

[19] 中华人民共和国国家知识产权局

[51] Int. Cl<sup>7</sup>

C12N 15/52

C07K 14/705 C12Q 1/68

C07K 16/18 A61K 38/00

# [12] 发明专利申请公开说明书

[21] 申请号 00809163.3

[43] 公开日 2002 年 11 月 13 日

[11] 公开号 CN 1379817A

[22] 申请日 2000.6.16 [21] 申请号 00809163.3

[30] 优先权

[32]1999.6.18 [33]US [31]60/140,264

[32]1999.9.14 [33]US [31]60/153,872

[32]1999.11.19 [33]US [31]60/166,573

[86] 国际申请 PCT/US00/16591 2000.6.16

[87] 国际公布 W000/78971 英 2000.12.28

[85] 进入国家阶段日期 2001.12.18

[71] 申请人 CV 治疗公司

地址 美国加利福尼亚州

共同申请人 华盛顿大学

[72] 发明人 理查德·M·罗恩 大卫·维德

约翰·F·奥伦

迈克尔·加维

[74] 专利代理机构 隆天国际专利商标代理有限公司

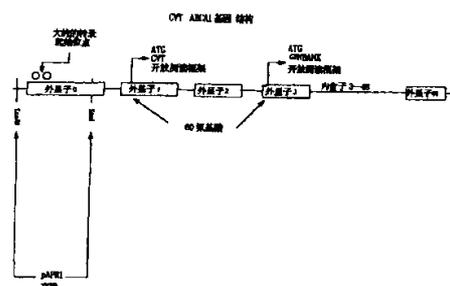
代理人 徐金国 黄韧敏

权利要求书 6 页 说明书 167 页 附图 10 页

[54] 发明名称 使用 ATP 结合框运输蛋白 ABC1 增加胆固醇排放并提高 HDL 的组合物及方法

[57] 摘要

本发明涉及到新型 ABC1 多肽和编码它的核酸分子。本发明也涉及到包含 ABC1 多核苷酸的重组载体, 宿主细胞, 和组合物, 以及生产 ABC1 多肽的方法。本发明也涉及到特异性结合 ABC1 多肽的抗体。此外, 本发明涉及到增加胆固醇排放的方法以及增加 ABC1 表达和活性的方法。本发明进一步涉及到鉴别调节 ABC1 表达的化合物的方法, 和在哺乳动物受体中检测 ABC1 多肽和多核苷酸的相对水平的方法。本发明也提供适于筛选化合物以测定该化合物 ABC1 表达调控活性的试剂盒和组合物, 以及适于测定化合物是否调节 ABC1 依赖性胆固醇排放的试剂盒和组合物。



## 权利要求书

---

1. 一种分离的多肽，其中包含 SEQ ID NO: 2。
2. 一种分离的多肽，其中包含的氨基酸序列与 SEQ ID NO: 2 所示氨基酸序列至少 98% 同源。
3. 一种分离的多核苷酸，其选自：编码权利要求 1 所述多肽的多核苷酸、包含 SEQ ID NO: 1 的多核苷酸、包含 SEQ ID NO: 1 中第 291-7074 位核苷酸的多核苷酸、以及编码权利要求 2 所述多肽的多核苷酸。
4. 一种分离的多核苷酸，其中包含的核苷酸序列互补于权利要求 3 所述的多核苷酸。
5. 一种分离的多核苷酸，其中包含的核苷酸序列与包含 SEQ ID NO: 1 的多核苷酸至少 90% 同源。
6. 一种分离的多核苷酸，其中包含的核苷酸序列与包含 SEQ ID NO: 1 的多核苷酸至少 95% 同源。
7. 一种分离的多核苷酸，其中包含的核苷酸序列互补于权利要求 5 所述的多核苷酸。
8. 一种组合物，其中包括权利要求 3 所述多核苷酸以及合适的载体。
9. 一种组合物，其中包括权利要求 4 所述多核苷酸以及合适的载体。
10. 一种组合物，其中包括权利要求 5 所述多核苷酸以及合适的载体。
11. 一种重组载体，其中包含权利要求 3 所述的多核苷酸。
12. 一种重组载体，其中包含权利要求 4 所述的多核苷酸。
13. 一种重组载体，其中包含权利要求 5 所述的多核苷酸。
14. 权利要求 11 中的重组载体，进一步还包括一异源启动子多核苷酸。
15. 权利要求 12 中的重组载体，进一步还包括一异源启动子多核苷酸。
16. 权利要求 13 中的重组载体，进一步还包括一异源启动子多核苷酸。

17. 权利要求 14 中的重组载体,其中所述的异源启动子是巨细胞病毒启动子。

18. 权利要求 17 中的重组载体,其中所述的载体是 pCEPhABC1。

19. 一种组合物,其中包括权利要求 11 所述的重组载体。

5 20. 一种组合物,其中包括权利要求 13 所述的重组载体。

21. 一种组合物,其中包括权利要求 14 所述的重组载体。

22. 一种宿主细胞,其中包含权利要求 11 所述的重组载体。

23. 一种宿主细胞,其中包含权利要求 13 所述的重组载体。

24. 一种宿主细胞,其中包含权利要求 14 所述的重组载体。

10 25. 一种分离的多肽,其中包含 SEQ ID NO: 8。

26. 一种分离的多核苷酸,其编码权利要求 25 所述多肽。

27. 一种分离的多核苷酸,其中包含 SEQ ID NO: 7。

28. 一种分离的多肽,其中包含 SEQ ID NO: 10。

29. 一种分离的多核苷酸,其编码权利要求 28 所述多肽。

15 30. 一种分离的多核苷酸,其中包含 SEQ ID NO: 9。

31. 一种重组载体,其中包含权利要求 27 所述的多核苷酸。

32. 一种重组载体,其中包含权利要求 30 所述的多核苷酸。

33. 一种在哺乳动物宿主细胞中制备 ABC1 蛋白的方法,该方法包括:

20 (a) 用含有编码 ABC1 的多核苷酸的重组表达载体转染哺乳动物宿主细胞,其中所述多核苷酸的量足以生产出可检测水平的 ABC1 蛋白;以及

(b) 纯化制备到的 ABC1 蛋白。

34. 权利要求 33 中的方法,其中所述的重组表达载体包括权利要求 3 所述的多核苷酸。

25 35. 一种用于在哺乳动物受体细胞内表达 ABC1 的方法,该方法包括下述步骤:向哺乳动物受体施用含有编码 ABC1 的多核苷酸的重组表达载体,

其中所述多核苷酸的量足以在哺乳动物受体细胞内表达 ABC1。

36. 权利要求 35 中的方法，其中所述的重组表达载体包括权利要求 3 所述的多核苷酸。

37. 一种适于增加哺乳动物受体细胞胆固醇排放的方法，该方法包括：  
5 向哺乳动物受体施用含有编码 ABC1 的多核苷酸的重组表达载体，其中所述多核苷酸的量足以增加所述细胞的胆固醇排放。

38. 权利要求 37 中的方法，其中所述的重组表达载体包括权利要求 3 所述的多核苷酸。

39. 权利要求 37 中的方法，其中所述的重组表达载体是一种病毒递送  
10 载体。

40. 权利要求 39 中的方法，其中所述的病毒递送载体是一种腺病毒载体。

41. 权利要求 39 中的方法，其中所述的病毒递送载体是一种慢病毒载体。

42. 权利要求 37 中的方法，其中所述的重组表达载体是一种非病毒递送  
15 载体。

43. 一种适于增加哺乳动物受体细胞胆固醇排放的方法，该方法包括：  
通过一非病毒递送系统向哺乳动物受体施用编码 ABC1 的多核苷酸，其中所述多核苷酸的量足以增加所述细胞的胆固醇排放。

44. 权利要求 43 中的方法，其中所述的编码 ABC1 的多核苷酸包括权利  
20 要求 3 所述多核苷酸。

45. 权利要求 43 中的方法，其中所述的非病毒载体选自： DNA-配体复  
合体、腺病毒-配体-DNA 复合物、腺联病毒-配体-DNA 复合物、DNA 直接  
注射、CaPO<sub>4</sub> 沉淀、基因枪技术、电穿孔、脂质体、和脂质转染法。

46. 一种适于增加哺乳动物受体细胞胆固醇排放的方法，该方法包括下  
25

述步骤：向哺乳动物受体施用治疗量的化合物，该化合物可增加所述细胞的 ABC1 表达。

47. 权利要求 46 中的方法，其中所述的化合物为 cAMP 类似物。

48. 权利要求 47 中的方法，其中所述的化合物选自 8-溴代 cAMP、N6-  
5 苯甲酰基、cAMP 和 8-硫代甲基 cAMP。

49. 权利要求 47 中的方法，其中所述的化合物能增加 cAMP 的合成。

50. 权利要求 49 中的方法，其中所述的化合物是弗司扣林。

51. 权利要求 46 中的方法，其中所述的化合物是磷酸二酯酶抑制剂。

52. 权利要求 51 中的方法，其中所述的化合物选自咯利普兰、茶叶碱、  
10 3-异丁基-1-甲基黄嘌呤、R020-1724、长春西汀、扎普司特、双嘧达莫、米力农、氨力农、pimobendan、西洛酰胺、依诺昔酮、潘乐新针、和维司力农

53. 一种适于增加哺乳动物受体细胞中 ABC1 基因表达的方法，该方法包括下述步骤：向哺乳动物受体施用 cAMP 类似物，该类似物的用量足以增加所述细胞中 ABC1 的表达。

15 54. 一种适于增加哺乳动物受体细胞胆固醇排放的方法，该方法通过下述操作实现：向哺乳动物受体施用能增强 ABC1 活性的化合物，该化合物的用量足以增加胆固醇的排放。

55. 一种筛选待测化合物测定其是否能促进培养物细胞的 ABC1-介导胆固醇排放的方法，该方法包括：

20 测定培养物中哺乳动物细胞样本的胆固醇排放水平，测定胆固醇排放调控水平；

将所筛选的待测化合物与所述细胞接触；

测定接触待测化合物后，细胞样本的胆固醇排放水平；

测定接触待测化合物后，细胞样本的 ABC1-介导胆固醇排放水平，从而  
25 测定出该待测化合物是否能促进细胞培养物的 ABC1-介导胆固醇排放水平。

56. 权利要求 55 中的方法，其中所述的培养物细胞源自细胞系。

57. 权利要求 56 中的方法，其中所述的细胞系选自：成纤维细胞、巨噬细胞、肝及肠细胞。

58. 权利要求 57 中的方法，其中所述的细胞系是 RAW264.7。

5 59. 权利要求 55 中的方法，其中所述的 ABC1-依赖性胆固醇排放水平用抗-ABC1 抗体测定，该抗体当与 ABC1 结合时能抑制 ABC1 的活性。

60. 权利要求 55 中的方法，其中所述的 ABC1-依赖性胆固醇排放水平用反义 ABC1 多核苷酸测定。

61. 权利要求 60 中的方法，其中所述的多核苷酸包含 SEQ ID NO: 57。

10 62. 一种用于检测哺乳动物受体细胞中 ABC1 表达相对水平的方法,该方法包括下述步骤:

从哺乳动物受体获取细胞样本;

测定该细胞样本中 ABC1 mRNA 表达水平;

15 将细胞样本中 ABC1 mRNA 表达水平与预定的 ABC1 mRNA 表达标准水平进行比较，从而检测出哺乳动物受体细胞中 ABC1 基因表达的相对水平。

63. 权利要求 62 中的方法，其中对哺乳动物受体细胞中 ABC1 表达相对水平的检测是通过测定哺乳动物冠心病易感程度来实现的。

64. 权利要求 62 中的方法，其中所述的 ABC1 mRNA 表达水平检测是通过反转录聚合酶链反应、RNA 印迹法转印、或 RNase 保护试验来实现的。

20 65. 一种用于检测哺乳动物受体细胞中 ABC1 蛋白相对水平的方法,该方法包括下述步骤:

从哺乳动物受体获取细胞样本;

测定细胞样本中 ABC1 蛋白的量;

25 将细胞样本中 ABC1 蛋白的量与预定的 ABC1 标准量进行比较，从而检测出哺乳动物受体细胞中 ABC1 蛋白的相对水平。

66. 权利要求 65 中的方法, 其中对哺乳动物受体细胞中 ABC1 蛋白相对水平的检测是通过测定哺乳动物冠心病易感程度来实现的。

67. 权利要求 65 中的方法, 其中所述的测定 ABC1 蛋白的试验包括免疫测定。

5        68. 权利要求 67 中的方法, 其中所述的 ABC1 蛋白的量是通过下述步骤测定的:

(a) 将细胞样本与一系列抗-ABC1 抗体接触, 以及 (b) 检测出与细胞样本相关的特异性结合的 ABC1 抗体。

69. 权利要求 68 中的方法, 其中所述的 ABC1 抗体用蛋白质印迹、免疫  
10 沉淀或 FACS 来检测。

70. 一种分离的抗体, 该抗体能特异性结合权利要求 1 或 2 所述的分离多肽。

71. 权利要求 70 中的抗体, 其中所述的抗体是单克隆抗体。

72. 权利要求 70 中的抗体, 其中所述的抗体是多克隆抗体。

15        73. 权利要求 70 中的抗体, 其中所述的抗体在与 ABC1 多肽结合时能抑制 ABC1 多肽的胆固醇排放活性。

74. 一种适于筛选化合物测定该化合物是否能调控 ABC1-依赖性胆固醇排放的试剂盒, 该试剂盒包括灭活的抗-ABC1 抗体和使用说明书, 该抗体的量足以完成至少 1 次测定。

20        75. 一种适于筛选化合物测定该化合物是否能调控 ABC1-依赖性胆固醇排放的试剂盒, 该试剂盒包括反义 ABC1 寡核苷酸和使用说明书, 该抗体的量足以完成至少 1 次测定。

76. 权利要求 75 中的试剂盒, 其中所述的反义 ABC1 寡核苷酸包含 SEQ  
ID NO: 53

# 说明书

---

## 使用ATP结合框运输蛋白ABC1增加 胆固醇排放并提高HDL的组合物及方法

5

本申请请求下列申请的优先权：1999年6月18日提交登记号为60/140264的美国临时申请、1999年9月14日提交登记号为60/153872的美国临时申请以及1999年11月19日提交登记号为60/166573的美国临时申请。

### 10 技术领域

本发明涉及新的ABC1多肽及编码该多肽的核酸分子。本发明还涉及重组载体、宿主细胞、和含有ABC1多肽的组合物，以及制备ABC1多肽的方法。本发明还涉及可特异性结合ABC1多肽的抗体。此外，本发明还涉及用于增加胆固醇排放量的方法以及用于增加ABC1表达量及增强其活性的方法。

15 本发明还涉及用来鉴定可调控ABC1表达的化合物的方法以及用于检测哺乳动物体内ABC1多肽及多核苷酸的相对水平的方法。本发明还提供了适于筛选化合物测定该化合物调控ABC1表达的活性的试剂盒和组合物，以及适于测定化合物是否能够调控ABC1-依赖性胆固醇的排放量的试剂盒和组合物。

### 20 发明背景

人体血浆及淋巴液中的循环脂类由胆固醇、胆固醇酯(cholesteryl esters)、甘油三酯及磷脂组成。这些酯类通过名为脂蛋白的大分子复合体运输，该复合体包括：一胆固醇酯和/或甘油三酯的核、一由磷脂及游离胆固醇组成的包膜，以及载脂蛋白(Scriver等人, Eds., *遗传病的代谢及分子基础*, 第7版., p.1841-1851 (McGraw-Hill, New York 1995))。载脂蛋白参与脂蛋白的组装和

25

分泌，以及脂蛋白修饰酶的活化，例如卵磷脂胆固醇酰基转移酶 (LCAT)。此外，载脂蛋白能提供结构完整，并可作为大量受体及膜休克蛋白的配体。血浆脂蛋白根据大小分为5种类型：乳糜微粒 (个体最大密度最低)、极低密度脂蛋白 (VLDL)、中等密度脂蛋白 (IDL)、低密度脂蛋白 (LDL) 以及  
5 高密度脂蛋白 (HDL)。

乳糜微粒、VLDLs、IDLs、以及LDLs将外源及内源胆固醇和三酰甘油运输到外周部位，在此处脂类参与各种代谢途径，作为细胞膜的组要成分。乳糜微粒由小肠粘膜合成，是将食物中胆固醇和三酰甘油运输到各种组织中的工具。VLDLs 在肝中生成，负责将肝中合成的胆固醇和三酰甘油运输到肝外  
10 组织，例如肌肉和脂肪组织。空腹血清中， VLDLs 含有占总血清胆固醇总量10-15%的血清胆固醇和大部分的甘油三酯。循环系统中， VLDLs通过脂蛋白脂酶的作用转化为LDLs。 LDLs是将胆固醇分送到所有组织的主要血浆载体，一般含有60 -70%的总空腹血清胆固醇。

相反，HDLs参与“反向胆固醇运输”，通过该途径外周部位的过量胆固醇被运回到肝中，以胆盐的形式排泄出去 (Glomset, J.A., *J. Lipid Res.*, 9, 155-167 (1968))。初生的HDLs在肝和小肠中从头合成，以富含蛋白的碟状颗粒的形式存在，其中不含胆固醇和胆固醇酯。实际上， HDLs 的主要功能是作为载脂蛋白的循环储备，主要为apo C-I, aco C-II及apo E。通过积累细胞来源的胆固醇酯新生或富含蛋白的HDLs转化为球形脂蛋白颗粒。通常，HDL  
20 含有20-30% 的总空腹血清胆固醇。

根据当前理论，细胞胆固醇向HDL的反向流动是通过下述两机制介导的：水扩散途径和载脂蛋白介导途径。这些不同途径之间的相对重要程度取决于细胞类型和代谢状态 (Oram等., 脂类研究杂志 *J LipidRes.*, 37:2743-2491 (1996); Rothblat 等人, *J. LipidRes.*, 40:7g1-796 (1999); Stein Ct al.,  
25 *Atherosclerosis*, 144:285-301 (1999))。对于多数细胞而言，水扩散途径是胆

固醇排放的主要途径 (Johnson等人, *Biochim. Biophys. Acta*, 1085:273-298 (1991))。在细胞外空间内通过被动运输, 该途径可实现胆固醇在细胞膜和脂蛋白受体之间的双向交换, 例如HDL, (Remaley等人, *Arterioscler. Thromb. Vasc. Biol.*, 17:1813-1821 (1997); Rothblat 等人, *J Lip. Res.*, 40:781-796 (1999))。上述交换也可发生在称为小窝的表面微区 (Fielding 等人, *Biochemistry*, 34:14288-14292 91995))。网状流动通过在LCAT的作用将细胞外腔室中的胆固醇转化为胆固醇酯来实现驱动。

或者, 在巨噬细胞和成纤维细胞中, 胆固醇和磷脂流动主要是通过脂蛋白介导的, 例如, apo A-I, apo A-II, 和 Apo E (Remaley, swpvu (1997); Fmncis, 等人, *J. Clin. Invest.*, 96:78-87 (1995); Vega 等人, *J Intern. Med.*, 226:5-15 (1989); Sakar 等人, *Biochim. Biophys. Acta*, 1438: 85-98 (1999); Hara 等人, *BioL Chem.*, 266:3080-3086 (1991); Fielding 等人, *J Liptd Res.* 38, 1503-1521 (1997); Oram 等人, *J. Lipid Res.*, 374.2743-2491 (1996))。上述脂蛋白介导的脂类排放途径在巨噬细胞及其他清除细胞中占绝对主导地位, 尤其是当这些细胞承载胆固醇和/或生长停滞时。载脂蛋白-介导的流动是一种活性运输途径, 需要载脂蛋白与细胞表面间的直接相互作用, 载脂蛋白的脂化以及随后来自细胞的脂类-载脂蛋白颗粒的解离 (Oram, supnz (1996); Mendez, A., *J Lipid Re*, 38, 1807-1821 (1997); Remaley, *supra* (1997); Mendcz, AJ., *J LtpidRes.*, 37,2510-2524 (1996))。一旦被清除出细胞, 富含胆固醇的HDL颗粒就被运输到肝, 或如上所述排出体外。

遗传缺陷或者作为另一疾病的次级效应而引发的脂蛋白功能和/或异常, 能够导致严重的生物学后果。除饮食外, 糖尿病、甲状腺机能减退以及肝病都能导致LDL-胆固醇和甘油三酯血浆水平的提高。LDL-胆固醇和甘油三酯水平的提高已被视为引发冠心病的主要危险因素, 而冠心病是美国及其他发达国家死亡的主要原因 (Hokanson 等人, *J Cardiovasc. Risk.*, 3:213-219 (19%);

The Expert Panel, JAAM, 269:3015-3023 (1993))。过量LDL-胆固醇积聚在动脉壁上会引发动脉粥样硬化病变的出现, 而该病变在心脏病的发病过程中起着主要作用。病变的形成是由于动脉壁释放出的基团氧化LDL而形成。根据该理论, LDL的氧化激发炎症应答, 将循环细胞吸引到病发部位。在这些细胞中包括巨噬细胞和其他含有清除受体的细胞, 该受体以非调控的方式积累胆固醇 (Brown等, *Ann. Rev. Biochem.*, 52:223-261 (1986))。内源胆固醇的大量储存导致泡沫细胞表型的转化, 该转化是引发血管病变的主要原因。随着噬斑的增大, 动脉壁收缩, 流向心脏的血排放量减小。

但是, 令人惊奇的是, 约有60%的心脏病发生在LDL-胆固醇并未升高的人群中。这些人当中, 约有45%的血中HDL-胆固醇低于平均水平, 这表明低的HDL-胆固醇水平是引发心脏病的重要因素。实际上, 近来研究发现HDL-胆固醇水平的降低是早期冠心病患者中最常见的脂蛋白异常 (Genest J, *Circulation*, 85:2025-2033 (1992); Genest 等人, *Arterioscler. Thromb.*, 13:1728-1737 (1993))。尽管HDL-胆固醇与冠心病之间的反向联系尚不清楚, 但是已经说明, HDL的心脏保护作用来自于其下述活性: 即能加速胆固醇从动脉粥样硬化病变中巨噬细胞泡沫细胞的排放。

与低HDL相关的心血管病的一个实例是丹尼尔病 (TD), 这是一种罕见的遗传疾病, 特征为循环HDL近乎或完全缺乏。除近乎为零的HDL水平外, TD患者的另一特征是: 胆固醇大量沉积及累积在几种组织中, 包括扁桃体、淋巴结、肝、脾、胸腺、肠及神经膜细胞 (Fredrickson, D.S., *J. Clin Invest.* 43:228-236 (1964); Assmann等人, *The Metabolic Basis of Inherited Disease*, (Mcraw-Hill, New York, 1995))。尽管此前未鉴定出该细胞机制, 但是近来的研究表明, TD患者的细胞有着脂蛋白介导的胆固醇及磷脂清除途径的缺陷 (Remaley 等人, *Arterioscler. Thromb. Vasc. Biol.*, 17, 1813-1821 (1997); Francis 等人, *J. Clin Invest.*, 96, 78-87 (1995); Rogler等人, *Arterioscler. Thromb. Vasc.*

*BioL*, 15,683-690 (1995))。这些结果导致这样的观点：TD患者中HDL的严重缺陷是由于新生apo A-I 无法获得脂类。因为新生的HDL无法成熟为富含脂的颗粒，因而就被快速代谢并清除出血浆，导致循环HDL水平近乎为零 (Remaley, *supra* (1997); Francis, *supra* (1995); Horowitz 等人,): *Clin. Invest.*, 91, 1743-1752 (1993); Schaefer等人, *Lip. Res.*, 22:217-228 (1981))。

其他与严重早期动脉粥样硬化及起因于HDL-胆固醇水平降低的冠心病高度危险相关的疾病是低 $\alpha$ 脂蛋白血症及家族性HDL缺陷综合征 (FHA)。患有这些疾病的病人通常具有正常水平的LDL-胆固醇及甘油三酯。此外，糖尿病、酒精中毒、甲状腺机能减退、肝病以及高血压等疾病也能导致血浆HDL-胆固醇水平的降低，尽管这些疾病中多数也还伴有LDL-胆固醇及甘油三酯水平的升高。

目前，冠心病的治疗主要集中在饮食操控和/或旨在通过抑制LDL分泌及促进LDL更新来降低LDL-胆固醇血浆水平的药物治疗。纤维酸 ( fibric acid) 的衍生物，例如clofibrate、gemfibrozil和 fenofibrate，能够通过活化脂蛋白脂酶促进VLDL的快速更新。烟酸可以通过抑制肝VLDL分泌降低VLDL和LDL的血浆水平。此外， HMG-CoaA 还原酶抑制剂，例如mevinolin、mevastatin、pravastatin、simvastatin、fluvastatin、和 lovastatin 通过抑制胆固醇的胞内合成来降低血浆LDL水平，从而增加了LDL的细胞摄入量。此外，胆酸-结合树脂，例如cholestyrene、colestipol 和 probucol 通过增加肝中LDL-胆固醇的代谢来降低LDL-胆固醇的水平。

但是，这些治疗方法多数具有低效和/或可能妨碍长期使用的副作用。例如，使用HMG-CoaA还原酶抑制剂会带来严重的毒性，因为其会抑制甲羟戊酸的合成，而该物质是除胆固醇外其他重要类异戊二烯化合物合成必不可少的。另外，吉非贝齐和烟酸具有严重的副作用，包括肾损伤、肌病、肌球蛋白尿和难以忍受的皮肤发红及瘙痒。此外，普罗布考在治疗冠心病中的作

用还不确定，这是由于使用它降低LDL-胆固醇的同时会带来导致HDL-胆固醇水平更低的副作用。

5 另外，治疗已经离析且具有低HDL-胆固醇水平的患者会带来特别困难的治疗难题。例如即使TD病患者的LDL-胆固醇已经降低了约50%，他们还是在心血管疾病发生上增加了4-6倍。尽管有证据显示吉非贝齐和烟酸可以同时增加HDL，但是通常，由于HDL水平的降低，旨在降低血浆LDL胆固醇水平的治疗对于患有冠心病的丹尼尔患者来说不起作用。类似地，低 $\alpha$ 脂蛋白血症、家族性HDL缺陷综合征、或者其他低水平HDL引发的心血管疾病都不会从旨在降低血浆LDL的治疗方法中获益。

10 与当前心血管疾病治疗相关的问题部分起因于对胆固醇出入细胞的运动过程还没有完全理解。另外，在胆固醇运动中其作用的蛋白质也还未完全认知。因此，仍有必要对胆固醇细胞生物学以及用于治疗心血管疾病及其他高胆固醇血症相关疾病的新方法做进一步的理解。此外，还有必要发展一种新方法，用于诊断心血管疾病以及用来筛查心血管疾病高危人群的新方法。

15 涉及胆固醇运输的基因及蛋白的鉴定可以用来开发用于治疗心脏病及其他高胆固醇血症和动脉粥样硬化的药物。此外，这类基因的鉴定还有助于开发筛选方法，筛选出可调控胆固醇运输相关基因表达的化合物。这类调控化合物的鉴定还有助于其他治疗药物的开发。另外，参与胆固醇运输的基因及蛋白的鉴定还可以用作心血管疾病及其他高胆固醇血症相关疾病的诊断指示  
20 剂。

## 发明内容

本发明提供了参与胆固醇排放的新的多肽及多核苷酸。具体而言，本发明提供了新的ATP-结合框 (ABC1)多肽及编码ABC1多肽新的多核苷酸。术语  
25 “ABC1”和““ABCA1”为同一ATP-结合框蛋白及基因的别称。本发明还

提供了ABC1 多肽、多肽片段及多态变体。在一个优选实施方案中，本发明提供了一种包括SEQ ID NO: 2的分离多肽。另一个优选实施方案中，本发明提供了一种分离多态，其氨基酸序列与SEQ IDNO: 2有至少98%的同源性。本发明还提供了来自丹尼尔疾病患者的ABC1多肽。一个优选实施方案中，本  
5 发明提供了一种包括SEQ ID NO: 8的分离多肽。另一个优选实施方案中，本发明提供了一种包括SEQ ID NO: 10的分离多肽。

此外，本发明还提供了ABC1多核苷酸、多核苷酸片段、以及多核苷酸变体。在一个优选实施方案中，本发明提供了编码包括SEQ ID NO: 2的多肽的多核苷酸。另一优选实施方案中，本发明提供了编码包括与SEQ ID NO: 2至  
10 少98%同源的氨基酸序列的多肽的多核苷酸。另外，另一个优选实施方案中，本发明提供了一分离多核苷酸，其中包括与编码包括SEQ ID NO: 2的多肽的多核苷酸互补的核苷酸序列，或者另一分离多核苷酸，其中包括与编码多肽的多核苷酸互补的核苷酸序列，所述多肽含有与SEQ ID NO: 2至少98%同源的氨基酸序列。

15 另一优选实施方案中，本发明还提供了一包括SEQ ID NO: 1的分离ABC1多核苷酸。另一优选实施方案中，本发明提供包括SEQ ID NO: 1中第291-7074位核苷酸的分离多核苷酸。另一优选实施方案中，本发明提供了包括与SEQ ID NO: 1至少90%同源的核苷酸序列的多核苷酸。更优选地，所述多核苷酸包括与SEQ ID NO: 1至少95%、96%、97%、98%或99%同源的核苷酸序列。  
20 另外，其他优选实施方案中，本发明提供了具有下述特征的分离多核苷酸：其中包括与含SEQ ID NO: 1的多核苷酸互补的核苷酸序列、其中包括与含SEQ ID NO: 1中第291-7074位核苷酸的多核苷酸互补的核苷酸序列，以及与SEQ ID NO: 1至少90%同源的核苷酸序列的多核苷酸互补的核苷酸序列。

本发明还提供了与ABC1基因5'端侧翼区对应的ABC1多核苷酸。一个优  
25 选实施方案中，本发明提供了一包括SEQ ID NO: 3的分离多核苷酸。另一实

5 施 方 案 中 ， 本 发 明 提 供 了 含 有 SEQ ID NO: 3 中 第 1-1532,1080-1643,1181-1643,1292-1643,或1394-1532核苷酸的分离多核苷酸。 优选地，包括SEQ ID NO: 3中第1394-1532位核苷酸的多核苷酸。另一优选实施方案中，本发明提供了即能在严格条件下与含SEQ ID NO: 3的多核苷酸杂交的分离多核苷酸。 另外，其他优选实施方案中，本发明提供了具有下述特征的多核苷酸：即能在严格条件下与含SEQ ID NO: 3 中第 1-1532,1080-1643,1181-1643,1292-1643， 或1394-1532的核苷酸的多核苷酸杂交。本发明的另一优选实施方案中，提供了包括与含SEQ ID NO: 3的多核苷酸至少80%同源的分离多核苷酸。更优选地，所述多核苷酸与含SEQ ID NO: 10 3的多核苷酸至少90%同源。甚至更优选，与含SEQ ID NO: 3的多核苷酸至少95%同源。另外的优选实施方案中，本发明提供了与含SEQ ID NO: 3第 1-1532,1080-1643,181-1643,1292-1643， 或1394-1532核苷酸的多核苷酸至少80%同源的分离多核苷酸。更优选地，所述多核苷酸与含SEQ ID NO: 3第 1-1532,1080-1643,1181-1643,1292-1643， 或1394-1532核苷酸的多核苷酸至少 15 90%同源， 甚至更优选地，所述多核苷酸与含SEQ ID NO: 3中第 1-1532,1080-1643,1181-1643,1292-1643， 或1394-1532核苷酸的多核苷酸至少 95%同源。此外，本发明提供了包括与上述ABC1 5'侧翼区互补的核苷酸序列的多核苷酸。一个优选实施方案中，本发明提供了一分离的多核苷酸，其中包括与含SEQ ID NO: 3多核苷酸互补的核苷酸序列。另一优选实施方案中， 20 本发明提供了一具有下述特征的分离核苷酸：即其中包括与含SEQ ID NO: 3 中第1-1532,1080-1643,1181-1643,1292-1643， 或1394-1532核苷酸的多核苷酸互补的核苷酸序列。

25 本 发 明 还 提 供 了 与 ABC1 基 因 3' 端 侧 翼 区 对 应 的 ABC1 多 核 苷 酸 。 一 个 优 选 实 施 方 案 中 ， 本 发 明 提 供 了 一 包 括 SEQ ID NO: 4、SEQ ID NO:5或SEQ ID NO:6的分离多核苷酸，及其互补序列。其他优选实施方案中，本发明提供了

一分离多核苷酸，其能在严格条件下与含SEQ ID NO: 4、SEQ ID NO:5或SEQ ID NO:6的分离多核苷酸杂交。另一优选实施方案中，本发明提供了与含SEQ ID NO: 4、SEQ ID NO:5或SEQ ID NO:6的分离多核苷酸至少80%同源的多核苷酸，及其互补序列。更优选地，本发明提供了与含SEQ ID NO: 4、SEQ ID NO:5或SEQ ID NO:6的分离多核苷酸至少90%同源的多核苷酸。甚至更优选地，本发明提供了与含SEQ ID NO: 4、SEQ ID NO:5或SEQ ID NO:6的分离多核苷酸至少95%同源的多核苷酸。

此外，本发明还提供了来自丹尼尔病患者的ABC1 多肽。另一优选实施方案中，本发明提供了一编码含SEQ ID NO: 8多肽的分离多核苷酸。另一优选实施方案中，本发明提供了一包括SEQ ID NO: 7的分离多核苷酸。另一优选实施方案中，本发明提供了一编码含SEQ ID NO: 10多肽的分离多核苷酸。另一优选实施方案中，本发明提供了包括SEQ ID NO: 9的分离多核苷酸。本发明进一步还提供了含与上述核苷酸互补的核酸序列。

另一方面，本发明提供了包括任一上述多核苷酸和一合适载体的组合物。一优选实施方案中，本发明提供了一种组合物，其中包括编码含SEQ ID NO: 2的多肽的分离多核苷酸、编码含SEQ ID NO: 1的多核苷酸、含SEQ ID NO: 1中第291-7074位核苷酸的多核苷酸、或编码含与SEQ ID NO: 2至少98%同源的氨基酸序列的多肽的多核苷酸，以及适当载体。另一优选实施方案中，所述组合物包括一分离的多核苷酸和适当的载体，所述多核苷酸中包括与含SEQ ID NO: 1多核苷酸至少90%同源的核苷酸序列。其他优选实施方案中，所述组合物包括含SEQ ID NO: 3的分离多核苷酸，或者含SEQ ID NO: 3中第1-1532,1080-1643,1181-1643, 1292-1643, 或 1394-1532位核苷酸的多核苷酸，以及适当载体。另一些优选实施方案中，本发明提供了一种组合物，其中包括能在严格条件下与含SEQ ID NO: 3的多核苷酸杂交的多核苷酸，或包括SEQ ID NO: 3中第1-1532,1080-1643,1181-1643, 1292-1643, 或 1394-1532位

核苷酸的多核苷酸，以及这样的组合物，其中包括与含SEQ ID NO: 3的多核苷酸至少 80 % 同源的多核苷酸，或者含 SEQ ID NO: 3 中第 1-1532,1080-1643,1181-1643, 1292-1643, 或 1394-1532位核苷酸的多核苷酸，以及适当载体。另外，本发明还提供了一种组合物，其中包括与上述任一多核苷酸互补的分离多核苷酸及适当载体。

此外，本发明还提供了其中含有任一上述ABC1多核苷酸序列的重组载体和宿主细胞。一个优选实施方案中，本发明提供了包含下列多核苷酸的重组载体：编码含SEQ ID NO: 2多肽的分离多核苷酸、含SEQ ID NO: 1的分离多核苷酸、含SEQ ID NO: 1中的第291-7074位核苷酸的分离多核苷酸、或者编码含与SEQ ID NO: 2至少98%同源的氨基酸序列的多肽的分离多核苷酸。另一优选实施方案中，重组载体中含有一分离多核苷酸，该多核苷酸与含SEQ ID NO: 1的多核苷酸至少90%同源，更优选至少95%同源。另一优选实施方案中，所述重组载体包括含SEQ ID NO: 7或SEQ ID NO: 9的分离多核苷酸。本发明进一步还提供了包含这样的分离多核苷酸的重组载体，即该多核苷酸包括与任一上述核苷酸互补的多核苷酸。另一实施方案中，所述重组载体包括任一上述多核苷酸以及还包括一异源启动子多核苷酸。一适当的异源启动子为巨细胞病毒启动子。在一具体的优选的实施方案中，所述的重组启动子为pCEPhABC1。

本发明还提供了一种包含分离多核苷酸的重组载体,该多核苷酸包含ABC1 5'侧翼区。一个优选实施方案中，本发明提供了包含这样的分离多核苷酸的重组载体，该分离多核苷酸包括SEQ ID NO: 3 或者包括SEQ ID NO: 3 中的第1-1532,1080-1643,1181-1643, 1292-1643, 或 1394-1532位核苷酸。其他实施方案中，本发明提供了包含这样的多核苷酸的重组载体，即该多核苷酸能在严格条件下与含SEQ ID NO: 3的多核苷酸杂交，或者该多核苷酸中包括SEQ ID NO: 3中的第1-1532,1080-1643,1181-1643,1292-1643, 或 1394-1532

位核苷酸，以及这样的重组载体，其中包括与这些多核苷酸至少80%同源的多核苷酸。本发明还包含这样的多核苷酸的重组载体，即该多核苷酸包括与上述任一多核苷酸互补的核苷酸序列。另一优选实施方案中，所述重组载体包括任一上述多核苷酸以及进一步包括至少一编码异源多肽的多核苷酸。合适的异源多肽包括荧光素酶、半乳糖苷酶、氯霉素乙酰转移酶转移酶、以及绿色荧光蛋白。优选地，该异源多肽为荧光素酶蛋白。在一具体优选的实施方案中，重组载体为pAPRI。

另外，本发明还提供了含有任一上述重组载体的宿主细胞。本发明还提供了含任一上述重组载体及适当载体的组合物。

10 本发明还提供了在哺乳动物宿主细胞中制备ABC1蛋白的方法以及在哺乳动物受试者体内表达ABC1蛋白的方法。所述在哺乳动物宿主细胞中制备ABC1蛋白的方法包括下述步骤：(a)用含有编码ABC1的多核苷酸的重组表达载体转染哺乳动物宿主细胞，其中多核苷酸的量足以产生可检测水平的ABC1蛋白，以及(b)纯化产生的ABC1蛋白。所述在哺乳动物受试者体内表  
15 达ABC1蛋白的方法包括下述步骤：向哺乳动物受试者施用含编码ABC1多核苷酸的重组载体，其中所述多核苷酸的量足以在哺乳动物受试者体内表达ABC1蛋白。

此外，本发明还提供了用于增加哺乳动物受试者细胞胆固醇排放量的组合物和方法。一个优选实施方案中，所述方法包括向哺乳动物受试者施用含  
20 编码ABC1的多核苷酸的表达载体，其中多核苷酸的量足以增加细胞的胆固醇排放量。合适的重组载体包括含有下述分离多核苷酸的载体：编码含SEQ ID NO: 2的多肽的多核苷酸、含SEQ ID NO: 1的多核苷酸、含SEQ ID NO: 1中第291-7074位核苷酸的分离多核苷酸、编码含包括与SEQ ID NO: 2至少98%同源的氨基酸序列的多肽的分离多核苷酸。优选的表达载体包括病毒载体、具  
25 体为腺病毒载体和慢病毒载体。其他实施方案中，本发明提供了非病毒递送

系统，包括DNA-配体复合体、腺病毒-配体-DNA复合体、DNA直接注射、CaPO<sub>4</sub>沉淀、基因枪技术、电穿孔、脂质体和脂质转染法。

5 在另一优选实施方案中，提供了用于增加哺乳动物受试者细胞胆固醇排放量的方法，该方法包括向哺乳动物受试者施用治疗量的化合物，该化合物能增强细胞内ABC1的表达。一种适当的方法包括向哺乳动物受试者施用cAMP类似物。适当的cAMP类似物包括8-溴cAMP、N6-苯甲酰基cAMP、及8-硫代甲基cAMP。另一适当的方法包括向哺乳动物受试者施用可增加cAMP合成的化合物，例如弗司扣林。另一合适的方法包括向哺乳动物受试者施用可抑制cAMP降解的化合物，例如磷酸二酯酶抑制剂。合适的磷酸二酯酶抑制剂包括咯利普兰、茶叶碱、3-异丁基-1-甲基黄嘌呤、R020-1724、长春西汀、扎普司特、双嘧达莫、米力农、氨力农、pimobendan、西洛酰胺、依诺昔酮、潘乐新针、和维司力农。

此外，另一适于增加哺乳动物受试者细胞胆固醇排放量的方法包括向受试者施用至少一种核受体的配体，其用量足以增加胆固醇的排放量。适当的配体包括LXR、RXR、FXR、SXR 和PPAR受体的配体。一个优选的实施方案中，该方法包括向哺乳动物受试者施用LXR 核受体的配体。适当的LXR配体包括20 (S) 羟基胆固醇、22 (R) 羟基胆固醇、24 (S) 羟基胆固醇、25-羟基胆固醇和24 (S), 25 .环氧胆固醇。优选地，所述LXR 配体为20 (S)羟基胆固醇。另一优选的实施方案中，所述方法包括向哺乳动物受试者施用RXR核受体的配体。合适的RXR配体包括9-顺式视黄酸、维生素A醇、视黄醛、所有反式视黄酸、13-顺式视黄酸、阿维acitretin、fenretinide、阿维A酯、CD 495、CD564、TTNN, TTNNPB、TTAB、和LGD 1069。所述RXR配体优选为9-顺式视黄酸。另一优选的实施方案中，所述方法包括向哺乳动物受试者施用PPAR核受体的配体。一合适的配体选自噻唑烷二酮。另一优选的实施方案中，所述方法包括施用至少两种核受体的配体。在一个具体的优选实施方案

中，所述配体为20 (S) 羟基胆固醇和9-顺式视黄酸。

此外，另一适于哺乳动物受试者细胞胆固醇排放量的方法包括向哺乳动物受试者施用类二十烷酸，其用量足以增加胆固醇的排放量。合适的类二十烷酸包括前列腺素E2、前列腺素J2、和前列腺环素 (前列腺素12)。

5 另一实施方案中，本发明提供了用于增加哺乳动物受试者细胞胆固醇排放量的方法，该方法包括向哺乳动物受试者施用可增强ABC1活性的化合物，其用量足以增加细胞的胆固醇排放量。

本发明还提供了适于增加哺乳动物受试者ABC1基因表达的方法。一个优选实施方案中，所述方法包括向哺乳动物施用至少一种核受体的配体，其用  
10 量足以增加哺乳动物受试者ABC1基因的表达。合适的配体包括LXR、RXR、FXR、SXR 和PPAR受体的配体。另一优选实施方案中，所述方法包括向哺乳动物施用cAMP 类似物，其用量足以增加ABC1的基因表达。另一优选实施方案中，所述方法包括向哺乳动物施用可增加cAMP 合成的化合物，其用量足以增加ABC1的基因表达。

15 此外，本发明还提供了一种用于筛选待测化合物是否具有ABC1 表达调控活性的方法，该包括下列步骤： (a) 可操作地将一报道cDNA与哺乳动物ABC1基因表达调控部分连接生成一重组的报道构建体； (b) 将所述重组报道构建体转染入宿主细胞群； (c)测定宿主细胞样本中的报道基因表达水平； (d) 将所述宿主细胞与所筛选的待测化合物接触； (e)在接触待测化合物后，  
20 测定宿主细胞样本中的报道基因表达水平；以及 (f) 比较暴露待测化合物后引发的报道基因表达的相对改变，从而测定出所述的 ABC1 表达调控活性。所述的重组报道构建体包括可操作地连接到哺乳动物ABC1基因表达调控部分的报道基因，例如本发明提供的任一ABC1 5'侧翼区序列。一个优选实施方案中，所述的ABC1基因表达调控部分包括SEQ ID NO: 3。另一优选的实施方案中，所述的ABC1基因表达调控部分包括SEQ ID NO: 3中的第1-1532、  
25

1080-1643、1181-1643、1292-1643、1394-1643，或1394-1532位核苷酸。合适的报道cDNAs包括荧光素酶、p-半乳糖苷酶、氯霉素和绿色荧光蛋白cDNA。所述宿主细胞优选为哺乳动物细胞。在该方法的一个具体的优选实施方案中，所述的重组报道构建体为pAPR1。

5       另外，本发明还提供了筛选待测化合物测定其是否能促进培养物细胞的ABC1-介导胆固醇排出量的方法，该方法包括：(a)测定培养物中哺乳动物细胞样本的胆固醇排出水平以测定胆固醇排出的调控水平；(b)将所筛选的待测化合物与所述细胞接触；(c)测定接触待测化合物后细胞样本的胆固醇排放水平 (d) 测定接触待测化合物后细胞样本的ABC1-介导胆固醇排放水平，从而测定出该待测化合物是否能促进细胞培养物的ABC1-介导的胆固醇排放。上述细胞可以来自原代培养细胞或细胞系。合适的细胞系包括成纤维细胞、巨噬细胞、肝及肠细胞系。所述细胞系优选为RAW 264.7。一个优选的实施方案中，用抗-ABC1抗体测定ABC1-介导的胆固醇排放，在与ABC1结合时该抗体能抑制其活性。另一优选实施方案中，用反义ABC1多核苷酸测定ABC1-  
10 介导的胆固醇排放。一个具体的优选实施方案中，所述反义多核苷酸含有SEQ ID NO: 57。

此外，本发明还提供了检测哺乳动物受试者细胞中ABC1相对表达水平的方法。这类方法可用来测定受试者的冠心病易感程度。一种用于检测哺乳动物受试者细胞中ABC1相对表达水平的方法，该方法包括(a)从哺乳动物受试者获得细胞样本，(b)测定该细胞样本中ABC1 mRNA表达水平；以及(c)将细胞的ABC1 mRNA表达水平与预定的ABC1 mRNA表达标准水平进行比较，从而检测出哺乳动物受试者细胞中ABC1 基因表达的相对水平。适于测量ABC1 mRNA 表达水平的方法包括，例如RT-PCR、RNA印迹法、以及RNase 保护测定。

25       本发明还提供了检测哺乳动物受试者细胞中ABC1蛋白的相对表达水平

的方法。这类方法可用来测定受试者的冠心病易感程度。一种用于检测哺乳动物受试者细胞中ABC1蛋白的相对表达水平的方法。该方法包括 (a) 从哺乳动物受试者获得细胞样本, (b)测定该细胞样本中ABC1 蛋白的表达水平; 以及 (c) 将细胞样本ABC1蛋白的表达水平与预定的ABC1 蛋白标准量进行比较, 从而检测出哺乳动物受试者者细胞中ABC1 蛋白的相对水平。本领域现有的各种免疫测定方法都可以用来测定ABC1 蛋白的量.例如, ABC1 蛋白量可以通过下述步骤测定: (a)将一系列抗-ABC1抗体与细胞样本接触以及 (b) 检测ABC1 抗体与该样本的结合。适于检测ABC1抗体的方法包括western 印迹、免疫沉淀反应、及FACS。

10 另一方面, 本发明提供了能特异性结合上述ABC1多肽的抗体。一个优选实施方案中, 本发明提供了一种能够与含SEQ ID NO: 2的分离多核苷酸特异性结合的分选抗体。另一个优选实施方案中, 本发明提供了一种可与分离多肽特异性结合的抗体, 所述分离多肽包含与SEQ ID NO: 2至少98%同源的氨基酸序列。所述抗体可以是单克隆抗体或者也可以是多克隆抗体。另一实施  
15 方案中, 当所述抗体与ABC1多肽结合时, 可抑制该ABC1多肽的胆固醇运输活性。

此外, 本发明还提供了适于筛选化合物测定该化合物的ABC1表达调控活性的试剂盒, 其中包括一可操作连接到哺乳动物ABC1 基因表达调控部分的报道cDNA以及使用说明书, 所述cDNA的量足以完成至少一次测定。一个优  
20 选实施方案中, 所述试剂盒还包括用于检测报道基因的工具。另一优选实施方案中, 哺乳动物ABC1 基因表达调控部分包括SEQ ID NO: 3。另一优选实施方案中, 所述哺乳动物ABC1 基因表达调控部分包括SEQ ID NO: 3中的第1-1532、1080-1643、1181-1643、1292-1643、1394-1643或1394-1532位核苷酸。合适的报道cDNAs 包括荧光素酶、 $\beta$ -半乳糖苷酶、氯霉素乙酰转移酶  
25 转移酶、以及绿色荧光蛋白cDNA。优选地, 所述报道cDNA为荧光素酶蛋白。

在一具体优选的实施方案中，所述的重组报道构建体为pAPRI。

本发明还提供了适于筛选化合物测定该化合物是否能调控ABC1-依赖性胆固醇排放的试剂盒。一个优选实施方案中，所述试剂盒包括一灭活的抗-ABC1抗体以及使用说明书，所述抗体的含量足以完成至少一次测定。另一个优选实施方案中，所述试剂盒包括反义ABC1寡核苷酸以及使用说明书，所述寡核苷酸的含量足以完成至少一次测定。一个具体的优选实施方案中，所述的反义ABC1寡核苷酸包括SEQ ID NO: 53。

## 附图说明

10 图1 A-D是以图形形式显示的来自丹吉尔疾病患者的成纤维细胞 (1B, D) 及正常成纤维细胞 (1A, C)在apo A-I及HDL存在下胆固醇的流量及对照组胆固醇的流量的结果。开放圈代表的是来自并未暴露于HDL或apo A-I的细胞的胆固醇流量，封闭圈代表的是来自暴露于HDL的细胞的胆固醇流量，封闭菱形代表的是来自暴露于apo A-I的细胞的胆固醇流量；

15 图2是以基因表达微阵列图形形式显示的来自丹尼尔疾病患者 (TD1) 的细胞及正常细胞中发现的基因表达的对比，其中总共有58,800个人类cDNAs与来自cAMP处理的TD1细胞cDNA (使用Cy3染色标记)的mRNA制备的cDNA进行杂交及与来自cAMP处理的正常细胞cDNA (使用Cy5染色标记) 的mRNA制备的cDNA进行杂交；

20 图3显示的是一个重组表达载体pCEPhABC1 (其含有人类ABC1基因的开放阅读框) 的酶切图谱的示意图；

图4 显示的是一个人类ABC1基因结构的示意图，其也显示了已公开的人类ABC1氨基酸序列 (GenBank, 登记号#AJ012376)及目前已经公开的且有权利要求保护的人类ABC1 氨基酸序列 ("CVT") (其在N末端有60个额外的氨基酸) 的对比；

25

图5是以图形形式显示的ABC1转移抑制剂4,4-均二苯乙烯二异硫氰酸-2,2'二磺酸 (DIDS)及二硫溴苯甲酰 (BSP)的抑制效果,其与apo A-I介导的胆固醇流量相关,其中开放圈代表的是在BSP存在下apo A-I介导的胆固醇流量,封闭圈显示的是在DIDS存在下apo A-I介导的胆固醇流量;

5 图6是以图形形式显示的在apo A-I介导的胆固醇流出中的反义ABC1寡聚核苷的抑制效果,其显示了在没有反义寡聚核苷下培养的细胞中的apo A-I介导的胆固醇流出,在暴露于30 $\mu$ M $\beta$ 球蛋白反义寡聚核苷下培养的细胞中的apo A-I介导的胆固醇流出,及在暴露于30 $\mu$ M ABC1反义寡聚核苷下培养的细胞中的apo A-I介导的胆固醇流出;

10 图7是一个以图形形式显示的由使用人类ABC1的表达质粒(pCEPhABC1)稳定转染的鼠巨噬细胞RAW 264.7的ABC1基因的过量表达导致的对apo A-I介导的胆固醇流量的刺激,其也显示了在对照的亲本细胞中(无pCEPhABC1)及在对照的使用pCEPhABC1 (L3, L5, L6)转染的克隆细胞中的apo A-I介导的胆固醇流量;

15 图8是一个以图形形式显示的反转录酶聚合链式反应 (RT-PCR)分析的在正常细胞中及在来自丹尼尔疾病患者 (TD1及TD2)的细胞中其任一已暴露于白蛋白(封闭柱),暴露于8-溴-cAMP(开放柱),已载胆固醇(黑色柱),或者已载胆固醇及随后暴露于apo A-I(阴影线柱)的ABC1基因表达的水平;

20 图9是一个以图形形式显示的反转录酶聚合链式反应 (RT-PCR)分析的在暴露于乙醇 (0.1%v/v),9-顺-视黄酸 (9-*cis* RA; 10 $\mu$ M),20 (S)羟基胆固醇 (20 (S)-OH; 10 $\mu$ M),或者9-顺RA及20 (S)-OH (每一10 $\mu$ M)的RAW 264.7细胞中ABC1基因表达的水平;

25 图10显示的是免疫沉降反应分析的发现于正常成纤维细胞 (NL1,10A)及来自于丹尼尔疾病患者的细胞 (TD1,10B)在没有添加剂的存在下(对照组),或者在以下试剂:8-Br-cAMP (1mM),胆固醇 (30  $\mu$ g/ml),或者胆固醇及

8-Br-cAMP (分别为30 $\mu$ g/ml及1mM) 的存在下细胞表面ABC1蛋白的水平;

图11是一含有位于荧光素受体基因开放阅读框上游的ABC1基因的5'侧翼区的重组表达载体pAPRI的酶切图谱的示意图;

图12是以图形形式显示的使用pAPRI在EtOH (对照组),20 (S)-OH (10 $\mu$ M),  
5 9-*cis* RA (10 $\mu$ M), 或者20 (S)-OH及9-*cis* RA (每一均为10 $\mu$ M)存在下转染RAW  
264.7细胞中诱导产生的荧光素受体基因表达的水平;

图13是ABC1基因5'侧翼区的示意图.

