

使得在真核生物中组织特异性表达可能的调节元件的其他实例为来自编码仅在某些细胞类型中表达的蛋白质的那些基因的启动子或增强子的启动子序列或活化剂序列。

使得在真核生物中代谢性特异表达可能的调节元件的实例为受缺氧、氧化胁迫、葡萄糖缺乏、磷酸浓度或热休克调节的启动子。

使得在真核生物中细胞周期特异性表达可能的调节元件的实例为下面基因的启动子：cdc25A、cdc25B、cdc25C、细胞周期蛋白 A、细胞周期蛋白 E、cdc2、E2F-1 到 E2F-5、B-myb 或 DHFR (wicker J. 和 Müller R.. 1997, *Trends Genet.* 13: 3-6)。在下面情况尤其优选使用细胞周期调节的启动子：根据本发明的多肽或核酸的表达将被限制于增殖细胞。

为了能够导入如上述的核酸，或者为该核酸的非功能突变变体的核酸并从而能够通过转染、转化或者感染在真核或原核细胞中表达该多肽，该核酸可以作为质粒，病毒或非病毒载体的部分存在。这里适宜的病毒载体尤其为：杆状病毒、牛痘病毒、腺病毒、腺相关病毒、反转录病毒和疱疹病毒。这里适宜的非病毒载体尤其为：病毒体、脂质体、阳离子脂质或聚赖氨酸偶联的 DNA 或裸露的 DNA。

适于根据本发明的用途的质粒、穿梭载体、噬菌粒和粘粒也是本领域技术人员公知的并且通常可从商业的实验室供应商得到。

可应用于基因治疗的载体的实例为病毒载体，例如腺病毒载体、反转录病毒载体或基于 RNA 病毒的复制子的载体 (Lindemann 等人, 1997, *Mol. Med.* 3: 466-76; Springer 等人, 1998, *Mol. Cell.* 2: 549-58, Khromykh, 2000, *Curr. Opin. MolTher.* 2: 555-569)。真核表达载体适于以分离的形式用于基因治疗，因为裸露的 DNA 当局部应用或者通过血液供应时可以例如，穿入活细胞中。

与本领域状况相比，该融合构建体令人惊奇地允许肝脏失调和/或其他上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

在本发明的另一方面还涉及含有根据本发明的核酸和/或其变体的细胞。该细胞优选根据本发明的载体转化。该细胞优选含有核酸，其中该核酸是根据本发明的核酸或者其变体的非功能突变变体。具体地，该细胞含有含有核酸的载体，其中该核酸是根据本发明的核酸或者其变体的非功能突变变体。该细胞优选含有编码具有与根据本发明的核酸或者其变体互补的序列的核酸的核酸。此外，该细胞优选含有含有编码根据本发明的核酸的抗体或者该抗体的片段的核酸的载体。根据本发明的细胞可以例如是肝细胞，含有至少一种上面提到的核酸，或者用一种上面描述的载体转化的细胞。细胞可以是原核或真核细胞、异体或自体细胞。原核细胞的实例为大肠杆菌，真核细胞的实例包括原代肝细胞、肝细胞细胞系如 HepG2 和 Hep3B 细胞、酵母细胞，例如酿酒酵母或昆虫细胞。

与本领域状况相比，根据本发明的细胞令人惊奇地允许肝脏失调和/或其他上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

在本发明的优选的实施方案中，细胞为转基因的胚胎非人干细胞，其含有如上述的至少一种根据本发明的核酸、至少一种载体、至少一种敲除基因构建体和/或至少一种表达载体。

细胞和/或干细胞的转化方法是本领域技术人员熟知的并且包括，例如，电穿孔或微注射。

本发明的另一方面涉及转基因非人哺乳动物的提供，该哺乳动物含有化合物，所述化合物选自根据本发明的核酸和/或其变体、为该核酸的非功能突变变体的核酸、具有与上述核酸之一互补的序列的核酸、以载体、击倒或敲除基因构建体和表达载体形式存在的上述核酸之一。

转基因动物通常显示出核酸和/或多肽的组织特异性增加的表达并且可用于肝脏失调和/或上皮癌，如例如 HCC 的分析，以及用于这些失调的治疗策略的开发和评价。转基因动物还可以用于产生根据本发明的多肽。该动物产生的多肽可以例如在动物的体液中富集。根据

本发明的多肽可以例如从体液如奶中分离。

与本领域状况相比，该转基因非人哺乳动物令人惊奇地允许肝脏失调和/或其他上皮癌的改良的、更敏感、更早、更快、和/或非侵入性分析和/或诊断。

制备转基因动物，尤其是转基因小鼠的方法同样是本领域技术人员从 DE 196 25 049 和 US 4.736.866、US 5.625.122、US 5.698.765、US 5.583.278 和 US 5.750.825 公知的并包括可以例如，通过将根据本发明的表达载体直接注射到胚胎或精母细胞或通过表达载体注射到受精卵的原核或者通过将表达载体转染到胚胎干细胞或通过核转移到适宜的受体细胞中产生转基因动物 (Polites 核 Pinkert, DNA 微注射核转基因动物产生, 15 到 68 页 Pinkert, 1994, 转基因动物技术: 实验室手册 (Transgenic animal technology: a laboratory handbook), Academic Press, 伦敦, UK; Houdebine, 1997, Harwood Academic Publishers, 阿姆斯特丹, 荷兰; Doetschman, 胚胎干细胞中的基因转移, 115 到 146 页, Pinkert, 1994, 如前; Wood, 反转录病毒介导的基因转移, 147 到 176 页, Pinkert, 1994, 如前; Monastersky, 基因转移技术、备选技术和应用 (Gene Transfer Technology; Alternative Techniques and Applications), 177 到 220 页, Pinkert, 1994, 如前)。

如果上述核酸被整合到所谓的“靶定载体”或“敲除”基因构建体 (Pinkert, 1994, 如前), 那么胚胎干细胞转染和同源重组后可能, 例如, 产生敲除小鼠, 其通常作为杂合小鼠, 显示出核酸的降低的表达, 而纯合小鼠不再表现出该核酸的表达。这样产生的动物也可用于肝脏失调, 如例如 HCC, 和/或上皮癌的分析。

敲除基因构建体是本领域技术人员可以从美国专利 5.625.122、US 5.698.765、US 5.583.278 和 US 5.750.825 公知的。

本发明的另一方面涉及抗体或其片段的提供, 其中该抗体或抗体片段针对根据本发明的多肽、其功能变体或针对编码该多肽, 或者其变体的核酸。

与本领域状况相比, 这些抗体或其片段令人惊奇地允许肝脏失调和/或其他上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断和/或改良的、持续的和/或更有效的治疗。

术语“抗体”或“抗体片段”根据本发明还被理解为通过遗传工程制备和任选被修饰的抗体或其抗原结合部分，如，例如，嵌合抗体、人源化抗体、多功能抗体、双特异性或寡特异性抗体、单链抗体、F(ab)或F(ab)<sub>2</sub>片段（见，例如，EP-B 1-0 368 684、US 4.816. 567、US 4.816. 397、W088/01649、W093/06213、W0 98/24884）。根据本发明的抗体可以例如用于根据本发明的失调如肝脏失调，例如 HCC，和/或上皮癌的预防和/或治疗。

本发明还涉及例如用于根据本发明的失调的诊断和/或预防和/或治疗的对根据本发明的核酸或者其变体编码的多肽或其功能变体特异的抗体或抗体片段，优选多克隆或单克隆抗体的产生方法。根据本领域技术人员公知的方法，通过用根据本发明的核酸或它们的变体，或者用长度至少为 6 个氨基酸，优选至少 8 个氨基酸，尤其至少 12 个氨基酸的根据本发明的多肽或其部分或者其功能变体，如合适在例如，弗氏佐剂和/或氢氧化铝的存在下，免疫哺乳动物，例如，兔，来实施该方法（见，例如，Harlow 和 Lane, 1998, 使用抗体：实验室手册（Using Antibodies: A Laboratory Manual），冷泉港出版社，纽约，美国，第 5 章，53-135 页）。由于免疫学反应导致在动物中形成的多克隆抗体用根据公知方法容易地从血液分离并例如，通过柱层析纯化。可以根据 Winter 和 Milstein（Winter 和 Milstein, 1991, Nature 349: 293-299）的公知方法制备单克隆抗体。

本发明还涉及针对上述多肽并与该多肽特异反应的抗体或抗体片段，其中上述多肽的部分或者自身是免疫原性的或者可以通过偶联适宜的载体，如，例如，牛血清白蛋白或者锁眼帽贝血蓝蛋白以增加它们的免疫原性而获得免疫原性。该抗体是多克隆的或者单克隆的；优选其是单克隆抗体。

此外，本发明涉及从如例如，Knappik 等人（2000, J. Molec. Biol. 296: 57-86）或者 Chadd 和 Chamow（2001 Curr. Opin. Biotechnol. 12: 188-94）重组载体表达文库产生和/或制备对根据本发明的多肽特异的抗体或抗体片段。

在本发明的另一实施方案中，提供了阵列，其中该阵列含有选自根据本发明的多肽、其功能变体、编码该多肽的核酸、该核酸的非功能突变变体和针对该多肽的抗体或抗体片段的至少两种化合物。备选



地，该阵列含有至少一种根据本发明的组分与以前描述的与瘤性或代谢性肝脏失调或上皮癌相关的组分的组合。

在本发明的意义内，术语“阵列”指固相或凝胶样载体，其上粘附或结合以一维、二维或三维排列的至少两种化合物。这种阵列是本领域技术人员公知的并且通常在显微镜载玻片，特别是有涂层载玻片如涂有聚阳离子、硝酸纤维素或生物素的载玻片，盖片、和膜例如基于硝酸纤维素或尼龙的膜上产生。

上面提到的阵列包括根据本发明的结合多肽或其功能变体或者编码该多肽的核酸或其变体、根据本发明的融合蛋白或者针对根据本发明的多肽或其变体的抗体或抗体片段或者表达根据本发明的多肽或者其功能变体的细胞或者表达根据本发明的至少一种核酸或者其变体的至少两种细胞。编码这些或者其变体的核酸也可以是阵列的部分。这种阵列可用于肝脏失调，优选 HCC，和/或上皮癌的分析 and/或诊断。

本发明还涉及产生根据本发明的阵列的方法，其中根据本发明的至少两种化合物结合到载体材料。

产生（例如，基于固相化学和光不稳定的保护基团）这种阵列的方法是公知的（US 5,744,305）。这种阵列也可以与物质或者物质文库接触并试验相互作用，例如结合或构象的改变。

本发明还涉及制备固定在载体材料上的阵列的方法，该阵列用于根据本发明的失调如肝脏失调，优选 HCC 的分析 and/或诊断，其中如上述的至少两种核酸、至少两种多肽或至少两种抗体或抗体片段，和/或至少两种细胞，或者至少一种上述组分与其他与瘤性和代谢性肝脏失调或上皮癌相关的组分的组合，被用于制备。通过这种方法产生的阵列可用于根据本发明的失调的诊断。

本发明的另一方面涉及诊断试剂，其含有至少一种化合物、所述化合物选自根据本发明的多肽、或者其功能变体、编码该多肽的核酸、优选根据 SEQ ID 10 到 19 的核酸、或一种上面提到的核酸的变体、和根据本发明的抗体或抗体片段，与适宜的添加剂或助剂组合或在一起。

与本领域状况相比，该诊断试剂令人惊奇地允许肝脏失调和/或其他上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断。

在本发明的意义内，“适宜的添加剂”或“助剂”是本领域技术

人员公知的并且包括，例如，基于生理盐溶液、软化水、明胶或甘油的蛋白质稳定试剂。备选地，根据本发明的核酸或多肽可以为了稳定而被冻干。

在另一实施方案中，可以产生基于根据本发明的核酸序列的诊断试剂盒。这种试剂盒可被特别设计以检测由于所描述的失调而导致循环系统细胞改变并从而可以在来自受试患者的血清中被检测到。诊断试剂盒的其他实例包括酶联免疫吸附测定法 (ELISA)、放射免疫测定法 (RIA)，和特定抗体与根据本发明的多肽的免疫反应的检测，包括特定应答免疫细胞的检测。

在优选的实施方案中，根据本发明的诊断试剂含有探针，优选 DNA 探针。

例如，可能根据本发明制备基于聚合酶链式反应 (PCR) 的诊断试剂。在限定的条件下，优选使用对根据本发明的核酸特异的引物作为 DNA 探针，对本发明的核苷酸序列特异的 PCR 将用于监视为了诊断或治疗目的从患者分离的样品中根据本发明的特定核酸的存在，特别是量。这使得进一步可能通过适宜的探针，从适宜的基因或 cDNA 文库，例如，从肝脏失调特异的或肝脏特异的基因库分离得到所描述的核酸 (见，例如 J. Sambrook 等人，1989，分子克隆：实验室手册第二版，冷泉港实验室，冷泉港，NY，第 8 章，8.1 到 8.81 页，第 9 章，9.47 到 9.58 页和第 10 章 10.1 到 10.67 页)。

适宜的探针为，例如，长度为约 50-1000 个核苷酸，优选约 10 到约 100 个核苷酸，优选约 100 到约 200 个核苷酸，尤其约 200-500 个核苷酸的 DNA 或 RNA 片段，其序列可来自根据 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 的多肽，和其功能变体，和编码该多肽优选根据 SEQ ID 10 到 SEQ ID 19 的核酸，和其变体。

备选地，优选可能通过衍生的核酸序列合成适于用作聚合酶链式反应的引物的寡核苷酸。使用该寡核苷酸，上述核酸或其部分可被扩增并从 cDNA，例如，HCC-特异的 cDNA 分离。适宜的引物为，例如，长度为约 10 到 100 个核苷酸，优选约 15 到 50 个核苷酸，尤其为 17 到 30 个核苷酸的 DNA 片段，其序列可以来自根据 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 的多肽，这些多肽来自根据 SEQ ID 10 到 SEQ ID 19 的核酸。这些引物的设计和合成是本领域技术人员公知的。该引物可

以还含有限制位点，例如，本领域技术人员公知的适于将扩增的序列整合到载体的限制位点，或者其他衔接或突出端序列，例如具有标记分子如粘附荧光标记的序列。

在本发明的另一方面提供了根据本发明的失调的诊断方法，其中至少一种选自根据序列 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 的多肽、其功能变体、编码该多肽的核酸、前述核酸之一的变体和针对该多肽的抗体或其抗体片段的化合物在患者的样品中被鉴定并且与参比文库或参比样品的至少一种化合物比较。

在本方法的优选实施方案中，肝脏的失调是选自肝硬变、酒精肝疾病、慢性肝炎、威尔逊病、血色素沉着、肝细胞癌、良性肝肿瘤和病灶性结节性增生症的失调。

在本发明的优选实施方案中，上皮癌是不同于肝脏的器官的腺癌，该器官优选为选自肺、胃、肾、结肠、前列腺、皮肤和乳腺的器官。

与本领域的状况相比，该诊断试剂令人惊奇地允许肝脏失调和/或其他上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断。

优选通过如上述的非侵入性方法从患者分离样品。

例如，通过 ELISA 测定法的特定去调节的基因蛋白质的血清检测是一种应用，备选一种或一组去调节的基因产物的抗体，基于患病组织或来自患病个体的血清中表达的基因产物的表达水平的组合从该去调节的基因产物推出诊断得分。

根据本发明的优选的诊断试剂含有所描述的多肽或者上面更详细描述的多肽的免疫原性部分。该多肽或者其部分优选结合例如，硝酸纤维素或尼龙的固相，可以例如与所研究的体液，例如，血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液体外接触，从而能够例如与例如患者的血液中存在的自身免疫抗体反应。然后，例如，通过标记的抗人 IgG 抗体检测抗体-肽复合物。标记包括，例如，酶，如过氧化物酶，其催化变色或化学发光反应。然后可容易而快速地通过该颜色检测自身免疫抗体的存在和所存在的量。

此外，诊断试剂可以用于检测从患者分离的样品中存在的内源抗体或者其片段，该抗体或者其片段针对根据本发明的多肽。这种自身

免疫抗体的检测可以通过本领域公知的方法实现，例如通过使用根据本发明的多肽或者其功能变体或者其部分作为探针的免疫亲和测定法实现。优选地，这种自身免疫抗体的存在表明该患者患有根据本发明的失调。

另一种诊断试剂是本发明的主题，该诊断试剂含有根据本发明的抗体自身。通过这些抗体，可能，例如，容易而快速地研究组织样品关于所涉及的根据本发明的多肽是否以增量存在，以便从而得到包括肝脏失调，例如 HCC 的可能的疾病的指征。在这种情况下，根据本发明的抗体优选被直接例如用上面已经描述的酶或荧光分子标记，或者更普遍地例如，用特异的第二抗体间接检测这些抗体。从而该特定抗体-肽复合物可例如，通过酶促变色反应容易且快速地检测。

在本发明另一方面，提供了鉴定相对于参比文库或参比样品，在来自患者的样品中差别表达的至少一种根据 SEQ ID 10 到 SEQ ID 19 的核酸，或者其变体的方法，所述方法包括下面的步骤：

- (a) 检测从患者分离的样品中至少一种根据 SEQ ID 10 到 SEQ ID 19 的核酸，或者其变体的表达，
- (b) 将步骤 (a) 中检测的所述核酸的表达与参比文库或参比样品中相同核酸的表达比较，
- (c) 鉴定与参比文库或参比样品相比，在从患者分离的样品中差别表达的所述核酸。

与本领域状况相比，该方法令人惊奇地允许根据本发明的核酸的差别表达的改良的、更敏感、更早、更快和/或非侵入性鉴定，该鉴定为根据本发明的失调的诊断提供了有用的基础。

优选鉴定至少 2 种、至少 3 种、至少 4 种、至少 5 种、至少 6 种，或者至少 7 种核酸。

在该方法的另一个优选实施方案中，所述一种或几种核酸通过基于 PCR 的检测或者通过杂交测定法检测。

在该方法的另一优选的实施方案中，通过选自基于固相的筛选方法、杂交、扣除杂交、差别展示和 RNA 酶保护测定法的方法比较所述核酸的表达。

在该方法另一优选的实施方案中，从患者分离的样品选自肝脏组织、肝脏细胞、来自受到癌性转化的另一器官的组织、来自该器官的

细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液、和粪便。

参比样品优选从选自相同患者的未患病样品或者来自另一受试者的非患病样品的来源分离。适宜参比样品的选择是本领域技术人员公知的。具体地，参比样品可选自肝脏组织、肝脏细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液、和粪便。

在该方法的另一优选的实施方案中，参比文库是表达文库或者数据库，其含有关于样品中至少一种根据本发明的核酸的非患病表达的克隆或数据，所述样品优选可以选自肝脏组织、肝脏细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液、和粪便。

在本发明的另一方面，提供了肝脏失调，和/或另一上皮癌的诊断方法，该方法包括下面的步骤：

- a) 检测从患者分离的样品中至少一种根据 SEQ ID 10 到 SEQ ID 19 和/或 SEQ ID 47 的核酸，或者其变体的表达，
- b) 将步骤 (a) 中检测的所述核酸的表达与参比文库或参比样品中相同核酸的表达比较，
- c) 鉴定与参比文库或参比样品相比，在从患者分离的样品中差别表达的所述核酸，
- d) 将步骤 (c) 中鉴定的所述核酸与病理学参比样品或病理学参比文库中差别表达的所述核酸匹配，

其中所匹配的一种或几种核酸指出该患者患有肝脏失调，和/或其他上皮癌。

与本领域的状况相比，该诊断方法令人惊奇地允许肝脏失调和/或其他上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断。

优选鉴定了至少 2 种、至少 3 种、至少 4 种、至少 5 种、至少 6 种、或至少 7 种核酸。

在该方法的另一优选的实施方案中，所述一种或几种核酸通过基于 PCR 的检测或者通过杂交测定法检测。

在该方法的另一优选的实施方案中，通过选自基于固相的筛选方法、杂交、扣除杂交、差别展示和 RNA 酶保护测定法的方法比较所述核酸的表达。

在该方法另一优选的实施方案中，从患者分离的样品选自肝脏组织、肝脏细胞、来自受到癌性转化的另一器官的组织、来自该器官的细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液、和粪便。

参比样品优选从选自相同患者的未患病样品或者来自另一受试者的非患病样品的来源分离。适宜参比样品的选择是本领域技术人员公知的。具体地，参比样品可选自肝脏组织、肝脏细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液、和粪便。

在该诊断方法的另一优选的实施方案中，参比文库是表达文库或者数据库，其含有关于样品中至少一种根本发明的核酸的非患病表达的克隆或数据，所述样品优选可以选自肝脏组织、肝脏细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液、和粪便。

在该诊断方法的另一优选的实施方案中，病理学参比样品分离自另一患者的患病样品。后一患者已经被诊断为含有将要被诊断的根据本发明的失调。适宜的病理学参比样品的选择是本领域技术人员公知的。具体地，病理学参比样品可以选自肝脏组织、肝脏细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液、和粪便。

在该诊断方法的另一优选的实施方案中，病理学参比文库为数据库，该数据库含有关于相对于参比样品或参比文库中的对照表达，从患有将要在该发明性方法中诊断的根据本发明的失调的至少一个患者（排除处于诊断中的患者）分离的样品中至少一种根据本发明的核酸的差别表达的数据。该病理学参比数据库优选还涉及差别表达文库，该文库含有相对于参比样品或参比文库中的对照表达，从患有将要在该发明方法中诊断的根据本发明的失调的至少一个患者（排除处于诊断中的患者）分离的样品中差别表达的根据本发明的核酸。适宜的病理学参比文库的选择是本领域技术人员公知的。

优选地，该肝脏失调是选自肝硬变、酒精肝疾病、慢性肝炎、威尔逊病、血色素沉着、肝细胞癌、良性肝肿瘤和病灶性结节性增生症的失调。具体地，上皮癌是不同于肝脏的器官的腺癌，该器官优选选自肺、胃、肾、结肠、前列腺、皮肤和乳腺。

术语“检测核酸”在本发明的意义内指优选从样品中存在的其他组分的背景中揭露、显示、分开或允许识别根据本发明的核酸的方法。

这些方法是本领域技术人员公知的并且包括基于原位杂交、PCR 扩增、凝胶电泳、RNA 印迹、固相阵列（基因芯片）的方法、核酸酶保护方法（如在 Alberts, 等人（2002）细胞分子生物学，第四版（Garland, 纽约，美国）中描述和引用的）。

术语“将步骤（a）中检测的所述核酸的表达与参比文库或参比样品中相同核酸的表达比较”在本发明的意义内指通过实验方法如差别展示、扣除杂交、RNA 酶保护测定法，或者特别是 DNA 芯片杂交在定量或定性水平上比较两组所述核酸的表达。而且，这里还包括比较步骤（a）中检测的所述核酸的实验数据与参比文库中相同核酸的表达。

术语“鉴定与参比文库或参比样品相比，在从患者分离的样品中差别表达的所述核酸”在本发明的意义内被理解为表示选择与参比文库或参比样品相比差别表达的所述核酸，该选择满足下面的标准：与参比文库或参比样品相比所检测的核酸的差别表达水平被上调高约 2 倍，优选高约 5 倍，更优选高约 10 倍。

术语“将步骤（c）中鉴定的所述核酸与病理学参比样品或病理学参比文库中差别表达的所述核酸匹配”在本发明的意义内被理解为指步骤（c）中鉴定的所述核酸与病理学参比样品或病理学参比文库中差别表达的所述核酸相比。然后，也在病理学参比样品或病理学参比文库中差别表达的步骤（c）中鉴定的所述核酸被匹配，即所述相同对被鉴定和分配。因为病理学参比样品或病理学参比文库中所述核酸的差异表达表明根据本发明的失调，所以与样品中差异表达的这种对应指出该患者患有该失调。

优选地，通过如上述的非侵入性或优选最小侵入性方法（包括静脉穿刺）从患者分离样品。

根据本发明的诊断方法允许肝脏失调和/或上皮癌的更早检测，和/或该失调的非侵入性诊断，该检测或诊断基于相对于参比样品和/或参比文库，在从患有肝脏失调和/或上皮癌的动物和/或人类患者分离的样品中检测的根据本发明的核酸的基本上一致的表达模式。该方法的额外优点是其还提供了表征肝脏失调的不同亚型，如例如 HCC 的亚型的新的诊断参数。

术语根据本发明的核酸的“基本一致的表达模式”指如果所比较

的患者或受试者分别处于相同的或者相当的病理状况或者健康状况下，从患者到患者或者受试者到受试者基本上可以再现的表达模式。

在本发明的另一方面提供了鉴定至少一种根据 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 的多肽，或者其功能变体的方法，相对于参比文库或参比样品，在来自患者的样品中该多肽或其变体差别表达，所述方法包括下面的步骤：

