



(12)发明专利申请

(10)申请公布号 CN 111183234 A

(43)申请公布日 2020.05.19

(21)申请号 201880065116.X

C·舒尔曼

(22)申请日 2018.10.11

(74)专利代理机构 深圳永慧知识产权代理事务
所(普通合伙) 44378

(30)优先权数据

62/570,985 2017.10.11 US

代理人 黄鑫

(85)PCT国际申请进入国家阶段日

2020.04.05

(51)Int.Cl.

C12Q 1/6883(2006.01)

(86)PCT国际申请的申请数据

PCT/US2018/055377 2018.10.11

(87)PCT国际申请的公布数据

W02019/075181 EN 2019.04.18

(71)申请人 雷杰纳荣制药公司

地址 美国纽约

(72)发明人 辛玉蓉 J·格罗马达 程希平

F·杜威

T·特斯洛维奇·多斯塔尔

权利要求书6页 说明书75页

序列表90页 附图48页

(54)发明名称

在表达PNPLA3 I148M变异的患者的肝病治
疗中对HSD17B13的抑制

(57)摘要

本公开提供了将人受试者鉴定为用于通过抑制HSD17B13治疗或抑制肝病的候选者的方法。本公开还提供了通过施用HSD17B13的抑制剂来治疗呈PNPLA3 Ile148Met+的受试者的方法。本公开还提供了检测受试者中的PNPLA3 Ile148Met变体和功能性HSD17B13的方法。本公开还提供了鉴定具有针对肝病的保护效应的受试者的方法。本公开还提供了用于治疗肝病的HSD17B13的抑制剂。

1. 一种用于将人受试者鉴定为用于治疗或抑制肝病的候选者的方法,所述方法包括:
确定来自所述受试者的样品是否包含:
 - i) 编码包含I148M变异的含patatin样磷脂酶结构域3 (PNPLA3) 蛋白的第一核酸和编码功能性HSD17B13蛋白的第二核酸;和/或
 - ii) 包含I148M变异的PNPLA3蛋白和功能性HSD17B13蛋白;以及当检测到i) 中定义的所述第一和第二核酸二者和/或ii) 中定义的一种所述蛋白质时,将所述受试者鉴定为用于通过抑制HSD17B13治疗或抑制肝病的候选者。
2. 根据权利要求1所述的方法,其中所述第一核酸分子包含基因组DNA、mRNA或从mRNA获得的cDNA。
3. 根据权利要求2所述的方法,其中:
所述基因组DNA在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子;
所述mRNA在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子;
所述mRNA在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子;
所述cDNA在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子;或者
所述cDNA在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。
4. 根据权利要求2或权利要求3所述的方法,其中:
所述基因组DNA包含根据SEQ ID NO:31的核苷酸序列,或与SEQ ID NO:31具有至少90%序列同一性并编码包含所述I148M变异的PNPLA3蛋白的核苷酸序列;
所述mRNA包含根据SEQ ID NO:34的核苷酸序列,或与SEQ ID NO:34具有至少90%序列同一性并编码包含所述I148M变异的PNPLA3蛋白的核苷酸序列;
所述mRNA包含根据SEQ ID NO:35的核苷酸序列,或与SEQ ID NO:35具有至少90%序列同一性并编码包含所述I148M变异的PNPLA3蛋白的核苷酸序列;
所述cDNA包含根据SEQ ID NO:38的核苷酸序列,或与SEQ ID NO:38具有至少90%序列同一性并编码包含所述I148M变异的PNPLA3蛋白的核苷酸序列;或者
所述cDNA包含根据SEQ ID NO:39的核苷酸序列,或与SEQ ID NO:39具有至少90%序列同一性并编码包含所述I148M变异的PNPLA3蛋白的核苷酸序列。
5. 根据权利要求1至4中任一项所述的方法,其中检测所述第一核酸包括:
对所述第一核酸的至少一部分进行测序,其中所述部分包含编码所述I148M变异的密码子;或者
将所述第一核酸与探针或引物杂交,所述探针或引物与所述第一核酸的一部分特异性杂交,其中所述部分包含编码所述I148M变异的密码子。
6. 根据权利要求5所述的方法,其中所述探针或引物是等位基因特异性探针或引物,并且其中所述探针或引物任选地包含标记。
7. 根据权利要求1至6中任一项所述的方法,其还包括确定所述受试者对于所述I148M变异是纯合的还是杂合的。
8. 根据权利要求1至7中任一项所述的方法,其中所述第二核酸包含基因组DNA、mRNA或从mRNA获得的cDNA。
9. 根据权利要求8所述的方法,其中:

所述基因组DNA在对应于根据SEQ ID NO:1的位置12,667的位置处包含腺嘌呤;

所述基因组DNA包含根据SEQ ID NO:1的核苷酸序列,或与SEQ ID NO:1具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列

所述mRNA包含根据SEQ ID NO:3的核苷酸序列,或与SEQ ID NO:3具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;

所述mRNA包含根据SEQ ID NO:4的核苷酸序列,或与SEQ ID NO:4具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;

所述mRNA包含根据SEQ ID NO:7的核苷酸序列,或与SEQ ID NO:7具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;

所述mRNA包含根据SEQ ID NO:11的核苷酸序列,或与SEQ ID NO:11具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;

所述cDNA包含根据SEQ ID NO:12的核苷酸序列,或与SEQ ID NO:12具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;

所述cDNA包含根据SEQ ID NO:13的核苷酸序列,或与SEQ ID NO:13具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;

所述cDNA包含根据SEQ ID NO:16的核苷酸序列,或与SEQ ID NO:16具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;或者

所述cDNA包含根据SEQ ID NO:20的核苷酸序列,或与SEQ ID NO:20具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列。

10. 根据权利要求1至9中任一项所述的方法,其中检测所述第二核酸包括:

对所述第二核酸进行测序;或者

将所述第二核酸与探针或引物杂交,所述探针或引物与所述第二核酸的一部分特异性杂交,其中所述部分在对应于根据SEQ ID NO:1的位置12,667的位置处包含腺嘌呤。

11. 根据权利要求10所述的方法,其中所述探针或引物是等位基因特异性探针或引物,并且其中所述探针或引物任选地包含标记。

12. 根据权利要求1至11中任一项所述的方法,其还包括确定所述受试者对于所述样品中编码功能性HSD17B13蛋白的所述第二核酸是纯合的还是杂合的。

13. 根据权利要求1至12中任一项所述的方法,其还包括向所述受试者施用HSD17B13的抑制剂。

14. 根据权利要求1至13中任一项所述的方法,其中所述肝病是酒精性肝病。

15. 根据权利要求14所述的方法,其中所述酒精性肝病包括由饮酒引起的肝硬化、脂肪变性或肝细胞癌中的一种或多种。

16. 根据权利要求1至13中任一项所述的方法,其中所述肝病是非酒精性肝病。

17. 根据权利要求16所述的方法,其中所述非酒精性肝病包括非酒精性脂肪性肝病 (NAFLD) 或非酒精性脂肪性肝炎 (NASH)。

18. 根据权利要求16所述的方法,其中所述非酒精性肝病包括并非由饮酒导致的肝硬化、脂肪变性或肝细胞癌中的一种或多种。

19. 一种治疗或抑制肝病的方法,所述方法包括向表达包含I148M变异的含patatin样磷脂酶结构域3 (PNPLA3) 蛋白的人肝病患者施用羟基类固醇17-β脱氢酶13 (HSD17B13) 的抑

制剂,使得在所述患者中治疗或抑制所述肝病。

20. 根据权利要求19所述的方法,其中所述肝病是酒精性肝病。

21. 根据权利要求20所述的方法,其中所述酒精性肝病包括由饮酒引起的肝硬化、脂肪变性或肝细胞癌中的一种或多种。

22. 根据权利要求19所述的方法,其中所述肝病是非酒精性肝病。

23. 根据权利要求22所述的方法,其中所述非酒精性肝病包括非酒精性脂肪性肝病 (NAFLD) 或非酒精性脂肪性肝炎 (NASH)。

24. 根据权利要求22所述的方法,其中所述非酒精性肝病包括并非由饮酒导致的肝硬化、脂肪变性或肝细胞癌中的一种或多种。

25. 根据权利要求19至24中任一项所述的方法,其中所述患者是肥胖的。

26. 根据权利要求19至25中任一项所述的方法,其中所述患者患有脂肪肝。

27. 根据权利要求19至26中任一项所述的方法,其中已经通过在来自所述受试者的样品中检测变体PNPLA3蛋白,确定所述患者表达所述变体PNPLA3蛋白。

28. 根据权利要求19至27中任一项所述的方法,其中所述变体PNPLA3蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸。

29. 根据权利要求26或权利要求28所述的方法,其中所述样品中的所述变体PNPLA3蛋白包含根据SEQ ID NO:42的氨基酸序列,或与SEQ ID NO:42具有至少90%序列同一性并包含所述I148M变异的氨基酸序列。

30. 根据权利要求27至29中任一项所述的方法,其中所述变体PNPLA3蛋白通过氨基酸测序或通过免疫测定来检测。

31. 根据权利要求19至26中任一项所述的方法,其中已经通过在来自所述受试者的样品中检测编码所述变体PNPLA3蛋白的核酸分子,确定所述受试者表达所述变体PNPLA3蛋白。

32. 根据权利要求31所述的方法,其中所述变体PNPLA3蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸。

33. 根据权利要求32所述的方法,其中所述变体PNPLA3蛋白包含根据SEQ ID NO:42的氨基酸序列,或与SEQ ID NO:42具有至少90%序列同一性并包含所述I148M变异的氨基酸序列。

34. 根据权利要求31至33中任一项所述的方法,其中编码所述变体PNPLA3蛋白的核酸分子包括基因组DNA、mRNA或从mRNA获得的cDNA。

35. 根据权利要求31至34中任一项所述的方法,其中所述核酸分子包含基因组DNA,所述基因组DNA在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子。

36. 根据权利要求34或权利要求35所述的方法,其中所述基因组DNA包含根据SEQ ID NO:31的核苷酸序列,或与SEQ ID NO:31具有至少90%序列同一性并编码包含所述I148M变异的PNPLA3蛋白的核苷酸序列。

37. 根据权利要求31至34中任一项所述的方法,其中所述核酸分子包含mRNA,所述mRNA在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子。

38. 根据权利要求34或37所述的方法,其中所述mRNA包含根据SEQ ID NO:34的核苷酸序列,或与SEQ ID NO:34具有至少90%序列同一性并编码包含所述I148M变异的PNPLA3蛋

白的核苷酸序列。

39. 根据权利要求31至34中任一项所述的方法,其中所述核酸分子包含mRNA,所述mRNA在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子。

40. 根据权利要求34或权利要求39所述的方法,其中所述mRNA包含根据SEQ ID NO:35的核苷酸序列,或与SEQ ID NO:35具有至少90%序列同一性并编码包含所述I148M变异的PNPLA3蛋白的核苷酸序列。

41. 根据权利要求31至34中任一项所述的方法,其中所述核酸分子包含从mRNA获得的cDNA,所述cDNA在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子。

42. 根据权利要求34或权利要求41所述的方法,其中所述cDNA包含根据SEQ ID NO:38的核苷酸序列,或与SEQ ID NO:38具有至少90%序列同一性并编码包含所述I148M变异的PNPLA3蛋白的核苷酸序列。

43. 根据权利要求31至34中任一项所述的方法,其中所述核酸分子包含从mRNA获得的cDNA,所述cDNA在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。

44. 根据权利要求34或权利要求43所述的方法,其中所述cDNA包含根据SEQ ID NO:39的核苷酸序列,或与SEQ ID NO:39具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列。

45. 根据权利要求31至34中任一项所述的方法,其中通过对所述核酸的至少一部分进行测序来检测所述核酸,其中所述部分包含编码所述I148M变异的密码子。

46. 根据权利要求31至44中任一项所述的方法,其中通过所述核酸与探针或引物的杂交来检测所述核酸,所述探针或引物与所述核酸的一部分特异性杂交,其中所述部分包含编码所述I148M变异的密码子。

47. 根据权利要求46所述的方法,其中所述探针或引物是等位基因特异性探针或引物。

48. 根据权利要求46或权利要求47所述的方法,其中所述探针或引物包含标记。

49. 根据权利要求19至48中任一项所述的方法,其中所述患者对于编码所述变体PNPLA3蛋白的基因是纯合的。

50. 根据权利要求19至48中任一项所述的方法,其中所述患者对于编码所述变体PNPLA3蛋白的基因是杂合的。

51. 根据权利要求49或权利要求50所述的方法,其中所述患者还对于编码功能性HSD17B13蛋白的基因是纯合的。

52. 根据权利要求49或权利要求50所述的方法,其中所述患者还对于编码功能性HSD17B13蛋白的基因是杂合的。

53. 根据权利要求51或权利要求52所述的方法,其中所述功能性HSD17B13蛋白是HSD17B13蛋白,其具有野生型HSD17B13蛋白的至少10%活性。

54. 根据权利要求52所述的方法,其中所述患者对于编码所述功能性HSD17B13蛋白的基因和编码HSD17B13功能丧失变体的基因是杂合的。

55. 根据权利要求54所述的方法,其中所述功能性HSD17B13蛋白是HSD17B13蛋白,其具有野生型HSD17B13蛋白的至少10%活性。

56. 根据权利要求19至55中任一项所述的方法,其中所述HSD17B13抑制剂还能够抑制HSD17B1、HSD17B2、HSD17B3、HSD17B4、HSD17B6、HSD17B7、HSD17B8、HSD17B10、HSD17B11、

HSD17B12、HSD17B13、HSD17B14、DHRS3或RDH10中的一种或多种。

57.一种用于在人受试者中治疗肝病的HSD17B13的抑制剂,所述人受试者具有包含I148M变异的PNPLA3蛋白并具有功能性HSD17B13蛋白。

58.根据权利要求57所述的使用的抑制剂,其中所述人受试者已被测试为对于包含I148M变异的PNPLA3蛋白和对于功能性HSD17B13蛋白是阳性的。

59.根据权利要求57或58所述的使用的抑制剂,其中所述治疗包括确定所述人受试者是否具有包含I148M变异的PNPLA3蛋白和功能性HSD17B13蛋白。

60.根据权利要求56至58中任一项所述的使用的抑制剂,其中已经通过使用如权利要求1至18中任一项所定义的方法将所述人受试者鉴定为用于通过抑制HSD17B13治疗或抑制肝病的候选者。

61.根据权利要求57至60中任一项所述的抑制剂,其中所述变体PNPLA3蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸。

62.根据权利要求61所述的抑制剂,其中所述变体PNPLA3蛋白包含根据SEQ ID NO:42的氨基酸序列,或与SEQ ID NO:42具有至少90%序列同一性并包含所述I148M变异的氨基酸序列。

63.根据权利要求57至60中任一项所述的抑制剂,其中编码所述变体PNPLA3蛋白的核酸分子是基因组DNA。

64.根据权利要求63所述的抑制剂,其中所述基因组DNA在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子。

65.根据权利要求64所述的抑制剂,其中所述基因组DNA包含根据SEQ ID NO:31的核苷酸序列,或与SEQ ID NO:31具有至少90%序列同一性并编码包含所述I148M变异的PNPLA3蛋白的核苷酸序列。

66.根据权利要求56至60中任一项所述的抑制剂,其中编码所述变体PNPLA3蛋白的核酸分子是mRNA。

67.根据权利要求66所述的抑制剂,其中所述mRNA在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子。

68.根据权利要求67所述的抑制剂,其中所述mRNA包含根据SEQ ID NO:34的核苷酸序列,或与SEQ ID NO:34具有至少90%序列同一性并编码包含所述I148M变异的PNPLA3蛋白的核苷酸序列。

69.根据权利要求66所述的抑制剂,其中所述mRNA在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子。

70.根据权利要求69所述的抑制剂,其中所述mRNA包含根据SEQ ID NO:35的核苷酸序列,或与SEQ ID NO:35具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列。

71.根据权利要求57至60中任一项所述的抑制剂,其中编码所述变体PNPLA3蛋白的核酸分子是cDNA。

72.根据权利要求71所述的抑制剂,其中所述cDNA在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子。

73.根据权利要求72所述的抑制剂,其中所述cDNA包含根据SEQ ID NO:38的核苷酸序

列,或与SEQ ID NO:38具有至少90%序列同一性并编码包含所述I148M变异的PNPLA3蛋白的核苷酸序列。

74.根据权利要求71所述的抑制剂,其中所述cDNA在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。

75.根据权利要求74所述的抑制剂,其中所述cDNA包含根据SEQ ID NO:39的核苷酸序列,或与SEQ ID NO:39具有至少90%序列同一性并编码包含所述I148M变异的PNPLA3蛋白的核苷酸序列。

76.根据权利要求57至75中任一项所述的抑制剂,其中所述肝病是酒精性肝病。

77.根据权利要求76所述的抑制剂,其中所述酒精性肝病包括由饮酒引起的肝硬化、脂肪变性或肝细胞癌中的一种或多种。

78.根据权利要求57至75中任一项所述的抑制剂,其中所述肝病是非酒精性肝病。

79.根据权利要求78所述的抑制剂,其中所述非酒精性肝病包括非酒精性脂肪性肝病(NAFLD)或非酒精性脂肪性肝炎(NASH)。

80.根据权利要求78所述的抑制剂,其中所述非酒精性肝病包括并非由饮酒导致的肝硬化、脂肪变性或肝细胞癌中的一种或多种。

81.根据权利要求57至80中任一项所述的抑制剂,其中所述人受试者对于功能性HSD17B13是纯合的或杂合的。

在表达PNPLA3 I148M变异的患者的肝病治疗中对HSD17B13的抑制

[0001] 对序列表的引用

[0002] 本申请包括作为文本文件电子提交的序列表,所述文本文件的名称为18923801002SEQ,于2018年10月10日创建,其大小为238千字节。所述序列表通过引用并入本文。

技术领域

[0003] 本公开整体涉及精准医疗的领域。更具体地,本公开涉及鉴定受试者并用羟基类固醇17- β 脱氢酶13 (HSD17B13) 的抑制剂治疗此类受试者的方法,所述受试者为含patatin样磷脂酶结构域3 (PNPLA3) Ile148Met阳性,并且患有肝病或易患肝病。

背景技术

[0004] 在整个说明书中引用了各种参考文献,包括专利、专利申请、登记号、技术文章和学术文章。每个参考文献通过引用以其整体并出于所有目的并入本文。

[0005] 慢性肝病和肝硬化是美国发病率和死亡率的主要原因,2014年计有38,170人死亡(占死亡总数的1.5%) (Kochanek等人,Nat'l.Vital Stat.Rep.,2016,65,1-122)。在美国,肝硬化最常见的病因是酒精性肝病、慢性丙型肝炎和非酒精性脂肪性肝病 (NAFLD),在2004年至2013年间,共占等待肝移植患者的约80% (Wong等人,Gastroenterology,2015,148,547-555)。美国NAFLD的估计患病率在19%至46%之间 (Browning等人,Hepatology,2004,40,1387-1395;Lazo等人,Am.J.Epidemiol.,2013,178,38-45;和Williams等人,Gastroenterology,2011,140,124-131),并随着时间的推移而上升 (Younossi等人,Clin.Gastroenterol.Hepatol.,2011,9,524-530),可能与其主要风险因素肥胖率增加有关 (Cohen等人,Science,2011,332,1519-1523)。尽管丙型肝炎的治疗已取得了显著进展,但目前还没有针对酒精性或非酒精性肝病和肝硬化的循证治疗。

[0006] 先前的全基因组关联研究 (GWAS) 已鉴定了与慢性肝病风险增加相关的序列变异。最可靠的关联是与由基因PNPLA3编码的含patatin样磷脂酶结构域3中常见的错义变体的关联。最初发现这种变体 (rs738409,p.Ile148Met) 与肝甘油三酯水平升高有关 (Romeo等人,Nat.Genet.,2008,40,1461-5),并且随后与非酒精性脂肪性肝炎 (NASH) (Rotman等人,Hepatology,2010,52,894-903;Sookoian等人,J.Lipid Res.,2009,50,2111-2116) 和肝硬化 (Shen等人,J.Lipid Res.,2015,56,167-175) 相关。编码跨膜6超家族成员2的TM6SF2中的错义变体也增加了非酒精性脂肪性肝病 (NAFLD) 的风险 (Kozlitina等人,Nat.Genet.,2014,46,352-6;Liu等人,Nat.Comm.,2014,5,4309;和Sookoian等人,Hepatology,2015,61,515-25)。PNPLA3和TM6SF2中的变体究竟如何导致肝病还有待于完全阐明 (Smagris等人,J.Biol.Chem.,2016,291,10659-76;Mahdessian等人,Proc.Natl.Acad.Sci.USA,2014,111,8913-8;Huang等人,J.Biol.Chem.,2011,286,37085-93;和Pirazzi等人,J.Hepatol.,2012,57,1276-82)。迄今为止,还没有鉴定出防止慢性肝病的基因变体。

发明内容

[0007] 本公开提供了用于将人受试者鉴定为用于治疗或抑制肝病的候选者的方法,所述方法包括:确定来自所述受试者的样品是否包含:i) 编码包含I148M变异的含patatin样磷脂酶结构域3 (PNPLA3) 蛋白的第一核酸和编码功能性HSD17B13蛋白的第二核酸;和/或ii) 包含I148M变异的PNPLA3蛋白和功能性HSD17B13蛋白;以及当检测到i) 中定义的第一和第二核酸二者和/或ii) 中定义的一种蛋白质时,将受试者鉴定为用于通过抑制HSD17B13治疗或抑制肝病的候选者。

[0008] 在一些实施方案中,第一核酸分子包含基因组DNA、mRNA或从mRNA获得的cDNA。

[0009] 在一些实施方案中,基因组DNA在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子;mRNA在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子;mRNA在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子;cDNA在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子;或者cDNA在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。