## 具体实施方式

10 本发明提供了能增加细胞胆固醇排放的新的多肽。具体地, 本发明提供了新的ATP-结合框I (ABC1) 多肽, 已证实该多肽能增加胆固醇排放。ABC1是ATP-结合框蛋白家族的成员, 该家族位于细胞膜上, 通过ATP水解转运多种底物通过胞膜。值得注意的是, 术语"ABC1"和"ABCA1"二者都是指同样的  
15 本领域获得一定仍可。迄今为止, 在人类基因组中, 已鉴定出了该家族的30多个成员。这些同源蛋白包括通道-样结构和一个或两个结构域, 通过通道-样结构分子被运输通过细胞膜, 所述结构域能够与ATP结合以将ATP水解产生的能量结合到转运中。家族成员包括: 多种药物抗性因子 (MDR/P-糖蛋白; Chen 等人, Cell, 47: 381-389 (1986); Stride 等人, Mol. Pharmacol., 49: 962-971  
20 (1996))、抗原提呈相关运输物 (Neefjes 等人, Science, 261: 769-771 (1993); Shepherd 等人, Cell, 74: 577-584 (1993))、以及囊性纤维化跨膜传导调节蛋白 (Chang 等人, J. Biol. Chem., 269: 18572-18575 (1994); Rommens 等人, Science, 245: 1059-1065 (1989))。通常, ABC运输家族的成员包括4个结构域和一高度疏水节段, 上述4个结构域位于两个对称的等分部分内, 这两部分  
25 通过一长带电区域相连。每一等分部分包括一个含6个跨膜节段的疏水结构域

和一疏水核苷酸结合结构域，后者内含多种ATP酶中都常见的高度保守Walker A和B 序列基元 (Hyde 等人, *Nature*, 346: 362-365 (1990); Luciani 等人, *Genomics*, 21: 150-159 (1994))。运输物的活性取决于核苷酸结合结构域与ATP之间的相互作用，通过两对称等分部分之间的连接区域内的残基的磷酸化实现调控 (Becq 等人, *J. Biol. Chem.*, 272 2695-26952699 (1997))。

本发明所述的系列证据表明ABC1是胞内胆固醇存储的载脂蛋白-介导代谢的枢纽蛋白。首先，本发明的研究显示ABC1在丹尼尔病中是缺陷性的，丹尼尔病是一种以HDL-胆固醇代谢异常为特征的遗传病。如前所述，以及本发明的实施例1，丹尼尔病的基因缺陷导致载脂蛋白介导的胆固醇排放出细胞的途径缺失，从而导致胆固醇排放活动显著减少及HDL-胆固醇水平降低 (Oram 等人, *J. Lipid Res.*, 37: 2743-2491 (1996); Francis 等人, *J. Clin. Invest.*, 96: 78-87 (1995))。患有丹尼尔病的家族的遗传连锁分析将缺陷基因分配到染色体9q31的间隔部位 (Rust 等人, *Nature Genetics*, 20: 96-98 (1998))。公共数据库检索发现ABC1定位于染色体9q22-9q31，这比Rust 等人中公开的间隔部位更宽，但是包括Rust 等人中公开的间隔部位 (Luciani 等人, *supra* (1994))。以这些数据为基础，制作了人ABC1基因的辐射杂合图谱 (radiation hybrid mapping)，该图谱将该基因置于两个marker之间，所述两marker成直角地位于Rust 等人报道的人染色体9q31的7-cM区域内。此外，如实施例2所示，微排列分析显示，丹尼尔患者细胞内ABC1基因的表达量在正常细胞的2.5倍以下。这些研究将丹尼尔病中的该缺陷基因鉴定为ABC1。此外，本发明进一步研究将ABC1活性与胆固醇排放活动联系起来。首先，研究表明ABC1运输活性的抑制剂，例如 4,4-diisothiocyanostilbene acid (DIDS) 和 sulphobromophthaleine (BSP)，也抑制成纤维细胞的apoAI-介导胆固醇排放 (见实施例6)。另外，用反义ABC1寡核苷酸抑制ABC1基因表达也显示能够抑制成纤维细胞的apoAI-介导胆固醇排放 (见实施例7)。相反，将ABC1基因转染

入鼠单核细胞的转染试验表明ABC1的过量表达导致apoAI-介导排放的增加(实施例8)。最后,用野生型和丹尼尔患者mRNA进行的RT-PCR表明:正常的皮肤成纤维细胞中ABC1 mRNA 的表达受与胆固醇排放相关的细胞条件的调控,但是在丹尼尔患者的成纤维细胞中这种调控不存在(实施例9)。这些发现确证了ABC1在胆固醇排放中发挥着主要的作用。

可以得出这样的结论:ABC1在胞内胆固醇向胞膜的外周小叶易位的过程发挥作用。由于ABC1缺乏或缺陷性ABC1导致的胞内胆固醇运输缺陷会引发特定膜结构域内胆固醇的缺乏,该结构域与apoAI和其他载脂蛋白特异性相互作用(Stangl 等人, *J. Biol. Chem.*, 273: 31002-31008 (1998); Babitt 等人, *J. Biol. Chem.*, 272: 13242-13249 (1997))。胆固醇无法递送到apoAI导致生成胆固醇-缺陷的HDL颗粒,这样的颗粒被迅速清除出胞浆(Bojanovski 等人, *J. Clin. Invest.*, 80: 1742-1747 (1987))。

## 定义

提供下列定义是为了有助于本说明书全文使用的某些术语的理解。

在本发明中,“分离”是指物质离开初始环境(例如,如果是天然生成即为自然环境),从而“可操作”地改变了其天然状态。例如,分离的多核苷酸可以是载体或组合物的一部分,或者可以取自细胞,并且仍然是“分离”,因为载体、组合物或具体细胞都并非该多核苷酸的初始环境。

本发明中使用的“多核苷酸”包括合成和天然来源的DNA和RNA。所述多核苷酸可以以单链-、双链-DNA或RNA、或者RNA/DNA杂合体的形式存在。因此,本发明所述的多核苷酸可由任一聚核糖核苷酸或聚脱氧核糖核苷酸组成,其可以是未经修饰的RNA 或DNA或者修饰后的RNA或DNA。例如,多核苷酸可以由下述成分组成:单链及双链DNA,包括单链和双链区域、或单链、双链和三链区域的DNA混合物,单链及双链RNA,包括单链和双链区域的RNA混合物,包括DNA和RNA的杂合分子,该DNA和RNA可以是单链或更

常见为双链或三链或者单链和双链区域的混合。此外，所述多核苷酸可以由包含RNA或DNA或这三者的三链区域组成。为了稳定或其他原因，多核苷酸可以包括一个或多个修饰后的碱基或者修饰后的DNA或RNA骨架。“修饰后的”碱基包括，例如，三苯甲基化碱基及次黄嘌呤核甙等非常见碱基。可以对DNA和RNA进行多种修饰；因而，“多核苷酸”包括多核苷酸的化学、酶学或代谢性修饰形式。“多核苷酸”还可包括通常称为寡核苷酸的短的多核苷酸。

“多肽”一词是指包括相互间通过肽键或修饰后肽键相连的两个或多个氨基酸。”多肽”是指通常称为肽的短氨基酸序列以及通常称为蛋白质的长氨基酸序列。多肽可以包含20个基因编码氨基酸以外的氨基酸。另外，多肽的修饰可以通过天然过程、例如处理及其他翻译后修饰，或者通过本领域已知的化学修饰技术来实现。一种给定的多肽可以包括多种类型的修饰。另外，同种类型的修饰可以在多肽的一个或多个位点以同等或不同程度出现。修饰可以发生在多肽的任一部分，包括多肽骨架、氨基酸侧链、以及氨基或羧基端。修饰包括，但不限于乙酰化、酰化、ADP-核糖基化、酰胺化、达夫反应甲酰化、 $\gamma$ -羧基化、糖基化、羟基化、碘化、甲基化、myristoylation、氧化、磷酸化、prenylation、硫酸化 和selenoylation，以及核苷酸或核苷酸衍生物、脂类和脂类衍生物，或者phosphatidylinositol的共价连接。其他修饰包括交联、环化、生成pyroglutamate、生成GPI锚定、蛋白水解、外消旋作用、和t-RNA-介导的氨基酸添加，例如精氨酸化及泛素化。参见，例如，蛋白质-结构及分子特性，第二版，T. E. Creighton, W. H. Freedman and Co., New York (1993); Wold, F., 翻译后蛋白修饰: 透视和展望，蛋白质的翻译后共价修饰，B. C. Johnson, Ed., Academic Press, New York (1983); Seifter 等人, Meth. Enzymol., 182: 626-646 (1990); and Rattan 等人, 蛋白合成: 翻译后修饰及老化, Ann. N. Y. Acad. Sci., 663: 48-62 (1992)). 本发明所述多肽可以任一合适

的方法制备。这类多肽可以包括：分离后的天然生成多肽、重组制备的多肽、合成多肽、或者这些方法联用制备的多肽。制备这类多肽的方法已为本领域技术人员熟知。

本发明中的“多核苷酸”还可包括在严格条件下能够与SEQ ID NO: 3或  
5 SEQ ID NO: 3中第1-1532,1080-1643,1181-1643,1292-1643,1394-1643，或  
1394-1532位核苷酸杂交的多核苷酸或其互补链。本发明所述多肽还包括能够在严格条件下与SEQ ID NO: 4、SEQ ID NO: 5或SEQ ID NO: 6杂交的多核苷酸，或其互补链。

“严格条件”是指在下述溶液中42℃孵育过夜，该溶液中包括50%甲酰胺、  
10 5x SSC (750 mM NaCl, 75 mM 柠檬酸钠), 50 mM 磷酸钠 (pH 7.6)、5x  
Denhardt's溶液、10% dextran sulfate和20μg/ml 变性剪切鲑鱼精DNA，然后在约65℃下用0.1x SSC洗涤滤膜。

本发明使用的“互补”是指核苷酸间的杂合或碱基配对，例如双链多核苷酸的两条链之间或者寡核苷酸引物与待扩增或待测序的单链多核苷酸之  
15 间。当通过适当核苷酸插入、缺失或取代使两个单链核苷酸分子实现最大限度对齐时，如果其中一条链上的核苷酸可与另一条链上的约80%的核苷酸配对，那么就可以认为这两个核苷酸分子是互补的。

“同源性”在本领域中是指两个或多个多肽序列之间或者两个或多个多核苷酸序列之间的相互关系，通过比较两序列测定而出。

20 “同源性”或“相似性”还有本领域技术人员认可含义，即指多肽或多核苷酸序列之间的相关程度，通过将所述序列的链配对测定出来。“同源性”和“相似性”可以用大量众所周知的方法测算出来，其中包括下列文献中公开的方法：计算分子生物学，Lesk, A. M., ed., Oxford University Press, New York, (1988); 生物计算：信息学及基因组计划，Smith, D. W., ed.,  
25 Academic Press, New York, (1993); 序列数据的电脑分析，Part I, Griffin,

A. M., and Griffin, H. G., eds., Humana Press, New Jersey, (1994); 分子生物学序列分析, von Heinje, G., Academic Press, (1987); 和引物序列分析, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, (1991); 以及 Carillo, H., and Lipton, D., SIAM J Applied Math 48: 1073 (1988))。

- 5 测定两序列间同源性或相似性的常用方法包括但不限于下述文献中公开的方法: “巨型计算机导论”, Martin J. Bishop, ed., Academic Press, San Diego, (1994), and Carillo, H., and Lipton, D., SIAM J Applied Math 48: 1073 (1988)。测定同源性的优选方法是为待测序列间最大匹配而设计的。用于多核苷酸或多肽对齐的方法被编制成电脑程序, 其中包括GCG 程序包 (Devereux, J.,
- 10 等人, Nucleic Acids Research (1984) 12 (1): 387 (1984))、BLASTP、BLASTN、FASTA (Atschul, S. F. 等人, J. Molec. Biol. 215: 403 (1990)、Bestfit 程序 (Wisconsin 序列分析包、Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711 (using the
- 15 local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2: 482-489 (1981))。

例如, 当用任一序列对齐程序测定具体序列是否90%同源于参照序列, 建立参数以便计算出参照多肽或多核苷酸全序列的相同百分比, 允许参照多核苷酸中核苷酸总数10%同源性缺口。

- 用于测定查询序列 (本发明所述序列) 和主题序列之间最大程度匹配的
- 20 优选方法, 也称整体序列对齐, 可以用以Brutlag 等人 (Comp. App. Biosci. 6: 237-245 (1990))中算法为基础的FASTDB电脑程序测定。“序列”一词包括核苷酸和氨基酸序列。在序列对齐中, 查询序列和主题序列要么都是核苷酸序列要么都是氨基酸序列。上述整体序列对齐的结果为同源性百分比。用FASTDB 查究 DNA 序列计算同源性百分比使用的优选参数为:
- 25 Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30,

Randomization Group Length=0, 及 Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, 以及 Window Size=500 或 核苷酸碱基中的查询序列长度, 无论那个较短。氨基酸对齐中计算同源性及类似性百分比使用的参数为: Matrix=PAM 150, k-tuple=2, Mismatch Penalty= 1, Joining Penalty=20.

- 5 Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty=0.05. 及 Window Size=500 or 氨基酸残基中的查询序列长度, 无论那个较短。

试举一例, 核苷酸序列与SEQ ID NO: 1中所含序列至少90% “同源”的多核苷酸是指, 除了每总长100个核苷酸中至多可允许10个点突变外, 该多核苷酸的核苷酸序列与SEQ ID NO: 1中包含的序列相同。换句话说, 为了获得核苷酸序列与SEQ ID NO: 1至少90%同源的多核苷酸, 至多允许SEQ ID NO: 1所含序列核苷酸的10%被缺失、插入或被其他核苷酸取代。这些改变可以发生在该多核苷酸的任一部位, 也可以单个散布在核苷酸中或者发生在SEQ ID NO: 1中一个或多个连续族中。

- 15 类似地, 氨基酸序列与SEQ ID NO: 2中所含序列至少98% “同源”的多肽是指, 除了可以在每总长100个氨基酸中至多可允许2个点突变外, 该多肽的氨基酸序列与SEQ ID NO: 2中包含的序列相同。换句话说, 为了获得氨基酸序列与SEQ ID NO: 2至少98%同源的多肽, 至多允许SEQ ID NO: 2中所含氨基酸残基的2%被缺失、插入或被其他核苷酸取代。这些改变可以发生在所述多肽的任一部位, 也可以单个散布在残基中或者发生在SEQ ID NO: 2中一个或多个连续族中。

“具有生物活性的多肽”是指该多肽活性与本发明所述多肽活性(例如胆固醇运输活性)类似, 但不必相同, 可用一特定生物学测定试验测算, 具有或无剂量依赖性。当存在剂量依赖性时, 与本发明所述多肽相比, 就给定活性而言, 多肽活性不必相同, 而只需剂量依赖性地基本近似(即, 侯选多

肽与本发明所述多肽相比具有更强活性，或者减弱程度小于25-倍，优选地，活性减弱小于10倍，以及最优选活性减弱小于3倍)。

“多肽变体”是指虽有别于本发明所述的ABC多肽但是基本保留了其特性的多肽。通常，变体与含SEQ ID NO: 2的多肽整体上非常近似，甚至多个部位上相同。优选地，所述多肽变体保留了其生物活性，即胆固醇运输活性。变体包括，但不限于，拼接变体及等位突变体，以及添加、缺失及取代突变体。

类似地，“多核苷酸变体”是指虽有别于本发明所述的多核苷酸但是基本保留了其基本特征的多核苷酸。所述变体可以包括编码区域、非编码区域或者二者均有的改变。因此，例如，ABC1多核苷酸变体的核苷酸序列虽然不同于SEQ ID NO:1,但是其编码的多肽仍具有胆固醇运输活性。另外，例如，一多核苷酸变体的核苷酸序列不同于SEQ ID NO: 3，但仍保留了启动子活性。特别有的多核苷酸变体是其中的改变仅产生沉默取代、添加或缺失而不改变其编码的多肽的特性或活性。优选由于遗传密码子简并用沉默取代中制备的核苷酸变体。另外，其中10-20个、5-10个、1-5个或1-2个氨基酸被任一组合的取代、添加或缺失的突变体也是优选的。可以因多种原因而制备多核苷酸变体，例如为适应特定宿主而优化密码子表达 (例如将任mRNA中的密码子改变为大肠杆菌等细菌宿主优选的密码子)。

“等位突变体”是指这样的天然突变体，即位于生物染色体一给定座位上的基因的多个不同形式的其中之一 (Genes II, Lewin, B., 编著 John Wiley & Sons, New York (1985).) 这些等位突变体可以在多核苷酸和/或多肽水平上变化。替代地，非天然生成的变体可以用诱变技术或直接合成来制备。

术语“保守的氨基酸取代”是指用标准的氨基酸残基取代天然的氨基酸残基，以保证对该位置上的氨基酸残基的极性或电荷几乎或根本没有影响。例如，用任一非极性残基代替多肽中的一非极性残基而发生的保守性取代。

保守性取代的另一实例是用另一酸性残基代替一酸性残基。保守取代产生的ABC1多肽，其功能及化学特性类似于天然生成的ABC1多肽。

“直向同源”是指与从不同种中鉴定出的多肽所对应的多肽。例如，鼠和人的ABC1多肽就可认为是直向同源。

5 本发明使用的“载体”是指用来将编码信息传输给宿主细胞的任一分子(例如，核酸、质粒或病毒)。

“表达载体”是指适于转化宿主细胞的载体，其中包括能够指导和/或调控所插入的异源核酸序列的表达。表达包括，但不限于，转录、翻译以及有内含子存在时的RNA拼接。

10 本发明使用的“转录调控区域”或“表达调控部分”是指基因的任一区域，包括但不限于启动子、增强子及阻遏物。

本发明使用的“启动子”是指位于结构基因起始密码子上游(即5'端)的非转录序列(通常约100~1000 bp)，该序列可调控结构基因的转录。

15 本发明使用的“增强子”是指DNA的顺式-作用单元，通常长约10-300 bp，它可以作用于启动子增强转录。增强子的取向和位置相对独立。它们在转录单元的5'和3'端都已发现。

“宿主细胞”是指已被转化或转染的细胞，或者能够被外源多核苷酸序列转化或转染的细胞。该术语包括亲代细胞的子代，无论子代在形态或遗传构成上与原来的亲代是否相同，只要所选基因还在就可。

20 本发明使用的“可操作连接”是指侧翼序列的重排，其中所述侧翼序列被装配或组装后仍能行使其常规功能。因此，操作连接到编码序列上的侧翼序列能启动编码基因的复制、转录和/或翻译。例如，当将编码序列与启动子操作连接时，所述启动子能够指导该编码序列的转录。侧翼序列不必与编码序列邻接，只要其正确发挥功能即可。因此，例如，插入的转录但非翻译序列  
25 可以位于启动子和编码序列之间，这时所述启动子序列仍可认为是“可操作

连接”到编码序列上。

本发明使用的“转染”是指外来或外源DNA被细胞摄入的过程，以及当外源DNA被导入细胞膜内时就可以说一个细胞已被“转染”。大量的转染技术已为本领域技术人员所熟知。参见，例如，Graham 等人，1973，*Virology* 52: 456; Sambrook 等人，分子克隆，实验室手册 (Cold Spring Harbor Laboratories, 1989); Davis 等人，分子生物学基本方法 (Elsevier, 1986); and Chu 等人，1981，*Gene* 13: 197。这些技术可以将一个或多个外源DNA基因导入到合适的宿主细胞中。

### **ABC1 多肽**

10 本发明涉及新的人ABC1 多肽。一个实施方案中，所述 ABC1多肽包含 SEQ ID NO: 2中所示的氨基酸序列。与其他人报道的人ABC1蛋白不同，SEQ ID NO: 2所示的ABC1多肽在氨基末端添加了另外的60个氨基酸，从而使 ABC1 蛋白包括2261个氨基酸而非2201个氨基酸 (参见，Langmann 等人 in *Biochem. Biophys. Res. Comm.*, 257,29-33 (1999))。另外，SEQ ID NO: 2 所  
15 示的ABC1多肽与其他报道过的序列有几个氨基酸残基不同。具体地，本发明 SEQ ID NO: 2 所示的ABC1多肽发生了下述氨基酸残基的取代：第159位残基 K取代了R，765位I取代了V，823位M取代了I，1495位I取代了T，1588位L取代了P，1914位K取代了R，以及2018位L取代了P。为了与公开出版的记数法一致，上述氨基酸数字采用Lawn 等人，*J. Clin. Invest.*, 104: R25-31 (1999)，  
20 而非SEQ ID NO: 2所示的数字。如下面进一步讨论的那样，所述的序列差异可能起因于这样一个事实：前一个ABC1 cDNA用以PCR为基础的策略克隆自小鼠，而随后报道的人ABC1 cDNA 序列是从小鼠蛋白质序列中推导出来的。经SDS-PAGE 测定，所述ABC1 蛋白的分子量约为240 kD。

本发明还涉及这样的ABC1多肽，其中包含的氨基酸序列优选与SEQ ID  
25 NO: 2全长氨基酸序列至少98%同源，更优选与SEQ ID NO: 2全长氨基酸序列

至少99%同源,最优选,所述多肽与SEQ ID NO: 2全长氨基酸序列100%同源。如前所述,“同源性”是指多肽序列之间的序列相关程度,下文将进一步描述。

5 此类相关的ABC1多肽包括取代、缺失及插入变体,以及等位突变体、拼接变体、片段、衍生物或直源同系物。优选的多肽或多肽片段包括那些具有ABC1生物活性的多肽及片段。具体而言,优选能介导反向胆固醇运输活性的多肽及片段。具有改良反向胆固醇运输活性的多肽及片段也是优选的。

取代、缺失及插入变体是指包含这样氨基酸序列的ABC1多肽,即与SEQ ID NO: 2中所示的氨基酸序列相比,所述氨基酸序列中包含一处或多处氨基  
10 酸序列的取代、缺失和/或添加。优选实施方案中,所述变体有约1~5、或约1~10、或约1~20、或约1~40、或约1~65个氨基酸的取代、添加和/或缺失。例如,所述变体可以在多肽的任一部位以及羧基端和/或氨基端添加1个或多个氨基酸残基,只要所述变体仍保留生物功能即可。另外,例如,可以从包括羧基端或/或氨基端在内的多肽任一区域缺失1个或多个氨基酸,而生物功能  
15 基本保留 (Ron 等人, J. Biol. Chem., 268 2984-29842988 (1993), Dobeli 等人, J. Biotechnology, 7: 199-216 (1988))。氨基酸取代可以是保守性的,也可以是非保守性的,或者二者任一组合,只要ABC1保留其生物活性即可。此外,所述取代也可以用非保守性氨基酸残基,所取代残基可以是也可以不是遗传密码子编码的氨基酸,另外也可用带有取代基团的氨基酸残基。

20 ABC1多肽的合适变体可以用众所周知的技术来测定。例如,可以通过鉴定ABC1分子中的可改变但不破坏其生物活性的区域对合适的ABC1变体进行测定。另外,如本领域认可的那样,甚至对于生物活性或结构而言特别重要的区域也可进行保守性的氨基酸取代,而不会破坏其生物活性或者给多肽结构带来不良影响。通过鉴定ABC1多肽中的那些区域对于其活性而言是无关紧  
25 要的,就可以测定出可改变而不破坏生物活性的氨基酸残基。(Bowie 等人,

Science, 247: 1306-1310 (1990))。例如，通过比较来自不同物种的ABC1 多肽，可以测定出ABC1分子的种间保守的氨基酸残基及区域。保守的氨基酸残基可能对于生物功能和/或结构而言是重要的。反过来讲，对于ABC1分子中种间非保守因而被自然选择容许的区域而言，其中的改变几乎不可能会影响生物活性和/或结构。因此，非保守区域内发生了添加、缺失或取代的ABC1多肽应该是合适的变体。甚至在相对保守的区域内，可用化学上近似的氨基酸取代天然生成残基而保留活性。

此外，可以通过下述方法鉴定出合适的ABC1变体：用结构-功能试验测定出ATP-结合框家族其他成员，例如ABCR和ABC-C中对于活性或结构而言重要的残基。这类试验可以通过与其他ATP-结合框蛋白中对于活性或结构而言重要的氨基酸残基对应的方式，预测出ABC1变体中的重要氨基酸残基。例如，在其他ATP-结合框蛋白的结构-功能试验的基础上，能够在与核苷酸结合及胆固醇运输相关的区域内发现ABC1中的重要氨基酸残基。合适的变体包括，例如，其中用上述方法预测到的ABC1多肽中重要氨基酸残基被化学上类似氨基酸取代的多肽。

合适的ABC1变体也可以用下述方法测定：为了鉴定出对于多肽功能至关重要的区域，用基因工程技术在特定位置导入氨基酸改变。氨基酸改变可以用例如定点诱变或丙氨酸-扫描诱变方法实现 (Cunningham 等人, Science, 244: 1081-1085 (1989))。然后用例如任一本发明上述胆固醇排放测定试验，对得到的ABC1变体的生物学活性进行测试。特定氨基酸残基取代破坏了胆固醇排放活性，这样的变体不是合适的ABC1 变体。

其他鉴定合适变体的方法也为本领域已知。另外，本领域技术人员应当能识别出蛋白特定氨基酸位置上可容许的氨基酸改变 (Bowie 等人, supra (1990))。例如，通常已知大部分被包埋或内部 (位于蛋白的三级结构内)氨基酸残基需要非极性侧链，而表面或外部侧链通常极少有此特征。而且， 已知

容许的保守氨基酸取代包括下列取代：脂肪族或疏水氨基酸Ala、Val、Leu和Ile的取代、羟基残基Ser和Thr的取代、酸性残基Asp和Glu的取代、酰胺残基Asn和Gln的取代、碱性残基Lys、Arg和His的取代、芳香族残基Phe、Tyr和Trp的取代以及小型氨基酸Ala、Ser、Thr、Met及Gly的取代。

5           ABC1变体可以天然生成或者人工构建。天然生成的变体的实例有等位突变体和拼接变体。等位突变体是指即位于生物体或生物群染色体一给定座位上的基因的多种不同形式的其中之一 (Lewin, B., ed., Genes II, John Wiley & Sons, New York (1985))。等位突变体可以在多核苷酸和/或多肽水平上变化。拼接变体是指一核酸分子，通常为RNA，通过对RNA转录本或相应多肽  
10  中内含子序列进行选择处理而生成的。或者，可以人工构建ABC1变体。例如，可以用定点诱变技术构建ABC1变体。另外，例如，ABC1变体也可以从编码所述变体的相应核酸分子制备得到，该变体的DNA序列从SEQ ID NO: 1所示的野生型DNA序列变化而来的。

          多肽片段是指其氨基酸序列短于SEQ ID NO: 2所示的氨基酸全序列。优  
15  选的多肽片段包括具有ABC1生物活性的片段。具体而言，是那些介导反向胆固醇运输或者具有改良反向胆固醇运输活性的片段也是优选的。ABC1 片段可以是所述多肽任一区域缺失1个或多个氨基酸而得到的，包括羧基端和/或氨基端，只要生物功能保持不变。ABC1多肽片段可以天然生成，例如通过选择性拼接或体内蛋白酶活性，或者用已知方法人工构建。

20           本发明还涉及ABC1多肽衍生物，所述衍生物是指本发明所述的化学方法修饰得到的ABC1 多肽、变体、或片段。 所述衍生物是用不同于天然生成ABC1多肽的方式制备而成的，即与该多肽结合的分子的类型或部位不同。衍生物还可以包括通过天然连接到ABC1多肽上的1个或多个基团缺失形成的多肽。此外，可以将包含SEQ ID NO: 2所示氨基酸序列的ABC1多肽，以及上述  
25  的ABC1变体和片段，与同源多肽融合生成同源二聚体，或者与一异源多肽融

合生成异源二聚体。

本发明的另一方面涉及突变的ABC1多肽及其片段，对应于分离自丹尼尔患者的多肽。在一个优选实施方案中，所述ABC1多肽包含SEQ ID NO: 8。从丹尼尔患者 (TD1)分离出该蛋白质，并按照实施例1和5所述测序。SEQ ID NO: 8所示氨基酸序列与野生型序列相似，不同之处为第537位 (残基序号同文献Lawn 等人, J. Clin. Invest., 104: r25-31 (1999), 对应于SEQ ID NO: 8中的第597位)上的谷氨酰胺取代成为精氨酸。该残基位于氨基末端亲水结构域内，靠近第1个推测的跨膜结构域。所述取代改变了该蛋白所述区域的氨基酸的电荷，生成的ABC1蛋白其胆固醇排放活性明显减弱，如图1所示。

10 另一个优选实施方案中，所述ABC1多肽包含SEQ ID NO: 10。从丹尼尔患者 (TD2)分离出该蛋白质，并按照实施例1和5所述测序。SEQ ID NO: 10所示氨基酸序列与野生型序列相似，不同之处为第527位 (残基序号同文献Lawn 等人, sufi-a (1999), 对应于SEQ ID NO: 10中的第587位)上的色氨酸被精氨酸所取代。与TD1多肽类似，该取代改变了ABC1蛋白氨基末端亲水结构域中氨基酸残基的电荷。得到的突变ABC1蛋白的胆固醇排放活性也明显  
15 减弱，如图1所示。

### **ABC1 多核苷酸**

本发明另一方面涉及到分离的编码新的ABC1多肽的多核苷酸和其变体。例如，本发明提供分离的编码全长的ABC1多肽的多核苷酸，包含野生型ABC1全长的cDNA的多核苷酸、包含全长的野生型ABC1的  
20 编码序列，和包含ABC1的非编码5'和3'序列的多核苷酸，以及有关ABC1变体的多核苷酸。本发明也提供分离的编码突变型ABC1多肽的多核苷酸，例如丹尼尔患者的那些。

在一项优选的实施方案中，分离的多核苷酸包括编码包含SEQ ID  
25 NO:2的多肽的核苷酸序列。重要地，与Langmann等人已公开的序列

(Langmann 等人, *Biochem. Biophys. Res. Comm.*, 257,29-33 (1999) (GenBank 登记号 AJ012376)对照, 该序列编码基于在外显子 3 发现的推测的起始甲硫氨酸的 2201 个氨基酸蛋白, 目前要求保护的核苷酸序列包含 50 个外显子和 2261 个氨基酸蛋白的密码 (参见图 4)。本发明相应的核苷酸序列包含一种编码序列, 它包括在对应于下列 60 个氨基末端的氨基酸的 5'末端额外的 180 个核苷酸: MACWPQLRLLLWKNLTFRRRQTCQL LLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA。

假设从此位置上游存在框内终止密码子 6 至 9 个核苷酸, 最近预测起始位点是能产生连续开放阅读框架的第一甲硫氨酸密码子。这种具有相关的 ABC 转运蛋白序列 ABCR 和 ABC-C (也称为 ABC3)的新 ABC1 cDNA 的排列显示出高度的相似性, 该排列亦包含额外的 60 个氨基酸的开放阅读框, 这意味着同源的 ABC 转运蛋白开始于与为人 ABC1 建议的氨基末端延伸序列相关的序列。很可能早期公开的人 ABC1 起始位点是以已公开的小鼠 ABC1 cDNA 序列 (Luciani 等人, *Genomics*, 21150-159 (1994); GenBank 登记号: X75926)来预测的, 该序列在延伸区包含额外的核苷酸 "n", 这样以致于最近公开的甲硫氨酸不在框内。然而, 如果忽略小鼠序列中的 "n"核苷酸, 小鼠和人的延伸区序列是相同的。根据这些结果, 全长人 ABC1 蛋白包含 2261 个氨基酸是很可能的, 而不是如 Langmann 等人和其它人先前提出的 2201 个氨基酸。因此, Langmann 等人未提出完全的人 ABC1 的开放阅读框架。

另一优选的实施方案中, 分离的多核苷酸包括全长的 ABC1 cDNA, 它包含至少部分的或者非编码的 5'和 3'序列。优选地, 这种多肽包含 SEQ ID NO: 1 所示的核苷酸序列。SEQ ID NO: 1 显示的 10.4kb 人 ABC1 cDNA 序列包含 6783 个核苷酸加 5'和 3'未翻译区的开放阅读框架。起始密码子位于位置 291, 终止密码子位于位置 7047。SEQ ID NO: 1 表示的本发明

的 ABC1 cDNA 在几个方面不同于已公开的 ABC1 cDNA (GenBank 登记号 AJO12376)。首先, 本发明 ABC1 cDNA 包括 5'端额外的 350 核苷酸和 3'端额外的 3136 核苷酸 (不包括 poly (A)尾)。本发明 ABC1 序列也不同于已公开的在编码区 10 个核苷酸取代的 ABC1 cDNA。十个区别当中, 7 个核苷酸差异引起氨基酸变化。为了与已公开的符号保持一致, 下列核苷酸和氨基酸号码是 Lawn 等人, *J. Clin. Invest.*, 104: R25-31 (1999)和 GenBank 登记号 AJO12376 的那些, 而不是 SEQ ID NO: 1 的那些。核苷酸和氨基酸变化如下: (1)在核苷酸 414, A 替 G; (2)在核苷酸 596, A 替 G (在氨基酸 159, K 替 R); (3)在核苷酸 705, T 替 C; (4)在核苷酸 1980, A 替 C; (5)在 2413, A 替 G (在氨基酸 765, I 替 V); (6)在 2589, G 替 A (在氨基酸 823, M 替 I); (7)在 4604, T 替 C (在氨基酸 1495, I 替 T); (8)在 4883, T 替 C (在氨基酸 1588, L 替 P); (9)在 5861, A 替 G (在氨基酸 1914, K 替 R); 和 (10)在 6443, T 替 C (在氨基酸 2108, L 替 P)。这些氨基酸变化中的五种是保守的氨基酸变化并且可能表示多态性或序列错误。在两种情况下, 本序列推测不同于 GenBank 序列的重要的氨基酸差异。差异导致在残基 1588 和残基 2108 是亮氨酸而不是脯氨酸。有趣地, 在两个位置, 在分析的三个 TD 样品中的每一个以及非常保守的小鼠 ABC1 蛋白序列, 亦发现推测的亮氨酸。

本发明也涉及到包含核苷酸序列的 ABC1 多核苷酸, 该核苷酸序列在它们整个长度内, 优选具有与包含 SEQ ID NO: 1 的多核苷酸至少 80% 的同源性。更优选地, 本核苷酸在其完全长度内具有与包含 SEQ ID NO: 1 的多核苷酸至少 90% 的同源性。甚至更优选地, 多核苷酸在其整个长度内具有与包含 SEQ ID NO: 1 的多核苷酸至少 95% 的同源性。最优选地, 多核苷酸在其整个长度内具有与包含 SEQ ID NO: 1 的多核苷酸 100% 的同源性。这种涉及的 ABC1 多核苷酸包括取代, 缺失, 和插入变体, 以

及等位变体，剪接变体，片段，衍生物，和直向同源体 (orthologs)，其中已经取代、缺失、插入或衍生一个或多个核苷酸。优选的多核苷酸包括那些编码具有生物学活性的 ABC1 多肽和多核苷酸，例如胆固醇排放活性。

5           另一优选的实施方案中，分离的多核苷酸包括 ABC1 全部编码序列。在特别优选的实施方案中，多核苷酸包括如 SEQ ID NO: 1 的 291-7074 核苷酸所示的序列。这种分离的多核苷酸包含 6783 个核苷酸的 ABC1 开放阅读框架并且编码 2261 个氨基酸的多肽，如上所述。

          又一项优选的实施方案中，分离的多核苷酸包括编码 ABC1 变体多肽  
10   的核苷酸序列。尤其是，分离的多核苷酸包括编码包含氨基酸序列的多肽的核苷酸序列，该氨基酸序列至少 98% 同源于 SEQ ID NO: 2 的氨基酸序列。也优选含有编码多肽的核苷酸序列的分离的多核苷酸，该多肽含有与 SEQ ID NO: 2 氨基酸序列至少 99% 同源的氨基酸序列。因此，本发明包括那些编码上述 ABC1 多肽的多核苷酸，包括介绍过的取代、缺失和插  
15   入变体，以及 ABC1 等位变体，剪接变体，片段，衍生物，融合多肽，和直向同源体 (orthologs)。优选的多核苷酸是编码具有 ABC1 生物活性的多肽的那些多核苷酸。尤其优选的是编码介导逆向胆固醇转运的那些多肽的多核苷酸。编码具有改善的逆向胆固醇转运活性的多肽的多核苷酸也是优选的。

20           本发明的又一方面涉及到分离的编码来自丹尼尔患者的突变型 ABC1 多肽的多核苷酸。一项优选的实施方案中，多核苷酸编码 SEQ ID NO: 8 多肽，该多肽自患者 TD1 分离并如上所述。另一项优选的实施方案中，多核苷酸包括 SEQ ID NO: 7 提出的核苷酸序列。SEQ ID NO: 7 提出的核苷酸序列包含全部开放阅读框架，以及 5' 和 3' 侧翼序列。可译框架  
25   编码 2261 个氨基酸的多肽，在其它取代之间，该肽包含导致位置 537 的

A 至 G 取代的核苷酸取代 (采用 Lawn 等人的编号, *supra* (1999))。

涉及到编码突变型 ABC1 多肽的多核苷酸的另一项优选的实施方案中, 多核苷酸编码 SEQ ID NO: 10 多肽, 该多肽自丹尼尔病人 TD2 分离并也在上文介绍过。又一项优选的实施方案中, 多核苷酸包括 SEQ ID NO: 9 提出的核苷酸序列。SEQ ID NO: 9 提出的核苷酸序列包含全部开放阅读框架, 以及 5'和 3'侧翼序列。开放阅读框架编码 2261 个氨基酸多肽, 在其它的取代之间, 包含导致残基 527 的 Arg 至 Tryp 取代的多核苷酸取代 (采用 Lawn 等人的编号, *supra* (1999))。

本发明的另一方面涉及到分离的包括 ABC1 非编码 5'侧翼区和 3'侧翼区多核苷酸。一项实施方案中, 分离的多核苷酸包括 ABC1 非编码 5'侧翼区。优选地, 5'侧翼区包含, 但是并非限制, ABC1 启动子区。从而, 在一项优选的实施方案中, 多核苷酸包括 SEQ ID NO: 3 显示的序列。如通过异源报道基因证实, 实施例 15 讨论的, SEQ ID NO: 3 提出的多核苷酸包含 AB1 基因的转录调控区。如图 13 所示, SEQ ID NO: 3 提出的多核苷酸是包含几个转录调控因子的 1643b.p.非编码区, 包括位置 1522、1435 和 1383 的 TATA 框, 以及转录因子结合位点, 包括几个推定的 SP1 位点, 和几个核受体半位点。此外, 在位置 1483-1500 发现识别的固醇反应元素。进一步地, 实施例 17 进行的进一步的异源报道基因试验揭示了保留启动子活性的 SEQ ID NO: 3 的几个不连续部分。从而, 另一项优选的实施方案, 多核苷酸包含 SEQ ID NO: 3 的 1-1532, 1080-1643, 1181-1643, 1292-1643, 或 1394-1643 核苷酸。在一项特别优选的实施方案中, 这种多核苷酸包括 SEQ ID NO: 3 的 1394-1532 核苷酸, 该序列已经显示具有 ABC1 启动子活性 (参见实施例 17)。又一项优选的实施方案中, 这种多核苷酸包括 SEQ ID NO: 3 的 1480-1510 核苷酸, 它显示出调节 ABC1 对 LXR 配体的转录反应。

5'侧翼多核苷酸也包括在严格条件下与 SEQ ID NO: 3 中提出的核苷酸序列杂交的核苷酸序列，其中这种核苷酸序列具有 ABC1 启动子活性。又一项实施方案中，这种多核苷酸包括在严格条件下与包含 SEQ ID NO: 3 的 1-1532, 1080-1643, 1181-1643, 1292-1643, 或 1394-1643 核苷酸杂交的核苷酸序列，其中该核苷酸序列具有 ABC1 启动子活性。

又一项实施方案中，分离的多核苷酸包括 ABC1 3'侧翼区。已经鉴别几种 3'未翻译区，它可以表示 ABC1 转录聚腺苷酸化的可变位点。优选地，3'侧翼区包含调节序列。例如全长的 3'UTR (SEQ ID NO: 6) 包含 46 个序列 (AA)nCU/UC (AA)n，它已经显示对于 Vigilin 的结合是必需的。Vigilin，一种具有 14K 同源域的普遍存在的蛋白，是雌激素-诱导的卵黄生成素 mRNA 3'-未翻译区粘合蛋白 (*J. Biol. Chem.*, 272: 12249-12252 (1997))。除了结合 HDL 之外，Vigilin 已经显示出结合 mRNAs 的 3'侧翼区，并提高 mRNA 转录物的半衰期 (*Mol. Cell. Biol.*, 18: 3991-4003 (1998))。因此，可以改变 3'侧翼区，例如，增加 Vigilin 的结合，因此增加 ABC1 mRNA 的半衰期。优选地，分离的多核苷酸包括 SEQ ID NO: 4 显示的序列。同样优选地，分离的多核苷酸包括 SEQ ID NO: 5 显示的序列。另一优选的实施方案中，分离的多核苷酸包括 SEQ ID NO: 6 显示的序列。在其它优选的实施方案中，多核苷酸包括在严格条件下与 SEQ ID NO: 4, SEQ ID NO: 5 或 SEQ ID NO: 6 提出的核苷酸序列杂交的序列。

本发明也包括有关的 ABC1 5'和 3'侧翼多核苷酸。因此，本发明涉及到包括核苷酸序列的多核苷酸，该核苷酸序列在其完全长度内具有与包含 SEQ ID NO: 3 的多核苷酸、包含 SEQ ID NO: 4 的多核苷酸、包含 SEQ ID NO: 5 的多核苷酸、包含 SEQ ID NO: 6 的多核苷酸，或者包含 SEQ ID NO: 3 的 1-1532, 1080-1643, 1181-1643, 1292-1643, 或 1394-1643 核苷酸的多核苷酸至少 80%的同源性。优选地，这种多核苷酸在其整个长度内对

于任何一种前述的侧翼多核苷酸，具有至少 80%，更优选至少 90%，甚至更优选至少 95%，并且最优选 100%的同源性。优选的多核苷酸包括具有例如转录调控活性的生物活性的那些核苷酸。

可以理解的是，本发明进一步涉及到分离的互补于任何一种上述的多核苷酸序列的多核苷酸。如本文使用的，术语"互补"指核苷酸之间杂交或碱基配对，这样，例如，在双链多核苷酸的两条链之间。当一链的核苷酸，以适当的核苷酸插入、缺失或取代最优排列，与另一链至少 80%的核苷酸配对时，两条单链 DNA 核苷酸分子称之为互补。

本发明的另一方面涉及到包括上述的新型 ABC1 核苷酸和适当的载体的组合物。一项实施方案中，该组合物包括编码包含 SEQ ID NO: 2 多肽的多核苷酸，包含 SEQ ID NO: 1 的多核苷酸，包含 SEQ ID NO: 1 的核苷酸 291-7074 的多核苷酸，或者编码包含氨基酸序列的多肽的多核苷酸，该氨基酸序列至少 98%同源于氨基酸序列 SEQ ID NO: 2，以及适当的载体。另一项实施方案中，这种组合物包括在其整个长度内具有与包含 SEQ ID NO: 1 的多肽至少 80%，优选 90%，或者更优选 95%同源性的多核苷酸，以及适当的载体。

另一项实施方案中，这种组合物包括含 ABC1 侧翼序列的多核苷酸和适当的载体。优选地，这种组合物包括含 SEQ ID NO: 3 的多核苷酸或者含 SEQ ID NO: 3 核苷酸片段 1-1532，1080-1643，1181-1643，1292-1643，或 1394-1643 和适当的载体。同样优选地，这种组合物包括在其整个长度内具有与含任何一种描述过的 5'侧翼序列至少 80%，90%，或 95%同源性的多核苷酸，以及适当的载体。又一项实施方案中，这种组合物包括含 ABC 3'侧翼序列和适当载体的多核苷酸。优选的组合物包括含 SEQ ID NO: 4 的多核苷酸，含 SEQ ID NO: 5 的多核苷酸，或含 SEQ ID NO: 6 的多核苷酸，以及在其整个长度内具有与这些多核苷酸的任何一种至少 80%，

优选 90%，或更优选 95%同源性的多核苷酸，和适当的载体。本发明的又一项组合物包括突变型 ABC1 多核苷酸。优选地，这种组合物包括含 SEQ ID NO: 7 的多核苷酸或含 SEQ ID NO: 9 的多核苷酸和适当的载体。

此外，本发明的组合物可包括，任意组合地，两种或多种上述 ABC1  
5 多核苷酸和适当的载体。任何适当的含水载体可用于这种组合物。优选地，这种载体使得组合物在预计的温度稳定，例如室温或贮藏温度（即 4℃至 20℃），并且近似为中性 pH。适当载体的例子对于本领域熟练人员是已知的，并且包括 Tris-EDTA 缓冲剂和 DEPC-H<sub>2</sub>O。

#### ABC1 载体和宿主细胞

10 本发明亦涉及到包括一种或多种上述 ABC1 多核苷酸的重组载体，以包含 ABC1 多核苷酸的载体的基因工程的宿主细胞，以及通过重组技术产生 ABC1 多肽。如所述的，本发明提供包括一种或多种上述的野生型 ABC1 多核苷酸的重组载体。在优选的实施方案中，这种重组载体包括多核苷酸，它编码含 SEQ ID NO: 2 的多肽，含 SEQ ID NO: 1 的多核苷酸，  
15 和含 SEQ ID NO: 1 核苷酸 291-7074 的多核苷酸。另一优选的实施方案中，重组载体包括编码包含至少 98%同源于 SEQ ID NO: 2 氨基酸序列的氨基酸序列。同样，另一优选的实施方案中，重组载体包括至少 80%同源于包含 SEQ ID NO: 1 的多核苷酸的变体多核苷酸。又一项优选的实施方案中，重组载体包括互补于这些氨基酸的任何一种的多核苷酸。

20 另一实施方案中，这种重组载体包括自丹尼尔疾病患者分离的突变型 ABC1 多核苷酸。在一项优选的实施方案中，重组载体包括含 SEQ ID NO: 8 的多核苷酸。另一项优选的实施方案中，重组载体包括含 SEQ ID NO: 10 的多核苷酸。重组载体也包括互补于这些序列的多核苷酸序列。

这种重组载体亦理解可包括，任意组合地，两种或多种上述的野生型、  
25 变体型或变异型 ABC1 多核苷酸。

一种分离的 ABC1 多核苷酸，例如任何上述野生型、变体型或变异型多核苷酸，采用公知的连接和克隆技术，插入载体。在几种标准实验室手册中，包括 Davis 等人，*Basic Methods in Molecular Biology* (1986); Sambrook 等人，*Molecular Cloning: A Laboratory Manual*, 2<sup>nd</sup> Ed., (Cold Spring Harbor Press, Cold Spring Harbor, N. Y. (1989)); Ausubel 等人 eds., *Current Protocols in Molecular Biology*, (Wiley 和 Sons (1994)); Goedel ed., *Gene Expression Technology* (Methods in Enzymology (1991)); Murray, ed., *Gene Transfer 和 Expression Protocols* (Human Press, Clifton, NJ), 已经介绍了克隆技术。

10 可以采用任何适合于 ABC1 多核苷酸插入的载体。通常选择在采用的特有的宿主细胞中是功能的载体 (即, 这种载体与宿主细胞机 (machinery) 相匹配, 这样以致于可发生基因的扩增和/或表达)。优选地, 载体与细菌、昆虫或哺乳动物宿主细胞相匹配。同样优选地, 这种载体是一种表达载体 (对于表达载体的述评, 参见 Goedel, D. V. ed., *Methods Enzymol.*,  
15 *Academic Press Inc.*, San Diego, CA (1990))。例如, 这种载体可以是噬菌体、质粒、病毒或逆转录病毒载体。逆转录病毒载体可以是复制型或复制缺损型。在后一种情形下, 病毒增殖通常会仅在宿主细胞中产生。

一般地, 采用任意的这些宿主细胞的表达载体包含适于质粒维持和外源的核苷酸序列克隆和表达的序列。这样的序列, 总的来说称为"侧翼  
20 序列", 应该优选包括一种或多种下列的核苷酸序列: 启动子, 一种或多种增强子序列, 复制起点, 转录终止序列, 含有剪接供体和受体的完全内含子序列, 编码前导序列以便多肽分泌的前导序列, 核糖体结合位点, 聚腺苷酸化序列, 用于插入编码待表达多肽的核酸的多接头区, 和可选择的标记因子。

25 侧翼序列可以是同源的 (即, 源自与宿主细胞相同的种和/或菌株),

5 异源的 (即, 来自除了宿主细胞种或菌株之外的种), 杂交体 (即, 一种以上来源的侧翼序列的组合), 或者合成的。同样, 侧翼序列是一般功能为调节 ABC1 多肽表达的天然序列。侧翼序列的来源可以是任何原核或真核生物, 任何脊椎或无脊椎生物, 或者任何植物, 只要这种侧翼序列在宿主细胞机中是功能的, 且可由其激活。