- (a) 检测从患者分离的样品中至少一种根据 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 的多肽，或者其功能变体的表达，
- (b) 将步骤 (a) 中检测的所述多肽的表达与参比文库或参比样品中所述多肽的表达比较，
- (c) 鉴定与参比文库或参比样品相比，在从患者分离的样品中差别表达的所述多肽。

与本领域状况相比，该方法令人惊奇地允许根据本发明的多肽的差别表达的改良的、更敏感、更早、更快和/或非侵入性鉴定，该鉴定为根据本发明的失调的诊断提供了有用的基础。

优选鉴定至少 2 种、至少 3 种、至少 4 种、至少 5 种、至少 6 种，或者至少 7 种多肽。

优选地，通过如上述的非侵入性或最小侵入性方法（包括静脉穿刺）从患者分离样品。

在该方法的另一实施方案中，样品为如更上面定义的产品。参比样品优选为如上定义的参比样品。

在该方法的另一优选实施方案中，参比文库是表达文库或者数据库，其含有关于样品中至少一种根据本发明的多肽的非患病表达的克隆或数据，所述样品优选可以选自肝脏组织、肝脏细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液、和粪便。这种数据库作为根据本发明的 cDNA 微阵列表达分析的结果而产生并且是本领域技术人员公知的。上面已经描述了可以根据本发明使用的其他参比文库。

在本发明的另一方面，提供了诊断肝脏失调和/或上皮癌的方法，该方法包括步骤：

- a) 检测从患者分离的样品中至少一种根据 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 的核酸，和/或其功能变体的表达，



- b) 将步骤(a)中检测的所述多肽的表达与参比文库或参比样品中所述多肽的表达比较,
- c) 鉴定与参比文库或参比样品相比,在从患者分离的样品中差别表达的所述多肽,
- d) 将步骤(c)中鉴定的所述多肽与病理学参比样品或病理学参比文库中差别表达的所述多肽匹配,

其中所匹配的一种或几种多肽指出该患者患有肝脏失调,和/或其他上皮癌。

与本领域的状况相比,该诊断方法令人惊奇地允许肝脏失调和/或其他上皮癌的改良的、更敏感、更早、更快、和/或非侵入性的诊断。

优选鉴定了至少2种、至少3种、至少4种、至少5种、至少6种、或至少7种多肽。

在本发明的意义内术语“检测多肽”指优选从样品中存在的其他组分的背景中揭露、显示、分开和/或允许识别根据本发明的多肽的方法。这些方法是本领域技术人员公知的并且包括凝胶电泳、层析技术、免疫印迹分析、免疫组织化学、基于酶的免疫测定法、质谱、高压液相色谱、表面胞质团共振,和/或如上述的抗体和蛋白质阵列(Ausubel, F. A. 等人, 编辑, 1990, 当前分子生物学方案(Current Protocols in Molecular Biology), Greene Publishing and Wiley-Interscience, 纽约, USA, 第10章; Myszka 和 Rich 2000, Pharm. Sci. Technol. Today 3: 310-317)。优选地,通过例如用物理剪切或超声方法破碎细胞从样品制备蛋白质和多肽。蛋白质用还原剂处理和加热变性和稳定化并且该蛋白质在电泳聚丙烯酰胺凝胶上按大小分级分离。

在本发明的意义内术语“将步骤(a)中检测的所述多肽的表达与参比文库或参比样品中相同多肽的表达比较”指通过实验方法如二维凝胶电泳、层析分离技术、免疫印迹分析、表面胞质团共振、免疫组织化学、和基于酶的免疫测定法在定量或定性水平上比较两组所述多肽的表达。在二维凝胶电泳中,根据本领域中技术人员熟知的方法,所有多肽首先根据第一电泳维中的等电点分离然后通过大小分离。此外,还包括将步骤1中所检测的至少一种多肽的数据与如上定义的参

比文库中多肽的表达比较。

术语“鉴定与参比文库或参比样品相比，在从患者分离的样品中差别表达的所述多肽”在本发明的意义内被理解为表示选择与参比文库或参比样品相比差别表达的所述多肽，该选择满足下面的标准：与参比文库或参比样品相比所检测的多肽的差别表达水平被上调高约2倍，优选高约5倍，更优选高约10倍。

术语“将步骤(c)中鉴定的所述多肽与病理学参比样品或病理学参比文库中差别表达的所述多肽匹配”在本发明的意义内被理解为指步骤(c)中鉴定的所述多肽与病理学参比样品或病理学参比文库中差别表达的所述多肽相比。然后，也在病理学参比样品或病理学参比文库中差别表达的步骤(c)中鉴定的所述多肽被匹配，即所述相同对被鉴定和分配。因为病理学参比样品或病理学参比文库中所述多肽的差异表达表明根据本发明的失调，所以与样品中差异表达的这种对应指出该患者患有该失调。

优选地，通过如上述的非侵入性或优选最小侵入性方法（包括静脉穿刺）从患者分离样品。

在该方法的另一实施方案中，样品为更上面定义的产品。参比样品优选为如上定义的参比样品。

在该诊断方法的另一优选的实施方案中，参比文库是表达文库或者数据库，其含有关于样品中至少一种根本发明的核酸的非患病表达的克隆或数据，所述样品优选可以选自肝脏组织、肝脏细胞、血液、血清、血浆、腹水、胸膜积液、脑脊液、唾液、尿、精液、和粪便。

根据本发明的数据库的例子以及根据本发明可用的实验参比文库如上所述。

在本发明诊断方法的另一优选实施方案中，病理学参比样品是如上所述的参比样品。

在该诊断方法的另一优选的实施方案中，病理学参比文库为数据库，该数据库含有关于相对于参比样品或参比文库中的对照表达，从患有将要在该发明方法中诊断的根据本发明的失调的至少一个患者（排除处于诊断中的患者）分离的样品中根据本发明的多肽的差别表达的数据。该病理学参比数据库还涉及差别表达文库，该文库含有相对于参比样品或参比文库中的对照表达，从患有将要在该发明方法中

诊断的根据本发明的失调的至少一个患者（排除处于诊断中的患者）分离的样品中差别表达的根据本发明的多肽。适宜的病理学参比文库的选择是本领域技术人员公知的。

优选地，该肝脏失调是选自肝硬化、酒精肝疾病、慢性肝炎、威尔逊病、血色素沉着、肝细胞癌、良性肝肿瘤和病灶性结节性增生症的失调。具体地，上皮癌是不同于肝脏的器官的腺癌，该器官优选选自肺、胃、肾、结肠、前列腺、皮肤和乳腺。

根据本发明的诊断方法允许肝脏失调和/或上皮癌的更早检测，和/或该失调的非侵入性诊断，该检测或诊断基于相对于参比文库，在从患有肝脏失调和/或上皮癌的动物和/或人类患者分离的样品中检测的根据本发明的多肽的基本上一致的表达模式。该方法的额外优点是其还提供了表征肝脏失调的不同亚型，如例如 HCC 的亚型的新的诊断参数。

术语根据本发明的多肽的“基本一致的表达模式”指如果所比较的患者或受试者分别处于相同的或者相当的病理状况或者健康状况下，从患者到患者或者受试者到受试者基本上可以再现的表达模式。

在本发明的另一方面，提供了含有选自根据本发明的多肽、其功能变体、编码前述多肽之一的核酸、前述核酸之一的变体、为前述核酸之一的非功能突变变体的核酸、具有与前述核酸之一互补的序列的核酸、含有前述核酸之一的载体、含有前述核酸之一的细胞、含有前述载体的细胞、针对前述多肽之一的抗体或该抗体的片段、含有编码前述抗体的核酸的载体、含有含有编码前述抗体的核酸的载体的细胞和含有含有编码前述抗体片段的核酸的载体的细胞中的至少一种化合物的药物组合物，其与适宜的添加剂或助剂组合或在一起。在优选的实施方案中，该药物组合物含有至少一种根据本发明的细胞，其与适宜的添加剂或助剂组合或在一起。

当与本领域肝脏失调和/或其他上皮癌的疗法的状态相比时，根据本发明的药物组合物令人惊奇地提供了改良的、持续的和/或更有效的治疗。

按本发明理解的药物组合物包括可用于预防和/或治疗肝脏失调和/或上皮癌的药物。药物组合物包括，例如，通过细胞系统，优选真核系统中的特异抗体基因片段的表达产生的稳定的重组抗体。重组

抗体治疗剂例如，通过注射到患病肝脏区或静脉或动脉血管系统中或者肝门静脉而被递送。可以在规则的时间间隔重复注射以实现治疗效能。根据本发明的治疗剂还可以与其他化学品、抗体或其他治疗应用组合使用以提高效能。

本发明还涉及产生用于根据本发明的失调，例如，HCC 的治疗和/或预防的药物组合物的方法，其中选自根据本发明的多肽、其功能变体、编码前述多肽之一的核酸、前述核酸之一的变体、为前述核酸之一的非功能突变变体的核酸、具有与前述核酸之一互补的序列的核酸、含有前述核酸之一的载体、含有前述核酸之一的细胞、含有前述载体的细胞、针对前述多肽之一的抗体或该抗体的片段、含有编码前述抗体的核酸的载体、含有含有编码前述抗体的核酸的载体的细胞和含有含有编码前述抗体片段的核酸的载体的细胞中的至少一种组分与适宜的添加剂组合或混合。

本发明还涉及通过用于根据本发明的失调，例如，HCC 的治疗和/或预防的方法产生的药物组合物，其含有选自根据本发明的多肽、其功能变体、编码前述多肽之一的核酸、前述核酸之一的变体、为前述核酸之一的非功能突变变体的核酸、具有与前述核酸之一互补的序列的核酸、含有前述核酸之一的载体、含有前述核酸之一的细胞、含有前述载体的细胞、针对前述多肽之一的抗体或该抗体的片段、含有编码前述抗体的核酸的载体、含有含有编码前述抗体的核酸的载体的细胞和含有含有编码前述抗体片段的核酸的载体的细胞中的至少一种组分，如果适宜，该组分与适宜的添加剂或助剂在一起。本发明还涉及该药物组合物的用途，用于肝脏失调，例如，HCC 和上皮癌的预防和/或治疗。

优选地，该药物组合物用于治疗选自肝硬化、酒精肝疾病、慢性肝炎、威尔逊病、血色素沉着、肝细胞癌、良性肝肿瘤和病灶性结节性增生症的失调。具体地，该药物组合物用于治疗上皮癌，该上皮癌是不同于肝脏的器官的腺癌，该器官优选为选自肺、胃、肾、结肠、前列腺、皮肤和乳腺的器官。

也可以以本领域技术人员公知的常规方式，例如，通过口服应用或者静脉内注射根据本发明的药物组合物实施治疗。从而可能施用含有适宜的添加剂或助剂，如，例如，生理盐溶液、软化水、稳定剂、

蛋白酶抑制剂的药物组合物。

通过使用自体或异体细胞可以实现治疗，该治疗基于细胞的使用，该细胞表达至少一种根据本发明的多肽、其功能变体或编码该多肽的核酸，或者其变体。优选的细胞包括肝细胞，例如，肝细胞的原代培养物、肝脏定居干细胞或祖细胞，或者血细胞。该细胞可与适宜的载体材料应用于组织，优选血液或者注射到肝脏。这种疗法优选基于如下观念：当根据本发明的多肽表达和/或释放时，该多肽刺激需要该治疗的患者中的免疫应答。

优选地，该治疗方法抑制根据本发明的至少一种多肽的功能和/或表达和/或根据本发明的至少一种核酸的功能和/或表达。表达和/或功能的这种抑制优选显著降低了所靶定核酸/多肽的表达和/或功能。该表达和/或功能的抑制优选消除了所靶定核酸和/或多肽的表达和/或功能。所靶定核酸和/或多肽的表达和/或功能的降低或消除可以使用本领域技术人员公知的用于确定多肽/核酸的表达和/或功能的常规测定法确定。具体地，确定该功能的这种测定法包括比较药物组合物施用之前和之后所靶定的核酸和/或多肽的生物学活性的方法。优选地，用于确定表达的这种测定法包括比较药物组合物施用之前和之后所靶定的核酸和/或多肽的表达水平的方法。

优选通过使用具有与根据本发明的核酸之一互补的序列的核酸完成这种治疗，即，降低或消除根据本发明的转录的核酸的翻译从而抑制所靶定核酸/多肽的功能和/或表达的反义分子或 RNA 干扰分子。优选地，具有互补序列的这种核酸可以以载体或者含有这种核酸的细胞的形式使用。在多肽水平上，可以尤其通过使用针对根据本发明的多肽的抗体或抗体片段实施治疗。该抗体或抗体片段可以直接施用于患者或者优选地编码该抗体的核酸包含在载体中，该载体优选包含在细胞中。该细胞或者载体然后可以施用于需要这种治疗的患者。

当与肝脏失调，和/或其他上皮癌的疗法的本领域状态相比时，根据本发明的治疗方法令人惊奇地提供了改良的、持续的和/或更有效的治疗。

本发明还涉及治疗肝脏失调患者的方法，其中选自根据本发明的多肽、其功能变体、编码该多肽的核酸、前述核酸之一的变体、为前述核酸之一的非功能突变变体的核酸、具有与前述核酸之一互补的序

列的核酸、含有前述核酸之一的载体、含有前述核酸之一的细胞、含有前述载体的细胞、针对前述多肽之一的抗体或该抗体的片段、含有编码前述抗体的核酸的载体、含有含有编码前述抗体的核酸的载体的细胞和含有含有编码前述抗体片段的核酸的载体的细胞中的至少一种组分，任选与适宜的添加剂和/或助剂组合，以治疗有效量被施用于需要这种治疗的患者。

优选地，该治疗方法针对选自肝硬变、酒精肝疾病、慢性肝炎、威尔逊病、血色素沉着、肝细胞癌、良性肝肿瘤和病灶性结节性增生症的失调。该治疗方法尤其针对上皮癌，该上皮癌是不同于肝脏的器官的腺癌，该器官优选为选自肺、胃、肾、结肠、前列腺、皮肤和乳腺的器官。

使用这种化合物或细胞的方法已经在上面详细描述。

术语“治疗有效量”指对患者施用导致如上定义的“有效治疗”的化合物的量。化合物的治疗有效量的确定是本领域技术人员公知的。

这种治疗方法允许如上描述的肝脏失调和/或上皮癌的有效治疗。

在本发明的另一方面提供了刺激患有肝脏失调和/或上皮癌的患者对根据本发明的多肽，或者其功能变体的免疫应答的方法，其中选自根据本发明的多肽、其功能变体、编码前述多肽之一的核酸、前述核酸之一的变体、含有前述核酸之一的载体、含有前述核酸之一的细胞、含有前述载体的细胞中的至少一种组分以有效刺激该患者中免疫应答的量施用于需要这种治疗的患者。

当与肝脏失调和/或其他上皮癌的疗法的本领域状况相比时，根据本发明的刺激免疫应答的方法令人惊奇地提供了改良的、持续的和/或更有效的免疫。

在本发明的另一方面，提供了防止患者发生肝脏失调和/或上皮癌的方法，其中选自根据本发明的多肽、其功能变体、编码前述多肽之一的核酸、前述核酸之一的变体、具有与前述核酸之一互补的序列的核酸、为前述核酸之一的非功能突变变体的核酸、含有前述核酸之一的载体、含有前述核酸之一的细胞和含有前述载体的细胞中的至少一种组分被以治疗有效量施用于需要这种预防性治疗的患者。

当与肝脏失调和/或其他上皮癌的疗法的本领域状况相比时，根据本发明的预防方法令人惊奇地提供了改良的、持续的和/或更有效的预防措施。

优选地，该预防方法和/或刺激免疫应答的方法针对选自肝硬变、酒精肝疾病、慢性肝炎、威尔逊病、血色素沉着、肝细胞癌、良性肝肿瘤和病灶性结节性增生症的肝脏失调。具体地，该预防方法和/或刺激免疫应答的方法优选针对上皮癌，该上皮癌是不同于肝脏的器官的腺癌，该器官优选为选自肺、胃、肾、结肠、前列腺、皮肤和乳腺的器官。

在本发明的另一方面涉及鉴定至少一种药理学活性化合物的方法，该方法包括下面的步骤：

- a) 提供根据 SEQ ID 1 到 9 或 SEQ ID47 的至少一种多肽，或者其功能变体，
- b) 将所述多肽与怀疑是药理学活性的化合物接触，
- c) 测定步骤 (a) 的所述多肽与怀疑为药理学活性的化合物的相互作用，
- d) 鉴定直接或间接与步骤 (a) 的所述多肽相互作用的怀疑具有药理学活性的所述化合物。

优选地，所述多肽以选自：所述多肽附着到柱子、所述多肽粘附到阵列、所述多肽包含在电泳凝胶中、所述多肽粘附到膜和所述多肽被细胞表达等的形式提供。

优选通过选自基于酶或荧光的细胞报道分子测定法的方法测定相互作用，其中检测了怀疑具有药理学活性的化合物与包含步骤 (a) 的所述多肽的重组融合蛋白之间的相互作用。还优选可以通过表面胞质团共振、HPLC 和质谱测定相互作用。直接或间接相互作用优选选自诱导所述多肽的表达、抑制所述多肽的表达、激活所述多肽的功能、抑制所述多肽的功能。

术语“药理学活性物质”在本发明的意义内被理解为指在适宜条件下可以与根据 SEQ ID 1 到 9 和/或 SEQ ID47 的多肽，或者其功能变体（被 SEQ ID 10 到 19）编码）相互作用的所有那些分子、化合物和/或组合物和物质混合物，如果适宜与适当的添加剂和/或助剂一起。可能的药理学活性物质是简单化学（有机或无机）分子或化合物，但是

也可以包括肽、蛋白质或它们的复合物。药学活性物质的实例是来自化合物文库的有机分子，已经分析了这些化合物的药学活性。由于它们的相互作用，药学活性物质可以影响该多肽的体内或体外表达和/或功能或者备选地仅仅结合上述多肽或者以共价或非共价方式进行其他的相互作用。

可以根据本发明使用的适宜的试验系统基于用双杂交系统 (Fields 和 Stemglanz, 1994, Trends in Genetics, 10, 286-292; Colas 和 Brent, 1998 TIBTECH, 16, 355-363) 鉴定相互作用。在该试验系统中，细胞被表达融合蛋白的表达载体转化，该融合蛋白由至少一种根据本发明的多肽和转录因子如 Gal4 或 LexA 的 DNA-结合结构域组成。所转化的细胞还含有报道基因，其启动子含有相应 DNA-结合结构域的结合位点。另一表达载体表达由已知的或未知的多肽和活化结构域 (例如，来自 Gal4 或单纯疱疹病毒 VP16) 组成的另一融合蛋白，通过转化该表达载体，如果第二种融合蛋白与所研究的根据本发明的多肽相互作用，那么报道基因的表达可被极大地增加表达中的这种增加可以用于鉴定新的相互作用配偶体 (例如通过制备来自例如肝脏组织或者患病肝脏组织的 cDNA 文库，目的是构建第二种融合蛋白)。在优选的实施方案中，相互作用配偶体是根据 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 (被 SEQ ID 10 到 19 编码) 的多肽或者其功能变体的抑制剂。该试验系统还可用于筛选抑制根据本发明的多肽和相互作用配偶体之间相互作用的物质。这种物质降低正在表达根据本发明的多肽和相互作用配偶体的融合蛋白的细胞中报道基因的表达 (Vidal 和 Endoh, 1999, Trends in Biotechnology, 17: 374-81)。这样，可以快速鉴定可用于肝脏失调和/或上皮癌的治疗和/或预防的新的活性化合物。

用于鉴定对蛋白质的表达发挥影响的药学活性物质的测定法是技术人员熟知的 (见，例如，Sivaraja 等人，2001, US 6.183. 956)。这样，表达根据 SEQ ID 2 的多肽或者功能变体的细胞可以被作为分析体外基因表达的试验系统培养，优选肝脏细胞。使用本领域技术人员公知的方法在例如 mRNA 或蛋白质水平上分析基因表达。这样，将一种或多种假定的药学活性物质加入到细胞培养物后，测量根据 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 (被 SEQ ID 10 到 19 编码) 的多肽的



量并与对照培养物中相应的量比较。这例如通过特异针对 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47 (被 SEQ ID 10 到 19 编码) 的多肽或者其功能变体的抗体进行, 该抗体可用于检测细胞裂解物中存在的该多肽。通过本领域技术人员公知的方法使用, 例如, ELISA 或蛋白质印迹可以定量表达的多肽的量。关于这点, 可能实施作为高通量方法的分析并分析非常多的物质作为根据 SEQ ID 1 到 SEQ ID 9 和/或 SEQ ID 47(被 SEQ ID 10 到 19 编码) 的多肽的表达的调制剂的适合性(Sivara ja 等人, 2001, US 6.183. 956)。关于这点, 所要分析的物质可以来自物质文库(见, 例如, DE19816414, DE19619373), 其可以含有数千种物质, 这些物质通常是非常异源的。

现在将通过附图和实施例进一步阐明本发明, 这些附图和实施例代表本发明的优选实施方案和特征而不限制本发明。

#### 附图说明

图 1: HCC 中 RNA 表达水平

提供了 HCC 对未患病肝脏在 cDNA 微阵列实验中表达值的概括性箱形图(boxplot)。该箱形图是表达值比例的  $\log_2$  的图解表示, 中位值通过每个箱子中的水平线指出。每个箱子的范围表明 iqr; 触须表明 1.5 倍 iqr。不落在该范围内的比例用小圆指出。对于根据本发明的每种核酸, 在 HCC 中上升的表达是明显的。除了 OBc15 (SEQ ID 11), 表达值以相似比例一致地上升, 患者和对照样品中 OBc15 的表达差异是最显著的。

图 2: 当与正常组织和其他类型的癌相比时 HCC 中 OBc15 的表达特异性

通过使用引物 OBc15-p8, SEQ ID 66; OBc15-p9, SEQ ID 67; 和 OBc15-p10, SEQ ID 68, Taqman 荧光标记的基因特异探针的掺入和水解监视 OBc15 特异 PCR 产物的量。通过定量 RT-PCR (Q-PCR) 进行的 HCC=A, FNH=B 中 OBc15 (SEQ ID 11) 表达的定量评定与正常组织 (C=非癌性(正常)肝脏; D=肺正常; F=结肠正常; H=睾丸正常; J=肌肉正常; K=皮肤正常; L=心脏正常; M=肾正常) 和其他癌 (E=肺癌; G=结肠癌; I=睾丸癌) 中的表达模式比较。如 Wilcoxon 检验 (选项对=“假”) 进行的 Mann-Whitney-U 检验 (非正态分布数据的非参数检验) 提供了较大的两组 (HCC) (= Wilcoxon 值, W) 的等级

的总和并显示出所有组织样品中 Obc15 分布中的显著差异 (P-值), 如表 7 中所示。(HCC=肝细胞癌; FNH=病灶性结节性增生症; NNL=非瘤性(正常)肝脏; Lung N=肺正常; Col N=结肠正常; Tst. N=睾丸正常; Ms. N=肌肉正常; Skin N=皮肤正常; Hrt. N=心脏正常; Kdny. N=肾正常) 和其他癌 (Lung C=肺癌; Col. C=结肠癌; Tst. C=睾丸癌)。

表 7: 各种组织样品中 Obc15 的分布

数据	W	P-值
HCC 对 FNH	71	0.0005468
HCC 对 NNL	54	0.001504
HCC 对 Lung N	54	0.001504
HCC 对 Lung C	36	0.01053
HCC 对 Col. N	54	0.001504
HCC 对 Col. C	54	0.001504
HCC 对 Tst. N	72	0.0002734
HCC 对 Tst. C	54	0.001504
HCC 对 Ms. N	72	0.0002734
HCC 对 Skin N	54	0.001504
HCC 对 Hrt. N	54	0.001504
HCC 对 Kdny. N	54	0.001504

图 3: 表示独立 HCC 样品和对照中核酸 (SEQ ID 10 到 19) 表达的 RT-PCR 数据

在使用每种 cDNA 模板的平行实验中包括应用“管家”基因甘油醛磷酸脱氢酶 (GAPDH) 以控制 cDNA 质量。将进行 30-40 个 PCR 循环的 5 到 10% 的 RT-PCR 反应产物加到琼脂糖凝胶, 这里描绘出溴化乙锭染色的 DNA 凝胶。从 HCC 文库纯化的 DNA 作为根据本发明的每种核酸的阳性对照 (C)。在该分析中包括两个独立的 HCC 样品 (H) 与一种未患病肝脏样品 (N) 以代表根据本发明的核酸。M=分子量标准参比

物 (100 bp 序列梯)。

#### 图 4A/B: 通过 RNA 印迹验证差别基因表达

在如该图上指示的 RNA 印迹放射自显影图的图像中, 来自三个未患病肝脏 (L) 和 2 个 HCC 组织 (H) 的库的 RNA 样品的独立评定证实了 OBc11 (SEQ ID 10) 和 OBc15 (SEQ ID 11) 的增加的表达。来自对每种序列 (A, 顶部; 特定信号) 特异的反义链探针和作为阴性对照 (B, 底部) 的相应有义链探针的结果阐明了反义探针杂交的特异性。