[0010] 在一些实施方案中,基因组DNA包含根据SEQ ID NO:31的核苷酸序列,或与SEQ ID NO:31具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列;mRNA包含根据SEQ ID NO:34的核苷酸序列,或与SEQ ID NO:34具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列;mRNA包含根据SEQ ID NO:35的核苷酸序列,或与SEQ ID NO:35具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列;cDNA包含根据SEQ ID NO:38的核苷酸序列,或与SEQ ID NO:38具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列;或者cDNA包含根据SEQ ID NO:39的核苷酸序列,或与SEQ ID NO:39具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列。

[0011] 在一些实施方案中,检测第一核酸包括:对第一核酸的至少一部分进行测序,其中所述部分包含编码I148M变异的密码子;或者将第一核酸与探针或引物杂交,所述探针或引物与第一核酸的一部分特异性杂交,其中所述部分包含编码I148M变异的密码子。

[0012] 在一些实施方案中,探针或引物是等位基因特异性探针或引物,并且其中所述探针或引物任选地包含标记。

[0013] 在一些实施方案中,所述方法还包括确定受试者对于I148M变异是纯合的还是杂合的。

[0014] 在一些实施方案中,第二核酸包含基因组DNA、mRNA或从mRNA获得的cDNA。

[0015] 在一些实施方案中,基因组DNA在对应于根据SEQ ID NO:1的位置12,667的位置处包含腺嘌呤;基因组DNA包含根据SEQ ID NO:1的核苷酸序列,或与SEQ ID NO:1具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;mRNA包含根据SEQ ID NO:3的核苷酸序列,或与SEQ ID NO:3具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;mRNA包含根据SEQ ID NO:4的核苷酸序列或与SEQ ID NO:4具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;mRNA包含根据SEQ ID NO:7的核苷酸序列或与SEQ ID NO:7具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;mRNA包含根据SEQ ID NO:11的核苷酸序列或与SEQ ID NO:11具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;cDNA包含根据SEQ ID NO:12的核苷酸序列或与

SEQ ID NO:12具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;cDNA包含根据SEQ ID NO:13的核苷酸序列或与SEQ ID NO:13具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;cDNA包含根据SEQ ID NO:16的核苷酸序列或与SEQ ID NO:16具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列;或者cDNA包含根据SEQ ID NO:20的核苷酸序列或与SEQ ID NO:20具有至少90%序列同一性并编码功能性HSD17B13蛋白的核苷酸序列。

[0016] 在一些实施方案中,检测第二核酸包括:对第二核酸进行测序;或者将第二核酸与探针或引物杂交,所述探针或引物与第二核酸的一部分特异性杂交,其中所述部分在对应于根据SEQ ID NO:1的位置12,667的位置处包含腺嘌呤。

[0017] 在一些实施方案中,探针或引物是等位基因特异性探针或引物,并且其中所述探针或引物任选地包含标记。

[0018] 在一些实施方案中,所述方法还包括确定受试者对于样品中编码功能性HSD17B13蛋白的第二核酸是纯合的还是杂合的。

[0019] 在一些实施方案中,所述方法还包括向受试者施用HSD17B13的抑制剂。

[0020] 在一些实施方案中,所述肝病是酒精性肝病。在一些实施方案中,酒精性肝病包括由饮酒引起的肝硬化、脂肪变性或肝细胞癌中的一种或多种。

[0021] 在一些实施方案中,所述肝病是非酒精性肝病。在一些实施方案中,非酒精性肝病包括非酒精性脂肪性肝病 (NAFLD) 或非酒精性脂肪性肝炎 (NASH)。在一些实施方案中,非酒精性肝病包括并非由饮酒导致的肝硬化、脂肪变性或肝细胞癌中的一种或多种。

附图说明

[0022] 并入本说明书且构成其一部分的附图说明了若干个方面,并连同描述一起用于解释本公开的原理。

[0023] 图1示出了来自发现和复制群组的测序欧洲血统个体的基线特征。

[0024] 图2示出了发现群组中以 $P < 1.0 \times 10^{-7}$ 与血清转氨酶水平相关的单核苷酸变体。

[0025] 图3示出了在三个单独的欧洲血统群组中,来自发现群组的35个全外显子组的重要单核苷酸变体的复制和联合综合分析。

[0026] 图4示出了发现群组中13个全外显子组重要复制性单核苷酸变体与肝病表型的关联。

[0027] 图5示出了来自达拉斯肝脏和儿科肝脏研究 (Dallas Liver and Pediatric Liver Studies) 的多种族基因分型病例和对照的基线特征。

[0028] 图6(图A和图B) 示出了HSD17B13周围区域的GHS发现群组中丙氨酸转氨酶 (ALT;A) 和天冬氨酸转氨酶 (AST;B) 水平的区域关联图。

[0029] 图7示出了HSD17B13 rs72613567剪接变体的纯合参考 (T/T)、杂合 (T/TA) 和纯合交替 (TA/TA) 载体中PNPLA3的表达。

[0030] 图8示出了63个PNPLA3 rs738409载体 (C/C和C/G) 在三种HSD17B13 rs72613567基因型 (T/T、T/TA、TA/TA) 中的表达差异。

[0031] 图9示出了对PNPLA3 rs738409 (p.I148M) 与HSD17B13rs72613567之间遗传相互作用的分析。

[0032] 图10(图A和图B)示出了HSD17B13 rs72613567:TA减小与PNPLA3 p.I148M相关的肝损伤风险。

[0033] 图11(图A至图F)示出了PNPLA3 rs738409(p.I148M)和HSD17B13 rs72613567基因型的原始和残差化ALT水平。

[0034] 图12(图A至图F)示出了PNPLA3 rs738409(p.I148M)和HSD17B13 rs72613567基因型的原始和残差化AST水平。

[0035] 图13(图A至图F)示出了HSD17B13剪接变体的纯合参考(T/T)、杂合(T/TA)和纯合交替(TA/TA)载体中四个另外新型HSD17B13转录物(E-H)的mRNA表达。

[0036] 图14(图A和图B)示出了GHS发现群组中与血清转氨酶水平的单核苷酸变体关联的曼哈顿图(左)和分位数-分位数图(右)。

[0037] 图15(图A和图B)示出了HSD17B13 rs72613567:TA与酒精性和非酒精性肝病表型的风险降低相关。

[0038] 图16(图A和图B)示出了HSD17B13 rs72613567:TA减小与PNPLA3 p.I148M相关的肝损伤风险。

[0039] 图17(图A和图B)示出了HSD17B13 rs72613567:TA与从单纯性脂肪变性进展到脂肪性肝炎和纤维化的风险降低相关。

[0040] 图18(图A至图G)示出了新型HSD17B13转录物的表达、亚细胞定位和酶活性。

[0041] 图19(图A和图B)示出了HSD17B13 rs72613567:TA减小与PNPLA3 I148M相关的酒精性和非酒精性肝病风险。每个条形上方的数字表示对照/病例。

[0042] 本公开的另外的优点将部分陈述在以下描述中,并且部分将从所述描述中显而易见或者可以通过实践本文公开的实施方案来得知。本公开的优点将借助所附权利要求书中具体指出的要素和组合来实现并达成。应了解上文一般性描述与下文详述两者均仅具有示例性和说明性,并且不限制如所要求的实施方案。

具体实施方式

[0043] 在整个说明书和权利要求书中使用了与公开内容的方面相关的各种术语。除非另有说明,否则此类术语将被赋予其在本领域中的普通含义。其他具体定义的术语将以与本文提供的定义一致的方式来解释。

[0044] 除非另外明确说明,否则决不意图将本文陈述的任何方法或方面解释为要求以特定顺序执行其步骤。因此,在权利要求书或说明书中,当方法权利要求没有确切地说明步骤是限于特定顺序时,在任何方面决非意图推断顺序。这适用于任何可能的用于解释的非表达基础,包括关于步骤安排或操作流程的逻辑事项、从语法组织或标点符号得到的清晰含义、或者在说明书中描述的方面的数量或类型。

[0045] 如本文所用,单数形式“一个/一种(a)”、“一个/一种(an)”和“所述(the)”包括复数个/复数种指示物,除非上下文另外明确说明。

[0046] 如本文所用,术语“受试者”和“患者”可互换地使用。受试者可以包括任何动物,包括哺乳动物。哺乳动物包括但不限于农场动物(例如马、牛、猪)、伴侣动物(例如狗、猫)、实验室动物(例如小鼠、大鼠、兔)和非人灵长类动物。在一些实施方案中,受试者是人类。

[0047] 如本文所用,“核酸”、“核酸分子”、“核酸序列”、“多核苷酸”或“寡核苷酸”可以包

含任何长度的核苷酸的聚合形式,可以包含DNA和/或RNA,并且可以是单链、双链或多链。核酸的一条链也指它的补体。

[0048] 如本文所用,当在给定氨基酸或核酸序列或位置的编号的情况下使用时,短语“对应于”或其语法变体是指当给定氨基酸或核酸序列与参考序列(例如,与本文中的参考序列,它是例如(功能性的或表现为功能性的转录物)HSD17B13的核酸分子或多肽)进行比较时,指定参考序列的编号。换句话说,给定聚合物的残基(例如,氨基酸或核苷酸)编号或残基(例如,氨基酸或核苷酸)位置是相对于参考序列指定的,而不是依据残基在给定氨基酸或核酸序列内的实际数字位置指定的。例如,可以通过引入空位来优化两个序列之间的残基匹配,将给定氨基酸序列与参考序列进行比对。在这些情况下,尽管存在空位,但给定氨基酸或核酸序列中的残基的编号是相对于其所比对的参考序列进行的。

[0049] 例如,短语“编码HSD17B13功能丧失变体蛋白的核酸分子,所述变体蛋白在对应于根据SEQ ID NO:2的位置12,667的位置处包含胸腺嘧啶”(和类似短语)意味着,如果将所检查的HSD17B13基因组DNA的核酸序列与根据SEQ ID NO:2的核苷酸序列比对,则所检查的HSD17B13基因组DNA在对应于SEQ ID NO:2的位置12,667的位置处包含胸腺嘧啶。

[0050] 例如,通过在给定HSD17B13蛋白与SEQ ID NO:2的核酸序列之间进行序列比对,可以容易地鉴定编码HSD17B13功能丧失变体蛋白的核酸分子,所述变体蛋白在对应于根据SEQ ID NO:2的位置12,667的位置处包含胸腺嘧啶。同样,通过在给定PNPLA3蛋白与SEQ ID NO:42或SEQ ID NO:43的氨基酸序列之间进行序列比对,可以容易地鉴定在对应于根据SEQ ID NO:42的位置148的位置处或对应于根据SEQ ID NO:43的位置144的位置处具有甲硫氨酸的PNPLA3 Ile148Met蛋白。存在多种可以用于进行序列比对的计算算法,以便鉴定特定的核酸分子和蛋白质,所述核酸分子和蛋白质在对应于特定SEQ ID NO的位置的特定位置处具有特定的核苷酸或氨基酸。例如,用于鉴定序列同一性百分比的程序可以用于进行序列比对。核酸内的核酸序列或多肽内的氨基酸序列的具体延伸段之间的同一性百分比(或互补百分比)可以使用以下来确定:BLAST程序(基本局部比对搜索工具)和PowerBLAST程序(Altschul等人,J.Mol.Biol.,1990,215,403-410;Zhang和Madden,Genome Res.,1997,7,649-656)或CLUSTALW软件(Sievers等人,2014,Methods Mol.Biol.,1079,105-116),或通过使用Gap程序(Wisconsin序列分析包,Unix第8版,Genetics Computer Group,University Research Park,麦迪逊市,威斯康辛州),所述Gap程序使用默认设置且使用Smith和Waterman的算法(Adv.Appl.Math.,1981,2,482-489)。然而,还可以手动比对序列。在本文中,如果提到序列同一性百分比,则较高的序列同一性百分比优于较低的百分比。

[0051] 本公开提供了将人受试者鉴定为用于通过抑制HSD17B13治疗或抑制肝病的候选者的方法;治疗或抑制肝病的方法,所述方法包括施用HSD17B13的抑制剂;检测受试者中的PNPLA3 Ile148Met(在本文也称为“I148M”)和功能性HSD17B13的方法;鉴定具有针对肝病的保护效应的受试者的方法;和用于治疗肝病的HSD17B13的抑制剂。

[0052] 本公开提供了将人受试者分类为用于通过抑制HSD17B13治疗或抑制肝病的候选者的方法;治疗或抑制肝病的方法,所述方法包括施用HSD17B13的抑制剂;检测受试者中的PNPLA3 Ile148Met(在本文也称为“I148M”)和功能性HSD17B13的方法;对具有针对肝病的保护效应的受试者进行分类的方法;和用于治疗肝病的HSD17B13的抑制剂。

[0053] 根据本公开,已观察到编码17- β 羟基类固醇脱氢酶13(一种肝脂滴蛋白)的

HSD17B13中的剪接变体(rs72613567:TA)与ALT($P=4.2 \times 10^{-12}$)和AST($P=6.2 \times 10^{-10}$)水平的降低可重复地相关。还观察到这种变体以等位基因剂量依赖性方式与酒精性和非酒精性肝病(对于每个rs72613567:TA等位基因,分别地:降低38%,95%置信区间(CI)19%-52%;和降低16%,95%CI 9%-22%)和肝硬化(对于每个rs72613567:TA等位基因,对于酒精性和非酒精性肝硬化,分别地:降低44%,95%CI 22-59%;和降低26%,95%CI 12%-38%)的风险降低相关。这种关联在两个独立的群组中得到证实.rs72613567:TA与非酒精性脂肪性肝炎(NASH)组织学特征的严重程度降低相关(在患有脂肪性肝病的个体中,对于每个rs72613567:TA等位基因,在非酒精性脂肪性肝炎(NASH)中,降低23%,95%CI10%-34%),并减轻了与PNPLA3 p.I148M相关的肝损伤.rs72613567:TA导致缺乏针对类固醇底物的酶活性的截短的同种型。因此,HSD17B13的功能丧失变体与酒精性和非酒精性肝病以及从脂肪变性进展到NASH的风险降低相关。美国专利申请公开号US2018/0216084(对应于PCT公开号WO 2018/136702)通过引用以其整体并入本文。

[0054] 本公开提供了用于将人受试者鉴定为用于通过抑制羟基类固醇17- β 脱氢酶13(HSD17B13)治疗或抑制肝病的首选者的方法,所述方法包括确定来自所述受试者的样品是否包含编码包含I148M变异的含patatin样磷脂酶结构域3(PNPLA3)蛋白的第一核酸和编码功能性HSD17B13蛋白的第二核酸,和/或包含I148M变异的PNPLA3蛋白和功能性HSD17B13蛋白,以及当检测到第一和第二核酸二者和/或检测到两种蛋白质时,将受试者鉴定为用于通过抑制HSD17B13治疗或抑制肝病的首选者。

[0055] 本公开还提供了将人受试者分类为用于通过抑制HSD17B13治疗或抑制肝病的首选者的方法;治疗或抑制肝病的方法,所述方法包括施用HSD17B13的抑制剂;检测受试者中的PNPLA3 Ile148Met(在本文也称为“I148M”)和功能性HSD17B13的方法;对具有针对肝病的保护效应的受试者进行分类的方法;和用于治疗肝病的HSD17B13的抑制剂。

[0056] 本公开还提供了治疗或抑制肝病的方法,所述方法包括向表达包含I148M变异的含patatin样磷脂酶结构域3(PNPLA3)蛋白的人肝病患者施用羟基类固醇17- β 脱氢酶13(HSD17B13)的抑制剂,使得在患者中治疗或抑制肝病。

[0057] 在本文所述的方法中,检测、表达或采用各种PNPLA3和HSD17B13蛋白以及编码它们的核酸分子(例如基因组DNA、mRNA和来源于mRNA的cDNA)。更详细地描述了这些PNPLA3和HSD17B13蛋白以及编码它们的核酸分子。

[0058] 两种野生型PNPLA3蛋白的氨基酸序列在SEQ ID NO:40和SEQ ID NO:41中列出。具有SEQ ID NO:40的野生型PNPLA3蛋白的长度为481个氨基酸,而具有SEQ ID NO:41的野生型PNPLA3蛋白的长度为477个氨基酸。具有SEQ ID NO:40的野生型PNPLA3蛋白在位置148处具有异亮氨酸。具有SEQ ID NO:41的野生型PNPLA3蛋白在位置144处具有异亮氨酸。

[0059] 在一些实施方案中,变体PNPLA3 Ile148Met蛋白包含与根据SEQ ID NO:42的氨基酸序列具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列,并且在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3 Ile148Met蛋白包含与根据SEQ ID NO:42的氨基酸序列具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列,并且在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3 Ile148Met蛋白

包含根据SEQ ID NO:42的氨基酸序列或由其组成。

[0060] 在一些实施方案中,变体PNPLA3 Ile144Met蛋白包含与根据SEQ ID NO:43的氨基酸序列具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列,并且在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3 Ile144Met蛋白包含与根据SEQ ID NO:43的氨基酸序列具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列,并且在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3 Ile144Met蛋白包含根据SEQ ID NO:43的氨基酸序列或由其组成。

[0061] 在一些实施方案中,变体PNPLA3 Ile148Met和变体PNPLA3Ile144Met蛋白是上述蛋白质的片段,其中所述片段在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸,或者在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,所述片段包含编码的多肽(如具有SEQ ID NO:42或SEQ ID NO:43的氨基酸序列的多肽)的至少约10、至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约55、至少约60、至少约65、至少约70、至少约75、至少约80、至少约85、至少约90、至少约95、至少约100、至少约150或至少约200个连续氨基酸残基。在这方面,较长的片段比较短的片段更优选。在一些实施方案中,所述片段包含编码的多肽的至少约10、至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约55、至少约60、至少约65、至少约70、至少约75、至少约80、至少约85、至少约90、至少约95或至少约100个连续氨基酸残基。在这方面,较长的片段比较短的片段更优选。

[0062] 编码野生型PNPLA3蛋白的基因组DNA分子的核酸序列在SEQ ID NO:30中列出。具有SEQ ID NO:30的野生型PNPLA3基因组DNA分子在位置5109处包含胞嘧啶。具有SEQ ID NO:30的野生型PNPLA3基因组DNA分子在位置5107至5109处包含密码子ATC。

[0063] 在一些实施方案中,变体PNPLA3基因组DNA分子包含编码PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白的核酸序列或由其组成,所述PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白分别包含与SEQ ID NO:42或SEQ ID NO:43具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列,并且在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸或者在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3基因组DNA分子包含编码PNPLA3Ile148Met蛋白或PNPLA3 Ile144Met蛋白的核酸序列或由其组成,所述PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白分别包含与SEQ ID NO:42或SEQ ID NO:43具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列,并且在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸或者在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3基因组DNA分子包含编码PNPLA3Ile148Met蛋白或PNPLA3 Ile144Met蛋白的核酸序列或由其组成,所述PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白分别包含根据SEQ ID NO:42或SEQ ID NO:43的氨基酸序列或由其组成。

[0064] 在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白或变体PNPLA3 Ile144Met蛋白的变体PNPLA3基因组DNA分子包含与SEQ ID NO:31具有至少约90%、至少约91%、至少约

92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成,并且在对应于根据SEQ ID NO:31的位置5109的位置处包含鸟嘌呤,或者在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含密码子ATG。在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白或变体PNPLA3Ile144Met蛋白的变体PNPLA3基因组DNA分子包含与SEQ ID NO:31具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成,并且在对应于根据SEQ ID NO:31的位置5109的位置处包含鸟嘌呤,或者在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含密码子ATG。在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白或变体PNPLA3Ile144Met蛋白的变体PNPLA3基因组DNA分子包含根据SEQ ID NO:31的核苷酸序列或由其组成。

[0065] 在一些实施方案中,变体PNPLA3基因组DNA分子包含少于整个基因组DNA序列。在一些实施方案中,变体PNPLA3基因组DNA分子包含SEQ ID NO:31的至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400、至少约500、至少约600、至少约700、至少约800、至少约900、至少约1000、至少约2000、至少约3000、至少约4000、至少约5000、至少约6000、至少约7000、至少约8000、至少约9000、至少约10000、至少约11000或至少约11500个连续核苷酸或由其组成。在一些实施方案中,变体PNPLA3基因组DNA分子包含SEQ ID NO:31的至少约1000至至少约2000个连续核苷酸或由其组成。

[0066] 在一些实施方案中,变体PNPLA3基因组DNA分子包含SEQ ID NO:31的至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400、至少约500、至少约600、至少约700、至少约800、至少约900、至少约1000、至少约1000、至少约1100、至少约1200、至少约1300、至少约1400、至少约1500、至少约1600、至少约1700、至少约1800、至少约1900、至少约2000、至少约2100、至少约2200、至少约2300、至少约2400或至少约2500个连续核苷酸或由其组成。

[0067] 两种野生型PNPLA3 mRNA分子的核酸序列在SEQ ID NO:32和SEQ ID NO:33中列出。具有SEQ ID NO:32的野生型PNPLA3 mRNA分子在位置444处包含胞嘧啶。具有SEQ ID NO:32的野生型PNPLA3mRNA分子在位置442至444处包含密码子AUC。具有SEQ ID NO:33的野生型PNPLA3 mRNA分子在位置432处包含胞嘧啶。具有SEQ ID NO:33的野生型PNPLA3 mRNA分子在位置430至432处包含密码子AUC。

[0068] 在一些实施方案中,变体PNPLA3 mRNA分子包含编码PNPLA3Ile148Met蛋白或PNPLA3 Ile144Met蛋白的核酸序列或由其组成,所述PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白分别包含与SEQ ID NO:42或SEQ ID NO:43具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列,并且在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸或者在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3 mRNA分子包含编码PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白的核酸序列或由其组成,所述PNPLA3Ile148Met蛋白或PNPLA3 Ile144Met蛋白分别包含与SEQ ID NO:42或SEQ ID NO:43具有至少约95%、至少约96%、至少约97%、至