载体应该优选包括至少一种可选择的在宿主中增殖的标记。一种可选择的标记是编码对于生长在选择性培养基中的宿主细胞的存活和生长必需蛋白的基因因子 (gene element)。适合的选择标记基因编码这样的蛋白: (a)赋予原核宿主细胞对如氨苄西林、四环素或卡那霉素的抗生素或其它毒素的抗药性; (b) 补充细胞营养缺陷的缺失; 或者 (c) 提供复合培养基中没有的必需的营养物。优选的选择性标记包括真核细胞培养物 zeocin、G418、潮霉素、新霉素抗药性和在 *E. coli* 和其它细菌中四环素、卡那霉素或氨苄西林抗药性。

15 其它的适当的选择基因包括用于扩增表达基因的那些。扩增是这样的方法, 其中对于对生长有决定性的蛋白生产极需的基因, 在重组细胞连续生殖的染色体, 是串联复制的。适合于哺乳动物细胞的选择性标记实例包括二氢叶酸还原酶 (DHFR)和胸苷激酶。哺乳动物细胞转化体置于选择压力下, 这里仅仅转化体独一无二地适合存活, 凭借位于载体中的选择基因。培养基内的选择剂浓度连续地变化, 从而导致选择基因和 ABC1 基因均扩增, 在这样条件下通过培养转化的细胞, 强制实行选择压。结果, 20 从扩殖的 DNA 合成增加数量的 ABC1 多肽。

载体也应优选包含转录终止序列, 它一般位于多肽编码区的 3'端, 并且有助于终止转录。通常地, 原核细胞内的转录终止序列是接着多 T 序列的 G-C 丰富片段。这种序列可以作为商业上载体部分购买, 或者采用 25 公知的核酸合成方法合成。

这种载体也应优选包含核糖体结合位点，它对于 mRNA 的翻译起始是必需的，并且由 Shine-Dalgarno 序列 (原核生物)或 Kozak 序列 (真核生物)表征。该因子对于启动子一般位于 3'端，对于待表达的 ABC1 多肽的编码序列位于 5'端。Shine Dalgarno 序列是变化的，但是一般是多嘌呤 (即，具有高 A-G 含量)。已经鉴别了许多 Shine-Dalgarno 序列，每一种采用公知的方法可轻易地合成。

载体也优选包含通过宿主生物识别并可操作地连接于编码多核苷酸上的启动子。这种启动子是诱导的启动子或者组成型启动子。诱导的启动子在它们的控制下引发自 DNA 升高水平的转录，响应培养条件的一些变化，例如存在或缺乏营养物或温度改变。相反，组成型启动子引发连续基因产品产生；从而，对于基因表达很少或无控制。适当的启动子可操作地连接多核苷酸，通过从源 DNA 经限制酶切消化移动启动子并插入载体希望的启动子序列。如描述过的，可以采用天然启动子以直接扩增和/或表达多核苷酸。从而，重组载体可包括任何一种上述的野生型、变体型或突变型 ABC1 多核苷酸和 ABC1 启动子，例如在 SEQ ID NO: 3 发现的。重组载体也可包括，任意组合地，两种或多种以上介绍的野生型、变体型或突变型 ABC1 多核苷酸和 ABC1 启动子。

优选地，与天然 ABC1 启动子相比，可以采用异源启动子，如果它允许较大转录和较高收率的 ABC1 蛋白，并且如果它与适用的已经选择的宿主细胞相容。异源启动子可以单独，或者组合天然 ABC1 启动子使用。从而，在一项优选的实施方案中，重组载体包括任何一项上述的野生型 ABC1 多核苷酸和异源启动子。另一项优选的实施方案中，重组载体包括任何一种以上介绍的变体型 ABC1 多核苷酸和异源启动子。还一项优选的实施方案中，重组载体包括任何一种以上介绍的突变型 ABC1 多核苷酸和异源启动子。优选的实施方案也包括重组载体，它包含以上介绍的两

种或多种野生型、变体型、突变型 ABC1 多核苷酸的组合以及异源启动子。

适用于原核生物宿主的异源启动子包括，但并非限制， $\beta$ -内酰胺酶和乳糖启动子体系 (Villa-Kamaroff 等人, 1978, *Proc. Natl. Acad. Sci. U.S.A.*, 75: 3727-31), 碱活性磷酸酶, 色氨酸 (trp) 启动子体系和杂交体启动子, 例如 tac 启动子 (Villa-Kamaroff 等人, 1978, *Proc. Natl. Acad. Sci. U.S.A.*, 75: 3727-31)。它们的序列业已出版, 从而使得本领域熟练人员能够连接它们至希望的 DNA 序列, 采用对于提供任何有用的限制位点必需的接头或衔接头。其它适合的异源启动子对本领域熟练人员是已知的。

适合的用于哺乳动物宿主细胞的异源启动子也是公知的, 并且包括, 但并非限制, 得自病毒基因组的那些, 例如多瘤病毒、禽痘病毒、腺病毒 (例如腺病毒 2)、牛乳头瘤病毒、禽肉瘤病毒、巨细胞病毒、逆转录病毒、乙型肝炎病毒和猴病毒 40 (SV40)。其它适合的哺乳动物启动子包括异源哺乳动物启动子, 例如, 热激启动子和肌动蛋白启动子。另外适合的启动子包括, 但是并非限制: SV40 早期启动子和晚期启动子区 (Bernoist 和 Chambon, 1981, *Nature* 290: 304-10); 包含于肉瘤病毒 3'长末端重复序列的启动子 (Yamamoto, 等人, 1980, *Cell* 22: 787-97); 疱疹胸苷激酶启动子 (Wagner 等人, 1981, *Proc. Natl. Acad. Sci. U.S.A.* 78:1444-45); 和金属硫蛋白 (metallothionine) 基因的调节序列 (Brinster 等人, 1982, *Nature* 296: 39-42)。优选地, 启动子是巨细胞病毒或 SV40 启动子。从而, 在特别优选的实施方案中, 重组载体包括一种上述的野生型 ABC1 多核苷酸, 一种上述的变体型 ABC1 多核苷酸和巨细胞病毒启动子。又一项特别优选的实施方案中, 重组载体包括, 任意组合地, 两种或多种上述的野生型、变体型或突变型 ABC1 多核苷酸和巨细胞病毒启动子。

载体亦优选包含增强子序列以提高如 ABC1 多核苷酸的转录。适合于

活化真核生物启动子的增强子包括病毒增强子，例如 SV40，巨细胞病毒早期启动子和腺病毒增强子。

本发明的表达载体可以自起始载体，例如商业上可得到的一种载体构建，它可以包含或不包含一切希望的侧翼序列。这里载体中不含有本文介绍的一种或多种侧翼序列时，它们可以各个获得并连接至载体。用于得到每个侧翼序列的方法对于本领域熟练人员是公知的。

优选的载体与细菌、昆虫和哺乳动物宿主细胞相容。优选用于细腻的载体包括，例如， pQE70、pQE60 和 pQE-9 (Quiagen, Inc.)， pBluescript 载体、Phagescript 载体、pNH16A、pNH18A、pNH46A (Stratagene Cloning Systems, Inc.)， ptrc99a、pKK223-3、pDR540、pRIT5 (pharmacia Biotech, Inc.)， 和 pCEP4 (Invitrogen Corp., Carlsbad, CA)。优选的真核生物载体包括，但是并非限制， pWLNEO、pSV2CAT、pOG44、pXTI 和 pSG (Stratagene)、pSVK3、pBPV、pMSG 和 pSVL (Pharmacia)和 pGL3 (Promega, Madison, WI)。其它适合的载体对于本领域熟练技术人员是明显的。

在一项特别优选的实施方案中，重组载体包括 pCEPhABC1，它在实施例 4 中介绍并如图 3 所示。重组载体 pCEPhABC1 包括质粒 pCEP4 (Invitrogen)，一种包含巨细胞病毒启动子和增强子的表达载体。载体 pCEPhABC1 进一步包括可操作连接这种异源巨细胞病毒启动子的 ABC1 多核苷酸。这种 ABC1 多核苷酸包括 SEQ ID NO: 1，它包含全长的 ABC1 cDNA，包括非编码的 5'侧翼 (即，天然 ABC1 启动子)和 3'侧翼序列。

此外，本发明提供包含 ABC1 侧翼序列多核苷酸的重组载体。一项实施方案中，这种重组载体包括一种包含优选含有启动子活性的 ABC1 5'侧翼序列多核苷酸。从而，在一项优选实施方案中，这种重组载体包括含 SEQ ID NO: 3 的多核苷酸。另一项优选的实施方案中，这种重组载体包括含 SEQ ID NO: 3 的 1-1532，1080-1643，1181-1643，1292-1643，或

1394-1643 核苷酸的多核苷酸。一项特别优选的实施方案中，多核苷酸包括 SEQ ID NO: 3 的 1394-1532 核苷酸。同样，在另一项实施方案中，重组载体包括在严格条件下与 SEQ ID NO: 3 提出的或者包含 SEQ ID NO: 3 的 1-1532, 1080-1643, 1181-1643, 1292-1643, 或 1394-1643 核苷酸杂交的多核苷酸。在又一项实施方案中，这种多核苷酸包括在其整个长度内具有与含 SEQ ID NO: 3 的多核苷酸或者包含 SEQ ID NO: 3 的 1-1532, 1080-1643, 1181-1643, 1292-1643, 或 1394-1643 核苷酸至少 80%，更优选 90%，甚至更优选至少 95%同源性的核苷酸序列。重组载体亦包括互补于任何上述的 5'侧翼序列的多核苷酸。

10 另一项实施方案中，重组载体包括含 ABC1 3'侧翼序列的多核苷酸。一项优选的实施方案中，重组载体包括含 SEQ ID NO: 4 的多核苷酸。另一项优选的实施方案中，重组载体包括含 SEQ ID NO: 5 的多核苷酸。在同样优选的实施方案中，重组载体包括含 SEQ ID NO: 6 的多核苷酸。同样地，另一项实施方案中，重组载体包括在严格条件下与 SEQ ID NO: 4, 15 SEQ ID NO: 5, 或 SEQ ID NO: 6 的多核苷酸杂交的多核苷酸。又一项实施方案中，该多核苷酸包含一种核苷酸序列，它在其整个长度内具有与含 SEQ ID NO: 4, SEQ ID NO: 5, 或 SEQ ID NO: 6 的多核苷酸至少 80%，更优选于少 90%，甚至更优选至少 95%的同源性。重组载体也包含与任何上述的 3'侧翼序列互补的多核苷酸。

20 一种分离的 ABC1 侧翼序列核苷酸，例如任何一种上述的 5'或 3'侧翼序列，采用公知的连接和克隆技术插入载体。可以采用任何一种先前描述的载体。优选地，这种载体与细菌、昆虫或哺乳动物宿主细胞相容。同样优选地，这种载体是表达载体。例如，这种载体可以是噬菌体、质粒、病毒或逆转录病毒载体。

25 除了 ABC1 侧翼序列之外，这种载体可以包含一种或多种下列侧翼核

5 核苷酸序列：启动子，一种或多种增强子序列，复制起点，转录终止序列，含有剪接供体和受体的完全内含子序列，编码前导序列以便多肽分泌的前导序列，核糖体结合位点，聚腺苷酸化序列，用于插入编码待表达多肽的核酸的多接头区，和可选择的标记因子。任何以前介绍过的侧翼核苷酸序列是适合的。侧翼序列可以是同源的、异源的、杂交体或者合成的。同样，侧翼序列是一般功能为调节 ABC1 多肽表达的天然序列。侧翼序列的来源可以是任何原核或真核生物，任何脊椎或无脊椎生物，或者任何植物，只要这种侧翼序列在宿主细胞机中是功能的，且可由其激活。

10 优选的载体是与细菌、昆虫和哺乳动物相容的那些载体。先前已经介绍了适当的载体，它包括用于细菌的 pQE70, pQE60, pQE9, pBluescript 载体, Phagescript 载体, pNH16A, pNH18A, pNH46A, ptrc99a, pKK223-3, pDR540, pRIT5, 和 pCEP4, 以及用于真核生物的 pWLNEO, pSV2CAT, pOG44, pXTI, pSG, pSVK3, pBPV, pMSG, pSVL, 和 pGL3。

15 在一项特别优选的实施方案中，这种重组载体包括含 5'侧翼区的 ABC1 基因的多核苷酸，并且进一步包括至少一种编码异源多肽的多核苷酸。这种异源多肽可操作地与 ABC1 5'侧翼序列连接。这种 ABC1 5'侧翼序列优选含有 ABC1 启动子。从而，优选地，5'侧翼序列包括 SEQ ID NO: 3 提出的序列。同样优选地，这种 5'侧翼序列包括 SEQ ID NO: 3 的 1-1532, 1080-1643, 1181-1643, 1292-1643, 或 1394-1643 核苷酸。异源的多核苷酸 20 优选是完全蛋白或蛋白的生物学活性片段。这种载体也包含多于一种的异源多核苷酸。更优选地，这种异源多核苷酸编码报道蛋白。这种情形下，重组载体优选不包含任何额外的启动子序列。适合报道蛋白的例子包括荧光素酶、 $\beta$ -半乳糖苷酶、氯霉素乙酰转移酶 转移酶，以及绿色荧光蛋白。优选地，报道多肽是荧光素酶。从而，在一项特别优选的实施方案中，25 重组载体包括 SEQ ID NO: 3 提出的 5'侧翼序列和荧光素酶报道基因多核

苷酸。其它同样优选的实施方案中，该重组载体包括含 SEQ ID NO: 3 的 1-1532, 1080-1643, 1181-1643, 1292-1643, 或 1394-1643 核苷酸的多核苷酸以及荧光素酶报道基因多核苷酸。

可以从包含报道基因多核苷酸的起始载体，例如商业上可得到的载体，构建包括 ABC1 5'侧翼序列的表达载体。适合的表达载体的例子包括含有荧光素酶报道基因的 pGL3-Basic (Promega, Madison, WI), 和 pDGal-Basic (Clontech, Palo Alto, CA)。优选的载体是 pGL3-Basic 荧光素酶报道载体，它是非启动子。包含 ABC1 启动子的 5'侧翼序列，例如 SEQ ID NO: 3, 可以采用公知的方法，包括本文描述的方法 (参见实施例 15) 接入一种上述的表达载体。因此，在特别优选的实施方案中，重组载体是 pAPRI, 一种在 pGL3 载体中构建的包含 SEQ ID NO: 3 和荧光素酶报导基因的报导基因 (参见图 11)。

本发明也涉及到包括任何一种上面介绍过的重组载体的宿主细胞。在已经构建载体之后，可以插入全部载体至适合的宿主细胞，以便扩增和/或多肽表达。宿主细胞可以是原核生物宿主细胞 (例如 *E. coli*) 或者真核生物宿主细胞 (例如酵母细胞，昆虫细胞，或脊椎动物细胞)。宿主细胞，当在适当条件下培养时，合成 ABC1 多肽，或者选择性地，报道基因多肽，它随后可测量到。宿主细胞可以是哺乳动物宿主细胞，例如灵长类细胞系或啮齿类细胞系，包括转化细胞系。通常的二倍体细胞，由原始组织体外培养物衍生的细胞菌株，以及原始外植体，亦是适合的。候选宿主细胞在选择基因时可以是基因型缺失的，或者可以包含显性开放选择基因。

大量的适合的宿主细胞在本领域是公知的，并且可以从美国典型培养物保藏中心 (ATCC), Manassas, VA 得到。适当的哺乳动物宿主细胞包括，但并不是限制，中国仓鼠卵巢细胞 (CHO), CO DHFR-细胞 (Urlaub 等人, 1980, *Proc. Natl. Acad. Sci.*, 97: 4216-20), 人胚胎肾 (HEK)293

或 293T 细胞，或 3T3 细胞，猴 COS-1 和 COS-7 细胞系，和 CV-1 细胞系。其它适合的哺乳动物细胞系包括但是并非限制，小鼠成神经细胞瘤 N2A 细胞，HeLa 细胞，小鼠 1-929 细胞，源自 Swiss、Balb-c 或 NIH 小鼠的 3T3 系，BHK 或 HAK 仓鼠细胞系，Thp-1，HepG2，和小鼠 RAW 细胞系。这些细胞系中的每一种对于蛋白表达领域熟练人员是已知和可得到的。优选的宿主细胞是小鼠单核细胞系 RAW 264.7，实施例 8 描述了它的用途。

适当的细菌宿主细胞包括各种菌株的 *E. coli* (例如，HB101，DH50，DH10，和 MC1061)，作为生物工程领域的宿主细胞它是公知的。在这种方法中也可以采用各种菌株的 *B. subtilis*，*Pseudomonas spp.*，其它的 *Bacillus spp.*，*Streptomyces spp.*，等等。同样，作为表达 ABC1 多肽的许多菌株的酵母细胞也是可利用的。优选的酵母细胞包括，例如，啤酒糖酵母和巴斯德毕赤氏酵母。此外，昆虫细胞体系是适当的宿主细胞。例如，Kitts 等人于 1993, *Biotechniques*, 14: 810-17; Lucklow 于 1993, *Curr. Opin. Biotechnol.* 4: 564-72; 和 Lucklow 等人于 1993, *J. Virol.*, 67: 4566-79 中介绍了昆虫细胞体系。优选的昆虫细胞是 Sf-9 和 Hi5 (Invitrogen)。

本发明也提供一种在哺乳动物宿主细胞中表达 ABC1 蛋白的方法，它包括以下步骤：(a) 以一种包含多核苷酸编码 ABC1 的重组表达载体，以能够产生可检测水平的 ABC1 蛋白的充分量转染哺乳动物宿主细胞；和 (b) 纯化产生的 ABC1 蛋白。在一项优选的实施方案中，重组表达载体包括编码含 SEQ ID NO: 2 的多肽的多核苷酸。另一项优选的实施方案中，这种多核苷酸包含 SEQ ID NO: 1。又一优选实施方案中，这种多核苷酸包含 SEQ ID NO: 1 的核苷酸 291-7074。还有一项优选实施方案中，这种多核苷酸编码一种多肽，它至少 98% 同源于一含 SEQ ID NO: 2 的多肽。

重组 ABC1 载体至哺乳动物宿主细胞的引入可通过本领域公知的和

标准实验室手册介绍的方法实现，例如 Sambrook, *supra*。优选地，这种重组载体以一种带电脂质引入沉淀物或复合体内的宿主细胞。适于 ABC1 载体引入的方法包括磷酸钙转染，DEAE-葡聚糖介导的转染，阳离子的脂质-介导的转染，电穿孔，转导，传染，或其它本领域已知的方法。例如，  
5 这些方法描述在 Davis 等人的 *Basic Methods in Molecular Biology* (1986)。如果重组 ABC1 载体是病毒载体，采用适当的包装细胞系它可以在体外包装，接着转导入宿主细胞。

采用已知的方法，包括硫酸铵或乙醇沉淀、酸、提取、阴离子或阳离子交换色谱法、磷酸纤维素色谱法、疏水性相互作用色谱法、亲和层析法、  
10 羟磷灰石色谱法和外源凝集素色谱法，这种 ABC1 多肽可以从重组细胞培养物回收和纯化（参见，例如 Smith 和 Johnson, *Gene* 67: 31-40 (1988)）。优选地，采用使用抗-ABC1 抗体的亲和层析以纯化。

此外，本发明提供一种在哺乳动物受体中表达 ABC1 蛋白的方法，它包括给予哺乳动物受体充分量的重组表达载体，该表达载体含有编码  
15 ABC1 的多核苷酸，以在所述哺乳动物受体中表达 ABC1 蛋白。优选地，这种重组表达载体包括编码包含 SEQ ID NO: 2 多肽的多核苷酸，包含 SEQ ID NO: 1 的多核苷酸，包含 SEQ ID NO: 1 的 291-7074 核苷酸的多核苷酸，或者编码至少 98%同源于含 SEQ ID NO: 2 的多肽的多核苷酸的多核苷酸。重组 ABC1 载体至哺乳动物受体的引入可以通过本领域公知的方法实现，并且在下文中详细描述。通过从重组 ABC1 载体给药的受体取得的血样，分离单核群体，并且测量巨噬细胞中 ABC1 基因的表达，  
20 可以测量表达的 ABC1。ABC1 基因的表达可以采用本领域公知的方法（例如 RT-PCR）或者本文描述的方法测量（参见实施例 9 和 10）。ABC1 蛋白的水平可以通过从受体中采集血样，分离单核细胞和采用公知的方法在  
25 巨噬细胞中测量 ABC1 蛋白，例如本文实施例 11 描述的免疫沉淀法。

## 增加胆固醇排放的方法和化合物

本发明的另一方面，提供一种适于从哺乳动物细胞中增加胆固醇排放的方法。这种方法包括给予哺乳动物受体充分量的一种包含 ABC1 多核苷酸的重组表达载体，以从所述细胞中提高胆固醇排放。这种重组载体可以是任何一种以上介绍的包含任何一种先前介绍的野生型或变体型 ABC1 多核苷酸的载体，只要这种编码的 ABC1 多肽具有生物学活性（即，胆固醇转运活性）。优选地，该重组载体含有编码含 SEQ ID NO: 2 的多肽的多核苷酸，含 SEQ ID NO: 1 的多核苷酸，含 SEQ ID NO: 1 的 291-7074 的核苷酸，或者编码包含氨基酸序列的多肽的多核苷酸，该氨基酸序列至少与 SEQ ID NO: 2 的氨基酸序列 98%同源。同样地，优选地，重组载体包括变体型多核苷酸，它至少 80%、90%或 95%同源于含 SEQ ID NO: 1 的多核苷酸，只要这种编码的 ABC1 多肽具有胆固醇转运活性。

重组 ABC1 表达载体对哺乳动物受体的给药可以用于在所述受体中表达 ABC1 基因，以便治疗心血管疾病。具体地，这种方法通过将 ABC1 基因引入巨噬细胞和在具有心血管疾病动物的动脉损伤中发现的其它胆固醇-蓄积细胞，会达到其治疗作用。在靶细胞中的 ABC1 多核苷酸的表达会有效增强 ABC1 蛋白的生产。随后生产的 ABC1 蛋白会通过增加从巨噬细胞和在动脉损伤中发现的其它胆固醇-蓄积细胞的胆固醇排放至初生的 HDL 颗粒，以改善疾病。细胞排放会导到胆固醇自外周部位（例如动脉血小板的富含胆固醇核）的全面清除。同时发生的这些病理学损伤尺寸的减少，减小了导致心脏病突发和心绞痛的动脉阻滞的风险。这种方法也可预防性地用于防止胆固醇在动脉壁的蓄积。

充分量的 ABC1 表达载体是可以从哺乳动物受体细胞中增加胆固醇排放的 ABC1 载体，。通过在各种剂量的重组 ABC1 表达载体给药之前（对照水平）和之后，测量受体细胞内的胆固醇排放，并且确定与对照水

平相比引起胆固醇排放增加的剂量，从而确定这种数量。通过从受体中获得血样，分离单核细胞群，并且测量在受体巨噬细胞内胆固醇排放的数量，可以测量胆固醇排放。任何如此描述的试验可以用于测量胆固醇排放。

此外，胆固醇排放可以通过在重组 ABC1 表达载体给药之前 (对照) 5 和之后，测定受体质粒 HDL 胆固醇水平来测量。重组 ABC1 表达载体给药之后，一种观察到的受体血清中 HDL-胆固醇的升高表明胆固醇排放的增加。采用本领域公知的方法可以测定血清中 HDL-胆固醇水平。

此外，胆固醇排放可以通过在重组 ABC1 表达载体给药之前 (对照水平) 10 和之后，测量受体动脉壁内发现的动脉粥样硬化损害的大小来测量。动脉损害尺寸的减少，表明胆固醇排放的增加。测量动脉损伤的试验是本领域公知的。例如，采用 Lawn 等人于 *Nature*, 360: 670-672 (1992)) 描述的小鼠动脉粥样硬化模型，以及任何其它已知的模型，从动脉损害可以测量增加的胆固醇排放。采用 Lawn 等人介绍的 LDL 受体剔除小鼠，在 ABC1 载体给药之前和之后可以测量胆固醇排放。在食用致动脉粥样硬化的食物的 15 动物组中脂肪斑纹尺寸可以通过如 Lawn 等人介绍的油红 O 染色的主动脉部分测量。接受 ABC1 表达载体的组与接受对照载体的组相比，脂肪斑纹损害尺寸的减少表明从损害处胆固醇排放的增加。在人类中，在动脉壁发现的动脉粥样硬化损伤的尺寸，例如，可以采用血管造影术和非攻击性超声法测量。

20 或者，胆固醇排放可以通过在重组 ABC1 表达载体给药之前和之后，采血样并测量巨噬细胞内 ABC1 mRNA 或 ABC1 蛋白水平测定。可以进行常规的试验以确定增加的 ABC1 mRNA 浓度和胆固醇排放之间的联系。同样地，进行试验以使 ABC1 蛋白数量与胆固醇排放数量之间相关。采用这样的相关数据，重组 ABC1 表达载体给药之后，在受体细胞内观察 25 到的 ABC1 mRNA 和 ABC1 蛋白的增加可以用于表明胆固醇排放的增加。

ABC1 mRNA 和 ABC1 蛋白水平可以采用本文介绍的试验和其它对于 mRNA 和蛋白定量公知的技术测量。治疗剂量和制剂以下将进一步讨论。

适于核酸分子引入靶细胞的多病毒和非病毒方法对于本领域熟练技术人员是可得到的。例如，适于基因治疗的病毒送递载体包括，但是并非  
5 限制，腺病毒、单纯疱疹病毒、痘病毒（即牛痘）、肝炎病毒、细小病毒、乳多空病毒、甲病毒属、冠状病毒、弹状病毒、乳头瘤病毒、腺伴随病毒 (AAV)、脊髓灰质炎病毒和 RNA 病毒，例如逆转录病毒和新培斯病毒。ABC1 多核苷酸也可采用非病毒送递体系送递，例如裸 DNA 送递（直接  
10 注射）、受体-介导的送递（DNA-配体复合体）、电穿孔、腺病毒-配体-DNA 复合体、磷酸钙 (CaPO<sub>4</sub>) 沉淀、微粒轰击（基因枪技术）、脂质体介导的转运和脂转染。

细胞的基因修饰可以采用一种或多种基因治疗领域公知的技术完成 (Mulligan, R., 1993, *Science*, 260 (5110): 926-32)。基因治疗材料和方法可包括诱导的启动子，组织特异性增强子-启动子，为位点特异整合设计  
15 的 DNA 序列，能够提供优于亲本细胞的选择性优点的 DNA 序列，识别转化细胞的标记，负选择体系和表达控制体系（安全测量），细胞特异结合剂（适于细胞寻靶），细胞特异性内化因子，和通过载体增强表达的转录因子，以及载体制备方法。下述文献介绍了基因治疗技术实践的方法和材料的例子：美国专利 No.4,970,154（包括电穿孔技术），5,679,559（介  
20 绍了适于基因送递的含脂蛋白体系），5,676,954（包括脂质体载体），5,593,875（介绍适于磷酸钙转染的方法），和 4,945,050（介绍生物活性粒子在细胞以一定速率推进，于是这种粒子穿入细胞表面并掺入细胞内部的方法），和 PCT 公开 No.WO 96/40958（包括核配体）。

腺病毒载体已经证实特别适用于基因转运至真核细胞 (Rosenfeld, M., 等人, *Science*, 252: 431-4 (1991); 美国专利 No.5,631,236)。人类 Ad-  
25

介导的基因治疗第一次试验是囊性纤维化跨膜传导调节蛋白 (CFTR) 基因至肺的转运 (Crystal, R., 等人, 1994, *Nat. Genet.*, 8 (1): 42-51)。重组 Ad 给药至各种体内组织的实验途径已经包括气管内滴注法 (Rosenfeld, M., 等人, 1992, *Cell*, 68 (1): 143-55) 注射入肌肉 (Quantin. B., 等人, 1992, *Proc. Natl. Acad. Sci. U.S.A.*, 89 (7):2581-4), 外周静脉内注射 (Herz, J., 和 Gerard, R., 1993, *Proc. Natl. Acad. Sci. U.S.A.*, 90 (7):2812-6) 和定向接种至脑 (Le Gal La Salle, G., 等人, 1993, *Science*, 259 (5097):988-90)。接着, 腺病毒载体是本领域熟练人员广泛采用的且适于本发明的。

腺伴随病毒 (AAV) 最近已经作为有效用途的基因转运体系引入基因治疗。野生型 AAV 证明在整合入宿主细胞基因组中高水平感染性、宽宿主范围和特异性 (Hermonat, P., 和 Muzyczka, N., 1984, *Proc. Natl. Acad. Sci. U.S.A.*, 81 (20):6466-70)。单纯疱疹病毒 1-型 (HSV-1) 是一种优选的载体体系 (Geller, A., 等人, 1991, *Trends Neurosci.*, 14 (10):428-32; Glorioso, J., 等人, 1995, *Mol. Biotechnol.*, 4 (1):87-99; Glorioso, J., 等人, 1995, *Annu. Rev. Microbiol.*, 49:675-710)。痘病毒科的牛痘病毒, 亦发展为一种表达载体 (Smith, G., 和 Moss, B., 1983, *Gene*, 25 (1):21-8; Moss, B., 1992, *Biotechnology*, 20:345-62; Moss, B., 1992, *Curr. Top. Microbiol. Immunol.*, 158:25-38)。对于本领域熟练人员每一种上述载体是广泛采用的并且适用于本发明。

优选地, 这种病毒送递体系采用逆转录病毒载体。逆转录病毒载体能够感染大百分比的靶细胞并整合入细胞基因组 (Miller, A., 和 Rosman, G., *Biotechniques*, 7 (9):980-2,984-6,989-90 (1989); 美国专利 No.5,672,510)。研发逆转录病毒作为基因转运载体相对早于其它的病毒, 并且它首先成功地用于基因标记和转导腺苷脱氨基酶 (ADA)cDNA 至人淋巴细胞。优选地, 这种逆转录病毒载体是鼠或鸟逆转录病毒的衍生物,

或者是慢病毒载体。特别优选的逆转录病毒载体是慢病毒载体。其中可插入单外源基因的逆转录病毒载体包括，但并不是限制：Moloney 鼠白血病病毒 (MoMuLV)，Harvey 鼠肉瘤病毒 (HaMuSV)，鼠乳房瘤病毒 (MuMTV)，SIV，BIV，HIV 和鲁斯氏肉瘤病毒 (RSV)。为了选择性的标记，大量的额外的逆转录病毒载体可掺入多基因，这样以致于可以识别和产生转导细胞。

由于重组逆转录病毒是缺失的，它们需要协助以更产生感染的载体粒子。例如，可以通过采用包含编码所有逆转录病毒的结构基因辅助细胞系，在 LTR 内调节序列的控制下，提供这种协助。这种辅助质粒缺少使包装机制能够识别 RNA 转录核苷酸序列以衣壳化的核苷酸序列。因此，由于无基因组普查包装，辅助细胞系产生空病毒颗粒。适当的细胞系包括，但是并非限制， $\psi$ 2，PA317 和 PA12。如果引入逆转录病毒载体至这种包装信号完整的细胞，但是结构基因由其它目的基因取代，可包装该载体并生产载体病毒颗粒。然后可使用通过这种方法产生的载体病毒颗粒，以感染组织细胞系，例如 NIH 3T3 细胞，以产生大量的嵌合的逆转录病毒颗粒。

本发明的载体也可采用本领域熟练人员广泛采用的标准重组技术构建。在普通的分子生物学参考文献中可发现这样的技术，例如 *Molecular Cloning: 实验室手册* (Sambrook, 等人, 1989, Cold Spring Harbor Laboratory Press), *Gene Expression Technology* (Methods in Enzymology, Vol. 185, edited by D. Goeddel, 1991, Academic Press, San Diego, CA), 和 *PCR Protocols: 方法和应用指南* (Innis, 等人 1990, Academic Press, San Diego, CA)。

为了在靶细胞内获得 ABC1 多核苷酸转录，可采用能够在靶细胞驱动基因表达的转录调控区。这种转录调控区优选包括启动子和/或增加子，

它可操作地与 ABC1 多核苷酸连接。对于 ABC1 基因，这种启动子可以是同源或异源的，只要它在将插入构建体的细胞或组织类型内是活性的。选择的转录调控区域应该推动在靶细胞内高水平的基因表达。优选地，采用巨噬细胞特异性启动子，例如消除受体 A 型基因，基质金属蛋白启动子 (MMP-12)，或者热带巨噬细胞慢病毒属启动子 (Fabunmi 等人，*Atherosclerosis*, 148:375-386 (1999))。特别优选的启动子是 5'区的清除受体 A 型基因，它含有强巨噬细胞启动子，可以用于驱动 ABC1 基因的转录。此外，增加细胞内内源 ABC1 多肽表达的方法是插入启动子区一种或多种增强子因子，这里增强子因子有助于提高 ABC1 基因的活性。同样地，选择使用的增强子因子基于人们希望活性基因的组织。从而，例如，会选择已知赋予启动子在动脉组织发现的细胞内激活的增强子因子。其它的适用于本发明的转录调控区包括，但是并非限制，人巨细胞病毒 (CMV) 立即早期增强子/启动子，SV40 早期增强子/启动子，JC 多瘤病毒启动子，白蛋白启动子，PGK 和偶联至 CMV 增强子的 $\alpha$ -肌动蛋白启动子 (Doll, R. 等人, 1996, *Gene Ther.*,3 (5):437-47)。

其它的载体成分可选择性地包括为位点特异整合设计的 DNA 分子 (例如，适用于同源重组的内源序列)、组织特异性启动子、能够提供优于母细胞的选择性优点的 DNA 分子，用作识别转化细胞的 DNA 分子，负选择体系，细胞特异性结合剂 (例如适于细胞寻靶)，细胞特异性内化因子，增强载体表达的转录因子，和使载体能够生产的因子。

在一项实施方案中，载体可包括适于位点特异整合的靶 DNA。靶 DNA 是一种核苷酸序列，它互补于 (同源的)目的基因区，例如 ABC1 基因。通过同源重组，这种插入基因组的 DNA 序列可通过连接其至靶 DNA 指向 ABC1 基因。适于插入的 DNA 序列包括，例如，互相影响或者控制 ABC1 多肽表达的 DNA 区，例如侧翼序列。因此，并不是通过编码 ABC1

基因本身的 DNA 转染，而是借助于与 DNA 调节部分偶联的靶 DNA 实现希望的 ABC1 多肽表达，这种 DNA 调节部分以可识别的用于 ABC1 多肽转录的信号提供内源基因序列 (Sauer, *Curr. Opin. Biotechnol.*, 5: 521-27 (1994); Sauer, *Methods Enzymol.*, 225: 890-900 (1993))。

5        又一项实施方案中，对于控制靶细胞内 ABC1 基因表达，可包括调节因子。随适当的效应子接通这样的因子。这样，希望时可表达治疗的 ABC1 多肽。一种常规的控制方法包括使用小分子二聚物或 rapalogs，以二聚包含小分子结合域和能够引发生物学方法的域的嵌合蛋白，例如 DNA-结合蛋白或转录活化蛋白 (参见 PCT 公开 No.WO 96/41865, WO 97/31898, 10 和 WO97/31899)。蛋白的二聚作用可以用于引发转基因的转录。另一适当的控制方法或基因开关包括使用米啡司酮 (RU486)，它是孕甾酮拮抗剂。修饰的孕甾酮受体配体结合域与孕甾酮拮抗剂的结合，通过形成接着变成结合 DNA 的核的两种转录因子二聚物，激活转录。修饰配体-结合域以剔除受体结合天然配体的能力。美国专利 No.5,364,791 和 PCT 公开 15 No.WO96/40911 和 WO97/10337 进一步介绍了修饰过的甾类激素受体体系。又一项控制方法采用阳性四环素控制的反式激活蛋白。该体系包括变异的破伤风阻抑蛋白 DNA-结合域 (变异的破伤风 R-4 氨基酸变化，它导致逆转四环素-调节的反式激活蛋白，即它在四环素的存在下结合破伤风操纵基因)，它与激活转录的多肽连接。美国专利 No.5,464,758, 5,650,298, 20 和 5,654,168 介绍了这类体系。美国专利 No.5,741,679 和 5,834,186 介绍了另外的表达控制体系和核酸构建体。

含 ABC1 多核苷酸的病毒送递载体通过改变病毒包被产生靶向特异性，这样以致于它包含对于在靶细胞发现的其它分子特异的配体。这会允许载体特异地与希望的细胞类型结合。这种配体可以是任何目的化合物， 25 它会与靶细胞内发现的分子，例如细胞表面的受体，特异结合。优选地，

仅仅发现该受体在靶细胞而不是其它细胞上。例如，可采用消除受体 A 的配体以指导病毒递载体至巨噬细胞。或者，可以改变这种病毒包被，这样以致于它包含抗体或抗体片段，例如 Fab，或 F(ab')<sub>2</sub>，它识别并与靶细胞上的抗原表位结合。通过插入另外的编码配体至病毒基因组的多核苷酸，可改变这种病毒包被。本领域的熟练人员会知晓其它的特异多核苷酸序列，它可插入病毒基因组以允许定向特异递送含 ABC1 多核苷酸的载体。

此外，可以给予裸 ABC1 多核苷酸。ABC1 多核苷酸和重组 ABC1 表达载体，例如以上介绍的那些，可以作为药物组合物给药。这样的组合物包括有效量的 ABC1 多核苷酸或重组 ABC1 表达载体，如本文以前定义的，以及药学上可接受的选择适于给药模式的配制剂。适合的制剂原料优选是在采用的浓度对受体无毒，并且如下文描述。

包括 ABC1 多核苷酸或 ABC1 重组表达载体的药物组合物可包含适于改性、维护或保存，例如组合物的 pH、充分渗透压浓度、粘度、澄明度、颜色、等渗性、气味、无菌、稳定性、溶解或稀释速率、吸收或渗入。适合的制剂原料包括，但是并非限制，氨基酸（例如甘氨酸、谷氨酰胺、天门冬酰胺、精氨酸或赖氨酸），抗菌剂，抗氧化剂（例如抗坏血酸，亚硫酸钠或亚硫酸氢钠），缓冲剂（例如硼酸盐、碳酸氢盐、Tris-HCl、柠檬酸盐、磷酸盐或其它有机酸），湿胀剂（例如甘露醇或甘氨酸），螯合剂（例如乙二胺四乙酸 (EDTA)），络合剂（例如咖啡因、聚乙烯吡咯烷酮、β-环糊精或羟丙基-β-环糊精），充填剂，单糖，二糖，和其它碳水化合物（例如葡萄糖、甘露糖或糊精），蛋白（例如血清白蛋白、明胶或免疫球蛋白），着色，调味和稀释剂，乳化剂，亲水聚合物（例如聚乙烯吡咯烷酮），低分子量多肽，成盐平衡离子（例如钠盐），防腐剂（例如苯扎氯铵、苯甲酸、水杨酸、硫汞撒、苯乙醇、尼泊金甲酯、尼泊金丙酯、洗必太、山梨

酸或过氧化氢), 溶剂 (例如甘油、丙二醇或聚乙二醇), 糖醇 (例如甘露醇或山梨醇), 悬浮剂, 表面活性剂或润湿剂 (例如普罗尼克 (pluronic); PEG; 脱水山梨醇酯; 聚山梨醇酯例如聚山梨醇酯 20 或聚山梨醇酯 80; 粹通; 氨丁三醇; 卵磷脂; 胆固醇或四丁酚醛), 稳定性增强剂 (例如蔗糖或山梨醇), 毒性增强剂 (例如碱金属卤化物-优选氯化钠或钾, 或者甘露醇山梨醇), 送递载体, 稀释剂, 赋形剂和/或药学上的助剂。参见 *Remington's Pharmaceutical Sciences* (18<sup>th</sup> Ed., A. R. Gennaro, ed., Mack Publishing Company 1990)。

药学上活性化合物 (即 ABC1 多核苷酸或 ABC1 载体) 可以按照药剂学常规方法加工以生产适于向患者, 包括人和哺乳动物, 给药的药剂。因此, 包含 ABC1 多核苷酸或 ABC1 重组表达载体的药物组合物可以制备成固体形式 (包括颗粒剂、粉剂或栓剂) 或者液体形式 (例如溶液、悬浮液或乳剂)。适于口服给药的固体剂型包括胶囊剂、片剂、丸剂、粉剂和颗粒剂。在这样的固体剂型中, 活性化合物可以与至少一种惰性稀释剂例如蔗糖、乳糖或淀粉混合。通常实践中, 这样的剂型也可包括, 惰性稀释剂之后的另外的物质, 例如润滑剂, 如硬脂酸镁。在胶囊剂、片剂和丸剂的情形下, 该剂型也包括缓冲剂。片剂和丸剂可以肠衣另外制备。

适于口服或非肠道给药的液体剂型可包括药学上可接受的乳剂、溶液、悬浮液、糖浆和酏剂, 它们含有本领域通常使用的惰性稀释剂, 例如水。这类组合物也可包括助剂, 例如润湿增甜剂, 调味剂和香味剂。例如, 适合于注射的载体是水、生理盐水溶液, 或者人工脑脊髓液, 可能以通常在于非肠道给药的组合物中的其它物质补充。中性的缓冲的盐水或以血清白蛋白混合物盐水是进一步的示范载体。其它示范性药物组合物包括 pH 约 7.0-8.5 的 Tris 缓冲剂, 或者 pH 约 4.0-5.5 的乙酸盐缓冲剂, 它可进一步包括山梨醇或适当的取代物。

适于以包括 ABC1 多核苷酸或 ABC1 表达载体治疗心血管疾病的给药方案是以多种因素为基础，包括心血管疾病的类型、年龄、体重、性别、患者的体格状况、病情的严重性、给药途径和采用的具体化合物。因此，给药方案可广泛变化，但是可以采用标准方法常规测定。例如，待给药的 ABC1 多核苷酸或 ABC1 表达载体的数量是足以增加哺乳动物受体细胞胆固醇排放的数量。可以测定这种数量，例如，通过在 ABC1 多核苷酸或 ABC1 表达载体给药之前和之后，测量受体血浆 HDL 胆固醇水平。充分数量的 ABC1 多核苷酸或 ABC1 表达载体是能够增加受体血浆 HDL-胆固醇水平的数量。从而，临床医师可以滴定测定 (titer) 剂量并且修正给药途径以获得最优治疗作用。常规的剂量范围是大约 0.1? g/kg 至大约 100 mg/kg，或者更多，取决于以上述及的因素。

给药频率取决于所采用制剂中 ABC1 多核苷酸或载体的药理学参数。通常地，临床医师会给予组合物直到剂量达到实现预计的作用。因此，随时间过去该组合物可以一次，两次或多次剂量 (它含或不含同样数量的希望的分子)，或者经移植装置或导管连续输入。进一步求精的适当剂量可由本领域熟练人员常规确定，并且在它们完成的工作范围之内。通过采用适当的量-效数据可以确定适当的剂量。可以体内、离体、试管内转染哺乳动物受体的细胞。

可以采用任何一种本领域熟练人员公知的技术完成 ABC1 多核苷酸或包含 ABC1 多核苷酸的重组载体至靶细胞的体内给药。例如，美国专利 No. 5672344 介绍了一种体内病毒-介导的基因转移体系，它包括重组神经营养 HSV-1 载体。以上介绍的 ABC1 多核苷酸和重组 ABC1 载体组合物可以通过经口、经颊、非肠道、直肠或局部给药以及通过吸入喷雾转染。本文使用的术语"非肠道"包括皮下、静脉内、肌内、胸骨内、输液法或腹膜内。

对于口服给药,这种包含 ABC1 多核苷酸或重组 ABC1 载体的药物组合物可以是这样的剂型,例如,胶囊剂、片剂、悬浮液、或液体。这种药物组合物优选制备成包含给定数量的 DNA 或病毒载体粒子的剂量单位。例如,这些可包含大约  $10^3$ - $10^{15}$  数量的病毒颗粒,优选大约  $10^6$ - $10^{12}$  病毒颗粒。适合于人或其它哺乳动物的每天剂量可根据患者的症状和其它因素广泛变化,但是,再一次,可以采用常规方法确定。

含有 ABC1 DNA 或重组 ABC1 载体的药物组合物也可以直肠给药。适于载体直肠给药的栓剂可以通过混合载体和适当的非刺激性赋形剂制备,例如在普通温度下是固体但在直肠温度下是液体的可可脂或聚乙二醇,从而在直肠熔化和释放载体。

药物组合物也可配制成吸入剂。例如,ABC1 多核苷酸或载体也可配制成适于吸入的干粉。同样地,ABC1 多核苷酸或载体吸入溶液可以以适于气溶胶运输的抛射剂配制。又一实施方案中,溶液可以成雾状。

也可以注射包含 ABC1 DNA 或重组 ABC1 载体的药物组合物。可注射的制剂,例如无菌可注射的含水或油状的悬浮液,可以按照已知的方法采用适当的分散或润湿剂和悬浮剂配制。无菌可注射的制剂也可以是无毒非肠道可接受的稀释剂或溶剂的无菌可注射的溶液或悬浮液,例如 1,3-丁二醇溶液。特别适于非肠道注射的载体是无菌蒸馏水,其中 ABC1 多核苷酸或载体配制成无菌、等渗、适当防腐的溶液。其它可接受的载体和溶剂中,可以采用 Ringer's 溶液,和等渗的氯化钠溶液。此外,通常采用无菌、不挥发性油作为溶剂或悬浮介质。为此目的,可以采用任何温和的不挥发性油,包括合成的单-或二脂酸甘油酯。此外,例如油酸的脂肪酸发现用于可注射的制剂。又一项制剂可包括预计的 ABC1 分子和一种例如可注射的微球、生物-可侵蚀的颗粒、聚合化合物、珠或脂质体的助剂的配方,它能提供控制和缓释的 ABC1 产品 (参见,例如

PCT/US93/00829; Eppstein 等人, *Proc. Natl. Acad. Sci.*, 82: 3688-3692 (1985))。

同样地, 包含 ABC1 DNA 或重组 ABC1 载体的药物组合物可以局部给药。对于局部给药, 该载体可包括制剂的 0.001% 至 10% w/w, 例如, 5 以重量计 1% 至 2%, 尽管它可包括制剂的差不多 10%w/w, 但优选不超过 5%w/w, 更优选 0.1% 至 1%。适于局疗给药的制剂包括适于经皮肤渗入的液体或半-液体制剂 (例如, 搽剂、洗剂、软膏、霜或糊剂) 和适于向眼、耳或鼻给药的滴剂。本发明载体的活性成分适当的局部给药是每天给药一至四次, 优选两或三次。

10 当本发明的核酸和/或载体作为唯一活性药剂给药时, 它们也可联合一种或多种本发明的载体或其它药剂使用。当组合给药时, 治疗剂可以配制为在同一时间或不同时间给药的分离组合物, 或者该治疗剂可以作为单一组合物给予。

本发明的另一项实施方案中, 通过"生产细胞系"植入, 体内接近转染 15 靶细胞至靶细胞群 (Culver, K., 等人, 1994, *Hum. Gene Ther.*, 5 (3):343-79; Culver, K., 等人, *Cold Spring Harb. Symp. Quant. Biol.*, 59:685-90; Oldfield, E., 1993, *Hum. Gene Ther.*, 4 (1):39-69)。改造这种生产细胞系, 以生产包含 ABC1 多核苷酸的病毒载体并且在靶细胞附近释放其病毒颗粒, 即优选在动脉粥样硬化损伤处发现的巨噬细胞。部分释放的病毒 20 颗粒接触靶巨噬细胞并感染那些细胞, 从而送递 ABC1 多核苷酸至靶巨噬细胞。靶细胞感染之后, 发生 ABC1 多核苷酸的表达, 假如巨噬细胞具有功能 ABC1 蛋白。

另一项实施方案中, 本发明通过离体引入 ABC1 多核苷酸或重组 ABC1 表达载体, 提供治疗心血管疾病的方法。在这样的情形下, 已经从 25 患者移出的细胞、组织或器官暴露于 ABC1 组合物, 之后细胞、组织或

器官随后移植回患者。例如，一种方法包括从患有心血管疾病的受体中取出血样，丰富该样品的单核细胞，以及使分离的单核细胞与包含 ABC1 多核苷酸和任选包含的定向特异基因的重组表达载体接触。选择性地，可以以生长因子，例如 GM-CSF 处理这种单核细胞，以在再引入细胞于受体之前刺激细胞生长。当再引入时，该细胞会特异性地寻靶细胞群，它们开始从该细胞群中被分离出。这样，可以采用转运活性的 ABC1 多肽以促进受体胆固醇排放。

另一种离体给药的方法包括通过包含病毒的细胞的皮肤移植，引入 ABC1 多核苷酸或重组 ABC1 载体至哺乳动物受体。优选地，逆转录病毒适用于这种给药方法。如果采用强的管家基因启动子推动转录作用，采用成纤维细胞起源的细胞，可以实现外源基因在植入物中的长期表达。例如，可以采用二氢叶酸还原酶 (DHFR) 基因启动子。以包含逆转录病毒构建体的病毒颗粒可以感染例如成纤维细胞的细胞，该构建体包含 ABC1 多核苷酸，加能够特异寻靶的基因，例如消除受体 A，和强启动子。感染过的细胞可以包埋入能够移植入受体真皮结缔组织的胶原基质。由于逆转录病毒增生和逃避基质，它会特异性感染靶细胞群。这样，移植导致显示转运紊乱的细胞内增加数量的胆固醇排放活性。

另一方案中，包含 ABC1 多核苷酸的重组表达载体可以采用体外技术给药，例如美国专利 No. 5,399,346 介绍的。例如，通过移植某些已经采用如本文介绍的那些方法基因改造，以表达 ABC1 多肽的细胞，送递 ABC1 多肽。为了最小化服用 ABC1 多肽的患者中可能的免疫学反应，外源种的多肽给药时会发生，产生 ABC1 多肽的天然细胞优选是人类起源并生产人 ABC1 多肽。从而，这种产生 ABC1 多肽的重组细胞优选以包含编码人 ABC1 多肽的基因的表达载体转化。该细胞可以是自身的或异源的。选择性地，细胞可以被无限增殖化。为了减少免疫学反应的机会，细胞可

以包入胶囊内以避免周围组织的渗入。胶囊化物质一般是生物相容的、半渗透性的聚合壳或膜，它允许蛋白产物的释放，防止患者免疫系统或者来自周围组织的其它不利因素破坏该细胞。转染的细胞可以采用以上介绍的方法对患者给药。

5            胶囊化活细胞的技术本领域是已知的，装胶囊细胞制剂和它们植入患者可以常规完成。例如， Baetge 等人 (PCT 公开 No.WO 95/05452 和 PCT/US94/09299) 介绍了包含基因工程细胞的膜胶囊，适于生物学活性分子的有效送递。这种胶囊是生物相容的并且易于获取。这种胶囊将重组 DNA 分子转染的细胞装胶囊，该分子包含 DNA 编码序列，以便生物学活  
10 性分子可操作地与刚植入哺乳动物宿主时不服从体内负调节的启动子连接。该装置供给分子从活细胞至受体内特异位点的送递。此外，参见美国专利 Nos. 4,892,538; 5,011,472; 和 5,106,627。PCT 公开 No.WO91/10425 介绍了胶囊化活细胞的体系 (Aebischer 等人)。也参见，PCT 公开 No. WO91/10470 (Aebischer 等人); Winn 等人, 1991, *Exper. Neurol.* 113:322-29;  
15 Aebischer 等人, 1991, *Exper. Neurol.* 111:269-75; 和 Tresco 等人, 1992, *ASAIO* 38:17-23。