#### 图 5: HCC 对 NNL 中 OBc15 RNA 定位

原位杂交分析检测肝细胞癌 (HCC) 和非瘤性肝 (NLL) 样品中 OBc15 RNA。放射性同位素标记的反义探针 (as) 与组织切片上 OBc15 RNA 特异杂交并通过放射自显影感光乳剂对切片的显影来检测。黑点是来自乳剂的显影的银颗粒, 指示对 OBc15 RNA 的特异杂交。互补的有义探针 (s) 尽管具有与反义探针的化学相似性但是不能原位杂交 OBc15 RNA。因此, 有义探针作为板 A 和 C 中的阴性对照而仅检测到背景信号。OBc15 RNA 在板 B 中所示的 NNL 中略微检测到并在 HCC 原位中清楚地指示, 如通过板 D 中许多银颗粒点所证明的。每个板以 200 倍 (200X) 的放大率显示。

#### 图 6: OBc15 RNA 表达的 siRNA-介导的击倒

将 HepG2 细胞用对 OBc15 RNA 序列特异的 siRNA 寡核苷酸或者用作为阴性对照的具有相同组成但是被扰乱的序列转染 (表 10)。这些特异性寡核苷酸与 OBc15 RNA 相互作用并使 OBc15 RNA 去稳定从而减少该 RNA 在肝癌细胞中的水平, 该过程称为 OBc15 RNA 水平的击倒 (knockdown)。在平行转染中使用阴性对照扰乱的寡核苷酸以为随后的实验读出提供对照参比 RNA。使用 Q-PCR 确定与来自三个独立实验 (A、B 和 C) 的扰乱的寡核苷酸转染的 (对照) 细胞相比, 特异寡核苷酸转染的细胞 (实验的) 中 OBc15 RNA 和成视网膜细胞瘤蛋白质 1 (RB1) mRNA 的表达水平。Y 轴代表 OBc15 mRNA 剩余活性的 log<sub>2</sub> 百分数值 (白色, 左柱); 而 RB 1 mRNA log<sub>2</sub> 比例值指示 OBc15 siRNA 转染的细胞与对照寡核苷酸转染的 HepG2 细胞相比 RB 1 mRNA 的水平的增长倍数 (黑色, 右柱)。特异 siRNA 寡核苷酸介导的 OBc15 RNA 的降低是明显的。实验细胞而不是对照细胞中 RB 1 mRNA 的上升的水

平表明 OBc15 表达负调节该肿瘤抑制剂 mRNA 的水平。

#### 图 7: 组织中的 DAP3 蛋白质表达

将蛋白质提取物进行免疫印迹分析, 该分析使用对 DAP 和  $\beta$ -肌动蛋白蛋白质特异的抗体以确定人组织中这些蛋白质的表达水平。用缀合辣根过氧化物酶 (HRP) 的第二抗体孵育并用化学发光的 HRP 底物检测免疫复合物后, 通过光密度分析带的强度并将每种信号标准化为相应  $\beta$ -肌动蛋白信号的强度。每个泳道中代表的组织在表 8 中定义, 表 8 还包括这些组织中 DAP3 蛋白质水平的定量分析。这些分析指出 DAP3 蛋白质 (HCC 中被特异上调的 DAP3 mRNA 的功能产物) 在 HCC 中也高度过表达。

表 8: 图 7 中检查的组织和人组织提取物中 DAP3 蛋白质表达水平的光密度定量。

编号	组织	DAP-3	$\beta$ -肌动蛋白	DAP-3 标准化
1	脑	1.5	7.4	1.4
2	小脑	1.6	7.5	1.4
3	心脏	1.2	0.0	1.2
4	结肠	3.3	7.4	3.0
5	肺	0.0	6.7	0.0
6	胃	4.2	6.2	4.6
7	胰腺	16.3	6.2	17.8
8	肾	0.0	0.0	0.0
9	前列腺	0.9	4.2	1.5
10	子宫	1.4	9.2	1.0
11	HCC2	20.7	6.3	22.2
12	HCC2	31.1	8.0	26.3
13	HCC4	15.6	6.9	15.3
14	肝脏	1.9	3.5	3.7
15	骨骼肌	0.1	0.0	0.1
16	睾丸	0.5	6.0	0.6
17	脾脏	0.0	5.1	0.0
18	乳腺	0.4	8.1	0.3

### 图 8: HCC 去调节的基因的表达与肝癌细胞增殖的相关性

对静止细胞用血清刺激 8 小时 (黑色柱) 和 12 小时 (白色柱) 之后肝癌细胞 (Hep3B) 中根据本发明的靶标基因序列的增殖依赖的表达。提供了来自 cDNA 微阵列实验读出的血清-刺激的对静止表达的  $\log_2$ -转化的比例。与静止肝癌细胞相比增殖肝细胞中这些序列表达水平的实质性增加表明这些序列对于肝癌细胞生长在功能上是重要的。

#### 具体实施方式

##### 实施例

##### 实施例 1: HCC 扣除 cDNA 文库的制备

根据标准方法 (Chomczynski 和 Sacchi, 1987, Anal. Biochem. 162: 156-159) 使用 TRIZOL 试剂 (Invitrogen) 从三个病理学家证实的 HCC 肿瘤样品和三个病理学家证实的未患病人肝脏样品分离 RNA。用于产生 cDNA 文库的组织来自被通知关于该组织用于研究目的, 包括商业研究的特定知情同意的患者。按照 Clontech 实验室的“PCR 选择 cDNA 扣除试剂盒”中的使用说明中描述的反转录酶和 DNA 聚合酶将 mRNA 转化成双链 cDNA。为了富集 HCC 中特异增加和减少的 cDNA, 通过根据该试剂盒中提供的使用说明的扣除抑制杂交 (SSH) 和 Diatchenko 等人 (1996, Proc. Natl. Acad. Sci. USA 93: 6025-6030) 所描述的除去在参比肝脏库和 HCC 中以共同和相似水平表达的 cDNA。在两个方面实施 SSH 步骤 (从 HCC cDNA 扣除未患病的肝脏 cDNA 和从未患病的肝脏 cDNA 扣除 HCC cDNA) 从而所得 cDNA 分子代表在 HCC 中被上调的和被下调的核酸序列但是不代表未差别表达的那些核酸序列。另外, 产生标准化但是未扣除的 HCC cDNA 文库以更好地代表 HCC 组织中的稀有 mRNA 转录物。这些 cDNA 通过连接到该质粒而被单独克隆到 pCRII 载体 (Invitrogen) 然后通过电泳转化到大肠杆菌 XL-1-Blue 电穿孔-感受态细胞 (Stratagene) 中。按照载体和感受态细胞的供应商所描述的克隆。将克隆的差别表达的 cDNA 涂在选择 (氨苄西林) 培养基上以分离单个克隆。从每个 SSH 文库分离 960 个克隆, 从标准化 HCC 文库分离到 576 个克隆, 并在 96-孔微量滴定板中建立培养物。这些 cDNA 克隆一起提供了对人 HCC 组织特异的 mRNA 表达的

独特代表。

### 实施例 2: HCC cDNA 微阵列的制备和杂交

建立 SSH cDNA 文库克隆的 1ml 培养物并将通过使用对 cDNA 插入片段侧翼的载体序列特异的引物进行的 PCR 扩增 cDNA 插入片段插入。M13 正向 (5'-GTAAAACGACGGCCAG-3'; SEQ ID 20) 和 M13 反向引物 (5'-CAGGAAACAGCTATGAC-3'; SEQ ID 21) 用于克隆插入片段的 PCR 扩增。将 50 微升细菌培养物在 95℃ 热变性 10 分钟, 离心除去碎片, 并且在标准 PCR [1X Amplitaq PCR 缓冲液, 2.5 mM MgCl<sub>2</sub>, 37.5 nM 每种引物, 0.5 mM 每种 dATP、dCTP、dGTP 和 dTTP 和 1.5 单位 Amplitaq DNA 聚合酶 (Applied Biosystems) ] 中包含 2 μl 上清液。反应条件为 95℃ 5 分钟然后是 35 个如下循环: 94℃ 30 秒, 60℃ 30 秒, 72℃ 60 秒; 然后 72℃ 7 分钟并冷却到 4℃。通过在含有 0.4 μg/ml 溴化乙锭的 1% 琼脂糖凝胶上 5% PCR 的电泳并在 1× Tris 乙酸 EDTA (TAE; 40mM Tris-乙酸盐, 1mM EDTA, pH 7.5) 中运行来证实 cDNA 插入片段的扩增。使用 Genetic Microsystems 417 cDNA 阵列机器人将每种 SSH 克隆扩增的插入序列固定到唾液酸化 (sialinized) 的玻璃显微载玻片 (GAPS Coming) 上以产生定制的 HCC cDNA 微阵列。将 cDNA 插入片段点到玻片的方案是根据 Hedge 等人 (2000, Biotechniques 29: 548-560) 所公布的只是所点的 PCR 产物直接来自 PCR 微量滴定板而不纯化或者调节 cDNA 缓冲液。除了 SSH cDNA 克隆插入片段, 许多对照 DNA 被点到微阵列上作为杂交反应的对照。此外, 相应于以前报导与癌相关的基因的约 2000 个可以公开得到的 cDNA 克隆购买于德国基因组研究中心 (German Genome Research Center (RZPD)), 如上述的将这些克隆扩大 (expand)、扩增并点到这些微阵列上。为了制备杂交探针, 使用反转录酶根据标准方法 (Hedge 等人, 2000, Biotechniques 29: 548-560) 将 20 微克来自额外的病理学证实的肝脏失调的 RNA 和来自相同量的合并的未患病的肝脏 RNA 分别转化成 cy5-荧光-标记的和 cy3-荧光标记的 cDNA (cy5-CTP 和 cy3-CTP, Pharmacia)。使用该方案, 这些标记的 cDNA 竞争性地杂交到 HCC 微阵列。在 5× SCC (0.75 M 柠檬酸钠, 75 mM 柠檬酸钠, pH 7.0); 0.1% SDS (十二烷基硫酸钠) 和 1% BSA (牛血清白蛋白) 中 42℃ 预杂交 45

分钟后, 在含有 50%甲酰胺、 $5 \times \text{SSC}$ , 和 0.1%SDS 的缓冲液中  $42^\circ\text{C}$  进行过夜杂交。杂交的载玻片在严格条件中洗涤(每个在  $1 \times \text{SSC}$ 、0.1% SDS 中  $42^\circ\text{C}$  洗涤两次, 每次 2 分钟; 每个在  $0.1 \times \text{SSC}$ 、0.1% SDS 中室温下洗涤两次, 每次 4 分钟; 每个在  $0.05 \times \text{SSC}$  中室温下洗涤两次, 每次 2 分钟), 干燥并用 Genetic Microsystem 418 cDNA 微阵列扫描仪和相关的 Imogene 4.1 成像分析软件根据生产商的推荐分析。

### 实施例 3: 根据本发明的核酸和多肽的差别表达的独立验证

如上详述的从人类患者样品分离 RNA。用于该分析的 HCC 样品不是来自与用于产生 HCC SSH 文库或者 cDNA 微阵列芯片杂交(见上面的实施例, 表 3A/3B, 4 和图 1)相同的患者。除了 HCC 样品, 从独立的未患病的肝脏样品制备 RNA 以评价未患病肝脏组织中根据本发明的核酸的表达。此外, 从另外的未患病的和癌组织制备 RNA 以评价其他正常人组织和其他人类癌中根据本发明的核酸的表达。使用本领域中公知的标准方法 (Sambrook 等人, 分子克隆 (Molecular Cloning), 第二版, 1989, 冷泉港出版社, NY, 美国, 5.52-5.55 页) 通过 Superscript 反转录酶 (Invitrogen) 在 dATP、dCTP、dGTP、和 dTTP (每种 0.4 mM)、7.5 nM 随机 6-核苷酸引物 (六聚体)、10 mM 二硫苏糖醇和 1 单位 RNA 酶抑制剂中将  $1 \mu\text{g}$  RNA 转变成单链 cDNA。然后在 PCR 实验中通过用对根据本发明的每种核酸特异的引物对从 cDNA 扩增这些序列来确定根据本发明的核酸的存在与否。用于该分析的引物在下面的表 9 中给出。

表 9: RT-PCR 引物与它们各自 SEQ ID 号

克隆	SEQ ID	引物 1 (SEQ ID)	引物 2 (SEQ ID)
OBcl1	10	5'-CAGGTGAATTTCAAAGG AGGATTTACTCAC-3' (22)	5'-GTGAGTAAATCCTCCTT TGAAATTCACCTG-3' (23)
OBcl5	11	5'-GCAAGCCAGGAAGAGT CGTCACG-3' (24)	5'-TGCCAGGAACTTCTTG CTTGATGC-3' (25)
IK2	12	5'-AGTAACCAAGTTGAGATG AAGCACGTC-3' (26)	5'-CAGAAGAGCAACAAGA ATGGTATCCTGC-3' (27)
IK5	13	5'-AACTTGAGTTCTATTTAC CTTGAC-3' (28)	5'-TTGCTTGGGTCATCTAA AGAC-3' (29)
DAP3	14	5'-ACTCACGTGCAAGGATG ATG-3' (30)	5'-AGCTCTCGGACTCTCAA CTG-3' (31)
LOC5	15	5'-CTTCTCCTATGACTGATC CTACTATG-3' (32)	5'-CAGGATGCAGAACTCAC CCTG-3' (33)
SEC1 4L2	16	5'-GCAGATTTCCCGTGGCT CCTC-3' (34)	5'-GTTGGGCAGCACCTCTG TCATC-3' (35)
SSP29	17	5'-CTGTGACATTCCGCCTTC CTTC-3' (36)	5'-CCACGCTACTGCAAGAA TCTTAC-3' (37)
HS16	18	5'-AGAAGTTCAACCTGGAG AGATGG-3' (38)	5'-CAAGGAAGCTAGGAATG ACAGGAG-3' (39)
IK3	19	5'-GCAAAGCCAAATTCATG TACTCT-3' (40)	5'-CAGATACGAACAGTGAA TGAAATACG-3' (41)

这些引物也可用于根据本发明的失调的诊断，但是技术人员也可以设计对根据本发明的给定核酸特异的其他引物。该 PCR 包括 0.5% cDNA、1 × Amplitaq PCR 缓冲液、2.5 mM MgCl<sub>2</sub>、37.5 nM 每种引物、0.5 mM 每种 dATP、dCTP、dGTP 和 dTTP 和 1.5 单位 Amplitaq DNA 聚合酶 (Applied Biosystems)。按照每种引物对所需优化 PCR 条件，通常为：94℃ 3 分钟和 30 个如下循环：94℃ 15 秒，60℃ 30 秒，72℃ 60 秒，然后冷却到 4℃。通过在含有 0.5 μg/ml 溴化乙锭的 1% 琼脂糖凝胶上 5-10% PCR 的电泳并在 1 × Tris 乙酸 EDTA (TAE) 缓冲液上运行来证实 cDNA 插入片度的扩增。包括 cDNA 合成前样品的 RNA 酶处理以及反转录酶的缺失的 RT-PCR 标准对照常规地展示了这些反应的特异性。



根据在凝胶上是否观察到正确的分子大小的离散带为反应的表达 (+) 或无表达 (-) 打分。在这些条件下非常弱的或者模糊的带记为 (+/-)。HCC 和未患病肝脏中这些验证研究的概述在表 6 中给出。在图 3 中提供了代表独立 HCC 和未患病肝脏样品中这些分析的数据。

定量 RT-PCR (Q-PCR) 也证实了相对于未患病肝脏在肝癌和其他肝脏失调中根据本发明的序列的过表达。对于这些研究, 使用如在实施例 5 中详细描述并在图 2 和 6 中图解的使用 TaqMan 水解引物策略和 SYBR Green 插入染料策略。

在图 4 中图解了根据本发明的核酸的差别表达的另外的独立的确证。在该情况中, 将来自两个 HCC 样品和未患病肝脏的 15  $\mu\text{g}$  RNA 在 1% 琼脂糖凝胶上进行变性电泳分离, 该琼脂糖凝胶含有 2.2M 甲醛和 1  $\times$  MOPS 缓冲液 (10 mM 4-吗啉丙烷磺酸、1 mM EDTA、5mM 乙酸钠, pH 7.0), 电泳在 1  $\times$  MOPS 缓冲液中运行。使用 RNA (northern) 印迹技术将按大小分级的变性 RNA 转移到尼龙膜 (GeneScreen, New England Nuclear) 并将该变性 RNA 用紫外光交联到膜, 所有的都根据技术人员熟知的方法 (Sambrook 等人, 分子克隆, 第二版, 1989, 冷泉港出版社, NY, Press, NY, USA, 7.39-7.52 页)。通过前面实施例中描述的 PCR 扩增从 OBc11 和 OBc15 (SEQ ID No. 10 和 11) 的 SSH 克隆分离 cDNA 克隆插入片段。在 1  $\times$  标记缓冲液: 0.5 mM ATP、CTP、GTP、10 mM 二硫苏糖醇中  $\alpha$ - $^{32}\text{P}$ -UTP 存在下, 使用 SP6 和 T7 RNA 聚合酶, 和 20 单位适宜的 RNA 聚合酶 37 $^{\circ}\text{C}$  35 分钟从该模板合成相应于这些序列的单链放射性标记的 RNA 探针。所得反义探针与相应的 mRNA 序列互补并从而预期特异杂交 RNA 印迹上的 mRNA 序列。相反地, 有义探针序列与该 mRNA 的序列匹配从而不杂交该 mRNA。相同的 RNA 印迹在 15ml 250 mM 磷酸二氢钠、250 mM 磷酸氢二钠、7%SDS、1 Mm EDTA 和 1% BSA 中 68 $^{\circ}\text{C}$  下预杂交至少 30 分钟。为了杂交, 除去预杂交缓冲液并用含有上述有义和每种反义 RNA 探针的相同缓冲液代替, 在 68 $^{\circ}\text{C}$  过夜。RNA 印迹被在严格条件 (2X SSC, 0.1% SDS 室温下洗涤两次, 每次 15 分钟; 1X SSC, 0.1% SDS 68 $^{\circ}\text{C}$  两次, 每次 10 分钟) 下洗涤, 干燥并暴露于 x-射线薄膜以产生放射自显影照片。如在图 4 中所看到的, 每种反义探针特异杂交离散的 HCC RNA 但是仅仅微弱的或者根本不杂交未患病的肝脏 RNA。这些结果的特异性通过缺少来自 OBc11 和

OBc15 的相应有义探针的特异性信号阐明。此外,使用 OBc11 反义探针,不同分子量的 RNA 是明显的。该结果最可能代表离散的 mRNA 种类,其可能由备选剪接产生。基于如下发现预期这些种类:在 GenBank 序列数据库中报导相应于该序列的一些不同大小的 cDNA 克隆。

此外,原位杂交揭示当与 NNL 组织样品相比时 HCC 中的强 OBc15 RNA 表达。根据 Fickert 等人(Am J Pathol. 2002 年 2 月; 160(2): 491-9.) 的方案,如上述的合成用于 RNA 印迹的  $S^{35}$ -标记的探针。用于体外转录的模板从含有 OBc15 3' cDNA 的质粒扩增。用于产生体外转录模板的引物(MWG Biotech, Munich, 德国)为 OBc15-p6 正向引物(5'-aatctgcaagccaggaagagt-3', SEQ ID 48) 和 M13for (5'-gtaaaacgacggccag-3', SEQ ID 20), 它们针对包括两个外显子的横跨 OBc15 RNA 的 365 个碱基的 T7 反义探针(SEQ ID 11 从核苷酸 95 到 484); 和 M13rev(5'-caggaacagctatgac-3', SEQ ID 21) 和 OBc15-p7 反向引物(5'-tctagtttcagttttgatgatattttg-3', SEQ ID 49), 它们针对包括两个外显子的横跨 OBc15 序列的 436 个碱基的 SP6 有义探针(SEQ ID No 11 从核苷酸 436 到 1)。为了扩增这些模板,用 Applied Biosystems Gene Amp PCR System 270 进行 PCRs, 其包括 10 pM 正向引物、10 pM 反向引物、1 pM dNTPs (Invitrogen)、PCR 缓冲液 II、5m M  $MgCl_2$  (Applied Biosystems, Foster City, CA)、217 ng 模板质粒 DNA、2.5 U AmpliTaq 聚合酶(Applied Biosystems, Foster City, CA), 通常为: 94℃ 3 分钟, 94℃ 30 秒, 50℃ 30 秒, 72℃ 50 秒, 最后 3 步重复 25 次, 然后 72℃ 3 分钟用于最后的延伸。PCR 产物中 DNA 的量通过光谱测定法使用 Smart Spec 3000 (BIO-RAD, Hercules, CA) 确定。使用转录缓冲液(Boehringer Mannheim, 德国)中的每种模板 200ng、100 mM 二硫苏糖醇(DTT)、1mM 每种 rNTP、RNA 酶抑制剂(Eppendorf, Hamburg, 德国)、 $\alpha$ - $S^{35}$ -UTP (Amersham Bioscience)、RNA-聚合酶 SP6/T7 (Boehringer Mannheim), 在 37℃ 下实施体外转录测定法 2 小时。除去未掺入的核苷酸(无 RNA 酶的 MicroSpin S-200 HR 柱, AmershamBioscience, Buckinghamshire, UK), 模板 DNA 用 2 单位无 RNA 酶的 DNA 酶在 37℃ 消化 10 分钟。为了得到 150bp 的平均探针大小, 使用水解缓冲液(400 mM  $NaHCO_3$ , 600 mM  $Na_2CO_3$ , 100m M DTT) 在 60℃ 水解 42 分钟并在 0.1M 乙酸钠、10 mM DTT 和 1%冰乙酸

中中和。转录探针用 LiCl/异丙醇沉淀并重悬在含有 25 mM DTT 的 50% 甲酰胺中。HCC 的和非瘤性正常肝切片的石蜡包埋的组织学证实的样品在 Microm HM 355S 切片机 (Microm, Waldorf, 德国) 上以 2.5 微米切割并以每个载玻片 2 个切片在 Superfrost 载玻片 (Menzel-Gläser, Braunschweig, 德国) 上封片。所有切片过夜干燥, 将干燥的切片加热到 60°C 1 小时并在二甲苯中脱蜡 30 分钟。用 100%、90%、70% 和 50% 的梯度乙醇进行脱水, 然后在 Tris 缓冲液 (TBS 缓冲液) 中洗涤 4 次, 每次 3 分钟, 然后将切片在含有 4% 低聚甲醛的磷酸缓冲液 (PBS 缓冲液) 中固定。几次 PBS 洗涤后, 将切片在 0.2M HCl 中变性 10 分钟并在 TBS 中再洗涤 4 次, 每次 3 分钟。37°C 下在含有 2 mM CaCl<sub>2</sub> 的 TBS 中 20 µg/ml 无 RNA 酶的蛋白酶 K (F. Hoffman La Roche Ltd. Basel, 瑞士) 中消化蛋白质 20 分钟。通过将载玻片在 4°C TBS 中孵育 5 分钟终止反应。随后, 将切片在 TBS 缓冲液中室温下再次洗涤 3 次, 每次 4 分钟, 并在含有乙酸酐的 0.1M Tris 缓冲液 (pH8) 中孵育。切片在 50%、70%、90%、100% 梯度乙醇并最后在氯仿中脱水, 然后空气干燥 2 小时。为了杂交, 将标记的探针 (1x10<sup>6</sup> cpm/切片; 探针放射活性用 LKBWallac 确定, 1211 RACKBETA 液体闪烁计数器) 在 50 µl/切片杂交缓冲液中稀释, 该杂交缓冲液含有 12.5 mM pH6.8 的磷酸缓冲液、12.5 mM Tris、0.4M NaCl、3 mM EDTA、1.25 × Denhardtts 溶液、50% 甲酰胺、12.5% 葡聚糖硫酸盐、0.1M DTT、100 nM S-rATP (Boehringer Mannheim)、60ng 酵母 tRNA, 和 20ng 聚腺苷酸 (Boehringer Mannheim)。切片在 52°C 在含有 2 × 标准柠檬酸盐溶液 (SSC) pH7 和 50% 甲酰胺的湿室中过夜杂交。然后, 切片用甲酰胺缓冲液 (10mM 磷酸缓冲液 pH 6.8、10 mM Tris-HCl pH 7.7、0.3M NaCl、5mM EDTA、0.1 × Denhardtts 溶液、0.07% β-巯基乙醇, 和 50% 甲酰胺) 洗涤两次, 分别为 1 小时和 2 小时。此后切片在 10 mM Tris-HCl pH 7.4、0.5M NaCl、2.5 mM EDTA 和 0.07% β-巯基乙醇中洗涤两次, 每次 15 分钟。在含有 20 µg/ml RNA 酶 A (Boehringer Mannheim) 的相同缓冲液中 38°C 下实施 RNA 酶处理 30 分钟, 然后在 37°C 用甲酰胺洗涤缓冲液洗涤过夜。切片随后在 2 × SSC 和 0.07% β 巯基乙醇中 45°C 下洗涤 30 分钟, 然后在 0.1 × SSC 和 0.07% β 巯基乙醇中 45°C 下再洗涤 30 分钟。此后切片在 50%、70%、90%、100% 的梯度乙醇中脱水并空气干