少约98%或至少约99%序列同一性的氨基酸序列,并且在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸或者在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3 mRNA分子包含编码PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白的核酸序列或由其组成,所述PNPLA3Ile148Met蛋白或PNPLA3 Ile144Met蛋白分别包含根据SEQ ID NO:42或SEQ ID NO:43的氨基酸序列或由其组成。

[0069] 在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白的变体PNPLA3 mRNA分子包含与SEQ ID NO:34具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成,并且在对应于根据SEQ ID NO:34的位置444的位置处包含鸟嘌呤,或者在对应于根据SEQ ID NO:34的位置442至444的位置处包含密码子AUG。在一些实施方案中,编码变体PNPLA3Ile148Met蛋白的变体PNPLA3 mRNA分子包含与SEQ ID NO:34具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成,并且在对应于根据SEQ ID NO:34的位置444的位置处包含鸟嘌呤,或者在对应于根据SEQ ID NO:34的位置442至444的位置处包含密码子AUG。在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白的变体PNPLA3 mRNA分子包含根据SEQ ID NO:34的核苷酸序列或由其组成。

[0070] 在一些实施方案中,编码变体PNPLA3 Ile144Met蛋白的变体PNPLA3 mRNA分子包含与SEQ ID NO:35具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成,并且在对应于根据SEQ ID NO:35的位置432的位置处包含鸟嘌呤,或者在对应于根据SEQ ID NO:35的位置430至432的位置处包含密码子AUG。在一些实施方案中,编码变体PNPLA3Ile144Met蛋白的变体PNPLA3 mRNA分子包含与SEQ ID NO:35具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成,并且在对应于根据SEQ ID NO:35的位置432的位置处包含鸟嘌呤,或者在对应于根据SEQ ID NO:35的位置430至432的位置处包含密码子AUG。在一些实施方案中,编码变体PNPLA3 Ile144Met蛋白的变体PNPLA3 mRNA分子包含根据SEQ ID NO:35的核苷酸序列或由其组成。

[0071] 在一些实施方案中,变体PNPLA3 mRNA分子包含的核苷酸少于整个变体PNPLA3 mRNA序列。在一些实施方案中,变体PNPLA3mRNA分子包含SEQ ID NO:34或SEQ ID NO:35的至少约5、至少约8、至少约10、至少约12、至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400、至少约500或至少约600个连续核苷酸或由其组成。在一些实施方案中,变体PNPLA3 mRNA分子包含SEQ ID NO:34或SEQ ID NO:35的至少约200至至少约500个连续核苷酸或由其组成。在这方面,较长的mRNA分子比较短的mRNA分子更优选。在一些实施方案中,变体PNPLA3 mRNA分子包含SEQ ID NO:34或SEQ ID NO:35的至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400或至少约500个连续核苷酸或由其组成。在这方面,较长的mRNA分子比较短的mRNA分子更优选。在一些实施方案中,此类变体PNPLA3mRNA分子包括编码对应于根据SEQ ID NO:42的位置148的位置处的甲硫氨酸的密码子,或编码对应于根据SEQ ID NO:43的位置144的位置处的甲硫氨酸的密码子。在一些实施方案中,此类变体PNPLA3 mRNA分子在对应于根据SEQ ID NO:34的位置444的位置处包括鸟嘌呤,或在对应于根据SEQ ID NO:35的位置432的位置处

包括鸟嘌呤。在一些实施方案中,此类变体PNPLA3 mRNA分子在对应于根据SEQ ID NO:34的位置442至444的位置处包括密码子AUG,或在对应于根据SEQ ID NO:35的位置430至432的位置处包括密码子AUG。

[0072] 两种野生型PNPLA3 cDNA分子的核酸序列在SEQ ID NO:36和SEQ ID NO:37中列出。具有SEQ ID NO:36的野生型PNPLA3 cDNA分子在位置444处包含胞嘧啶。具有SEQ ID NO:36的野生型PNPLA3cDNA分子在位置442至444处包含密码子ATC。具有SEQ ID NO:37的野生型PNPLA3 cDNA分子在位置432处包含胞嘧啶。具有SEQ ID NO:37的野生型PNPLA3 cDNA分子在位置430至432处包含密码子ATC。

[0073] 在一些实施方案中,变体PNPLA3 cDNA分子包含编码PNPLA3Ile148Met蛋白或PNPLA3 Ile144Met蛋白的核酸序列或由其组成,所述PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白分别包含与SEQ ID NO:42或SEQ ID NO:43具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列,并且在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸或者在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3 cDNA分子包含编码PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白的核酸序列或由其组成,所述PNPLA3Ile148Met蛋白或PNPLA3 Ile144Met蛋白分别包含与SEQ ID NO:42或SEQ ID NO:43具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列,并且在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸或者在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3 cDNA分子包含编码PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白的核酸序列或由其组成,所述PNPLA3Ile148Met蛋白或PNPLA3 Ile144Met蛋白分别包含根据SEQ ID NO:42或SEQ ID NO:43的氨基酸序列或由其组成。

[0074] 在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白的变体PNPLA3 cDNA分子包含与SEQ ID NO:38具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成,并且在对应于根据SEQ ID NO:38的位置444的位置处包含鸟嘌呤,或者在对应于根据SEQ ID NO:38的位置442至444的位置处包含密码子ATG。在一些实施方案中,编码变体PNPLA3Ile148Met蛋白的变体PNPLA3 cDNA分子包含与SEQ ID NO:38具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成,并且在对应于根据SEQ ID NO:38的位置444的位置处包含鸟嘌呤,或者在对应于根据SEQ ID NO:38的位置442至444的位置处包含密码子ATG。在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白的变体PNPLA3 cDNA分子包含根据SEQ ID NO:38的核苷酸序列或由其组成。

[0075] 在一些实施方案中,编码变体PNPLA3 Ile144Met蛋白的变体PNPLA3 cDNA分子包含与SEQ ID NO:39具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成,并且在对应于根据SEQ ID NO:39的位置432的位置处包含鸟嘌呤,或者在对应于根据SEQ ID NO:39的位置430至432的位置处包含密码子ATG。在一些实施方案中,编码变体PNPLA3Ile144Met蛋白的变体PNPLA3 cDNA分子包含与SEQ ID NO:39具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成,并且在

对应于根据SEQ ID NO:39的位置432的位置处包含鸟嘌呤,或者在对应于根据SEQ ID NO:39的位置430至432的位置处包含密码子ATG。在一些实施方案中,编码变体PNPLA3 Ile144Met蛋白的变体PNPLA3 cDNA分子包含根据SEQ ID NO:39的核苷酸序列或由其组成。

[0076] 在一些实施方案中,变体PNPLA3 cDNA分子包含的核苷酸少于整个变体PNPLA3 cDNA序列。在一些实施方案中,变体PNPLA3cDNA分子包含SEQ ID NO:38或SEQ ID NO:39的至少约5、至少约8、至少约10、至少约12、至少约15、至少约20、至少约25、至少约30、至少约39、至少约40、至少约45、至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400、至少约500或至少约600个连续核苷酸或由其组成。在一些实施方案中,变体PNPLA3 cDNA分子包含SEQ ID NO:38或SEQ ID NO:39的至少约200至至少约500个连续核苷酸或由其组成。在这方面,较长的cDNA分子比较短的cDNA分子更优选。在一些实施方案中,变体PNPLA3 cDNA分子包含SEQ ID NO:38或SEQ ID NO:39的至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400或至少约500个连续核苷酸或由其组成。在这方面,较长的cDNA分子比较短的cDNA分子更优选。在一些实施方案中,此类变体PNPLA3cDNA分子包括编码对应于根据SEQ ID NO:42的位置148的位置处的甲硫氨酸的密码子,或编码对应于根据SEQ ID NO:43的位置144的位置处的甲硫氨酸的密码子。在一些实施方案中,此类变体PNPLA3 cDNA分子在对应于根据SEQ ID NO:38的位置444的位置处包括鸟嘌呤,或在对应于根据SEQ ID NO:39的位置432的位置处包括鸟嘌呤。在一些实施方案中,此类变体PNPLA3 cDNA分子在对应于根据SEQ ID NO:38的位置442至444的位置处包括密码子ATG,或在对应于根据SEQ ID NO:39的位置430至432的位置处包括密码子ATG。

[0077] 与功能性HSD17B13蛋白相关的四种HSD17B13同种型蛋白的氨基酸序列在SEQ ID NO:21 (同种型A)、SEQ ID NO:22 (同种型B)、SEQ ID NO:25 (同种型E) 和SEQ ID NO:29 (同种型I) 中列出。具有SEQ ID NO:21 (同种型A) 的HSD17B13蛋白的长度是300个氨基酸。具有SEQ ID NO:22 (同种型B) 的HSD17B13蛋白的长度是264个氨基酸。具有SEQ ID NO:25 (同种型E) 的HSD17B13蛋白的长度是324个氨基酸。具有SEQ ID NO:29 (同种型I) 的HSD17B13蛋白的长度是271个氨基酸。

[0078] 在一些实施方案中,与功能性HSD17B13蛋白相关的HSD17B13同种型蛋白包含与根据SEQ ID NO:21 (同种型A) 的氨基酸序列具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能性HSD17B13蛋白相关的HSD17B13同种型蛋白包含与根据SEQ ID NO:21 (同种型A) 的氨基酸序列具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能性HSD17B13蛋白相关的HSD17B13同种型蛋白包含根据SEQ ID NO:21 (同种型A) 的氨基酸序列或由其组成。

[0079] 在一些实施方案中,与功能性HSD17B13蛋白相关的HSD17B13同种型蛋白包含与根据SEQ ID NO:22 (同种型B) 的氨基酸序列具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能性HSD17B13蛋白相关的HSD17B13同种型蛋白包含与根据SEQ ID NO:22 (同种型B) 的氨基酸序列具有至少约95%、至少约96%、至少

约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能性HSD17B13蛋白相关的HSD17B13同种型蛋白包含根据SEQ ID NO:22(同种型B)的氨基酸序列或由其组成。

[0080] 在一些实施方案中,与功能性HSD17B13蛋白相关的HSD17B13同种型蛋白包含与根据SEQ ID NO:25(同种型E)的氨基酸序列具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能性HSD17B13蛋白相关的HSD17B13同种型蛋白包含与根据SEQ ID NO:25(同种型E)的氨基酸序列具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能性HSD17B13蛋白相关的HSD17B13同种型蛋白包含根据SEQ ID NO:25(同种型E)的氨基酸序列或由其组成。

[0081] 在一些实施方案中,与功能性HSD17B13蛋白相关的HSD17B13同种型蛋白包含与根据SEQ ID NO:29(同种型I)的氨基酸序列具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能性HSD17B13蛋白相关的HSD17B13同种型蛋白包含与根据SEQ ID NO:29(同种型I)的氨基酸序列具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能性HSD17B13蛋白相关的HSD17B13同种型蛋白包含根据SEQ ID NO:29(同种型I)的氨基酸序列或由其组成。

[0082] 与功能丧失rs72613567HSD17B13蛋白(SEQ ID NO:2)相关的五种HSD17B13同种型蛋白的氨基酸序列在SEQ ID NO:23(同种型C)、SEQ ID NO:24(同种型D)、SEQ ID NO:26(同种型F)、SEQ ID NO:27(同种型G)和SEQ ID NO:28(同种型H)中列出。具有SEQ ID NO:23(同种型C)的HSD17B13蛋白的长度是261个氨基酸。具有SEQ ID NO:24(同种型D)的HSD17B13蛋白的长度是274个氨基酸。具有SEQ ID NO:26(同种型F)的HSD17B13蛋白的长度是284个氨基酸。具有SEQ ID NO:27(同种型G)的HSD17B13蛋白的长度是238个氨基酸。具有SEQ ID NO:28(同种型H)的HSD17B13蛋白的长度是298个氨基酸。

[0083] 在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含与根据SEQ ID NO:23(同种型C)的氨基酸序列具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含与根据SEQ ID NO:23(同种型C)的氨基酸序列具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含根据SEQ ID NO:23(同种型C)的氨基酸序列或由其组成。

[0084] 在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含与根据SEQ ID NO:24(同种型D)的氨基酸序列具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含与根据SEQ ID NO:24(同种型D)的氨基酸序列具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能丧失相关的HSD17B13变体蛋

白包含根据SEQ ID NO:24 (同种型D) 的氨基酸序列或由其组成。

[0085] 在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含与根据SEQ ID NO:26 (同种型F) 的氨基酸序列具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含与根据SEQ ID NO:26 (同种型F) 的氨基酸序列具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含根据SEQ ID NO:26 (同种型F) 的氨基酸序列或由其组成。

[0086] 在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含与根据SEQ ID NO:27 (同种型G) 的氨基酸序列具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含与根据SEQ ID NO:27 (同种型G) 的氨基酸序列具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含根据SEQ ID NO:27 (同种型G) 的氨基酸序列或由其组成。

[0087] 在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含与根据SEQ ID NO:28 (同种型H) 的氨基酸序列具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含与根据SEQ ID NO:28 (同种型H) 的氨基酸序列具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,与功能丧失相关的HSD17B13变体蛋白包含根据SEQ ID NO:28 (同种型H) 的氨基酸序列或由其组成。

[0088] 在一些实施方案中,与功能性HSD17B13蛋白相关的HSD17B13同种型蛋白和与功能丧失相关的HSD17B13变体蛋白是上述蛋白的片段。在一些实施方案中,所述片段包含编码的多肽(如具有SEQ ID NO:21、SEQ ID NO:22、SEQ ID NO:23、SEQ ID NO:24、SEQ ID NO:25、SEQ ID NO:26、SEQ ID NO:27、SEQ ID NO:28或SEQ ID NO:29的氨基酸序列的多肽)的至少约10、至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约55、至少约60、至少约65、至少约70、至少约75、至少约80、至少约85、至少约90、至少约95、至少约100、至少约150或至少约200个连续氨基酸残基。在这方面,较长的片段比较短的片段更优选。在一些实施方案中,所述片段包含编码的多肽的至少约10、至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约55、至少约60、至少约65、至少约70、至少约75、至少约80、至少约85、至少约90、至少约95或至少约100个连续氨基酸残基。在这方面,较长的片段比较短的片段更优选。

[0089] 功能性HSD17B13基因组DNA分子的核酸序列在SEQ ID NO:1中列出。具有SEQ ID NO:1的功能性HSD17B13基因组DNA分子在位置12,667处包含腺嘌呤。

[0090] 在一些实施方案中,功能性HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:21 (同种型A) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,功能

性HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:21(同种型A)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,功能性HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:21(同种型A)的氨基酸序列或由其组成。

[0091] 在一些实施方案中,功能性HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:22(同种型B)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,功能性HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:22(同种型B)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,功能性HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:22(同种型B)的氨基酸序列或由其组成。

[0092] 在一些实施方案中,功能性HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:25(同种型E)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,功能性HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:25(同种型E)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,功能性HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:25(同种型E)的氨基酸序列或由其组成。

[0093] 在一些实施方案中,功能性HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:29(同种型I)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,功能性HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:29(同种型I)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,功能性HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:29(同种型I)的氨基酸序列或由其组成。

[0094] 在一些实施方案中,功能性HSD17B13基因组DNA分子包含与SEQ ID NO:1具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,功能性HSD17B13基因组DNA分子包含与SEQ ID NO:1具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,功能性HSD17B13基因组DNA分子包含根据SEQ ID NO:21的核酸序列或由其组成。

[0095] 编码与功能丧失相关的HSD17B13变体蛋白的变体HSD17B13基因组DNA分子的核酸

序列在SEQ ID NO:2中列出。具有SEQ ID NO:2的变体HSD17B13基因组DNA分子在位置12, 667处包含胸腺嘧啶。

[0096] 在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:23(同种型C)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:23(同种型C)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:23(同种型C)的氨基酸序列或由其组成。

[0097] 在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:24(同种型D)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:24(同种型D)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:24(同种型D)的氨基酸序列或由其组成。

[0098] 在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:26(同种型F)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:26(同种型F)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:26(同种型F)的氨基酸序列或由其组成。

[0099] 在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:27(同种型G)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:27(同种型G)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:27(同种型G)的氨基酸序列或由其组成。

[0100] 在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白

的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:28(同种型H)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:28(同种型H)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,变体HSD17B13基因组DNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:28(同种型H)的氨基酸序列或由其组成。

[0101] 在一些实施方案中,变体HSD17B13基因组DNA分子包含与SEQ ID NO:2具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,变体HSD17B13基因组DNA分子包含与SEQ ID NO:2具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,变体HSD17B13基因组DNA分子包含根据SEQ ID NO:2的核酸序列或由其组成。

[0102] 在一些实施方案中,功能性HSD17B13基因组DNA和变体HSD17B13基因组DNA分子包含少于整个基因组DNA序列。在一些实施方案中,功能性HSD17B13基因组DNA和变体HSD17B13基因组DNA分子包含SEQ ID NO:1(功能性HSD17B13基因组DNA)或SEQ ID NO:2(变体HSD17B13基因组DNA)的至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400、至少约500、至少约600、至少约700、至少约800、至少约900、至少约1000、至少约2000、至少约3000、至少约4000、至少约5000、至少约6000、至少约7000、至少约8000、至少约9000、至少约10000、至少约11000或至少约11500个连续核苷酸或由其组成。在一些实施方案中,功能性HSD17B13基因组DNA和变体HSD17B13基因组DNA分子包含SEQ ID NO:1(功能性HSD17B13基因组DNA)或SEQ ID NO:2(变体HSD17B13基因组DNA)的至少约1000至至少约2000个连续核苷酸或由其组成。

[0103] 在一些实施方案中,功能性HSD17B13基因组DNA和变体HSD17B13基因组DNA分子包含SEQ ID NO:1(功能性HSD17B13基因组DNA)或SEQ ID NO:2(变体HSD17B13基因组DNA)的至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400、至少约500、至少约600、至少约700、至少约800、至少约900、至少约1000、至少约1000、至少约1100、至少约1200、至少约1300、至少约1400、至少约1500、至少约1600、至少约1700、至少约1800、至少约1900、至少约2000、至少约2100、至少约2200、至少约2300、至少约2400或至少约2500个连续核苷酸或由其组成。

[0104] 编码与功能性HSD17B13蛋白相关的同种型蛋白的四种HSD17B13 RNA转录物的核酸序列在SEQ ID NO:44(转录物A)、SEQ ID NO:45(转录物B)、SEQ ID NO:48(转录物E)和SEQ ID NO:52(转录物I)中列出。

[0105] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:21(同种型A)具有至少约90%、至少约91%、至少约92%、至少约93%、

至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:21(同种型A)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:21(同种型A)的氨基酸序列或由其组成。

[0106] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:22(同种型B)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:22(同种型B)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:22(同种型B)的氨基酸序列或由其组成。

[0107] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:25(同种型E)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:25(同种型E)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:25(同种型E)的氨基酸序列或由其组成。

[0108] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:29(同种型I)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:29(同种型I)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的

核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:29 (同种型I) 的氨基酸序列或由其组成。

[0109] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:44 (转录物A) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:44 (转录物A) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含根据SEQ ID NO:44 (转录物A) 的核酸序列或由其组成。

[0110] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:45 (转录物B) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:45 (转录物B) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含根据SEQ ID NO:45 (同种型B) 的核酸序列或由其组成。

[0111] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:48 (转录物E) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:48 (转录物E) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含根据SEQ ID NO:48 (转录物E) 的核酸序列或由其组成。

[0112] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:52 (转录物I) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:52 (转录物I) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 RNA转录物包含根据SEQ ID NO:52 (转录物I) 的核酸序列或由其组成。

[0113] 编码与功能丧失相关的同种型蛋白的五种HSD17B13 RNA转录物的核酸序列在SEQ ID NO:46 (转录物C)、SEQ ID NO:47 (转录物D)、SEQ ID NO:49 (转录物F)、SEQ ID NO:50 (转录物G) 和SEQ ID NO:51 (转录物H) 中列出。

[0114] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ

ID NO:23 (同种型C) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:23 (同种型C) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:23 (同种型C) 的氨基酸序列或由其组成。

[0115] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:24 (同种型D) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:24 (同种型D) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:24 (同种型D) 的氨基酸序列或由其组成。

[0116] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:26 (同种型F) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:26 (同种型F) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:26 (同种型F) 的氨基酸序列或由其组成。

[0117] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:27 (同种型G) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:27 (同种型G) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:27 (同种型G) 的氨基酸序列或由其组成。

[0118] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包

含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:28 (同种型H) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:28 (同种型H) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:28 (同种型H) 的氨基酸序列或由其组成。

[0119] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:46 (转录物C) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:46 (转录物C) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含根据SEQ ID NO:46 (转录物C) 的核酸序列或由其组成。

[0120] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:47 (转录物D) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:47 (转录物D) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含根据SEQ ID NO:47 (转录物D) 的核酸序列或由其组成。

[0121] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:49 (转录物F) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:49 (转录物F) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含根据SEQ ID NO:49 (转录物F) 的核酸序列或由其组成。

[0122] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:50 (转录物G) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:50 (转录物G) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能

丧失相关的同种型蛋白的HSD17B13 RNA转录物包含根据SEQ ID NO:50 (转录物G) 的核酸序列或由其组成。

[0123] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:51 (转录物H) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含与SEQ ID NO:51 (转录物H) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 RNA转录物包含根据SEQ ID NO:51 (转录物H) 的核酸序列或由其组成。