          另一种适于编码 ABC1 多核苷酸的送递体系是"非病毒"送递体系。已经使用或建议基因治疗的技术包括 DNA-配体复合体、腺病毒-配体-DNA 复合体、DNA 直接注射、CAPOT 沉淀、基因枪技术、电穿孔、脂转染和  
20 胶体分散 (Mulligan, R., 1993, *Science*, 260 (5110): 926-32)。这些方法的任何一种是本领域熟练人员广泛采用的，并且适用于本发明。其它适合的方法是本领域熟练人员广泛采用的，本发明可理解成可以采用任何可利用的转染方法完成。本领域熟练人员经变化的成果已经采用几种这类的方法学 (Mulligan, R., 1993, *Science*, 260 (5110): 926-32)。

25            优选地，这种非病毒送递体系是胶体分散体系。胶体分散体系包括大

分子复合体、毫微囊、微球、珠和包含水包油乳剂、胶束、混合胶束和脂质体的脂基体系。本发明优选的胶体体系是脂质体，它是用作体外和体内送递载体的人工膜载体。脂质体是自我装配的胶体颗粒，其中由两亲分子例如磷脂酰丝氨酸或磷脂酰胆碱组成的类脂双分子层，将部分周围介质的装胶囊，这样以致于类脂双分子层围绕亲水内部。构建单层脂质体或多层脂质体，这样以致于内部包含希望的化学药物，或者当在本发明时包含分离的 DNA 分子。例如，已经显示，大的单层脂质体载体 (LUV)，它的大小范围为 0.2-4.0 $\mu\text{m}$ ，可将相当百分比的包含巨大分子 (例如 RNA、DNA 和完整的病毒颗粒)含水缓冲剂装胶囊。一旦胶囊化含水内部，这些大分子可以以生物学活性形式送递至哺乳动物细胞 (Fraley, R. 和 Papahadjopoulos, D. 1981, *Trends Biochem. Sci.*, 6:7780)。为了使脂质体是一种有效基因传递载体，应该具备以下特征: (1)目的基因高效率装胶囊同是不危害它们的生物学活性; (2)与非靶细胞相比与靶细胞优选和大量结合; (3)高效送递水容量的载体至靶细胞细胞质; 和 (4)基因信息的准确和有效的表达 (Mannino, R., 等人, 1988, *Biotechniques*, 6 (7):682-90)。

脂质体组合物通常是磷脂，特别是高相变温度磷脂的组合，通常与类固醇，尤其是胆固醇组合。也可以采用其它的磷脂或其它的类脂。脂质体的物理特征取决于 pH、离子强度和二价离子的存在。用于脂质体生产的类脂的例子包括磷脂酰基化合物，例如磷脂酰甘油、磷脂酰胆碱、磷脂酰丝氨酸、磷脂酰乙醇胺、鞘类磷脂、脑苷脂和神经节苷脂。特别有效的是二酰基磷脂酰甘油，这里类脂部分包含 14-18 个碳原子，特别是 16-18 个碳原子，并且是饱和的。例证性的磷脂包括卵磷脂酰胆碱、二棕榈酰磷脂酰胆碱和二硬脂酰磷脂酰胆碱。

基于构造上和机械因素已经分类了脂质体的寻靶。构造上的分类基于选择性水平，例如，器官特异性、细胞特异性和细胞器特异性。机械导向

可以基于它是被动还是主动区分。被动寻靶利用脂质体的天然的趋向以在包含正弦样的毛细管的器官中分散网状内皮系统 (RES)细胞。另一方面, 主动寻靶, 包括通过偶联脂质体至特异性配体, 例如多克隆或单克隆抗体、糖、糖脂或蛋白, 或者通过脂质体的成分或大小以便实现对器官和细胞类型而不是天然存在的定位位点的寻靶。优选地, 该配体是可用于引导脂质体至特异细胞表面配体的多克隆或单克隆抗体。这种配体也可以是抗体片段, 例如 Fab 或 F(ab')<sub>2</sub>, 只要它与靶细胞上抗原表位有效结合。优选地, 抗体或抗体片段识别仅仅在靶细胞上发现的抗原。例如, 某些特异性表达在巨噬细胞上的抗原, 如清除受体 A, 可以用于直接导向抗体-ABC1 脂质体至巨噬细胞目的。

大量的方法可以用于共价连接或者多克隆或者单克隆抗体至脂质体双分子层。同样, 类脂基团可以掺入脂质体的类脂双分子层, 以便维持导向配体稳定缔合脂质体双分子层。此外, 可以采用各种连接基团以便结合类脂链与导向配体。

本文表明的研究显示核受体配体能够引起胆固醇排放的增加。相应地, 在另一项实施方案中, 本发明提供一种适于从哺乳动物受体细胞增加胆固醇排放的方法, 包括给予哺乳动物受体足够增加所述细胞胆固醇排放数量的至少一种核受体配体。包含核受体配体的药物组合物可以采用上述药物组合物配制和给药方法制备和给药。充分数量的核受体配体是这种配体数量能够增加胆固醇排放。通过在各种剂量的配体给药之前和之后, 测量胆固醇排放, 并且确定引起胆固醇排放增加的剂量, 从而确定这种数量。胆固醇排放可以采用先前介绍的试验测量。

核受体是配体-激活的转录因子, 它通过转导少量的、亲脂激素的作用至转录反应, 在脊椎动物发育和成人生理学起重要作用。存在几种族的核受体, 包括过氧化物酶体增生活化受体 (PPARs)、肝 X 受体 (LXR)、

维生素 A X 受体 (RXR)、farnesoid X 受体 (FXR)和类固醇和生物异源受体 (SXR)。PPAR 族包括三种密切相关的基因产物 PPAR $\alpha$ 、PPAR $\gamma$ 和 PPAR $\beta/\delta$ 。已经推断 PPAR $\alpha$ 为细胞内和细胞外类脂代谢的关键调节物。当结合脂肪酸时, PPAR $\alpha$ 刺激过氧化物酶体的增生并且诱导包括涉及脂肪酸 $\beta$ -氧化的几种酶的合成。PPAR $\alpha$ 受体亦是 fibrates (规定减少甘油三酯水平的药物)的分子靶 (Isseman 等人, *Nature*, 347, 645 (1990))。Fibrates 充当 PPAR 配体以调节大量的影响脂蛋白和脂肪酸代谢的基因的转录。此外, PPAR $\gamma$ 配体, 例如属于噻唑烷二酮化合物, 已经显示增加人 HDL 和减小甘油三酯水平。

10 LXR 是调节过量胆固醇分解代谢的氧甾醇受体。LXR $\alpha$ 显示作为杂二聚物结合 RXR 至 CYP7a 基因内的 DNA 效应元件, 它编码在胆固醇至胆汁酸的转化中负责限制步骤的酶。研究显示当食用富胆固醇食物时, 由于随着胆固醇饮食它们不能增加 CYP7a 基转录, 缺乏 LXR $\alpha$ 的小鼠在其肝脏中累积巨大的数量胆固醇酯 (Peet 等人, *Cell*, 93: 693 (1998))。零 LXR $\alpha$ 小鼠的进一步地研究显示 LXR $\alpha$ 也涉及几种其它参予胆固醇和脂肪酸体内稳态的基因的调节。密切相关的核受体的生物学作用, LXR $\beta$ , 它在几种器官中表达并由相同的氧甾醇如 LXR $\alpha$ 激活, 仍然不清楚 (Song 等人, *Proc. Natl. Acad. Sci.*, 91:10809 (1994); Seol, *Mol. Endocrinol.*, 9:72 (1995))。

20 FXR, 它发展地涉及到 LXR $\alpha$ , 也涉及胆固醇体内稳态。正如 LXR $\alpha$ , FXR 与 RXR 受体起杂二聚的作用 (Schwartz 等人, *Curr. Opin. Lipidol.*, 9: 113 (1998); Vlahcevic 等人, *Gastroenterology*, 113: 1949 (1997))。FXR 由合成的类维生素 A TNPB 和过生理学浓度的全反式视黄酸激活 (Zavacki 等人, *Proc. Natl. Acad. Sci.*, 94: 7909 (1997))。近来的研究表明  
25 FXR 是核胆汁酸受体。第一, FXR 经过胆汁酸循环充分表达在器官中,

包括肝、肠和肾 (Seol 等人, *Mol. Endocrinol.*, 9: 72 (1995); Forman 等人, *Cell*, 81: 687 (1995))。同样, FXR 近来已经显示充当生理学的浓度的几种胆汁进受体, 其中鹅去氧胆酸 (CDCA)是最有效的 (Kliwer 等人, *Science*, 284:757-284 (1999); Makishima 等人, *Science*, 284: 1362-1363 (1999))。CDCA 已知调节参予胆汁盐体内稳态的几种基因的表达, 包括编码 CYP&a 和肠胆汁酸-结合蛋白的那些。

如实施例 13 详细介绍的, LXR、RXR 和 PPAR 核受体的配体显示在胆固醇-负荷的小鼠巨噬细胞中增加 AI-诱导的胆固醇排放。例如, 9 顺-视黄酸 (30ng/ml)给药在这些细胞内 apoAI-诱导的胆固醇排放上产生大约 3 倍的增长。同样地, 22 (R)羟基胆甾醇 (5 $\mu$ g/ml)给药在 apoAI-诱导的胆固醇排放上产生大约 3 倍的增长。接受 fenbirate (3 $\mu$ g/ml)的细胞在胆固醇排放上产生大约 2 倍增长。这些结果表明可以调整核受体以从巨噬细胞增加载脂蛋白-介导的胆固醇排放。此外, 如实施例 13 介绍的, 9-顺-RA 以剂量依赖性方式介导的来自巨噬细胞的胆固醇排放。其它的核受体激活剂, 例如苯扎贝特, 显示增加胆固醇排放 (数据未显示)。

因此, 在通过核受体配体给药的增加胆固醇排放的方法中, 优选来自 LXR, RXR, PPAR,FXR, 和 SXR 核受体配体的配体。在其中 LXR 配体用于增加胆固醇排放的优选的实施方案中, 更优选来自 20 (S)羟基胆甾醇, 22 (R)羟基胆甾醇, 24-羟基胆甾醇, 25-羟基胆甾醇, 和 24 (S), 25 环氧胆固醇 LXR 配体的配体。最优选地, LXR 配体是 20 (S)羟基胆甾醇。在其中 RXR 配体用于增加胆固醇排放的优选的实施方案中, 更优选来自 9-顺式视黄酸、维生素 A 醇、维生素 A 醛、一切反式视黄酸、13-顺式视黄酸、阿维 A (acitretin)、fenretinide、阿维 A 酯、CD 495、CD 495、CD564、TTNN、TTNNPB、TTAB、LGD 1069 的配体。最优选地, RXR 配体是 9-顺视黄酸。另一优选的 PPAR 配体用于增加胆固醇排放的实施方案中,

配体优选噻唑烷二酮化合物。

另一优选的实施方案中，将超过一种的核受体配体给予哺乳动物受体，以增加胆固醇排放。优选地，当给予受体两种或多种核受体配体时，这类配体是 LXR 和 RXR 配体。更优选地，核受体配体是 20 (S)羟基胆固醇和 9-顺式视黄酸。

又一项实施方案中，本发明提供适于从哺乳动物受体细胞增加胆固醇排放的方法，包括给予哺乳动物受体足够增加胆固醇排放量的二十烷类 (eicosanoid)。包含二十烷类的药物组合物可以采用以上介绍的药物组合物配制和给药方法制备和给药。充分数量的二十烷类是增加胆固醇排放的数量。通过在各种剂量的二十烷类给药之前和之后，测量胆固醇排放，并且确定引起胆固醇排放增加的剂量，从而确定这种数量。胆固醇排放可以采用先前介绍的试验测量。

如实施例 14 介绍的，二十烷类显示在胆固醇负荷的小鼠巨噬细胞内增加 apoAI 诱导的胆固醇排放。例如，PGI<sub>2</sub> (25nm)的给药在这些细胞内的 apoAI-诱导的胆固醇排放上产生大约 2-倍的增长。同样地，PGE<sub>1</sub> (25nm) 给药在 apoAI-诱导的胆固醇排放上产生大约 3-倍的增长。这些结果显示二十烷类选自前列腺素 E<sub>2</sub>、前列腺环素 (前列腺素 I<sub>2</sub>)，和前列腺素 J<sub>2</sub> 二十烷基。

### 增加 ABC1 表达/活性的方法和化合物

假若 ABC1 起促进胆固醇排放作用，一种增加胆固醇排放的方法是增加细胞内 ABC1 的表达。从而，本发明也提供通过给予哺乳动物受体治疗剂量的在所述细胞中增加 ABC1 表达的化合物，增加哺乳动物受体细胞胆固醇排放的方法。治疗量的化合物是增加 ABC1 表达的数量的化合物。通过在各种剂量的该化合物给药之前和之后，测量 ABC1 的基因表达，并且确定引起 ABC1 基因表达增加的剂量，从而确定这种数量。

ABC1 基因表达可以通过从受体中采血样，分离单核细胞群，并采用本领域已知和本文介绍的方法，例如 RT-PCR，测定 ABC1 mRNA 的浓度。

5 一项优选的实施方案中，该方法包括给予 cAMP 类似物以增加 ABC1 的基因表达。如图 8 所示，cAMP 增加了正常成纤维细胞内近 10 倍的 ABC1 mRNA 表达。优选地，这种 cAMP 类似物选自 8-溴代 cAMP，N6-苯甲酰基 cAMP，和 8-硫代甲基 cAMP。另一优选的实施方案中，该方法包括给予增加 cAMP 合成的化合物以增加 ABC 基因表达。优选地，这种化合物是弗司扣林。又一项优选的实施方案中，该方法包括给予抑制 cAMP 断裂的化合物，以增加 ABC1 基因表达。这种化合物的例子是磷酸二酯酶  
10 抑制剂。优选地，这种磷酸二酯酶抑制剂选自咯利普兰、茶碱、3-异丁基-1-甲基黄嘌呤、R020-1724、长春西汀、扎普司特、双嘧达莫、米力农、氨力农、pimobendan、西洛酞胺、依诺昔酮、潘乐新针和维司力农磷酸二酯酶抑制剂。

另一优选的实施方案中，这种方法包括给予哺乳动物受体足够增加  
15 ABC1 基因表达量的核受体配体。如实施例 17 介绍和图 12 所示，核受体配体可以向上调节 ABC1 基因表达。采用 pAPR1 转染研究，它包括在 ABC1 启动子控制下的荧光素酶报道基因，显示在 LXR 配体和 RXR 核受体存在时激活 ABC1 启动子。特别地，与 EtOH 对照相比，在 20 OH-胆固醇 (OH-chol) 存在时以 pAPR1 转染的巨噬细胞在荧光素酶报道活性产  
20 生 19 倍的增长，在 9-顺 RA 存在时在荧光素酶活性产生 16 倍的增长，在两种配体都存在时在荧光素酶活性产生 280 倍的增长。结果表明两种甾醇和类维生素 A 引起源自 ABC1 启动子的强转录反应。此外，在两类化合物之间存在明显的协同作用，由于可通过在以两种配体处理的细胞中观察到荧光素酶活性的惊人增长。按照本发明的方法，优选地，配体选自 LXR，  
25 RXR，PPAR，FXR，和 SXR 配体。

除了增加细胞水平的 ABC1 蛋白之外,可通过增强 ABC1 蛋白的活性促进反向胆固醇转运。从而,另一项实施方案中,本发明提供适于从哺乳动物受体细胞增加胆固醇排放的方法,它包括以足够增加胆固醇排放的数量,给予哺乳动物受体治疗量的增加 ABC1 活性的化合物。包括这类化合物的药物组合物可以采用以上介绍的药物组合物配制和给药方法制备和给药。治疗量的化合物是可增加胆固醇排放的数量的化合物。通过在各种剂量的化合物给药之前和之后,采用先前介绍的方法,测量胆固醇排放,并且确定引起胆固醇排放增加的剂量,从而确定这种数量。为了确定是否胆固醇排放的增长归因于 ABC1 活性的增加,测定位于细胞样品中该化合物给药之前和之后的 ABC1 蛋白数量,采用本文介绍的方法 (参见实施例 11)。对于两种测量 (即,化合物给药之前和之后),胆固醇排放活性总数由 ABC1 蛋白浓度分开,以确定在标准浓度 ABC1 蛋白中发现的胆固醇活性总数。相对于蛋白浓度在标准化胆固醇活性上观察到增加,表明这种增加归因于 ABC1 活性的增加。

### 15 鉴别治疗化合物的方法

本发明的另一方面涉及到筛选化合物以确定该化合物是否调节 (即,向上调节或向下调节)ABC1 基因表达的方法。这类化合物可以用于发展增加 ABC1 表达、从而促进胆固醇排放和升高 HDL-胆固醇血液浓度的治疗化合物。相应地,提供可用于治疗心血管疾病的化合物的鉴别方法。一项实施方案中,本发明提供筛选待试化合物以便 ABC1 表达调控活性的方法,包括以下步骤: (a)以表达调控部分的哺乳动物 ABC1 基因可操作地连接报道 cDNA,以产生重组报道构建体; (b)重组报道构建体转染至宿主细胞群; (c)在转染宿主细胞的样品中分析报道基因表达水平; (d)将转染过宿主细胞与待筛选的待试化合物接触; (e)在与待试化合物接触之后分析在转染宿主细胞的样品中报道基因表达的水平; 和 (f)在由暴露于待试化

合物引起的该水平的报道基因表达上比较有关变化，从而测定 ABC1 表达调控活性。

首先，构建包含可操作地连接于 ABC1 基因表达调控部分的异源报道基因的重组报道构建体。采用公知的连接和克隆技术，例如本文和标准实验室手册介绍的那些，包括 Davis 等人, *Basic Methods in Molecular Biology* (1986); Sambrook 等人, *Molecular Cloning : A Laboratory Manual*, 第 2 版 (Cold Spring Harbor Press, Cold Spring Harbor, N. Y. (1989)); 和 Ausubel 等人 eds., *Current Protocols in Molecular Biology*, (Wiley and Sons (1994)), 这种 ABC1 表达调控多核苷酸和报道基因可以被插入载体。或者，ABC1 表达调控多核苷酸可以插入商业上可得到的报道构建体，例如先前介绍过的那些。可以采用任何适于 ABC1 多核苷酸和报道基因插入的载体。选择的载体在具体采用的宿主细胞中应该是功能的。优选地，这种载体与哺乳动物宿主细胞是相容的。

优选地，ABC1 基因表达调控部分是 ABC1 的 5'侧翼区，包含 ABC1 启动子活性。在一项优选的实施方案中，ABC1 基因表达调控部分包括 SEQ ID NO: 3。另一项优选的实施方案中，ABC1 基因表达调控部分包括 SEQ ID NO: 3 的 1-1532,1080-1643,1181-1643,1292-1643,1394-1643，或 1394-1532 核苷酸。同样，优选的异源报道基因选自编码荧光素酶、 $\beta$ -半乳糖苷酶、氯霉素乙酰基转移酶和绿荧光蛋白多核苷酸。更优选地，异源报道基因是编码荧光素酶蛋白的多核苷酸。在特别优选的实施方案中，这种重组报道构建体是 pAPR1，如图 11 所示。

下一步，这种重组报道构建体转染至宿主细胞群。重组报道构建体可以采用任何一种以前介绍过的转染方法，以及实施例 8 和 15 介绍过的方法，引入宿主细胞。例如，报道基因可以采用磷酸钙转染、DEAE-葡聚糖介导的转染、阳离子类脂-介导的转染、电穿孔、转导、感染，或者任何

一种其它已知和介绍过的方法转染（参见，例如，Davis 等人，*Basic Methods in Molecular Biology* (1986)）。当在适当条件下培养时，宿主细胞可以是任何合成报道基因多肽的细胞，它随后可被测量。优选地，宿主细胞是哺乳动物宿主细胞。更优选地，这种哺乳动物宿主细胞是巨噬细胞、成纤维细胞、肝细胞或肠细胞。更优选地，这种宿主细胞选自 RAW 264.7 细胞，Thp-1 细胞，和 HepG2 细胞。报道构建体的浓度和转染持续时间可以改变，根据转染方法和采用的宿主细胞的类型和浓度。适当浓度的报道构建体和转染时间的测定在本普通技术人员技能之内。

转染之后，分析转染宿主细胞的未暴露于待试化合物的样品，以测定报道基因表达水平。在未暴露的转染宿主细胞样品中发现的报道基因水平提供对照测量。转染宿主细胞裂解，采用本领域任何一种公知的方法分析报道基因表达的水平。根据转染中采用的报道构建体，用于测量报道基因表达水平的试验不同。例如，如果采用荧光素酶报导基因，如实施例 15 所述，采用发光计以光单位测量细胞裂解物中荧光素酶的活性。

转染宿主细胞不同的样品与待筛选的待试化合物接触，随后测量在这些细胞中发现的报道基因表达水平。优选地，转染宿主细胞与待试化合物接触大约 4-48 小时。更具体地，转染宿主细胞与待试化合物接触大约 8-36 小时。甚至更优选地，转染宿主细胞与待试化合物接触，以待试化合物接触转染宿主细胞大约 24 小时。用于测量报道基因表达，在未暴露（对照）细胞中，应该采用同样的用于测量这种报道基因表达水平的试验，以便测量在暴露于待试化合物的细胞内测量报道基因表达水平。

最后，在未暴露对照细胞中发现的报道基因表达水平，与在暴露于待试化合物的细胞内发现的报道基因表达水平相比，以确定这种待试化合物是否具有 ABC1 表达调控活性。如颖在两种细胞样品中的基因表达水平相同或者大约相同，这种待试化合物不能调控 ABC1 基因表达。相对于

在对照细胞中发现的报道基因表达水平,在暴露于待试化合物的细胞中较高水平的报道基因表达表明这种待试化合物向上调节 ABC1 基因表达。相对于在对照细胞中发现的报道基因表达水平,在暴露于待试化合物的细胞中较低水平的报道基因表达表明这种待试化合物向下调节 ABC1 基因表达。

本发明另一方面涉及到筛选待试化合物,以确定该化合物是否促进 ABC1-介导的胆固醇排放的方法。这种方法包括: (a)在保持在培养基内的哺乳动物细胞样品中分析胆固醇排放的水平以测定胆固醇排放的对照水平, (b)以待筛选的待试化合物接触哺乳动物细胞; (c)在与待试化合物接触之后分析细胞样品中胆固醇排放的水平;和 (d)在与待试化合物接触之后,分析细胞样品中 ABC1-依赖性的胆固醇排放,从而确定这种待试化合物是否促进 ABC1-介导的胆固醇从培养物内的细胞中排放。

采用本领域已知和本文介绍的方法可以测定培养的细胞样品内胆固醇排放的水平(参见实施例 1)。可以采用能在培养物内培养的任何哺乳动物细胞,以测量胆固醇排放。该细胞可以源自原始培养物或者固定化的细胞系。方便起见,通过以含有载体的兼嗜性逆转录病毒转染,可以固定化细胞,该载体具有人乳头瘤病毒 16、致癌基因 E6 和 E7 以及选择性标记基因插入物的载体,如实施例 1 所介绍。优选地,培养的细胞是成纤维细胞,巨噬细胞,肝细胞,或肠细胞。更优选地,这种培养细胞是 RAW 264.7 细胞。

#### 优选实施方案的详细说明

测定并未与待试化合物接触的细胞样品中的胆固醇排放量水平以获得一胆固醇排放量的对照水平。此外,测定与待试化合物接触的哺乳细胞样品中胆固醇排放量的水平以确定被待试化合物影响的胆固醇排放量的水平。同样,测定在一与待试化合物接触的哺乳动物细胞样品中 ABC1 介导的的胆固醇排

放量水平以确定被待试化合物影响的ABC1介导的胆固醇排放量水平。更为适宜的是，细胞与待试化合物接触大约8-24小时后才进行胆固醇排放量或ABC1介导的胆固醇排放量的分析测试。ABC1介导的胆固醇排放量水平可以使用其在结合后抑制了ABC1的活性的抗ABC1抗体进行试验。作为一种选择，ABC1接触的胆固醇排放量水平可以使用可以抑制ABC1的表达的反义ABC1多聚核苷酸进行测试。例如，ABC1接触的胆固醇排放量水平可以使用包含有SEQ ID NO: 57 (参见实施例7)的反义ABC1多聚核苷酸进行测试。该细胞应该与抗ABC1抗体或反义ABC1多聚核苷酸同时与待试化合物接触及接触的持续时间相同。

10 如果对照组的胆固醇排放量水平与测试试剂接触的细胞中胆固醇的排放量水平相同或大约相同,该待试化合物并未提高胆固醇的排放量水平。如果与待试化合物接触的细胞中发现的胆固醇的排放量水平高于对照组的胆固醇的排放量水平表明该待试化合物提高了胆固醇的排放量。单独与待试化合物接触的细胞中发现的胆固醇的排放量与与待试化合物及抗ABC1抗体或反义ABC1多聚核苷酸接触的细胞中发现的胆固醇排放量水平的差异表明了经过ABC1介导的胆固醇的排放量。例如，如果对照的胆固醇排放量水平为1.0且与待试化合物接触的细胞中发现的胆固醇的排放量水平为1.1,待试化合物以10%提高了胆固醇的排放量。如果与待试化合物及抗ABC1抗体或反义ABC1多聚核苷酸接触的细胞中发现的胆固醇的排放量水平为1.0，由待试化合物导

15

20 致的胆固醇的排放量全部为ABC1所介导的。

#### 检测冠心病易感性的方法

本发明同样涉及一种检测在哺乳动物受体包括人类中ABC1基因或蛋白表达的相对水平。假定ABC1在胆固醇排放量中的作用,哺乳动物受体中ABC1基因或蛋白的表达相对于预定的标准水平的ABC1基因或蛋白的表达的水平

25 的减少可以被用来显示受体中冠心病的易感性。因此，本发明提供了一种检

测在哺乳动物受体中ABC1基因表达的相对水平的方法，包括以下主要步骤：  
(a) 从哺乳动物细胞受体中获取一测试细胞样品； (b) 分析在测试细胞样品  
中ABC1 mRNA 的表达；及 (c)比较在测试细胞样品中的ABC1 mRNA 的表  
达与预定的标准水平的ABC1 mRNA 的表达，因此来确定在哺乳动物受体中  
5 ABC1基因表达的相对水平。

测试细胞样品首先获自一包括人类的哺乳动物受体。测试细胞样品可以  
是一血样，其中单核细胞数量已被浓缩。使用已知的细胞分离技术基于例如，  
细胞大小，细胞密度，或细胞亲和力对单核细胞进行浓缩。然后分析在测试  
细胞样品中的ABC1 mRNA的表达。ABC1 mRNA的表达水平使用任一已知技  
10 术进行mRNA的预备及检测来进行分析，其中包括如实施例2及9所述的方法。  
例如，ABC1 mRNA的浓缩可以通过反转录多聚酶链式反应，RNA印迹法或  
者核糖核酸酶保护性测试进行确定。ABC1 mRNA浓缩应与测试细胞样品中  
总量的mRNA浓缩标准一致。最后，测试细胞样品中ABC1基因的表达与预定  
标准水平的ABC1 mRNA的表达进行比较。预定标准水平的ABC1 mRNA的表  
15 达可以通过确定取自有代表性意义的哺乳动物受体中的细胞样品中发现的平  
均浓度及分配值来获得，其中哺乳动物受体可以与所获得的测试细胞样品的  
受体为同一种而且哺乳动物受体并没有冠心病，丹尼尔疾病或其他与低  
HDL-胆固醇相关的疾病，且其被认为在正常的范围内具有胆固醇排放量活性  
(也就是说，为正常范围内的HDL-胆固醇水平所指示)。相对于预定标准水平  
20 的ABC1 mRNA 表达，哺乳动物受体的测试细胞样品中ABC1 mRNA 的表  
达的减少水平的确定可被用于指示在哺乳动物受体中冠心病的易感性。

同样，ABC1蛋白减少水平的检测可被用于指示胆固醇排放量的减少及冠  
心病的易感性。因此，本发明的另一实施方案提供了检测在哺乳动物细胞受  
体中ABC1蛋白相对水平的检测方法。这些方法包括以下步骤： (a) 从哺乳  
25 动物细胞受体中获取一测试细胞样品； (b) 分析在测试细胞样品中ABC1 蛋

白的量;及 (c)比较在测试细胞样品中的ABC1蛋白的量与预定的标准水平的ABC1蛋白的量, 因此来确定在哺乳动物受体中ABC1蛋白量的相对水平。

ABC1蛋白的量使用任一测量蛋白的已知技术进行分析。更为适宜的是, ABC1蛋白的量使用免疫测定进行测量。在其中一个实施方案中, ABC1蛋白的量由下列步骤进行确定: (a) 将细胞样品与抗ABC1抗体进行接触且 (b) 检测与抗ABC1抗体相关的细胞样品。例如, ABC1蛋白可以与抗血清接触, 该抗血清抗的合成多肽相应的氨基酸序列为 KNQTVVDAVLTSFLQDEKVKES 如实施例11所示位于C末端。使用本领域的几种已知技术检测抗ABC1抗体, 其包括例如蛋白质印迹, 免疫沉降反应及 FACS, 其中检测可以通过放射性元素, 颜色反应或荧光标记进行。一种优选的在细胞样品中测量ABC1蛋白的数量的方法是免疫沉降,其中生物素酰化的ABC1蛋白与抗ABC1抗体接触, 且结合的抗ABC1抗体使用抗生物素蛋白链菌素辣根过氧化物酶进行检测。

测试细胞样品中的ABC1蛋白数量与预定标准水平的ABC1蛋白数量进行比较。一个预定标准数量的ABC1蛋白可以通过确定在获自哺乳动物受体的细胞样品中发现的ABC1蛋白的平均浓度进行确定。其中哺乳动物受体可以与所获得的测试细胞样品的受体为同一种而且哺乳动物受体并没有冠心病, 丹尼尔疾病或其他与低HDL-胆固醇相关的疾病,且其被认为在正常的范围内具有胆固醇排放量活性 (也就是说, 为正常范围内的HDL-胆固醇水平所指示)。

## 20 ABC1抗体

正如这儿所使用的, 术语“抗体”(Ab)或“单克隆抗体”(Mab)意指包括完整的分子及抗体片段 (例如有Fab及 F(ab')<sub>2</sub>片段), 其可以与蛋白特异性结合。Fab 及 F(ab')<sub>2</sub>片段缺少完整抗体的Fc片段,从循环中快速清除, 且可以相对于完整抗体具有较少非特异性组织结合 (Wahl 等., J Nucl. Med., 24:316-325 (1983)). 这样, 这些片段为优选的, 正如一个FAB产品或其他的免

疫球蛋白表达文库一样。此外，本发明的抗体还包括嵌合，单链及人源化的抗体。

另外的实施方案包括嵌合抗体,例如，人源化的鼠单克隆抗体。这些人源化抗体可以通过已知技术进行准备且提供了一个当抗体被用于人类时免疫原性减少的优点。在一个实施方案中，人源化的单克隆抗体包括有鼠抗体（或正好是其抗原结合位点）的可变区域及获自人类抗体的恒定区域。作为一种选择，人源化抗体片段可以包括鼠单克隆抗体抗原结合位点及一个获自人类抗体的可变区（缺少抗原结合位点）。嵌合抗体的生产程序及进一步加工为单克隆抗体的生产程序包括如Riechmann等. (Nature 332:323, 1988), Liu等 (PNAS 84:3439, 1987), Larrick等 (Dio/Technolog}~ 7:934, 1989), 及 Winter 及Harris (TIPS 14:139, May, 1993)所述那样。

生产抗体的方法包括使用用从下列多核苷酸翻译的多肽免疫一非人动物例如转基因鼠：含有SEQ ID NO: 1所示序列的多核苷酸、含有如SEQ ID NO: 1所示序列的291-7074核苷酸的多聚核苷酸、或含有与含有SEQ ID NO: 1所示序列的多聚核苷酸至少有90%同源性的核苷酸序列的多聚核苷酸，其中直接抗该自所述多聚核苷酸翻译的多肽的抗体在所述动物细胞中产生。已经开发出这样的实验程序用于在非人动物中产生人类抗体。这些抗体可以是部分人源化的，但最好全是人源化的。非人动物（例如转基因鼠）其中被引入以编码一个或多个人类免疫球蛋白链的基因物质是可以获取的。这样的转基因鼠可以以不同的方法对其进行遗传学改变。基因操作最终的结果是在至少一部分（最好是全部）产生于免疫的动物中的抗体中人类免疫球蛋白多肽链取代内源的免疫球蛋白链。本发明还提供了使用转译自上述多聚核苷酸的多肽进行免疫的转基因鼠中产生的抗体。

鼠中一个或多个内源的免疫球蛋白基因通过已准备的不同方法使其失活。人类免疫球蛋白基因被引入鼠中以取代失活的鼠基因。动物中所产生的

抗体包含有为引入该动物中的人类的遗传物质编码的人类免疫球蛋白多肽链。这样的转基因动物的生产及用途技术的实施方案可以参见美国专利5,814,318, 5,569,825, 及5,545,806, 其这儿均被引入作为参考。

5 单克隆抗体可以通过传统技术进行生产,例如,对通过完成免疫程序后获自转基因动物的脾细胞进行无限增殖。脾细胞与骨髓瘤细胞融合以通过传统技术产生杂交瘤细胞。

生产一种杂交瘤细胞系的方法其包括使用具有转译自上述多聚核苷酸的至少7个连续的氨基酸残基的多肽的免疫原免疫这样的一个转基因动物;从免疫动物中收获脾细胞; 将收获的脾细胞与骨髓瘤细胞进行融合, 这样产生一个杂交瘤细胞; 且鉴别这样的一个可以产生结合一转译自上述多聚核苷酸的10 之一的多肽的单克隆抗体的杂交瘤细胞系。这样的杂交瘤细胞系及单克隆抗体包括在本发明的保护范围之内。杂交瘤细胞系分泌的单克隆抗体通过传统的技术进行纯化。

特异性结合于一个ABC1多肽的抗体可以抑制ABC1多肽的活性。更适宜15 的是, 抗体的结合抑制了ABC1多肽胆固醇转移活性。该抗体也可以通过使用相应于胆固醇转移必须区域的多肽免疫非人动物生产。抗体已按照上述任一胆固醇排放量测试方法被测试确定是否其抑制了胆固醇的排放量。这种失活抗体可被用于进行体外实验, 例如这儿所述的任一胆固醇排放量测试方法, 以确定是否待试化合物提高了AbC1介导的胆固醇排放量。失活抗体同样可被20 用于进行体外实验以检测在一个哺乳动物受体细胞中ABC1蛋白的相对水平。失活抗体同样适用于筛选一个能够确定该试剂是否调控ABC1依赖性的胆固醇排放量的试剂盒。

#### 鉴别治疗性化合物的试剂盒

本发明同样包括一个试剂盒,其适用于筛选一个化合物以确定该化合物25 的ABC1表达调控的活性。该试剂盒包括, 以足以实施至少一种检测方法的量,

将一种重组报道构建体作为独立包装的试剂，所述重组报道构建体包括一个可操作地连接至一个哺乳动物ABC1基因的表达调控区域的报道CDNA。试剂盒中还包括有所包装试剂的使用指导。哺乳动物ABC1基因的表达调控区域包括有一个5'侧翼序列。在一个实施方案中，哺乳动物ABC1基因的表达调控区域包括有如SEQ ID NO: 3所示序列。在另一个实施方案中，哺乳动物ABC1基因的表达调控区域包括有如SEQ ID NO: 3所示序列的核苷酸1-1532，1080-1643，1181-1643，1292-1643，1394-1643，或者1394-1532。报道cDNA可以是任一适宜的报道基因，包括在一个具体的实施方案中，重组记录构建体为pAPRI。在另一个实施方案中，试剂盒进一步包括用于检测报道蛋白的方法。这样，试剂盒包含有试剂，例如缓冲剂及底物用于进行蛋白检测

正如这儿所使用到的，术语 "包装" 意指一个固体基质或例如玻璃，塑胶 (例如，聚乙烯，聚丙烯或聚碳酸酯)，纸张，金属薄片及其类似物，其能够在确定限度内支持本发明的重组载体。这样，例如，包装可以是一个玻璃瓶其可被用于包括毫克级数量的预期载体。

"使用说明" 通常包括描述该试剂浓度及至少一种测试方法参数，例如试剂的相对数量及要混合的样品，试剂样品混合液的保存时间，温度，缓冲条件及其类似情况等内容。

此外，本发明同样包括一个适用于筛选一个能够确定该试剂是否调控ABC1依赖性的胆固醇排放量的试剂盒。在一个实施方案中，该试剂盒包括，将足以进行至少一种试验的量的失活的抗ABC1抗体作为独立包装的试剂，同样包括包装试剂的使用指导。

在另一个实施方案中，该试剂盒包括足以进行一种试验的量的一个反义ABC1寡聚核苷酸及使用指导。最好是该反义ABC1寡聚核苷酸包括有如SEQ ID NO: 53所示序列。

下面的实施例进一步说明本发明但不应当被认为是以任何方式限制本发明。

### 实施例1

本实施例证明了丹尼尔疾病 (TD)的患者缺乏apo A-I-介导的脂质排放。

5        细胞培养物: 人类成纤维细胞获自来自两个正常受体 (NL1)及三个不相关的丹尼尔疾病 (TD)患者的皮肤移植体。TD 1细胞获自一个具有严重低血浆HDL胆固醇及apo A-I水平且临床症状为典型的丹尼尔疾病的53岁的妇女。TD2细胞获自一个具有临床症状, 形态学上及生物化学特征为典型的丹尼尔疾病(其包括具有较低水平的的血浆HDL胆固醇及apo A-I (*Francis, 等, J. Clin. Invest., 96, 78-87 (1995)*)) 的56岁的男人。TD3 细胞获自一个具有丹尼尔疾病的18岁的男孩其呈现出桔黄色扁桃体残余, 不对称的运动神经性疾病, 5 mg/dL 的血浆HDL胆固醇及16mg/dL的HDL胆固醇 (*Lawn等., J Clin. Invest., 104, R25-R31 (1999)*)。正常细胞及TD受体细胞如*Oram*等, *油脂研究杂志, 40: 1769-1781 (1999)*所述进行无限增殖培养。简而言之,使用含有人类乳头瘤病毒16癌基因E6及E7及一个新毒素抗性选择标记的插入子的载体的双噬型逆转录病毒转染这些细胞。对照细胞仅使用载体进行感染(模拟感染)。混合细胞在G418的存在下以两种途径进行选择, 然后从培养基中除去G418。成纤维细胞被用于在第5及第16途径之间(初步)或第6至第14途径之间(增殖阶段)。无限增殖的正常及TD细胞接种于一个16-mm 的孔或一个35-mm的器皿中且使其在Dulbecco's修饰的Eagle培养基 (DMEM)及在实验之前加上10%的牛胎血浆 (FBS), 在其中这些细胞进行生长培养。RAW 264.7 鼠单核细胞(美国典型培养物收集中心, Rockville, MD)同样被培养在含有10%的FBS的DMEM中。

25        测量脂质排放量的方法: Apo A-I-介导的胆固醇排放物及磷脂按照*Francis, 等., J. Clin. Invest., 96: 78-87 (1995)*等所述的方法进行试验。来自正常及TD

受体的培养的成纤维细胞通过在  $0.2-0.5\mu\text{Ci/ml}$  [ $^3\text{H}$ ]胆固醇 (40-60 Ci/mmol, Amersham公司., Arlington Heights, IL)存在情况下进行培养标记。当细胞生长至60-80%丰度时放射性胆固醇被加入到含有生长培养基的血清中。3天后,细胞使用PBS/BSA洗涤两次且同时使用捕获培养及载入胆固醇以扩增阿朴脂蛋白介导的脂排放物。其在不含血清的DMEM中加入2 mg/ml 不含脂肪酸的牛血清白蛋白 (DMEM~BSA) (Sigma Chemical Co., St. Louis, MO)及30  $\mu\text{g/ml}$ 非放射性胆固醇中通过孵育该细胞48小时而获得。

RAW 264.7细胞通过食腐动物受体载入胆固醇且如Smith等.,生化杂志., 271:30647-30655 (1996)所述那样使用乙酰基LDL孵育两小时。简而言之,在24孔盘中的RAW 264.7细胞是载入胆固醇的且以0.5 ml DMEM补充以50  $\mu\text{l/ml}$  1M葡萄糖, 10  $\mu\text{l/ml}$  200 mM 谷氨酸盐及2% BSA 及50 $\mu\text{g/ml}$  乙酰化的低密度脂蛋白 (AcLDL)及[ $^3\text{H}$ ]-胆固醇(其已经在37 $^{\circ}\text{C}$  下使用AcLDL 预孵育30分钟以产生一最终的浓度为0.33 $\mu\text{Ci/ml}$  [ $^3\text{H}$ ]-的胆固醇)进行过夜培养标记的。随后细胞使用含有1% BSA的PBS洗涤5次且在DMEM/BSA过夜(16-18小时)孵育以平衡胆固醇的分布。

平衡胆固醇后,细胞使用PBS/BSA漂洗四次且在排放物培养之前使用DMEM/BSA在37 $^{\circ}\text{C}$ 下至少孵育1小时。排放物培养基 (DMEM/BSA)其或者仅含有白蛋白(对照组),白蛋白及 HDL (40 $\mu\text{g}$  蛋白/ml),或者白蛋白加上 apo A-I (10  $\mu\text{g/ml}$ , Biodesign International, Kennebunk, ME)且细胞被孵育 4,24或48小时。磷脂通过在包括有10  $\mu\text{Ci/ml}$ [ $^3\text{H}$ ]维生素B复合体 (75-85 Ci/mmol, Amersham Corp.)的DMEM/BSA中过夜平衡培养进行标记。培养基中发现的放射能在1 2,000g 下离心15分钟后使用闪烁计数进行测量。细胞中的放射能通过0.5 ml 的0.2M NaOH (Smith 等,生化杂志, 271:30647-30655 (1996))中溶解或者在己烷:异丙醇 (3:2 v/v) 中如Francis, 等, *J. Clin. Invest.*, 96, 78-87 (1995)所述那样进行萃取后通过闪烁计数进行测量。含有标记磷脂的细胞使

用1 ml的异丙醇抽提1小时然后按照上述的己烷对异丙醇的比例进行抽提。胆固醇或者磷脂排放物按照在培养基中的氘脂质计数与从细胞及培养基中恢复的总的氘脂质计数的百分比进行表达 ( $\text{cpm 培养基} / \text{cpm (培养基 + 溶解产物)} \times 100$ ).

5 正如图1A及C所示, 额外的HDL或者apo A-I结果在于胆固醇从获自正常细胞受体的载满胆固醇的成纤维细胞中的去除。然而, 在TD细胞中, HDL去除胆固醇的能力稍微有所减少且apo A-I除去胆固醇的能力全部缺失。图1显示的是通过将培养基与白蛋白一起孵育48小时后正常及TD成纤维细胞释放大约34%的细胞 $^{[3}\text{H}]$ -胆固醇进入到培养基中去。额外的进入到白蛋白培养基中  
10 去的HDL显著地提高了来自正常及TD成纤维细胞的 $^{[3}\text{H}]$ -胆固醇排放物量, 尽管对于TD细胞来讲范围较小 (图 1B, D)。apo A-I的增加提高了来自于正常成纤维细胞 (图1A, C)的 $^{[3}\text{H}]$ -胆固醇排放物量, 但是其对于来自于TD成纤维细胞 (图1B, D)的 $^{[3}\text{H}]$ -胆固醇排放物量有较小的作用或没有作用。

## 实施例2

15 该实施例证明了在TD细胞中, 相对于正常细胞中与正常细胞相比, 有175个基因显示出其表达量至少减少2.5倍, 且375个基因显示出其表达量至少增加2.5倍。不同的基因表达可以使用基因表达微阵列 (GEM) 分析来自正常个体 (非TD) 及具有TD的患者中的cDNAs来进行测定。

细胞培养物:使用如实施例1所示的来自正常个体及TD患者的无限增殖  
20 细胞培养物。 保存在DMEM/BSA且24小时补充有1mM 8-溴环腺苷磷酸盐 (8-Br-cAMP, 西格玛化学试剂公司., St Louis, MO)。

mRNA的抽提及cDNA合成: 来自正常及TD成纤维细胞的mRNA可以通过从具有三孢细胞 (生命技术公司, Bethesda, MD, Cat. # 15596-026)的总RNA中抽提而进行准备。mRNA通过Oligotex mRNA试剂盒 (Qiagen 公司.,  
25 瓦伦西亚, CA, Cat. # 70022) 按其提供的操作规程进行分离。按照DeRisi 等,

科学, 24:680-686 (1997)所述的方法, mRNA 通过Cy3或Cy5荧光染色进行反转录以产生荧光标记的cDNA。结果来自TD的cDNA使用Cy3荧光染色且来自正常细胞的cDNA使用Cy5进行荧光染色 (Incyte Genomics, Palo Alto, CA).

微阵列分析: 为了分析来自不同的具有TD或正常的个体的细胞中基因表达的分化, 如上述准备的Cy3及Cy5荧光标记的cDNA样品杂交至一系列显微滑片上的基因微阵列 (GEMs) (Incyte Genomics, Palo Alto, CA), 六个滑片上的每一个包含有大约9,800个人类cDNA样品及200个对照组样品其整个微阵列具有58,800的部分cDNAs。因此, 考虑到估计的多余, 近似有30-50%的已表达的人类基因被表现出来。准备自TD1细胞的Cy3标记的cDNA与来自正常细胞的Cy5标记的cDNA的杂交使得TD细胞的相对RNA含量与正常细胞的已表达基因的对比得以进行。此外, 获自TD2细胞的Cy3-标记的cDNA与来自正常细胞的Cy5标记的cDNA进行杂交至相同的微阵列上以确定在不同的TD患者当中基因表达的变化。

结果: 使用GemTools软件 (Incyte Genomics, Palo Alto, CA)分析数据且将表达为TD细胞相对于正常细胞mRNA的比例。结果显示多数基因在TD1及正常细胞当中表达量大致相同。正如图2 (对角线右侧上部)所示那样仅有175个基因在TD1细胞中相对于正常细胞来说有超过2.5倍的低表达量及有375个基因在TD1细胞中相对于正常细胞来说有超过2.5倍的过表达量 (对角线的右侧下面)。TD细胞中过量表达的基因可以包括那些其结果是丹尼尔突变的已分化调控的基因, 其或者作为一种对该疾病病理学的补偿性的反应。在这些基因当中, 其相关于已观察到的TD表型, 且其在TD细胞中是高表达量的, 其包括有干扰素P (IFN-P), 巨噬细胞炎症蛋白-2 $\alpha$ , 粒细胞趋药性蛋白-2, IL-11, 前列腺素内过氧化物合成酶-2 (COX-2), 血小板反应蛋白, 及单细胞趋药性蛋白1, 3, 及4 (Lawn等, *J. Clin. Invest.*, 104:IR25-R31 (1999)).