燥。最后，载玻片在 Ilford K2 光乳剂 (Ilford Ltd. Mobberly, Cheshire, UK) 中包被。10、14 和 17 天的曝光后，用 Kodak D 19 显影剂 (Eastman Kodak, Rochester NY) 进行显影。切片用苏木精复染色并在水性封片介质 (Aquatex-EM Science, Gibbstown, NJ) 中封片。显影暗点，来自乳剂的银颗粒指示对 OBc15 RNA 的特异杂交 (图 5)。互补有义探针尽管具有与反义探针的化学相似性但是不能原位杂交 OBc15 RNA。所以有义探针作为阴性对照 (图 5, 板 A 和 C 中)，其中仅仅检测到背景信号。OBc15 RNA 在 NNL 中略微检测到 (图 5 中所示, 板 B 中) 并在 HCC 原位清楚的指示，如通过大量银颗粒点所证明的 (图 5, 板 D 中)。

此外，蛋白质表达分析指出 DAP3 蛋白质 (HCC 中特异上调的 DAP3 mRNA 的功能产物) 也在 HCC 中特异地上调。为了检测各种组织中 DAP3 蛋白质表达，使用来自冰冻组织 (贮存在液氮中) 的蛋白质提取物进行标准蛋白质印迹分析，见图 7。使用冰冻切片机 (cyrocut, Leica CM3050) 得到 50  $\mu\text{m}$  切片 (HCC, 正常肝脏和各种器官样品)，其中在每个切割过程之前、之间和之后通过切片的 H&E 染色检验所细查的组织的同时性和同质性。将组织切片重悬在补加 2  $\mu\text{g}/\text{ml}$  亮抑蛋白酶肽、2  $\mu\text{g}/\text{ml}$  胃蛋白酶抑制剂、2  $\mu\text{g}/\text{ml}$  抑酶肽、1mM 苯甲基磺酰氟 (PMSF)、和 2 mM 二硫苏糖醇的冰冷却的 RIPA 缓冲液 (50 mM Tris-HCl pH 7.4, 250 mM NaCl, 0.1% SDS, 1% 脱氧胆酸盐, 1% NP-40) 中然后通过冰上超声 (5 秒的 2 次猝发) 处理匀浆。在冰上孵育 20 分钟后，通过在微离心机中 4 $^{\circ}\text{C}$  下以 13 000 rpm 15 分钟的两个离心步骤使裂解物澄清并收集上清液。通过 Bradford 测定法 (Biorad) 使用牛血清白蛋白作为标准确定蛋白质浓度。等量蛋白质 (通常 10-30  $\mu\text{g}$ ) 在 12% SDS-PAGE 凝胶上分离并通过半干印迹 (TE 70, Amersham) 电泳转移到聚偏氟乙烯 (PVDF) 膜 (Hybond-P, Amersham)。将膜在封闭溶液 [TBS-T (25 mM Tris-HCl pH 7.4, 137 mM NaCl, 3 mM KCl, 含有 0.1% Tween-20) 中 5% 奶] 中室温下封闭 1 小时并在搅拌下与第一抗体溶液 (在 TBS-T/1% 奶中制备) 在 4 $^{\circ}\text{C}$  孵育过夜。使用对下面的抗原特异的抗体: DAP3 (1: 1000; BD Transduction Laboratories) 和  $\beta$ -肌动蛋白 (1: 5000, Sigma)。除去第一抗体溶液并在 TBS-T 中几次洗涤后，将膜与缀合 HRP (辣根过氧化物酶) 的第二抗体 (兔抗鼠: 1: 1000; Dako) 在室温

下孵育 1 小时。在 TBS-T 中洗涤几次后,通过化学发光(ECL, Amersham)并使 x-光片曝光进行检测(图 7)。使用 ChemImager 5500 软件(Alpha Innotech)以光密度法分析带的强度并将每种信号标准化到相应于  $\beta$ -肌动蛋白信号的强度(表 8)。

这些数据提供了 HCC 中根据本发明的核酸和多肽的去调节表达的独立验证。根据本发明的核酸和多肽的表达在未患病肝脏中或者不存在或者仅以非常低的水平被观察到,从而证实了通过杂交到 cDNA 微阵列鉴定的这些核酸的差别表达。这些结果提供了令人惊奇的证明,即根据本发明的核酸和多肽可用于诊断、预防和/或治疗根据本发明的失调。

**实施例 4: 在正增殖的肝癌(肝细胞瘤)细胞系中根据本发明的序列增加**

在 5%CO<sub>2</sub>, 37℃ 的湿润培养箱中补加 10%胎牛血清(FBS)的 DMEM 中培养人肝癌细胞系(HepG2, Hep3B)。细胞分离到约 20%融合并随后通过在没有血清时培养 3 天使细胞静止。在饥饿期后,通过向培养基中加入 10%FBS 刺激细胞增殖。在诱导细胞生长之前和之后(0、8 和 12 小时)采样用于制备 RNA 和用于通过 FACS(荧光激活细胞分类术)分析确定细胞在细胞周期中的位置。因此,为了通过碘化丙锭(PI)染色确定细胞周期分布,将细胞通过胰蛋白酶消化收获,以磷酸缓冲液(PBS)洗涤两次并最后重悬在 500  $\mu$ l PBS 中。随后,加入 5ml 预冷却的甲醇。在 -20℃ 孵育 10 分钟后,将细胞悬浮物直接用于 FACS 分析,之后在 PBS 中洗涤 3 次,重悬在 500  $\mu$ l 碘化丙锭(PI)染色缓冲液(DNA-Prep Stain, Part No. 6604452; Beckman Coulter)中并在 37℃ 孵育 15 分钟。最后,加入 70  $\mu$ l 1M NaCl 并在 EPICS XL-MCL 流式细胞仪(Beckman Coulter)上分析前将样品保持在避光保护的冰上。从异步(asynchronous)细胞群体制备的细胞用作参比。

所分离的 RNA 用于通过 cDNA 微阵列分析监测静止的与正在增殖的肝癌细胞中基因的表达。如实施例 2 中描述的用荧光染料标记后,将 RNA 在特定开发的 HCC-特异性 cDNA 微阵列芯片上杂交,该微阵列芯片还含有对照基因,已知这些对照基因以独立于细胞周期的方式表达。最后,使用 ImaGene 4.1 和 GeneSight 软件包分析数据。加入血

清前分离的 0 小时样品所得到的信号用作参比。在图 8 中提供了来自 cDNA 实验读出的血清刺激的与静止表达的  $\log_2$ -转化的比例。

这些数据指出根据本发明的序列与人肝脏肿瘤细胞增殖相关。与本领域的状况相比，这些核酸和多肽因此令人惊奇地允许肝脏失调和/或上皮癌的改良的、更敏感的、更早、更快和/或非侵入性诊断。

实施例 5: 根据本发明的序列的升高的表达在肝脏失调，特别是肝癌中的功能上的重要角色

详细的序列分析揭示了 OBc15 RNA 和真核生物非编码 RNA 之间的序列相似性。此外，用多种方法检测来自该 RNA 的蛋白质产物的多种尝试都没有揭示这种产物。因此，该 RNA 可能不被翻译成多肽而是其自身具有功能（例如，调节的）性质。使用根据 TransMessenger 转染试剂手册（TransMessenger Transfection Reagent Handbook）（Qiagen, 10/2002）的方案，实施了用小干涉 RNA（siRNA）寡核苷酸（siRNA 介导的 OBc15 RNA 的击倒）减少正在增殖的人肝癌细胞中 OBc15 RNA 的水平。设计双链小干涉 RNA（siRNA）寡核苷酸探针（表 10）用于相应于 OBc15（SEQ ID 11）的 RNA 水平的原位耗尽并且该双链小干涉 RNA 寡核苷酸探针由 Qiagen 提供。

表 10: 双链小干涉 RNA（siRNA）寡核苷酸探针

名称	序列	SEQ ID*
OBc15 siRNA fw	5' r(UCUGCAAGCCAGGAAGAGU)d(TT) 3'	50
OBc15 siRNA rev	5' r(ACUCUUCCUGGCUUGCAGA)d(TT) 3'	51
OBc15 siRNA fw1	5' r(CCUCCAGAACUGUGAUCCA)d(TT) 3'	52
OBc15 siRNA rev1	5' r(UGGAUCACAGUUCUGGAGG)d(TT) 3'	53
DAP3 siRNA fw	5' r(CUACAAAUGAGCGCUUCCU)d(TT) 3'	54
DAP3 siRNA rev	5' r(AGGAAGCGCUCAUUUGUAG)d(TT) 3'	55

[\*关于这些 siRNA 核糖-寡核苷酸的所附信息中以 SEQ ID 号列出的序列不包括每条序列末端的两个 3'脱氧核糖核苷酸（dT）“尾巴”，因为在这些列表中不可能指定核糖核苷酸/脱氧核糖核苷酸嵌合分子]。

接种 HepG2 细胞（每孔密度为  $3 \times 10^4$  个细胞）并在  $37^\circ\text{C}$  孵育 24 小时并根据生产商的使用说明(Invitrogen 方案)将这些细胞用  $1.5 \mu\text{l}$  Oligofectamine 试剂 (Invitrogen) 和  $2.5 \mu\text{l}$   $20 \mu\text{M}$  双链 siRNA 寡核苷酸母液转染。孵育 24 小时后，如实施例 2 中描述的分离总 RNA 并反转录成 cDNA。通过将荧光标记的引物或者包括 Taqman 探针水解系统和荧光双链 DNA 插入分子如 SYBR 绿的各种基于荧光的指示剂的掺入来监视 PCR 产物。根据生产商的使用说明 (GeneAmp®5700 序列检测系统, 用户手册; PE Biosystems) 进行实验。因此, 如下使用 5700 序列检测系统 (Applied) 进行基于 TaqMan 方法的实时定量 RT-PCR 分析: 如实施例 3 中描述的将 500ng 总 RNA 反转录并且该 cDNA 模板的 1:4 稀释液用于 Q-PCR (相应于 6.25ng RNA),  $30 \mu\text{l}$  终体积中包括 5-8pmol/ $\mu\text{l}$  每种引物。根据生产商的使用说明使用 Q-PCR 的温度, 使用 40 个循环, 进行一式三份的反应。

使用 7000 序列检测系统 (Applied) 进行基于 SYBR-绿的方法的实时 Q-PCR 分析。根据生产商的使用说明用 SYBR-Green Universal PCR Master Mix (Applied) 进行 PCR, 使用相应于如上 6.25 ng RNA 的 cDNA、和每种引物的经验确定的量 (RB 和  $\beta$ -肌动蛋白, 反应样品中 10 pmol 每种引物), 终体积  $30 \mu\text{l}$ 。根据使用说明手册使用 SYBR-RT-PCR 的温度。这些反应也循环 40 次并进行一式三份反应。通过 Q-PCR 使用 GAPDH 或者  $\beta$ -肌动蛋白 RNA 水平的平行 Q-PCR 确定作为参比在基于如前述的 TaqMan (所用 GAPDH 引物= GAPDH-p1, SEQ ID 56; GAPDH-P2, SEQ ID 57; GAPDH-p3, SEQ ID 58) (所用  $\beta$ -肌动蛋白引物=bActin-p1, SEQ ID 59; bActin-p2, SEQ ID 60; bActin-p3, SEQ ID 61) 或者 SYBR Green (用作 SYBR green 分析的参比的  $\beta$ -肌动蛋白引物= bActin-p4, SEQ ID 62; bActin-p5, SEQ ID 63) 的分析中确定靶标 RNA 水平 (在该情况下为 OBc15 RNA) 击倒的百分比。根据 Pfaffl (Nucleic Acids Research (2001) 5 月 1 日, 29 (9): e45) 描述的方法确定 RNA 水平的改变。

在其中肝细胞中 OBc15 RNA 的水平被击倒的实验中, 已确定当 OBc15 RNA 水平降低时, 编码该肿瘤抑制基因成视网膜细胞瘤蛋白质 1 (RB1) 的 mRNA 的水平被以剂量依赖的方式上调数倍 (图 6) (所用

RB1 Q-PCR 引物=RB1-p1, SEQ ID 64; RB 1-p2, SEQ ID 65)。可得到清楚的结论: HCC 中 OBc15 RNA 的上升的表达可以提供 RB1 的负调节并因此促进肿瘤细胞生长。从而, 用 siRNA 寡核苷酸降低正在增殖的人肝癌细胞中 OBc15 RNA 的水平(击倒)证明了肝脏失调, 特别是肝癌中 OBc15 RNA 的上升的表达的功能上的重要作用。

另一实验, 其中设计了用以击倒肝癌细胞中 DAP3 mRNA (SEQ ID 14) 的 siRNA 寡核苷酸, 提供了令人惊奇的形态学影响(表 10 中提供了用于 DAP3 siRNA 击倒研究的 oligo 序列)。在 DAP3 siRNA oligo 处理的细胞中而不是同样地但是使用其他 siRNA oligo (如扰乱的序列 siRNA oligo 对照) 处理的细胞中, 观察到细胞形态学的极大改变, 包括细胞体积的变大。这些被处理的细胞保持粘附到培养基质, 但是还观察到使用这些实施例中描述的标准方法不能从这些被处理的细胞中提取 RNA 和蛋白质。在同样处理的肝癌细胞中用相似的 siRNA oligo 处理却没有这种效果, 这证明这些观测是肝癌细胞中 DAP3 mRNA 水平的击倒所特有的。因此 DAP3 mRNA 的过表达对于肝癌细胞的活力可能是关键的。这些观察进一步证明 DAP3 在肝肿瘤细胞中的功能上的重要作用。

这些结果进一步令人惊奇的证明根据本发明的核酸和/或多肽可用于诊断、预防和/或治疗根据本发明的失调。

#### 实施例 6: 使用 HCC 特异探针的诊断方法

可以建立优选基于聚合酶链式反应 (PCR) 的根据本发明的失调的诊断方法。本发明的核酸序列的标准 PCR 检测足够鉴定, 例如, 患者血流中循环的 HCC 肿瘤细胞。然而, 肿瘤活检材料(如来自细针活检)中本发明的核酸序列的表达的检测将也是该诊断方法的优选的指示剂。本发明的核酸序列, 例如, OBc15 (SEQ ID 11), 在多数患病组织中没有检测到并且在例如 HCC 中相对特异的表达。还证明了肝硬化和 HCC 中该核酸的上升的表达, 指出这种方法用于肝脏疾病诊断的潜在的辨别力(图 1, 2, 5 和表 5A/B)。

PCR 诊断将优选需要约 1 pg, 优选至少 100 ng, 更优选至少 1 μg 从患者材料分离的 RNA。在优选的利用中, 将根据标准方法从来自例如优选通过最小侵入性静脉穿刺方法得到的循环血液的白细胞级分



离 RNA。在该优选的情况下，该方法将检测血液循环系统中 HCC 肿瘤细胞的存在。可以类似地从肝脏活检材料分离 RNA。

对于 OBc15 的特异检测，例如，PCR 诊断法可以包括对 OBc15 核酸特异的几种引物，包括用于从患者样品产生的 RNA 合成 cDNA 的特异反义引物（引物 OBc15-p1；5'-GCCACAGGTTGAACACTTAATTTG-3'；SEQ ID 42；来自 SEQ ID 11 的核苷酸 350-327）。类似地，特异 PCR 引物为例如，OBc15-p2（5'-AGGAAGAGTCGTCACGAGAACC-3'；SEQ ID 43；来自 SEQ ID 11 的核苷酸 107-128）和 OBc15-p3（5'-ATAATGCTGTGCTTAGTTTATTGCC-3'；SEQ ID 44；来自 SEQ ID 11 上的核苷酸 313-289）。通过提供对 OBc15 核酸插入片段特异并在引物 OBc15-p2 和-p3 内部（巢式）的另外的引物组（例如：OBc15-p4；5'-GATCGTGGACATTTCAACCTC-3'；SEQ ID 45；来自 SEQ ID 11 上的核苷酸 147-167；和 OBc15-p5；5'-TCTTGCTTGATGCTTTGGTC-3'；SEQ ID 46；来自 SEQ ID 11 上的核苷酸 280-261）可以提高敏感性、特异性和质量控制。使用 TaqMan Q-PCR，利用例如：OBc15-8，SEQ ID 66（5'-ATCTGCAAGCCAGGAAGAGTC-3'）；OBc15-p9，SEQ ID 67（5'-CTTGCTTGATGCTTTGGTCTGT-3'）；和 OBc15-p10，SEQ ID 68，（5'-CCAGACCATGCAGGA ACTCTGATCGTGGAC-3'）也可以在如图 2 中图解的检测策略中实现 OBc15 mRNA 水平的定量评定。

可以从患者 RNA 样品制备 cDNA，之后用无 RNA 酶的 DNA 酶-1（Roche）消化 RNA 以消除基因组 DNA 的潜在污染。通过包含用于从 OBc15 基因的不同外显子的序列扩增的引物，从而来自基因组 DNA 模板的 PCR 产物（由此不反映相应于 OBc15 的 mRNA 的表达）将大于 RNA 特异的 PCR 产物，从而进一步控制污染可能性。通过例如 OBc15 特异的 OBc15-p1（约 1  $\mu$ M），利用反转录酶[如约 2 单位/反应的 Maloney 鼠白血病病毒反转录酶（Roche）]在适宜的缓冲液如 50 mM Tris-HCl、6 mM MgCl<sub>2</sub>、40 mM KCl、和 10 mM 二硫苏糖醇，pH 8.5 中引发 cDNA 合成。cDNA 合成反应中还需要的是 dATP、dCTP、dGTP 和 dTTP，每种约 1 mM，RNA 酶抑制剂，如约 1-10 个单位/反应的胎盘 RNA 酶抑制剂（Roche）。cDNA 合成将优选在 42℃ 下进行 30 到 60 分钟，然后在 95℃ 加热 10 分钟以变性 RNA 模板。所得 cDNA 可用作 PCR 的模板以检测血样（或者肝脏活检样品）中的 OBc15。优选还提供 OBc15 的 PCR 检

测所需的额外试剂,包括:10×Taq DNA 聚合酶缓冲液(500 mM Tris-C1 pH 8.3, 25 mM MgCl<sub>2</sub>, 0.1 % Triton X-100); dATP、dCTP、dGTP 和 dTTP 的混合物,每种终浓度为 0.2 mM; Taq DNA 聚合酶(2.5U/反应),和 OBc15 特异引物如 OBc15-p2、OBc15-p3、OBc15-p4、和 OBc15-p5 (0.1-1 μM 终浓度)。优选还包括用 OBc15 序列插入片段从质粒克隆 PCR 扩增这种 DNA 的阳性对照(1-10ng/反应)。该 PCR 可以例如进行 22-40 个循环,每个循环为 95℃ 30 秒,60℃ 30 秒,72℃ 60 秒。如上面所指出的,通过使用位于用最初的 PCR 引物组扩增的序列内的另外的 OBc15 引物组可以在该诊断方法中实现优选的另外的敏感性和特异性。在该情况中,随后将使用类似于第一次 PCR 反应中所用的那些条件只是优选使用引物 OBc15-p4 和 OBc15-p5 来扩增巢式序列,该反应中包括 1-10 μl 第一次 PCR 产物作为模板 DNA。备选地,优选使用第一引物组(OBc15-p2 和 OBc15-p3)实施反应 10-15 个循环,之后将 1-10 μl 该反应物包括在新的 PCR 反应中作为模板,该新的 PCR 反应使用引物 OBc15-p4 和 OBc15-p5 (并包括所有必须的 PCR 组分)。将优选利用如本领域公知的并且在前面实施例中描述的琼脂糖凝胶电泳检测 OBc15 特异的 PCR 产物。该诊断法中所包括的将优选为作为这种基于 PCR 的诊断性试验的可比较的液体或组织提取物作为对照。这可以包括来自未患病个体的血清或血浆和/或来自适当动物模型的血清、血浆或组织提取物。如果相对于对照,从患者分离的样品中根据本发明的核酸的 PCR 确定的表达如使用引物 OBc15-p4 和 OBc15-p5 得到的该反应的产物被上调,并且如果尤其该上调的表达与如图 1 中图解的失调特异性(平均)表达比基本上匹配,那么这种匹配指出该患者患有该失调。

也可以理解该方法的变通方案。可以利用热稳定的具有反转录酶活性的 DNA 聚合酶(如 Titan 一管(one-tube)或 Roche 的 Carboxydotherrmus DNA 聚合酶一组 RT-PCR 系统)在单个反应容器中顺序地或者同时实施 cDNA 合成和 PCR 扩增。备选地,通过荧光标记的引物或者 PCR 产物的各种基于荧光的指示剂(包括 Taqman 探针水解系统,如上述,和荧光双链 DNA 插入分子如 SYBR 绿)的掺入监视 PCR 产物。基于荧光的方法具有优势,因为 PCR 产物的积累可被连续监视以实现根据本发明的核酸的表达的灵敏的定量评定。这对于在根据本

发明的血液或组织中增加并且在未患病患者和组织中以低水平存在的核酸是尤其有利的，从而得到关于该核酸表达水平的定量信息。此外，关于该实例，核酸表达水平的准确定量促进例如肝硬变和 HCC 之间的差示诊断。该数据与所提供的标准数据的比较可指示存在或者不存在疾病，该比较为这种诊断方法提供了重要的优势。

该诊断策略的其他变通方案包括同时检测根据本发明的多种核酸和/或根据本发明的核酸与和该失调有关的其他核酸。还设想了根据本发明的核酸的基于杂交的诊断检测。在该情况下，mRNA 检测优选使用 RNA 印迹、RNA 酶保护或者在患者细胞或者组织活检样品上的原位杂交，该 mRNA 检测也是有效的。

通过类似方法和它们的变通方法，根据本发明的核酸和/或根据本发明的核酸与其他核酸可用于诊断根据本发明的失调。

#### 实施例 7: 通过根据本发明的多肽的抗体检测的诊断方法

根据本发明的失调的优选诊断方法基于针对根据本发明的多肽的抗体。例如，诊断方法可优选使用借助于酶联免疫吸附测定法 (ELISA) 的特异上调的基因蛋白质的血清检测。在简单形式中，该诊断测定法优选包括微量滴定板或微量滴定板的一条 (例如，用根据本发明的多肽如 OBc15. pr (SEQ ID 2) 或 DAP3 (SEQ ID 5) 特异的分离并纯化的抗体包被的)。抗体可以是例如亲和纯化的多克隆抗体，如例如通常在兔中产生的多克隆抗体，或者如根据本领域成熟的方法通常在小鼠中产生并纯化的单克隆抗体 (Cooper, H. M. 和 Paterson, Y., (2000), Current Protocols in Molecular Biology (Ansel, F. A. 等人, 编者) 11.12. 1-11.12. 9 页, Greene Publ. & Wiley Intersci., NY); (Fuller S. A. 等人, (1992), In Current Protocols in Molecular Biology (Ansel, F. A. 等人, 编者) 11.4. 1-11.9. 3 页, Greene Publ. & Wiley Intersci., NY)。优选地，该抗体可以是如 Knappik 等人 (2000, J. Molec. Biol. 296: 57-86) 或者 Chadd 和 Chamow (2001 Curr. Opin. Biotechnol. 12: 188-94) 所描述的从噬菌体展示文库淘选和纯化得到的重组抗体或者其片段。优选通过将抗 OBc15. pr 或者抗 DAP3. pr 抗体在标准包被溶液如磷酸缓冲液 (PBS) 中稀释到 1-100  $\mu\text{g}/\text{ml}$  实现抗体包被。抗体优选在 37°C 60 分