[0124] 在一些实施方案中,功能性HSD17B13 RNA转录物和变体HSD17B13 RNA转录物包含少于RNA转录物序列。在一些实施方案中,功能性HSD17B13 RNA转录物和变体HSD17B13 RNA转录物包含SEQ ID NO:44、SEQ ID NO:45、SEQ ID NO:48或SEQ ID NO:52 (功能性HSD17B13 RNA转录物) 或者SEQ ID NO:46、SEQ ID NO:47、SEQ ID NO:49、SEQ ID NO:50或SEQ ID NO:51 (变体HSD17B13 RNA转录物) 的至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400、至少约500、至少约600、至少约700、至少约800、至少约900、至少约1000、至少约2000或至少约2500个连续核苷酸或由其组成。在一些实施方案中,功能性HSD17B13 RNA转录物和变体HSD17B13 RNA转录物包含少于RNA转录物序列。在一些实施方案中,功能性HSD17B13 RNA转录物和变体HSD17B13 RNA转录物包含SEQ ID NO:44、SEQ ID NO:45、SEQ ID NO:48或SEQ ID NO:52 (功能性HSD17B13 RNA转录物) 或者SEQ ID NO:46、SEQ ID NO:47、SEQ ID NO:49、SEQ ID NO:50或SEQ ID NO:51 (变体HSD17B13 RNA转录物) 的至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400或至少约500个连续核苷酸或由其组成。在一些实施方案中,功能性HSD17B13 RNA转录物和变体HSD17B13 RNA转录物包含SEQ ID NO:44、SEQ ID NO:45、SEQ ID NO:48或SEQ ID NO:52 (功能性HSD17B13 RNA转录物) 或者SEQ ID NO:46、SEQ ID NO:47、SEQ ID NO:49、SEQ ID NO:50或SEQ ID NO:51 (变体HSD17B13 RNA转录物) 的至少约1000至至少约2000个连续核苷酸或由其组成。

[0125] 编码与功能性HSD17B13蛋白相关的同种型蛋白的四种HSD17B13 cDNA转录物的核酸序列在SEQ ID NO:53 (转录物A)、SEQ ID NO:54 (转录物B)、SEQ ID NO:57 (转录物E) 和SEQ ID NO:61 (转录物I) 中列出。

[0126] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:21 (同种型A) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:21 (同种型A) 具有至少约95%、至少约96%、至少约97%、至

少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:21 (同种型A)的氨基酸序列或由其组成。

[0127] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:22 (同种型B)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:22 (同种型B)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:22 (同种型B)的氨基酸序列或由其组成。

[0128] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:25 (同种型E)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:25 (同种型E)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:25 (同种型E)的氨基酸序列或由其组成。

[0129] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:29 (同种型I)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:29 (同种型I)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:29 (同种型I)的氨基酸序列或由其组成。

[0130] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:53 (转录物A)具有至少约90%、至少约91%、至少约92%、至

少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:53(转录物A)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含根据SEQ ID NO:53(转录物A)的核酸序列或由其组成。

[0131] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:54(转录物B)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:54(转录物B)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含根据SEQ ID NO:54(转录物B)的核酸序列或由其组成。

[0132] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:57(转录物E)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:57(转录物E)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含根据SEQ ID NO:57(转录物E)的核酸序列或由其组成。

[0133] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:61(转录物I)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:61(转录物I)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA转录物包含根据SEQ ID NO:61(转录物I)的核酸序列或由其组成。

[0134] 编码与功能丧失相关的同种型蛋白的五种HSD17B13 cDNA转录物的核酸序列在SEQ ID NO:55(转录物C)、SEQ ID NO:56(转录物D)、SEQ ID NO:58(转录物F)、SEQ ID NO:59(转录物G)和SEQ ID NO:60(转录物H)中列出。

[0135] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:23(同种型C)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ

ID NO:23 (同种型C) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:23 (同种型C) 的氨基酸序列或由其组成。

[0136] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:24 (同种型D) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:24 (同种型D) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:24 (同种型D) 的氨基酸序列或由其组成。

[0137] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:26 (同种型F) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:26 (同种型F) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:26 (同种型F) 的氨基酸序列或由其组成。

[0138] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:27 (同种型G) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:27 (同种型G) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:27 (同种型G) 的氨基酸序列或由其组成。

[0139] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:28 (同种型H) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包

含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:28(同种型H)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:28(同种型H)的氨基酸序列或由其组成。

[0140] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:55(转录物C)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:55(转录物C)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含根据SEQ ID NO:55(转录物C)的核酸序列或由其组成。

[0141] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:56(转录物D)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:56(转录物D)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含根据SEQ ID NO:56(转录物D)的核酸序列或由其组成。

[0142] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:58(转录物F)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:58(转录物F)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含根据SEQ ID NO:58(转录物F)的核酸序列或由其组成。

[0143] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:59(转录物G)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:59(转录物G)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含根据SEQ ID NO:59(转录物G)的核酸序列或由其组成。

[0144] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:60(转录物H)具有至少约90%、至少约91%、至少约92%、至少约93%、至

少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含与SEQ ID NO:60 (转录物H) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA转录物包含根据SEQ ID NO:60 (转录物H) 的核酸序列或由其组成。

[0145] 在一些实施方案中,HSD17B13 cDNA转录物包含少于cDNA转录物序列。在一些实施方案中,HSD17B13 cDNA转录物包含SEQ ID NO:53、SEQ ID NO:54、SEQ ID NO:57或SEQ ID NO:61或者SEQ ID NO:55、SEQ ID NO:56、SEQ ID NO:58、SEQ ID NO:59或SEQ ID NO:60的至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400、至少约500、至少约600、至少约700、至少约800、至少约900、至少约1000、至少约2000或至少约2500个连续核苷酸或由其组成。在一些实施方案中,HSD17B13 cDNA转录物包含少于cDNA转录物序列。在一些实施方案中,HSD17B13 cDNA转录物包含SEQ ID NO:53、SEQ ID NO:54、SEQ ID NO:57或SEQ ID NO:61或者SEQ ID NO:55、SEQ ID NO:56、SEQ ID NO:58、SEQ ID NO:59或SEQ ID NO:60的至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400或至少约500个连续核苷酸或由其组成。在一些实施方案中,HSD17B13 cDNA转录物包含SEQ ID NO:53、SEQ ID NO:54、SEQ ID NO:57或SEQ ID NO:61或者SEQ ID NO:55、SEQ ID NO:56、SEQ ID NO:58、SEQ ID NO:59或SEQ ID NO:60的至少约1000至至少约2000个连续核苷酸或由其组成。

[0146] 编码与功能性HSD17B13蛋白相关的同种型蛋白的四种HSD17B13 mRNA分子的核酸序列在SEQ ID NO:3 (转录物A)、SEQ ID NO:4 (转录物B)、SEQ ID NO:7 (转录物E) 和SEQ ID NO:11 (转录物I) 中列出。

[0147] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:21 (同种型A) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:21 (同种型A) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:21 (同种型A) 的氨基酸序列或由其组成。

[0148] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:22 (同种型B) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的

氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:22 (同种型B) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:22 (同种型B) 的氨基酸序列或由其组成。

[0149] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:25 (同种型E) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:25 (同种型E) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:25 (同种型E) 的氨基酸序列或由其组成。

[0150] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:29 (同种型I) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:29 (同种型I) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:29 (同种型I) 的氨基酸序列或由其组成。

[0151] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:3 (转录物A) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:3 (转录物A) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含根据SEQ ID NO:3 (转录物A) 的核酸序列或由其组成。

[0152] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:4 (转录物B) 具有至少约90%、至少约91%、至少约92%、至少约

93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:4 (转录物B) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含根据SEQ ID NO:4 (转录物B) 的核酸序列或由其组成。

[0153] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:7 (转录物E) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:7 (转录物E) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含根据SEQ ID NO:7 (转录物E) 的核酸序列或由其组成。

[0154] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:11 (转录物I) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:11 (转录物I) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 mRNA分子包含根据SEQ ID NO:11 (转录物I) 的核酸序列或由其组成。

[0155] 编码与功能丧失相关的同种型蛋白的五种HSD17B13 mRNA分子的核酸序列在SEQ ID NO:5 (转录物C)、SEQ ID NO:6 (转录物D)、SEQ ID NO:8 (转录物F)、SEQ ID NO:9 (转录物G) 和SEQ ID NO:10 (转录物H) 中列出。

[0156] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:23 (同种型C) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:23 (同种型C) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:23 (同种型C) 的氨基酸序列或由其组成。

[0157] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:24 (同种型D) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序

列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO: 24(同种型D)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:24(同种型D)的氨基酸序列或由其组成。

[0158] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:26(同种型F)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:26(同种型F)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:26(同种型F)的氨基酸序列或由其组成。

[0159] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:27(同种型G)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:27(同种型G)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:27(同种型G)的氨基酸序列或由其组成。

[0160] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:28(同种型H)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:28(同种型H)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:28(同种型H)的氨基酸序列或由其组成。

[0161] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:5(转录物C)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核

酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:5 (转录物C) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含根据SEQ ID NO:5 (转录物C) 的核酸序列或由其组成。

[0162] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:6 (转录物D) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:6 (转录物D) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含根据SEQ ID NO:6 (转录物D) 的核酸序列或由其组成。

[0163] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:8 (转录物F) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:8 (转录物F) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含根据SEQ ID NO:8 (转录物F) 的核酸序列或由其组成。

[0164] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:9 (转录物G) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:9 (转录物G) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含根据SEQ ID NO:9 (转录物G) 的核酸序列或由其组成。

[0165] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:10 (转录物H) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含与SEQ ID NO:10 (转录物H) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 mRNA分子包含根据SEQ ID NO:10 (转录物H) 的核酸序列或由其组成。

[0166] 在一些实施方案中,HSD17B13 mRNA分子包含的核苷酸少于整个mRNA序列。在一些实施方案中,HSD17B13 mRNA分子包含SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:7或SEQ ID

NO:11或者SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:9或SEQ ID NO:10的至少约5、至少约8、至少约10、至少约12、至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400、至少约500、至少约600、至少约700、至少约800或至少约900个连续核苷酸或由其组成。在一些实施方案中，HSD17B13 mRNA分子包含SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:7或SEQ ID NO:11或者SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:9或SEQ ID NO:10的至少约200至至少约500个连续核苷酸或由其组成。在这方面，较长的mRNA分子比较短的mRNA分子更优选。在一些实施方案中，HSD17B13mRNA分子包含SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:7或SEQ ID NO:11或者SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:9或SEQ ID NO:10的至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400或至少约500个连续核苷酸或由其组成。在这方面，较长的mRNA分子比较短的mRNA分子更优选。

[0167] 编码与功能性HSD17B13蛋白相关的同种型蛋白的四种HSD17B13 cDNA分子的核酸序列在SEQ ID NO:12 (转录物A)、SEQ ID NO:13 (转录物B)、SEQ ID NO:16 (转录物E) 和SEQ ID NO:20 (转录物I) 中列出。

[0168] 在一些实施方案中，编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成，所述HSD17B13同种型蛋白包含与SEQ ID NO:21 (同种型A) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中，编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成，所述HSD17B13同种型蛋白包含与SEQ ID NO:21 (同种型A) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中，编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成，所述HSD17B13同种型蛋白包含根据SEQ ID NO:21 (同种型A) 的氨基酸序列或由其组成。

[0169] 在一些实施方案中，编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成，所述HSD17B13同种型蛋白包含与SEQ ID NO:22 (同种型B) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中，编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成，所述HSD17B13同种型蛋白包含与SEQ ID NO:22 (同种型B) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中，编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成，所述HSD17B13同种型蛋白包含根据SEQ ID NO:22 (同种型B) 的氨基酸序列或由其组成。

[0170] 在一些实施方案中，编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成，所述HSD17B13同种型蛋白

包含与SEQ ID NO:25 (同种型E) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:25 (同种型E) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:25 (同种型E) 的氨基酸序列或由其组成。

[0171] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:29 (同种型I) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:29 (同种型I) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:29 (同种型I) 的氨基酸序列或由其组成。

[0172] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:12 (转录物A) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:12 (转录物A) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含根据SEQ ID NO:12 (转录物A) 的核酸序列或由其组成。

[0173] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:13 (转录物B) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:13 (转录物B) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含根据SEQ ID NO:13 (转录物B) 的核酸序列或由其组成。

[0174] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:16 (转录物E) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列

同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:16(转录物E)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含根据SEQ ID NO:16(转录物E)的核酸序列或由其组成。

[0175] 在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:20(转录物I)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:20(转录物I)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能性HSD17B13蛋白相关的同种型蛋白的HSD17B13 cDNA分子包含根据SEQ ID NO:20(转录物I)的核酸序列或由其组成。

[0176] 编码与功能丧失相关的同种型蛋白的五种HSD17B13 cDNA分子的核酸序列在SEQ ID NO:14(转录物C)、SEQ ID NO:15(转录物D)、SEQ ID NO:17(转录物F)、SEQ ID NO:18(转录物G)和SEQ ID NO:19(转录物H)中列出。

[0177] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:23(同种型C)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:23(同种型C)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:23(同种型C)的氨基酸序列或由其组成。

[0178] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:24(同种型D)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:24(同种型D)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:24(同种型D)的氨基酸序列或由其组成。

[0179] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:26(同种型F)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、

至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:26(同种型F)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:26(同种型F)的氨基酸序列或由其组成。

[0180] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:27(同种型G)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:27(同种型G)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:27(同种型G)的氨基酸序列或由其组成。

[0181] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:28(同种型H)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含与SEQ ID NO:28(同种型H)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含编码HSD17B13同种型蛋白的核酸序列或由其组成,所述HSD17B13同种型蛋白包含根据SEQ ID NO:28(同种型H)的氨基酸序列或由其组成。

[0182] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:14(转录物C)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:14(转录物C)具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含根据SEQ ID NO:14(转录物C)的核酸序列或由其组成。

[0183] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:15(转录物D)具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13

cDNA分子包含与SEQ ID NO:15 (转录物D) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含根据SEQ ID NO:15 (转录物D) 的核酸序列或由其组成。

[0184] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:17 (转录物F) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:17 (转录物F) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含根据SEQ ID NO:17 (转录物F) 的核酸序列或由其组成。

[0185] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:18 (转录物G) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:18 (转录物G) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含根据SEQ ID NO:18 (转录物G) 的核酸序列或由其组成。

[0186] 在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:19 (转录物H) 具有至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含与SEQ ID NO:19 (转录物H) 具有至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的核酸序列或由其组成。在一些实施方案中,编码与功能丧失相关的同种型蛋白的HSD17B13 cDNA分子包含根据SEQ ID NO:19 (转录物H) 的核酸序列或由其组成。

[0187] 在一些实施方案中,HSD17B13 cDNA分子包含的核苷酸少于整个cDNA序列。在一些实施方案中,HSD17B13 cDNA分子包含SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:16或SEQ ID NO:20或者SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:17、SEQ ID NO:18或SEQ ID NO:19的至少约5、至少约8、至少约10、至少约12、至少约15、至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400、至少约500、至少约600、至少约700、至少约800或至少约900个连续核苷酸或由其组成。在一些实施方案中,HSD17B13 cDNA分子包含SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:16或SEQ ID NO:20或者SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:17、SEQ ID NO:18或SEQ ID NO:19的至少约200至至少约500个连续核苷酸或由其组成。在这方面,较长的cDNA分子比较短的cDNA分子更优选。在一些实施方案中,HSD17B13 cDNA分子包含SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:16或SEQ ID NO:20或者

SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:17、SEQ ID NO:18或SEQ ID NO:19的至少约50、至少约60、至少约70、至少约80、至少约90、至少约100、至少约200、至少约300、至少约400或至少约500个连续核苷酸或由其组成。在这方面,较长的cDNA分子比较短的cDNA分子更优选。

[0188] 本文所述的探针和引物可以用于与本文所述的功能性或变体PNPLA3基因组DNA分子、mRNA分子或来源于mRNA分子的cDNA分子中的任一种杂交。引物可以用于例如扩增本文所述的功能性或变体PNPLA3基因组DNA分子、mRNA分子或来源于mRNA分子的cDNA分子中任一种的部分,使得可以对扩增产物进行例如检测或测序。

[0189] 例如,探针和引物可以用于与本文所述的野生型PNPLA3基因组DNA分子中的任一种(包括包含SEQ ID NO:30的野生型PNPLA3基因组DNA分子)杂交。探针和引物还可以用于与本文所述的野生型PNPLA3 mRNA分子中的任一种(包括包含SEQ ID NO:32或SEQ ID NO:33的野生型PNPLA3 mRNA分子)杂交。探针和引物还可以用于与本文所述的野生型PNPLA3 cDNA分子中的任一种(包括包含SEQ ID NO:36或SEQ ID NO:37的野生型PNPLA3 cDNA分子)杂交。

[0190] 探针和引物还可以用于与本文所述的变体PNPLA3基因组DNA分子中的任一种(包括包含SEQ ID NO:31的变体PNPLA3基因组DNA分子)杂交。探针和引物还可以用于与本文所述的变体PNPLA3mRNA分子中的任一种(包括包含SEQ ID NO:34或SEQ ID NO:35的变体PNPLA3 mRNA分子)杂交。探针和引物还可以用于与本文所述的变体PNPLA3 cDNA分子中的任一种(包括包含SEQ ID NO:38或SEQ ID NO:39的变体PNPLA3 cDNA分子)杂交。

[0191] 探针可以用于例如检测本文所述的功能性或变体HSD17B13基因组DNA分子、mRNA分子或来源于mRNA分子的cDNA分子中的任一种。引物可以用于例如扩增本文所述的功能性或变体HSD17B13基因组DNA分子、mRNA分子或来源于mRNA分子的cDNA分子中任一种的部分,使得可以对扩增产物进行例如检测或测序。

[0192] 例如,探针和引物可以用于与本文所述的功能性HSD17B13基因组DNA分子中的任一种(包括包含SEQ ID NO:1的功能性HSD17B13基因组DNA分子)杂交。探针和引物还可以用于与本文所述的功能性HSD17B13 RNA转录物中的任一种(包括包含SEQ ID NO:44、SEQ ID NO:45、SEQ ID NO:48或SEQ ID NO:52的功能性HSD17B13 RNA转录物)杂交。探针和引物还可以用于与本文所述的功能性HSD17B13 DNA转录物中的任一种(包括包含SEQ ID NO:53、SEQ ID NO:54、SEQ ID NO:57或SEQ ID NO:61的功能性HSD17B13 DNA转录物)杂交。探针和引物还可以用于与本文所述的功能性HSD17B13 mRNA分子中的任一种(包括包含SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:7或SEQ ID NO:11的功能性HSD17B13 mRNA分子)杂交。探针和引物还可以用于与本文所述的功能性HSD17B13cDNA分子中的任一种(包括包含SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:16或SEQ ID NO:20的功能性HSD17B13 cDNA分子)杂交。

[0193] 探针和引物还可以用于与本文所述的变体HSD17B13基因组DNA分子中的任一种(包括包含SEQ ID NO:2的变体HSD17B13基因组DNA分子)杂交。探针和引物还可以用于与本文所述的变体HSD17B13 RNA转录物中的任一种(包括包含SEQ ID NO:46、SEQ ID NO:47、SEQ ID NO:49、SEQ ID NO:50或SEQ ID NO:51的变体HSD17B13 RNA转录物)杂交。探针和引物还可以用于与本文所述的变体HSD17B13 DNA转录物中的任一种(包括包含SEQ ID NO:55、SEQ ID NO:56、SEQ ID NO:58、SEQ ID NO:59或SEQ ID NO:60的变体HSD17B13 DNA转录

物)杂交。探针和引物还可以用于与本文所述的变体HSD17B13 mRNA分子中的任一种(包括包含SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:9或SEQ ID NO:10的变体HSD17B13 mRNA分子)杂交。探针和引物还可以用于与本文所述的HSD17B13 cDNA分子中的任一种(包括包含SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:17、SEQ ID NO:18或SEQ ID NO:19的HSD17B13 cDNA分子)杂交。

[0194] 在一些实施方案中,本文所述的探针和/或引物包含与本文公开的核酸分子中的任一种特异性杂交的核酸序列,或其互补体。在一些实施方案中,探针或引物在严格条件下与本文公开的核酸分子中的任一种特异性杂交。本公开还提供了具有在温和条件下与本文公开的核酸分子中的任一种杂交的核酸序列或其互补体的核酸分子。

[0195] 促进DNA杂交的适当的严格条件包括,例如,在约45℃下6X氯化钠/柠檬酸钠(SSC),随后在50℃下用2X SSC洗涤(还参见Current Protocols in Molecular Biology, John Wiley&Sons,纽约(1989),6.3.1-6.3.6)。通常,用于杂交和检测的严格条件将是如下的那些:其中在pH 7.0至8.3时的盐浓度低于约1.5M Na离子,通常约0.01至1.0M Na离子浓度(或其他盐),并且温度对于短探针(例如10至50个核苷酸)而言是至少约30℃,并且对于较长探针(例如,大于50个核苷酸)而言是至少约60℃。严格条件还可以通过添加去稳定剂如甲酰胺来实现。示例性低严格条件包括在37℃下用30%至35%甲酰胺、1M NaCl、1%SDS(十二烷基硫酸钠)的缓冲溶液杂交,并在50℃至55℃下在1X至2X SSC(20X SSC=3.0M NaCl/0.3M柠檬酸三钠)中洗涤。示例性中等严格条件包括在37℃下在40%至45%甲酰胺、1.0M NaCl、1%SDS中杂交,并在55℃至60℃下在0.5X至1X SSC中洗涤。示例性高严格条件包括在37℃下在50%甲酰胺、1M NaCl、1%SDS中杂交,并在60℃至65℃下在0.1X SSC中洗涤。任选地,洗涤缓冲液可以包含约0.1%至约1%的SDS。杂交的持续时间通常小于约24小时,通常为约4至约12小时。洗涤时间的持续时间将至少是足以达到平衡的时间长度。