并未发现在或者TD1细胞或者TD2细胞当中表达于正常成纤维细胞的

RNA是完全缺失的。同样，在TD1细胞或者TD2细胞当中已表达基因分化的比较揭示了在TD患者个体之间有很小的差异性。例如，TD2细胞当中许多低调控的基因，相比较正常细胞来说，TD1细胞当中有92%属于低表达的基因。在TD1细胞或者TD2细胞当中相比较正常细胞来说超过2.5倍的低表达量的基因之一是编码ABC1蛋白的基因。获取ABC1基因应归功于其同族体的某些功能且同样因为该基因位于被报道作为TD基因区的近似的染色体区域。

### 实施例3

该实施例证明了ABC1基因位于人类染色体9q31上。

先前的基因连接分析绘制了TD基因至一个人类染色体9q31 (Rust 等., Nat. Genet., 20,96-98 (1998))的7-cM区域。此外，原位杂交分析揭示了ABC1基因位于更宽的染色体间隔9q22-9q31 (Luciani等., Genomics, 21, 150-159 (1994))上，使用PCR方法利用基因桥4组人/鼠放射性杂交子 (基因研究公司., Huntsville, AL),人类ABC1被确定位于标记WI-14706及W14062之间,相应于人类染色体9q31的7-cM区，来自93个人/鼠杂交子细胞系的DNA使用PCR方法进行扩增其中使用人类ABC1特异性引物LF: CCTCTCATFACACAAAACCA GAC (SEQ ID NO: 11)及 LR: GCTTTCTTTCACTTCTCATCCTG (SEQ ID NO: 12)。每一细胞系因人类ABC1扩增产品的正负而计分。获自这些数据的分析的ABC1图谱使用怀特黑德研究所/基因组研究软件MIT中心，通过国际互联网 (<http://carbon.wi.mit.edu:8000/cgi-bin/contig/rhmapper.pl>) 来完成。这些结果通过DNA印迹杂交至一个取自相同间隔的人类基因组/酵母人造染色体克隆 (基因研究公司) 来进一步证实。此外，公众数据库检索 (GeneMap '98; 生物技术信息国家中心) 及放射杂交子图谱剔除了绘制在来自于记录基因间隔的其他低表达基因微阵列数据。这些补充数据证明了ABC1基因位于人类染色体9q31上且进一步显示ABC1基因与丹尼尔疾病相关。

### 25 实施例4

这一实施例确定了野生型ABC1基因的核苷酸序列,其包括全部编码区域及侧翼区域。

使用ABI棱镜310遗传分析仪或者经过戴维斯测序 (Davis, CA)来完成对DNA的测序, 其中两条链都被完全测序。来自正常细胞受体的ABC1基因的开放阅读框的序列从获自构建于正常成纤维细胞RNA的表达质粒文库的全长CDNA克隆进行确定。为了构建该质粒文库, 按照Stratagene试剂盒技术方案 (Stratagene, La Jolla, CA)合成cDNA。简而言之, 使用一个位于XhoI位点的寡dT引物同时在5-甲基dCTP的存在下使用一个MMLV反转录酶从一条mRNA合成cDNA的第一条链。第二条链使用核糖核酸酶H及DNA多聚酶I在非修饰性的dNTPs的存在下进行合成。使用pfu DNA 多聚酶对cDNA进行末端钝化, 一个EcoRI连结子捆绑在cDNA上。然后使用Xho I消化cDNA, 在cDNA的3'末端产生一个XhoI末端。内在的XhoI位点在这种消化过程中是受保护的因为第一条链合成时所形成的半甲基化。合成的cDNA被克隆进入一个质粒pCEP4 (Invitrogen Corp., Carlsbad, CA #VO44~50)的HindIII及XhoI位点, 一个含有细胞巨化病毒启动子/增强子的表达载体。一个585 bp的ABC1探针由反转录聚合酶链式反应 (RT-PCR) 使用基于下列已知ABC1序列的引物进行合成, 这些序列为 5'-TCCTTGGGJTCAGGGGATTATC (SEQ ID NO: 13) 及 5'-CAATGTTTTTGTGGCTTCGGC (SEQ ID NO: 14)。使用该ABC1探针, 一个含有10.5 kb的人类ABC1 cDNA插入子的克隆使用克隆捕获选择试剂盒按照生产者的说明 (CLONTECH Laboratories, Inc., Palo Alto, CA)从文库中获得。这一克隆显示于图3为pCEPhABC1。该10.5 kb的ABC1 cDNA插入子序列显示于SEQ ID NO:1。序列确定证实含有人类ABC1开放阅读框的6783 核苷酸及5'及3' 非转译区域的pCEPhABC1, 其比已为Langmann 等在生化, 生物物理研究通报上, 257, 29-33 (1999) (GeneBank 登记号AJ012376)报道的cDNA序列有一个较大的开放阅读框。

## 实施例5

该实施例证明了在野生型ABC1基因及TD1, TD2, TD3基因序列之间的序列差异。

TD1, TD2 及 TD3 的 cDNA 的合成: 通过反转录聚合酶链式反应 (RT-PCR)使用上标选择 cDNA 系统及优势 cDNA 聚合酶混合物按照生产者的说明 (CLONTECH, Palo Alto, CA; Cat. #8417-1)来准备获自 TD1, TD2, 及 TD3 细胞的 cDNA, 其中使用来自正常人类 ABC1 基因序列的两组引物碱基对, 标名为 (1)sacIhabcf, 5'-AGTCGAGCTCCAAACATGTCAGCTGTTACT GGAAGTGGCC (SEQ ID NO: 15); habcr3851, 5'-TCTCTGGATTCTGGGT CTATGTCAG (SEQ ID NO: 16) 及 (2) habcf3585, 5'-GGGAGCCTTTGTG GAACTCTTTC (SEQ ID NO: 17); habcersall, 5'-ACTGGTCGACCATTGAA TTGCATTGCATTGAATAGTATCAG (SEQ ID NO: 18)。

使用这些引物扩增0.2-0.5  $\mu\text{g}$  polyA+ RNA其最终浓度为0.4 $\mu\text{M}$ ,其产生两个大约为3.5 kb的重叠模板。该模板使用QIAEX II系统 (QIAGEN, Inc., Valencia, CA; Cat. #20021)进行凝胶纯化且调整其最终浓度为100ng/ $\mu\text{l}$ 。

TD1, TD2, 及 TD3 cDNA 的测序: 8 $\mu\text{l}$ 上述产生的每一模板与基于野生型 ABC1序列在其最终浓度为0.5 $\mu\text{M}$ 时使用单一测序引物进行反应而测序,引物如下所述:

1F: 5'-TTTCCTGGTGGACAATGAA (SEQ ID NO: 19), 2F: 5'-AGTGA CATGCGACAGGAG (SEQ ID NO: 20); 3F: 5'-GATCTGGAAGGCATGTG G (SEQ ID NO: 21); 4F: 5'-CCAGGCAGCATTGAGCTG (SEQ ID NO: 22); 5F: 5'-GGCCTGGACAACAGCATA (SEQ ID NO: 23); 6F: 5'-GGACAACC TGTTTGAGAGT (SEQ ID NO: 24); 7F: 5'-AAGACGACCACCATGTCA (SEQ ID NO: 25); 8F: 5'-ATATGGGAGCTGCTGCTG (SEQ ID NO: 26); 9F: 5'-GGGCATGAGCTGACCTATGTGCTG (SEQ ID NO: 27); 10F: 5'-AAGA GACTGCTAATTGCC (SEQ ID NO: 28); 1 IF: 5'-AGCGACAAAATCAAGA

AG (SEQ ID NO: 29); 12F: 5'-TGGCATGCAATCAGCTCT (SEQ ID NO: 30); 13F: 5'-TCCTCCACCAATCTGCCT (SEQ ID NO: 31); 14F: 5'-TTCTT CCTCATTACTGTT (SEQ ID NO: 32); 15F: 5'-GATGCCATCACAGAGCT G (SEQ ID NO: 33); 16F: 5'-AGTGTCCAGCATCTAAA (SEQ ID NO: 34);  
5 1R: 5'-CAAAGTTCACAAATACTT (SEQ ID NO: 35); 2R: 5'-CTTAGGGCA CAATTCCACA (SEQ ID NO: 36); 3R: 5'-TGAAAGTTGATGATTTTC (SEQ ID NO: 37); 4R: 5'-TTTTTCACCATGTTCGATGA SEQ ID NO: 38); 5R: 5'-CTCCACTGATGAACTGC (SEQ ID NO: 39); 6R: 5'-GTTTCTTCAT TTGTTTGA (SEQ ID NO: 40); 7R: 5'-AGGGCGTGTCTGGGATTG (SEQ  
10 ID NO: 41); 8R: 5'-CAGAATCATTGGATCAG (SEQ ID NO: 42); 9R: 5'-C ATCAGAACTGCTCTGAG (SEQ ID NO: 43); 10R: 5'-AGCTGGCTTGTTT TGCTTT SEQ ID NO: 44), 11R: 5'-TGGACACGCCAGCTTCA (SEQ ID NO: 45), 12R: 5'-CTGCCATGCCACACACA (SEQ ID NO: 46), 13R: 5'-CT CATCACCCGCAGAAAG (SEQ ID NO: 47), 14R: 5'-CACACTCCATGAAG  
15 CGAG (SEQ ID NO: 48), 15R: 5'-TCCAGATAATGCGGGAAA (SEQ ID NO: 49), 16R: 5'-TCAGGATTGGCTTCAGGA (SEQ ID NO: 50), UTR1R: 5'-A AGTTTGAGCTGGATTTCTTG (SEQ ID NO: 51)。

结果: 核苷酸编号方式以Lawn等 (1999)所示进行。患者TD 1保留有全开放阅读框,且与野生型序列 (SEQ ID NO: 8)相比有两个实质性差别。其中一个  
20 是A 至G的替代,结果在于在一个全长为2201的氨基酸序列位点537处发生一个由谷氨酸到精氨酸的改变,正如Lawn等所述那样 (1999)。这些残基位于NH<sub>2</sub>-末端亲水区,临近第一次预测到的跨膜功能域,患者TD2同样保留有该开放阅读框,其在残基527位点发生一个从精氨酸到色氨酸的取代 (SEQ ID NO:10)。这样,TD 1及TD2 包含有一个取代其改变了该蛋白的相同区域氨基  
25 酸的电荷。 TD3 DNA包含有一个在其ABC1 cDNA一个等位基因的5697位点的核苷酸之前有一个14个核苷酸的插入子及在另一个等位基因的5062核苷酸

位点的一个138 bp的插入子。

TD1, TD2及TD3 DNAs 的基因组测序证实了在其各自的cDNAs中发现的改变。通过PCR方法扩增一个分离自在其来自TD1及TD2的cDNA中发现的突变的成纤维细胞的基因组DNA的一个156 bp的区域来生产该基因组序列。

- 5 基因组测序同样表明患者TD1是一个从谷氨酸到精氨酸的取代的纯合子。基因组DNA分析表明TD2是一个与含有检测到的取代的一个等位基因及另一个含有不确定缺陷的等位基因（其并未产生检测到的mRNA）的异质结合体化合物。没有任何一种取代突变发生于非TD个体的基因组DNA的超过80个的等位基因中。TD3插入子通过序列分析及使用插入点周围引物通过RTPCR来证实。
- 10 核苷酸5697前的14-bp的插入导致一个移码突变，其结果在于野生型氨基酸序列从一个在另一个ATP结合区域前的位点到未成熟蛋白中止点的突变。在另一个等位基因的核苷酸5062位点之前的138 bp的插入导致一个炎症终止密码子。

#### 实施例6

- 15 该实施例证明了在成纤维细胞中ABC1转移活性的抑制剂其同时也可以抑制apo AI-介导的胆固醇排放物。

为了证明是否ABC1的抑制可以作用阿朴脂蛋白介导的胆固醇排放物的产生过程，已报道有两种试剂作为ABC1抑制剂以监控阿朴脂蛋白介导胆固醇排放物的试验方式被测试。这种试剂4,4-均二苯乙烯二异硫氰酸-2,2'二磺酸 (DIDS)及二硫溴苯甲酰 (BSP)已报道在一个药剂依赖性的样品中可以抑制

20 ABC1的阴离子转移活性 (Becq 等.. *J. Biol. Chem.*, 272:2695-2699 (1997); Hamon 等, *Blood*, 90:2911-2915 (1997))。阿朴脂蛋白介导的胆固醇排放物测试如具有显著变化的实施例1描述那样进行。载入胆固醇及<sup>3</sup>H胆固醇标记的正常成纤维细胞 (n=3)使用或者不使用5μg/ml apo A-I 及或者0, 0.2mM,或者

25 者0.4μM DIDS下孵育6小时。此外，载入胆固醇及<sup>3</sup>H胆固醇标记的正常成纤

维细胞 (n=3)使用或者不使用5 $\mu$ g/ml apo A-I 及或者0, 0.2mM,或者0.4 $\mu$ M BSP下孵育6小时。<sup>[3H]</sup>胆固醇排放物量通过如实施例1所示闪烁计数来进行测量且计算为占总共出现在培养基中去的荧光标记的胆固醇的百分数。这些结果显示于图5其为一个在减去不含apo A-I的培养基的值之后apo A-I的存在下排放物的平均值+ SD (n=3)。图5显示的是DIDS及BSP二者同时抑制由阿朴脂蛋白A-I介导的氘示踪胆固醇6小时的排放物量。此外, 使用DIDS及BSP的相似的抑制发现于氘示踪的磷脂酰维生素B复合体的排放物量(数据并未显示)。这些测试的结果与如实施例1所示的具有TD的患者成纤维细胞的排放物缺陷的结果相似。

#### 10 实施例7

该实施例证明了在成纤维细胞中, ABC1 mRNA表达的反义抑制抑制了apo A-I介导的胆固醇排放量。

如实施例1所示使用<sup>[3H]</sup>胆固醇标记正常皮肤成纤维细胞。细胞然后使用寡聚核苷酸通过在或者30 $\mu$ M对照组的吗啉寡聚核苷酸(5'-CCTCTTTACCTCAGTTACAATTTATA-3' 其相应于一个 $\beta$ 球蛋白的地中海贫血症的mRNA的反义互补链 (SEQ ID NO: 52); 或者30 $\mu$ M ABC1反义吗啉寡聚核苷酸 (5'-CATGTTGTTTCATAGGGTGGGTAGCTC-3'; SEQ ID NO: 53)载入且重新接种在一个新的盘上。对照组细胞在使用<sup>[3H]</sup>胆固醇标记后以同样形式载入且在寡聚核苷酸缺失的情况下重新接种, Apo A-I-介导的排放量经过12小时后通过闪烁计数计算为出现在培养基上的荧光标记的胆固醇的总数的百分比。结果显示于图6是三个分别实验的平均值 $\pm$  SEM, 在每一试验中寡聚核苷酸缺失的情况下Apo A-I-特异性的的排放物量的值为同一规格。正如图6所示, 直接指导ABC1 mRNA的反义寡聚核苷酸可导致从正常成纤维细胞中胆固醇排放物量与对照的反义寡聚核苷酸( $\beta$ 球蛋白反义寡聚核苷酸)相对比有50%的减少量。

## 实施例8

该实施例证明了在单核细胞中，人类ABC1基因的过量表达结果使得apo A-I-介导的胆固醇排放物量增加。

RAW 264.7细胞的稳定转染: 鼠单核细胞RAW 264.7使用人类ABC1的 pCEPhABC1 表达质粒进行稳定转染。含有人类ABC1开放阅读框的 pCEPhABC1质粒的构建如实施例4描述的那样进行。大约有 $1 \times 10^6$  RAW 264.7 细胞使用 $2 \mu\text{g}$  pCEPhABC1 DNA及 $12 \mu\text{l}$  Geneporter 转染试剂 (基因治疗系统公司., 圣地亚哥, CA; Cat. #T201007)在 $0.8 \text{ml}$ 不含血清的DMEM中被转染5 小时,两天后,细胞分化的比例范围大约为1:2-1:50且通过加入 $150 \mu\text{g/ml}$ 潮霉素到培养基中去进行选择,两周后,潮霉素抗性克隆被挑选且进行扩大培养。

Apo A-I-介导的胆固醇排放物量测试方法: 母系RAW 264.7细胞及三个稳定表达人类ABC1的克隆细胞系 (L3, L5, 及L6)被培养且汇集。细胞载入胆固醇且通过使用 $0.5 \mu\text{Ci/ml}$  [ $^3\text{H}$ ]胆固醇及 $50 \mu\text{g/ml}$ 乙酰化的LDL如实施例1所示孵育24小时进行标记。经过在DMEM/BSA中过夜孵育来平均胆固醇的量, 然后细胞被洗涤且如实施例1加入排放物量培养基。Apo A-I介导的胆固醇排放物量如前述的在细胞培养基中氘示踪的胆固醇闪烁计数来进行测量, 其表达为获自细胞及培养基中的总的量的百分比。结果表示为在每一试验中来自母系RAW 264.7细胞的apo A-I-特异性排放物量的规范化值的三组分别试验的平均值 $\pm$  SEM。图7显示的是来自母系RAW 264.7细胞及L3, L5, 及 L6 转染细胞系的apo A-I-介导的胆固醇排放物量。正如这儿所看到的那样, 使用ABC1表达载体进行转染的结果在于一个4倍 (L6)至8倍 (L3及L5)的apo A-I-介导的胆固醇排放物量的增加。这些结果表明ABC1基因的过量表达可以很明显地提高巨噬细胞的胆固醇排放量。

## 实施例9

该实施例证明了在正常的皮肤成纤维细胞中而不是在TD成纤维细胞中

ABC1 mRNA表达是由与胆固醇排放量相关的细胞环境所调控。

为了确定是否ABC1在细胞固醇排放物中起到一个速率限定作用,在不同的相关于胆固醇排放量过程的细胞环境下测量ABC1的合成。具体地讲,是将正常成纤维细胞及TD成纤维细胞分别暴露于过量cAMP, 胆固醇或者Apo A-I 的环境下。正常皮肤成纤维细胞及TD1及TD2成纤维细胞的培养物如实施例1进行准备。ABC1 mRNA的量可以由RT-PCR进行测量。

细胞培养物: 正常皮肤成纤维细胞及TD1及TD2成纤维细胞的灭菌细胞培养物可以如实施例1进行准备。细胞在DMEM/10% FBS 培养物下分会合培养, 然后24或48小时后使用DMEM/BSA及指示添加剂替代该培养物。RNA如实施例2进行准备。

RT-PCR: 使用GeneAmp 5700序列检测系统 (Perkin-Elmer应用生物系统, 福斯特城, CA)进行定量PCR。简而言之, 500ng脱氧核糖核酸酶处理过的mRNA使用任意的六邻体引物以2.5 $\mu$ M的浓度进行逆转录。这种反应大约有5%可以使用SYBR绿核试剂盒 (PE Applied Biosystems, Foster City, CA; Cat. #4304886) 及人类ABC1引物LF: 5'-CCTCTCATTACACAAAAACCAGAC (SEQ ID NO: 11) 及 LR: 5'-GCTTTCTTTCACTTCTCATCCTG (SEQ ID NO: 12)通过PCR进行扩增以产生一个82 bp的片段其相应于人类ABC1的7018至7099的核苷酸位点。PCR循环条件如下: 10 分钟95 $^{\circ}$ C; 然后在95 $^{\circ}$ C下40个循环, 15秒钟; 及60 $^{\circ}$ C, 60秒钟。每一样品中的mRNA通过检测由结合至产生于每一循环的PCR双链扩增产品的SYBR绿色所致荧光的增加来量化检测。所有样品都以三倍的浓度运行且使其规范为 $\beta$ 球蛋白mRNA,每一平行反应扩增都使用到引物肌动蛋白F序列: 5'-TCACCCACACTGTGCCATCTACGA (SEQ ID NO:54)及肌动蛋白B序列: 5'-CAGCGGAACCGCTCATTGCCAATGG (SEQ ID NO: 55)。在相同的PCR平板中ABC1及 $\beta$ 球蛋白的标准曲线同时绘制。

8-Br-cAMP试验: 正常, TDI,及TD2成纤维细胞在DMEM/10% FBS 培养物下分会合培养, 然后使用DMEM/BSA中的1mM8-Br-cAMP对其进行处理24小时。

5 胆固醇测试: 正常, TDI,及TD2成纤维细胞在DMEM/10% FBS 培养物下分会合培养, 然后使用DMEM/BSA中的30 $\mu$ g/ml自由胆固醇对其进行处理48小时。然后在DMEM/BSA 下平衡18-24小时。

APO A-I 测试: 正常, TDI,及TD2成纤维细胞在DMEM/10% FBS 培养物下分会合培养, 然后使用DMEM/BSA中的30 $\mu$ g/ml自由胆固醇对其进行处理48小时。然后在DMEM/BSA及10 $\mu$ g/ml apo A-I下平衡18-24小时。

10 结果: 图8显示的是在正常成纤维细胞中ABC1 mRNA通过暴露于8-Br-cAMP增加将近10倍且在暴露于不含血清的培养基中的胆固醇其增加将近17倍。随后暴露于Apo A-I的胆固醇装载细胞的结果在于ABC1 mRNA表达量的显著减少。尽管其机理尚未清楚, 先前的研究已经显示胆固醇排放量在下列化合物的存在下是得以提高的: 例如有cAMP及胆固醇 (Hokiand 等, *J*  
15 *Dial. Chem.*, 268:25343-25349 (1993))。现在的结果显示在正常成纤维细胞中, ABC1 mRNA可以由这些已知的胆固醇排放物路径效应物进行诱导且因暴露于阿朴脂蛋白胆固醇受体而被抑制。其证明了ABC1的表达为相关于阿朴脂蛋白介导的胆固醇排放物相关的细胞环境所调控。在对照实验中, 来自TD患者的成纤维细胞并不为胆固醇排放物的效应物所调控。一开始, ABC1mRNA在  
20 TD1及TD2细胞中的cAMP诱导水平仅仅是相当于正常细胞的40%, 更进一步, 暴露于Apo A-I的胆固醇载入细胞或者并未改变ABC1的表达 (TD1细胞) 或者轻微增加了ABC1的表达 (TD2细胞)。这些结果反映出TD细胞中Apo A-I介导的胆固醇排放物量的缺损。有趣的是, 含血清培养基中的细胞生长抑制了ABC1将近可检测值的极限 (数据并未显示)。这些可以反映出这样的事实  
25 油脂排放物路径的功能要求细胞处于休眠期及其他的细胞状态其减少了细胞

的胆固醇需求量。因此，与细胞胆固醇排放物量增加相关的条件（也就是说，胆固醇载入，cAMP处理，血清驱除）等同样使得在正常成纤维细胞中ABC1 mRNA的表达量增加。相反地，暴露于apo A-I的载入胆固醇的正常成纤维细胞减少了ABC1的表达量。

## 5 实施例10

该实施例证明了LXR细胞核受体配体,例如20-羟基胆固醇, 及RXR受体配体例如9-顺式视黄酸 可以在鼠RAW 264.7细胞中提高ABC1基因的表达量。

LXR细胞核受体是一个形成与核受体RXR异二聚体的转录因子且可以通过连接至一类包括22-羟基胆固醇及20-羟基胆固醇 (Janowski 等, 自然, 383:728-731 (1996))而被激活以增强其目的基因的转录。同样地,其也为胆固醇诱导的基因转录的调控的候选物质。更进一步, 在显示在成纤维细胞及巨噬细胞对胆固醇载入所起反应而引起ABC1 mRNA及蛋白的增加的研究及另外的显示LXR及RXR在胆固醇载入巨噬细胞中因暴露于氧化的LDL而表达量增加的研究的启示下, LXR 及RXR核受体作为ABC1基因转录激活的最为推荐的候选物质。为了确定是否LXR及RXR受体在ABC1基因表达中起到一定的作用, 进行测量反应于20-羟基胆固醇及9-顺视黄酸的ABC1 mRNA水平。

鼠RAW 264.7细胞在DMEM/10% FBS 培养物下分会合培养, 然后使用不含血清的DMEM/BSA及或者9-顺视黄酸 (10 $\mu$ M),20-羟基胆固醇 (10 $\mu$ M)或两者一起 (总共20 $\mu$ M)对其进行处理24小时。对照细胞仅近使用乙醇载体进行处理 (0.1% v/v)。抽提RNA, 使用脱氧核糖核酸酶进行处理, 然后如实施例9所述的使用PE 生物系统 SYBR 绿色技术进行RE-PCR对ABC1 mRNA进行测量。图9显示的是使用20-羟基胆固醇或者9-顺视黄酸处理的结果是ABC1 mRNA 表达量的增加。此外, 图9还显示的是二者共同处理的结果在于显著的增强效应, 其大约比在上述单独处理的ABC1表达量增强将近6倍。这些结

果证明了核受体LXR及RXR的连接可以提高ABC1基因的表达量。

#### 实施例11

该实施例证明了质膜中ABC1蛋白的增加表达与脂质排放相关。

使用细胞表面标记及免疫沉降反应来确定是否在细胞质膜中发生了  
5 ABC1蛋白的增加表达，而这些是与胆固醇排放量相关的 (图10)。细胞表面  
ABC1的相对数量通过在完整细胞中表面蛋白与膜不渗透性试剂磺基-NHS-  
生物素的交联，然后的步骤是膜溶解，使用ABC1抗体进行免疫沉降，  
SDS-PAGE及用链霉亲核素检测等来确定。

细胞培养物: 正常及TD1成纤维细胞如实施例1所述进行无限增殖。正常  
10 及TD1细胞一起在对照条件下及在已知可以增加阿朴脂蛋白介导的胆固醇排  
放物的条件 (Oram, 等, J. Lip. Res., 40:1769-1781 (1999))下进行培养。对照  
细胞然后在DMEM/ 10% FBS下汇合培养然后在无添加剂的存在下在  
DMEM/BSA下孵育18小时 (对照组)。cAMP-处理的细胞在DMEM/10% FBS  
下汇合培养, 然后在DMEM/BSA下与1mM 8-Br-cAMP (cAMP)一起孵育18小  
15 时。载满胆固醇的细胞在DMEM/10% FBS下汇合生长然后与30 $\mu$ g/ml胆固醇  
一起孵育48小时及在没有添加剂 (胆固醇)下孵育18小时。使用cAMP处理的  
载满胆固醇的细胞在DMEM/10% FBS下汇合生长然后与30 $\mu$ g/ml胆固醇一起  
在DMEM/BSA下孵育48小时及在1mM 8-Br-cAMP (胆固醇 + cAMP)下处理  
18小时。

20 细胞表面标记: 为了选择性标记质膜ABC1, 细胞在0 $^{\circ}$ C 下使用含有1  
mg/ml sulfo-NHS-生物素 (Pierce, Rockford, IL; Cat. #21217)的PBS至生物素  
酰化的细胞表面蛋白 (参见Walker等., 生物化学, 50:14009-14014 (1993))孵  
育30分钟。

免疫沉降反应: 抗相应于一个位于人类ABC1的C末端的诱导多肽:  
25 KNQTVVDAVLTSFLQDEKVKES的合成多肽的ABC1的兔抗血清得以提高。

在含有1% Triton X-100 (Sigma, St. Louis, MO) 及蛋白酶抑制剂亮抑蛋白酶肽 (1 mM), 抑肽素 (1mM),及抑蛋白酶肽 (1mM)中溶解该细胞然后进行免疫沉降反应。细胞抽提物在4<sup>0</sup>C下与抗ABC1抗血清一起以1:200的稀释度下过夜培养, 然后与5 $\mu$ l 蛋白A包装的磁性小珠 (Dyna, Lake Success, NY; Cat. #1001.01)一起额外孵育1小时。抗原抗体复合体使用磁体进行沉淀, 小珠使用1% TritonX/PBS 洗涤两次, 然后用1%的乙酸洗涤该蛋白。

ABC1蛋白的检测: 抽提的生物素酰化的蛋白进行SDSPAGE (6%凝胶; 150V, 5小时)及转移至一个硝化纤维膜上 (200mA, 18小时)。硝化纤维使用抗生物素蛋白链菌素-辣根过氧化物歧化酶 (Amersham Pharmacia, Piscataway, NJ; Cat. #RPN 123 1) 稀释至300倍进行探针标记且按照操作手册 (Amersham Pharmacia, Piscataway, NJ)通过放大的化学光标记 (ECL)进行检测。为了测试可能的生物素酰化的胞内蛋白, 免疫沉降后上血清使用鼠单克隆抗体进行处理至一胞内蛋白 $\beta$ -COP 且免疫沉降的生物素酰化的 $\beta$ -COP 使用链霉抗生素蛋白染色印迹进行测试。结果什么也没检测到。

结果: 正如图10所示那样, 240 KDa ABC1蛋白以成对形式出现。ABC1蛋白部分位于正常 (10A) 及TD1 (10B) 成纤维细胞中的质膜上。相似的结果出现于另一个正常的成纤维细胞系及TD2成纤维细胞 (数据并未出现)。ABC1在当细胞在血清 (正常及TD1细胞) 中生长时轻微增强表达的细胞表面使用8-Br-cAMP 进行处理。正常及TD1细胞中血清损失及胆固醇载入明显地增强了细胞表面ABC1的表达,其也可以进一步通过cAMP 处理进行增强。这些结果表明在细胞表面ABC1的表达为其可以增强阿朴脂蛋白介导的脂排放量的条件所调控, 与其一致的是, 其在质膜上的位点对于脂运输功能起到一定的作用。发生在TD1及TD2细胞中的突变并未表现出ABC1处理及表达的严重受损。其意味着脂运输的次效应或与赖于其发生突变的NH<sub>2</sub>末端区域的辅助蛋白的相互作用。

## 实施例12

该实施例证明了可以抑制cAMP 5'及3'端降解的试剂,例如,磷酸二酯酶抑制剂,其可以提高阿朴脂蛋白A-I-介导的从巨噬细胞中排放物量。

正如图10所示, cAMP提高了ABC1的活性。进行本研究以确定在提高的cAMP的存在下从巨噬细胞中的胆固醇排放物量。cAMP水平的提高可以获自下列试剂的存在, 其可以刺激cAMP的合成或者抑制cAMP的降解。例如, rolipram是一种可以通过抑制磷酸二酯酶, 一组可以降解cAMP的酶来调控cAMP水平的化合物。提高的cAMP 对胆固醇排放物量的效应可以通过如实施例1所示的阿朴脂蛋白介导的胆固醇排放物量测试实验进行确定。简而言之, 悬浮于一浓度为 $1.25 \times 10^5$ 细胞/ml 的RAW 264.7细胞生长于一个丙酮酸酯补充的DMEM/10% FBS中。24小时后, 移去培养基替代以DMEM/BSA 及放射性标记的胆固醇 ( $1 \mu\text{Ci/ml } ^3\text{[H]-胆固醇}$ ) 及 $50 \mu\text{g/ml}$  乙酰化的LDL培养24小时。细胞然后保存在含有DMEM/BSA 及或者apo A-I 单独 ( $20\mu\text{g/ml}$ ), apo A-I 及 8-溴 3',5'cAMP ( $1\text{mM}$ ) 或者apo A-I 及 rolipram ( $50\mu\text{M}$ )的平衡培养基中保持24小时。经过12-24小时后,  $^3\text{[H]-胆固醇}$ 排放物量通过如实施例1所述的闪烁计数进行测量且计算为在培养基中出现的总的荧光标记的胆固醇的百分比。该结果表明没有接收到apo A-I的载满胆固醇的对照细胞表现出3%的胆固醇排放物量, 而接收到apo A-I的细胞仅表现出5%的胆固醇排放物量。接收到apo A-I及cAMP的载满胆固醇的细胞表现出32%的胆固醇排放物量, 证明了提高的cAMP提高了胆固醇排放物量。同样地, 接收到apo A-I及磷酸二酯酶抑制剂 (rolipram)的载满胆固醇的细胞表现出17%的胆固醇排放物量。

## 实施例13

该实施例证明了可以为核受体配体的试剂例如LXR, RXR, 及PRAR核受体可以提高阿朴脂蛋白A-I-介导的从巨噬细胞中排放物量。

为了确定是否核受体配体作用于阿朴脂蛋白介导的胆固醇排放物量,使

用如实施例12所示的apo A-I 介导的胆固醇排放量测试实验测试不同的配体。核受体家族包括有不同的成员,例如有肝受体LXR, 维生素A受体RXR, 及过氧化物酶体增殖剂活化的受体PPAR,其已被包含进入类脂化合物的代谢过程当中了 (Russell. D.W., 细胞, 97:539-542 (1999); Spiegelman, B.W., 细胞, 5 93:153-155 (1998); Janowski等, 自然, 383:728-731 (1996))。更进一步, 一些受体的配体已被发现可以增加血浆 HDL且基因表达(微阵列)数据已显示出对经过氧化LDL载入胆固醇反应的激素受体。按照上述试验, 9-顺视黄酸(RXR配体), 氧甾醇 (LXR 配体),及 fenfibrate (PPAR配体)被测试以确定其对胆固醇排放物的作用, 没有收到apo A-I的载满胆固醇的对照细胞显示出3 10 %的胆固醇排放量, 而接收到apo A-I的细胞显示出5%的胆固醇排放量。与之形成对比的是, 接收到apo A-I 及 9 顺视黄酸 (30 ng/ml) 的载满胆固醇的细胞显示出16%的胆固醇排放量。接收到apo A-I 及氧甾醇 (5  $\mu$ g/ml) 的细胞显示出14%的胆固醇排放量。接收到apo A-I 及fenfibrate (3 $\mu$ g/ml) 的细胞显示出10%的胆固醇排放量。这些结果表明,激素受体可以被调整以提高阿朴脂蛋白介导的从巨噬细胞中的胆固醇排放的速率。 15

更进一步, 当使用不同浓度的9-顺-RA (0.3ng/ml, 3.0ng/ml, 或者30ng/ml) 进行排放量测试时,结果显示9-顺-RA介导的从巨噬细胞中的胆固醇排放量是以药剂依赖性的形式.更明确一点讲,对照细胞 (仅apo A-I)显示出1890 c.p.m., 0.3 ng/ml 9-cis-RA 显示为 1522 c.p.m., 3.0 ng/ml 9-cis-RA 显示为 20 3568 c.p.m.及30 ng/ml 9-cis-RA 显示为 8597 c.p.m。此外, 使用相似的实验其中RAW 264.7细胞载入胆固醇48小时, 其他的核受体激活剂, 例如, 22-羟基胆固醇 (LXR配体) 及benzfibrate,显示出其可以提高胆固醇排放量 (数据并未显示)。

#### 实施例14

25 该实施例表明类二十烷酸,例如前列腺素E1及环前列腺素PGI<sub>2</sub>,可以提高

阿朴脂蛋白介导的从巨噬细胞中的排放物量。

类二十烷酸，例如前列腺素及环前列腺素，已被发现对于治疗血胆固醇过多症有一定的疗效。为了明确是否类二十烷酸作用于阿朴脂蛋白介导胆固醇排放物量，可以进行如实施例12所示的apo A-I介导的胆固醇排放测试实验测试PGE 1及 PG 12。该测试表明没有接收到apo A-I 的载满胆固醇的对照细胞表现出3%的胆固醇排放物量，而接收到apo A-I的细胞仅具有5%的排放物量。接收到apo A-I 及PG12 (25 nm)的载满胆固醇的细胞表现出10%的胆固醇排放物量。而接收到apo A-I 及PG1 (25 nm)的细胞表现出15%的胆固醇排放物量。这些结果证明类二十烷酸可以提高阿朴脂蛋白介导的从巨噬细胞中的胆固醇排放的速率。

#### 实施例15

该实施例证明了在ABC1启动子的控制下报道基因可被用于测试化合物调控ABC1基因表达的能力。

pGL3 荧光素酶记录载体系统 (Promega,麦迪逊, WI) 被用于产生一重组质粒来测量在ABC1启动子的控制下报道基因的表达。

报道质粒的构建: 质粒pGL3-Basic (Promega, Madison, WI; Cat. #E 1751) 被用作含有不含启动子的荧光素酶基因的对照质粒。含有ABC1启动子及荧光素酶基因的报道构建体通过无性繁殖一来自ABC1基因 (hAPR1 5' 启动子, 相应于SEQ ID NO: 3的核苷酸1080-1643) 5' 侧翼区的基因组片段的克隆至一个GL3-基础的质粒的SacI位点以产生一个质粒GL-6a制得。下一步, 使用SpeI 及 Acc65I消化质粒GL-6a。一个切割自一个λ亚克隆的BsiWI-SpeI片段,其代表了一个相应于SEQ ID NO: 3的核苷酸1-1534 的ABC1基因组序列被结合至由该切割所产生的剩余的载体/ABC1启动子片段。结果的质粒, pAPR1,在1.75 kb 的人类ABC1启动子序列的转录控制下编码荧光素记录蛋白基因。

报道构建体的转染: 上述的控制或者pAPR1质粒被转染进入保存在含有

10%的牛胎血清的DMEM 中的RAW 264.7细胞的汇合培养物中。 一个12孔碟子的每一孔中使用或者PGL3-Basic, pGL3-启动子或者pAPR1 DNA (1 $\mu$ g), 荧光素酶质粒DNA (1 $\mu$ g), 及12  $\mu$ l Geneporter试剂 (基因治疗系统, 圣地亚哥, CA; Cat~ #201007)转染5小时。此外, 0.1  $\mu$ g的pCMV $\beta$ 质粒DNA (Clontech, Palo Alto, CA, Cat. #6177-1)被加入以作为转染效率的对照。5小时后, 在乙酰化LDL (100  $\mu$ g/ml)的存在或不存在下的不含血清的DMEM/BSA取代该培养基且孵育24小时。

考虑到大量筛选的便利性,培养细胞使用报道质粒按照下列程序可以被稳定转染。首先, 5 $\times 10^6$  RAW 264.7细胞在一个60mm的碟中使用9 $\mu$ g pAPR1质粒及在一个10 ml 不含血清的DMEM中使用50 $\mu$ l Geneporter转染试剂 (基因治疗系统, 圣地亚哥, CA)的pCMVscript (Stratagene, Lalolla, CA)中转染5小时。其结果是,转染培养基被完全培养基所取代, 且细胞在37 $^{\circ}$ C下过夜培养。随后细胞被转移到另一个盘中其稀释范围为1:5到1:1000且孵育在含有800  $\mu$ g/ml G418 (Life Technologies, Bethesda, MD)的选择性培养基中孵育20天。挑选可视的菌落, 扩大培养, 然后如下述方法进行荧光素酶活性的测试。

使用这种方法,所鉴别出来的荧光素酶活性呈阳性的5个无性繁殖细胞系被用于进行高量测试。

荧光素酶测试: 转染后,每一孔中的细胞溶解于 70 $\mu$ l 的 1X细胞溶解试剂 (Promega, Madison, WI, Cat. #E397 1), 然后进行一个冷冻-溶解循环, 溶解产物通过在12,000g 下离心5分钟进行清除。离心后, 100 $\mu$ l的荧光素酶测试试剂 (Promega, Madison, WI; Cat. #E1501) 被加入到10  $\mu$ l 的细胞溶解产物中去。每一溶解产物的荧光素酶活性使用照度计以发光单位进行测量。此外, 每一溶解产物的 $\beta$ -牛乳糖活性使用化学发光试验试剂 (其由Galacto-光试剂盒) 按照生产者的指导 (Tropix Inc., Bedford, MA: Cat. #BL100G)进行测量。每一溶解产物的规范化的荧光素酶活性值可以通过确定的 $\beta$ -牛乳糖除以

荧光素酶活性值且记录为相对光量单位。

结果：使用pAPR1转染的细胞中检测到的荧光素酶活性是使用仅含有荧光素酶cDNA 的基于pGL3质粒转染的对照细胞中荧光素酶活性的3.3倍。这些结果表明ABC1的转录调控区域已被取代。 当pAPR1转染细胞在100 µg/ml 乙酰化的LDL中转染24小时后,其荧光素酶活性比未经乙酰化LDL处理的细胞中的荧光素酶活性高出3.25倍。这些结果显示基因组ABC1序列含有一个“胆固醇 反应性”元素其发现于5'侧翼区其可以调控野生型ABC1基因的胆固醇载入调控。这些记录系统同样可被用于测试其他的化合物以确定是否该化合物可以调控ABC1的表达。

#### 10 实施例16

该实施例证明了另外一个可因其能够在内源性ABC1启动子的控制下使用报道基因进行调控ABC1基因表达而被用于进行化合物测试的试验。

该测试涉及到构建一个重组质粒载体其含有一个无启动子的报道基因及一个选择性标记基因。该载体被线性化且转染进入一个报道基因被整合进入内源ABC1启动子的细胞基因组下游的细胞。使用这种测试,报道基因的表达由内源ABC1启动子反应于一个测试化合物而驱动。

报道质粒的构建：含有一个无启动子的报道基因的重组载体被制成一个起始为一个含有外显子0 (其包括ABC1 起始位点)及部分内含子1的的ABC1基因的7 kb 的EcoRI 基因组片段。

20 使用定点诱变, 一个SaII限制性位点可以产生于两个已知起始位点的下游内含子0序列。重组载体通过插入一个含有无启动子报道基因的DNA片段例如荧光素酶, 及一个无启动子选择性标记例如嘌呤霉素抗性至一个SaII位点而产生。一个内源的核糖体进入信号可以插入到报道基因和标记基因之间以至于该基因能够被以正确的方向进行转录。含有两个Eco47III 位点的重组载体, 其中一个被删除, 使用, 例如定点诱变。剩余的Eco47III 位点, 位于

内含子0的上游位点，被用于线性化该载体。

报道构建体的转染：含有报道基因及标记基因的线性化的重组载体被引入重组细胞，这些细胞包括人类细胞，使用的转染方法是本领域的公知技术。例如，线性化的载体可以使用如实施例15所示的方法进行转染。线性化的重组载体含有允许该载体在内源ABC1基因位点整合进入细胞基因组中的ABC1序列。在细胞培养基中加入的合适的抗生素允许只有这些细胞能够进行选择其中报道基因及标记基因已被以正确方向整合进入内源ABC1启动子的下游位点。例如，如果该载体含有一个嘌呤霉素抗性基因插入到报道基因的下游，转染细胞应该可以在嘌呤霉素的存在下生长。只有这些细胞具有一个正确整合进入的嘌呤霉素抗性基因，且可以因此编码一个功能性的蛋白，其可以在嘌呤霉素的存在下存活。因此，转染细胞应该在这样的条件下生长其可以在合适的抗生素的存在下诱导ABC1启动子的活性。存活细胞能够被无限增殖培养且使用PCR进行DNA测序或使用DNA印迹分析以测试基因组序列是否被正确整合。

结果产生的含有一个报道基因的细胞在内源ABC1启动子的控制下可被用于确定是否测试试剂可以调控ABC1的表达。ABC1调控化合物的活性可以通过分析暴露于测试试剂的化合物的细胞中的报道基因的表达水平而确定。例如，整合有荧光素酶基因的细胞可被用于确定ABC1调控一个测试化合物的活性，其方法是通过测定暴露于该试剂的细胞中发现的荧光素酶活性的数量来进行。

### 实施例17

该实施例证明了核受体配体在ABC1启动子的控制下正调报道基因的表达。

为了确定是否LXR- $\alpha$ ，LXR- $\beta$  及 RXR $\alpha$ 核受体的配体能够调控ABC1基因的表达，在ABC1启动子的控制下包含有荧光素报道基因的pAPR1质粒被转

染进入使用至少一种该核受体的配体处理过的RAW 264.7细胞 (图12).

报道构建体的构建与转染: 报道构建体 pAPR1及对照的报道构建体 pGL3-Basic 可以如实施例15获得. RAW 264.7 细胞保存在培养物中且使用或者pGL3-Basic (1 $\mu$ g)或 pAPR1 (1 $\mu$ g)如实施例15所述进行转染。转染的RAW  
5 264.7细胞使用或者乙醇 (EtOH) (0.1% v/v), 20 (S)-羟基胆固醇 (20 (S) OH-chol) (10 $\mu$ M), 9-*cis*视黄酸 (9-*cis* RA) (10 $\mu$ M) 或者20 (S) OH-chol 及 9-*cis* RA (总共20 $\mu$ M)二者一起处理24小时。测量荧光素酶活性且以实施例15所述相对光单位记录。

结果: 该研究的结果显示于图12。使用PGL3-Basic转染的对照组细胞并未  
10 未显示出荧光素活性 (并未显示出数据)。使用pAPR1转染的细胞在20 OH-chol的存在下荧光素酶报道活性增加了19倍。在9-*cis* RA的存在下荧光素酶报道活性增加了16倍。在与对照组EtOH 相比较的两个配体的存在下荧光素酶报道活性增加了280倍。这些结果显示固醇及维生素A二者都可以在pAPR1.中从ABC1 5'侧翼序列诱发一个转录反应。更进一步讲, 这两组化合物  
15 具有明显的协同效应, 其可以由经过两种配体一起处理的细胞中发现的荧光素酶活性的显著增强而看到。已知LXR- $\alpha$ 及RXR $\alpha$ 受体形成活性异聚体。这样, 两种核受体同时产生的配体诱导活性可以在转录过程中发现其协同性的增强。

这些数据证明了羟基固醇,例如20 (S)羟基胆固醇及维生素A例如9-顺视  
20 黄酸,活化ABC1启动子, 表明这些及相关的试剂可用于在巨噬细胞中增强ABC1的表达的治疗性化合物的开发以除去周围剩余的胆固醇。此外,该ABC1启动子/报道基因筛选方法可被用于筛选其它试剂其可以增强ABC1的表达以开发出更好的治疗性化合物。

#### 实施例18

25 该实施例证明了ABC1启动子区域的进一步的特征,包括LXR 反应元素

的鉴别。

5 为了确定ABC1的5'侧翼区的哪一部分保持有响应于核配体的转录活性，含有5'侧翼区的不同部分及荧光素酶的报告基因的不同的质粒被转染进入到使用至少一种核受体配体进行处理的RAW 264.7 细胞中去。使用这种体系，可以鉴别出一个相应于SEQ ID NO: 3的核苷酸1480-1510位点的固醇反应元素。该固醇反应元件包括一个直接重复序列的4元件TGACCGatagTAACCT。固醇反应元件使用定点诱变及接合转移试验技术进行确认。

报道构建体的构建: 报道构建体pAPR1及对照的报道构建体pGL3-Basic如实施例15那样获取。含有或者SEQ ID NO: 3的核苷酸1-1532, 1080-1643, 10 1181-1643, 1292-1643, 或1394-1643 的报道构建体同样被构建。一个含有SEQ ID NO: 3 的1080-1643 位核苷酸的报道构建体 (GL-6a)如实施例15进行构建。一个含有SEQ ID NO: 3 的1-1532 位核苷酸的报道构建体使用Spe I及Nhe I消化pAPR I及凝胶载体片段的重新结合进行构建。含有1181至1643核苷酸的报道构建体通过使用StvI 首次消化GL-6a进行构建,使用Klenow酶对粘 15 末端进行钝化,使用Sac I消化结果产生的载体且分离出462个碱基对的钝化的Sac I粘末端片段。然后其被克隆进入一个由使用Acc65 I消化GL-6a 细胞所获取的载体。使用Klenow酶对粘末端进行钝化,使用Sac I消化及凝胶分离结果产生的载体片段。一个含有1292-1643位核苷酸位点的报道构建体通过使用Acc65 I 连续消化GL-6a 进行构建。使用Klenow酶对末端进行钝化,使用Sac 20 II消化T4聚合酶钝化的末端且重新连接至一个凝胶分离的载体片段上。含有核苷酸1394-1643位点的报道构建体通过使用Acc65 I 消化GL-6a 进行构建,使用Klenow酶对末端进行钝化,接着使用Apa I 进行消化, T4聚合酶钝化末端, 且重新连接至一个凝胶分离的载体片段上。

报道构建体的转染: RAW 264.7 细胞保持在这样的一个培养物中且使用 25 或者pGL3-Basic (1µg), pAPR 1 (1µg)或如实施例15所述方法其他任意一种报

道构建体进行转染。已转染的RAW 264.7细胞使用或者乙醇 (EtOH) (0.1% v/v), 20 (S)-羟基胆固醇 (20 (S) OH-chol) (10 $\mu$ M), 22 (R)-羟基胆固醇 (22 (R) OH-chol) (10 $\mu$ M), 9-顺视黄酸 (9-cis RA) (10 $\mu$ M), 或20 (S) OH-chol及9-cis RA (总共20 $\mu$ M)一起处理24小时。测量荧光素酶活性然后记录为如实施例15所述

5 的相对光单位。

定点诱变: 相应于如SEQ ID NO: 3所示序列的1480-1510的核苷酸位点的固醇反应元件如上述使用定点诱变的方法进行突变。具体地讲, 含有直接重复序列4元件TGACCGatagTAACCT 的反应元件使用GeneEditor系统按照生产商提供的操作程序被突变为CTGCACatagTAACCT 。

10 凝胶阻滞测试: 使用如Ohisson 等, *细胞*, 45:3544 (1986)所使用的方法从RAW 264.7 细胞中准备核苷酸抽提物。相应于LXR反应元件(TCGAGTGACCGATAGTAACCTCTCGA; SEQ ID NO: 56)及其突变配对子(TCGAGCTGCACATAGTAACCTCTCGA; SEQ ID NO: 57)的<sup>32</sup>P-标记的寡聚核苷酸 (5 ng)分别使用5 $\mu$ g核内蛋白在20mM HEPES, pH 7.9, 60mM KCL,

15 1mM Mgcl<sub>2</sub>, 1mM DTT, 66.6 $\mu$ g/ml poly (dIdC);及10%甘油在至LXR $\alpha$  及LXR $\beta$  (Santa Cruz Biotechnology, Cat. No. SC-1591, Santa Cruz, CA)的1 $\mu$ g的抗血清或至RXR (Santa Cruz Biotechnology, Cat. No. SC-774, Santa Cruz, CA)的抗血清存在或不存的情况下在室温下孵育30分钟。蛋白-DNA复合体在4% 非变性的聚丙烯酰胺凝胶下在150V在0.5X TBE缓冲液的存在下1.5小时。蛋白

20 -DNA复合体通过干凝胶自动射线照相术进行检测。

结果: 使用含有相应于 (即, pAPR1)的核苷酸位点1-1643,及SEQ ID NO: 3的核苷酸位点区1-1532, 1080-1643, 1181-1643, 1292-1643,或 1394-1643的5' 侧翼区的不同的报道构建体进行的转染均产生相同的结果。每一构建体在20 (S) OH-chol 或 22 (R) OH-chol的存在下相对于对照组的EtOH均产生3-4

25 倍增加的荧光素酶报道活性。同样, 每一构建体在9-cis RA 的存在下产生一

个8-10倍增加的荧光素酶报道活性。此外，使用任一构建体进行的转染在甾醇配体 (either (20 (S) OH -Chol或22 (R) OH-chol))及类维生素A配体 (9-cis RA) 一起存在的情况下相比较EtOH 对照组荧光素酶活性增加了25至50倍。这表明其相互之间产生了协同相互作用。上述每一构建体均证明了相应于配体测试的荧光素酶活性的比较水平。其表明即使较短的5'侧翼区也包含有固醇及类维生素A的转录调控序列。更明确地讲，这些结果证明了固醇及类维生素A的转录调控序列位于相应于如SEQ ID NO: 3所示序列的核苷酸位点1394-1532位点的5'侧翼区。

这些结果通过使用含有相应于如SEQ ID NO: 3所示序列的1080-1643 核苷酸位点的野生型序列的报道构建体及含有相应于如SEQ ID NO: 3所示序列的1080-1643 核苷酸位点的突变型序列的报道构建体进行的荧光素酶活性测试可以进一步确认，其中发现位于核苷酸1480-1500位点的固醇反应元件如上述发生了突变。使用野生型序列进行的转染结果产生了一个转录反应，其可以通过在或者20 (S) OH-chol 或9-cis RA分别存在的情况下荧光素酶活性的增加（其在两种配体共同存在的情况下产生一种协同效应）进行测量。与之相比较的是，使用突变序列进行的转染并未在20 (S) OH-chol 或者22 (R) OH-chol的存在下产生一个转录反应。突变序列的转染结果是对9-cis RA 的反应降低,其相对于所观察到的野生序列转染所产生的荧光素酶活性的增加8-10倍而仅增加了4-5倍。突变序列的转染同样抑制了在20 (S) OR-chol 及9-cis RA一起存在的情况下所观察到的协同增效的转录反应。这些结果可以由使用固醇契合序列 (核苷酸 1480-1510)及其突变的配对物一起进行的凝胶阻滞测试进一步证实。凝胶阻滞测试显示出当分离自RAW 264.7细胞的核结合蛋白被结合在固醇契合序列上而其并未结合在突变序列上。更进一步，使用野生型固醇契合序列在LXR抗血清的存在下核蛋白的孵育结果产生超级阻滞复合体的形成 (也就是抗体-蛋白-DNA复合体)，其表明作为固醇反应元件的序列结

合在核受体LXR上，作为一种对比，核蛋白与野生型固醇反应元件在RXR抗血清的存在下一起孵育并未产生该超级阻滞复合体的形成，其表明RXR并未结合在该序列上。这些结果都显示出破坏了核蛋白结合至契合序列上的突变同样也抑制了对LXR配体的转录反应及减少了对RXR配体的转录反应。更进一步，在LXR及RXR抗血清的存在下的核体结合研究证实了发现于核苷酸位点1480-1510的契合序列是一个LXR反应元件，该元件同样调控对9-cis RA的部分反应。

这儿所引用的所有参考书目，包括专利文献及出版物都全部被引入作为参考资料。尽管这儿仅描述了本发明的重点部分,但是很显然要实现本发明也会用到本领域的许多公知技术,本发明的实施例进一步扩张了本发明。当然并不限于这儿所描述的内容。

本发明包括所有对其发明点中心思想的修饰，本发明的保护范围如下述的权利要求进行界定。

<110> 罗恩, 理查德 M.

维德, 大卫

加维, 迈克尔

奥曼, 约翰 F.

<120> 使用ATP结合框运输蛋白ABC1增加胆固醇排放并提高HDL的组合物及方法

<130> 99,395-A

<140> 09/595,526

<141> 2000-06-16

<150> US 60/140,264

<151> 1999-06-18

<150> US 60/153,872

<151> 1999-09-14

<150> US 60/166,573

<151> 1999-11-19

<160> 57

<170> PatentIn Ver. 2.0

<210> 1

<211> 10442

<212> DNA

<213> Homo sapiens

<220>

<221> 未确定

<222> (1) .. (10442)

<223> 所有的n都是未知的.