钟内，或者室温或 4℃ 下过夜结合到微量滴定板孔（如 Nunc Maxisorp 免疫板）的吸附表面。样品结合到所包被的孔之前，优选将孔在浓缩的蛋白质溶液（如磷酸缓冲液中的 5% 牛血清白蛋白或者重悬在相同缓冲液中的 5% 无脂干奶粉）中室温下孵育 15-60 分钟完全将孔封闭以防止非特异结合。优选地，患者样品材料然后被应用于微量滴定板孔，稀释到封闭溶液中以增加检测的特异性。该样品可以是例如根据本领域中熟知的方法从组织活检或手术切除制备的血浆或者血清或者蛋白质提取物 (Smith, J. A. (2001) In, Current Protocols in Molecular Biology, Ausubel, F. A. 等人, 编者) 10.0.1-10.0.23 页, Greene Publ. & Wiley Intersci., NY)。具体地，将患者样品与抗体包被的孔在室温或者 4℃ 下接触 30-120 分钟（或者更长）。优选通过用例如标准洗涤缓冲液如 0.1M Tris 缓冲液与 0.02-0.1% Tween20 充分洗涤除去非-特异相互作用的蛋白质。优选洗涤 3-10 分钟并重复 3-5 次。患者样品中 DAP3. pr 多肽的检测为例如通过与第二种独立的抗-DAP3. pr 抗体随后的结合反应来实现，该抗-DAP3. pr 抗体如上描述的产生，识别标准二位“三明治”型 ELISA 中 DAP3. pr 多肽的不同表位。第二抗-OBc15. pr 抗体或 DAP3. pr 抗体的结合为例如通过将孔在抗体（例如，封闭溶液中 1-100 μg/ml 的浓度）中室温下孵育 30-60 分钟，然后通过如前面步骤中的充分洗涤来实现。第二抗体可以优选直接偶联到在适当底物存在下能够产生生色或者生荧光反应产物的酶，如碱性磷酸酶。备选地，例如，偶联到酶的抗物种且抗同种型特异的第三抗体可以用于产生反应产物，优选可以在标准分光光度板读出仪器中检测该反应产物。对于反应产物显影，所洗涤的（如上）的抗体-抗原-酶复合物被优选在室温下暴露于生色底物，如来自 Roche 的 AttoPhos 约 10 分钟，反应可以用低 pH 缓冲液如 50 mM Tris-HCl pH 5.5 终止，或者直接测定。通过例如测量每孔中酶反应产物的量以及分光光度计中适当波长（在该情况中为 420nm）处的激发来确定被特异结合的 OBc15. pr 多肽或 DAP3. pr 多肽的量。优选在平板读出器中在发射波长（在该情况中为 560nm）下进行检测。该诊断法中优选包括 OBc15 蛋白质标准或 DAP3 蛋白质标准，如例如纯化的重组 OBc15. pr 多肽或者 DAP3. pr 多肽。在 ELISA 中优选平行包括该蛋白质标准的稀释系列作为反应的标准和推导出如本领域中熟知的用于多肽

表达水平比较的蛋白质标准曲线。在该诊断法中应该优选提供特定肝脏失调的浓度范围相应的指示。此外，优选还应该包括可比较的液体或组织提取物以作为这种 ELISA 试验的对照。这可以优选包括来自未患病个体的血清或血浆和/或来自适宜的动物模型的血清、血浆或组织提取物。这种 ELISA 检测诊断法在本领域中是普通的（见例如，Hauschild 等人，2001，Cancer Res. 158: 169-77）。通过 ELISA 确定的样品：对照蛋白质水平与 ELISA 确定的失调特异蛋白质表达比值相比较，该比值优选在患有根据本发明的失调的患者的病理学家证实的组织中相对于对照样品确定。如果样品：对照的蛋白质水平基本上匹配失调特异的蛋白质表达比，那么该匹配优选表明该患者患有该失调。优选地，为一种以上的根据本发明的多肽实施这种诊断。

此外，该诊断法可以检测存在于从患者分离的样品中的针对根据本发明的多肽的内源性抗体，或者其功能变体或者其片段，该抗体或者其片段针对根据本发明的多肽。这种自身免疫抗体的检测可以通过本领域中公知的方法，例如通过免疫亲和测定法如上面详述的 ELISA，使用根据本发明的多肽或者其功能变体或者其部分作为探针来完成。这种自身免疫抗体的存在指出该患者患有根据本发明的失调。

此外或者备选地，可以制备基于根据本发明的至少一种多肽的组织化学检测的相关诊断试剂盒。在这种试剂盒中，可以包括例如对根据本发明的多肽特异的一种或几种纯化的抗体以及优选地检测该一种或几种抗体与患者细胞或组织切片结合所必需的试剂。这些试剂包括，例如，针对根据本发明的多肽或其功能变体的特异抗物种和亚型的特异第二抗体，该第二抗体优选偶联到能够催化例如生色底物的酶或者偶联到荧光团（如例如，得克萨斯红）。优选还包括酶的底物以及洗涤和孵育缓冲液。这种试剂盒的另外的任选的组分可以是阳性对照组织的切片，例如肝脏或组织或来自特异表达该多肽的细胞的被包装的沉淀的切片作为阳性组织对照。所提供的使用说明将包括根据本发明的多肽检测的抗原回收的优选的和/或备选方法或者，例如，指出应该使用冰冻的而不是福尔马林固定并石蜡包埋的组织材料。在该情况中，应该优选包括固定冰冻的组织样品切片的推荐方法，例如，浸在冰-冷却的丙酮中 10 分钟。其他使用说明将优选提供关于用于基因产物检测的抗体的浓度以及例如，组织暴露于所提供的免疫学试剂

的推荐的和建议的孵育时间和温度。还可以包括抗体孵育的优选反应缓冲液，如含有包含 3%正常绵羊血清的磷酸缓冲液的 0.01%-0.1% 吐温 20。还将优选包括在特定抗体中孵育之前和之后组织切片洗涤的特定条件，如，例如，用含有磷酸缓冲液的 0.1%吐温 20 洗涤 4 次，每次 5 分钟。这种免疫组织化学检测方案是本领域技术人员公知的。通常，该试剂盒将优选包括作为用户指南的从阳性和阴性组织样品所得到的特定免疫组织化学染色的一组图像，尤其是指出哪种结果指示患者患有所要诊断的失调。这种试剂盒的利用将优选排除、支持或证实前面提到的根据本发明的肝脏失调、肝癌或上皮癌的诊断。

如上面关于基于核酸的诊断方法所说明的，基于根据本发明的多肽的检测和/或定量的诊断法可以包括一种或多种这种多肽。而且，在这些诊断法中可以使用这些多肽与根据本发明的失调有关的其他肽同时检测。

对于本领域技术人员显而易见的是，可以对本发明的组合物和方法进行各种修改。从而，如果这些修改和变化位于所附权利要求和它们的等价权利要求的范围内，那么本发明将覆盖这些修改和变化。此处引用的所有出版物被完整地并入作为参考。

## 序列表

<110> ORIDIS BIOMED Forschungs- und Entwicklungs GmbH  
 Guelly, Christian  
 Buck, Charles R.  
 Zatloukal, Kurt

<120> 多肽与编码这些多肽的核酸和它们用于预防、诊断或治疗肝脏失调  
 和上皮癌用途

<130> Oridis Biomed

<140> 1223FPC  
 <141> 2003-09-22

<160> 73

<170> PatentIn version 3.1

<210> 1  
 <211> 654  
 <212> PRT  
 <213> 人

<400> 1

Met Tyr Ser Ser Ser Cys Glu Thr Thr Arg Asn Thr Thr Gly Ile Glu  
 1                   5                   10                   15

Glu Ser Thr Asp Gly Met Ile Leu Gly Pro Glu Asp Leu Ser Tyr Gln  
                   20                   25                   30

Ile Tyr Asp Val Ser Gly Glu Ser Asn Ser Ala Val Ser Thr Glu Asp  
                   35                   40                   45

Leu Lys Glu Cys Leu Lys Lys Gln Leu Glu Phe Cys Phe Ser Arg Glu  
           50                   55                   60

Asn Leu Ser Lys Asp Leu Tyr Leu Ile Ser Gln Met Asp Ser Asp Gln  
 65                   70                   75                   80

Phe Ile Pro Ile Trp Thr Val Ala Asn Met Glu Glu Ile Lys Lys Leu  
                   85                   90                   95

Thr Thr Asp Pro Asp Leu Ile Leu Glu Val Leu Arg Ser Ser Pro Met  
100 105 110

Val Gln Val Asp Glu Lys Gly Glu Lys Val Arg Pro Ser His Lys Arg  
115 120 125

Cys Ile Val Ile Leu Arg Glu Ile Pro Glu Thr Thr Pro Ile Glu Glu  
130 135 140

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

Glu Phe Ala His Asn Ser Asn Trp Tyr Ile Thr Phe Gln Ser Asp Thr  
165 170 175

Asp Ala Gln Gln Ala Phe Lys Tyr Leu Arg Glu Glu Val Lys Thr Phe  
180 185 190

Gln Gly Lys Pro Ile Met Ala Arg Ile Lys Ala Ile Asn Thr Phe Phe  
195 200 205

Ala Lys Asn Gly Tyr Arg Leu Met Asp Ser Ser Ile Tyr Ser His Pro  
210 215 220

Ile Gln Thr Gln Ala Gln Tyr Ala Ser Pro Val Phe Met Gln Pro Val  
225 230 235 240

Tyr Asn Pro His Gln Gln Tyr Ser Val Tyr Ser Ile Val Pro Gln Ser  
245 250 255

Trp Ser Pro Asn Pro Thr Pro Tyr Phe Glu Thr Pro Leu Ala Pro Phe  
260 265 270

Pro Asn Gly Ser Phe Val Asn Gly Phe Asn Ser Pro Gly Ser Tyr Lys  
275 280 285



Thr Asn Ala Ala Ala Met Asn Met Gly Arg Pro Phe Gln Lys Asn Arg  
290 295 300

Val Lys Pro Gln Phe Arg Ser Ser Gly Gly Ser Glu His Ser Thr Glu  
305 310 315 320

Gly Ser Val Ser Leu Gly Asp Gly Gln Leu Asn Arg Tyr Ser Ser Arg  
325 330 335

Asn Phe Pro Ala Glu Arg His Asn Pro Thr Val Thr Gly His Gln Glu  
340 345 350

Gln Thr Tyr Leu Gln Lys Glu Thr Ser Thr Leu Gln Val Glu Gln Asn  
355 360 365

Gly Asp Tyr Gly Arg Gly Arg Arg Thr Leu Phe Arg Gly Arg Arg Arg  
370 375 380

Arg Glu Asp Asp Arg Ile Ser Arg Pro His Pro Ser Thr Ala Glu Ser  
385 390 395 400

Lys Ala Pro Thr Pro Lys Phe Asp Leu Leu Ala Ser Asn Phe Pro Pro  
405 410 415

Leu Pro Gly Ser Ser Ser Arg Met Pro Gly Glu Leu Val Leu Glu Asn  
420 425 430

Arg Met Ser Asp Val Val Lys Gly Val Tyr Lys Glu Lys Asp Asn Glu  
435 440 445

Glu Leu Thr Ile Ser Cys Pro Val Pro Ala Asp Glu Gln Thr Glu Cys  
450 455 460

Thr Ser Ala Gln Gln Leu Asn Met Ser Thr Ser Ser Pro Cys Ala Ala  
465 470 475 480

Glu Leu Thr Ala Leu Ser Thr Thr Gln Gln Glu Lys Asp Leu Ile Glu  
 485 490 495

Asp Ser Ser Val Gln Lys Asp Gly Leu Asn Gln Thr Thr Ile Pro Val  
 500 505 510

Ser Pro Pro Ser Thr Thr Lys Pro Ser Arg Ala Ser Thr Ala Ser Pro  
 515 520 525

Cys Asn Asn Asn Ile Asn Ala Ala Thr Ala Val Ala Leu Gln Glu Pro  
 530 535 540

Arg Lys Leu Ser Tyr Ala Glu Val Cys Gln Lys Pro Pro Lys Glu Pro  
 545 550 555 560

Ser Ser Val Leu Val Gln Pro Leu Arg Glu Leu Arg Ser Asn Val Val  
 565 570 575

Ser Pro Thr Lys Asn Glu Asp Asn Gly Ala Pro Glu Asn Ser Val Glu  
 580 585 590

Lys Pro His Glu Lys Pro Glu Ala Arg Ala Ser Lys Asp Tyr Ser Gly  
 595 600 605

Phe Arg Gly Asn Ile Ile Pro Arg Gly Ala Ala Gly Lys Ile Arg Glu  
 610 615 620

Gln Arg Arg Gln Phe Ser His Arg Ala Ile Pro Gln Gly Val Thr Arg  
 625 630 635 640

Arg Asn Gly Lys Glu Gln Tyr Val Pro Pro Arg Ser Pro Lys  
 645 650

<210> 2

<211> 72

&lt;212&gt; PRT

&lt;213&gt; 人

&lt;400&gt; 2

Met Gly Val Glu Leu Met Met Glu Leu Glu Pro Leu Gln Gly Asn Glu  
 1                   5                   10                   15

Glu Thr Arg Ala Leu Phe Met Pro Arg Glu Asp Thr Ala Arg Pro Gln  
                   20                   25                   30

Ser Ala Ser Gln Glu Glu Ser Ser Arg Glu Pro Asp His Ala Gly Thr  
           35                   40                   45

Leu Ile Val Asp Ile Ser Thr Ser Arg Thr Val Ile Gln Asn Ala Tyr  
       50                   55                   60

Val Ser Leu Glu Glu Thr Leu Lys  
 65                   70

&lt;210&gt; 3

&lt;211&gt; 367

&lt;212&gt; PRT

&lt;213&gt; 人

&lt;400&gt; 3

Met Leu Pro Pro Arg Arg Leu Gln Thr Leu Leu Arg Gln Ala Val Glu  
 1                   5                   10                   15

Leu Gln Arg Asp Arg Cys Leu Tyr His Asn Thr Lys Leu Asp Asn Asn  
           20                   25                   30

Leu Asp Ser Val Ser Leu Leu Ile Asp His Val Cys Ser Arg Arg Gln  
           35                   40                   45

Phe Pro Cys Tyr Thr Gln Gln Ile Leu Thr Glu His Cys Asn Glu Val  
       50                   55                   60

Trp Phe Cys Lys Phe Ser Asn Asp Gly Thr Lys Leu Ala Thr Gly Ser  
65 70 75 80

Lys Asp Thr Thr Val Ile Ile Trp Gln Val Asp Pro Asp Thr His Leu  
85 90 95

Leu Lys Leu Leu Lys Thr Leu Glu Gly His Ala Tyr Gly Val Ser Tyr  
100 105 110

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

Asp Cys Ser Glu Leu Trp Leu Trp Asn Val Gln Thr Gly Glu Leu Arg  
130 135 140

Thr Lys Met Ser Gln Ser His Glu Asp Ser Leu Thr Ser Val Ala Trp  
145 150 155 160

Asn Pro Asp Gly Lys Arg Phe Val Thr Gly Gly Gln Arg Gly Gln Phe  
165 170 175

Tyr Gln Cys Asp Leu Asp Gly Asn Leu Leu Asp Ser Trp Glu Gly Val  
180 185 190

Arg Val Gln Cys Leu Trp Cys Leu Ser Asp Gly Lys Thr Val Leu Ala  
195 200 205

Ser Asp Thr His Gln Arg Ile Arg Gly Tyr Asn Phe Glu Asp Leu Thr  
210 215 220

Asp Arg Asn Ile Val Gln Glu Asp His Pro Ile Met Ser Phe Thr Ile  
225 230 235 240

Ser Lys Asn Gly Arg Leu Ala Leu Leu Asn Val Ala Thr Gln Gly Val  
245 250 255

His Leu Trp Asp Leu Gln Asp Arg Val Leu Val Arg Lys Tyr Gln Gly  
 260 265 270

Val Thr Gln Gly Phe Tyr Thr Ile His Ser Cys Phe Gly Gly His Asn  
 275 280 285

Glu Asp Phe Ile Ala Ser Gly Ser Glu Asp His Lys Val Tyr Ile Trp  
 290 295 300

His Lys Arg Ser Glu Leu Pro Ile Ala Glu Leu Thr Gly His Thr Arg  
 305 310 315 320

Thr Val Asn Cys Val Ser Trp Asn Pro Gln Ile Pro Ser Met Met Ala  
 325 330 335

Ser Ala Ser Asp Asp Gly Thr Val Arg Ile Trp Gly Pro Ala Pro Phe  
 340 345 350

Ile Asp His Gln Asn Ile Glu Glu Glu Cys Ser Ser Met Asp Ser  
 355 360 365

<210> 4

<211> 188

<212> PRT

<213> 人

<400> 4

Met Asp Val Asn Ile Ala Pro Leu Arg Ala Trp Asp Asp Phe Phe Pro  
 1 5 10 15

Gly Ser Asp Arg Phe Ala Arg Pro Asp Phe Arg Asp Ile Ser Lys Trp  
 20 25 30

Asn Asn Arg Val Val Ser Asn Leu Leu Tyr Tyr Gln Thr Asn Tyr Leu  
 35 40 45

Val Val Ala Ala Met Met Ile Ser Ile Val Gly Phe Leu Ser Pro Phe

50    55    60

Asn Met Ile Leu Gly Gly Ile Val Val Val Leu Val Phe Thr Gly Phe  
 65    70    75    80

Val Trp Ala Ala His Asn Lys Asp Val Leu Arg Arg Met Lys Lys Arg  
     85    90    95

Tyr Pro Thr Thr Phe Val Met Val Val Met Leu Ala Ser Tyr Phe Leu  
     100    105    110

Ile Ser Met Phe Gly Gly Val Met Val Phe Val Phe Gly Ile Thr Phe  
     115    120    125

Pro Leu Leu Leu Met Phe Ile His Ala Ser Leu Arg Leu Arg Asn Leu  
     130    135    140

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

Pro Met Gly Ile Val Leu Asp Ala Leu Glu Gln Gln Glu Glu Gly Ile  
     165    170    175

Asn Arg Leu Thr Asp Tyr Ile Ser Lys Val Lys Glu  
     180    185

<210> 5  
 <211> 398  
 <212> PRT  
 <213> 人

<400> 5

Met Met Leu Lys Gly Ile Thr Arg Leu Ile Ser Arg Ile His Lys Leu  
 1    5    10    15

Asp Pro Gly Arg Phe Leu His Met Gly Thr Gln Ala Arg Gln Ser Ile  
     20    25    30

Ala Ala His Leu Asp Asn Gln Val Pro Val Glu Ser Pro Arg Ala Ile  
 35 40 45

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

Gly Gln His Tyr Asn Ile Ser Pro Gln Asp Leu Glu Thr Val Phe Pro  
 65 70 75 80

His Gly Leu Pro Pro Arg Phe Val Met Gln Val Lys Thr Phe Ser Glu  
 85 90 95

Ala Cys Leu Met Val Arg Lys Pro Ala Leu Glu Leu Leu His Tyr Leu  
 100 105 110

Lys Asn Thr Ser Phe Ala Tyr Pro Ala Ile Arg Tyr Leu Leu Tyr Gly  
 115 120 125

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

Cys Ala Lys Gln Asp Trp Leu Ile Leu His Ile Pro Asp Ala His Leu  
 145 150 155 160

Trp Val Lys Asn Cys Arg Asp Leu Leu Gln Ser Ser Tyr Asn Lys Gln  
 165 170 175

Arg Phe Asp Gln Pro Leu Glu Ala Ser Thr Trp Leu Lys Asn Phe Lys  
 180 185 190

Thr Thr Asn Glu Arg Phe Leu Asn Gln Ile Lys Val Gln Glu Lys Tyr  
 195 200 205

Val Trp Asn Lys Arg Glu Ser Thr Glu Lys Gly Ser Pro Leu Gly Glu  
 210 215 220

Val Val Glu Gln Gly Ile Thr Arg Val Arg Asn Ala Thr Asp Ala Val  
225 230 235 240

Gly Ile Val Leu Lys Glu Leu Lys Arg Gln Ser Ser Leu Gly Met Phe  
245 250 255

His Leu Leu Val Ala Val Asp Gly Ile Asn Ala Leu Trp Gly Arg Thr  
260 265 270

Thr Leu Lys Arg Glu Asp Lys Ser Pro Ile Ala Pro Glu Glu Leu Ala  
275 280 285

Leu Val His Asn Leu Arg Lys Met Met Lys Asn Asp Trp His Gly Gly  
290 295 300

Ala Ile Val Ser Ala Leu Ser Gln Thr Gly Ser Leu Phe Lys Pro Arg  
305 310 315 320

Lys Ala Tyr Leu Pro Gln Glu Leu Leu Gly Lys Glu Gly Phe Asp Ala  
325 330 335

Leu Asp Pro Phe Ile Pro Ile Leu Val Ser Asn Tyr Asn Pro Lys Glu  
340 345 350

Phe Glu Ser Cys Ile Gln Tyr Tyr Leu Glu Asn Asn Trp Leu Gln His  
355 360 365

Glu Lys Ala Pro Thr Glu Glu Gly Lys Lys Glu Leu Leu Phe Leu Ser  
370 375 380

Asn Ala Asn Pro Ser Leu Leu Glu Arg His Cys Ala Tyr Leu  
385 390 395

<210> 6

<211> 261



&lt;212&gt; PRT

&lt;213&gt; 人

&lt;400&gt; 6

Met Ala Gly Pro Glu Leu Leu Leu Asp Ser Asn Ile Arg Leu Trp Val  
 1                   5                   10                   15

Val Leu Pro Ile Val Ile Ile Thr Phe Phe Val Gly Met Ile Arg His  
                   20                   25                   30

Tyr Val Ser Ile Leu Leu Gln Ser Asp Lys Lys Leu Thr Gln Glu Gln  
                   35                   40                   45

Val Ser Asp Ser Gln Val Leu Ile Arg Ser Arg Val Leu Arg Glu Asn  
                   50                   55                   60

Gly Lys Tyr Ile Pro Lys Gln Ser Phe Leu Thr Arg Lys Tyr Tyr Phe  
 65                   70                   75                   80

Asn Asn Pro Glu Asp Gly Phe Phe Lys Lys Thr Lys Arg Lys Val Val  
                   85                   90                   95

Pro Pro Ser Pro Met Thr Asp Pro Thr Met Leu Thr Asp Met Met Lys  
                   100                   105                   110

Gly Asn Val Thr Asn Val Leu Pro Met Ile Leu Ile Gly Gly Trp Ile  
                   115                   120                   125

Asn Met Thr Phe Ser Gly Phe Val Thr Thr Lys Val Pro Phe Pro Leu  
                   130                   135                   140

Thr Leu Arg Phe Lys Pro Met Leu Gln Gln Gly Ile Glu Leu Leu Thr  
 145                   150                   155                   160

Leu Asp Ala Ser Trp Val Ser Ser Ala Ser Trp Tyr Phe Leu Asn Val  
                   165                   170                   175

Phe Gly Leu Arg Ser Ile Tyr Ser Leu Ile Leu Gly Gln Asp Asn Ala  
 180 185 190

Ala Asp Gln Ser Arg Met Met Gln Glu Gln Met Thr Gly Ala Ala Met  
 195 200 205

Ala Met Pro Ala Asp Thr Asn Lys Ala Phe Lys Thr Glu Trp Glu Ala  
 210 215 220

Leu Glu Leu Thr Asp His Gln Trp Ala Leu Asp Asp Val Glu Glu Glu  
 225 230 235 240

Leu Met Ala Lys Asp Leu His Phe Glu Gly Met Phe Lys Lys Glu Leu  
 245 250 255

Gln Thr Ser Ile Phe  
 260

<210> 7

<211> 403

<212> PRT

<213> 人

<400> 7

Met Ser Gly Arg Val Gly Asp Leu Ser Pro Arg Gln Lys Glu Ala Leu  
 1 5 10 15

Ala Lys Phe Arg Glu Asn Val Gln Asp Val Leu Pro Ala Leu Pro Asn  
 20 25 30

Pro Asp Asp Tyr Phe Leu Leu Arg Trp Leu Arg Ala Arg Ser Phe Asp  
 35 40 45

Leu Gln Lys Ser Glu Ala Met Leu Arg Lys His Val Glu Phe Arg Lys  
 50 55 60

Gln Lys Asp Ile Asp Asn Ile Ile Ser Trp Gln Pro Pro Glu Val Ile  
65 70 75 80

Gln Gln Tyr Leu Ser Gly Gly Met Cys Gly Tyr Asp Leu Asp Gly Cys  
85 90 95

Pro Val Trp Tyr Asp Ile Ile Gly Pro Leu Asp Ala Lys Gly Leu Leu  
100 105 110

Phe Ser Ala Ser Lys Gln Asp Leu Leu Arg Thr Lys Met Arg Glu Cys  
115 120 125

Glu Leu Leu Leu Gln Glu Cys Ala His Gln Thr Thr Lys Leu Gly Arg  
130 135 140

Lys Val Glu Thr Ile Thr Ile Ile Tyr Asp Cys Glu Gly Leu Gly Leu  
145 150 155 160

Lys His Leu Trp Lys Pro Ala Val Glu Ala Tyr Gly Glu Phe Leu Cys  
165 170 175

Met Phe Glu Glu Asn Tyr Pro Glu Thr Leu Lys Arg Leu Phe Val Val  
180 185 190

Lys Ala Pro Lys Leu Phe Pro Val Ala Tyr Asn Leu Ile Lys Pro Phe  
195 200 205

Leu Ser Glu Asp Thr Arg Lys Lys Ile Met Val Leu Gly Ala Asn Trp  
210 215 220

Lys Glu Val Leu Leu Lys His Ile Ser Pro Asp Gln Val Pro Val Glu  
225 230 235 240