[0196] 在杂交反应中,特异性通常随杂交后洗涤而变,关键因素是最后一次洗涤溶液的离子强度和温度。对于DNA-DNA杂合体, T_m 可以从Meinkoth和Wahl,Anal.Biochem.,1984,138,267-284的方程约计: $T_m = 81.5^\circ\text{C} + 16.6(\log M) + 0.41(\%GC) - 0.61(\% \text{甲酰胺}) - 500/L$;其中M是单价阳离子的摩尔浓度,%GC是DNA中鸟苷和胞嘧啶核苷酸的百分比,%甲酰胺是杂交溶液中甲酰胺的百分比,并且L是杂合体以碱基对计的长度。 T_m 是50%互补靶序列与完全匹配的探针杂交的温度(在确定的离子强度和pH下)。对于每1%错配, T_m 下降约1℃;因此,可以调整 T_m 、杂交条件和/或洗涤条件以与具有所需同一性的序列杂交。例如,如果寻找具有 $\geq 90\%$ 同一性的序列,则 T_m 可以降低10℃。通常,选择的严格条件在确定的离子强度和pH下比特定序列与其互补体的热熔点(T_m)低约5℃。然而,高严格条件可以利用在比热熔点(T_m)低1℃、2℃、3℃或4℃的温度下杂交和/或洗涤;中等严格条件可以利用在比热熔点(T_m)低6℃、7℃、8℃、9℃或10℃的温度下杂交和/或洗涤;低严格条件可以利用在比热熔点(T_m)低11℃、12℃、13℃、14℃、15℃或20℃的温度下杂交和/或洗涤。使用所述方程、杂交和洗涤组合物以及所需 T_m ,本领域普通技术人员将理解,固有地描述了杂交和/或洗涤溶液严格性的变化。如果所需错配程度导致低于45℃(水溶液)或32℃(甲酰胺溶液)的 T_m ,则增加SSC浓度使得可以使用较高温度是最佳的。

[0197] 本文所述的探针可以与标记连接或融合以帮助检测。标记可以是可直接检测的(例如,荧光团)或可间接检测的(例如,半抗原、酶或荧光团淬灭剂)。此类标记可以是可通

过光谱学、光化学、生物化学、免疫化学或化学方式检测的。此类标记包括,例如,可以用放射计数装置测量的放射性标记;颜料、染料或可以用分光光度计目视观察或测量的其他色原体;可以用自旋标记分析仪测量的自旋标记;和荧光标记(例如荧光团),其中输出信号通过激发合适的分子加合物来产生,并且所述标记可以通过用染料吸收的光激发来可视化,或者可以用标准荧光计或成像系统来测量。标记还可以是例如化学发光物质,其中输出信号通过对信号化合物的化学修饰来产生;含金属的物质;或酶,其中发生信号的酶依赖性二次产生,如由无色底物形成有色产物。术语“标记”还可以指“标签”或半抗原,其可以选择性地结合缀合分子,使得缀合分子在随后与底物一起加入时,用于产生可检测的信号。例如,可以使用生物素作为标签,并且然后使用辣根过氧化物(HRP)的抗生物素蛋白或抗生蛋白链菌素缀合物来结合所述标签,并且然后使用量热底物(例如,四甲基联苯胺(TMB))或荧光底物来检测HRP的存在。可以用作标签以促进纯化的示例性标记包括但不限于myc、HA、FLAG或3XFLAG、6XHis或聚组氨酸、谷胱甘肽-S-转移酶(GST)、麦芽糖结合蛋白、表位标签或免疫球蛋白的Fc部分。许多标记包括例如,颗粒、荧光团、半抗原、酶及其量热、荧光和化学发光底物和其他标记。

[0198] 探针或引物可以包括任何合适的长度,其非限制性实例包括长度为至少约5、至少约8、至少约10、至少约11、至少约12、至少约13、至少约14、至少约15、至少约16、至少约17、至少约18、至少约19、至少约20、至少约21、至少约22、至少约23、至少约24或至少约25个核苷酸。在一些实施方案中,探针或引物包含长度为至少约18个核苷酸至长度为约25个核苷酸。探针或引物可以包含长度为约10至约35、约10至约30、约10至约25、约12至约30、约12至约28、约12至约24、约15至约30、约15至约25、约18至约30、约18至约25、约18至约24或约18至约22个核苷酸。在一些实施方案中,探针或引物的长度为约18至约30个核苷酸。可选地,在一些实施方案中,探针包含至少约20、至少约25、至少约30、至少约35、至少约40、至少约45、至少约50、至少约55、至少约60、至少约65、至少约70、至少约75、至少约80、至少约85、至少约90、至少约95或至少约100个核苷酸或由其组成。

[0199] 在一些实施方案中,探针和/或引物可以与核酸分子的至少约15个连续核苷酸杂交,所述核酸分子与本文所述的野生型PNPLA3或HSD17B13核酸分子或者变体PNPLA3或HSD17B13核酸分子中的任一种至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或100%相同。

[0200] 在一些实施方案中,探针或引物包含DNA。在一些实施方案中,探针或引物包含RNA。

[0201] 本文所述的探针和引物还可以是改变特异性探针和改变特异性引物。改变特异性探针或改变特异性引物可以包含与编码变体PNPLA3蛋白的核酸序列互补和/或杂交或特异性杂交,但不与编码野生型PNPLA3蛋白的核酸序列互补和/或杂交或特异性杂交的核酸序列。在这种情况下,“特异性杂交”意指探针或引物(例如,改变特异性探针或改变特异性引物)不与编码野生型PNPLA3蛋白的核酸分子杂交。在本文中,术语“特异性杂交”意指探针或引物仅与指定的核酸分子杂交,而不与另一种核酸分子杂交。因此,与编码包含I148M变异的PNPLA3蛋白的核酸分子特异性杂交的探针或引物不与编码不包含I148M变异的PNPLA3蛋白的核酸分子杂交。改变特异性探针或改变特异性引物还可以包含与编码野生型PNPLA3蛋白的核酸序列互补和/或杂交或特异性杂交,但不与编码变体PNPLA3蛋白的核酸序列互补

和/或杂交或特异性杂交的核酸序列。在这种情况下,“特异性杂交”意指探针或引物(例如,改变特异性探针或改变特异性引物)不与编码变体PNPLA3蛋白的核酸分子杂交。

[0202] 改变特异性探针或改变特异性引物还可以包含与编码变体HSD17B13蛋白的核酸序列互补和/或杂交或特异性杂交,但不与编码功能性HSD17B13蛋白的核酸序列互补和/或杂交或特异性杂交的核酸序列。在这种情况下,“特异性杂交”意指探针或引物(例如,改变特异性探针或改变特异性引物)不与编码功能性HSD17B13蛋白的核酸分子杂交。例如,在这种情况下,“特异性杂交”意指探针或引物不与编码非活性/功能丧失HSD17B13蛋白的核酸分子杂交。改变特异性探针或改变特异性引物还可以包含与编码功能性HSD17B13蛋白的核酸序列互补和/或杂交或特异性杂交,但不与编码变体HSD17B13蛋白的核酸序列互补和/或杂交或特异性杂交的核酸序列。在这种情况下,“特异性杂交”意指探针或引物(例如,改变特异性探针或改变特异性引物)不与编码变体HSD17B13蛋白的核酸分子杂交。

[0203] 在一些实施方案中,改变特异性探针或改变特异性引物包含与PNPLA3核酸序列的一部分互补和/或杂交或特异性杂交的核酸序列,所述PNPLA3核酸序列的一部分在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸,或在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,改变特异性探针或改变特异性引物包含与变体PNPLA3基因组DNA的一部分互补和/或杂交或特异性杂交的核酸序列,所述变体PNPLA3基因组DNA的一部分在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子。在一些实施方案中,改变特异性探针或改变特异性引物包含与变体PNPLA3 mRNA的一部分互补和/或杂交或特异性杂交的核酸序列,所述变体PNPLA3 mRNA的一部分在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子。在一些实施方案中,改变特异性探针或改变特异性引物包含与变体PNPLA3 mRNA的一部分互补和/或杂交或特异性杂交的核酸序列,所述变体PNPLA3 mRNA的一部分在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子。在一些实施方案中,改变特异性探针或改变特异性引物包含与变体PNPLA3 cDNA的一部分互补和/或杂交或特异性杂交的核酸序列,所述变体PNPLA3 cDNA的一部分在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子。在一些实施方案中,改变特异性探针或改变特异性引物包含与变体PNPLA3 cDNA的一部分互补和/或杂交或特异性杂交的核酸序列,所述变体PNPLA3 cDNA的一部分在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。

[0204] 在一些实施方案中,改变特异性探针或改变特异性引物包含与HSD17B13基因组DNA的一部分互补和/或杂交或特异性杂交的核酸序列,所述HSD17B13基因组DNA的一部分在对应于根据SEQ ID NO:1的位置12,667的位置处包含腺嘌呤。在一些实施方案中,改变特异性探针或改变特异性引物包含与HSD17B13基因组DNA的一部分互补和/或杂交或特异性杂交的核酸序列,所述HSD17B13基因组DNA的一部分在对应于根据SEQ ID NO:2的位置12,667的位置处包含胸腺嘧啶。

[0205] 在一些实施方案中,与探针或引物杂交的核酸分子的部分包含约10至约200、约10至约150、约10至约100、约10至约50、约10至约40、约10至约30或约10至约20个核苷酸,并且包含对应于含有编码特定变异(例如,PNPLA3的I148M或变体HSD17B13蛋白与对应的野生型HSD17B13蛋白不同的部分)的密码子位置的密码子。在一些优选的实施方案中,与探针或引物杂交的核酸分子的部分包含约10至约50、约10至约40、约10至约30或约10至约20个核

苷酸,并且包含对应于含有编码特定变异(例如,PNPLA3的I148M或变体HSD17B13蛋白与对应的野生型HSD17B13蛋白不同的部分)的密码子位置的密码子。

[0206] 本文所述的试剂盒可以包含检测和/或扩增测定试剂,所述测定试剂可以用于检测和/或扩增本文所述的野生型PNPLA3和/或HSD17B13核酸分子中的任一种和/或本文所述的变体PNPLA3和/或HSD17B13核酸分子中的任一种。在一些实施方案中,用于这种检测和/或扩增的试剂盒可以含有本文所述的试剂(例如探针和引物)中的任一种。在一些实施方案中,基础试剂盒可以包括容器,所述容器具有针对本文公开的核酸分子中任一种中的基因座的至少一种探针或引物或至少两种探针或引物,如改变特异性探针或改变特异性引物。试剂盒还可以任选地包括使用说明。试剂盒还可以包含其他任选的试剂盒组分,如例如以下中的一个或多个:针对每个所扩增基因座的等位基因梯、足够量的用于扩增的酶、促进扩增的扩增缓冲液、促进酶活性的二价阳离子溶液、用于扩增过程中链延伸的dNTP、用于制备用于电泳的扩增材料的上样溶液、作为模板对照的基因组DNA、确保材料在分离介质中如所预期迁移的尺寸标记物,以及用于训练使用者并限制使用中的错误的方案和手册。试剂盒中各种试剂的量也可以根据许多因素而变化,如过程的最佳灵敏度。提供用于手动应用的测试试剂盒或用于自动化样品制备、反应装置、检测器或分析仪的测试试剂盒也在这些传授内容的范围内。在一些实施方案中,试剂盒包括至少一种用于检测的标记探针(例如,改变特异性探针)。在一些实施方案中,本文公开的试剂盒中的任一种还可以包括进行退火反应所需的产品和试剂,以及说明书。

[0207] 本公开提供了用于检测本文所述的野生型PNPLA3蛋白中的任一种的存在的方法。本公开还提供了用于检测本文所述的变体PNPLA3蛋白中的任一种的存在的方法。本公开还提供了用于检测本文所述的野生型PNPLA3核酸分子(例如,基因组DNA分子、mRNA分子和cDNA分子)中的任一种的存在的方法。本公开还提供了用于检测本文所述的变体PNPLA3核酸分子(例如,基因组DNA分子、mRNA分子和cDNA分子)中的任一种的存在的方法。

[0208] 本公开还提供了用于检测本文所述的功能性HSD17B13蛋白中的任一种的存在的方法。本公开还提供了用于检测本文所述的变体HSD17B13蛋白中的任一种的存在的方法。本公开还提供了用于检测本文所述的功能性HSD17B13核酸分子(例如,基因组DNA分子、RNA转录物、cDNA转录物、mRNA分子和cDNA分子)中的任一种的存在的方法。本公开还提供了用于检测本文所述的变体HSD17B13核酸分子(例如,基因组DNA分子、RNA转录物、cDNA转录物、mRNA分子和cDNA分子)中的任一种的存在的方法。

[0209] 在本文所述的方法中的任一种的一些实施方案中,在受试者或患者中检测或寻求检测功能性HSD17B13蛋白或编码所述功能性HSD17B13蛋白的核酸分子。在一些实施方案中,受试者或患者包含功能性HSD17B13蛋白。在一些实施方案中,功能性HSD17B13蛋白是本文所述的功能性HSD17B13蛋白之一(其可以由本文所述的编码所述功能性HSD17B13蛋白的核酸分子中的任一种编码)。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有根据SEQ ID NO:40的氨基酸序列的HSD17B13蛋白的至少90%、至少80%、至少70%、至少60%、至少50%、至少40%、至少30%、至少20%、至少10%、至少5%或至少1%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有根据SEQ ID NO:40的氨基酸序列的HSD17B13蛋白的至少90%、至少80%、至少70%、至少60%或至少50%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少90%、至少

80%、至少70%、至少60%、至少50%、至少40%、至少30%、至少20%、至少10%、至少5%或至少1%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少70%、至少60%、至少50%、至少40%、至少30%、至少20%、至少10%、至少5%或至少1%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少50%、至少40%、至少30%、至少20%、至少10%、至少5%或至少1%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少30%、至少20%、至少10%、至少5%或至少1%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少90%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少80%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少70%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少60%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少50%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少40%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少30%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少20%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少10%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少5%。在一些实施方案中,功能性HSD17B13蛋白的生物活性是具有SEQ ID NO:40的HSD17B13蛋白的至少1%。在一些实施方案中,HSD17B13蛋白的活性(例如,功能性)可以通过例如进行氧化还原酶活性测定来确定。

[0210] 应理解,群体内的基因序列和由此类基因编码的mRNA和蛋白质可以由于多态性(如单核苷酸多态性)而变化。本文提供的序列仅仅是示例性的序列。变体PNPLA3和HSD17B13基因组DNA、mRNA、cDNA和多肽的其他序列也是可能的。

[0211] 生物样品可以来源于受试者的任何细胞、组织或生物流体。所述样品可以包括任何临床相关组织,如骨髓样品、肿瘤活组织检查、细针抽吸物或体液样品(如血液、龈沟液、血浆、血清、淋巴、腹水、囊液或尿液)。在一些实施方案中,所述样品包括颊粘膜拭子。在本文公开的方法中使用的样品将基于测定形式、检测方法的性质以及用作样品的组织、细胞或提取物而有所不同。生物样品可以根据所采用的测定进行不同的处理。例如,当检测变体PNPLA3核酸分子时,可以采用设计为针对基因组DNA分离或富集样品的初步处理。多种技术可以用于此目的。当检测变体PNPLA3 mRNA的水平时,可以使用不同的技术富集具有mRNA的生物样品。可以使用各种方法来检测mRNA的存在或水平,或者特定变体基因组DNA基因座的存在。

[0212] 在一些实施方案中,通过对蛋白质的至少一部分进行测序以确定所述蛋白质是否包含编码本文所述的变体PNPLA3蛋白或HSD17B13蛋白(例如,功能性或变体)中的任一种的氨基酸序列,来检测特定PNPLA3蛋白或HSD17B13蛋白(例如,功能性或变体)的存在或不存在。在一些实施方案中,通过进行免疫测定(如ELISA)以确定样品中是否存在本文所述的变体PNPLA3蛋白或HSD17B13蛋白(例如,功能性或变体)中的任一种,来检测特定PNPLA3蛋白或HSD17B13蛋白(例如,功能性或变体)的存在或不存在。

[0213] 在一些实施方案中,测序的蛋白质的部分包含约5至约100、约5至约50、约5至约40、约5至约30、约5至约20或约5至约10个氨基酸,并且包含对应于含有变异(例如,PNPLA3的I148M或变体HSD17B13蛋白与对应的野生型HSD17B13蛋白不同的部分)的位置的位置。在一些优选的实施方案中,测序的蛋白质的部分包含约5至约20或约5至约10个氨基酸,并且包含对应于含有变异(例如,PNPLA3的I148M或变体HSD17B13蛋白与对应的野生型HSD17B13蛋白不同的部分)的位置的位置。

[0214] 蛋白质测序技术的说明性非限制性实例包括但不限于质谱分析和埃德曼降解(Edman degradation)。免疫测定的说明性实例包括但不限于免疫沉淀、蛋白质印迹、免疫组织化学、ELISA、免疫细胞化学、流式细胞术和免疫PCR。使用各种技术(例如量热技术、荧光技术、化学发光技术或放射性技术)可检测地标记的多克隆或单克隆抗体适用于免疫测定。

[0215] 在一些实施方案中,通过对核酸分子的至少一部分进行测序以确定所述核酸分子是否包含根据本文所述的变体PNPLA3核酸分子或HSD17B13核酸分子(例如,功能性或变体)中的任一种的核酸序列,来检测特定PNPLA3核酸分子或HSD17B13核酸分子(例如,功能性或变体基因组DNA、mRNA、cDNA、RNA转录物或cDNA转录物)的存在或不存在。

[0216] 在一些实施方案中,测序的核酸分子的部分包含约10至约200、约10至约150、约10至约100、约10至约50、约10至约40、约10至约30或约10至约20个核苷酸,并且包含对应于含有编码特定变异(例如,PNPLA3的I148M或变体HSD17B13蛋白与对应的野生型HSD17B13蛋白不同的部分)的密码子的位置的密码子。在一些优选的实施方案中,测序的核酸分子的部分包含约10至约50、约10至约40、约10至约30或约10至约20个核苷酸,并且包含对应于含有编码特定变异(例如,PNPLA3的I148M或变体HSD17B13蛋白与对应的野生型HSD17B13蛋白不同的部分)的密码子的位置的密码子。

[0217] 在一些实施方案中,检测受试者中本文所述的特定PNPLA3核酸分子或HSD17B13核酸分子中的任一种(例如,功能性或变体基因组DNA分子、mRNA分子、cDNA分子、RNA转录物或cDNA转录物中的任一种)存在或不存在的方法包括:对从受试者获得的生物样品进行测定,所述测定确定所述生物样品中的核酸分子是否包含本文所述的特定PNPLA3核酸分子或HSD17B13核酸分子中的任一种(例如,功能性或变体基因组DNA分子、mRNA分子、cDNA分子、RNA转录物或cDNA转录物中的任一种)。在一些实施方案中,生物样品包含细胞或细胞裂解物。此类方法还可以包括,例如,从受试者获得生物样品,任选地将mRNA逆转录成cDNA,并进行测定。此类测定可以包括,例如,确定本文所述的特定核酸分子的特定位置的同一性。

[0218] 例如,所述测定可以包括使用包含与PNPLA3核酸序列的一部分或与其相邻的部分互补和/或杂交或特异性杂交的核酸序列的改变特异性探针或改变特异性引物,所述PNPLA3核酸序列的一部分在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸,或在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,所述测定可以包括使用包含与变体PNPLA3基因组DNA的一部分或与其相邻的部分互补和/或杂交或特异性杂交的核酸序列的改变特异性探针或改变特异性引物,所述变体PNPLA3基因组DNA的一部分在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子。在一些实施方案中,所述测定可以包括使用包含与变体PNPLA3 mRNA的一部分或与其相邻的部分互补和/或杂交或特异性杂交的核酸序列的改变特异性探针或改变特异性引物,所述变

体PNPLA3 mRNA的一部分在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子。在一些实施方案中,所述测定可以包括使用包含与变体PNPLA3 mRNA的一部分或与其相邻的部分互补和/或杂交或特异性杂交的核酸序列的改变特异性探针或改变特异性引物,所述变体PNPLA3 mRNA的一部分在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子。在一些实施方案中,所述测定可以包括使用包含与变体PNPLA3 cDNA的一部分或与其相邻的部分互补和/或杂交或特异性杂交的核酸序列的改变特异性探针或改变特异性引物,所述变体PNPLA3 cDNA的一部分在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子。在一些实施方案中,所述测定可以包括使用包含与变体PNPLA3cDNA的一部分或与其相邻的部分互补和/或杂交或特异性杂交的核酸序列的改变特异性探针或改变特异性引物,所述变体PNPLA3cDNA的一部分在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。

[0219] 在一些实施方案中,所述测定可以包括使用包含与HSD17B13基因组DNA的一部分或与其相邻的部分互补和/或杂交或特异性杂交的核酸序列的改变特异性探针或改变特异性引物,所述HSD17B13基因组DNA的一部分在对应于根据SEQ ID NO:1的位置12,667的位置处包含腺嘌呤。在一些实施方案中,所述测定可以包括使用包含与HSD17B13基因组DNA的一部分或与其相邻的部分互补和/或杂交或特异性杂交的核酸序列的改变特异性探针或改变特异性引物,所述HSD17B13基因组DNA的一部分在对应于根据SEQ ID NO:2的位置12,667的位置处包含胸腺嘧啶。

[0220] 在一些实施方案中,所述测定包括:对来自受试者的生物样品中存在的本文所述核酸分子的至少一部分进行测序,其中测序的部分包括本文公开的位置。例如,测序的部分可以是在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸,或在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸的PNPLA3核酸序列的一部分。在一些实施方案中,测序的部分可以是在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子的变体PNPLA3基因组DNA的部分。在一些实施方案中,测序的部分可以是在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子的变体PNPLA3 mRNA的部分。在一些实施方案中,测序的部分可以是在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子的变体PNPLA3 mRNA的部分。在一些实施方案中,测序的部分可以是在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子的变体PNPLA3 cDNA的部分。在一些实施方案中,测序的部分可以是在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子的变体PNPLA3 cDNA的部分。

[0221] 在一些实施方案中,测序的部分可以是在对应于根据SEQ ID NO:1的位置12,667的位置处包含腺嘌呤的HSD17B13基因组DNA的部分。在一些实施方案中,测序的部分可以是在对应于根据SEQ ID NO:2的位置12,667的位置处包含胸腺嘧啶的HSD17B13基因组DNA的部分。