<400> 1

```

ggccgggacc cgagagccg agccgacct tctctcccgg gctgcccag ggcagggcgg 60
ggagctccgc gcaccaacag agccggttct cagggcgtt tgctccttgt tttttcccgg 120
gttctgtttt ctccccttct ccggaaggct tgtcaagggg taggagaaag agacgcaaac 180
acaaaagtgg aaaacagtta atgaccagcc acggggcgtc ctgctgtgag ctctggccgc 240
tgccctccag ggctcccag ccacacgctg ggcgtgctgg ctgagggaac atggcttggt 300
ggcctcagct gaggttgctg ctgtggaaga acctcacttt cagaagaaga caaacatgtc 360

```

agctgttact ggaagtggcc tggcctctat ttatcttctc gatcctgatc tctgttcggc 420  
 tgagctaccc accctatgaa caacatgaat gccattttcc aaataaagcc atgccctctg 480  
 caggaacact tccttggggtt caggggatta tctgtaatgc caacaacccc tgtttccggt 540  
 acccgactcc tggggaggct cccggagttg ttggaaactt taacaaatcc attgtggctc 600  
 gcctgttctc agatgctcgg aggettcttt tatacagcca gaaagacacc agcatgaagg 660  
 acatgcgcaa agttctgaga acattacagc agatcaagaa atccagctca aacttgaagc 720  
 ttcaagatth cctgggtggac aatgaaacct tctctggggtt cctatatcac aacctctctc 780  
 tcccaaagtc tactgtggac aagatgctga gggctgatgt cattctccac aaggatthtt 840  
 tgcaaggcta ccagttacat ttgacaagtc tgtgcaatgg atcaaaatca gaagagatga 900  
 ttcaacttgg tgaccaagaa gtttctgagc tttgtggcct accaaaggag aaactggctg 960  
 cagcagagcg agtacttctg tccaacatgg acatcctgaa gccaatcctg agaacactaa 1020  
 actctacatc tcccttcccg agcaaggagc tggctgaagc cacaaaaaca ttgctgcata 1080  
 gtcttgggac tctggcccag gagctgttca gcatgagaag ctggagtgc atgacagcag 1140  
 aggtgatggt totgaccaat gtgaacagct ccagctctc cacccaaact taccaggctg 1200  
 tgtctcgtat tgtctgctgg catcccagg gaggggggct gaagatcaag tctctcaact 1260  
 ggtatgagga caacaactac aaagccctct ttggaggcaa tggcactgag gaagatgctg 1320  
 aaaccttcta tgacaactct acaactcctt actgcaatga tttgatgaag aatttggagt 1380  
 ctagtctctt ttcccgcatt atctggaaag ctctgaagcc gctgctcgtt ggaagatcc 1440  
 tgtatacacc tgacactcca gccacaaggc aggtcatggc tgaggtgaac aagaccttcc 1500  
 aggaactggc tgtgttccat gatctggaag gcatgtggga ggaactcagc cccaagatct 1560  
 ggaccttcat ggagaacagc caagaaatgg accttgtccg gatgctgttg gacagcaggg 1620  
 acaatgacca cttttgggaa cagcagttgg atggcttaga ttggacagcc caagacatcg 1680  
 tggcgtthtt ggccaagcac ccagaggatg tccagtccag taatggttct gtgtacacct 1740  
 ggagagaagc tttcaacgag actaaccagg caatccggac catatctcgc ttcattggagt 1800  
 gtgtcaacct gaacaagcta gaacctatag caacagaagt ctggctcatc aacaagtcca 1860  
 tggagctgct ggatgagagg aagttctggg ctggtattgt gttcactgga attactccag 1920  
 gcagcattga gctgccccat catgtcaagt acaagatccg aatggacatt gacaatgtgg 1980  
 agaggacaaa taaaatcaag gatgggtact gggacctgg tctctgagct gacctctttg 2040  
 aggacatgog gtaogtctgg gggggcttct cctacttgca ggatgtgggt gagcaggcaa 2100  
 tcatcagggt gctgacgggc accgagaaga aaactgggtg ctatatgcaa cagatgccct 2160  
 atccctgtta cgttgatgac atctttctgc ggggtgatgag cgggtcaatg cccctcttca 2220  
 tgacgctggc ctggatttac tcagtggctg tgatcatcaa gggcatcgtg tatgagaagg 2280

aggcacggct gaaagagacc atgcggatca tgggcctgga caacagcata ctctggttta 2340  
 gctggttcat tagtagcctc attcctcttc ttgtgagcgc tggcctgcta gtggtcatcc 2400  
 tgaagttagg aaacctgctg cctacagtg atcccagcgt ggtgtttgtc ttctgtccg 2460  
 tgtttgctgt ggtgacaatc ctgcagtgtc tctgattag cacactcttc tccagagcca 2520  
 acctggcagc agcctgtggg ggcacatctc acttcacgct gtacctgccc tacgtcctgt 2580  
 gtgtggcatg gcaggactac gtgggcttca cactcaagat cttcgctagc ctgctgtctc 2640  
 ctgtggcttt tgggtttggc tgtgagtact ttgccctttt tgaggagcag ggcattggag 2700  
 tgcagtggga caacctgttt gagagtcctg tggaggaaga tggcttcaat ctcaccactt 2760  
 cgatctccat gatgctgttt gacaccttcc tctatggggg gatgacotgg tacattgagg 2820  
 ctgtctttcc aggccagtac ggaattccca ggccctggta ttttcttgc accaagtcc 2880  
 actggtttgg cgaggaaagt gatgagaaga gccaccctgg ttccaaccag aagagaatgt 2940  
 cagaaatctg catggaggag gaaccaccc acttgaagct gggcgtgtcc attcagaacc 3000  
 tggtaaaagt ctaccgagat gggatgaagg tggctgtcga tggcctggca ctgaattttt 3060  
 atgagggcca gatcacctcc ttctggggc acaatggagc ggggaagacg accaccatgt 3120  
 caatcctgac cgggttgctc cccccgacct cgggcaccgc ctacatctg ggaaaagaca 3180  
 ttcgctctga gatgagcacc atccggcaga acctgggggt ctgtccccag cataactgctc 3240  
 tgtttgacat gctgactgtc gaagaacaca tctggttcta tgccogcttg aaagggctct 3300  
 ctgagaagca cgtgaaggcg gagatggagc agatggccct ggatgttggg ttgccatcaa 3360  
 gcaagctgaa aagcaaaaaca agccagctgt cagggtggaat gcagagaaag ctatctgtgg 3420  
 ccttggcctt tgtcggggga tctaagggtg tcattctgga tgaaccaca gctggtgtgg 3480  
 acccttactc ccgcagggga atatgggagc tgctgctgaa ataccgaca ggccgcacca 3540  
 ttattctctc tacacaccac atggatgaag cggacgtcct gggggacagg attgccatca 3600  
 tctccatgg gaagctgtgc tgtgtgggct cctccctgtt totgaagaac cagctgggaa 3660  
 caggctacta cctgacctg gtcaagaaag atgtggaatc ctccctcagt tctgcagaa 3720  
 acagtagtag cactgtgtca tacctgaaaa aggaggacag tgtttctcag agcagttctg 3780  
 atgctggcct gggcagcgac catgagagtg acacgctgac catcgatgtc tctgctatct 3840  
 ccaacctcat caggaagcat gtgtctgaag cccggctggg ggaagacata gggcatgagc 3900  
 tgacctatgt gctgccatat gaagctgcta aggagggagc ctttgtggaa ctctttcatg 3960  
 agattgatga ccgctctca gacctgggca tttctagtta tggcatotca gagacgacct 4020  
 tggagaanaa attcctcaag gtggccgaag agagtggggg ggatgctgag acctcagatg 4080  
 gtaccttgc agcaagacga aacaggcggg ccttcgggga caagcagagc tgtcttcgcc 4140  
 cgttcoactga agatgatgct gctgatcaa atgattctga catagacca gaatccagag 4200

agacagactt gctcagtggg atggatggca aagggtoccta ccaggtgaaa ggctggaaac 4260  
 ttacacagca acagtttgtg gcccttttgt ggaagagact gctaattgcc agacggagtc 4320  
 ggaaaggatt ttttgetcag attgtcttgc cagctgtgtt tgtctgcatt gcccttgtgt 4380  
 tcagcctgat cgtgccaccc tttggcaagt accccagcct ggaacttcag ccctggatgt 4440  
 acaacgaaca gtacacattt gtcagcaatg atgctcctga ggacacggga accctggaac 4500  
 tcttaaacgc cctcaccaaa gaccctggct tcgggaccog ctgtatggaa ggaaacccaa 4560  
 tcccagacac gccctgccag gcaggggagg aagagtggac cactgccccca gttccccaga 4620  
 ccatcatgga cctcttccag aatgggaact ggacaatgca gaacccttca cctgcatgcc 4680  
 agtgtagcag cgacaaaatc aagaagatgc tgctgtgtg tccccaggg gcaggggggc 4740  
 tgcctcctcc acaaagaaaa caaacactg cagatatcct tcaggacctg acaggaagaa 4800  
 acatttcgga ttatctgggtg aagacgtatg tgcagatcat agccaaaagc ttaaagaaca 4860  
 agatctgggt gaatgagttt aggtatggcg gcttttccct ggggtgtcagt aatactcaag 4920  
 cacttctctc gagtcaagaa gttaatgatg ccatcaaaca aatgaagaaa cacctaaagc 4980  
 tggccaagga cagttctgca gatcgatttc tcaacagctt gggaagattt atgacaggac 5040  
 tggacaccag aaataatgtc aaggtgtggt tcaataacaa gggctggcat gcaatcagct 5100  
 ctttctgaa tgtcatcaac aatgccattc tccgggcaa cctgcaaaag ggagagaacc 5160  
 ctagccatta tgggaattact gctttcaatc atcccctgaa totcaccaag cagcagctct 5220  
 cagaggtggc totgatgacc acatcagtgg atgtccttgt gtccatctgt gtcattctttg 5280  
 caatgtcctt cgtcccagcc agctttgtcg tattctctgat ccaggagcgg gtcagcaaag 5340  
 caaacacct gcagttcatc agtggagtga agcctgtcat ctactggctc tctaattttg 5400  
 tctgggatat gtgcaattac gttgtccctg ccacactggc cattatcatc ttcattctgct 5460  
 tccagcagaa gtccatgtg tctccacca atctgcctgt gctagccctt ctacttttgc 5520  
 tgtatgggtg gtcaatcaca cctctcatgt acccagcctc ctttgtgttc aagatccccca 5580  
 gcacagccta tgtggtgctc accagcgtga acctcttcat tggcattaat ggcagcgtgg 5640  
 ccacctttgt gctggagctg ttcaccgaca ataagctgaa taatatcaat gatatcctga 5700  
 agtccgtggt cttgatcttc ccacatttt gctgggacg agggctcatc gacatgggtga 5760  
 aaaaccaggc aatggctgat gccctggaaa ggtttgggga gaatcgctt gtgtcaccat 5820  
 tatcttggga cttggtggga cgaaacctct tcgccatggc cgtggaaggg gtggtgttct 5880  
 tctcattac tgttctgac cagtacagat tcttcatcag gccagacct gtaaagcaa 5940  
 agctatctcc tctgaatgat gaagatgaag atgtgaggcg ggaaagacag agaattcttg 6000  
 atggtggagg ccagaatgac atcttagaaa tcaaggagtt gacgaagata tatagaagga 6060  
 agcggagcc tgetgttgac aggatttgcg tggcattcc tctggtgag tgcttgggc 6120

tcoctgggagt taatggggct ggaaaatcat caactttcaa gatgttaaca ggagatacca 6180  
ctgttaccag aggagatgct ttccttaaca aaaatagtat cttatcaaac atccatgaag 6240  
tacatcagaa catgggctac tgccotoagt ttgatgcat cacagagctg ttgactggga 6300  
gagaacacgt ggagttcttt gcccttttga gaggagtccc agagaaagaa gttggcaagg 6360  
ttggtgagtg ggcgattcgg aaactgggcc togtgaagta tggagaaaaa tatgctggta 6420  
actatagtgg aggcaacaaa cgcaagctct ctacagccat ggctttgatc ggcgggcctc 6480  
ctgtggtggt tctggatgaa cccaccacag gcatggatcc caaagcccgg cggttcttgt 6540  
ggaattgtgc cctaagtgtt gtcaaggagg ggagatcagt agtgcttaca tctcatagta 6600  
tggaagaatg tgaagctctt tgcactagga tggcaatcat ggtcaatgga aggttcaggt 6660  
gccttggcag tgtccagcat ctaaaaaata ggtttggaga tggttataca atagttgtac 6720  
gaatagcagg gtccaacccg gacctgaagc ctgtccagga tttctttgga cttgcatttc 6780  
ctggaagtgt tctaaaagag aaacaccgga acatgctaca ataccagctt ccatcttcat 6840  
tatcttctct gcccaggata ttcagcatcc tctcccagag caaaaagcga ctccacatag 6900  
aagactactc tgtttctcag acaacacttg accaagtatt tgtgaacttt gccaaggacc 6960  
aaagtgatga tgaccactta aaagacctct cattacacaa aaaccagaca gtagtggacg 7020  
ttgcagttct cacatctttt ctacaggatg agaaagtgaa agaaagctat gtatgaagaa 7080  
tcoctgttcat acgggggtggc tgaagtaaa gaggaactag actttccttt gcacatgtg 7140  
aagtgttgtg gagaaaagag ccagaagttg atgtgggaag aagtaaactg gatactgtac 7200  
tgatactatt caatgcaatg caattcaatg caatgaaaac aaaattccat tacaggggca 7260  
gtgcctttgt agcctatgtc ttgtatggct ctcaagtgaa agacttgaat ttagtTTTTT 7320  
acctataoct atgtgaaact ctattatgga acccaatgga catatggggt tgaactcaca 7380  
cttttttttt ttttttgttc ctgtgtatc tcatgggggt tgcaacaata attcatcaag 7440  
taatcatggc cagcgattat tgatcaaaat caaaaggtaa tgcacatcct cattcactaa 7500  
gccatgcat gcccaggaga ctggtttccc ggtgacacat ccattgctgg caatgagtgt 7560  
gccagagtta ttagtgcaa gtttttcaga aagtttgaag caccatgggtg tgcacatgctc 7620  
acttttgtga aagctgctct gctcagagtc tatcaacatt gaatatcagt tgacagaatg 7680  
gtgcatgog tggetaacat cctgctttga ttccctctga taagctgttc tgggtggcagt 7740  
aacatgcaac aaaaatgtgg gtgtctctag gcacgggaaa cttggttcca ttgttatatt 7800  
gtcctatgct togagccatg ggtctacagg gtcacctta tgagactctt aaatatactt 7860  
agatcctggg aagaggcaaa gaatcaacag ccaaactgct ggggctgcaa gctgctgaag 7920  
ccagggcatg ggattaaaga gattgtgctg tcaaacctag ggaagcctgt gccatttgt 7980  
cctgactgtc tgctaacatg gtacactgca totcaagatg tttatctgac acaagtgtat 8040

tatttctggc tttttgaatt aatctagaaa atgaaaagat ggagttgtat tttgacaaaa 8100  
atgtttgtac tttttaatgt tatttggaaat ttttaagttot atcagtgact tctgaatcct 8160  
tagaatggcc tctttgtaga accctgtggt atagaggagt atggccactg cccactatt 8220  
tttattttct tatgtaagtt tgcataatcag tcatgactag tgcctagaaa gcaatgtgat 8280  
ggtcaggatc tcatgacatt atatttgagt ttctttcaga tcatttagga tactcttaat 8340  
ctcacttcat caatcaaata ttttttgagt gtatgctgta gctgaaagag tatgtacgta 8400  
cgtataagac tagagagata ttaagtctca gtacacttcc tgtgccatgt tattcagctc 8460  
actggtttac aaatataggt tgtcttgtgg ttgtaggagc cactgtaac aatattgggc 8520  
agcctttttt tttttttttt aattgcaaca atgcaaaagc caagaaagta taagggtcac 8580  
aagtttaaac aatgaattct tcaacagggg aaacagctag cttgaaaact tgctgaaaaa 8640  
cacaacttgt gtttatggca tttagtacct tcaataaatt ggctttgcag atattggata 8700  
ccccattaaa tctgacagtc tcaaattttt catctcttca atcaactagtc aagaaaaata 8760  
taaaaacaac aaatacttcc atatggagca tttttcagag ttttctaacc cagtcttatt 8820  
tttctagtca gtaaacattt gtaaaaatac tgtttacta atacttactg ttaactgtct 8880  
tgagagaaaa gaaaaatatg agagaactat tgtttgggga agttcaagtg atctttcaat 8940  
atcattacta acttcttoca ctttttccaa aatttgaata ttaacgctaa aggtgtaaga 9000  
cttcagattt caaattaatc tttctatatt ttttaaattt acagaatatt atataacca 9060  
ctgctgaaaa agaaaaaat gattgtttta gaagttaaag tcaatattga ttttaaatat 9120  
aagtaatgaa ggcatatttc caataactag tgatattgca tcgttgcaat ttacagtatc 9180  
ttcaaaaata cagaatttat agaataattt ctctcattt aatatttttc aaaatcaaag 9240  
ttatggtttc ctcatthttac taaaatogta ttctaattct tcattatagt aaatctatga 9300  
gcaactcctt acttcggttc ctctgatttc aaggccatat tttaaaaaat caaaaggcac 9360  
tgtgaactat tttgaagaaa acacgacatt ttaatacaga ttgaaaggac ctcttctgaa 9420  
gctagaaca atctatagtt atacatcttc attaactctg tgttaccttt taaaatagta 9480  
atthtttaca ttttctgtg taaaccta atgtggtagaa atthtttacca actctatact 9540  
caatcaagca aaatttctgt atattcctg tggaatgtac ctatgtgagt ttcagaaatt 9600  
ctcaaaaatac gtgttcaaaa atthttgtct ttgcatcttt gggacacctc agaaaactta 9660  
ttaacaactg tgaatatgag aaatacagaa gaaaataata agccctctat acataaatgc 9720  
ccagcacaat tcattgttaa aaaacaacca aacctcacac tactgtattt cattatctgt 9780  
actgaaagca aatgctttgt gactatthaa tgttgcatat cattcattca ctgtatagta 9840  
atcattgact aaagccattt gctgtgtttt cttcttgtgg ntgnatata caggthaaat 9900  
atthttccaaa gagccatgtg tcatgtaata ctgaaccctt tgatattgag acattaattt 9960





Lys Thr Phe Gln Glu Leu Ala Val Phe His Asp Leu Glu Gly Met Trp  
 405 410 415  
 Glu Glu Leu Ser Pro Lys Ile Trp Thr Phe Met Glu Asn Ser Gln Glu  
 420 425 430  
 Met Asp Leu Val Arg Met Leu Leu Asp Ser Arg Asp Asn Asp His Phe  
 435 440 445  
 Trp Glu Gln Gln Leu Asp Gly Leu Asp Trp Thr Ala Gln Asp Ile Val  
 450 455 460  
 Ala Phe Leu Ala Lys His Pro Glu Asp Val Gln Ser Ser Asn Gly Ser  
 465 470 475 480  
 Val Tyr Thr Trp Arg Glu Ala Phe Asn Glu Thr Asn Gln Ala Ile Arg  
 485 490 495  
 Thr Ile Ser Arg Phe Met Glu Cys Val Asn Leu Asn Lys Leu Glu Pro  
 500 505 510  
 Ile Ala Thr Glu Val Trp Leu Ile Asn Lys Ser Met Glu Leu Leu Asp  
 515 520 525  
 Glu Arg Lys Phe Trp Ala Gly Ile Val Phe Thr Gly Ile Thr Pro Gly  
 530 535 540  
 Ser Ile Glu Leu Pro His His Val Lys Tyr Lys Ile Arg Met Asp Ile  
 545 550 555 560  
 Asp Asn Val Glu Arg Thr Asn Lys Ile Lys Asp Gly Tyr Trp Asp Pro  
 565 570 575  
 Gly Pro Arg Ala Asp Pro Phe Glu Asp Met Arg Tyr Val Trp Gly Gly  
 580 585 590  
 Phe Ala Tyr Leu Gln Asp Val Val Glu Gln Ala Ile Ile Arg Val Leu  
 595 600 605  
 Thr Gly Thr Glu Lys Lys Thr Gly Val Tyr Met Gln Gln Met Pro Tyr  
 610 615 620  
 Pro Cys Tyr Val Asp Asp Ile Phe Leu Arg Val Met Ser Arg Ser Met  
 625 630 635 640  
 Pro Leu Phe Met Thr Leu Ala Trp Ile Tyr Ser Val Ala Val Ile Ile  
 645 650 655

Lys Gly Ile Val Tyr Glu Lys Glu Ala Arg Leu Lys Glu Thr Met Arg  
660 665 670  
Ile Met Gly Leu Asp Asn Ser Ile Leu Trp Phe Ser Trp Phe Ile Ser  
675 680 685  
Ser Leu Ile Pro Leu Leu Val Ser Ala Gly Leu Leu Val Val Ile Leu  
690 695 700  
Lys Leu Gly Asn Leu Leu Pro Tyr Ser Asp Pro Ser Val Val Phe Val  
705 710 715 720  
Phe Leu Ser Val Phe Ala Val Val Thr Ile Leu Gln Cys Phe Leu Ile  
725 730 735  
Ser Thr Leu Phe Ser Arg Ala Asn Leu Ala Ala Ala Cys Gly Gly Ile  
740 745 750  
Ile Tyr Phe Thr Leu Tyr Leu Pro Tyr Val Leu Cys Val Ala Trp Gln  
755 760 765  
Asp Tyr Val Gly Phe Thr Leu Lys Ile Phe Ala Ser Leu Leu Ser Pro  
770 775 780  
Val Ala Phe Gly Phe Gly Cys Glu Tyr Phe Ala Leu Phe Glu Glu Gln  
785 790 795 800  
Gly Ile Gly Val Gln Trp Asp Asn Leu Phe Glu Ser Pro Val Glu Glu  
805 810 815  
Asp Gly Phe Asn Leu Thr Thr Ser Ile Ser Met Met Leu Phe Asp Thr  
820 825 830  
Phe Leu Tyr Gly Val Met Thr Trp Tyr Ile Glu Ala Val Phe Pro Gly  
835 840 845  
Gln Tyr Gly Ile Pro Arg Pro Trp Tyr Phe Pro Cys Thr Lys Ser Tyr  
850 855 860  
Trp Phe Gly Glu Glu Ser Asp Glu Lys Ser His Pro Gly Ser Asn Gln  
865 870 875 880  
Lys Arg Met Ser Glu Ile Cys Met Glu Glu Glu Pro Thr His Leu Lys  
885 890 895  
Leu Gly Val Ser Ile Gln Asn Leu Val Lys Val Tyr Arg Asp Gly Met  
900 905 910

Lys Val Ala Val Asp Gly Leu Ala Leu Asn Phe Tyr Glu Gly Gln Ile  
                   915                                  920                                  925  
 Thr Ser Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Met Ser  
                   930                                  935                                  940  
 Ile Leu Thr Gly Leu Phe Pro Pro Thr Ser Gly Thr Ala Tyr Ile Leu  
 945                                  950                                  955                                  960  
 Gly Lys Asp Ile Arg Ser Glu Met Ser Thr Ile Arg Gln Asn Leu Gly  
                                   965                                  970                                  975  
 Val Cys Pro Gln His Asn Val Leu Phe Asp Met Leu Thr Val Glu Glu  
                   980                                  985                                  990  
 His Ile Trp Phe Tyr Ala Arg Leu Lys Gly Leu Ser Glu Lys His Val  
                   995                                  1000                                  1005  
 Lys Ala Glu Met Glu Gln Met Ala Leu Asp Val Gly Leu Pro Ser Ser  
                   1010                                  1015                                  1020  
 Lys Leu Lys Ser Lys Thr Ser Gln Leu Ser Gly Gly Met Gln Arg Lys  
 1025                                  1030                                  1035                                  1040  
 Leu Ser Val Ala Leu Ala Phe Val Gly Gly Ser Lys Val Val Ile Leu  
                   1045                                  1050                                  1055  
 Asp Glu Pro Thr Ala Gly Val Asp Pro Tyr Ser Arg Arg Gly Ile Trp  
                   1060                                  1065                                  1070  
 Glu Leu Leu Leu Lys Tyr Arg Gln Gly Arg Thr Ile Ile Leu Ser Thr  
                   1075                                  1080                                  1085  
 His His Met Asp Glu Ala Asp Val Leu Gly Asp Arg Ile Ala Ile Ile  
                   1090                                  1095                                  1100  
 Ser His Gly Lys Leu Cys Cys Val Gly Ser Ser Leu Phe Leu Lys Asn  
 1105                                  1110                                  1115                                  1120  
 Gln Leu Gly Thr Gly Tyr Tyr Leu Thr Leu Val Lys Lys Asp Val Glu  
                   1125                                  1130                                  1135  
 Ser Ser Leu Ser Ser Cys Arg Asn Ser Ser Ser Thr Val Ser Tyr Leu  
                   1140                                  1145                                  1150  
 Lys Lys Glu Asp Ser Val Ser Gln Ser Ser Ser Asp Ala Gly Leu Gly  
                   1155                                  1160                                  1165

Ser Asp His Glu Ser Asp Thr Leu Thr Ile Asp Val Ser Ala Ile Ser  
 1170 1175 1180  
 Asn Leu Ile Arg Lys His Val Ser Glu Ala Arg Leu Val Glu Asp Ile  
 1185 1190 1195 1200  
 Gly His Glu Leu Thr Tyr Val Leu Pro Tyr Glu Ala Ala Lys Glu Gly  
 1205 1210 1215  
 Ala Phe Val Glu Leu Phe His Glu Ile Asp Asp Arg Leu Ser Asp Leu  
 1220 1225 1230  
 Gly Ile Ser Ser Tyr Gly Ile Ser Glu Thr Thr Leu Glu Glu Ile Phe  
 1235 1240 1245  
 Leu Lys Val Ala Glu Glu Ser Gly Val Asp Ala Glu Thr Ser Asp Gly  
 1250 1255 1260  
 Thr Leu Pro Ala Arg Arg Asn Arg Arg Ala Phe Gly Asp Lys Gln Ser  
 1265 1270 1275 1280  
 Cys Leu Arg Pro Phe Thr Glu Asp Asp Ala Ala Asp Pro Asn Asp Ser  
 1285 1290 1295  
 Asp Ile Asp Pro Glu Ser Arg Glu Thr Asp Leu Leu Ser Gly Met Asp  
 1300 1305 1310  
 Gly Lys Gly Ser Tyr Gln Val Lys Gly Trp Lys Leu Thr Gln Gln Gln  
 1315 1320 1325  
 Phe Val Ala Leu Leu Trp Lys Arg Leu Leu Ile Ala Arg Arg Ser Arg  
 1330 1335 1340  
 Lys Gly Phe Phe Ala Gln Ile Val Leu Pro Ala Val Phe Val Cys Ile  
 1345 1350 1355 1360  
 Ala Leu Val Phe Ser Leu Ile Val Pro Pro Phe Gly Lys Tyr Pro Ser  
 1365 1370 1375  
 Leu Glu Leu Gln Pro Trp Met Tyr Asn Glu Gln Tyr Thr Phe Val Ser  
 1380 1385 1390  
 Asn Asp Ala Pro Glu Asp Thr Gly Thr Leu Glu Leu Leu Asn Ala Leu  
 1395 1400 1405  
 Thr Lys Asp Pro Gly Phe Gly Thr Arg Cys Met Glu Gly Asn Pro Ile  
 1410 1415 1420

Pro Asp Thr Pro Cys Gln Ala Gly Glu Glu Glu Trp Thr Thr Ala Pro  
 1425                      1430                      1435                      1440  
 Val Pro Gln Thr Ile Met Asp Leu Phe Gln Asn Gly Asn Trp Thr Met  
                                  1445                      1450                      1455  
 Gln Asn Pro Ser Pro Ala Cys Gln Cys Ser Ser Asp Lys Ile Lys Lys  
                                  1460                      1465                      1470  
 Met Leu Pro Val Cys Pro Pro Gly Ala Gly Gly Leu Pro Pro Pro Gln  
                                  1475                      1480                      1485  
 Arg Lys Gln Asn Thr Ala Asp Ile Leu Gln Asp Leu Thr Gly Arg Asn  
                                  1490                      1495                      1500  
 Ile Ser Asp Tyr Leu Val Lys Thr Tyr Val Gln Ile Ile Ala Lys Ser  
 1505                      1510                      1515                      1520  
 Leu Lys Asn Lys Ile Trp Val Asn Glu Phe Arg Tyr Gly Gly Phe Ser  
                                  1525                      1530                      1535  
 Leu Gly Val Ser Asn Thr Gln Ala Leu Pro Pro Ser Gln Glu Val Asn  
                                  1540                      1545                      1550  
 Asp Ala Ile Lys Gln Met Lys Lys His Leu Lys Leu Ala Lys Asp Ser  
                                  1555                      1560                      1565  
 Ser Ala Asp Arg Phe Leu Asn Ser Leu Gly Arg Phe Met Thr Gly Leu  
                                  1570                      1575                      1580  
 Asp Thr Arg Asn Asn Val Lys Val Trp Phe Asn Asn Lys Gly Trp His  
 1585                      1590                      1595                      1600  
 Ala Ile Ser Ser Phe Leu Asn Val Ile Asn Asn Ala Ile Leu Arg Ala  
                                  1605                      1610                      1615  
 Asn Leu Gln Lys Gly Glu Asn Pro Ser His Tyr Gly Ile Thr Ala Phe  
                                  1620                      1625                      1630  
 Asn His Pro Leu Asn Leu Thr Lys Gln Gln Leu Ser Glu Val Ala Leu  
                                  1635                      1640                      1645  
 Met Thr Thr Ser Val Asp Val Leu Val Ser Ile Cys Val Ile Phe Ala  
                                  1650                      1655                      1660  
 Met Ser Phe Val Pro Ala Ser Phe Val Val Phe Leu Ile Gln Glu Arg  
 1665                      1670                      1675                      1680

Val Ser Lys Ala Lys His Leu Gln Phe Ile Ser Gly Val Lys Pro Val  
 1685 1690 1695  
 Ile Tyr Trp Leu Ser Asn Phe Val Trp Asp Met Cys Asn Tyr Val Val  
 1700 1705 1710  
 Pro Ala Thr Leu Val Ile Ile Ile Phe Ile Cys Phe Gln Gln Lys Ser  
 1715 1720 1725  
 Tyr Val Ser Ser Thr Asn Leu Pro Val Leu Ala Leu Leu Leu Leu Leu  
 1730 1735 1740  
 Tyr Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala Ser Phe Val Phe  
 1745 1750 1755 1760  
 Lys Ile Pro Ser Thr Ala Tyr Val Val Leu Thr Ser Val Asn Leu Phe  
 1765 1770 1775  
 Ile Gly Ile Asn Gly Ser Val Ala Thr Phe Val Leu Glu Leu Phe Thr  
 1780 1785 1790  
 Asp Asn Lys Leu Asn Asn Ile Asn Asp Ile Leu Lys Ser Val Phe Leu  
 1795 1800 1805  
 Ile Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Met Val Lys  
 1810 1815 1820  
 Asn Gln Ala Met Ala Asp Ala Leu Glu Arg Phe Gly Glu Asn Arg Phe  
 1825 1830 1835 1840  
 Val Ser Pro Leu Ser Trp Asp Leu Val Gly Arg Asn Leu Phe Ala Met  
 1845 1850 1855  
 Ala Val Glu Gly Val Val Phe Phe Leu Ile Thr Val Leu Ile Gln Tyr  
 1860 1865 1870  
 Arg Phe Phe Ile Arg Pro Arg Pro Val Asn Ala Lys Leu Ser Pro Leu  
 1875 1880 1885  
 Asn Asp Glu Asp Glu Asp Val Arg Arg Glu Arg Gln Arg Ile Leu Asp  
 1890 1895 1900  
 Gly Gly Gly Gln Asn Asp Ile Leu Glu Ile Lys Glu Leu Thr Lys Ile  
 1905 1910 1915 1920  
 Tyr Arg Arg Lys Arg Lys Pro Ala Val Asp Arg Ile Cys Val Gly Ile  
 1925 1930 1935

Pro Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys  
 1940 1945 1950  
 Ser Ser Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Arg Gly  
 1955 1960 1965  
 Asp Ala Phe Leu Asn Lys Asn Ser Ile Leu Ser Asn Ile His Glu Val  
 1970 1975 1980  
 His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Thr Glu Leu  
 1985 1990 1995 2000  
 Leu Thr Gly Arg Glu His Val Glu Phe Phe Ala Leu Leu Arg Gly Val  
 2005 2010 2015  
 Pro Glu Lys Glu Val Gly Lys Val Gly Glu Trp Ala Ile Arg Lys Leu  
 2020 2025 2030  
 Gly Leu Val Lys Tyr Gly Glu Lys Tyr Ala Gly Asn Tyr Ser Gly Gly  
 2035 2040 2045  
 Asn Lys Arg Lys Leu Ser Thr Ala Met Ala Leu Ile Gly Gly Pro Pro  
 2050 2055 2060  
 Val Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Lys Ala Arg  
 2065 2070 2075 2080  
 Arg Phe Leu Trp Asn Cys Ala Leu Ser Val Val Lys Glu Gly Arg Ser  
 2085 2090 2095  
 Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr  
 2100 2105 2110  
 Arg Met Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu Gly Ser Val  
 2115 2120 2125  
 Gln His Leu Lys Asn Arg Phe Gly Asp Gly Tyr Thr Ile Val Val Arg  
 2130 2135 2140  
 Ile Ala Gly Ser Asn Pro Asp Leu Lys Pro Val Gln Asp Phe Phe Gly  
 2145 2150 2155 2160  
 Leu Ala Phe Pro Gly Ser Val Leu Lys Glu Lys His Arg Asn Met Leu  
 2165 2170 2175  
 Gln Tyr Gln Leu Pro Ser Ser Leu Ser Ser Leu Ala Arg Ile Phe Ser  
 2180 2185 2190

Ile Leu Ser Gln Ser Lys Lys Arg Leu His Ile Glu Asp Tyr Ser Val

2195

2200

2205

Ser Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Asp Gln

2210

2215

2220

Ser Asp Asp Asp His Leu Lys Asp Leu Ser Leu His Lys Asn Gln Thr

2225

2230

2235

2240

Val Val Asp Val Ala Val Leu Thr Ser Phe Leu Gln Asp Glu Lys Val

2245

2250

2255

Lys Glu Ser Tyr Val

2260

<210> 3

<211> 1643

<212> DNA

<213> Homo sapiens

<400> 3

gaattccttg ctggtggctc cacatgcact tccagggcct gcttggctct tctatgggctc 60  
tgtcctgagt gttgatagaa cactgatgt gagtacctgg gcttgagcgt ggccctggaga 120  
tcctgttgac tgtagcatgg agggggcttg tcagctgaat gtctgtatgc aggtgggtggg 180  
agttctggaa tatgatggag ctggaggtgg gaagagaagt aggcttgggg cagctctctc 240  
atgccacctc attotggcca aaactcaggt caaactgtga agagtctaaa tgtgaatctg 300  
cccttcaagg tggctacaaa ggtatctttg tcaaggtagg agacctgtg gcctccacgt 360  
gcacttcag ggccctgcttg gcctcttcta cgggtctgtc ctgagtcttc tatgaatctc 420  
ccttcagggc agattcatat ttagactctt cacagtttga cctgagtttt ggccagaata 480  
aggtgacatt tagtttggtg gcttgatgaa tgacttaaat atttagacat atggtgtgta 540  
ggcctgcatt cctactcttg cctttttttt tgcccctcca gtgttttggg tagttttgct 600  
cccctacag ccaaaggcaa acagataagt tggaggtctg gaggggctac ataattttac 660  
acgactgcaa ttctctggct gcacttcaca aatgtataca aactaaatac aagtcctgtg 720  
ttttatcac agggaggctg atcaatataa tgaaattaa agggggctgg tcccatattg 780  
ttctgtgttt ttgtttgttt gtttcttttt ttgtttttgt ggccctcttc ctctcaattt 840  
atgaagagaa gcagtaagat gttcctctcg ggtcctctga gggacctggg gagctcaggc 900  
tggaatctc caaggcagta ggtcgcctat caaaaatcaa agtccagggtt tgtgggggga 960  
aaacaaaagc agcccattac ccagaggact gtccgccttc ccctcaccctc agcctaggcc 1020

ttgaaagga aacaaaagac aagacaaaat gattggcgtc ctgagggaga ttcagcctag 1080  
 agctctctct cccccaatcc ctccctccgg ctgaggaaac taacaaagga aaaaaaatt 1140  
 gcgaaagca ggatttagag gaagcaaatt cactgggtgc ccttggctgc cgggaacgtg 1200  
 gactagagag tctgcgggcg agccccgagc ccagcgcttc ccgcgcgctct taggccggcg 1260  
 ggccccggcg ggggaagggg acgcagaccg cggaccctaa gacacctgct gtaccctcca 1320  
 cccccacccc acccacctcc ccccaactcc ctagatgtgt cgtggggcggc tgaacgtcgc 1380  
 ccgtttaagg ggcgggcccc ggctccacgt gctttctgct gagtgactga actacataaa 1440  
 cagaggccgg gaacggggcg gggaggaggg agagcacagg ctttgaccga tagtaacctc 1500  
 tgcgctcggg gcagccgaat ctataaaag aactagtccc ggcaaaaacc ccgtaattgc 1560  
 gagcgagagt gagtggggcc gggaccgcga gagccgagcc gaccttctc tcccgggctg 1620  
 cggcagggca gggcggggag etc 1643

<210> 4

<211> 748

<212> DNA

<213> Homo sapiens

<220>

<221> 未确定

<222> (1) .. (748)

<223> 所有的n都是未知的.

<400> 4

attccaanaa cattttcng catctgtggt tgccaactca caatgtcttt cattttctga 60  
 acttaccnc caaatgaagn tcggcagca attatgtagt agcnactctt aggcccccg 120  
 cttacactta tgcttcggc tngttgtgtg ggaattggac ggataccatt tcacncagga 180  
 aacagatatg nccatgatta cgccaagtta ttaggtgcn cgatagaata ctcaagcttg 240  
 gaattcgcgg ccgcagtoga cggacccccg gaaagattc ctctcattac acaaaaacca 300  
 gacagtagtg gacgttgag ttctcacatc ttctctacag gatgagaaag tgaagaaag 360  
 ctatgtatga agaatcctgt tcatacggg tggctgaaag taaagaggaa ctagactttc 420  
 ctttgacca tgtgaagtgt tgtggagaaa agagccagaa gttgatgtgg gaagaagtaa 480  
 actggatact gtactgatac tattcaatgc aatgcaattc aatgcaatga aaacaaaatt 540  
 ccattacagg ggcagtgcct ttgtagccta tgtcttgat ggctctcaag tgaagaactt 600  
 gaatttagtt ttttacctat acctatgtga aactctatta tggaaaccaa tggacatatg 660  
 ggtttgaact cacacttttt ttttttgtt cctgtgtatt ctattgggg ttgcaacaat 720

aattcatcaa gtaaaaaaaaa aaaaaaaaa

748

<210> 5

<211> 2011

<212> DNA

<213> Homo sapiens

<400> 5

agaatcctgt tcatactggg gtggcttgaa agtaaattgga ggaactagac tttcctttgc 60  
accatgtgaa gtgttgtgga gaaaagagcc agaagttgat gtgggaagaa gtaaactgga 120  
tactgtactg atactattca atgcaatgca attcaatgca atgaaaacaa aattccatta 180  
caggggcaag tgcctttgta gcccatgtct tgtatggctc tcaagtgaaa gacttgaatt 240  
tagtttttta cctataccta tgtgaaactc tattatggaa cccaatggac atatggggtt 300  
gaactcacac tttttttttt ttttgttctt gtgtattctc attgggggtg caacaataat 360  
tcatcaagta atcatggcca gcgattattg atcaaaatca aaaggtaatg cacatcctca 420  
ttcactaagc catgccatgc ccaggagact ggtttcccggtg tgacacatcc attgctggca 480  
atgagtgtgc cagagttatt agtgccaagt ttttcagaaa gtttgaagca ccatgggtgtg 540  
tcatgctcac ttttgtgaaa gctgctctgc tcagagtcta tcaacattga atatcagttg 600  
acagaatggt gccatgcgtg gctaacatcc tgctttgatt ccctctgata agctgttctg 660  
gtggcagtaa catgcaacaa aaatgtgggt gtctctaggg acgggaaaact tggttccatt 720  
gttatattgt cctatgcttc gagccatggg tctacaggggt catccttatg agactcttaa 780  
atatacttag atcctggtaa gaggcaaga atcaacagcc aaactgctgg ggctgcaagc 840  
tgctgaagcc agggcatggg attaaagaga ttgtgcgttc aaacctaggg aagcctgtgc 900  
ccatttgtcc tgactgtctg ctaacatggt acactgcac tcaagatggt tatctgacac 960  
aagtgtatta tttctggctt tttgaattaa tctagaaaat gaaaagatgg agttgtattt 1020  
tgacaaaaat gtttgtactt ttaaatgta tttggaattt taagttctat cagtgacttc 1080  
tgaatcotta gaatggcctc tttgtagaac cctgtgggtat agaggagtat ggccactgcc 1140  
ccactatfff tattttctta tgtaagtttg catatcagtc atgactagtg cctagaaagc 1200  
aatgtgatgg tcaggatctc atgacattat atttgagttt ctttcagatc atttaggata 1260  
ctcttaatat cacttcatca atcaaatatt ttttgagtgt atgctgtagc tgaagagta 1320  
tgtaogtacg tataagacta gagagatatt aagtctcagt acacttctctg tgccatgtta 1380  
ttcagctcac tggtttacia atataggttg tcttgtgggt gtaggagccc actgtaacaa 1440  
tattgggcag cctttttttt tttttttttt aattgcaaca atgcaaaagc caagaaagta 1500  
taagggtcac aagtctaac aatgaattct tcaacagggg aaacagctag ctagaaaact 1560

tgctgaaaac acaacttggtg tttatggcat ttagtacctt caaataattg gctttgcaga 1620  
 tattggatac cccattaaat ctgacagtct caaatttttc atctcttcaa tcactagtca 1680  
 agaaaaatat aaaaacaaca aatacttcca tatggagcat ttttcagagt tttctaacc 1740  
 agtcttattt ttctagtcag taaacatttg taaaaatact gtttcactaa tacttactgt 1800  
 taactgtctt gagagaaaag aaaaatatga gagaactatt gtttggggaa gttcaagtga 1860  
 tctttcaata tcattactaa cttcttccac tttttccaaa atttgaatat taacgctaaa 1920  
 ggtgtaagga cttcagattt caaattaatc tttctatatt ttttaaattt acagaatatt 1980  
 atataocca ctgctgaaaa aaaaaaaaaa a 2011

<210> 6

<211> 3366

<212> DNA

<213> Homo sapiens

<220>

<221> 未确定

<222> (1) .. (3366)

<223> 所有的n都是未知的.

<400> 6

agaatcctgt tcatacgggg tggctgaaag taaagaggaa ctagactttc ctttgcacca 60  
 tgtgaagtgt tgtggagaaa agagccagaa gttgatgtgg gaagaagtaa actggatact 120  
 gtactgatac tattcaatgc aatgcaatc aatgcaatga aaacaaaatt ccattacagg 180  
 ggcagtgcct ttgtagccta tgtcttgtat ggctotcaag tgaaagactt gaatttagtt 240  
 ttttaoctat acctatgtga aactctatta tggaaaccaa tggacatatg ggtttgaact 300  
 cacacttttt tttttttttt gttcctgtgt attctcattg gggttgcaac aataattcat 360  
 caagtaatca tggccagcga ttattgatca aaatcaaaaag gtaatgcaca tcctcattca 420  
 ctaagccatg ccatgcccag gagactgggt tcccggtgac acatccattg ctggcaatga 480  
 gtgtgccaga gttattagt ccaagttttt cagaaagttt gaagcaccat ggtgtgtcat 540  
 gctcaacttt gtgaaagctg ctctgctcag agtctatcaa cattgaatat cagttgacag 600  
 aatggtgcca tgcgtggcta acatcctgct ttgattccct ctgataagct gttctgggtg 660  
 cagtaacatg caacaaaaat gtgggtgtct ctaggcacgg gaaacttgg tccattgtta 720  
 tattgtccta tgcttcgagc catgggtcta cagggtcatc cttatgagac tcttaaatat 780  
 acttagatcc tggtaagagg caaagaatca acagccaaac tgctggggct gcaagctgct 840  
 gaagccaggg catgggatta aagagattgt gcgttcaaac ctagggaagc ctgtgcccat 900

ttgtcctgac tgtctgctaa catggtacac tgcatctcaa gatgtttato tgacacaagt 960  
 gtattatttc tggctttttg aattaatcta gaaaatgaaa agatggagtt gtattttgac 1020  
 aaaaatgttt gtacttttta atgttatttg gaattttaag ttctatcagt gacttctgaa 1080  
 tcottagaat ggctctttg tagaacctg tggatagag gagtatggcc actgccccac 1140  
 tttttttatt ttcttatgta agtttgata tcagtcatga ctagtgccta gaaagcaatg 1200  
 tgatggtcag gatctcatga cattatattt gagtttcttt cagatcattt aggatactct 1260  
 taatctcact tcatcaatca aatatttttt gagtgtatgc tgtagctgaa agagtatgta 1320  
 cgtacgtata agactagaga gatattaagt ctacgtacac ttctgtgccc atgttattca 1380  
 gctcactggt ttacaaatat aggttgtctt gtggtttag gagcccactg taacaatatt 1440  
 gggcagcctt tttttttttt ttttaattgc aacaatgcaa aagccaagaa agtataaggg 1500  
 tcacaagttt aaacaatgaa ttcttcaaca gggaaaacag ctagcttgaa aacttgctga 1560  
 aaaacacaac ttgtgtttat ggcatttagt acottcaaat aattggcttt gcagatattg 1620  
 gataccocat taaatctgac agtctcaaat tttcatctc ttcaatcact agtcaagaaa 1680  
 aatataaaaa caacaaatac ttccatattg agcatttttc agagttttct aaccagttct 1740  
 tttttttcta gtcagtaaac atttgtaaaa atactgtttc actaatactt actgttaact 1800  
 gtcttgagag aaaagaaaaa tatgagagaa ctattgtttg gggaagttca agtgatcttt 1860  
 caatatcatt actaacttct tccacttttt ccaaaatttg aatattaacg ctaaaggtgt 1920  
 aagacttcag atttcaaatt aatctttcta ttttttttaa atttacagaa tattatataa 1980  
 cccactgctg aaaaagaaaa aatgattgt tttagaagtt aaagtcaata ttgattttta 2040  
 atataagtaa tgaaggcata tttccaataa ctagtगतat ggcatcgttg caatttacag 2100  
 tatcttcaaa aatacagaat ttatagaata atttctctc atttaattt tttcaaaatc 2160  
 aaagttatgg tttctcatt ttactaaaat cgtattctaa ttcttcatta tagtaaactc 2220  
 atgagcaact ccttacttcg gttcctctga tttcaaggcc atattttaaa aatcaaaaag 2280  
 gcactgtgaa ctattttgaa gaaaacacga cattttaata cagattgaaa ggacctctc 2340  
 tgaagctaga aacaatctat agttatacat cttcattaat actgtgttac cttttaaaat 2400  
 agtaattttt tacattttcc tgtgtaaacc taattgtggg agaaattttt accaactcta 2460  
 tactcaatca agcaaaattt ctgtatattc cctgtggaat gtacctatgt gagtttcaga 2520  
 aattctcaaa atacgtgttc aaaaatttct gcttttgcac ctttgggaca cctcagaaaa 2580  
 cttattaaca actgtgaata tgagaaatac agaagaaaat aataagccct ctatacataa 2640  
 atgcccagca caattcattg ttaaaaaaca accaaacctc acactactgt atttcattat 2700  
 ctgtactgaa agcaaatgct ttgtgactat taaatgttgc acatcattca ttcactgtat 2760  
 agtaatcatt gactaaagcc atttgtgtg ttttctctt gtggntgnat atatcaggta 2820

```

aaatattttc caaagagcca tgtgtcatgt aatactgaac cctttgatat tgagacatta 2880
at ttggaccc ttggtattat ctactagaat aatgtaatac tgnagaaata ttgctctaata 2940
tctttcaaaa tgggtcatcc cccttaaaan gttctatttc cataaggatt tagcttgctt 3000
atcccttctt ataccctaag atgaagctgt ttttgtgctc tttgttcate attggccctc 3060
attccaagca ctttacgctg tctgtaatgg gatctat tttt tgcactggaa tatctgagaa 3120
ttgcaaaaact agacaaaagt ttcacaacag atttctaagt taaatcattt tcattaaaag 3180
gaaaaaagaa aaaaaatttt gtatgtcaat aactttatat gaagtattaa aatgcatatt 3240
tctatgtttg aatataatga gtcacaaaat aaagctgtga cagttctggt aaaaaaaaaa 3300
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3360
aaaaaa 3366

```

<210> 7

<211> 10474

<212> DNA

<213> Homo sapiens

<220>

<221> 未确定

<222> (1) .. (10474)

<223> 所有的n都是未知的.