Tyr Gly Gly Thr Met Thr Asp Pro Asp Gly Asn Pro Lys Cys Lys Ser  
245 250 255

Lys Ile Asn Tyr Gly Gly Asp Ile Pro Arg Lys Tyr Tyr Val Arg Asp  
 260 265 270

Gln Val Lys Gln Gln Tyr Glu His Ser Val Gln Ile Ser Arg Gly Ser  
 275 280 285

Ser His Gln Val Glu Tyr Glu Ile Leu Phe Pro Gly Cys Val Leu Arg  
 290 295 300

Trp Gln Phe Met Ser Asp Gly Ala Asp Val Gly Phe Gly Ile Phe Leu  
 305 310 315 320

Lys Thr Lys Met Gly Glu Arg Gln Arg Ala Gly Glu Met Thr Glu Val  
 325 330 335

Leu Pro Asn Gln Arg Tyr Asn Ser His Leu Val Pro Glu Asp Gly Thr  
 340 345 350

Leu Thr Cys Ser Asp Pro Gly Ile Tyr Val Leu Arg Phe Asp Asn Thr  
 355 360 365

Tyr Ser Phe Ile His Ala Lys Lys Val Asn Phe Thr Val Glu Val Leu  
 370 375 380

Leu Pro Asp Lys Ala Ser Glu Glu Lys Met Lys Gln Leu Gly Ala Gly  
 385 390 395 400

Thr Pro Lys

<210> 8

<211> 251

<212> PRT

<213> 人

<400> 8

Met Asp Met Lys Arg Arg Ile His Leu Glu Leu Arg Asn Arg Thr Pro

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

195	200	205
Val Ser Glu Glu Glu Glu Glu Phe Gly Leu Asp Glu Glu Asp Glu Asp		
210	215	220
Glu Asp Glu Asp Glu Glu Glu Glu Glu Gly Gly Lys Gly Glu Lys Arg		
225	230	235 240
Lys Arg Glu Thr Asp Asp Glu Gly Glu Asp Asp		
	245	250
<p>&lt;210&gt; 9          &lt;211&gt; 151          &lt;212&gt; PRT          &lt;213&gt; 人</p> <p>&lt;400&gt; 9</p>		
Met Pro Arg Gly Ser Arg Ser Arg Thr Ser Arg Met Ala Pro Pro Ala		
1	5	10 15
Ser Arg Ala Pro Gln Met Arg Ala Ala Pro Arg Pro Ala Pro Val Ala		
	20	25 30
Gln Pro Pro Ala Ala Ala Pro Pro Ser Ala Val Gly Ser Ser Ala Ala		
	35	40 45
Ala Pro Arg Gln Pro Gly Leu Met Ala Gln Met Ala Thr Thr Ala Ala		
50	55	60
Gly Val Ala Val Gly Ser Ala Val Gly His Thr Leu Gly His Ala Ile		
65	70	75 80
Thr Gly Gly Phe Ser Gly Gly Ser Asn Ala Glu Pro Ala Arg Pro Asp		
	85	90 95
Ile Thr Tyr Gln Glu Pro Gln Gly Thr Gln Pro Ala Gln Gln Gln Gln		
	100	105 110

Pro Cys Leu Tyr Glu Ile Lys Gln Phe Leu Glu Cys Ala Gln Asn Gln  
 115 120 125

Gly Asp Ile Lys Leu Cys Glu Gly Phe Asn Glu Val Leu Lys Gln Cys  
 130 135 140

Arg Leu Ala Asn Gly Leu Ala  
 145 150

<210> 10

<211> 6497

<212> DNA

<213> 人

<400> 10

ccgggtggag gggcaaggcg agtgtgtgtc cttatcctag caattggggc ggggcctgt 60  
 gagccagttg gagttgcggc ggcgggaacg attgggctga gcagaggacg acatgttgct 120  
 tttcgtggag catttatggg gttaagtggc atgggatttc tgtttctgat agtaaatage 180  
 aggtagcacc taaaggaact gttttaaatc ctaatgcaa agtatggcaa gaaattgctc 240  
 ctggaataac tgatgccacc ccagtaactc atggaactga aagctcttgg catgaaatag 300  
 cagctacacc aggtgctcat cctgagggta atgcagagct ctcagaagat atatgtaaag 360  
 aatatgaagt aatgtattct tcactttgtg aaaccacaag aaatactaca ggcattgaag 420  
 aatcaactga tgggatgatt ttaggaccag aagatctgag ttaccacaata tatgatgttt 480  
 ccggagaaaag caattcagca gtttctacag aagacctaaa agaattgtctg aagaacaat 540  
 tagaattctg ttttccagc gaaaatttgt caaaggatct ttacttgata tctcaaatgg 600  
 atagtgatca gttcatccca atttgacag ttgccaacat ggaagaaata aaaaagttga 660  
 ctacagacc ctagcttaatt ettgaagtgt taagatcttc tccatggta caagttgatg 720  
 agaagggtga gaaagtgaga ccaagtcata agcgttgat tgtaattctt agagagattc 780  
 ctgaaacaac accaatagag gaagtgaaag gtttgttcaa aagtgaaaac tgccccaaag 840

tgataagctg tgagtttgca cacaatagca actggtatat cactttccag tcagacacag 900  
 atgcacaaca ggcttttaaa tacttaagag aagaagttaa aacatttcag ggcaagccaa 960  
 ttatggcaag gataaaagcc atcaatacat tttttgctaa gaatggttat cgattaatgg 1020  
 attctagtat ctatagtcac cccattcaaa ctcaagcaca gtatgcctcc ccagtcttta 1080  
 tgcagcctgt atataatcct caccaacagt actcggcteta tagtattgtg cctcagtctt 1140  
 ggtctccaaa tcttacacct tactttgaaa caccactggc tcccttccc aatggtagtt 1200  
 ttgtgaatgg ctttaattcg ccaggatctt ataaaacaaa tgctgctgct atgaatatgg 1260  
 gtcgaccatt ccaaaaaaat cgtgtgaagc ctcagtttag gtcactggt ggttcagaac 1320  
 actcaacaga gggtctgta tccctggggg atggacagtt gaacagatat agttcaagaa 1380  
 actttccagc tgaacggcat aacccacag taactgggca tcaggagcaa acttaccttc 1440  
 agaaggagac ttccacttg caggtggaac agaatgggga ctatggtagg ggcaggagaa 1500  
 ctctcttcag aggtcgaaga cgacgagaag atgacaggat ctcaagacct catcctcaa 1560  
 cagctgaate aaaggtcca acaccaaagt ttgacttatt agcctcaat tttccacct 1620  
 tacctggaag ttcacaaaga atgccaggtg aactcgttt ggagaatagg atgtctgatg 1680  
 ttgttaaagg tgtctacaaa gaaaaggata atgaagagtt gacaattagt tgcccagtgc 1740  
 ctgcagatga gcagacagaa tgcactctg cccagcaact caatatgagt accagttctc 1800  
 catgtgctgc tgagcttact gcattaagca caactcagca agaaaaggat ctaatagaag 1860  
 attcctctgt tcagaaggat ggtctcaatc agacaactat accagtttct cctccaagta 1920  
 ctacaaagcc atcgagggca agtactgctt caccatgtaa taataacata aatgcagcta 1980  
 cagctgtggc tetacaggaa cccgaaagt taagttatgc tgaagtgtgc cagaagcccc 2040  
 ctaaagagcc atcttcagtt cttgtgcagc cactacggga acttcgctcc aatgtggtgt 2100  
 ctcccacaa aatgaagac aatggagctc ctgagaactc cgttgagaaa ccacatgaga 2160  
 agccagaagc aagggttagt aaggattatt ctggcttccg aggcaatata atccccaggg 2220  
 gagcagcagg aaaaatcagg gaacagagac gccagtttag ccatagggct atacctcagg 2280



gagtgactcg acgtaatggc aaagagcaat atgtgccacc cagatcacca aagtaaaaaa 2340  
caacaaaact attcaaaaac ttcactctct tcccattaaa cttgaaactgt ggctatattg 2400  
aactgttttg gaggggaggg ggtagccagg aaggaaacaa gagaaagtac gtccatttca 2460  
ttatggattt tggagttgtg agtgatagga tcccataaatt catctctaatt gtggttttta 2520  
aatgctggag gattccaatc aatataaata tatatatata tatacacaca catatataaa 2580  
aagtataatt tttctatttt tgtttttggt ttttaattgc agagatttgc tgccaggaat 2640  
caattttgag ggctcagatt tagcttggaa gaaaaaaaaag aaacatacat ccttcagtat 2700  
aggagatgag ggaatgagag aaaatatttt ttgaagaagc atttctgtaa aattagaaat 2760  
tacttttttt aatctattta aagtttgget tgaagaatgc catctctgac tatatggcct 2820  
tgtattgcaa agcagatcag tggctgggggt gcctgtgtg ggtgtgagtg tgtacaagag 2880  
cgattgaagc caaatctggt gtcattgtag taaatgattt gaaaactgaa tgtaataact 2940  
gagtagattt tttttctag ttgaaattt agtctgtctt tttgacctta ctaatatctc 3000  
atttaacaag ttgtaaact ctgattgtac ttagagatgt gactaccaat cagtttgata 3060  
ctcaaggaaa gggggttatt caagaaattg aaaatttcat cttggacctc agtgcacg 3120  
tcaaatgat ttcagaggtt taaacttccc tgtgattccc cctgaatacc cccaaaatga 3180  
gaaacaaaat ttttttctt actccatttg ttactctctg ttctttgact gccaccac 3240  
agaaaagcaa aataaccaac taactactca attgtgtgtt tgtaattgct ttgagcagtc 3300  
tagtcaaact atataaattg ttctaaattt cagaattgaa cattgaagta ttaactctc 3360  
tgttcacaca tttagaattt tagctcccaa gatggtaggg cagactgacc gtacagtaat 3420  
ttatttgctg ttagtgtaa agattaagca tagtaactga ctcttaagtg ttaaataatg 3480  
tagaagtaaa aaaatttttt ttaaaggett aatttgggag gggggactta tttctgttta 3540  
cagtgattta ccttccttc cctctctctt cccccacac ccaacaaaat acagtttgga 3600  
attcactgaa acagtaccag caagtcatga gatttttag taaagatgag aaagatggtt 3660  
gaagaaaatt agtgcataat ttctcagtga ataaagtgt agctctcata tactaaatag 3720

acaagtttac atgctgttat ttagaaaatg actaaaatat taaaaaccgt gttgtgtaa 3780  
 tctgttttaa gtcataccat gttcagagtt ctatgtaagg tgggttttat ttttcttta 3840  
 agggatagtt tgtaatagta agaactgtcc catatgttag taaattacat atgtacaaat 3900  
 tgaaactgta aattgtgaac actggaaagc accattgtga catagagtaa acatcttagt 3960  
 aatatattaa agtgaatgta aatgggtggt aaaattacat tactgtgaaa ttcattctcc 4020  
 aactctaagt taagctttgg agatacatgt tagtggttaa ctgtaagag ctttgaaaac 4080  
 actgcacata tctgtacaag ccagaattac tatttctttg acttattatt agcttggcag 4140  
 ttgcttttga tttgattggt ttatgacatg gtatactact atatttactc agtttgaaac 4200  
 tattcatttc tacacactat ttttaaaaat tgctactag gtgaaacata acaataaaaac 4260  
 tacctgtgct gaaatttggg ggaagtttag gtcctttaa aaaacatatt aatcattgac 4320  
 tacatctatg ataaaagtgc ttattttggt ttactaagat aatgcagttg gtggaaatga 4380  
 taaacgtttt aagtgttaac atcctttgaa tgcgttggat ttcagagaat aacattttg 4440  
 taaaaatcac ttggaagga ttataaactt aattactgca cttaaaatga aacattactt 4500  
 tttttaaaca atgtgtcaca aatgtagtgc tgtattactt gtatgcttgt gtgacttact 4560  
 gttagtcag ctctaaaaat ttaaaggttg taattgaaat acaagaaaag agccttcttt 4620  
 tagaagaaag caagtatatt tttgcttita ctccaatgt tatttaaagt agaaatttaa 4680  
 tttgtagata taacctttaa aaattttctc attaagacaa tgtttttaat ttaatttgc 4740  
 tcattacatc taatagttcc catttgatgg catgtatagg gaagagtgag agagtgtgtg 4800  
 tgtgtgatg tgtgtgtaat atttatata attcacagta tgtattttagc atttatttta 4860  
 ttacagcaga tttaaagttt gtatctaaat aatgcctatg agttgtgtga agctcttggc 4920  
 ttttttcca acgttacttt gtaactaatg aggggtgatg ttcattgtag tttatttatt 4980  
 tggttcttta gatggaggaa tttaaaaaat caaattttcc tcttcacctt tatgacttga 5040  
 catttcttg atctgttggg ggctaaaagt aggtataaat gatattgaat gttgggtata 5100  
 gtgatactct gccatagtcc ttaactgcatg aagagaacaa gagtcacaca agttcaccac 5160

tttgcacttc atagagaagg tacatagaga cattgcaaaa cctgtctcca tttgctatcc 5220  
 tgataattaa ggttttcata atacctaggg cctgtctctg agtaatttta attttgccaa 5280  
 atacactgac atttaaaata gtgatecate taaatTTTTT tcagctgggt tttgaggaat 5340  
 ataagagctt tcaatgataa aggtttgttg tagttgtctt atgtgctgaa tttgcagatg 5400  
 atcagatgct gtgcagaatt ctgatttatt tttgtttcet aaaattaaga tagcttgaat 5460  
 attatttcaac attccttttt cttttttaa taaacaggtt tgctttggaa aggcttaatg 5520  
 atggaatggt agcatcttca ctagggtaaa gaagaacaaa aagaatgttg ctggaacgta 5580  
 aatagtatt taaaagttaa tgaacacttc tctagtttct ttagttatgg ccttaataat 5640  
 tagtctcttg gcttaaatgt ccaactggtt tactttgaca cagttgaaca acaactgggg 5700  
 taagtctctg gtatttaggc tggcaatata tatattaacc atattttaa agtaccaatt 5760  
 ttgtttttac agaaaagata aaactcaaaa gagaacagtg tttccttct gaggggcttt 5820  
 tataaattat taactataat atatgatgga ttttttcta atttttata tttccttaca 5880  
 attttggtgg ccattaattt aactttagge ttttgggcat atgctagtct gagcttcoga 5940  
 aaagatacat atatgtttcc cttttcatta gctgaatgag gatattttaa gaagttgaaa 6000  
 gagaatttat tttcaagttg tgagtaaate ctcctttgaa attcaactga ttattagata 6060  
 acttaaagtt tattttttaa agctgacaac tttttatgaa tcttcagatt gacagttcet 6120  
 aaaagcgtaa ctcagatatt aatgggctgt gtattaaatg gttttatttt cagttttgca 6180  
 gcacagaaca ctgttgaat atccatatca acttgatttt tttaacctaa ttcaggtgtc 6240  
 ctttgcactc cttaaagtgt ggggggtgggg gtcagagcca gttatccgge ttctgttttg 6300  
 tcgattgctt agatttgctc ctgttgctca aactgttacc cccaaaattg gtgtgacaca 6360  
 tgctcatgca taaaatgtta aatgagtac atccttgat ttgtatttgt tttcaacatc 6420  
 gccaaaggtgc tatgggaaat taacaaaatt agaaaaaaaa taaaattatt aaaaagcaaa 6480  
 aaaaaaaaaa aaaaaaa 6497

<211> 484  
 <212> DNA  
 <213> 人  
  
 <400> 11  
 atgggggtgg aactcatgat ggaattggag cctttacaag ggaatgaaga gacaagagct 60  
 ctctttatgc cacgtgagga tacagcaagg ccccaatctg caagccagga agagtctca 120  
 cgagaaccag accatgcagg aactctgac gtggacattt caacctccag aactgtgac 180  
 caaaatgcat atgtatcttt ggaagaaact ctgaagtaaa ggccggaata ttctttgttt 240  
 aaaacattaa aaacaaaaca gaccaaagca tcaagcaaga agtttcctgg caataaacta 300  
 agcacagcat tattttttaa ggaacacaaa ttaagtgttc aacctgtgga aaattgtac 360  
 tttctcctg aattatgttg ttatcaaaga aaaaaattgg gaagcatggc aaaatatcat 420  
 caaaactgaa actagaatta aactaaatta aaataaaaaa aaaaaaaaaa aaaaaaaaaa 480  
 aaaa 484

<210> 12  
 <211> 1904  
 <212> DNA  
 <213> 人

<400> 12  
 ctacgtgcaa aagcagaatg ggaaggctaa gggacagctt cccgatctaa actattggat 60  
 aaacttcaga cctatttacc accatcagtg atgcttcccc cacggcgttt acagactctc 120  
 ctgcggcagg cgggtggaact acaaagggat cgggtgcctat atcacaatac caaacttgat 180  
 aataatctag attctgtgtc tctgcttata gaccatgttt gtagtaggag gcagttccca 240  
 tgttatacgc agcagatact taoggagcat tgtaatgaag tgtggttctg taaattctct 300  
 aatgatggca ctaaactagc aacaggatca aaagatacaa cagttatcat atggcaagtt 360  
 gatccggata cacacctgct aaaactgctt aaacattag aaggacatgc ttatggcggt 420  
 tcttatattg catggagtcc agatgacaac tatcttgttg cttgtggccc agatgactgc 480  
 tetgagcttt ggctttggaa tgtacaaaca ggagaactaa ggacaaaaat gagccagtct 540

catgaagaca gtttgacaag tgtggccttg aatccagatg ggaagcgctt tgtgactgga 600  
 ggtcagcgtg ggcagttcta tcagtgtgac ttagatggta atctccttga ctctgggaa 660  
 ggggtaagag tgcaatgcct ttggtgcttg agtgatggaa agactgttct ggcatcagat 720  
 acacaccagc gaattcgggg ctataacttc gaggacctta cagataggaa catagtacaa 780  
 gaagatcadc ctattatgtc ttttactatt tcaaaaaatg gccgattagc tttgttaaat 840  
 gtagcaactc agggagtcca tttatgggac ttgcaagaca gagttttagt aagaaagtat 900  
 caaggtgtta cacaagggtt ttatacaatt cattcatggt ttggaggcca taatgaagac 960  
 ttcacgcta gtggcagtga agatcacaaag gtttacatct ggcacaaacg tagtgaactg 1020  
 ccaattgcgg agctgacagg gcacacacgt acagtaaact gtgtgagctg gaaccacag 1080  
 attccatcca tgatggccag cgcctcagat gatggcactg ttagaatatg gggaccagca 1140  
 ccttttatag accaccagaa tattgaagag gaatgcagta gcatggatag ttgatggtga 1200  
 atttgagca gacgacctct gtttaacta aaattagtcg tattttaatg gcttgggatt 1260  
 tgggtcaaac aaacatgatt gatagctgga cagacatgct cgtcatgaaa aaagaaccat 1320  
 ttctgaagcc cgattggggc caaacattta caccttgctt catagtaacc agttgagatg 1380  
 aagcaogtcg ttagaacggt gttggacacc atgttgaatt attccccat cggttgtgaa 1440  
 gaactgtgct acattcaggc ttaccattg aactcagtat atatattttt ttccttctg 1500  
 tctttgtct ggcaggatac cattcttggt gctctctgt gtaatgaagt ttaatgctt 1560  
 gtttgaaaaa ctttatttaa cagtttagaa ggcttgatag aaagagtgca ttagtctgaa 1620  
 gagtatacat tggataggaa agaatttctt tctttgttt ctccaaatct ttccgctta 1680  
 tttagcttga gatctttgca gcttggttca tggattctag ccttgcccgt tgcgcagtat 1740  
 atactgatcc agatgataaa ccagtgaact atgtcaaaag cactctcaat attacattg 1800  
 acaaaaagtt ttgtactttt cacatagctt gttgccccgt aaaagggtta acagcacaat 1860  
 tttttaaaaa taaattaaga agtatttaaa aaaaaaaaaa aaaa 1904

<211> 2088

<212> DNA

<213> 人

<400> 13

cgctgtcaac tctccaactc agctcagctg atcggttgcc gccgcgcgcg ccgccagatt 60

ctggaggcga agaacgcaaa gctgagaaca tggacgttaa tatcgcacca ctccgcgcct 120

gggacgattt ctcccggtt tccgatcctt ttgcccgccc ggaactcagg gacatttcca 180

aatggaacaa ccgcgtagtg agcaacctgc tctattacca gaccaactac ctggtggtgg 240

ctgccaatgat gatttocatt gtggggtttc tgagtcacct caacatgac ctgggaggaa 300

tctgtgtggt gctggtgttc acagggtttg tgtgggcagc ccacaataaa gaagtccttc 360

gccggatgaa gaagcgcctc cccacgacgt tcttatggt ggtcatgttg gcgagctatt 420

tccttatctc catgtttgga ggagtcctgg tctttgtgtt tggcattact tttcctttgc 480

tgttgatggt tatccatgca tctgtgagac ttcggaacct caagaacaaa ctggagaata 540

aatggaagg aataggtttg aagaggacac ccatgggcat tctcctggat gccctagaac 600

agcaggaaga aggcacacac agactcactg actatatcag caaagtgaag gaataaacat 660

aacttacctg agctaggggt gcagcagaaa ttgagttgca gcttgcacct gtccagacct 720

atgttctgct tgcgtttttg aaacaggagg tgcacgtacc acccaattat ctatggcagc 780

atgcatgtat aggcogaact attatcagct ctgatgttcc agagagaaga cctcagaaac 840

cgaaagaaaa ccaccacct cctattgtgt ctgaagtttc acgtgtgttt atgaaatcta 900

atgggaaatg gatcacacga tttctttaag ggaattaaaa aaaataaaag aattacggct 960

tttacagcaa caatacgatt atcttatagg aaaaaaaaa atcattgtaa agtatcaaga 1020

caatacgagt aatgaaaag gctgttaaag tagatgacat catgtgttag cctgttccta 1080

aatccctaga attgtaagt gtgggatata aattagtttt tattattctc ttaaaaatca 1140

aagatgatct ctatcacttt gccacctgtt tgatgtgcag tggaaactgg ttaagccagt 1200

tgttcatact tcctttacaa atataaagat agctgtttag gatattttgt tacatttttg 1260

taaatttttg aatgctagt aatgtgtttt caccagcaag tatttgttgc aaactaatg 1320

tcattttcct taagatggtt acagctatgt aacctgtatt attctggacg gacttattaa 1380  
 aatacaaca gacaaaaaat aaaacaaaac ttgagttcta tttacottgc acattttttg 1440  
 ttgttacagt gaaaaaatg gtccaagaaa atgtttgcca tttttgcatt gtttcgtttt 1500  
 taactggaac atttagaaag aaggaaatga atgtgcattt tattaattcc ttaggggcac 1560  
 aaggaggaca ataatagctg atcttttgaa atttgaaaaa cgtctttaga tgaccaagca 1620  
 aaaagacttt aaaaaatggt aatgaaaatg gaatgcagct actgcagcta ataaaaaatt 1680  
 ttagatagca attgttaca ccatatgect ttatagctag acattagaat tatgatagca 1740  
 tgagtttata cattctatta tttttctccc ctttctcatg tttttataaa taggtaataa 1800  
 aaaatgtttt gcctgccaat tgaatgattt cgtagctgaa gtagaaacat ttaggtttct 1860  
 gtagcattaa attgtgaaga caactggagt ggtacttact gaagaaactc tctgtatgic 1920  
 ctagaataag aagcaatgat gtgctgettc tgatttttct tgcattttaa attctcagcc 1980  
 aacctacagc catgatcttt agcacagtga taccacatg acttcacaga catggtctag 2040  
 aatctgtacc cttaccaca tatgaagaat aaaattgatt aaaggtta 2088