[0222] 在一些实施方案中,所述测定包括:a) 将生物样品与引物(或改变特异性引物)接触,所述引物杂交至与本文鉴定的核酸分子的部分相邻(例如,与在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸,或在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸的PNPLA3核酸序列的一部分相邻;与在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子的变体PNPLA3基因组DNA的一部分相邻;与在对应于根据SEQ

ID NO:34的位置442至444的位置处包含AUG密码子的变体PNPLA3 mRNA的一部分相邻;与在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子的变体PNPLA3 mRNA的一部分相邻;与在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子的变体PNPLA3 cDNA的一部分相邻;与在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子的变体PNPLA3cDNA的一部分相邻;与在对应于根据SEQ ID NO:1的位置12,667的位置处包含腺嘌呤的HSD17B13基因组DNA的一部分相邻;或与在对应于根据SEQ ID NO:2的位置12,667的位置处包含胸腺嘧啶的HSD17B13基因组DNA的一部分相邻)的区域;b) 将引物至少延伸至核酸分子中对应于超出改变位点(例如,在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸,或在对应于根据SEQ ID NO:43的位置144处的位置处包含甲硫氨酸的PNPLA3核酸序列的部分;在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子的变体PNPLA3基因组DNA的部分;在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子的变体PNPLA3 mRNA的部分;在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子的变体PNPLA3 mRNA的部分;在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子的变体PNPLA3 cDNA的部分;在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子的变体PNPLA3 cDNA的部分;在对应于根据SEQ ID NO:1的位置12,667的位置处包含腺嘌呤的HSD17B13基因组DNA的部分;或在对应于SEQ ID NO:2的位置12,667处包含胸腺嘧啶的HSD17B13基因组DNA的部分)的核苷酸位置的位置;和c) 确定引物的延伸产物是否包含本文所述的变体或野生型PNPLA3或HSD17B13核酸分子中的任一种的核酸序列。

[0223] 在一些实施方案中,仅分析PNPLA3基因组DNA。在一些实施方案中,仅分析PNPLA3 mRNA。在一些实施方案中,仅分析从PNPLA3 mRNA获得的PNPLA3 cDNA。在一些实施方案中,仅分析HSD17B13基因组DNA。在一些实施方案中,仅分析HSD17B13mRNA。在一些实施方案中,仅分析从HSD17B13 mRNA获得的HSD17B13 cDNA。在一些实施方案中,仅分析HSD17B13 RNA转录物。在一些实施方案中,仅分析从HSD17B13 RNA转录物获得的HSD17B13 cDNA。

[0224] 在一些实施方案中,所述测定包括将生物样品与引物或探针接触,并确定是否发生了杂交,所述引物或探针在严格条件下与本文所述的特定变体PNPLA3核酸分子或变体HSD17B13核酸分子中的任一种(例如,变体基因组DNA分子、mRNA分子、cDNA分子、RNA转录物或cDNA转录物中的任一种)特异性杂交,并且不与对应的功能性核酸分子杂交。

[0225] 在一些实施方案中,所述测定包括将生物样品与引物或探针接触,并确定是否发生了杂交,所述引物或探针在严格条件下分别与本文所述的特定变体PNPLA3核酸分子中的任一种(例如,变体基因组DNA分子、mRNA分子、cDNA分子、RNA转录物或cDNA转录物中的任一种)或编码功能性HSD17B13蛋白的核酸分子中的任一种(例如,编码功能性HSD17B13蛋白的基因组DNA分子、mRNA分子、cDNA分子、RNA转录物或cDNA转录物中的任一种)特异性杂交,并且不与编码野生型PNPLA3或变体HSD17B13的对应核酸分子杂交。

[0226] 在一些实施方案中,所述测定包括RNA测序(RNA-Seq)。在一些实施方案中,所述测定还包括通过逆转录酶聚合酶链反应(RT-PCR)将mRNA逆转录成cDNA。

[0227] 此类探针和引物可以在高严格杂交条件下与靶序列特异性杂交。探针和引物可以与靶序列具有连续核苷酸的完全核酸序列同一性,但可以通过常规方法设计与靶核酸序列不同并保留特异性检测和/或鉴定靶核酸序列的能力的探针。因此,探针和引物可以与靶核

酸分子共享至少约80%、至少约85%、至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或100%序列同一性或互补性。

[0228] 当探针在允许探针与核酸分子结合的条件下与生物样品中的核酸分子杂交时,这种结合可以被检测到,并允许指示生物样品中特定变体或野生型PNPLA3或变体或功能性HSD17B13基因座的存在,或者特定变体或野生型PNPLA3或变体或功能性HSD17B13 mRNA或cDNA的存在或水平。已描述了对已结合探针的这种鉴定。特异性探针可以包含与变体或野生型PNPLA3或者变体或功能性HSD17B13基因的特定区域至少约80%、约80%至约85%、约85%至约90%、约90%至约95%以及约95%至约100%相同(或互补)的序列。特异性探针可以包含与变体或野生型PNPLA3或者变体或功能性HSD17B13 mRNA的特定区域至少约80%、约80%至约85%、约85%至约90%、约90%至约95%以及约95%至约100%相同(或互补)的序列。特异性探针可以包含与变体或野生型PNPLA3或者变体或功能性HSD17B13 cDNA的特定区域至少约80%、约80%至约85%、约85%至约90%、约90%至约95%以及约95%至约100%相同(或互补)的序列。

[0229] 在一些实施方案中,为了确定生物样品的特定核酸互补体是否包含编码特定功能性或变体PNPLA3蛋白或HSD17B13蛋白的核酸序列,可以使用引物对对生物样品进行核酸扩增方法,所述引物对包括来源于与编码感兴趣位点的位置(例如,本文所述的位置中的任一个)相邻的5'侧翼序列的第一引物,和来源于与编码相同感兴趣位点的位置相邻的3'侧翼序列的第二引物,以产生诊断特定功能性或变体PNPLA3蛋白或HSD17B13蛋白的存在的扩增子。例如,关于PNPLA3,扩增子可以包含编码对应于根据SEQ ID NO:42的位置148的位置的核苷酸序列。关于HSD17B13,扩增子可以包含对应于根据SEQ ID NO:31的位置5107至5109的核苷酸序列。在一些实施方案中,扩增子的长度范围可以从引物对加上一个核苷酸碱基对的组合长度到可由DNA扩增方案产生的任何扩增子长度。这个距离的范围可以从一个核苷酸碱基对直到扩增反应的极限,或约两万个核苷酸碱基对。任选地,引物对位于区域的侧翼,所述区域包括编码感兴趣位点的位置和在编码感兴趣位点的位置每一侧的至少1、2、3、4、5、6、7、8、9、10或更多个核苷酸。类似的扩增子可以从mRNA和/或cDNA序列产生。

[0230] 用于制备和使用探针和引物的代表性方法描述于例如Molecular Cloning:A Laboratory Manual,第2版,第1-3卷,ed.Sambrook等人,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.1989(下文中,“Sambrook等人,1989”);Current Protocols in Molecular Biology,ed.Ausubel等人,Green Publishing and Wiley-Interscience,New York,1992(定期更新)(下文中,“Ausubel等人,1992”);和Innis等人,PCR Protocols:A Guide to Methods and Applications,Academic Press:San Diego,1990)中。PCR引物对可以来源于已知的序列,例如,通过使用计划用于该目的的计算机程序,如以下中的PCR引物分析工具:Vector NTI第10版(Informax公司,马里兰州贝塞斯达);PrimerSelect(DNASTAR公司,麦迪逊市,威斯康辛州);和Primer3(0.4.0.COPYRGT.版,1991,Whitehead Institute for Biomedical Research,马萨诸塞州剑桥市)。另外,可以使用指南目视扫描序列并手动地鉴定引物。

[0231] 任何核酸杂交或扩增或测序方法均可以用于特异性检测功能性或变体PNPLA3或HSD17B13基因座的存在和/或功能性或变体PNPLA3或HSD17B13 mRNA或由mRNA产生的cDNA

的水平。在一些实施方案中,核酸分子可以用作引物来扩增功能性或变体PNPLA3或HSD17B13核酸的区域,或者所述核酸分子可以用作探针,所述探针例如在严格条件下与包含功能性或变体PNPLA3或HSD17B13基因座的核酸分子或者包含功能性或变体PNPLA3或HSD17B13 mRNA或由mRNA产生的cDNA的核酸分子特异性杂交。

[0232] 本领域有多种技术可用,包括例如核酸测序、核酸杂交和核酸扩增。核酸测序技术的说明性实例包括但不限于链终止子(Sanger)测序和染料终止子测序。

[0233] 其他方法包括除了测序之外的核酸杂交方法,包括使用针对纯化的DNA、扩增的DNA和固定细胞制剂的标记引物或探针(荧光原位杂交(FISH))。在一些方法中,靶核酸可以在检测之前或同时扩增。核酸扩增技术的说明性实例包括但不限于聚合酶链反应(PCR)、连接酶链反应(LCR)、链置换扩增(SDA)和基于核酸序列的扩增(NASBA)。其他方法包括但不限于连接酶链反应、链置换扩增和嗜热SDA(tSDA)。

[0234] 任何方法都可以用于检测未扩增或扩增的多核苷酸,包括例如杂交保护测定(HPA)、扩增过程的实时定量评价,以及确定样品中最初存在的靶序列的数量(但这不是基于实时扩增)。

[0235] 还提供了用于鉴定核酸的方法,所述方法不一定需要序列扩增,并且基于例如染色体材料的DNA(DNA:DNA)印迹杂交、原位杂交(ISH)和荧光原位杂交(FISH)的方法。DNA印迹可以用于检测特定的核酸序列。在此类方法中,将从样品提取的核酸片段化、在基质凝胶上电泳分离,并转移到膜过滤器。用感兴趣的序列互补的标记探针对使过滤器结合的核酸进行杂交。检测与过滤器结合的杂交探针。在任何此类方法中,所述方法可以包括使用本文描述或举例说明的探针中的任一种进行杂交。

[0236] 在杂交技术中,可以采用严格的条件,使得探针或引物将与其靶标特异性杂交。在一些实施方案中,在严格条件下,多核苷酸引物或探针将以比其他序列(例如,对应的功能性或变体PNPLA3或HSD17B13基因座、功能性或变体PNPLA3或HSD17B13 mRNA,或者功能性或变体PNPLA3或HSD17B13 cDNA)可检测更高的程度与其靶序列(例如,功能性或变体PNPLA3或HSD17B13基因座、功能性或变体PNPLA3或HSD17B13 mRNA,或者功能性或变体PNPLA3或HSD17B13 cDNA)杂交,如超过背景至少2倍、至少3倍、至少4倍或更高,包括超过背景10倍多。在一些实施方案中,在严格条件下,多核苷酸引物或探针将以比其他序列高至少2倍的可检测更高的程度与其靶序列杂交。在一些实施方案中,在严格条件下,多核苷酸引物或探针将以比其他序列高至少3倍的可检测更高的程度与其靶序列杂交。在一些实施方案中,在严格条件下,多核苷酸引物或探针将以比其他序列高至少4倍的可检测更高的程度与其靶序列杂交。在一些实施方案中,在严格条件下,多核苷酸引物或探针将以比其他序列高超过背景10倍多的可检测更高程度与其靶序列杂交。严格条件是序列依赖性的并且在不同情况下将有所不同。通过控制杂交和/或洗涤条件的严格性,可以鉴定与探针100%互补的靶序列(同源探测)。可选地,可以调整严格条件以允许序列中的一些错配,使得检测到较低程度的同一性(异源探测)。

[0237] 在一些实施方案中,检测步骤包括:扩增编码感兴趣位点(例如,本文所述的位置中的任一个)的核酸分子的至少一部分;用可检测标记标记核酸分子;将标记的核酸与包含探针的支持物接触,其中探针包含在严格条件下与编码感兴趣位点(例如,本文所述的位置中的任一个)的核酸序列杂交的核酸序列;并检测可检测标记。

[0238] 在一些实施方案中,检测步骤包括:扩增编码PNPLA3或HSD17B13蛋白的核酸分子的至少一部分,其中扩增的核酸分子编码包含感兴趣位点(例如,本文所述的位置中的任一个)的氨基酸序列;用可检测标记标记核酸分子;将标记的核酸与包含探针的支持物接触,其中探针包含在严格条件下与编码感兴趣位点(例如,本文所述的位置中的任一个)的核酸序列杂交的核酸序列;并检测可检测标记。本文公开的核酸分子中的任一种都可以被扩增。例如,本文公开的基因组DNA、cDNA或mRNA分子中的任一种都可以被扩增。在一些实施方案中,核酸分子是mRNA,并且所述方法还包括在扩增步骤之前将mRNA逆转录成cDNA。

[0239] 在一些实施方案中,检测步骤包括:将编码PNPLA3或HSD17B13蛋白的核酸分子与包含可检测标记的探针接触,其中探针包含在严格条件下与编码变体PNPLA3或HSD17B13蛋白的核酸序列杂交的核酸序列,并检测可检测标记。在一些实施方案中,检测步骤包括:将编码PNPLA3或HSD17B13蛋白的核酸分子与包含可检测标记的探针接触,其中探针包含在严格条件下与编码感兴趣位点(例如,本文所述的位置中的任一个)的核酸序列杂交的核酸序列,并检测可检测标记。在一些实施方案中,核酸分子存在于从人受试者获得的细胞中,使得检测是根据原位杂交技术进行的。

[0240] 可以用于本文公开方法中的其他测定包括例如,逆转录聚合酶链反应(RT-PCR)或定量RT-PCR(qRT-PCR)。可以用于本文公开方法中的仍然其他测定包括例如,RNA测序(RNA-Seq),随后检测生物样品中变体mRNA或cDNA的存在和数量。

[0241] 在一些实施方案中,检测步骤包括扩增编码特定PNPLA3或HSD17B13蛋白的核酸分子的至少一部分,用可检测标记标记扩增的核酸分子,将标记的核酸分子与包含探针的支持物接触,其中探针包含包括例如在严格条件下与编码特定PNPLA3或HSD17B13蛋白的核酸序列特异性杂交的核酸序列,并检测可检测标记。在一些实施方案中,检测步骤包括扩增编码特定PNPLA3或HSD17B13蛋白的核酸分子的至少一部分,用可检测标记标记扩增的核酸分子,将标记的核酸分子与包含探针的支持物接触,其中探针包含包括例如在严格条件下与编码感兴趣位点(例如,在对应于根据SEQ ID NO:42的位置148的位置处编码甲硫氨酸或在对应于根据SEQ ID NO:43的位置144的位置处编码甲硫氨酸的PNPLA3核酸序列的部分;在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子的变体PNPLA3基因组DNA的部分;在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子的变体PNPLA3 mRNA的部分;在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子的变体PNPLA3 mRNA的部分;在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子的变体PNPLA3cDNA的部分;在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子的变体PNPLA3 cDNA的部分;在对应于根据SEQ ID NO:1的位置12,667的位置处包含腺嘌呤的HSD17B13基因组DNA的部分;或在对应于根据SEQ ID NO:2的位置12,667的位置处包含胸腺嘧啶的HSD17B13基因组DNA的部分)的核酸序列特异性杂交的核酸序列,并检测可检测标记。如果核酸包括mRNA,则所述方法还可以包括在扩增步骤之前将mRNA逆转录成cDNA。在一些实施方案中,确定步骤包括将编码特定PNPLA3或HSD17B13蛋白的核酸分子与包含可检测标记的探针接触,并检测可检测标记。

[0242] 本公开提供了用于将人受试者鉴定为用于通过抑制羟基类固醇17-β脱氢酶13(HSD17B13)治疗或抑制肝病的候选者的方法,所述方法包括确定来自所述受试者的样品是否包含编码包含I148M变异的含patatin样磷脂酶结构域3(PNPLA3)蛋白的第一核酸和编码

功能性HSD17B13蛋白的第二核酸,和/或包含I148M变异的PNPLA3蛋白和功能性HSD17B13蛋白,以及当检测到第一和第二核酸二者和/或检测到两种蛋白质时,将受试者鉴定为用于通过抑制HSD17B13治疗或抑制肝病的候选者。在一些实施方案中,受试者是肥胖的。在一些实施方案中,受试者患有脂肪肝。在一些实施方案中,第一核酸分子包含基因组DNA。在一些实施方案中,基因组DNA在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子。在一些实施方案中,基因组DNA包含根据SEQ ID NO:31的核苷酸序列。在一些实施方案中,第一核酸分子包含mRNA。在一些实施方案中,mRNA在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:34的核苷酸序列。在一些实施方案中,mRNA在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:35的核苷酸序列。在一些实施方案中,第一核酸分子包含从mRNA获得的cDNA。在一些实施方案中,cDNA在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:38的核苷酸序列。在一些实施方案中,cDNA在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:39的核苷酸序列。在一些实施方案中,检测第一核酸包括对第一核酸的至少一部分进行测序,并且所述部分包含编码I148M变异的密码子。在一些实施方案中,检测第一核酸包括将第一核酸与探针或引物杂交,所述探针或引物与第一核酸的一部分特异性杂交,其中所述部分包含编码I148M变异的密码子。在一些实施方案中,探针或引物是等位基因特异性探针或引物。在一些实施方案中,探针或引物包含标记。在一些实施方案中,所述方法还包括确定受试者对于I148M变异是纯合的还是杂合的。

[0243] 在一些实施方案中,第二核酸包含基因组DNA。在一些实施方案中,基因组DNA在对应于根据SEQ ID NO:1的位置12,667的位置处包含腺嘌呤。在一些实施方案中,基因组DNA包含SEQ ID NO:1。在一些实施方案中,第二核酸分子包含mRNA。在一些实施方案中,mRNA包含SEQ ID NO:3。在一些实施方案中,mRNA包含SEQ ID NO:4。在一些实施方案中,mRNA包含SEQ ID NO:7。在一些实施方案中,mRNA包含SEQ ID NO:11。在一些实施方案中,第二核酸分子包含从mRNA获得的cDNA。在一些实施方案中,cDNA包含SEQ ID NO:12。在一些实施方案中,cDNA包含SEQ ID NO:13。在一些实施方案中,cDNA包含SEQ ID NO:16。在一些实施方案中,cDNA包含SEQ ID NO:20。在一些实施方案中,检测第二核酸包括对第二核酸进行测序。在一些实施方案中,检测第二核酸包括将第二核酸与探针或引物杂交,所述探针或引物与第二核酸特异性杂交。在一些实施方案中,探针或引物是等位基因特异性探针或引物。在一些实施方案中,探针或引物包含标记。在一些实施方案中,所述方法还包括确定受试者对于样品中编码功能性HSD17B13蛋白的第二核酸是纯合的还是杂合的。

[0244] 本公开提供了鉴定作为HSD17B13抑制的候选者的受试者的方法,所述方法包括确定来自受试者的样品是否包含编码PNPLA3I1e148Met变体或PNPLA3 I1e144Met变体的核酸。本公开还提供了用于将人受试者鉴定为用于通过抑制HSD17B13治疗或抑制肝病的候选者的方法,所述方法包括确定来自所述受试者的样品是否包含编码包含I148M变异的PNPLA3蛋白的第一核酸和编码功能性HSD17B13蛋白的第二核酸,和/或包含I148M变异的PNPLA3蛋白和功能性HSD17B13蛋白,以及当检测到第一和第二核酸二者和/或检测到两种蛋白质时,将受试者鉴定为用于通过抑制HSD17B13治疗或抑制肝病的候选者。

[0245] 本公开还提供了对作为HSD17B13抑制的候选者的受试者进行分类的方法,所述方法包括确定来自受试者的样品是否包含编码PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体的核酸。本公开还提供了用于将人受试者分类为用于通过抑制HSD17B13治疗或抑制肝病的候选者的方法,所述方法包括确定来自所述受试者的样品是否包含编码包含I148M变异的PNPLA3蛋白的第一核酸和编码功能性HSD17B13蛋白的第二核酸,和/或包含I148M变异的PNPLA3蛋白和功能性HSD17B13蛋白,以及当检测到第一和第二核酸二者和/或检测到两种蛋白质时,将受试者分类为用于通过抑制HSD17B13治疗或抑制肝病的候选者。

[0246] 变体PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体可以是本文所述的变体PNPLA3 Ile148Met变体和PNPLA3 Ile144Met变体中的任一种。变体PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体可以通过本文所述的方法中的任一种来检测。在一些实施方案中,所述方法还包括确定受试者对于变体PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体是纯合的还是杂合的。在一些实施方案中,受试者对于变体PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体是纯合的。在一些实施方案中,受试者对于变体PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体是杂合的。在一些实施方案中,受试者对于变体PNPLA3 Ile148Met变体是纯合的。在一些实施方案中,受试者对于变体PNPLA3 Ile148Met变体是杂合的。在一些实施方案中,受试者对于变体PNPLA3 Ile144Met变体是纯合的。在一些实施方案中,受试者对于变体PNPLA3 Ile144Met变体是杂合的。

[0247] 在优选的实施方案中,受试者不包含编码HSD17B13蛋白中的功能丧失变异的任何基因。据信,HSD17B13蛋白中的功能丧失变异(包括本文所述或举例说明的那些)赋予了肝病保护效应,并且还据信,这种保护效应在变体PNPLA3 Ile148M变异的存在下得到增强。因此,据信其中编码HSD17B13蛋白的基因的两个拷贝(来自每条染色体)均编码功能丧失变异的受试者(例如,包含PNPLA3中的I148M变异的受试者)不太可能受益于HSD17B13抑制疗法。然而,据信表达至少部分功能性HSD17B13蛋白的受试者将受益于HSD17B13抑制疗法。因此,所述方法可以包括对编码HSD17B13的基因(在一条或两条染色体中)的状态进行分类,包括所述基因是否编码HSD17B13蛋白中的功能丧失变异,以及受试者是纯合的还是杂合的。