<400> 7

```

tctagaactg ggtaccagct gctagcaagc ttcggcacga gccgcagagc cgagccgacc 60
cttctctccc gggctgcggc agggcagggc ggggagctcc gcgcaccaac agagccgggt 120
ctcagggcgc tttgctcctt gtttttccc cggttctggt ttctcccctt ctccggaagg 180
cttgtcaagg ggtaggagaa agagacgcaa acacaaaagt ggaaaacagt taatgaccag 240
ccacgggcgt cctgctgtg agctctggcc gctgccttcc agggctcccg agccacacgc 300
tgggcgtgct ggctgagggg acatggcttg ttggcctcag ctgaggttgc tgctgtggaa 360
gaacctcact ttcagaagaa gacaaacatg tcagctgtta ctggaagtgg cctggcctct 420
at ttatcttc ctgatcctga tctctgttcg gctgagctac ccaccctatg aacaacatga 480
atgccatttt ccaaataaag ccatgccctc tgcaggaaca cttocttggg ttcaggggat 540
tatctgtaat gccacaacc cctgtttccg ttaccggact cctggggagg ctcccggagt 600
tgttgaaac ttaacaaat ccattgtggc tcgctgttc tcagatgctc ggaggttct 660
tttatacagc cagaaagaca ccagcatgaa ggacatgcgc aaagttctga gaacattaca 720
gcagatcaag aatccagct caaacttgaa gcttcaagat ttctggtgg acaatgaaac 780

```

cttctctggg ttcoatatc acaacctctc tctcccaaag tctactgtgg acaagatgct 840  
 gagggctgat gtcattotcc acaaggtatt tttgcaaggc taccagttac atttgacaag 900  
 tctgtgcaat ggatcaaaaat cagaagagat gattcaactt ggtgaccaag aagtttctga 960  
 gctttgtggc ctaccaaagg agaaactggc tgcagcagag cgagtacttc gttccaacat 1020  
 ggacatcctg aagccaatcc tgagaacact aaactotaca tctoccttcc cgagcaagga 1080  
 gctggctgaa gccacaaaaa cattgctgca tagtottggg actotggccc aggagctggt 1140  
 cagcatgaga agctggagtg acatgcgaca ggaggtgatg tttctgacca atgtgaacag 1200  
 ctccagctcc tccacccaaa tctaccaggc tgtgtctcgt attgtctgcg ggcatcccga 1260  
 gggagggggg ctgaagatca agtctctcaa ctggtatgag gacaacaact acaaagccct 1320  
 ctttgagggc aatggcactg aggaagatgc tgaaaccttc tatgacaact ctacaactcc 1380  
 ttactgcaat gatttgatga agaatttggg gtotagtctt ctttcccgca ttatctggaa 1440  
 agctctgaag ccgctgctcg ttgggaagat cctgtataca cctgacactc cagccacaag 1500  
 gcaggtcatg gctgaggtga acaagacctt ccaggaactg gctgtgttcc atgatctgga 1560  
 aggcatgtgg gaggaactca gccccaagat ctggaccttc atggagaaca gccaagaaat 1620  
 ggaccttgtc cggatgctgt tggacagcag ggacaatgac cacttttggg aacagcagtt 1680  
 ggatggctta gattggacag cccaagacat cgtggcgttt ttggccaagc acccagagga 1740  
 tgtccagtcc agtaatggtt ctgtgtacac ctggagagaa gctttcaacg agactaacca 1800  
 ggcaatccgg accatatctc gcttcatgga gtgtgtcaac ctgaacaagc tagaacccat 1860  
 agcaacagaa gtctggctca tcaacaagtc catggagctg ctggatgaga ggaagttctg 1920  
 ggctgggtatt gtgttcaactg gaattactcc aggcagcatt gagctgcccc atcatgtcaa 1980  
 gtacaagatc cgaatggaca ttgacaatgt ggagaggaca aataaaatca aggatgggta 2040  
 ctgggacctt ggtcctcgag ctgaccttct tgaggacatg cggtagctct gggggggctt 2100  
 cgcctacttg cgggatgtgg tggagcaggc aatcatcagg gtgctgacgg gcaccgagaa 2160  
 gaaaactggt gtctatatgc aacagatgcc ctatccctgt tacgttgatg acatctttct 2220  
 gcgggtgatg agccggtcaa tgcccctctt catgacgctg gcttgattt actcagtggc 2280  
 tgtgatcatc aagggcatcg tgtatgagaa ggaggcacgg ctgaaagaga ccatgcggat 2340  
 catgggcctg gacaacagca tactctgggt tagctgggtc attagtagcc tcattcctct 2400  
 tcttgtgagc gctggcctgc tagtggatc cctgaagtta ggaaacctgc tgccctacag 2460  
 tgatcccagc gtgggtggtt tcttctctgc cgtggttgct gtggtgacaa tcctgcagtg 2520  
 cttoctgatt agcacactct tctccagagc caacctggca gcagcctgtg ggggcatcat 2580  
 ctacttcacg ctgtacctgc cctacgtcct gtgtgtggca tggcaggact acgtgggctt 2640  
 cacactcaag atcttcgcta gctgctgtgc tctgtggct tttgggtttg gctgtgagta 2700

ctttgccctt tttgaggagc agggcattgg agtgcagtgg gacaacctgt ttgagagtcc 2760  
tgtggaggaa gatggottca atctcaccac ttcgatctcc atgatgctgt ttgacacctt 2820  
cctctatggg gtgatgacct ggtacattga ggctgtcttt ccaggccagt acggaattcc 2880  
caggccctgg tattttcctt gcaccaagtc ctactggttt ggcgaggaaa gtgatgagaa 2940  
gagccaccct ggttccaacc agaagagaaat gtcagaaatc tgcattggagg aggaaccac 3000  
ccacttgaag ctggggogtgt ccattcagaa cctggtaaaa gtctaccgag atgggatgaa 3060  
ggtggctgtc gatggcctgg cactgaattt ttatgagggc cagatcacct ccttcctggg 3120  
ccacaatgga gcggggaaga cgaccacat gtcaatcctg accgggttgt tcccccgac 3180  
ctcgggcacc gcctacatcc tgggaaaaga cattcgctct gagatgagca ccatccggca 3240  
gaacctgggg gtctgtcccc agcataacgt gctgtttgac atgctgactg tcgagaaca 3300  
catctggttc tatgcccgt tgaagggtct ctctgagaag cacgtgaagg cggagatgga 3360  
gcagatggcc ctggatgttg gtttgccatc aagcaagctg aaaagcaaaa caagccagct 3420  
gtcaggtgga atgcagagaa agctatctgt ggccttggcc tttgtcgggg gatctaaggt 3480  
tgtcattctg gatgaacca cagctggtgt ggacccttac tccgcaggg gaatatggga 3540  
gctgctgctg aaataccgac aaggccgcac cattattctc tctacacacc acatggatga 3600  
agcggacgtc ctgggggaca ggattgcat catotccat gggagctgt gctgtgtggg 3660  
ctcctcctg tttctgaaga accagctggg aacaggctac tacctgacct tggtaagaa 3720  
agatgtggaa tctccctca gttcctgcag aaacagtagt agcactgtgt catacctgaa 3780  
aaaggaggac agtgtttctc agagcagttc tgatgctggc ctgggcagcg accatgagag 3840  
tgacacgctg accatcgatg tctctgctat ctccaacctc atcaggaagc atgtgtotga 3900  
agccccgctg gtggaagaca tagggcatga gctgacctat gtgetgcat atgaagctgc 3960  
taaggagga gcccttgtgg aactcttca tgagattgat gaccggctct cagacctggg 4020  
catttctagt tatggcatct cagagacgac cctggaagaa atattcctca aggtggccga 4080  
agagagtggg gtggatgctg agacctcaga tggtaacctg ccagcaagac gaaacaggcg 4140  
ggccttcggg gacaagcaga gctgtcttcg ccggttcaact gaagatgatg ctgctgatcc 4200  
aaatgattct gacatagacc cagaatccag agagacagac ttgctcagtg ggatggatgg 4260  
caaaggtcc taccaggtga aaggctgga acttacacag caacagtttg tggccctttt 4320  
gtggaagaga ctgctaattg ccagacggag toggaaagga tttttgctc agattgtctt 4380  
gccagctgtg tttgtctgca ttgcccttgt gttcagcctg atcgtgccac cctttggcaa 4440  
gtaccccagc ctggaacttc agccctggat gtacaacgaa cagtacacat ttgtcagcaa 4500  
tgatgotcct gaggacacgg gaaccctgga actottaaac gccctacca aagacctgg 4560  
cttcgggacc cgctgatgg aaggaaacc aatcccagac acgccctgcc aggcagggga 4620

ggaagagtgg accactgccc cagtccccca yaccalcaly gacctcttcc agaatgggaa 4680  
ctggacaatg cagaaccctt cacctgcatg ccagtgtagc agcgacaaaa tcaagaagat 4740  
gctgcctgtg tgtccccag gggcaggggg gctgcctcct ccacaaagaa aacaaaacac 4800  
tgcagatata cttcaggacc tgacaggaag aaacatttctg gattatctgg tgaagacgta 4860  
tgtgcagatc atagccaaaa gcttaaagaa caagatctgg gtgaatgagt ttaggtatgg 4920  
cggcttttcc ctgggtgtca gtaatactca agcaacttct ccgagtcaag aagttaatga 4980  
tgccatcaaa caaatgaaga aacacctaaa gctggccaag gacagttctg cagatcgatt 5040  
tctcaacagc ttgggaagat ttatgacagg actggacacc agaaataatg tcaaggtgtg 5100  
gttcaataac aagggtctggc atgcaatcag ctctttcctg aatgtcatca acaatgccat 5160  
tctccgggcc aacctgcaaa agggagagaa ccctagccat tatggaatta ctgctttcaa 5220  
tcatcccctg aatctcacca agcagcagct ctacagagtg gctctgatga ccacatcagt 5280  
ggatgtcctt gtgtccatct gtgtcatctt tgcaatgtcc ttcgtcccag ccagctttgt 5340  
cgtattcctg atccaggagc gggtcagcaa agcaaaaacac ctgcagttca tcagtggagt 5400  
gaagcctgtc atctactggc tctctaattt tgtctgggat atgtgcaatt acgttgtccc 5460  
tgccacactg gtcattatca tcttcatctg cttccagcag aagtcctatg tgtcctccac 5520  
caatctgoc t gtgctagccc ttctactttt gctgtatggg tggtaaatca cacctctcat 5580  
gtaccagacc tcctttgtgt tcaagatccc cagcacagcc tatgtggtgc tcaccagcgt 5640  
gaacctcttc attggcatta atggcagcgt ggccaccttt gtgctggagc tgttcaccga 5700  
caataagctg aataatatca atgatatcct gaagtccgtg ttcttgatct tcccacattt 5760  
ttgcctggga cgagggctca tcgacatggt gaaaaaccag gcaatggctg atgcctgga 5820  
aaggtttggg gagaatcgct ttgtgtcacc attatcttgg gacttgggtg gacgaaacct 5880  
cttcgccatg gccgtggaag ggggtggtgtt cttcctcatt actgttctga tccagtacag 5940  
attcttcatc aggcccagac ctgtaaatgc aaagctatct cctctgaatg atgaagatga 6000  
agatgtgagg cgggaaagac agagaattct tgatggtgga ggccagaatg acatcttaga 6060  
aatcaaggag ttgacgaaga tatatagaag gaagcgggaag cctgctgttg acaggatttg 6120  
cgtgggcatt cctcctggtg agtgctttgg gctcctggga gttaatgggg ctggaaaatc 6180  
atcaactttc aagatgtaa caggagatac cactgttacc agaggagatg ctttcttaa 6240  
caaaaatagt atcttatcaa acatccatga agtacatcag aacatgggct actgcctca 6300  
gtttgatgcc atcacagagc tgttgactgg gagagaacac gtggagtctt ttgcctttt 6360  
gagaggagtc ccagagaaag aagttggcaa ggttgggtgag tgggcgattc ggaaactggg 6420  
cctcgtgaag tatggagaaa aatatgctgg taactatagt ggaggcaaca aacgcaagct 6480  
ctctacagcc atggctttga tcggcgggcc tcctgtggtg tttctggatg aaccaccac 6540

aggcattgat cccaaagccc ggcggttctt gtggaattgt gccctaagtg ttgtcaagga 6600  
 ggggagatca gtagtgctta catctcatag tatggaagaa tgtgaagctc tttgcaactag 6660  
 gatggcaatc atggtcaatg gaaggttcag gtgccttggc agtgtccagc atctaaaaaa 6720  
 taggttttga gatggttata caatagttgt acgaatagca gggccaacc cggacctgaa 6780  
 gcctgtccag gatttcttg gacttgcat tcttgggaagt gttctaaaag agaaacaccg 6840  
 gaacatgcta caataccagc tccatcttc attatctct ctggccagga tattcagcat 6900  
 cctctcccag agcaaaaagc gactccacat agaagactac tctgtttctc agacaacact 6960  
 tgaccaagta tttgtgaact ttgccaagga ccaaagtgat gatgaccact taaaagacct 7020  
 ctattacac aaaaaccaga cagtagtgga cgttgcaagt ctacatctt ttctacagga 7080  
 tgagaaagtg aaagaaagct atgtatgaag aatctgttc atacgggggtg gctgaaagta 7140  
 aagaggaact agactttctt ttgcaccatg tgaagtgttg tggagaaaag agccagaagt 7200  
 tgatgtggga agaagtaaac tggatactgt actgatacta ttcaatgcaa tgcaattcaa 7260  
 tgcaatgaaa acaaaattcc attacagggg cagtgccttt gtagcctatg tcttgtatgg 7320  
 ctctcaagtg aaagacttga attagtttt ttacctatac ctatgtgaaa ctctattatg 7380  
 gaaccaatg gacatatggg tttgaactca caotTTTTTT tttttttgt tctgtgtat 7440  
 totcattggg gttgcaaca taattcatca agtaatcatg gccagcgatt attgatcaaa 7500  
 atcaaaagt aatgcacatc ctattcact aagccatgcc atgccagga gactggtttc 7560  
 ccggtgacac atccattgct ggcaatgagt gtgccagagt tattagtgcc aagtttttca 7620  
 gaaagtttga agcaccatgg tgtgtcatgc tcaotTTTTGT gaaagctgct ctgctcagag 7680  
 totatcaaca ttgaatatca gttgacagaa tggtgccatg cgtggctaac atctgtcttt 7740  
 gattccctct gataagctgt tctggtggca gtaacatgca acaaaaatgt ggggtgtctct 7800  
 aggcacggga aacttggttc cattgttata ttgtcctatg ctctgagcca tgggtctaca 7860  
 gggtcactct tatgagactc ttaaataac ttagatcctg gtaagaggca aagaatcaac 7920  
 agccaaactg ctggggctgc aagctgctga agccagggca tgggattaaa gagattgtgc 7980  
 gttcaaacct agggaagcct gtgccattt gtctgactg tctgctaaca tggtaactg 8040  
 catctcaaga tgtttatctg acacaagtgt attatttctg gctttttgaa ttaatctaga 8100  
 aatgaaaag atggagttgt attttgaca aatgtttgt actttttaat gttatttga 8160  
 attttaagtt ctatcagtga cttctgaatc cttagaatgg cctctttgta gaacctgtg 8220  
 gtatagagga gtatggccac tgccccacta tttttatctt cttatgtaag tttgcatatc 8280  
 agtcatgact agtgectaga aagcaatgtg atggtcagga tctcatgaca ttatatttga 8340  
 gtttcttca gatcatttag gatactotta atctcacttc atcaatcaaa tttttttgga 8400  
 gtgtatgctg tagctgaaag agtatgtacg tacgtataag actagagaga tattaagtct 8460

cagtacactt cctgtgccat gttattcagc tcaactggttt acaaatatag gttgtcttgt 8520  
ggtttagga gccactgta acaatattgg gcagcctttt tttttttttt ttaattgcaa 8580  
caatgcaaaa gccaaagaaag tataagggtc acaagtttaa acaatgaatt cttcaacagg 8640  
gaaaacagct agcttgaaaa cttgctgaaa aacacaactt gtgtttatgg catttagtac 8700  
cttcaataa ttggctttgc agatattgga taccaccatta aatctgacag tctcaaattt 8760  
ttcatctctt caatcactag tcaagaaaaa tataaaaaaca acaataactt ccatatggag 8820  
catttttcag agttttctaa cccagtetta tttttctagt cagtaaacad ttgtaaaaat 8880  
actgtttcac taatacttac tgttaactgt cttgagagaa aagaaaaata tgagagaact 8940  
attgtttggg gaagttcaag tgatctttca atatcattac taacttcttc cactttttcc 9000  
aaaatttgaa tattaacgct aaaggtgtaa gacttcagat ttcaaattaa tctttctata 9060  
ttttttaaat ttacagaata ttatataacc cactgctgaa aaagaaaaaa atgattggtt 9120  
tagaagttaa agtcaatatt gattttaaat ataagtaatg aaggcatatt tccaataact 9180  
agtgatatgg catogttgca atttacagta tcttcaaaaa tacagaattt atagaataat 9240  
ttctcctcat ttaatatttt tcaaaatcaa agttatgggt tctcattttt actaaaatcg 9300  
tattctaatt cttcattata gtaaactctat gagcaactcc ttacttcggt tctctgatt 9360  
tcaaggccat attttaaaaa atcaaaaggc actgtgaact attttgaaga aaacacgaca 9420  
ttttaataca gattgaaagg acctctctg aagctagaaa caatctatag ttatacatct 9480  
tcattaatac tgtgttacct tttaaaatag taatttttta cattttctctg tgtaaaccta 9540  
attgtggtag aaatttttac caactctata ctcaatcaag caaaatttct gtatattccc 9600  
tgtggaatgt acctatgtga gtttcagaaa ttctcaaaat acgtgttcaa aaatttctgc 9660  
ttttgcatct ttgggacacc tcagaaaact tattaacaac tgtgaatatg agaaatacag 9720  
aagaaaataa taagcctctc atacataaat gccagcaca attcattggt aaaaaacaac 9780  
caaacctcac actactgtat ttcattatct gtactgaaag caaatgcttt gtgactatta 9840  
aatgttgcac atcattcatt cactgtatag taatcattga ctaaagccat ttgctgtggt 9900  
ttcttcttgt ggntgnatat atcaggtaa atattttcca aagagccatg tgtcatgtaa 9960  
tactgaacco tttgatattg agacattaat ttggaccctt ggtattatct actagaataa 10020  
tgtaaactg nagaaatatt gctctaattc tttcaaatg gtgcatcccc cttaaaangt 10080  
tctatttcca taaggattta gcttgcttat cccttcttat accctaagat gaagctgttt 10140  
ttgtgctctt tgttcatcat tggcctcat tccaagcact ttacgctgct tgtaatggga 10200  
tctatttttg cactggaata tctgagaatt gcaaaaactag acaaaaagttt cacaacagat 10260  
ttctaagtta aatcattttc attaaaagga aaaaagaaaa aaaattttgt atgtcaataa 10320  
ctttatatga agtattaataa tgcatatttc tatgttgtaa tataatgagt cacaaaataa 10380

agctgtgaca gttctgttaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 10440

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 10474

<210> 8

<211> 2261

<212> PRT

<213> Homo sapiens

<400> 8

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

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





	965		970		975
Val Cys Pro Gln His Asn Val Leu Phe Asp Met Leu Thr Val Glu Glu					
	980		985		990
His Ile Trp Phe Tyr Ala Arg Leu Lys Gly Leu Ser Glu Lys His Val					
	995		1000		1005
Lys Ala Glu Met Glu Gln Met Ala Leu Asp Val Gly Leu Pro Ser Ser					
	1010		1015		1020
Lys Leu Lys Ser Lys Thr Ser Gln Leu Ser Gly Gly Met Gln Arg Lys					
	1025		1030		1035
					1040
Leu Ser Val Ala Leu Ala Phe Val Gly Gly Ser Lys Val Val Ile Leu					
	1045		1050		1055
Asp Glu Pro Thr Ala Gly Val Asp Pro Tyr Ser Arg Arg Gly Ile Trp					
	1060		1065		1070
Glu Leu Leu Leu Lys Tyr Arg Gln Gly Arg Thr Ile Ile Leu Ser Thr					
	1075		1080		1085
His His Met Asp Glu Ala Asp Val Leu Gly Asp Arg Ile Ala Ile Ile					
	1090		1095		1100
Ser His Gly Lys Leu Cys Cys Val Gly Ser Ser Leu Phe Leu Lys Asn					
	1105		1110		1115
					1120
Gln Leu Gly Thr Gly Tyr Tyr Leu Thr Leu Val Lys Lys Asp Val Glu					
	1125		1130		1135
Ser Ser Leu Ser Ser Cys Arg Asn Ser Ser Ser Thr Val Ser Tyr Leu					
	1140		1145		1150
Lys Lys Glu Asp Ser Val Ser Gln Ser Ser Ser Asp Ala Gly Leu Gly					
	1155		1160		1165
Ser Asp His Glu Ser Asp Thr Leu Thr Ile Asp Val Ser Ala Ile Ser					
	1170		1175		1180
Asn Leu Ile Arg Lys His Val Ser Glu Ala Arg Leu Val Glu Asp Ile					
	1185		1190		1195
					1200
Gly His Glu Leu Thr Tyr Val Leu Pro Tyr Glu Ala Ala Lys Glu Gly					
	1205		1210		1215
Ala Phe Val Glu Leu Phe His Glu Ile Asp Asp Arg Leu Ser Asp Leu					

1220	1225	1230	
Gly Ile Ser Ser Tyr Gly Ile Ser Glu Thr Thr Leu Glu Glu Ile Phe			
1235	1240	1245	
Leu Lys Val Ala Glu Glu Ser Gly Val Asp Ala Glu Thr Ser Asp Gly			
1250	1255	1260	
Thr Leu Pro Ala Arg Arg Asn Arg Arg Ala Phe Gly Asp Lys Gln Ser			
1265	1270	1275	1280
Cys Leu Arg Pro Phe Thr Glu Asp Asp Ala Ala Asp Pro Asn Asp Ser			
1285	1290	1295	
Asp Ile Asp Pro Glu Ser Arg Glu Thr Asp Leu Leu Ser Gly Met Asp			
1300	1305	1310	
Gly Lys Gly Ser Tyr Gln Val Lys Gly Trp Lys Leu Thr Gln Gln Gln			
1315	1320	1325	
Phe Val Ala Leu Leu Trp Lys Arg Leu Leu Ile Ala Arg Arg Ser Arg			
1330	1335	1340	
Lys Gly Phe Phe Ala Gln Ile Val Leu Pro Ala Val Phe Val Cys Ile			
1345	1350	1355	1360
Ala Leu Val Phe Ser Leu Ile Val Pro Pro Phe Gly Lys Tyr Pro Ser			
1365	1370	1375	
Leu Glu Leu Gln Pro Trp Met Tyr Asn Glu Gln Tyr Thr Phe Val Ser			
1380	1385	1390	
Asn Asp Ala Pro Glu Asp Thr Gly Thr Leu Glu Leu Leu Asn Ala Leu			
1395	1400	1405	
Thr Lys Asp Pro Gly Phe Gly Thr Arg Cys Met Glu Gly Asn Pro Ile			
1410	1415	1420	
Pro Asp Thr Pro Cys Gln Ala Gly Glu Glu Glu Trp Thr Thr Ala Pro			
1425	1430	1435	1440
Val Pro Gln Thr Ile Met Asp Leu Phe Gln Asn Gly Asn Trp Thr Met			
1445	1450	1455	
Gln Asn Pro Ser Pro Ala Cys Gln Cys Ser Ser Asp Lys Ile Lys Lys			
1460	1465	1470	
Met Leu Pro Val Cys Pro Pro Gly Ala Gly Gly Leu Pro Pro Pro Gln			

1475	1480	1485
Arg Lys Gln Asn Thr Ala Asp Ile Leu Gln Asp Leu Thr Gly Arg Asn		
1490	1495	1500
Ile Ser Asp Tyr Leu Val Lys Thr Tyr Val Gln Ile Ile Ala Lys Ser		
1505	1510	1515
1520		
Leu Lys Asn Lys Ile Trp Val Asn Glu Phe Arg Tyr Gly Gly Phe Ser		
	1525	1530
		1535
Leu Gly Val Ser Asn Thr Gln Ala Leu Pro Pro Ser Gln Glu Val Asn		
	1540	1545
		1550
Asp Ala Ile Lys Gln Met Lys Lys His Leu Lys Leu Ala Lys Asp Ser		
	1555	1560
		1565
Ser Ala Asp Arg Phe Leu Asn Ser Leu Gly Arg Phe Met Thr Gly Leu		
	1570	1575
		1580
Asp Thr Arg Asn Asn Val Lys Val Trp Phe Asn Asn Lys Gly Trp His		
1585	1590	1595
1600		
Ala Ile Ser Ser Phe Leu Asn Val Ile Asn Asn Ala Ile Leu Arg Ala		
	1605	1610
		1615
Asn Leu Gln Lys Gly Glu Asn Pro Ser His Tyr Gly Ile Thr Ala Phe		
	1620	1625
		1630
Asn His Pro Leu Asn Leu Thr Lys Gln Gln Leu Ser Glu Val Ala Leu		
	1635	1640
		1645
Met Thr Thr Ser Val Asp Val Leu Val Ser Ile Cys Val Ile Phe Ala		
	1650	1655
		1660
Met Ser Phe Val Pro Ala Ser Phe Val Val Phe Leu Ile Gln Glu Arg		
1665	1670	1675
1680		
Val Ser Lys Ala Lys His Leu Gln Phe Ile Ser Gly Val Lys Pro Val		
	1685	1690
		1695
Ile Tyr Trp Leu Ser Asn Phe Val Trp Asp Met Cys Asn Tyr Val Val		
	1700	1705
		1710
Pro Ala Thr Leu Val Ile Ile Ile Phe Ile Cys Phe Gln Gln Lys Ser		
	1715	1720
		1725
Tyr Val Ser Ser Thr Asn Leu Pro Val Leu Ala Leu Leu Leu Leu Leu		

1730	1735	1740	
Tyr Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala Ser Phe Val Phe			
1745	1750	1755	1760
Lys Ile Pro Ser Thr Ala Tyr Val Val Leu Thr Ser Val Asn Leu Phe			
	1765	1770	1775
Ile Gly Ile Asn Gly Ser Val Ala Thr Phe Val Leu Glu Leu Phe Thr			
	1780	1785	1790
Asp Asn Lys Leu Asn Asn Ile Asn Asp Ile Leu Lys Ser Val Phe Leu			
	1795	1800	1805
Ile Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Met Val Lys			
	1810	1815	1820
Asn Gln Ala Met Ala Asp Ala Leu Glu Arg Phe Gly Glu Asn Arg Phe			
1825	1830	1835	1840
Val Ser Pro Leu Ser Trp Asp Leu Val Gly Arg Asn Leu Phe Ala Met			
	1845	1850	1855
Ala Val Glu Gly Val Val Phe Phe Leu Ile Thr Val Leu Ile Gln Tyr			
	1860	1865	1870
Arg Phe Phe Ile Arg Pro Arg Pro Val Asn Ala Lys Leu Ser Pro Leu			
	1875	1880	1885
Asn Asp Glu Asp Glu Asp Val Arg Arg Glu Arg Gln Arg Ile Leu Asp			
	1890	1895	1900
Gly Gly Gly Gln Asn Asp Ile Leu Glu Ile Lys Glu Leu Thr Lys Ile			
1905	1910	1915	1920
Tyr Arg Arg Lys Arg Lys Pro Ala Val Asp Arg Ile Cys Val Gly Ile			
	1925	1930	1935
Pro Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys			
	1940	1945	1950
Ser Ser Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Arg Gly			
	1955	1960	1965
Asp Ala Phe Leu Asn Lys Asn Ser Ile Leu Ser Asn Ile His Glu Val			
	1970	1975	1980
His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Thr Glu Leu			

1985                      1990                      1995                      2000  
 Leu Thr Gly Arg Glu His Val Glu Phe Phe Ala Leu Leu Arg Gly Val  
                             2005                      2010                      2015  
 Pro Glu Lys Glu Val Gly Lys Val Gly Glu Trp Ala Ile Arg Lys Leu  
                             2020                      2025                      2030  
 Gly Leu Val Lys Tyr Gly Glu Lys Tyr Ala Gly Asn Tyr Ser Gly Gly  
                             2035                      2040                      2045  
 Asn Lys Arg Lys Leu Ser Thr Ala Met Ala Leu Ile Gly Gly Pro Pro  
                             2050                      2055                      2060  
 Val Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Lys Ala Arg  
 2065                      2070                      2075                      2080  
 Arg Phe Leu Trp Asn Cys Ala Leu Ser Val Val Lys Glu Gly Arg Ser  
                             2085                      2090                      2095  
 Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr  
                             2100                      2105                      2110  
 Arg Met Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu Gly Ser Val  
                             2115                      2120                      2125  
 Gln His Leu Lys Asn Arg Phe Gly Asp Gly Tyr Thr Ile Val Val Arg  
                             2130                      2135                      2140  
 Ile Ala Gly Ser Asn Pro Asp Leu Lys Pro Val Gln Asp Phe Phe Gly  
 2145                      2150                      2155                      2160  
 Leu Ala Phe Pro Gly Ser Val Leu Lys Glu Lys His Arg Asn Met Leu  
                             2165                      2170                      2175  
 Gln Tyr Gln Leu Pro Ser Ser Leu Ser Ser Leu Ala Arg Ile Phe Ser  
                             2180                      2185                      2190  
 Ile Leu Ser Gln Ser Lys Lys Arg Leu His Ile Glu Asp Tyr Ser Val  
                             2195                      2200                      2205  
 Ser Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Asp Gln  
                             2210                      2215                      2220  
 Ser Asp Asp Asp His Leu Lys Asp Leu Ser Leu His Lys Asn Gln Thr  
 2225                      2230                      2235                      2240  
 Val Val Asp Val Ala Val Leu Thr Ser Phe Leu Gln Asp Glu Lys Val

2245

2250

2255

Lys Glu Ser Tyr Val

2260

<210> 9

<211> 10474

<212> DNA

<213> Homo sapiens

<220>

<221> 未确定

<222> (1) .. (10474)

<223> 所有的n都是未知的.

<400> 9

tctagaactg ggtaccagct gctagcaagc ttcggcacga gccgcagagc cgagccgacc 60  
cttctctccc gggctgcggc agggcagggc ggggagctcc ggcacccaac agagccgggt 120  
ctcagggcgc tttgctcctt gtttttccc cggttctggt ttctcccctt ctccggaagg 180  
cttgtcaagg ggtaggagaa agagacgcaa acacaaaagt ggaaaacagt taatgaccag 240  
ccaeggcgt ccctgctgtg agctctggcc gctgccttcc agggctcccg agccacacgc 300  
tgggcgtgct ggctgagggg acatggcttg ttggcctcag ctgagggtgc tgctgtggaa 360  
gaacctcact ttcagaagaa gacaaacatg tcagctgtta ctggaagtgg cctggcctct 420  
atztatcttc ctgatcctga tctctgttcg gctgagctac ccaccctatg aacaacatga 480  
atgccattht ccaaataaag ccatgccctc tgcaggaaca cttccttggg ttcaggggat 540  
tatctgtaat gccaacaacc cctgtttccg ttacccgact cctggggagg ctcccggagt 600  
tgttggaaac tttaaacaaat ccattgtggc tcgcctgttc tcagatgctc ggaggcttct 660  
ttatacagc cagaaagaca ccagcatgaa ggacatgcgc aaagtctga gaacattaca 720  
gcagatcaag aaatccagct caaactgaa gcttcaagat ttcctggtgg acaatgaaac 780  
cttctctggg ttctatatc acaacctctc tctcccaaag tctactgtgg acaagatgct 840  
gagggctgat gtcattctcc acaaggattt ttgcaaggc taccagttac atttgacaag 900  
tctgtgcaat ggatcaaaat cagaagagat gattcaactt ggtgaccaag aagtttctga 960  
gctttgtggc ctaccaaagg agaaactggc tgcagcagag cgagtacttc gttccaacat 1020  
ggacatcctg aagccaatcc tgagaacact aaactctaca tctcccttcc cgagcaagga 1080  
gctggctgaa gccacaaaaa cattgctgca tagtcttggg actctggccc aggagctggt 1140  
cagcatgaga agctggagtg acatgcgaca ggaggtgatg tttctgacca atgtgaacag 1200

ctccagctcc tccacccaaa tctaccagge tgtgtctcgt attgtctgcg ggcattcccga 1260  
gggagggggg ctgaagatca agtctctcaa ctggatagag gacaacaact acaaagccct 1320  
ctttggagge aatggcactg aggaagatgc tgaaaccttc tatgacaact ctacaactcc 1380  
ttactgcaat gatttgatga agaatttggg gtotagtoct ctttcccgca ttatctggaa 1440  
agctctgaag ccgctgctcg ttgggaagat cctgtataca cctgacactc cagccacaag 1500  
gcaggctcatg gctgaggtga acaagacctt ccaggaactg gctgtgttcc atgatctgga 1560  
aggcatgtgg gaggaactca gccccaagat ctggaccttc atggagaaca gccaaagaaat 1620  
ggaccttgtc cggatgctgt tggacagcag ggacaatgac cacttttggg aacagcagtt 1680  
ggatggctta gattggacag cccaagacat cgtggcgttt ttggccaagc acccagagga 1740  
tgtccagtcc agtaatggtt ctgtgtacac ctggagagaa gctttcaacg agactaacca 1800  
ggcaatccgg accatatctc gcttcatgga gtgtgtcaac ctgaacaagc tagaacccat 1860  
agcaacagaa gtctggotca tcaacaagtc catggagctg ctggatgaga ggaagttctg 1920  
ggctggtatt gtgttcaactg gaattactcc aggcagcatt gagctgcccc atcatgtcaa 1980  
gtacaagatc cgaatggaca ttgacaatgt ggagaggaca aataaaatca aggatgggta 2040  
ctgggaccct ggtcctcgag ctgaccctt tgaggacatg tggtaoctct gggggggett 2100  
cgcctacttg caggatgtgg tggagcaggc aatcatcagg gtgctgacgg gcaccgagaa 2160  
gaaaactggt gtctatatgc aacagatgcc ctatccctgt tacgttgatg acatctttct 2220  
gogggatgag agccggctca tgcocctott catgacgctg gcctggattt actcagtggc 2280  
tgtgatcatc aagggcatcg tgtatgagaa ggaggcacgg ctgaaagaga ccatgaggat 2340  
catgggcctg gacaacagca tactctgggt tagctgggtc attagtagoc tcattcctct 2400  
tcttgtagc gctggcctgc tagtgggtcat cctgaagtta ggaaacctgc tgcctacag 2460  
tgatcccagc gtgggtgttg tcttctgctc cgtgtttgct gtggtgacaa tctgacagt 2520  
cttctgatt agcacactct totocagagc caacctggca gcagcctgtg ggggcatcat 2580  
ctacttcaag ctgtacctgc cctacgtcct gtgtgtggca tggcaggact acgtgggett 2640  
cacactcaag atcttcgcta gcctgctgct tctgtggct tttgggttg gctgtgagta 2700  
ctttgcctt tttgaggagc agggcattgg agtgcagtgg gacaacctgt ttgagagtcc 2760  
tgtggaggaa gatggcttca atctcaccac ttgatctcc atgatgctgt ttgacacctt 2820  
cctctatggg gtgatgacct ggtacattga ggctgtctt ccaggccagt acggaattcc 2880  
caggccctgg tattttcctt gcaccaagtc ctactggtt ggagaggaaa gtgatgagaa 2940  
gagccacct ggttccaacc agaagagaat gtcagaaatc tgcatggagg aggaaccac 3000  
ccacttgaag ctgggcgtgt ccattcagaa cctggtaaaa gtctaccgag atgggatgaa 3060  
ggtggctgct gatggcctg cactgaattt ttatgagggc cagatcacct ccttctggg 3120

ccacaatgga gcggggaaga cgaccacccat gtcaatcctg accgggttgt tccccccgac 3180  
ctcgggcacc gcctacatcc tgggaaaaga cattcgctct gagatgagca ccatccggca 3240  
gaacctgggg gtctgtcccc agcataacgt gctgtttgac atgctgactg tcgaagaaca 3300  
catctggttc tatgcccgct tgaaagggtt ctctgagaag cacgtgaagg cggagatgga 3360  
gcagatggcc ctggatggtg gtttgccatc aagcaagctg aaaagcaaaa caagccagct 3420  
gtcaggtgga atgcagagaa agctatctgt ggccttggcc tttgtcgggg gatctaaggt 3480  
tgtcattctg gatgaacca cagctgggtg ggacccttac tcccgcaggg gaatatggga 3540  
gctgctgctg aaataccgac aaggccgcac cattattctc tctacacacc acatggatga 3600  
agcgggacgtc ctgggggaca ggattgccat catctcccat gggagctgt gctgtgtggg 3660  
ctcctccctg tttctgaaga accagctggg aacaggctac tacctgacct tgggtcaagaa 3720  
agatgtggaa tcctccctca gttcctgcag aacagtagt agcactgtgt catacctgaa 3780  
aaaggaggac agtgtttctc agagcagttc tgatgctggc ctgggcagcg accatgagag 3840  
tgacacgctg accatcgatg totctgctat ctccaacctc atcaggaagc atgtgtotga 3900  
agccccgctg gtggaagaca tagggcatga gctgacctat gtgctgcat atgaagctgc 3960  
taaggaggga gcctttgtgg aactotttca tgagattgat gaccggctct cagacctggg 4020  
catttctagt tatggcatct cagagacgac cctggaagaa atattcctca aggtggccga 4080  
agagagtggg gtggatgctg agacctcaga tgggtacctg ccagcaagac gaaacaggcg 4140  
ggccttcggg gacaagcaga gctgtottcg cccgttact gaagatgatg ctgctgatcc 4200  
aaatgattct gacatagacc cagaatccag agagacagac ttgctcagtg ggatggatgg 4260  
caaagggtcc taccaggtga aaggctggaa acttacacag caacagtttg tggccctttt 4320  
gtggaagaga ctgctaattg ccagacggag tcggaaagga tttttgctc agattgtctt 4380  
gccagctgtg tttgtctgca ttgcccttgt gttcagcctg atcgtgccac cctttggcaa 4440  
gtaccccagc ctggaacttc agccctggat gtacaacgaa cagtacacat ttgtcagcaa 4500  
tgatgctcct gaggacacgg gaaccctgga actcttaaac gccctacca aagaccctgg 4560  
cttcgggacc cgctgtatgg aaggaaacc aatcccagac acgccctgcc aggcagggga 4620  
ggaagagtgg accactgccc cagttcccca gaccatcatg gacctcttc agaatgggaa 4680  
ctggacaatg cagaaccctt cacctgcatg ccagtgtagc agcgacaaaa tcaagaagat 4740  
gctgcctgtg tgtccccag gggcaggggg gctgcctcct ccacaaagaa aacaaaacac 4800  
tgcagatata cttcaggacc tgacaggaag aaacattctg gattatctgg tgaagacgta 4860  
tgtgcagatc atagccaaaa gcttaaagaa caagatctgg gtgaatgagt ttaggtatgg 4920  
cggcttttcc ctgggtgtca gtaatactca agcacttctc ccgagtcaag aagttaatga 4980  
tgccatcaaa caaatgaaga aacacctaaa gctggccaag gacagttctg cagatcgatt 5040

tctcaacagc ttggaagat ttatgacagg actggacacc agaaataatg tcaaggtgtg 5100  
gttcaataac aagggtggc atgcaatcag ctctttcctg aatgtcatca acaatgccat 5160  
tctccgggcc aacctgcaa agggagagaa ccctagccat tatggaatta ctgctttcaa 5220  
tcatcccctg aatctcacca agcagcagct ctcagagggtg gctctgatga ccacatcagt 5280  
ggatgtcctt gtgtccatct gtgtcatctt tgcaatgtcc ttcgtcccag ccagctttgt 5340  
cgtattcctg atccaggagc gggtcagcaa agcaaaacac ctgcagttca tcagtggagt 5400  
gaagcctgtc atctactggc tctctaattt tgtctgggat atgtgcaatt acgttgtccc 5460  
tgccacactg gtcattatca tcttcatctg cttccagcag aagtcctatg tgtcctccac 5520  
caatctgcct gtgctagccc ttctaactttt gctgtatggg tggtaaatca caoctctcat 5580  
gtaccagcc tcctttgtgt tcaagatccc cagcacagcc tatgtggtgc tcaccagcgt 5640  
gaacctcttc attggcatta atggcagcgt ggcaccttt gtgctggagc tgttcaccga 5700  
caataagctg aataatatca atgatatact gaagtccgtg ttcttgatct tcccacattt 5760  
ttgcctggga cgagggtca tcgacatggt gaaaaaccag gcaatggctg atgccctgga 5820  
aaggtttggg gagaatcgtt ttgtgtcacc attatcttgg gacttgggtg gaogaaacct 5880  
cttcgccatg gccgtggaag ggggtggtgtt cttcctcatt actgttctga tccagtacag 5940  
attcttcatc aggcccagac ctgtaaatgc aaagctatct cctctgaatg atgaagatga 6000  
agatgtgagg cgggaaagac agagaattct tgatggtgga ggcagaatg acatcttaga 6060  
aatcaaggag ttgacgaaga tatatagaag gaagcgggaag cctgctggtg acaggatttg 6120  
cgtgggcatt cctcctgggtg agtgcttttg gctcctggga gttaatgggg ctggaaaatc 6180  
atcaactttc aagatgtaa caggagatac cactgttacc agaggagatg ctttcttaa 6240  
caaaaatagt atcttatcaa acatccatga agtacatcag aacatgggct actgcctca 6300  
gtttgatgcc atcacagagc tgttgactgg gagagaacac gtggagttct ttgccctttt 6360  
gagaggagtc ccagagaaag aagttggcaa ggttgggtgag tgggcgattc ggaaaactggg 6420  
octogtgaag tatggagaaa aatatgctgg taactatagt ggaggcaaca aacgcaagct 6480  
ctctacagcc atggctttga tcggcgggcc tctgtggtg tttctggatg aaccaccac 6540  
aggcatggat cccaaagccc ggcggttctt gtggaattgt gccctaagtg ttgtcaagga 6600  
ggggagatca gtagtgctta catctcatag tatggaagaa tgtgaagctc tttgcaactag 6660  
gatggcaatc atggtcaatg gaaggttcag gtgccttggc agtgtccagc atctaaaaaa 6720  
taggtttgga gatgggtata caatagttgt acgaatagca gggccaacc cggacctgaa 6780  
gcctgtccag gatttcttg gacttgcat tcttgggaagt gttctaaaag agaaacaccg 6840  
gaacatgcta caataccagc ttcatcttc attatcttct ctggccagga tattcagcat 6900  
cctctcccag agcaaaaagc gactccacat agaagactac tctgtttctc agacaacact 6960

tgaccaagta tttgtgaact ttgccaagga ccaaagtgat gatgaccact taaaagacct 7020  
 ctcatcac aaaaaccaga cagtagtgga cgttgagtt ctcatcctt ttctacagga 7080  
 tgagaaagt aaagaaagct atgtatgaag aatcctgttc atacgggggtg gctgaaagta 7140  
 aagaggaact agactttcct ttgcacatg tgaagtgttgg tggagaaaag agccagaagt 7200  
 tgatgtggga agaagtaaac tggatactgt actgatacta ttcaatgcaa tgcaattcaa 7260  
 tgcaatgaaa acaaaattcc attacagggg cagtgccttt gtagcctatg tcttgtatgg 7320  
 ctctcaagt aaagacttga atttagtttt ttacctatac ctatgtgaaa ctctattatg 7380  
 gaaccaatg gacatatggg tttgaactca cacttttttt tttttttgt tctgtgtat 7440  
 tctcattggg gttgcaacaa taattcatca agtaatcatg gccagcgatt attgatcaaa 7500  
 atcaaaaggt aatgcacatc ctcatcact aagccatgcc atgccagga gactggtttc 7560  
 ccggtgacac atccattgct ggcaatgagt gtgccagagt tattagtgcc aagtttttca 7620  
 gaaagtttga agcaccatgg tgtgtcatgc tcacttttgt gaaagctgct ctgctcagag 7680  
 tctatcaaca ttgaatatca gttgacagaa tggtgccatg cgtgggtaac atcctgcttt 7740  
 gattccctct gataagctgt tctgggtggca gtaacatgca acaaaaatgt ggggtgtctct 7800  
 aggcacggga aacttggttc cattgttata ttgtcctatg cttcgagcca tgggtctaca 7860  
 gggtcacct tatgagactc ttaaatac ttagatcctg gtaagaggca aagaatcaac 7920  
 agccaaactg ctggggctgc aagctgctga agccagggca tgggattaaa gagattgtgc 7980  
 gttcaaacct agggaagcct gtgccattt gtccctgactg tctgctaaca tggtagactg 8040  
 catctcaaga tgtttatctg acacaagtgt attatttctg gctttttgaa ttaatctaga 8100  
 aatgaaaag atggagtgt atttgacaa aatgtttgt actttttaat gttatttga 8160  
 attttaagtt ctatcagtga cttctgaatc cttagaatgg cctctttgta gaaccctgtg 8220  
 gtatagagga gtatggccac tgccccacta tttttatfff cttatgtaag tttgcatatc 8280  
 agtcatgact agtgcctaga aagcaatgtg atggtcagga totcatgaca ttatatttga 8340  
 gtttctttca gatcatttag gatactotta atctcacttc atcaatcaaa tattttttga 8400  
 gtgtatgctg tagctgaaag agtatgtacg taogtataag actagagaga tattaagtct 8460  
 cagtacactt cctgtgcat gttattcagc tcaactggtt acaaatatag gttgtcttgt 8520  
 ggtttagga gccactgta acaatattgg gcagcctttt ttttttttt ttaattgcaa 8580  
 caatgcaaaa gccaaagaaag tataagggtc acaagtttaa acaatgaatt cttcaacagg 8640  
 gaaaacagct agcttgaaaa cttgctgaaa aacacaactt gtgtttatgg catttagtac 8700  
 cttcaataa ttggcttgc agatattgga taccocatta aatctgacag tctcaaattt 8760  
 ttcatctctt caatcactag tcaagaaaa tataaaaaca acaataactt ccatatggag 8820  
 catttttcag agttttctaa cccagtotta tttttctagt cagtaaacad ttgtaaaaat 8880

actgtttcac taactttac tgttaactgt cttgagagaa aagaaaaata tgagagaact 8940  
attgtttggg gaagttcaag tgatctttca atatcattac taactttctc cactttttcc 9000  
aaaatttgaa tattaacgct aaaggtgtaa gacttcagat ttcaaattaa tctttctata 9060  
ttttttaaat ttacagaata ttatataacc cactgctgaa aaagaaaaaa atgattgttt 9120  
tagaagttaa agtcaatatt gatttttaaat ataagtaatg aaggcatatt tccaataact 9180  
agtgatatgg catcgttgca atttacagta tcttcaaaaa tacagaattt atagaataat 9240  
ttctoctcat ttaatatttt tcaaaatcaa agttatgggt tctcatttt actaaaatcg 9300  
tattotaatt cttcattata gtaaactctat gagcaactcc ttacttcggg tctctgatt 9360  
tcaaggccat attttaaaaa atcaaaaggc actgtgaact attttgaaga aaacacgaca 9420  
ttttaataca gattgaaagg acctottctg aagctagaaa caatctatag ttatacatct 9480  
tcattaatac tgtgttacct tttaaaatag taatttttta cattttctctg tgtaaacct 9540  
attgtggtag aaatttttac caactctata ctcaatcaag caaaatttct gtatattccc 9600  
tgtggaatgt acctatgtga gtttcagaaa ttotcaaaat acgtgttcaa aaatttctgc 9660  
ttttgcatct ttgggacacc tcagaaaact tattaacaac tgtgaatatg agaaatacag 9720  
aagaaaaata taagccctct atacataaat gccagcaca attcattggt aaaaaacaac 9780  
caaacctcac actactgtat ttcattatct gtactgaaag caaatgcttt gtgactatta 9840  
aatgttgcac atcattcatt cactgtatag taatcattga ctaaagccat ttgctgtggt 9900  
ttctttctgt ggntgnatat atcaggtaaa atattttcca aagagccatg tgtcatgtaa 9960  
tactgaacc tttgatattg agacattaat ttggaccctt ggtattatct actagaataa 10020  
tgtaatactg nagaaatatt gctctaattc tttcaaaatg gtgcatcccc cttaaaangt 10080  
tctattttcca taaggattta gcttgcttat cccttcttat accctaagat gaagctgttt 10140  
ttgtgctctt tgttcatcat tggccctcat tccaagcact ttacgctgtc tgtaatggga 10200  
tctatTTTTg cactggaata tctgagaatt gcaaaaactag acaaagttt cacaaacagat 10260  
ttotaagtta aatcattttc attaaaagga aaaaagaaaa aaaattttgt atgtcaataa 10320  
ctttatatga agtattaaaa tgcataattc tatgttgtaa tataatgagt cacaaaataa 10380  
agctgtgaca gttctgtaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 10440  
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 10474

<210> 10

<211> 2261

<212> PRT

<213> Homo sapiens

<400> 10

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

Ala Thr Lys Thr Leu Leu His Ser Leu Gly Thr Leu Ala Gln Glu Leu  
                           260                          265                          270  
 Phe Ser Met Arg Ser Trp Ser Asp Met Arg Gln Glu Val Met Phe Leu  
                           275                          280                          285  
 Thr Asn Val Asn Ser Ser Ser Ser Ser Thr Gln Ile Tyr Gln Ala Val  
                           290                          295                          300  
 Ser Arg Ile Val Cys Gly His Pro Glu Gly Gly Gly Leu Lys Ile Lys  
 305                          310                          315                          320  
 Ser Leu Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Leu Phe Gly Gly  
                           325                          330                          335  
 Asn Gly Thr Glu Glu Asp Ala Glu Thr Phe Tyr Asp Asn Ser Thr Thr  
                           340                          345                          350  
 Pro Tyr Cys Asn Asp Leu Met Lys Asn Leu Glu Ser Ser Pro Leu Ser  
                           355                          360                          365  
 Arg Ile Ile Trp Lys Ala Leu Lys Pro Leu Leu Val Gly Lys Ile Leu  
                           370                          375                          380  
 Tyr Thr Pro Asp Thr Pro Ala Thr Arg Gln Val Met Ala Glu Val Asn  
 385                          390                          395                          400  
 Lys Thr Phe Gln Glu Leu Ala Val Phe His Asp Leu Glu Gly Met Trp  
                           405                          410                          415  
 Glu Glu Leu Ser Pro Lys Ile Trp Thr Phe Met Glu Asn Ser Gln Glu  
                           420                          425                          430  
 Met Asp Leu Val Arg Met Leu Leu Asp Ser Arg Asp Asn Asp His Phe  
                           435                          440                          445  
 Trp Glu Gln Gln Leu Asp Gly Leu Asp Trp Thr Ala Gln Asp Ile Val  
                           450                          455                          460  
 Ala Phe Leu Ala Lys His Pro Glu Asp Val Gln Ser Ser Asn Gly Ser  
 465                          470                          475                          480  
 Val Tyr Thr Trp Arg Glu Ala Phe Asn Glu Thr Asn Gln Ala Ile Arg  
                           485                          490                          495  
 Thr Ile Ser Arg Phe Met Glu Cys Val Asn Leu Asn Lys Leu Glu Pro  
                           500                          505                          510

Ile Ala Thr Glu Val Trp Leu Ile Asn Lys Ser Met Glu Leu Leu Asp  
 515 520 525  
 Glu Arg Lys Phe Trp Ala Gly Ile Val Phe Thr Gly Ile Thr Pro Gly  
 530 535 540  
 Ser Ile Glu Leu Pro His His Val Lys Tyr Lys Ile Arg Met Asp Ile  
 545 550 555 560  
 Asp Asn Val Glu Arg Thr Asn Lys Ile Lys Asp Gly Tyr Trp Asp Pro  
 565 570 575  
 Gly Pro Arg Ala Asp Pro Phe Glu Asp Met Trp Tyr Val Trp Gly Gly  
 580 585 590  
 Phe Ala Tyr Leu Gln Asp Val Val Glu Gln Ala Ile Ile Arg Val Leu  
 595 600 605  
 Thr Gly Thr Glu Lys Lys Thr Gly Val Tyr Met Gln Gln Met Pro Tyr  
 610 615 620  
 Pro Cys Tyr Val Asp Asp Ile Phe Leu Arg Val Met Ser Arg Ser Met  
 625 630 635 640  
 Pro Leu Phe Met Thr Leu Ala Trp Ile Tyr Ser Val Ala Val Ile Ile  
 645 650 655  
 Lys Gly Ile Val Tyr Glu Lys Glu Ala Arg Leu Lys Glu Thr Met Arg  
 660 665 670  
 Ile Met Gly Leu Asp Asn Ser Ile Leu Trp Phe Ser Trp Phe Ile Ser  
 675 680 685  
 Ser Leu Ile Pro Leu Leu Val Ser Ala Gly Leu Leu Val Val Ile Leu  
 690 695 700  
 Lys Leu Gly Asn Leu Leu Pro Tyr Ser Asp Pro Ser Val Val Phe Val  
 705 710 715 720  
 Phe Leu Ser Val Phe Ala Val Val Thr Ile Leu Gln Cys Phe Leu Ile  
 725 730 735  
 Ser Thr Leu Phe Ser Arg Ala Asn Leu Ala Ala Ala Cys Gly Gly Ile  
 740 745 750  
 Ile Tyr Phe Thr Leu Tyr Leu Pro Tyr Val Leu Cys Val Ala Trp Gln  
 755 760 765