<210> 14

<211> 1650

<212> DNA

<213> 人

<400> 14

gccttttttg cagtctcagg acgggcgctt tggagccggc cccaggcagc gtgtgtcggc 60  
 cgcctagtct ggagaactag tcctcgactc acggtgaggg aatggaccga cacgggtatt 120  
 gtaccgctga gggaaaggag cgggactcog gacctccagg agtgcaagga tgatgctgaa 180  
 aggaataaca aggcttatct ctaggatcca taagttggac cctgggcgctt ttttacacat 240  
 ggggaccag gctcgccaaa gcattgctgc tcacctagat aaccaggttc cagttgagag 300  
 tccgagagct atttcccga ccaatgagaa tgaccocggcc aagcatgggg atcagcacga 360  
 gggtcagcac tacaacatct cccccagga tttggagaet gtatttccc atggccttc 420

tcctcgcttt gtgatgcagg tgaagacatt cagtgaagct tgcctgatgg taaggaaacc 480  
 agccctagaa cttctgcatt acctgaaaa caccagtttt gcttatccag ctatacgata 540  
 tctttctgtat ggagagaagg gaacaggaaa aaccctaagt ctttgccatg ttattcattt 600  
 ctgtgcaaaa caggactggc tgatactaca tattccagat gctcatcttt gggtgaaaa 660  
 ttgtcgggat cttctgcagt ccagctacaa caaacagcgc tttgatcaac ctttagaggc 720  
 ttcaacctgg ctgaagaatt tcaaaactac aaatgagcgc ttctgaacc agataaaagt 780  
 tcaagagaag tatgtctgga ataagagaga aagcactgag aaaggagtc ctctgggaga 840  
 agtggttgaa cagggcataa cacgggtgag gaacgccaca gatgcagttg gaattgtgct 900  
 gaaagagcta aagaggcaaa gttctttggg tatgtttcac ctctagtgg ccgtggatgg 960  
 aatcaatgct ctttggggaa gaaccactct gaaaagagaa gataaaagcc cgattgcccc 1020  
 cgaggaatta gcacttgctt acaacttgag gaaaatgatg aaaaatgatt ggcattggagg 1080  
 cgccatttg tggctttga gccagactgg gtctctcttt aagccccgga aagcctatct 1140  
 gccccaggag ttgctgggaa aggaaggatt tgatgccttg gateccttta ttccatcct 1200  
 ggtttccaac tataacccaa aggaatttga aagttgtatt cagtattatt tggaaaacaa 1260  
 ttggcttcaa catgagaaag ctctacaga agaagggaaa aaagagctgc tgttcttaag 1320  
 taacgcgaac cctcctctgc tggagcggca ctgtgcctac ctctaagcca agatcacagc 1380  
 atgtgaggaa gacagtggac atctgcttta tgctggacce agtaagatga ggaagtcggg 1440  
 cagtaacacag gaagaggagc caggcccttg tacctatggg attggacagg actgcagttg 1500  
 gctctggacc tgcattaataa tgggtttcac tgtgaatgog tgacaataag atattccctt 1560  
 gttcctaaaa ctttatatca gtttattgga tgtggttttt cacatttaag ataattatgg 1620  
 ctcttttctt aaaaaataaa atatctttct 1650

- <210> 15  
 <211> 1109  
 <212> DNA  
 <213> 人



<400> 15  
 actggaagac caggcagccc agctgaagge agtaagctcg gctcacagtc gcaggagagt 60  
 tctgggggtac acgggcaaag gggcttgaga aggccccggag gccaagccga agagaagcaa 120  
 ctgtgccccg gagaagagaa gctcgcccat tccagactgg gaaccagctt tcagtgaaga 180  
 tggcagggcc agaactgttg ctgactcca acatccgct ctgggtggtc ctaccateg 240  
 ttatcatcac tttcttgta ggcattgatc gccactaagt gtccatctg ctgcagagcg 300  
 acaagaagct caccagga caagtatctg acagtcaagt cctaattcga agcagagtcc 360  
 tcagggaaaa tggaaaatac attcccaaac agtctttctt gacacgaaaa tattatttca 420  
 acaaccaga ggatggattt ttcaaaaaa ctaaacggaa ggtagtgcca cttctccta 480  
 tgactgatcc tactatgttg acagacatga tgaaggga tgtaacaaat gtctccta 540  
 tgattcttat tgggtgatgg atcaacatga cattctcagg cttgtcaca accaaggtcc 600  
 catttccact gacctccgt ttaagccta tgttacagca aggaatcgag ctactcacat 660  
 tagatgcate ctgggtgagt tetgcatcct ggtacttctt caatgtattt gggttcgga 720  
 gcatttactc tetgattctg ggccaagata atgccgtga ccaatcacga atgatgcagg 780  
 agcagatgac gggagcagcc atggccatgc ccgagacac aaacaaagct ttcaagacag 840  
 agtgggaagc tttggagctg acggatcacc agtgggcaact agatgatgtc gaagaagagc 900  
 tcatggccaa agacctccac ttogaaggca tgttcaaaaa ggaattacag acctctattt 960  
 tttgaagacc gacagggat tagctgtgtc aggaacttgg agttgcactt aaccttghaa 1020  
 ctttgtttgg agctggcacc tcttgaaata aaaaggagga tgcacagctt ggcaggcatg 1080  
 caaaaaaaaa aaaaaaaaa aaaaaaaaa 1109

<210> 16  
 <211> 2818  
 <212> DNA  
 <213> 人

<400> 16  
 ccctactccg cctctogga tcctttaaga ggccgggctt ggctgccagc tccgcccgc 60

gggcaaaagg ctgggacttt actccgggtg gcgcgagga cgagtctgtg ctccatcage 120  
 tgccgcaccc gccgcctccc gccccaaac cccatecccg cggttgagcc acgatgagcg 180  
 gcagagtccg cgatctgagc cccaggcaga aggaggcatt ggccaagttt cgggagaatg 240  
 tccaggatgt gctgccggcc ctgccgaatc cagatgacta tttctctctg cgttggtctc 300  
 gagccagaag cttcgacctg cagaagtcgg aggccatgct ccggaagcat gtggagtctc 360  
 gaaagcaaaa ggacattgac aacatcatta gctggcagcc tccagaggtg atccaacagt 420  
 atctgtcagg gggatatgtg gctatgacc tggatggctg cccagtctgg tacgacataa 480  
 ttggacctct ggatgccaaag ggtctgctgt tctcagcctc caaacaggac ctgctgagga 540  
 ccaagatgcg ggagtgtgag ctgcttctgc aagagtgtgc ccaccagacc acaaagttgg 600  
 ggaggaaggt ggagaccatc accataatth atgactgcca ggggcttggc ctcaagcatc 660  
 tctggaagcc tgctgtggag gcctatggag agtttctctg catgtttgag gaaaattatc 720  
 ccgaaacact gaagcgtctt tttgttgta aagccccaa actgtttcct gtggcctata 780  
 acctcatcaa acccttctg agtgaggaca ctcgtaagaa gatcatggtc ctgggagcaa 840  
 attggaagga ggttttactg aaacatatca gcctgacca ggtgcctgtg gagtatgggg 900  
 gcacatgac tgacctgat ggaaacccca agtgcaaate caagatcaac tacgggggtg 960  
 acatccccag gaagtattat gtgcgagacc aggtgaaaca gcagtatgaa cacagcgtgc 1020  
 agatttcccg tggctcctcc caccaagtgg agtatgagat cctcttccct ggctgtgtcc 1080  
 tcaggtggca gtttatgtca gatggagcgg atgttggtt tgggatttct ctgaagacca 1140  
 agatgggaga gaggcagcgg gcaggggaga tgacagaggt gctgcccaac cagaggtaca 1200  
 actcccact ggtccctgaa gatgggaccc tcacctgcag tgatcctggc atctatgtcc 1260  
 tgcggtttga caaacctac agcttcattc atgccaagaa ggtcaatttc actgtggagg 1320  
 tctgcttcc agacaaagcc tcagaagaga agatgaaaca gctgggggca ggcacccccg 1380  
 aataacacct tctctatag caggcctggc cccctcagtg tctcctgtc aatttctacc 1440  
 cctttagca gtcattttcg cacaccctg aagccccaaag aaactgggct ggaggacaga 1500

cctcaggagc tttcatttca gttaggcaga ggaagagoga ctgcagtggg tctccgtgtc 1560  
 tatcaaatac ctaaggagtc cccaggagct ggctggcoat cgtgatagga tctgtctgtc 1620  
 ctgtaaactg tgccaacttc acctgtccag ggacagcga gctgggggtg gcggggggca 1680  
 tgtaccacag ggtggcagca gggaaaaaaaa ttagaaaagg gtgaaagatt gggacttaac 1740  
 acttcaggga agtcagctgc cggggagaaa ottgctccta aatgaacaca taagtttaga 1800  
 tcgcaatgag gagtagcagg gtagctggtt gctagagtta cggtagggat cagaaactct 1860  
 tccaaacatt ttagcaactga ggctggggta gcttttggct tttcccaggt ctcaggaggt 1920  
 ggctgagtc agcacacatc ttcccactcg gtagacagge tggectctcc ctcactttga 1980  
 gactttggca actcctgggc cacacggcct gcctctttga ttactaatga ttgtcagtga 2040  
 ctcagagctt cctgggactt cgggtacceca cccgctgttc tccatgcaaa caaagcgcca 2100  
 gggaaatgac ccacagggat cgcagctgca gggagggcca gggaggttg ggtgggagt 2160  
 gaatgctaaa agcagatcgt ccagtgcctt tttcagtgt accggcctct caccaagcag 2220  
 tctccatgt gagcaacccc gagacaaaa tgctaagtgg gatcaagaga gcagcactcg 2280  
 gagaggggtg ttgccagtct gagtgtcccg cgggtcccgc caaccogett cctgactgac 2340  
 ctgagcaagg tcttactaag cagtccatc tctgtgggag gcatgcaacg cgtgcaggga 2400  
 gttcaggtgc cggteggcgt agccaggcct ggaggcccc caggcaggag gccgccccaa 2460  
 ggccgggccc gcgtctcgca gactaggggc tgggggcggc cacagacggc ctcgaaacca 2520  
 cagcccttac cccaatccca cgagccccgc caacgaacca cagggtgctgg gctttagaga 2580  
 acatgggaag ggggccccag acctggcggg aacgccttc cctcagagcc aggccccggc 2640  
 cccgtctggg aagctcatct tgcaagctg agggagctca gggcaaaggc caggctagcg 2700  
 cggaccggaa ggggcccagg ctgcaogggc ctctgccaga acgtcagga catcccggcc 2760  
 tgggtttaca acgtgttag gaaaattaac caatgaataa agcaacgttc agtgcgca 2818

<210> 17  
 <211> 1475  
 <212> DNA

<213> 人

<400> 17

gtcgcacgcg cgcgcctccg ctcccgtgag taacttggct ccgggggctc cgetcgcctg 60

cccgcacgcc gcccgccacc caggaccgcg ccgcccgcct ccgcccctag caaaccttc 120

cgacggccct cgetcgcgaa gccgggacgc ctctccccc tccgccccg ccgcggaaag 180

ttaagtttga agagggggga agaggggaac atggacatga agaggaggat ccacctggag 240

ctgaggaacc ggaccccggc agctgttcga gaacttgtct tggacaattg caaatcaaat 300

gatggaaaaa ttgagggtt aacagctgaa tttgtgaact tagagttcct cagttaata 360

aatgtaggct tgatctcagt ttcaaatctc occaagctgc ctaaattgaa aaagcttgaa 420

ctcagtgaaa atagaatctt tggaggtctg gacatgttag ctgaaaaact tccaaatctc 480

acacatctaa acttaagtgg aaataaactg aaagatatca gcaccttga acctttgaaa 540

aagttagaat gtctgaaaag cctggacctc tttaactgtg aggttacc aa cctgaatgac 600

taccgagaga gtgtcttcaa gctcctgccc cagcttaact acttgatgg ctatgaccga 660

gaggaccagg aagcacctga ctcagatgcc gaggtggatg gtgtggatga agaggaggag 720

gacgaagaag gagaagatga ggaagacgag gacgatgagg atggtgaaga agaggagttt 780

gatgaagaag atgatgaaga tgaagatgta gaaggggatg aggacgacga tgaagtcagt 840

gaggaggaag aagaatttgg acttgatgaa gaagatgaag atgaggatga ggatgaagag 900

gaggaagaag gtgggaaagg tgaaaagagg aagagagaaa cagatgatga aggagaagat 960

gattaagacc ccagatgacc tgcagaaaca gaactgttca gtattggtt gactgctcat 1020

ggattttgta gctgtttaa aaaaaaaaa aggtagctgt gatacaaac ccaggacacc 1080

caccaccca aagagccaaa gaatagtcc tgtgacatc cgccttcctt ccatgtagtc 1140

cctcttgta atctaccacc aagcttgtgg acttcacccc aacaaaattg taagcgttgt 1200

taggttttgg tgtaagatc ttgctgtagc gtggatagct gtgattggtg agtcaaccgt 1260

ctgtggctac cagttacact gagattgtaa cagcattttt actttctgta caacaaaaa 1320

gctttgtaaa taaaatctta acattttggg tctgtttttt catgctttgc tttttaata 1380

ttattattat tttttttaca ttaggcacatt ttatgtgaca actgccaaaa aagtattttt 1440

aagaatttaa gcgaaataaa cagttactct ttggc 1475

<210> 18

<211> 841

<212> DNA

<213> 人

<400> 18

gcaaccactg cagctgggcc aagtcgctta gctcttoggt ggttgtcaca cgtccggagg 60

cctagccgctc gcgtacctag gatgccgctt ggaagccgaa gccgcacctc ccgcatggcc 120

cctccggcca gccgggcccc tcagatgaga gctgeacca gcccagcacc agtcgctcag 180

ccaccagcag cggcaccccc atctgcagtt ggctcttctg ctgctgcgcc ccggcagcca 240

ggtctgatgg ccagatggc aaccactgca gctggcgtgg ctgtgggctc tgctgtgggg 300

cacacattgg gtcacgcat tactgggggc ttcagtggag gaagtaatgc tgagcctgoc 360

aggcctgaca tcacttacca ggagcctcag ggaaccacgc cggcacagca gcagcagcct 420

tgccctatg agatcaaaca gtttctggag tgtgccaga accagggtga catcaagctc 480

tgtgagggtt tcaatgaggt gctgaaacag tgccgacttg caaacggatt ggccaatga 540

agaagttaa cctggagaga tggaaaatca gctctcataa ctaagttaat ttagtataaa 600

aatagaattg atagtgagg tataaagtgt aaccatcagt taaacctctc ctgtcattec 660

tagcttctt gcttcagaat tgaaatggaa gtgggggtgt ccctactctg tagaatctgg 720

gactgggcaa atgtttgtgt ggcctcctta aactagctgt tatgttatga ttttattctt 780

tgtgagttaa ttagaataaa gtcattttct tacaaaaaaaa aaaaaaaaaa aaaaaaaaaa 840

a 841

<210> 19

<211> 1486

<212> DNA

<213> 人

<400> 19  
 gggctcgtca gatataataa ttttacactt cagttttgat tggtagagaaa gtaccattc 60  
 tcttcaaata atcaaagata attattattt tgttttgttt ttggaatcaa cagggaggcg 120  
 caaagtataa agttgctgct aacatatata catatacatc catattttat aagggtgtct 180  
 atgtatata agacagtgtg tccacacaaa aaatagatac agttatcagt cagtcagttc 240  
 ttccatgatt tagttttttt aaacgtagaa aagctattgt aaacgtctct ttccatttat 300  
 tottaatttt ttgacatatt ggtatttctt taaagggaaa tgaggaatgc acatcagtga 360  
 ttgattgtca aacctcacc cctgatttcc tacctaactc acccccacct aaccaatcaa 420  
 tcacatccac aaattgtttt gttgtttgt tagtcaggtc tccaacagag tccaatattt 480  
 ctaacactct agtgcaataa aaattattat taaatagcta agaggtgtgc atgtgggaaa 540  
 ggtagtgca tateccttta ggaggggaga atgttgtaat atatcagcta tcgagttggt 600  
 taaaaaagt gtattcaatc gtatattgtc tatagtatgt gctatgaaat ttgcatttat 660  
 gatatgtaac aggggcaaag ccaaattcat gttactctgt tcagtcagaa acattttgtg 720  
 gcatacagca ttctgggaa gtgctgtact ttgtttcgtt ttggttttag tttgcattt 780  
 agagtgcctt ataattgatg cctattttaa tagcatttct ttttagcttt tggttcgtat 840  
 ttccattcac tgttcgtatc tgttacttcc tattaagca ttatctgttt accacatgta 900  
 caaaaactct ttgaataata tgcattccta gttttcagcc aagacgggga tgttagtgat 960  
 tgtaccagcc caaagcactt ggataatcag ggcccttctt cctttataa tcaatcatca 1020  
 acatcagaaa aagctacttg tttatattat attccttcc aaatccgctc tggaacatgc 1080  
 agtaactgca ccaaacttat tttagtaaca aatatcattg gcaactttgg aatatattg 1140  
 atattccatt aggatttttc taaaagggga aataaactat atatataat gtatcttacc 1200  
 cccaattctt ccaacagaat ttctatagga agccatggat gatggcataa gtttgcaca 1260  
 tattacatga ttttaataa tctcaaaat acccaaggaa ctcttaaga gttttggtat 1320  
 gagtatacta ctttggttta attttagctt catggatggt ctgcatggaa ggatttttgt 1380  
 tttccacatt ttccattgc tagcagagtg aaatccaaga gaccaaacat ttgcaagcat 1440

tgtatttgag cacttttgta aaaaacaaaa aaaaaaaaa aaaaaa	1486
<210> 20	
<211> 16	
<212> DNA	
<213> 噬菌体 M13mp18	
<400> 20	
gtaaaacgac ggccag	16
<210> 21	
<211> 17	
<212> DNA	
<213> 噬菌体 M13mp18	
<400> 21	
caggaacag ctatgac	17
<210> 22	
<211> 30	
<212> DNA	
<213> 人	
<400> 22	
caggtgaatt tcaaaggagg atttactcac	30
<210> 23	
<211> 30	
<212> DNA	
<213> 人	
<400> 23	
gtgagtaaat cctcctttga aattcacctg	30
<210> 24	
<211> 23	
<212> DNA	
<213> 人	
<400> 24	
gcaagccagg aagagtcgtc acg	23

<210> 25	
<211> 25	
<212> DNA	
<213> 人	
<400> 25	
tgccaggaaa cttcttgctt gatgc	25
<210> 26	
<211> 26	
<212> DNA	
<213> 人	
<400> 26	
agtaaccagt tgagatgaag cacgtc	26
<210> 27	
<211> 28	
<212> DNA	
<213> 人	
<400> 27	
cagaagagca acaagaatgg tatcctgc	28
<210> 28	
<211> 25	
<212> DNA	
<213> 人	
<400> 28	
aacttgagtt ctatttacct tgcaac	25
<210> 29	
<211> 21	
<212> DNA	
<213> 人	
<400> 29	
ttgcttgggt catctaaaga c	21
<210> 30	
<211> 20	



---

<212> DNA	
<213> 人	
<400> 30	
actcacgtgc aaggatgatg	20
<210> 31	
<211> 20	
<212> DNA	
<213> 人	
<400> 31	
agctctogga ctctcaactg	20
<210> 32	
<211> 26	
<212> DNA	
<213> 人	
<400> 32	
cttctcctat gactgatcct actatg	26
<210> 33	
<211> 21	
<212> DNA	
<213> 人	
<400> 33	
caggatgcag aactcacct g	21
<210> 34	
<211> 21	
<212> DNA	
<213> 人	
<400> 34	
gcagatttcc cgtggctcct c	21
<210> 35	
<211> 22	
<212> DNA	
<213> 人	

---

<400> 35	
gttgggcagc acctctgtca tc	22
<210> 36	
<211> 22	
<212> DNA	
<213> 人	
<400> 36	
ctgtgacatt ccgccttctc tc	22
<210> 37	
<211> 23	
<212> DNA	
<213> 人	
<400> 37	
ccacgtact gcaagaatct tac	23
<210> 38	
<211> 23	
<212> DNA	
<213> 人	
<400> 38	
agaagtcaa cctggagaga tgg	23
<210> 39	
<211> 24	
<212> DNA	
<213> 人	
<400> 39	
caaggaagct aggaatgaca ggag	24
<210> 40	
<211> 24	
<212> DNA	
<213> 人	
<400> 40	
gcaaagccaa attcatgtta ctct	24

<210> 41	
<211> 27	
<212> DNA	
<213> 人	
<400> 41	
cagatacgaag cagtgaatgg aaatagc	27
<210> 42	
<211> 24	
<212> DNA	
<213> 人	
<400> 42	
gccacaggtt gaacaattaa ttg	24
<210> 43	
<211> 22	
<212> DNA	
<213> 人	
<400> 43	
aggaagagtc gtcacgagaa cc	22
<210> 44	
<211> 25	
<212> DNA	
<213> 人	
<400> 44	
ataatgctgt gcttagttta ttgcc	25
<210> 45	
<211> 21	
<212> DNA	
<213> 人	
<400> 45	
gatcgtggac atttcaacct c	21
<210> 46	
<211> 20	

<212> DNA

<213> 人

<400> 46

tcttgcttga tgccttggtc

20

<210> 47

<211> 1254

<212> PRT

<213> 小鼠

<400> 47

Met Pro Gly Gly Ser Val Asn Ile Thr Cys Val Ala Val Gly Ser Pro  
1                   5                   10                   15

Met Pro Tyr Val Lys Trp Met Leu Gly Ala Glu Asp Leu Thr Pro Glu  
                  20                   25                   30

Asp Asp Met Pro Ile Gly Arg Asn Val Leu Glu Leu Asn Asp Val Arg  
                  35                   40                   45

Gln Ser Ala Asn Tyr Thr Cys Val Ala Met Ser Thr Leu Gly Val Ile  
50                   55                   60

Glu Ala Ile Ala Gln Ile Thr Val Lys Ala Leu Pro Lys Pro Pro Gly  
65                   70                   75                   80

Thr Pro Val Val Thr Glu Ser Thr Ala Thr Ser Ile Thr Leu Thr Trp  
                  85                   90                   95

Asp Ser Gly Asn Pro Glu Pro Val Ser Tyr Tyr Ile Ile Gln His Lys  
                  100                   105                   110

Pro Lys Asn Ser Glu Glu Pro Tyr Lys Glu Ile Asp Gly Ile Ala Thr  
                  115                   120                   125

Thr Arg Tyr Ser Val Ala Gly Leu Ser Pro Tyr Ser Asp Tyr Glu Phe  
130                   135                   140

Arg Val Val Ala Val Asn Asn Ile Gly Arg Gly Pro Ala Ser Glu Pro  
145 150 155 160

Val Leu Thr Gln Thr Ser Glu Gln Ala Pro Ser Ser Ala Pro Arg Asp  
165 170 175

Val Gln Ala Arg Met Leu Ser Ser Thr Thr Ile Leu Val Gln Trp Lys  
180 185 190

Glu Pro Glu Glu Pro Asn Gly Gln Ile Gln Gly Tyr Arg Val Tyr Tyr  
195 200 205

Thr Met Asp Pro Thr Gln His Val Asn Asn Trp Met Lys His Asn Val  
210 215 220

Ala Asp Ser Gln Ile Thr Thr Ile Gly Asn Leu Val Pro Gln Lys Thr  
225 230 235 240

Tyr Ser Val Lys Val Leu Ala Phe Thr Ser Ile Gly Asp Gly Pro Leu  
245 250 255

Ser Ser Asp Ile Gln Val Ile Thr Gln Thr Gly Val Pro Gly Gln Pro  
260 265 270

Leu Asn Phe Lys Ala Glu Pro Glu Ser Glu Thr Ser Ile Leu Leu Ser  
275 280 285

Trp Thr Pro Pro Arg Ser Asp Thr Ile Ala Ser Tyr Glu Leu Val Tyr  
290 295 300

Arg Asp Gly Asp Gln Gly Glu Glu Gln Arg Ile Thr Ile Glu Pro Gly  
305 310 315 320

Thr Ser Tyr Arg Leu Gln Gly Leu Lys Pro Asn Ser Leu Tyr Tyr Phe  
325 330 335

Arg Leu Ser Ala Arg Ser Pro Gln Gly Leu Gly Ala Ser Thr Ala Glu  
 340 345 350

Ile Ser Ala Arg Thr Met Gln Ser Met Phe Ala Lys Asn Phe His Val  
 355 360 365

Lys Ala Val Met Lys Thr Ser Val Leu Leu Ser Trp Glu Ile Pro Glu  
 370 375 380

Asn Tyr Asn Ser Ala Met Pro Phe Lys Ile Leu Tyr Asp Asp Gly Lys  
 385 390 395 400

Met Val Glu Glu Val Asp Gly Arg Ala Thr Gln Lys Leu Ile Val Asn  
 405 410 415

Leu Lys Pro Glu Lys Ser Tyr Ser Phe Val Leu Thr Asn Arg Gly Asn  
 420 425 430

Ser Ala Gly Gly Leu Gln His Arg Val Thr Ala Lys Thr Ala Pro Asp  
 435 440 445

Val Leu Arg Thr Lys Pro Ala Phe Ile Gly Lys Thr Asn Leu Asp Gly  
 450 455 460

Met Ile Thr Val Gln Leu Pro Asp Val Pro Ala Asn Glu Asn Ile Lys  
 465 470 475 480

Gly Tyr Tyr Ile Ile Ile Val Pro Leu Lys Lys Ser Arg Gly Lys Phe  
 485 490 495

Ile Lys Pro Trp Glu Ser Pro Asp Glu Met Glu Leu Asp Glu Leu Leu  
 500 505 510

Lys Glu Ile Ser Arg Lys Arg Arg Ser Ile Arg Tyr Gly Arg Glu Val  
 515 520 525

Glu Leu Lys Pro Tyr Ile Ala Ala His Phe Asp Val Leu Pro Thr Glu  
530 535 540

Phe Thr Leu Gly Asp Asp Lys His Tyr Gly Gly Phe Thr Asn Lys Gln  
545 550 555 560