[0248] 在一些实施方案中,所述方法还包括检测来自受试者的样品中编码功能性HSD17B13蛋白的核酸分子或基因的存在。核酸分子可以编码本文所述的功能性HSD17B13蛋白中的任一种。HSD17B13核酸分子可以通过本文所述的方法中的任一种来检测。在一些实施方案中,所述方法还包括确定受试者对于编码功能性HSD17B13蛋白的基因是纯合的还是杂合的。在一些实施方案中,受试者对于编码功能性HSD17B13蛋白的基因是纯合的。在一些实施方案中,受试者对于编码功能性HSD17B13的基因是杂合的。

[0249] 本公开还提供了包括支持物,所述支持物包含与本文公开的任何一种或多种探针连接的基质。固体支持物是分子(如本文公开的探针中的任一种)可以与之缔合的固态基质或支持物。固体支持物的一种形式是阵列。固体支持物的另一种形式是阵列检测器。阵列检测器是多个不同的探针以阵列、网格或其他有组织的模式与其耦合的固体支持物。

[0250] 用于固体支持物的固态基质可以包括可以与分子耦合的任何固体材料。这包括如以下的材料:丙烯酰胺、琼脂糖、纤维素、硝化纤维素、玻璃、聚苯乙烯、聚乙烯醋酸乙烯酯、聚丙烯、聚甲基丙烯酸酯、聚乙烯、聚氧化乙烯、聚硅酸盐、聚碳酸酯、特氟隆、碳氟化合物、尼龙、硅橡胶、聚酐、聚乙醇酸、聚乳酸、聚原酸酯、聚丙烯基延胡索酸酯、胶原、糖胺聚糖和聚

氨基酸。固态基质可以具有任何有用的形式,包括薄膜、膜、瓶子、盘子、纤维、编织纤维、成型聚合物、颗粒、珠子、微粒或其组合。固态基质和固体支持物可以是多孔的或无孔的。固态基质的形式是微量滴定盘,如标准96孔类型。在一些实施方案中,可以采用多孔载玻片,通常每个孔含有一个阵列。这一特征允许更好地控制测定重现性、提高处理量和样品处理,并易于自动化。在一些实施方案中,支持物是微阵列。

[0251] 在一些实施方案中,所述方法还包括确定受试者是否是肥胖的。在一些实施方案中,如果受试者的体重指数(BMI)超过 $30\text{kg}/\text{m}^2$,则所述受试者是肥胖的。肥胖可能是患有肝病或处于患上肝病的风险中的受试者的特征。在一些实施方案中,所述方法还包括确定受试者是否患有脂肪肝。脂肪肝可能是患有肝病或处于患上肝病的风险中的受试者的特征。在一些实施方案中,所述方法还包括确定受试者是否是肥胖的和是否患有脂肪肝。

[0252] 在一些实施方案中,所述方法还包括向受试者施用HSD17B13的抑制剂。本文描述了向受试者施用HSD17B13的抑制剂的方法。

[0253] 在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白或PNPLA3Ile144Met蛋白的核酸分子是本文所述的核酸分子中的任一种。在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白或PNPLA3Ile144Met蛋白的核酸分子是基因组DNA。在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白的基因组DNA是本文所述的基因组DNA分子中的任一种。在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白的核酸分子是基因组DNA。在一些实施方案中,基因组DNA在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子。在一些实施方案中,基因组DNA包含根据SEQ ID NO:31的核苷酸序列。在一些实施方案中,编码变体PNPLA3蛋白的基因组DNA通过核酸测序或探针杂交来检测,如本文所述。在一些实施方案中,包含ATG密码子的编码变体PNPLA3蛋白的基因组DNA通过核酸测序或探针杂交来检测,如本文所述。

[0254] 在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白或PNPLA3Ile144Met蛋白的核酸分子是mRNA。在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白的mRNA是本文所述的mRNA分子中的任一种。在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白的核酸分子是mRNA。在一些实施方案中,mRNA在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:34的核苷酸序列。在一些实施方案中,mRNA在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:35的核苷酸序列。在一些实施方案中,编码变体PNPLA3蛋白的mRNA通过核酸测序或探针杂交来检测,如本文所述。在一些实施方案中,包含AUG密码子的编码变体PNPLA3蛋白的mRNA通过核酸测序或探针杂交来鉴定,如本文所述。

[0255] 在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白或PNPLA3Ile144Met蛋白的核酸分子是cDNA。在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白的cDNA是本文所述的cDNA分子中的任一种。在一些实施方案中,编码PNPLA3Ile148Met蛋白的核酸分子是cDNA。在一些实施方案中,cDNA在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:38的核苷酸序列。在一些实施方案中,cDNA在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:39的核苷酸序列。

在一些实施方案中,编码变体PNPLA3蛋白的cDNA通过核酸测序或探针杂交来鉴定。在一些实施方案中,包含ATG密码子的编码变体PNPLA3蛋白的cDNA通过核酸测序或探针杂交来鉴定,如本文所述。

[0256] 在一些实施方案中,编码功能性HSD17B13蛋白的核酸分子是本文所述的核酸分子中的任一种。在一些实施方案中,编码功能性HSD17B13蛋白的核酸分子是基因组DNA。在一些实施方案中,编码功能性HSD17B13蛋白的基因组DNA是本文所述的基因组DNA分子中的任一种。在一些实施方案中,基因组DNA在对应于根据SEQ ID NO:1的位置12,667的位置处包含腺嘌呤。在一些实施方案中,基因组DNA包含SEQ ID NO:1。在一些实施方案中,功能性HSD17B13基因组DNA的存在通过核酸测序或探针杂交来确定,如本文所述。

[0257] 在一些实施方案中,编码功能性HSD17B13蛋白的核酸分子是mRNA。在一些实施方案中,编码功能性HSD17B13蛋白的mRNA是本文所述的mRNA分子中的任一种。在一些实施方案中,功能性HSD17B13核酸分子是mRNA。在一些实施方案中,mRNA包含SEQ ID NO:3。在一些实施方案中,mRNA包含SEQ ID NO:4。在一些实施方案中,mRNA包含SEQ ID NO:7。在一些实施方案中,mRNA包含SEQ ID NO:11。在一些实施方案中,功能性HSD17B13 mRNA的存在通过核酸测序或探针杂交来确定,如本文所述。

[0258] 在一些实施方案中,编码功能性HSD17B13蛋白的核酸分子是cDNA。在一些实施方案中,编码功能性HSD17B13蛋白的cDNA是本文所述的cDNA分子中的任一种。在一些实施方案中,cDNA包含SEQ ID NO:12。在一些实施方案中,cDNA包含SEQ ID NO:13。在一些实施方案中,cDNA包含SEQ ID NO:16。在一些实施方案中,cDNA包含SEQ ID NO:20。在一些实施方案中,功能性HSD17B13cDNA的存在通过核酸测序或探针杂交来确定,如本文所述。

[0259] 在一些实施方案中,所述方法还包括从受试者获得样品。在一些实施方案中,作为HSD17B13抑制的候选者的受试者患有肝病或易患肝病。在一些实施方案中,所述肝病是慢性肝病。在一些优选的实施方案中,慢性肝病是非酒精性脂肪性肝病 (NAFLD)、酒精性肝病 (ALD)、非酒精性脂肪性肝炎 (NASH)、肝硬化、脂肪变性或肝细胞癌。在一些优选的实施方案中,慢性肝病是非酒精性脂肪性肝病 (NAFLD)、酒精性肝病 (ALD)、非酒精性脂肪性肝炎 (NASH)、肝硬化或脂肪变性。在一些实施方案中,所述肝病是酒精性肝病。在一些实施方案中,酒精性肝病包括由饮酒引起的肝硬化、脂肪变性或肝细胞癌中的一种或多种。在一些实施方案中,所述肝病是非酒精性肝病。在一些实施方案中,非酒精性肝病包括非酒精性脂肪性肝病 (NAFLD) 或非酒精性脂肪性肝炎 (NASH)。在一些实施方案中,非酒精性肝病包括并非由饮酒导致的肝硬化、脂肪变性或肝细胞癌中的一种或多种。

[0260] 本公开还提供了检测受试者中的PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体和功能性HSD17B13的方法,所述方法包括:检测来自受试者的样品中PNPLA3 Ile148Met蛋白或编码PNPLA3Ile148Met蛋白的核酸分子的存在,或者检测来自受试者的样品中PNPLA3 Ile144Met蛋白或编码PNPLA3 Ile144Met蛋白的核酸分子的存在;以及检测来自受试者的样品中功能性HSD17B13蛋白或编码功能性HSD17B13蛋白的核酸分子的存在。变体PNPLA3 Ile148Met变体蛋白或核酸分子可以是本文所述的变体PNPLA3 Ile148Met变体蛋白或核酸分子中的任一种。变体PNPLA3 Ile144Met变体蛋白或核酸分子可以是本文所述的变体PNPLA3 Ile144Met变体蛋白或核酸分子中的任一种。功能性HSD17B13蛋白或核酸分子可以是本文所述的功能性HSD17B13蛋白或核酸分子中的任一种。

[0261] 在一些实施方案中,所述方法还包括确定受试者对于变体PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体是纯合的还是杂合的。在一些实施方案中,受试者对于变体PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体是纯合的。在一些实施方案中,受试者对于变体PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体是杂合的。在一些实施方案中,受试者对于变体PNPLA3 Ile148Met变体是纯合的。在一些实施方案中,受试者对于变体PNPLA3 Ile148Met变体是杂合的。在一些实施方案中,受试者对于变体PNPLA3 Ile144Met变体是纯合的。在一些实施方案中,受试者对于变体PNPLA3 Ile144Met变体是杂合的。

[0262] 在一些实施方案中,所述方法还包括确定受试者对于功能性HSD17B13是纯合的还是杂合的。在一些实施方案中,受试者对于功能性HSD17B13是纯合的。在一些实施方案中,受试者对于功能性HSD17B13是杂合的。

[0263] 在一些实施方案中,在样品中检测功能性HSD17B13蛋白的存在。功能性HSD17B13蛋白可以是本文所述的功能性HSD17B13蛋白中的任一种。在一些实施方案中,功能性HSD17B13蛋白包含根据SEQ ID NO:21、SEQ ID NO:22、SEQ ID NO:25或SEQ ID NO:29的氨基酸序列。在一些实施方案中,功能性HSD17B13蛋白通过氨基酸测序或免疫测定来检测,如本文所述。

[0264] 在一些实施方案中,在样品中检测功能性HSD17B13核酸分子的存在。功能性HSD17B13核酸分子可以是本文所述的功能性HSD17B13核酸分子中的任一种。在一些实施方案中,功能性HSD17B13核酸分子是基因组DNA。功能性HSD17B13基因组DNA分子可以是本文所述的功能性HSD17B13基因组DNA分子中的任一种。在一些实施方案中,基因组DNA在对应于根据SEQ ID NO:1的位置12,667的位置处包含腺嘌呤。在一些实施方案中,基因组DNA包含SEQ ID NO:1。在一些实施方案中,基因组DNA通过核酸测序或探针杂交来检测,如本文所述。

[0265] 在一些实施方案中,功能性HSD17B13核酸分子是mRNA。功能性HSD17B13 mRNA分子可以是本文所述的功能性HSD17B13mRNA分子中的任一种。在一些实施方案中,mRNA包含SEQ ID NO:3。在一些实施方案中,mRNA包含SEQ ID NO:4。在一些实施方案中,mRNA包含SEQ ID NO:7。在一些实施方案中,mRNA包含SEQ ID NO:11。在一些实施方案中,mRNA通过核酸测序或探针杂交来检测,如本文所述。

[0266] 在一些实施方案中,功能性HSD17B13核酸分子是cDNA。功能性HSD17B13 cDNA分子可以是本文所述的功能性HSD17B13 cDNA分子中的任一种。在一些实施方案中,cDNA包含SEQ ID NO:12。在一些实施方案中,cDNA包含SEQ ID NO:13。在一些实施方案中,cDNA包含SEQ ID NO:16。在一些实施方案中,cDNA包含SEQ ID NO:20。在一些实施方案中,cDNA通过核酸测序或探针杂交来检测,如本文所述。

[0267] 在一些实施方案中,在样品中检测PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白的存在。变体PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白可以是本文所述的变体PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白中的任一种。在一些实施方案中,变体PNPLA3 Ile148Met蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸,或在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3 Ile148Met蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸。在一些实施方案中,样品中的变体PNPLA3蛋白包含根据SEQ ID NO:42的氨基酸序列。在一些

实施方案中,变体PNPLA3 Ile148Met蛋白在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,样品中的变体PNPLA3蛋白包含根据SEQ ID NO:43的氨基酸序列。在一些实施方案中,变体PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白通过氨基酸测序或免疫测定来检测,如本文所述。

[0268] 在一些实施方案中,在样品中检测编码PNPLA3 Ile148Met蛋白或PNPLA3 Ile148Met蛋白的核酸分子的存在。编码PNPLA3Ile148Met蛋白或PNPLA3 Ile148Met蛋白的核酸分子可以是本文所述的编码PNPLA3 Ile148Met蛋白和PNPLA3 Ile148Met蛋白的核酸分子中的任一种。

[0269] 在一些实施方案中,编码PNPLA3 Ile148Met蛋白或PNPLA3Ile144Met蛋白的核酸分子是基因组DNA。编码PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白的基因组DNA可以是本文所述的基因组DNA分子中的任一种。在一些实施方案中,基因组DNA在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子。在一些实施方案中,基因组DNA包含根据SEQ ID NO:31的核苷酸序列。在一些实施方案中,包含ATG密码子的编码变体PNPLA3蛋白的基因组DNA通过核酸测序或探针杂交来检测,如本文所述。

[0270] 在一些实施方案中,编码PNPLA3 Ile148Met蛋白或PNPLA3Ile148Met蛋白的核酸分子是mRNA。编码PNPLA3 Ile148Met蛋白或PNPLA3 Ile148Met蛋白的mRNA分子可以是本文所述的mRNA分子中的任一种。在一些实施方案中,mRNA在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:34的核苷酸序列。在一些实施方案中,mRNA在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:35的核苷酸序列。在一些实施方案中,包含AUG密码子的编码变体PNPLA3蛋白的mRNA通过核酸测序或探针杂交来鉴定,如本文所述。

[0271] 在一些实施方案中,编码PNPLA3 Ile148Met蛋白或PNPLA3Ile144Met蛋白的核酸分子是cDNA。编码PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白的cDNA可以是本文所述的cDNA分子中的任一种。在一些实施方案中,cDNA在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:38的核苷酸序列。在一些实施方案中,cDNA在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:39的核苷酸序列。在一些实施方案中,包含ATG密码子的编码变体PNPLA3蛋白的cDNA通过核酸测序或探针杂交来鉴定,如本文所述。

[0272] 在一些实施方案中,所述方法还包括从受试者获得样品。

[0273] 本公开还提供了鉴定具有针对肝病的保护效应的受试者的方法,所述方法包括:检测来自受试者的样品中PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体的存在;并且检测来自受试者的样品中HSD17B13功能丧失变体的存在。本公开还提供了对具有针对肝病的保护效应的受试者进行分类的方法,所述方法包括:检测来自受试者的样品中PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体的存在;并且检测来自受试者的样品中HSD17B13功能丧失变体的存在。变体PNPLA3 Ile148Met变体和PNPLA3 Ile144Met变体可以是本文所述的变体PNPLA3 Ile148Met变体和PNPLA3 Ile144Met变体中的任一种。HSD17B13功能丧失变体可以是本文所述的HSD17B13功能丧失变体中的任一种。

[0274] 在一些实施方案中,所述方法还包括确定受试者对于变体PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体是纯合的还是杂合的。在一些实施方案中,受试者对于变体PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体是纯合的。在一些实施方案中,受试者对于变体PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体是杂合的。在一些实施方案中,受试者对于变体PNPLA3 Ile148Met变体是纯合的。在一些实施方案中,受试者对于变体PNPLA3 Ile148Met变体是杂合的。在一些实施方案中,受试者对于变体PNPLA3 Ile144Met变体是纯合的。在一些实施方案中,受试者对于变体PNPLA3 Ile144Met变体是杂合的。

[0275] 在一些实施方案中,所述方法还包括确定受试者对于功能性HSD17B13是纯合的还是杂合的。在一些实施方案中,受试者对于功能性HSD17B13是纯合的。在一些实施方案中,受试者对于功能性HSD17B13是杂合的。

[0276] 在一些实施方案中,通过在来自受试者的样品中检测PNPLA3Ile148Met蛋白或编码PNPLA3 Ile148Met蛋白的核酸分子,或者在来自受试者的样品中检测PNPLA3 Ile144Met蛋白或编码PNPLA3Ile144Met蛋白的核酸分子来检测受试者中的变体PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体;并且通过在来自受试者的样品中检测HSD17B13功能丧失变体蛋白或编码HSD17B13功能丧失变体蛋白的核酸分子来检测受试者中的HSD17B13功能丧失。在一些实施方案中,通过在来自受试者的样品中检测PNPLA3 Ile148Met蛋白或编码PNPLA3 Ile148Met蛋白的核酸分子来检测受试者中的变体PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体;并且通过在来自受试者的样品中检测HSD17B13功能丧失变体蛋白或编码HSD17B13功能丧失变体蛋白的核酸分子来检测受试者中的HSD17B13功能丧失。

[0277] 在一些实施方案中,在样品中检测PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白的存在。变体PNPLA3 Ile148Met蛋白和PNPLA3 Ile144Met蛋白可以是本文所述的变体PNPLA3 Ile148Met蛋白和PNPLA3 Ile144Met蛋白中的任一种。在一些实施方案中,变体PNPLA3 Ile148Met蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸,或在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3 Ile148Met蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸。在一些实施方案中,样品中的变体PNPLA3蛋白包含根据SEQ ID NO:42的氨基酸序列。在一些实施方案中,变体PNPLA3 Ile148Met蛋白在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,样品中的变体PNPLA3蛋白包含根据SEQ ID NO:43的氨基酸序列。在一些实施方案中,变体PNPLA3 Ile148Met蛋白通过氨基酸测序或免疫测定来检测,如本文所述。

[0278] 在一些实施方案中,在样品中检测编码PNPLA3 Ile148Met蛋白或PNPLA3 Ile144Met蛋白的核酸分子的存在。编码变体PNPLA3Ile148Met蛋白或PNPLA3 Ile144Met蛋白的核酸分子可以是本文所述的编码变体PNPLA3 Ile148Met蛋白和PNPLA3 Ile144Met蛋白的核酸分子中的任一种。

[0279] 在一些实施方案中,其中编码PNPLA3 Ile148Met蛋白或PNPLA3Ile144Met蛋白的核酸分子是基因组DNA。编码变体PNPLA3Ile148Met蛋白或PNPLA3 Ile144Met蛋白的基因组DNA分子可以是本文所述的变体PNPLA3 Ile148Met蛋白和PNPLA3 Ile144Met蛋白中的任一种。在一些实施方案中,基因组DNA在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子。在一些实施方案中,基因组DNA包含根据SEQ ID NO:31的核苷酸序列。在一

些实施方案中,包含ATG密码子的编码变体PNPLA3蛋白的基因组DNA通过核酸测序或探针杂交来检测。

[0280] 在一些实施方案中,编码变体PNPLA3 Ile148Met蛋白或PNPLA3Ile144Met蛋白的核酸分子是mRNA。编码变体PNPLA3 Ile148Met蛋白和PNPLA3 Ile144Met蛋白的mRNA分子可以是本文所述的变体PNPLA3 Ile148Met蛋白和PNPLA3 Ile144Met蛋白中的任一种。在一些实施方案中,mRNA在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:34的核苷酸序列。在一些实施方案中,mRNA在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:35的核苷酸序列。在一些实施方案中,包含AUG密码子的编码变体PNPLA3蛋白的mRNA通过核酸测序或探针杂交来鉴定,如本文所述。

[0281] 在一些实施方案中,编码PNPLA3 Ile148Met蛋白或PNPLA3Ile148Met蛋白的核酸分子是cDNA。编码变体PNPLA3 Ile148Met蛋白和PNPLA3 Ile148Met蛋白的cDNA可以是本文所述的变体PNPLA3 Ile148Met蛋白和PNPLA3 Ile148Met蛋白中的任一种。在一些实施方案中,cDNA在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:38的核苷酸序列。在一些实施方案中,cDNA在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:39的核苷酸序列。在一些实施方案中,包含ATG密码子的编码变体PNPLA3蛋白的cDNA通过核酸测序或探针杂交来鉴定,如本文所述。

[0282] 在一些实施方案中,在样品中检测HSD17B13功能丧失变体蛋白的存在。HSD17B13功能丧失变体可以是本文所述的HSD17B13功能丧失变体蛋白中的任一种。在一些实施方案中,HSD17B13功能丧失变体蛋白包含根据SEQ ID NO:23的氨基酸序列。在一些实施方案中,HSD17B13功能丧失变体蛋白包含根据SEQ ID NO:24的氨基酸序列。在一些实施方案中,HSD17B13功能丧失变体蛋白包含根据SEQ ID NO:26的氨基酸序列。在一些实施方案中,HSD17B13功能丧失变体蛋白包含根据SEQ ID NO:27的氨基酸序列。在一些实施方案中,HSD17B13功能丧失变体蛋白包含根据SEQ ID NO:28的氨基酸序列。在一些实施方案中,HSD17B13功能丧失变体蛋白通过氨基酸测序或免疫测定来检测,如本文所述。

[0283] 在一些实施方案中,在样品中检测编码HSD17B13功能丧失变体蛋白的核酸分子的存在。编码HSD17B13功能丧失变体蛋白的核酸分子可以是本文所述的编码HSD17B13功能丧失变体蛋白的核酸分子中的任一种。

[0284] 在一些实施方案中,编码HSD17B13功能丧失变体蛋白的核酸分子是基因组DNA。编码HSD17B13功能丧失变体蛋白的基因组DNA分子可以是本文所述的HSD17B13功能丧失变体蛋白中的任一种。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的基因组DNA在对应于根据SEQ ID NO:2的位置12,667的位置处包含胸腺嘧啶。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的基因组DNA包含SEQ ID NO:2。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的基因组DNA通过核酸测序或探针杂交来检测,如本文所述。