Asp Tyr Val Gly Phe Thr Leu Lys Ile Phe Ala Ser Leu Leu Ser Pro  
 770 775 780  
 Val Ala Phe Gly Phe Gly Cys Glu Tyr Phe Ala Leu Phe Glu Glu Gln  
 785 790 795 800  
 Gly Ile Gly Val Gln Trp Asp Asn Leu Phe Glu Ser Pro Val Glu Glu  
 805 810 815  
 Asp Gly Phe Asn Leu Thr Thr Ser Ile Ser Met Met Leu Phe Asp Thr  
 820 825 830  
 Phe Leu Tyr Gly Val Met Thr Trp Tyr Ile Glu Ala Val Phe Pro Gly  
 835 840 845  
 Gln Tyr Gly Ile Pro Arg Pro Trp Tyr Phe Pro Cys Thr Lys Ser Tyr  
 850 855 860  
 Trp Phe Gly Glu Glu Ser Asp Glu Lys Ser His Pro Gly Ser Asn Gln  
 865 870 875 880  
 Lys Arg Met Ser Glu Ile Cys Met Glu Glu Glu Pro Thr His Leu Lys  
 885 890 895  
 Leu Gly Val Ser Ile Gln Asn Leu Val Lys Val Tyr Arg Asp Gly Met  
 900 905 910  
 Lys Val Ala Val Asp Gly Leu Ala Leu Asn Phe Tyr Glu Gly Gln Ile  
 915 920 925  
 Thr Ser Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Met Ser  
 930 935 940  
 Ile Leu Thr Gly Leu Phe Pro Pro Thr Ser Gly Thr Ala Tyr Ile Leu  
 945 950 955 960  
 Gly Lys Asp Ile Arg Ser Glu Met Ser Thr Ile Arg Gln Asn Leu Gly  
 965 970 975  
 Val Cys Pro Gln His Asn Val Leu Phe Asp Met Leu Thr Val Glu Glu  
 980 985 990  
 His Ile Trp Phe Tyr Ala Arg Leu Lys Gly Leu Ser Glu Lys His Val  
 995 1000 1005  
 Lys Ala Glu Met Glu Gln Met Ala Leu Asp Val Gly Leu Pro Ser Ser  
 1010 1015 1020

Lys Leu Lys Ser Lys Thr Ser Gln Leu Ser Gly Gly Met Gln Arg Lys  
 1025                      1030                      1035                      1040  
 Leu Ser Val Ala Leu Ala Phe Val Gly Gly Ser Lys Val Val Ile Leu  
                             1045                      1050                      1055  
 Asp Glu Pro Thr Ala Gly Val Asp Pro Tyr Ser Arg Arg Gly Ile Trp  
                             1060                      1065                      1070  
 Glu Leu Leu Leu Lys Tyr Arg Gln Gly Arg Thr Ile Ile Leu Ser Thr  
                             1075                      1080                      1085  
 His His Met Asp Glu Ala Asp Val Leu Gly Asp Arg Ile Ala Ile Ile  
                             1090                      1095                      1100  
 Ser His Gly Lys Leu Cys Cys Val Gly Ser Ser Leu Phe Leu Lys Asn  
 1105                      1110                      1115                      1120  
 Gln Leu Gly Thr Gly Tyr Tyr Leu Thr Leu Val Lys Lys Asp Val Glu  
                             1125                      1130                      1135  
 Ser Ser Leu Ser Ser Cys Arg Asn Ser Ser Ser Thr Val Ser Tyr Leu  
                             1140                      1145                      1150  
 Lys Lys Glu Asp Ser Val Ser Gln Ser Ser Ser Asp Ala Gly Leu Gly  
                             1155                      1160                      1165  
 Ser Asp His Glu Ser Asp Thr Leu Thr Ile Asp Val Ser Ala Ile Ser  
                             1170                      1175                      1180  
 Asn Leu Ile Arg Lys His Val Ser Glu Ala Arg Leu Val Glu Asp Ile  
 1185                      1190                      1195                      1200  
 Gly His Glu Leu Thr Tyr Val Leu Pro Tyr Glu Ala Ala Lys Glu Gly  
                             1205                      1210                      1215  
 Ala Phe Val Glu Leu Phe His Glu Ile Asp Asp Arg Leu Ser Asp Leu  
                             1220                      1225                      1230  
 Gly Ile Ser Ser Tyr Gly Ile Ser Glu Thr Thr Leu Glu Glu Ile Phe  
                             1235                      1240                      1245  
 Leu Lys Val Ala Glu Glu Ser Gly Val Asp Ala Glu Thr Ser Asp Gly  
                             1250                      1255                      1260  
 Thr Leu Pro Ala Arg Arg Asn Arg Arg Ala Phe Gly Asp Lys Gln Ser  
 1265                      1270                      1275                      1280

Cys Leu Arg Pro Phe Thr Glu Asp Asp Ala Ala Asp Pro Asn Asp Ser  
1285 1290 1295  
Asp Ile Asp Pro Glu Ser Arg Glu Thr Asp Leu Leu Ser Gly Met Asp  
1300 1305 1310  
Gly Lys Gly Ser Tyr Gln Val Lys Gly Trp Lys Leu Thr Gln Gln Gln  
1315 1320 1325  
Phe Val Ala Leu Leu Trp Lys Arg Leu Leu Ile Ala Arg Arg Ser Arg  
1330 1335 1340  
Lys Gly Phe Phe Ala Gln Ile Val Leu Pro Ala Val Phe Val Cys Ile  
1345 1350 1355 1360  
Ala Leu Val Phe Ser Leu Ile Val Pro Pro Phe Gly Lys Tyr Pro Ser  
1365 1370 1375  
Leu Glu Leu Gln Pro Trp Met Tyr Asn Glu Gln Tyr Thr Phe Val Ser  
1380 1385 1390  
Asn Asp Ala Pro Glu Asp Thr Gly Thr Leu Glu Leu Leu Asn Ala Leu  
1395 1400 1405  
Thr Lys Asp Pro Gly Phe Gly Thr Arg Cys Met Glu Gly Asn Pro Ile  
1410 1415 1420  
Pro Asp Thr Pro Cys Gln Ala Gly Glu Glu Glu Trp Thr Thr Ala Pro  
1425 1430 1435 1440  
Val Pro Gln Thr Ile Met Asp Leu Phe Gln Asn Gly Asn Trp Thr Met  
1445 1450 1455  
Gln Asn Pro Ser Pro Ala Cys Gln Cys Ser Ser Asp Lys Ile Lys Lys  
1460 1465 1470  
Met Leu Pro Val Cys Pro Pro Gly Ala Gly Gly Leu Pro Pro Pro Gln  
1475 1480 1485  
Arg Lys Gln Asn Thr Ala Asp Ile Leu Gln Asp Leu Thr Gly Arg Asn  
1490 1495 1500  
Ile Ser Asp Tyr Leu Val Lys Thr Tyr Val Gln Ile Ile Ala Lys Ser  
1505 1510 1515 1520  
Leu Lys Asn Lys Ile Trp Val Asn Glu Phe Arg Tyr Gly Gly Phe Ser  
1525 1530 1535

Leu Gly Val Ser Asn Thr Gln Ala Leu Pro Pro Ser Gln Glu Val Asn  
 1540 1545 1550  
 Asp Ala Ile Lys Gln Met Lys Lys His Leu Lys Leu Ala Lys Asp Ser  
 1555 1560 1565  
 Ser Ala Asp Arg Phe Leu Asn Ser Leu Gly Arg Phe Met Thr Gly Leu  
 1570 1575 1580  
 Asp Thr Arg Asn Asn Val Lys Val Trp Phe Asn Asn Lys Gly Trp His  
 1585 1590 1595 1600  
 Ala Ile Ser Ser Phe Leu Asn Val Ile Asn Asn Ala Ile Leu Arg Ala  
 1605 1610 1615  
 Asn Leu Gln Lys Gly Glu Asn Pro Ser His Tyr Gly Ile Thr Ala Phe  
 1620 1625 1630  
 Asn His Pro Leu Asn Leu Thr Lys Gln Gln Leu Ser Glu Val Ala Leu  
 1635 1640 1645  
 Met Thr Thr Ser Val Asp Val Leu Val Ser Ile Cys Val Ile Phe Ala  
 1650 1655 1660  
 Met Ser Phe Val Pro Ala Ser Phe Val Val Phe Leu Ile Gln Glu Arg  
 1665 1670 1675 1680  
 Val Ser Lys Ala Lys His Leu Gln Phe Ile Ser Gly Val Lys Pro Val  
 1685 1690 1695  
 Ile Tyr Trp Leu Ser Asn Phe Val Trp Asp Met Cys Asn Tyr Val Val  
 1700 1705 1710  
 Pro Ala Thr Leu Val Ile Ile Ile Phe Ile Cys Phe Gln Gln Lys Ser  
 1715 1720 1725  
 Tyr Val Ser Ser Thr Asn Leu Pro Val Leu Ala Leu Leu Leu Leu Leu  
 1730 1735 1740  
 Tyr Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala Ser Phe Val Phe  
 1745 1750 1755 1760  
 Lys Ile Pro Ser Thr Ala Tyr Val Val Leu Thr Ser Val Asn Leu Phe  
 1765 1770 1775  
 Ile Gly Ile Asn Gly Ser Val Ala Thr Phe Val Leu Glu Leu Phe Thr  
 1780 1785 1790

Asp Asn Lys Leu Asn Asn Ile Asn Asp Ile Leu Lys Ser Val Phe Leu  
 1795 1800 1805  
 Ile Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Met Val Lys  
 1810 1815 1820  
 Asn Gln Ala Met Ala Asp Ala Leu Glu Arg Phe Gly Glu Asn Arg Phe  
 1825 1830 1835 1840  
 Val Ser Pro Leu Ser Trp Asp Leu Val Gly Arg Asn Leu Phe Ala Met  
 1845 1850 1855  
 Ala Val Glu Gly Val Val Phe Phe Leu Ile Thr Val Leu Ile Gln Tyr  
 1860 1865 1870  
 Arg Phe Phe Ile Arg Pro Arg Pro Val Asn Ala Lys Leu Ser Pro Leu  
 1875 1880 1885  
 Asn Asp Glu Asp Glu Asp Val Arg Arg Glu Arg Gln Arg Ile Leu Asp  
 1890 1895 1900  
 Gly Gly Gly Gln Asn Asp Ile Leu Glu Ile Lys Glu Leu Thr Lys Ile  
 1905 1910 1915 1920  
 Tyr Arg Arg Lys Arg Lys Pro Ala Val Asp Arg Ile Cys Val Gly Ile  
 1925 1930 1935  
 Pro Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys  
 1940 1945 1950  
 Ser Ser Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Arg Gly  
 1955 1960 1965  
 Asp Ala Phe Leu Asn Lys Asn Ser Ile Leu Ser Asn Ile His Glu Val  
 1970 1975 1980  
 His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Thr Glu Leu  
 1985 1990 1995 2000  
 Leu Thr Gly Arg Glu His Val Glu Phe Phe Ala Leu Leu Arg Gly Val  
 2005 2010 2015  
 Pro Glu Lys Glu Val Gly Lys Val Gly Glu Trp Ala Ile Arg Lys Leu  
 2020 2025 2030  
 Gly Leu Val Lys Tyr Gly Glu Lys Tyr Ala Gly Asn Tyr Ser Gly Gly  
 2035 2040 2045



<220>

<223> 人工序列描述: ABC1 扩增引物

<400> 11

cctctcatta cacaaaaacc agac 24

<210> 12

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 扩增引物

<400> 12

gctttctttc acttctcatc ctg 23

<210> 13

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 RT-PCR 引物

<400> 13

tccttggggtt caggggatta tc 22

<210> 14

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 RT-PCR 引物

<400> 14

caatgttttt gtggcttcgg c 21

<210> 15

<211> 40

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 RT-PCR 引物

<400> 15

agtcgagctc caaacatgtc agctgttact ggaagtggcc 40

<210> 16

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 RT-PCR 引物

<400> 16

tctctggatt ctgggtctat gtcag 25

<210> 17

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 RT-PCR 引物

<400> 17

gggagccttt gtggaactct ttc 23

<210> 18

<211> 41

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 RT-PCR 引物

<400> 18

actggtcgac cattgaattg cattgcattg aatagtatca g 41

<210> 19

<211> 19

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 19

tttcctgggtg gacaatgaa 19

<210> 20

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 20

agtgacatgc gacaggag 18

<210> 21

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 21

gatctggaag gcatgtgg 18

<210> 22

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 22

ccaggcagca ttgagctg 18

<210> 23

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 23

ggcctggaca acagcata

18

<210> 24

<211> 19

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 24

ggacaacctg tttgagagt

19

<210> 25

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 25

aagacgacca ccatgtca

18

<210> 26

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 26

atatgggagc tgctgctg

18

<210> 27

<211> 24

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 27

gggcatgagc tgacctatgt gctg

24

<210> 28

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 28

aagagactgc taattgcc

18

<210> 29

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 29

agcgacaaaa tcaagaag

18

<210> 30

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 30

tggcatgcaa tcagctct

18

<210> 31

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 31

tctctccacca atctgcct 18

<210> 32

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 32

ttcttctctca ttactgtt 18

<210> 33

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 33

gatgccatca cagagctg 18

<210> 34

<211> 17

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 34

agtgtccagc atctaaa 17

<210> 35

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 35

caaagttcac aaataactt 18

<210> 36

<211> 19

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 36

cttagggcac aattccaca 19

<210> 37

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 37

tgaaagttga tgattttc 18

<210> 38

<211> 19

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 38

tttttcacca tgtcgatga 19

<210> 39

<211> 17

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 39

ctccactgat gaactgc 17

<210> 40

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 40

gtttcttcat ttgtttga 18

<210> 41

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 41

agggcgtgtc tgggattg 18

<210> 42

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 42

cagaatcatt tggatcag 18

<210> 43

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 43

catcagaact gctctgag 18

<210> 44

<211> 19

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 44

agctggcttg ttttgcttt 19

<210> 45

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 45

tggacacgcc cagettca 18

<210> 46

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 46

cctgccatgc cacacaca 18

<210> 47

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 47

ctcatcaccc gcagaaag 18

<210> 48

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 48

cacactccat gaagcgag 18

<210> 49

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 49

tccagataat gcgggaaa 18

<210> 50

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 50

tcaggattgg cttcagga 18

<210> 51

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 测序引物

<400> 51

aagtttgagc tggatttctt g

21

<210> 52

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述:  $\beta$ -球蛋白反义寡核苷酸

<400> 52

cctcttacct cagttacaat ttata

25

<210> 53

<211> 26

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: ABC1 反义寡核苷酸

<400> 53

catgttgttc ataggggtggg tagctc

26

<210> 54

<211> 24

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述:  $\beta$ -肌动蛋白扩增引物

<400> 54

tcacccacac tgtgccatct acga

24

<210> 55

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述:  $\beta$ -肌动蛋白 扩增引物

<400> 55

cagcgggaacc gctcattgcc aatgg

25

<210> 56

<211> 26

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: 固醇反应元素寡核苷酸

<400> 56

tcgagtgacc gatagtaacc tctcga

26

<210> 57

<211> 26

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: 突变的固醇反应元素寡核苷酸

<400> 57

tcgagctgca catagtaacc tctcga

26

说明书附图

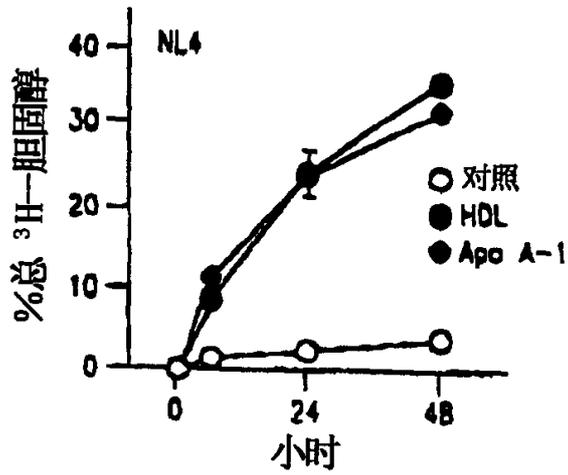


图 1A

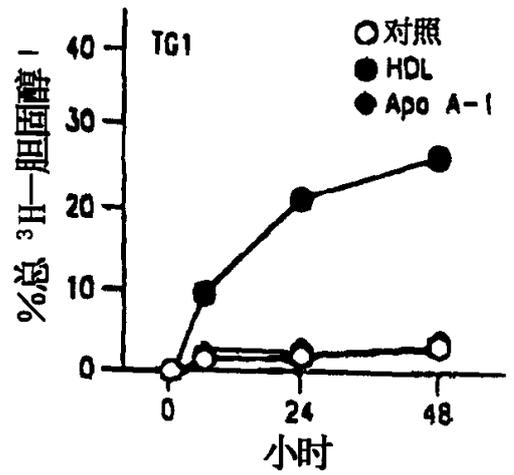


图 1B

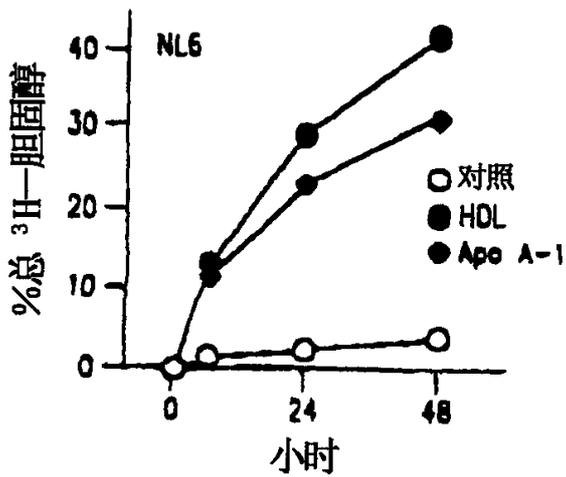


图 1C

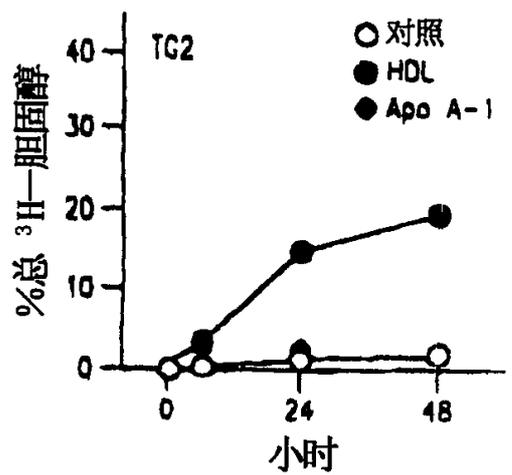
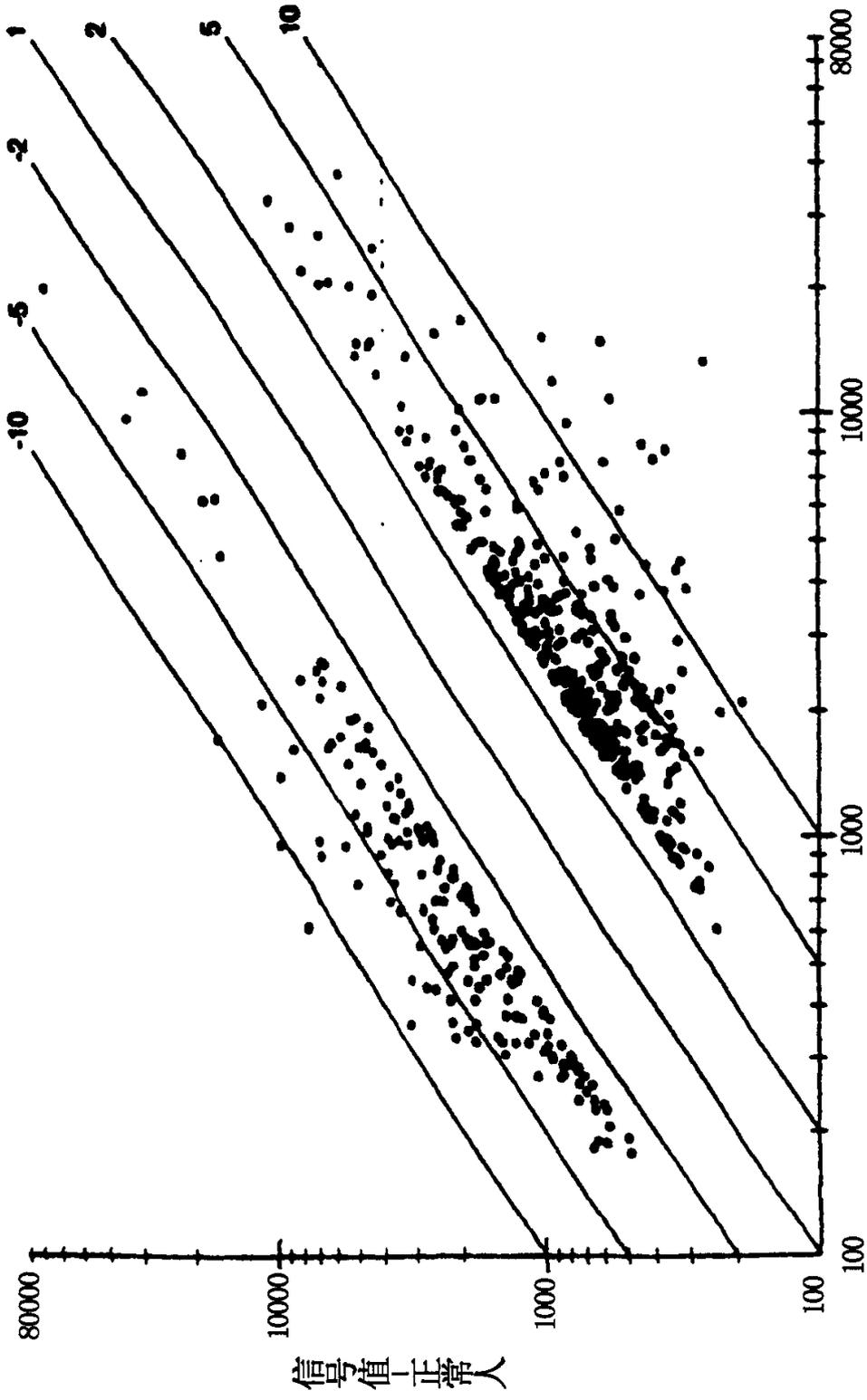


图 1D



信号值—丹尼尔病患者

图 2

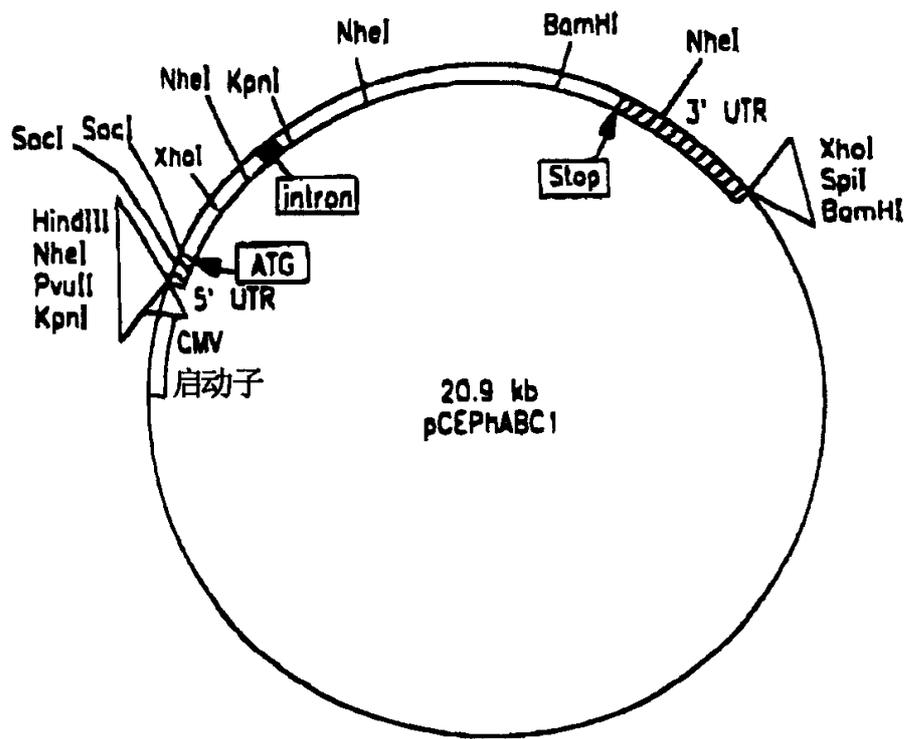
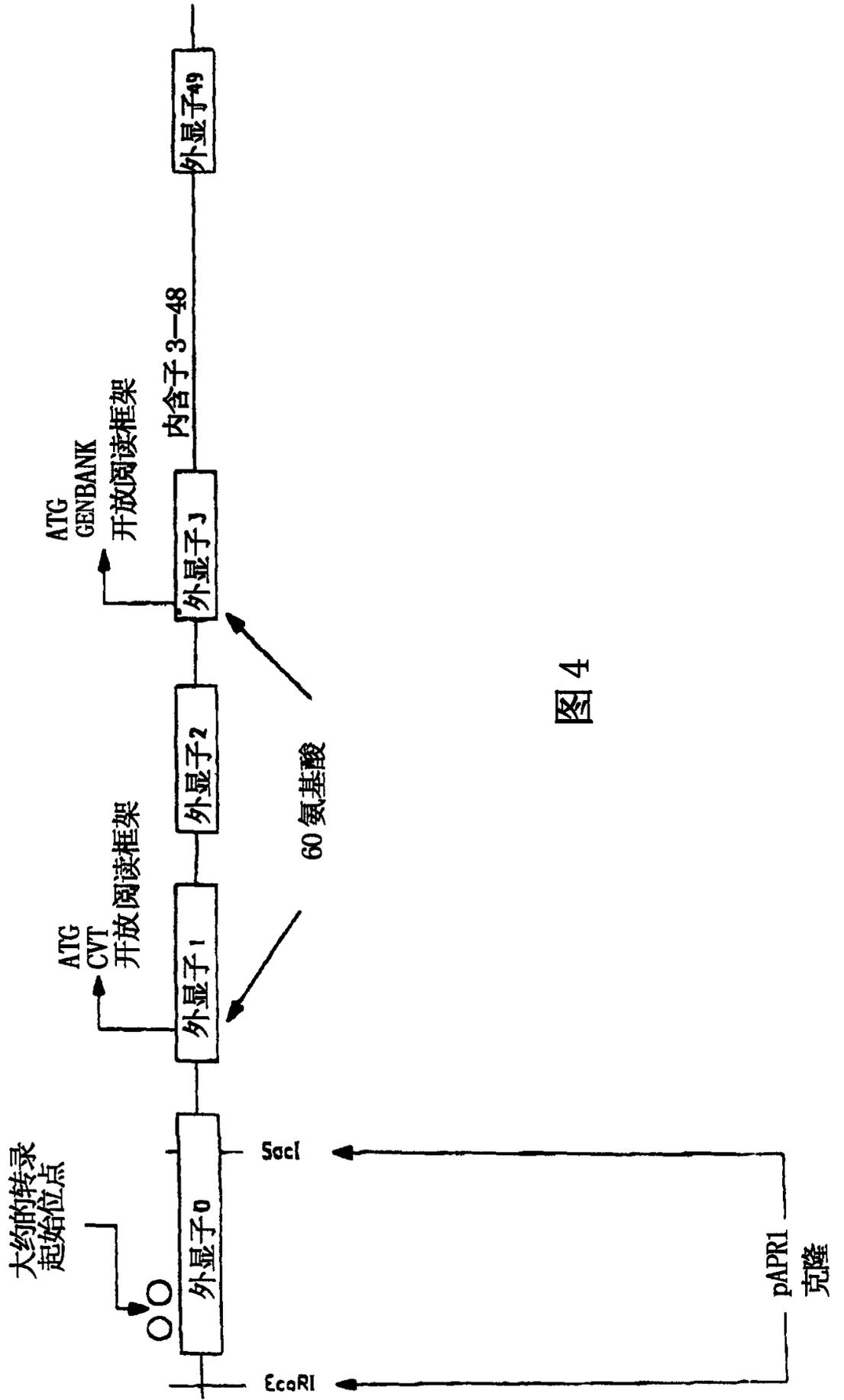


图 3

CVT ABCA1 基因结构



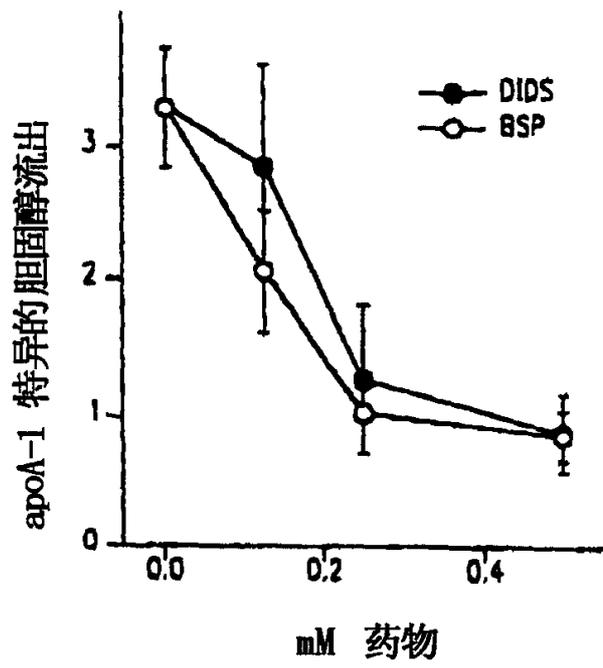


图 5

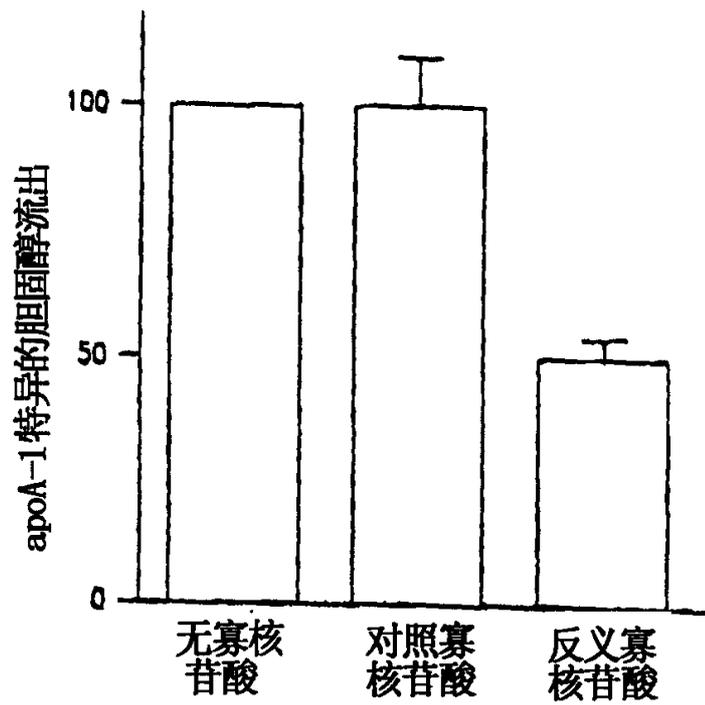


图 6

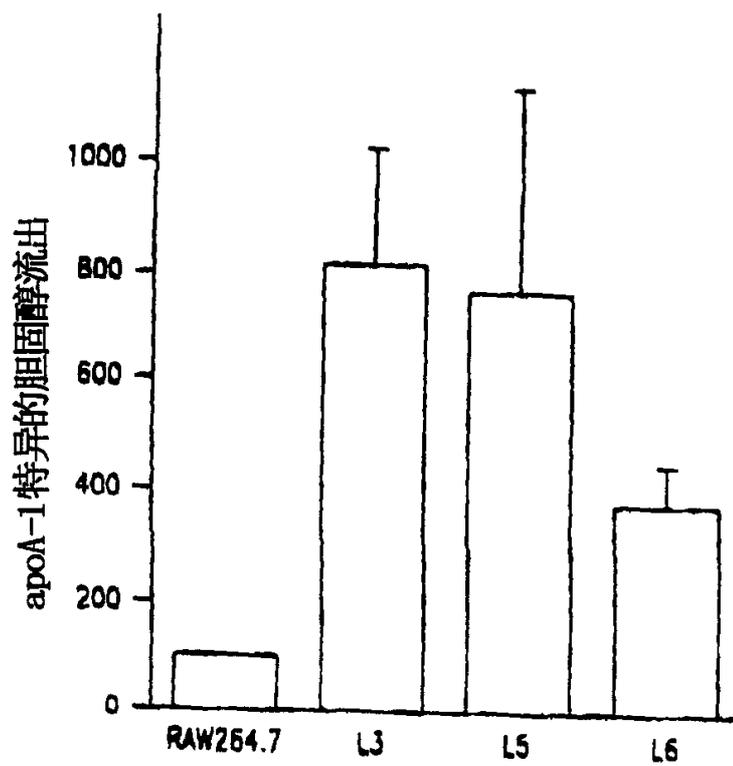


图 7

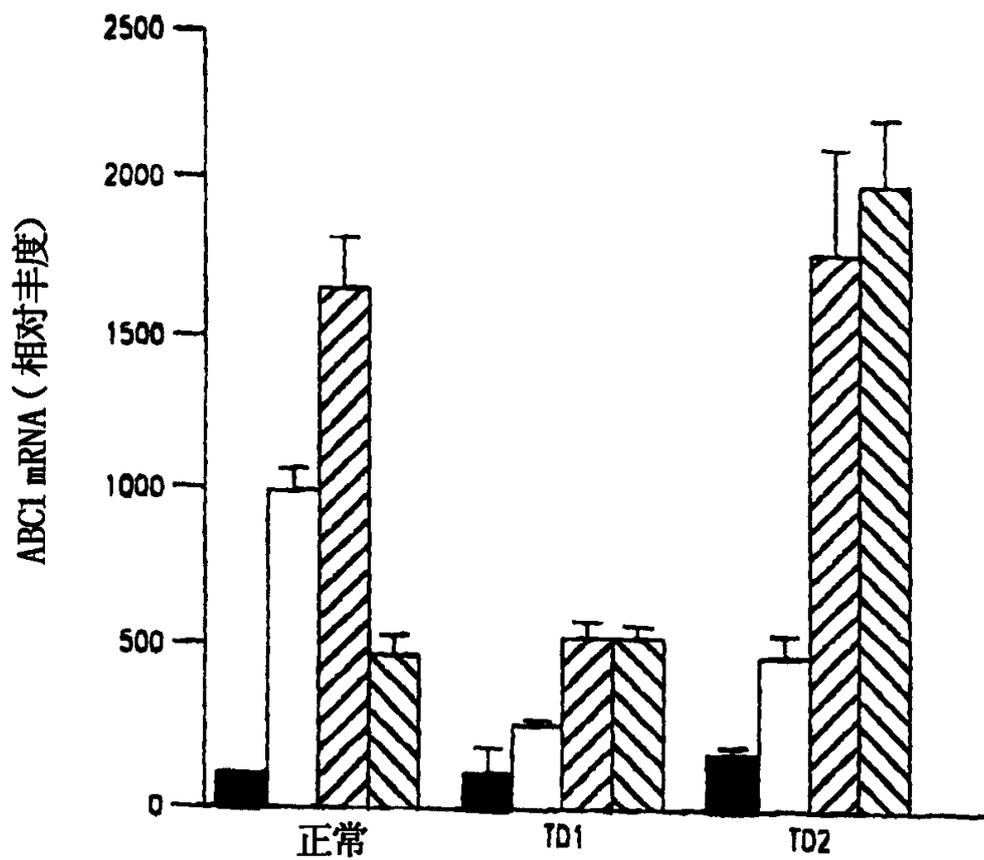


图 8

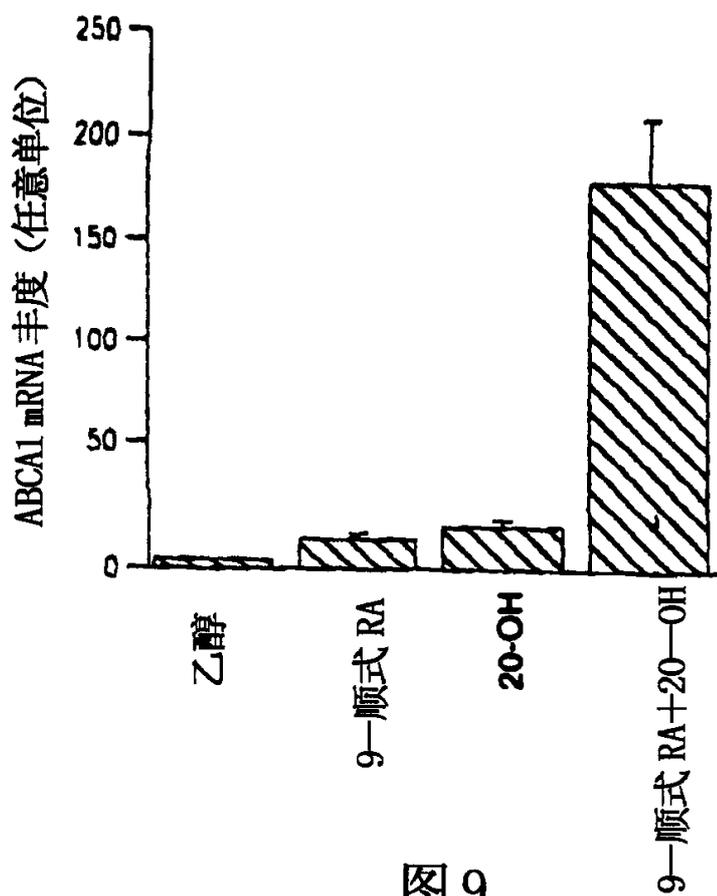


图 9

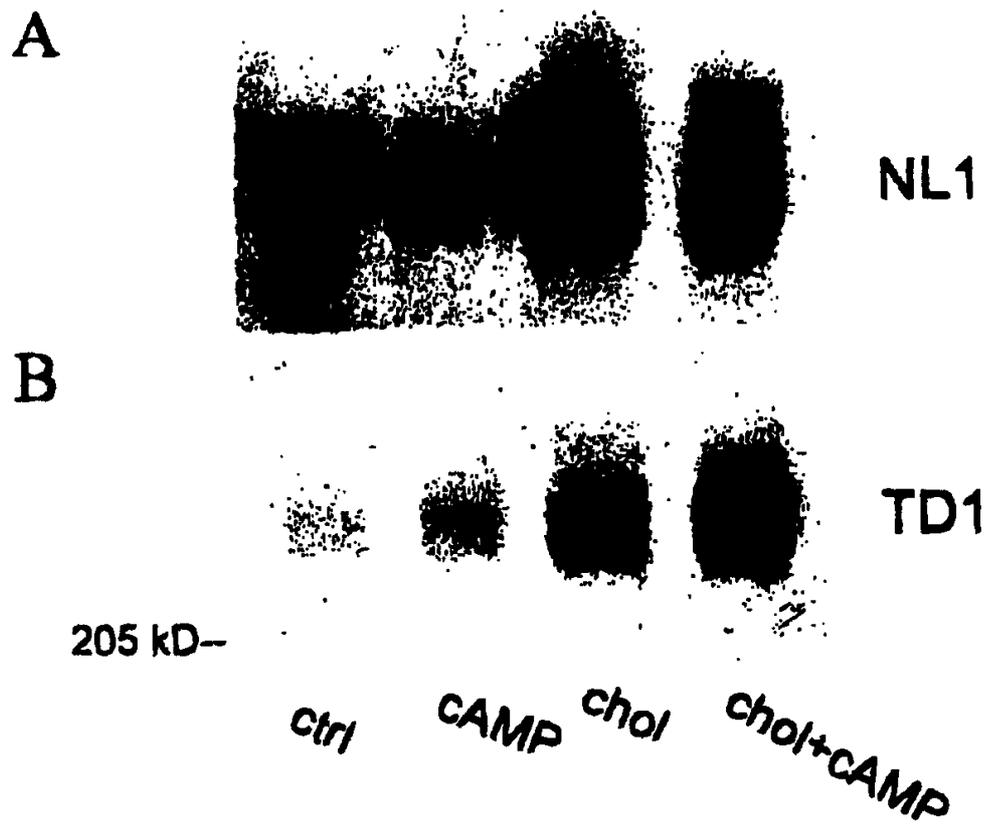


图 10

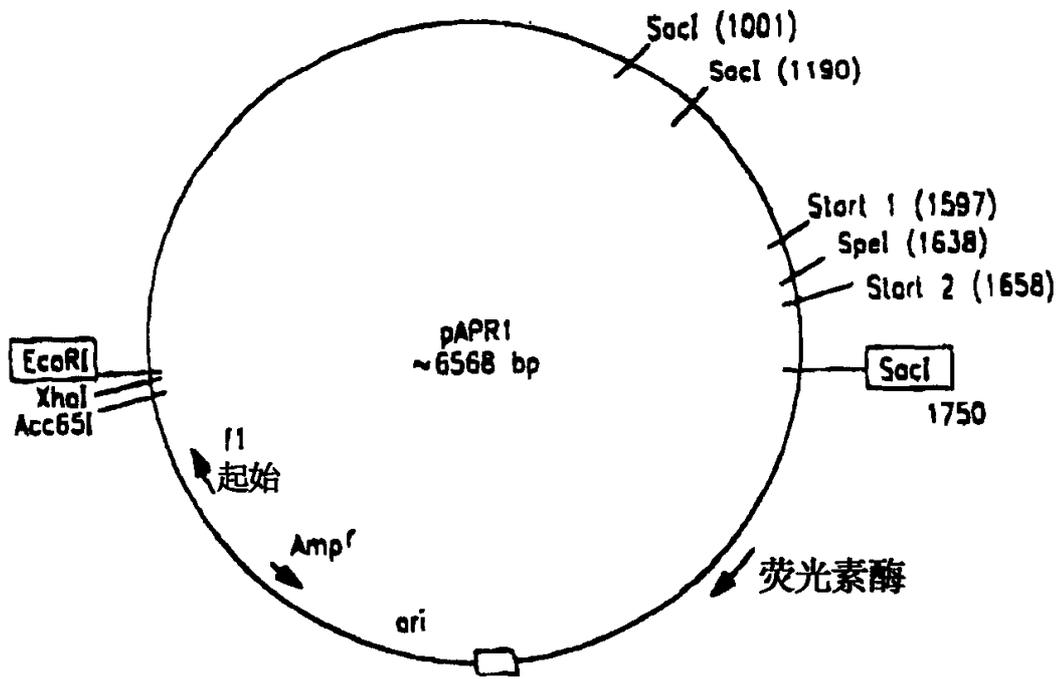


图 11

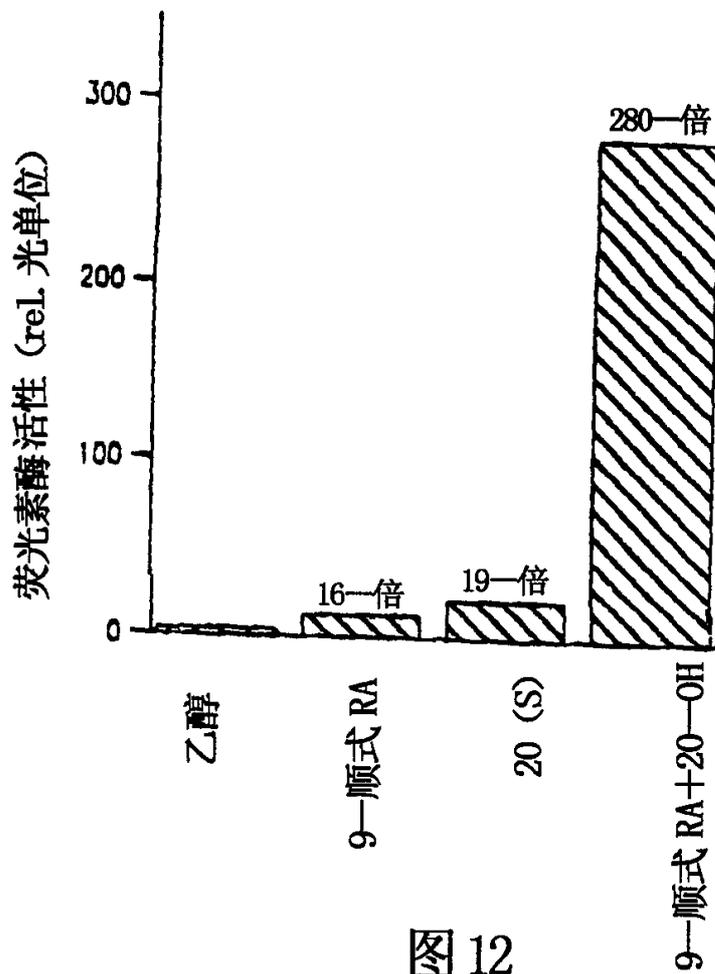


图 12

ABC1 基因 5' 侧翼区

1 GAATTCCTTGCTGGTCCACATGCACATTCACAGGSCCTGCCTTGGCCCTTCATATGGGCTCTGAGTGTTCATAGAACCCACTGATGTGCACTACCTGG  
 101 GCTTGAGCGTGGCCCTGGAGATCCCTGTGTGACTGTATGCAATGGAGGGGGCTTGTACAGCTGAATATGCTGTATGACAGSTGTGGGAGTTCCTGGAAATATGATGAATGGAG  
 201 CTGGAGGTGGCAAGAGAAAGTAGGCTTGGGGCAGCTCTCTCATGSCCACCCTCATTTCTGGCCCAAAACTGTCAGGTCAAACTGTGCAAGAGCTCTAAATGTAATCTG  
 301 CCCTTCAMGGTGGCTTACAAAGGTATCTTTGTCAAGGTAGGAGACCCTTGTGGCCCTCCACCTGTGCACCTTCCAGGGCCCTGCTTGGCCCTCTTCTACGGCTCTGTC  
 401 CTGAGTCTTCTATGAAATCTCCCTTACGGGCAGATTTCATATTTAGACTCTTCCACAGTTTGGACCTTNGTTTTTGGCCAGAAATAGGTTGACATTTAGATTTGTTG  
 501 GCTTGTATGATGACTTTAAATAATTTAGACATAATGTTGCTTGTAGGCCCTTCCATCTTACTCTTGGCTTTTTTTTTTGGCCCTCCAGTGTTTTGGGTAGTTTTGCT  
 601 CCCCCTACAGCCCAAGCCAAACAGATAGTTGGAGGCTCTGGAGTGGCTACATAAATTTTACACGACTGCAAATCTCTGCTGGCTGCACTTCCACAAATGTATACAA  
 701 AACTAAATACAGTCCCTGTGTTTTTATCA CAGGGAGGCTGATCAATAAATAAATTAATAAAGGGGGCTGGTCCCAZATTTGTTCTGAGGGACCTTGGGGAGCTCAGGC  
 801 GTTCTTTTTTTTTTGTGGCCCTCTTCTCTCAATTTATGAGAGAGAGCAGTAAAGATGTTCTCTGCGGTTCTCTGAGGGACCTTGGGGAGCTGTCGGCTTC  
 901 TGGGAATCTCCAAGGCAGTAGGTCCCTATCAAAATCAAGTCCAGGTTTTGTGGGGGGAACAACAANGCAGCCCAATTACCCAGAGGACTGTCGGCTTC  
 1001 CCTCACCCGCTTAGGCCCTTTGAAAGGAAACAAGAACAAGCAGACAAAATGATTTGGCTCTGAGGGGATTCAGCCCTAGAGCTCTCTCCGCCCAATCC  
 1101 CTCCCTCCGCTGAGGAACTAACAAAGGAAATAAATAATTTGCGGAAGCAGGAATTTAGAGGAAGCAAAATTCACACTGCTGTCCTTGGCTGCCGGAAAGCTG  
 1201 GACTAGAGAGTCTGCGGGCAGCCCCGAGCCCAAGCCCTTCCGGCCGCTTATAGGCCCGGGCCCGGGGAAAGGGGAACTCCAGACCCGCGGACCCCTAA  
 1301 GACACCTGCTGTAGCCCTGACCCCAAGCCCAAGCCCTCCGCCCAACTCCCTAGATGTTGTTGGCCGCTGAACTGTCGCCCTTAAAGGGGAGGGCC  
 1401 GGCTCCACGTCCTTTCTGGAGTGGTGAACATATAAACAAGGGCCCGGAAACGGGGCCGGGAGGAGAGCACAGGCTTTGAACTTAACTTAACTT  
 1501 TGGCTTGGTCCAGCCGGAATCTATAAAGGAACACTAGTCCCGGCCAAAAACCCCTTAATTTCCGAGCCGAGAGTGTAGTGGGGCCGGGAACCCGACAGAGCTCAGGCC  
 1601 GACCCCTCTCCCGGGCTGCGGAGGGCCAGGGCCGGGGAGCTC

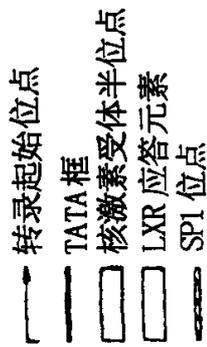


图 13

专利名称(译)	使用ATP结合框运输蛋白ABC1增加胆固醇排放并提高HDL的组合物及方法		
公开(公告)号	<a href="#">CN1379817A</a>	公开(公告)日	2002-11-13
申请号	CN00809163.3	申请日	2000-06-16
[标]申请(专利权)人(译)	CV治疗公司 华盛顿大学		
申请(专利权)人(译)	CV治疗公司 华盛顿大学		
当前申请(专利权)人(译)	CV治疗公司 华盛顿大学		
[标]发明人	理查德M罗恩 大卫维德 约翰F奥伦 迈克尔加维		
发明人	理查德·M·罗恩 大卫·维德 约翰·F·奥伦 迈克尔·加维		
IPC分类号	C12N15/09 A61K31/07 A61K31/11 A61K31/167 A61K31/202 A61K31/203 A61K31/216 A61K31/557 A61K31/5575 A61K31/5585 A61K31/56 A61K31/575 A61K31/7088 A61K35/14 A61K35/76 A61K38/00 A61K38/43 A61K39/00 A61K39/395 A61K45/00 A61K48/00 A61P1/16 A61P3/06 A61P3/10 A61P5/14 A61P9/00 A61P9/10 A61P9/12 A61P25/32 A61P43/00 C07K14/47 C07K14/705 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12P21/02 C12P21/08 C12Q1/02 C12Q1/68 C12R1/91 G01N33/15 G01N33/50 G01N33/53 G01N33/92 C12N15/52		
CPC分类号	A61K38/00 A61P1/16 A61P25/32 C07K14/705 C12N2799/022 C12N2799/027		
代理人(译)	徐金国		
优先权	60/140264 1999-06-18 US 60/153872 1999-09-14 US 60/166573 1999-11-19 US		
外部链接	<a href="#">Espacenet</a> <a href="#">SIPO</a>		

摘要(译)

本发明涉及到新型ABC1多肽和编码它的核酸分子。本发明也涉及到包含ABC1多核苷酸的重组载体,宿主细胞,和组合物,以及生产ABC1多肽的方法。本发明也涉及到特异性结合ABC1多肽的抗体。此外,本发明涉及到增加胆固醇排放的方法以及增加ABC1表达和活性的方法。本发明进一步涉及到鉴别调节ABC1表达的化合物的方法,和在哺乳动物受体中检测ABC1多肽和多核苷酸的相对水平的方法。本发明也提供适于筛选化合物以测定该化合物ABC1表达调控活性的试剂盒和组合物,以及适于测定化合物是否调节ABC1依赖性胆固醇排放的试剂盒和组合物。