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

His Ala Glu Ser Lys Met Tyr Ala Thr Ser Pro Tyr Ser Asp Pro Val  
580 585 590

Val Ser Met Asp Leu Asp Pro Gln Pro Ile Thr Asp Glu Glu Glu Gly  
595 600 605

Leu Ile Trp Val Val Gly Pro Val Leu Ala Val Val Phe Ile Ile Cys  
610 615 620

Ile Val Ile Ala Ile Leu Leu Tyr Lys Arg Lys Arg Ala Glu Ser Glu  
625 630 635 640

Ser Arg Lys Ser Ser Leu Pro Asn Ser Lys Glu Val Pro Ser His His  
645 650 655

Pro Thr Asp Pro Val Glu Leu Arg Arg Leu Asn Phe Gln Thr Pro Gly  
660 665 670

Met Ala Ser His Pro Pro Ile Pro Ile Leu Glu Leu Ala Asp His Ile  
675 680 685

Glu Arg Leu Lys Ala Asn Asp Asn Leu Lys Phe Ser Gln Glu Tyr Glu  
690 695 700

Ser Ile Asp Pro Gly Gln Gln Phe Thr Trp Glu His Ser Asn Leu Glu  
705 710 715 720

Val Asn Lys Pro Lys Asn Arg Tyr Ala Asn Val Ile Ala Tyr Asp His  
725 730 735

Ser Arg Val Leu Leu Ser Ala Ile Glu Gly Ile Pro Gly Ser Asp Tyr  
740 745 750

Val Asn Ala Asn Tyr Ile Asp Gly Tyr Arg Lys Gln Asn Ala Tyr Ile  
755 760 765

Ala Thr Gln Gly Ser Leu Pro Glu Thr Phe Gly Asp Phe Trp Arg Met  
770 775 780

Ile Trp Glu Gln Arg Ser Ala Thr Val Val Met Met Thr Lys Leu Glu  
785 790 795 800

Glu Arg Ser Arg Val Lys Cys Asp Gln Tyr Trp Pro Ser Arg Gly Thr  
805 810 815

Glu Thr His Gly Leu Val Gln Val Thr Leu Leu Asp Thr Val Glu Leu  
820 825 830

Ala Thr Tyr Cys Val Arg Thr Phe Ala Leu Tyr Lys Asn Gly Ser Ser  
835 840 845

Glu Lys Arg Glu Val Arg Gln Phe Gln Phe Thr Ala Trp Pro Asp His  
850 855 860

Gly Val Pro Glu His Pro Thr Pro Phe Leu Ala Phe Leu Arg Arg Val  
865 870 875 880

Lys Thr Cys Asn Pro Pro Asp Ala Gly Pro Met Val Val His Cys Ser  
885 890 895

Ala Gly Val Gly Arg Thr Gly Cys Phe Ile Val Ile Asp Ala Met Leu  
900 905 910



Glu Arg Ile Lys His Glu Lys Thr Val Asp Ile Tyr Gly His Val Thr  
 915 920 925

Leu Met Arg Ala Gln Arg Asn Tyr Met Val Gln Thr Glu Asp Gln Tyr  
 930 935 940

Ile Phe Ile His Asp Ala Leu Leu Glu Ala Val Thr Cys Gly Asn Thr  
 945 950 955 960

Glu Val Pro Ala Arg Asn Leu Tyr Ala Tyr Ile Gln Lys Leu Thr Gln  
 965 970 975

Ile Glu Thr Gly Glu Asn Val Thr Gly Met Glu Leu Glu Phe Lys Arg  
 980 985 990

Leu Ala Ser Ser Lys Ala His Thr Ser Arg Phe Ile Ser Ala Asn Leu  
 995 1000 1005

Pro Cys Asn Lys Phe Lys Asn Arg Leu Val Asn Ile Met Pro Tyr  
 1010 1015 1020

Glu Ser Thr Arg Val Cys Leu Gln Pro Ile Arg Gly Val Glu Gly  
 1025 1030 1035

Ser Asp Tyr Ile Asn Ala Ser Phe Leu Asp Gly Tyr Arg Gln Gln  
 1040 1045 1050

Lys Ala Tyr Ile Ala Thr Gln Gly Pro Leu Ala Glu Thr Thr Glu  
 1055 1060 1065

Asp Phe Trp Arg Met Leu Trp Glu His Asn Ser Thr Ile Val Val  
 1070 1075 1080

Met Leu Thr Lys Leu Arg Glu Met Gly Arg Glu Lys Cys His Gln  
 1085 1090 1095

Tyr Trp Pro Ala Glu Arg Ser Ala Arg Tyr Gln Tyr Phe Val Val  
 1100 1105 1110

Asp Pro Met Ala Glu Tyr Asn Met Pro Gln Tyr Ile Leu Arg Glu  
 1115 1120 1125

Phe Lys Val Thr Asp Ala Arg Asp Gly Gln Ser Arg Thr Val Arg  
 1130 1135 1140

Gln Phe Gln Phe Thr Asp Trp Pro Glu Gln Gly Val Pro Lys Ser  
 1145 1150 1155

Gly Glu Gly Phe Ile Asp Phe Ile Gly Gln Val His Lys Thr Lys  
 1160 1165 1170

Glu Gln Phe Gly Gln Asp Gly Pro Ile Ser Val His Cys Ser Ala  
 1175 1180 1185

Gly Val Gly Arg Thr Gly Val Phe Ile Thr Leu Ser Ile Val Leu  
 1190 1195 1200

Glu Arg Met Arg Tyr Glu Gly Val Val Asp Ile Phe Gln Thr Val  
 1205 1210 1215

Lys Met Leu Arg Thr Gln Arg Pro Ala Met Val Gln Thr Glu Asp  
 1220 1225 1230

Gln Tyr Gln Phe Cys Tyr Arg Ala Ala Leu Glu Tyr Leu Gly Ser  
 1235 1240 1245

Phe Asp His Tyr Ala Thr  
 1250

<210> 48

<211> 21

<212> DNA	
<213> 人	
<400> 48	
aatctgcaag ccaggaagag t	21
<210> 49	
<211> 27	
<212> DNA	
<213> 人	
<400> 49	
tctagtttca gttttgatga tattttg	27
<210> 50	
<211> 19	
<212> RNA	
<213> 人	
<400> 50	
ucugcaagcc aggaagagu	19
<210> 51	
<211> 19	
<212> RNA	
<213> 人	
<400> 51	
acucuuccug gcuugcaga	19
<210> 52	
<211> 19	
<212> RNA	
<213> 人	
<400> 52	
ccuccagaac ugugaacca	19
<210> 53	
<211> 19	
<212> RNA	
<213> 人	

<400> 53	
uggaucacag uucuggagg	19
<210> 54	
<211> 19	
<212> RNA	
<213> 人	
<400> 54	
cuacaaauga gcgcuuccu	19
<210> 55	
<211> 19	
<212> RNA	
<213> 人	
<400> 55	
aggaagcgc cauuuguag	19
<210> 56	
<211> 20	
<212> DNA	
<213> 人	
<400> 56	
ccacatcgt cagacacat	20
<210> 57	
<211> 17	
<212> DNA	
<213> 人	
<400> 57	
accaggcgcc caatacg	17
<210> 58	
<211> 28	
<212> DNA	
<213> 人	
<400> 58	
caaatcogtt gactccgacc ttcacctt	28

<210> 59	
<211> 20	
<212> DNA	
<213> 人	
<400> 59	
aaggccaacc gcgagaagat	20
<210> 60	
<211> 20	
<212> DNA	
<213> 人	
<400> 60	
gtcaccggag tccatcaaga	20
<210> 61	
<211> 32	
<212> DNA	
<213> 人	
<400> 61	
ccatgtacgt tgctatccag gctgtgctat cc	32
<210> 62	
<211> 25	
<212> DNA	
<213> 人	
<400> 62	
caactgggac gacatggaga aaatc	25
<210> 63	
<211> 22	
<212> DNA	
<213> 人	
<400> 63	
catggctggg gtgttgaagg tc	22
<210> 64	
<211> 23	

<212> DNA	
<213> 人	
<400> 64	
actctcacct cccatgttgc tea	23
<210> 65	
<211> 23	
<212> DNA	
<213> 人	
<400> 65	
gctatccgtg cactcctggt ctg	23
<210> 66	
<211> 21	
<212> DNA	
<213> 人	
<400> 66	
atctgcaagc caggaagagt c	21
<210> 67	
<211> 22	
<212> DNA	
<213> 人	
<400> 67	
cttgettgat gctttggtct gt	22
<210> 68	
<211> 30	
<212> DNA	
<213> 人	
<400> 68	
ccagaccatg caggaactct gatcgtggac	30
<210> 69	
<211> 21	
<212> DNA	
<213> 人	

---

<400> 69 atgccctgga tccctttatt c	21
<210> 70 <211> 23 <212> DNA <213> 人	
<400> 70 tcaccccgac ttctcatct tac	23
<210> 71 <211> 24 <212> DNA <213> 人	
<400> 71 aactccctta ttacactatc catt	24
<210> 72 <211> 24 <212> DNA <213> 人	
<400> 72 gtgttatgag gaaaagatta ggga	24
<210> 73 <211> 30 <212> DNA <213> 人	
<400> 73 tgcagccagg agaagcaaga gaacagaaat	30

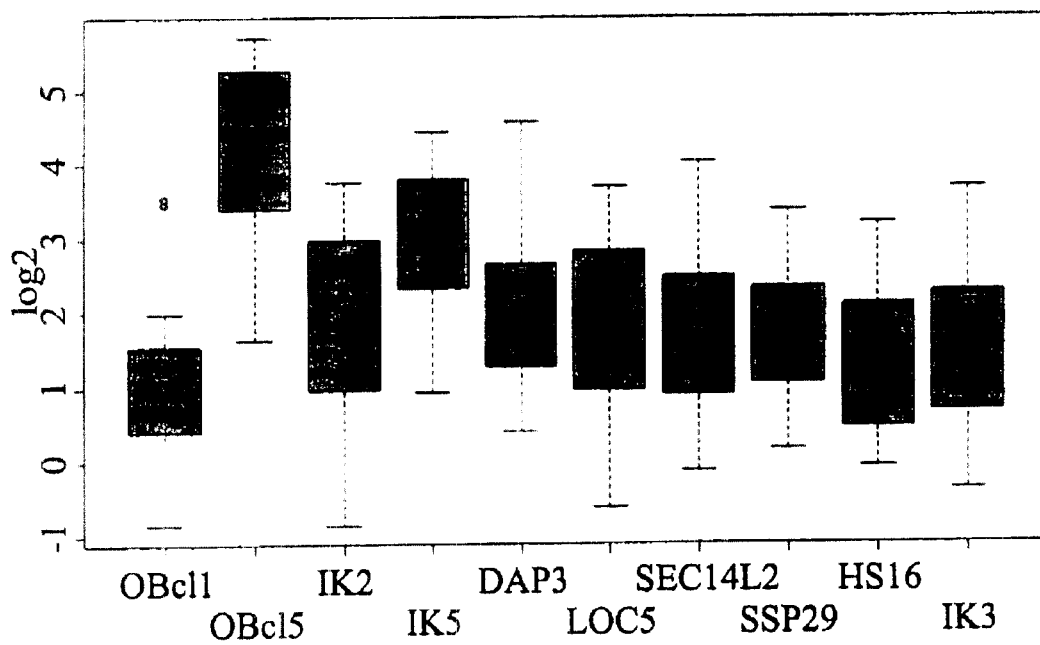


图 1



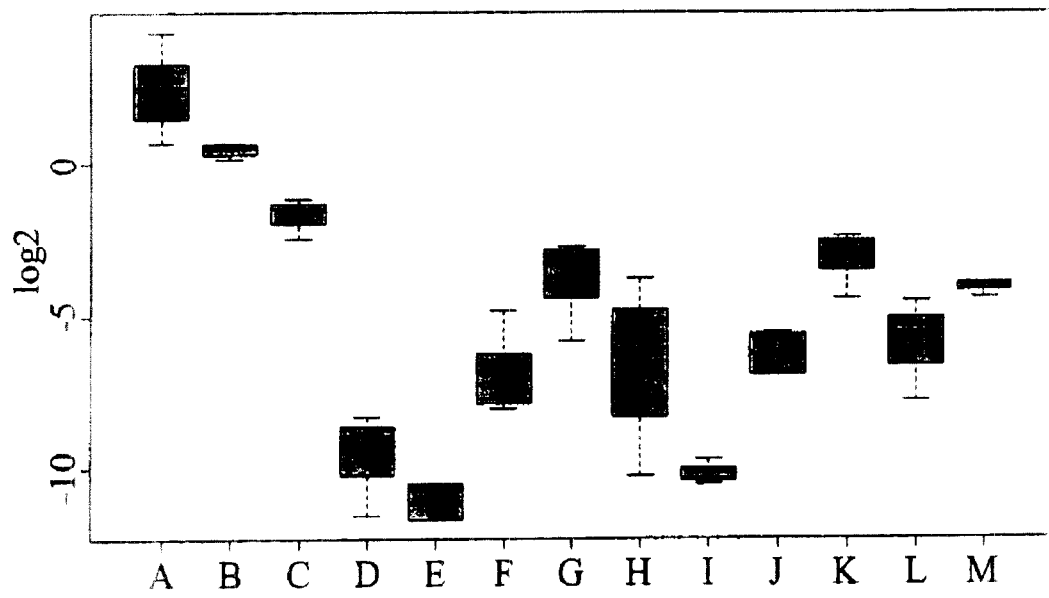


图 2

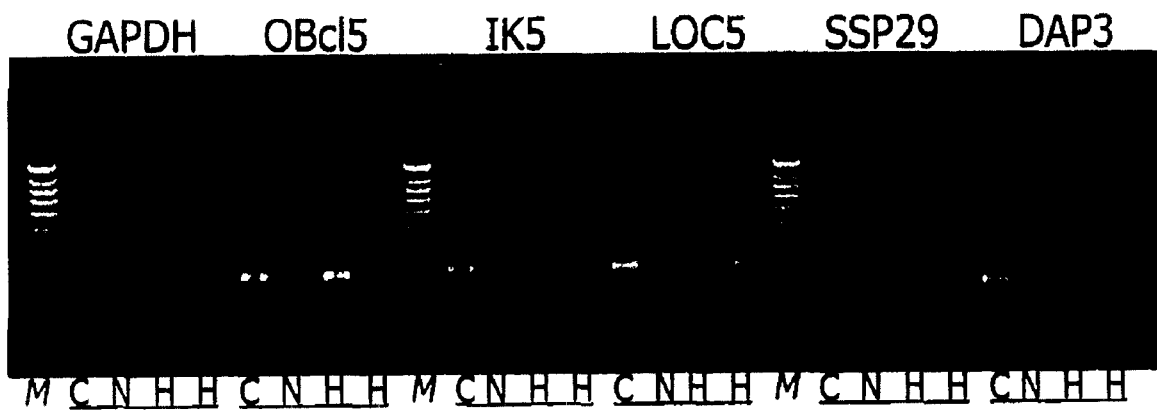


图 3

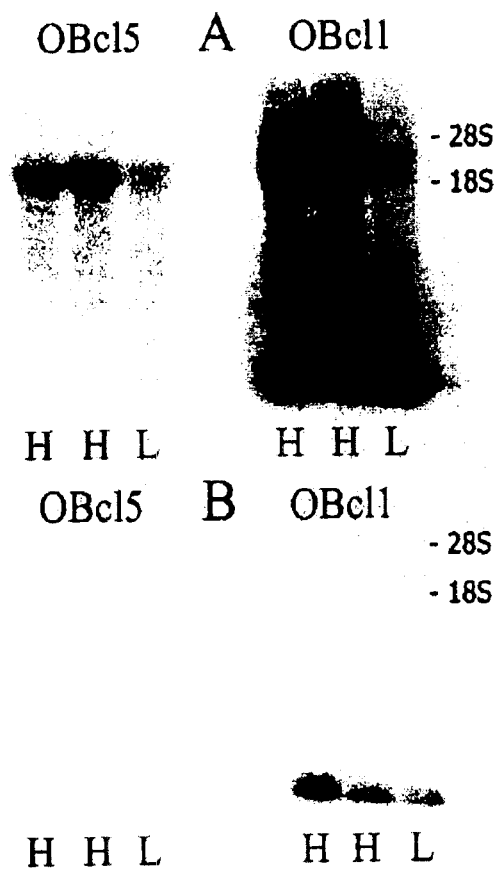


图 4

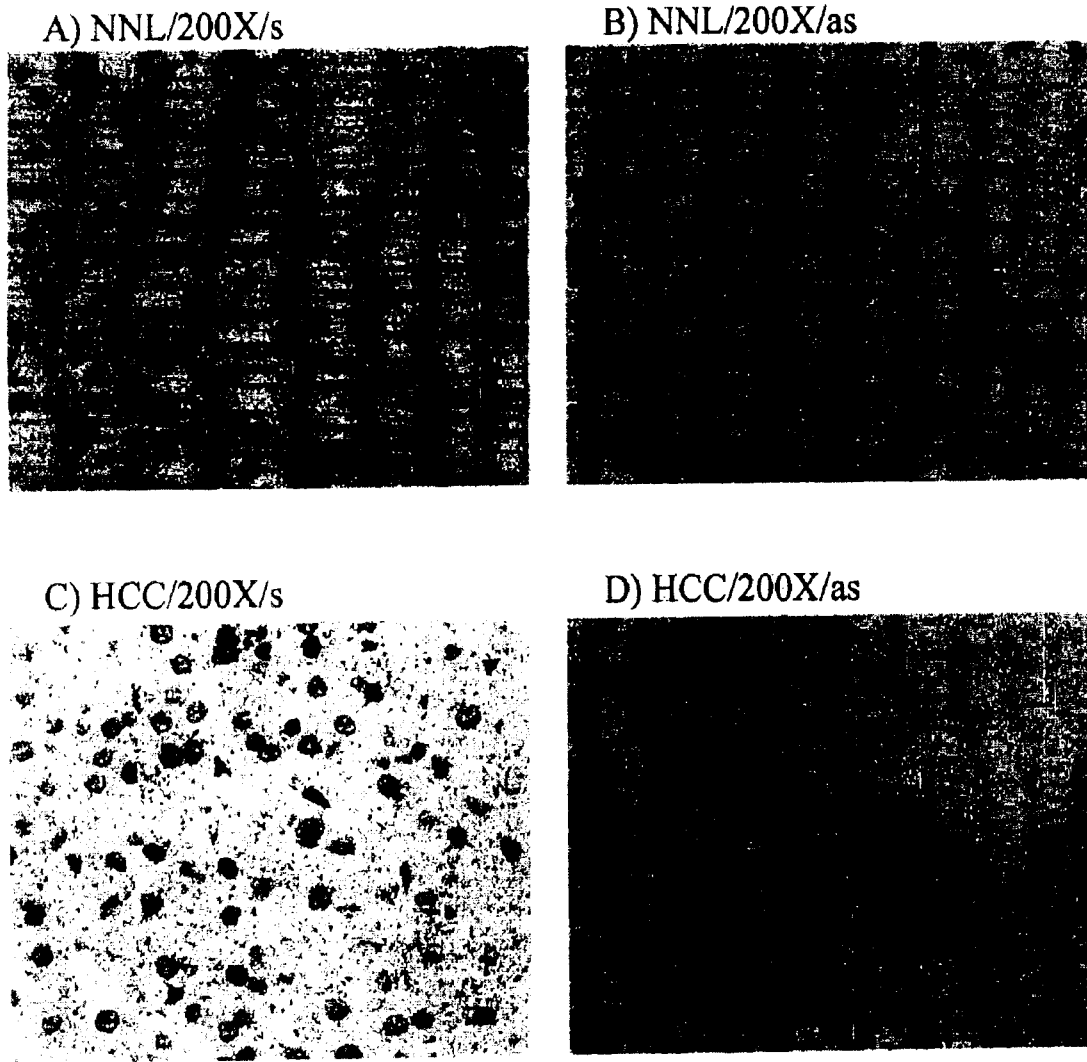


图 5

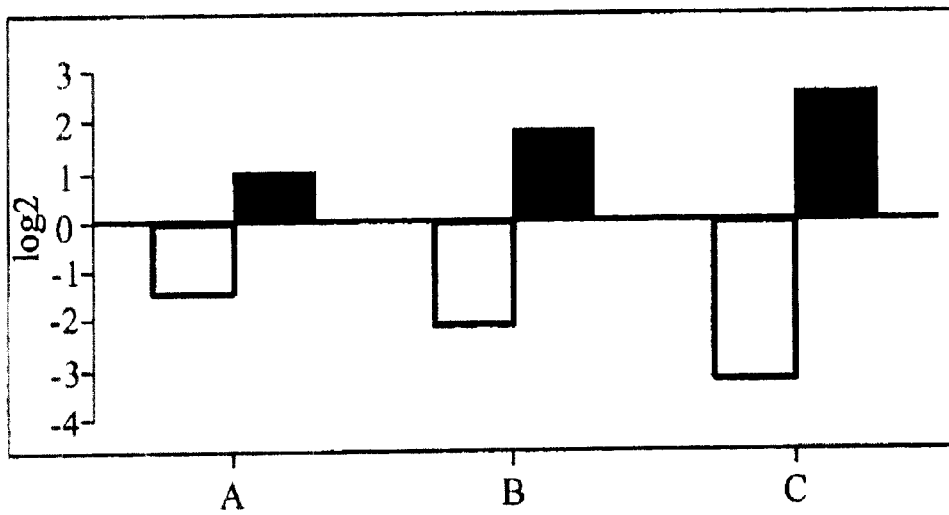


图 6

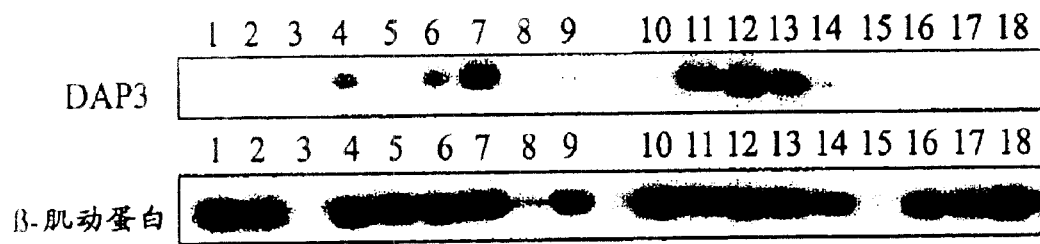


图 7

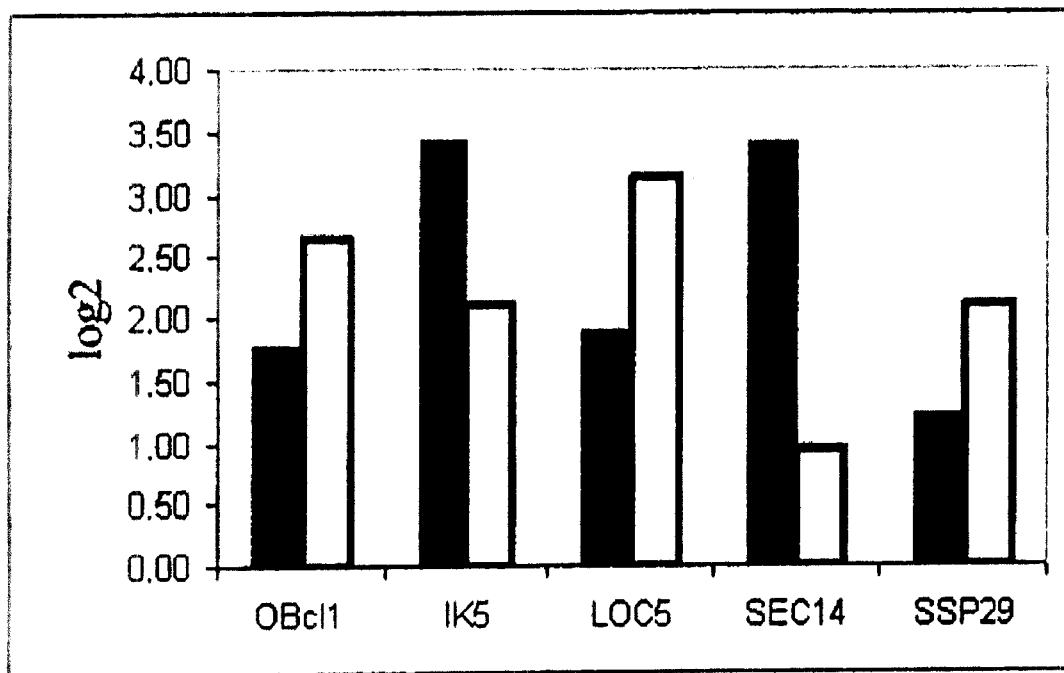


图 8