[0285] 在一些实施方案中,编码HSD17B13功能丧失变体蛋白的核酸分子是mRNA。编码HSD17B13功能丧失变体蛋白的mRNA分子可以是本文所述的编码HSD17B13功能丧失变体蛋白的mRNA分子中的任一种。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的mRNA包含SEQ ID NO:5。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的mRNA包含SEQ ID

N0:6。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的mRNA包含SEQ ID N0:8。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的mRNA包含SEQ ID N0:9。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的mRNA包含SEQ ID N0:10。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的mRNA通过核酸测序或探针杂交来检测,如本文所述。

[0286] 在一些实施方案中,编码HSD17B13功能丧失变体蛋白的核酸分子是cDNA。编码HSD17B13功能丧失变体蛋白的cDNA分子可以是本文所述的编码HSD17B13功能丧失变体蛋白的cDNA分子中的任一种。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的cDNA包含SEQ ID N0:5。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的cDNA包含SEQ ID N0:6。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的cDNA包含SEQ ID N0:8。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的cDNA包含SEQ ID N0:9。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的cDNA包含SEQ ID N0:10。在一些实施方案中,编码HSD17B13功能丧失变体蛋白的cDNA通过核酸测序或探针杂交来检测,如本文所述。

[0287] 在一些实施方案中,所述方法还包括从受试者获得样品。在一些实施方案中,所述肝病是慢性肝病。在一些实施方案中,慢性肝病是非酒精性脂肪性肝病 (NAFLD)、酒精性肝病 (ALD)、非酒精性脂肪性肝炎 (NASH)、肝硬化、脂肪变性或肝细胞癌。在一些实施方案中,所述肝病是酒精性肝病。在一些实施方案中,酒精性肝病包括由饮酒引起的肝硬化、脂肪变性或肝细胞癌中的一种或多种。在一些实施方案中,所述肝病是非酒精性肝病。在一些实施方案中,非酒精性肝病包括非酒精性脂肪性肝病 (NAFLD) 或非酒精性脂肪性肝炎 (NASH)。在一些实施方案中,非酒精性肝病包括并非由饮酒导致的肝硬化、脂肪变性或肝细胞癌中的一种或多种。

[0288] 本公开还提供了本文所述的方法中的任一种,所述方法还包括向受试者施用HSD17B13的抑制剂。在一些实施方案中,HSD17B13的抑制剂包含与内源性HSD17B13基因组DNA或mRNA杂交并降低受试者细胞中HSD17B13多肽的表达的功能性多肽、反义DNA、RNA、siRNA或shRNA。在一些实施方案中,HSD17B13抑制剂还可以抑制短链脱氢酶/还原酶 (SDR) 家族的一个或多个另外的成员,HSD17B13是所述家族的成员。此类其他成员包括但不限于HSD17B1、HSD17B2、HSD17B3、HSD17B4、HSD17B6、HSD17B7、HSD17B8、HSD17B10、HSD17B11、HSD17B12、HSD17B13、HSD17B14、HSD11B1、HSD11B2、HSD3B1、HSD3B2和HSD3B7,以及紧密同源物脱氢酶/还原酶3 (DHRS3) 和视黄醇脱氢酶10 (RDH10)。在一些实施方案中,施用HSD17B13的抑制剂以抑制受试者的肝病。在一些实施方案中,施用HSD17B13的抑制剂以治疗受试者的肝病。在一些实施方案中,所述肝病是慢性肝病。在一些实施方案中,慢性肝病是非酒精性脂肪性肝病 (NAFLD)、酒精性肝病 (ALD)、非酒精性脂肪性肝炎 (NASH)、肝硬化、脂肪变性或肝细胞癌中的一种或多种。在一些实施方案中,所述肝病是酒精性肝病。在一些实施方案中,酒精性肝病包括由饮酒引起的肝硬化、脂肪变性或肝细胞癌中的一种或多种。在一些实施方案中,所述肝病是非酒精性肝病。在一些实施方案中,非酒精性肝病包括非酒精性脂肪性肝病 (NAFLD) 或非酒精性脂肪性肝炎 (NASH)。在一些实施方案中,非酒精性肝病包括并非由饮酒导致的肝硬化、脂肪变性或肝细胞癌中的一种或多种。在一些实施方案中,受试者对于编码I148M变异的基因是纯合的。在一些实施方案中,受试者对于编码I148M变异的基因是杂合的。在一些实施方案中,受试者还对于编码功能性HSD17B13蛋白的基因是纯合的。在一些实施方案中,受试者还对于编码功能性HSD17B13蛋白的基因和编码HSD17B13功能丧失

变体的基因是杂合的。

[0289] 本公开还提供了治疗或抑制肝病的方法,所述方法包括向表达包含I148M变异的含patatin样磷脂酶结构域3(PNPLA3)蛋白的人肝病患者施用羟基类固醇17-β脱氢酶13(HSD17B13)的抑制剂,使得在患者中治疗或抑制肝病。在一些实施方案中,所述肝病是慢性肝病。在一些实施方案中,所述肝病是非酒精性脂肪性肝病(NAFLD)、酒精性肝病(ALD)、非酒精性脂肪性肝炎(NASH)、肝硬化、脂肪变性或肝细胞癌中的一种或多种。在一些实施方案中,所述肝病是酒精性肝病。在一些实施方案中,酒精性肝病包括由饮酒引起的肝硬化、脂肪变性或肝细胞癌中的一种或多种。在一些实施方案中,所述肝病是非酒精性肝病。在一些实施方案中,非酒精性肝病包括非酒精性脂肪性肝病(NAFLD)或非酒精性脂肪性肝炎(NASH)。在一些实施方案中,非酒精性肝病包括并非由饮酒导致的肝硬化、脂肪变性或肝细胞癌中的一种或多种。在一些实施方案中,患者是肥胖的。在一些实施方案中,患者患有脂肪肝。在一些实施方案中,已经通过在来自受试者的样品中检测变体PNPLA3蛋白确定患者表达变体PNPLA3蛋白(例如,包含I148M或I144M变异的PNPLA3蛋白)。在一些实施方案中,变体PNPLA3蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸。在一些实施方案中,样品中的变体PNPLA3蛋白包含根据SEQ ID NO:42的氨基酸序列。在一些实施方案中,变体PNPLA3蛋白通过氨基酸测序或通过免疫测定来检测。在一些实施方案中,已经通过在来自受试者的样品中检测编码变体PNPLA3蛋白的核酸分子(例如,编码包含I148M或I144M变异的PNPLA3蛋白的变体PNPLA3核酸分子)确定受试者表达变体PNPLA3蛋白。在一些实施方案中,变体PNPLA3蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3蛋白包含根据SEQ ID NO:42的氨基酸序列。在一些实施方案中,编码变体PNPLA3蛋白的核酸分子包括基因组DNA、mRNA或从mRNA获得的cDNA。在一些实施方案中,核酸分子包含基因组DNA,所述基因组DNA在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子。在一些实施方案中,基因组DNA包含根据SEQ ID NO:31的核苷酸序列。在一些实施方案中,核酸分子包含mRNA,所述mRNA在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:34的核苷酸序列。在一些实施方案中,核酸分子包含mRNA,所述mRNA在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:35的核苷酸序列。在一些实施方案中,核酸分子包含从mRNA获得的cDNA,所述cDNA在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:38的核苷酸序列。在一些实施方案中,核酸分子包含从mRNA获得的cDNA,所述cDNA在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:39的核苷酸序列。在一些实施方案中,通过对核酸的至少一部分进行测序来检测核酸,所述部分编码I148M变异。在一些实施方案中,通过与核酸的一部分特异性杂交的探针或引物的杂交来检测核酸,其中所述部分包含编码I148M变异的密码子。在一些实施方案中,探针或引物是等位基因特异性探针或引物。在一些实施方案中,探针或引物包含标记。在一些实施方案中,患者对于编码变体PNPLA3蛋白的基因是纯合的。在一些实施方案中,患者对于编码变体PNPLA3蛋白的基因是杂合的。在一些实施方案中,患者对于编码功能性HSD17B13蛋白的基因是纯合的。在一些实施方案中,患者对于编码功能性HSD17B13蛋白的基因是杂合的。在一些实施方案中,患者对于编码功

能性HSD17B13蛋白的基因和编码HSD17B13功能丧失变体的基因是杂合的。

[0290] 如本文所述,可以使用HSD17B13的抑制剂治疗人受试者的肝病,所述人受试者具有包含I148M变异的PNPLA3蛋白并具有功能性HSD17B13蛋白。在一些实施方案中,人受试者已被测试为对于包含I148M变异的PNPLA3蛋白和对于功能性HSD17B13蛋白是阳性的。在一些实施方案中,所述治疗包括确定人受试者是否具有包含I148M变异的PNPLA3蛋白和功能性HSD17B13蛋白。在一些实施方案中,已经通过使用如本文定义的方法中的任一种将人受试者鉴定为用于通过抑制HSD17B13治疗或抑制肝病的候选者。在一些实施方案中,变体PNPLA3蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3蛋白包含根据SEQ ID NO:42的氨基酸序列,或与SEQ ID NO:42具有至少90%序列同一性并包含I148M变异的氨基酸序列。在一些实施方案中,变体PNPLA3蛋白在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3蛋白包含根据SEQ ID NO:43的氨基酸序列,或与SEQ ID NO:43具有至少90%序列同一性并包含I144M变异的氨基酸序列。在一些实施方案中,编码变体PNPLA3蛋白的核酸分子是基因组DNA。在一些实施方案中,基因组DNA在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子。在一些实施方案中,基因组DNA包含根据SEQ ID NO:31的核苷酸序列,或与SEQ ID NO:31具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列。在一些实施方案中,编码变体PNPLA3蛋白的核酸分子是mRNA。在一些实施方案中,mRNA在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:34的核苷酸序列,或与SEQ ID NO:34具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列。在一些实施方案中,mRNA在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:35的核苷酸序列,或与SEQ ID NO:35具有至少90%序列同一性并编码包含I144M变异的PNPLA3蛋白的核苷酸序列。在一些实施方案中,编码变体PNPLA3蛋白的核酸分子是cDNA。在一些实施方案中,cDNA在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:38的核苷酸序列,或与SEQ ID NO:38具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列。在一些实施方案中,cDNA在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:39的核苷酸序列,或与SEQ ID NO:39具有至少90%序列同一性并编码包含I144M变异的PNPLA3蛋白的核苷酸序列。在一些实施方案中,所述肝病是慢性肝病。在一些实施方案中,慢性肝病是非酒精性脂肪性肝病 (NAFLD)、酒精性肝病 (ALD)、非酒精性脂肪性肝炎 (NASH)、肝硬化、脂肪变性或肝细胞癌。在一些实施方案中,所述肝病是酒精性肝病。在一些实施方案中,酒精性肝病包括由饮酒引起的肝硬化、脂肪变性或肝细胞癌中的一种或多种。在一些实施方案中,所述肝病是非酒精性肝病。在一些实施方案中,非酒精性肝病包括非酒精性脂肪性肝病 (NAFLD) 或非酒精性脂肪性肝炎 (NASH)。在一些实施方案中,非酒精性肝病包括并非由饮酒导致的肝硬化、脂肪变性或肝细胞癌中的一种或多种。在一些实施方案中,人受试者对于功能性HSD17B13是纯合的或杂合的。

[0291] 在一些实施方案中,HSD17B13的抑制剂降低或抑制HSD17B13基因表达或HSD17B13蛋白的功能。HSD17B13的抑制剂包括但不限于天然存在的和合成的配体、拮抗剂、激动剂、

抗体、肽、环肽、核酸、功能性多核苷酸、小有机分子等。功能性多核苷酸是具有特定功能,如结合靶分子或催化特定反应的核酸分子。功能性多核苷酸的实例包括但不限于反义分子、适体、核酶和三螺旋形成分子。功能性多核苷酸可以作为靶分子所具有的特定活性的抑制剂而起作用。反义分子被设计成通过规范或不规范的碱基配对与靶核酸分子相互作用。反义分子与靶分子的相互作用被设计成通过例如RNA酶-H介导的RNA-DNA杂合体降解来促进对靶分子的破坏。可选地,反义分子被设计成中断通常将发生在靶分子上的加工功能,如转录或复制。反义分子可以基于靶分子的序列来设计。存在许多通过鉴定靶分子的最易接近的区域来优化反义效率的方法。示例性方法包括但不限于体外选择实验和使用DMS和DEPC的DNA修饰研究。反义分子通常以小于或等于约10⁻⁶、小于或等于约10⁻⁸、小于或等于约10⁻¹⁰、或小于或等于约10⁻¹²的解离常数(kd)结合靶分子。有助于反义分子的设计和使用的的方法和技术以及反义分子的代表性样品可见于以下美国专利和申请的非限制性列表:5,135,917;5,294,533;5,627,158;5,641,754;5,691,317;5,780,607;5,786,138;5,849,903;5,856,103;5,919,772;5,955,590;5,990,088;5,994,320;5,998,602;6,005,095;6,007,995;6,013,522;6,017,898;6,018,042;6,025,198;6,033,910;6,040,296;6,046,004;6,046,319;6,057,437;和2018年3月21日提交的美国序列号62/645,941,所述文献各自通过引用以其整体并入本文。反义分子的实例包括但不限于反义RNA、小干扰RNA(siRNA)和短发夹RNA(shRNA)。例如,反义RNA、siRNA或shRNA可以被设计成靶向HSD17B13基因组DNA或mRNA特有的区域。在一些实施方案中,HSD17B13的抑制剂是反义分子。在一些实施方案中,HSD17B13的抑制剂是shRNA分子。在一些实施方案中,HSD17B13的抑制剂是siRNA分子。

[0292] 在本文所述的方法中的任一种中,施用HSD17B13的抑制剂可以导致肝病特定特征的减少或消除。在一些实施方案中,肝病的特征包括但不限于炎症和纤维化。

[0293] 本公开还提供了治疗呈PNPLA3 Ile148Met阳性(即,“PNPLA3 Ile148Met⁺”)或PNPLA3 Ile144Met阳性(即,“PNPLA3 Ile144Met⁺”)的受试者的方法,所述方法包括向受试者施用HSD17B13的抑制剂。本公开还提供了治疗或抑制肝病的方法,所述方法包括向表达包含I148M变异的PNPLA3蛋白的人肝病患者施用HSD17B13的抑制剂,使得在患者中治疗或抑制肝病。

[0294] 变体PNPLA3 Ile148Met阳性或PNPLA3 Ile144Met阳性受试者可以具有本文所述的变体PNPLA3蛋白中的任一种。在一些实施方案中,受试者对于功能性HSD17B13也是纯合的或杂合的。在一些实施方案中,受试者对于功能性HSD17B13是纯合的。在一些实施方案中,受试者对于功能性HSD17B13是杂合的。在一些实施方案中,受试者对于HSD17B13功能丧失变体是纯合的。受试者可以具有本文所述的功能性HSD17B13蛋白中的任一种。

[0295] 在一些实施方案中,已经通过在来自受试者的样品中检测PNPLA3蛋白确定呈PNPLA3 Ile148Met⁺或PNPLA3 Ile144Met⁺的受试者是PNPLA3 Ile148Met⁺或PNPLA3 Ile144Met⁺,其中变体PNPLA3蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸,或者在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。变体PNPLA3 Ile148Met阳性或PNPLA3 Ile144Met阳性受试者可以具有本文所述的变体PNPLA3蛋白中的任一种。

[0296] 在一些实施方案中,已经通过在来自受试者的样品中检测PNPLA3蛋白确定呈PNPLA3 Ile148Met⁺的受试者是PNPLA3 Ile148Met⁺,其中变体PNPLA3蛋白在对应于根据SEQ

ID NO:42的位置148的位置处包含甲硫氨酸。在一些实施方案中,样品中的变体PNPLA3蛋白包含根据SEQ ID NO:42的氨基酸序列。

[0297] 在一些实施方案中,已经通过在来自受试者的样品中检测PNPLA3蛋白确定呈PNPLA3 Ile144Met+的受试者是PNPLA3Ile144Met+,其中变体PNPLA3蛋白在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,样品中的变体PNPLA3蛋白包含根据SEQ ID NO:43的氨基酸序列。

[0298] 在一些实施方案中,如本文所述通过氨基酸测序或免疫测定来鉴定在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸,或在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸的变体PNPLA3蛋白。

[0299] 在一些实施方案中,已经通过在来自受试者的样品中检测编码PNPLA3蛋白的核酸分子确定呈PNPLA3 Ile148Met+或PNPLA3Ile144Met+的受试者是PNPLA3 Ile148Met+或PNPLA3 Ile144Met+,其中变体PNPLA3蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸,或者在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。变体PNPLA3 Ile148Met阳性或PNPLA3Ile144Met阳性受试者可以具有本文所述的变体PNPLA3核酸分子中的任一种。

[0300] 在一些实施方案中,已经通过在来自受试者的样品中检测编码PNPLA3蛋白的核酸分子确定呈PNPLA3 Ile148Met+的受试者是PNPLA3 Ile148Met+,其中变体PNPLA3蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸。在一些实施方案中,样品中的变体PNPLA3蛋白包含根据SEQ ID NO:42的氨基酸序列。

[0301] 在一些实施方案中,已经通过在来自受试者的样品中检测编码PNPLA3蛋白的核酸分子确定呈PNPLA3 Ile144Met+的受试者是PNPLA3 Ile144Met+,其中变体PNPLA3蛋白在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,样品中的变体PNPLA3蛋白包含根据SEQ ID NO:43的氨基酸序列。

[0302] 在一些实施方案中,编码变体PNPLA3蛋白的核酸分子是基因组DNA、mRNA或来源于mRNA的cDNA。

[0303] 在一些实施方案中,基因组DNA在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子。在一些实施方案中,基因组DNA包含根据SEQ ID NO:31的核苷酸序列。在一些实施方案中,包含ATG密码子的编码变体PNPLA3蛋白的基因组DNA通过核酸测序或探针杂交来鉴定,如本文所述。

[0304] 在一些实施方案中,mRNA在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:34的核苷酸序列。在一些实施方案中,mRNA在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:35的核苷酸序列。在一些实施方案中,包含AUG密码子的编码变体PNPLA3蛋白的mRNA通过核酸测序或探针杂交来鉴定,如本文所述。

[0305] 在一些实施方案中,cDNA在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:38的核苷酸序列。在一些实施方案中,cDNA在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:39的核苷酸序列。在一些实施方案中,包含ATG密码子的编码变体PNPLA3蛋白的cDNA通过核酸测序或探针杂交来鉴定,如本文所述。

[0306] HSD17B13的抑制剂的施用可以通过任何合适的途径进行,包括但不限于胃肠外、静脉内、口服、皮下、动脉内、颅内、鞘内、腹膜内、局部、鼻内或肌内。用于施用的药物组合物是合意地无菌的并且基本上是等渗的并且在GMP条件下制造。药物组合物可以以单位剂型(即单次施用的剂量)提供。药物组合物可以使用一种或多种生理和药学上可接受的载体、稀释剂、赋形剂或助剂来配制。配制品取决于所选择的施用途径。术语“药学上可接受”意指载体、稀释剂、赋形剂或助剂与配制品的其他成分相容并且对其接受者基本上无害。

[0307] 在一些实施方案中,受试者患有肝病或易患肝病。在一些实施方案中,所述肝病是慢性肝病。在一些实施方案中,慢性肝病是非酒精性脂肪性肝病 (NAFLD)、酒精性肝病 (ALD)、非酒精性脂肪性肝炎 (NASH)、肝硬化、脂肪变性或肝细胞癌。在一些实施方案中,所述肝病是酒精性肝病。在一些实施方案中,酒精性肝病包括由饮酒引起的肝硬化、脂肪变性或肝细胞癌中的一种或多种。在一些实施方案中,所述肝病是非酒精性肝病。在一些实施方案中,非酒精性肝病包括非酒精性脂肪性肝病 (NAFLD) 或非酒精性脂肪性肝炎 (NASH)。在一些实施方案中,非酒精性肝病包括并非由饮酒导致的肝硬化、脂肪变性或肝细胞癌中的一种或多种。

[0308] 本公开还提供了用肝病治疗剂治疗患者的方法,其中所述患者患有肝病。所述方法包括确定来自受试者的样品是否包含:i) 编码包含I148M变异的PNPLA3蛋白的第一核酸和编码功能性HSD17B13蛋白的第二核酸,和/或ii) 包含I148M变异的PNPLA3蛋白和功能性HSD17B13蛋白。在一些实施方案中,所述方法包括确定来自受试者的样品是否包含编码包含I148M变异的PNPLA3蛋白的第一核酸和编码功能性HSD17B13蛋白的第二核酸。在一些实施方案中,所述方法包括确定来自受试者的样品是否包含含有I148M变异的PNPLA3蛋白和功能性HSD17B13蛋白。

[0309] 在一些实施方案中,此确定通过从患者获得或已获得生物样品来进行。在一些实施方案中,所述方法还包括对生物样品进行或已进行基因分型测定,以确定所述患者是否具有编码包含I148M变异的PNPLA3蛋白的第一核酸和编码功能性HSD17B13蛋白的第二核酸。在一些实施方案中,所述方法还包括对生物样品进行或已进行测定,以确定所述患者是否具有包含I148M变异的PNPLA3蛋白和功能性HSD17B13蛋白。

[0310] 在一些实施方案中,当患者具有编码包含I148M变异的PNPLA3蛋白的核酸和/或具有包含I148M变异的PNPLA3蛋白,并且具有编码功能性HSD17B13蛋白的核酸和/或具有功能性HSD17B13蛋白时,则所述方法还包括向患者施用HSD17B13的抑制剂。在一些实施方案中,当患者具有编码包含I148M变异的PNPLA3蛋白的核酸和/或具有包含I148M变异的PNPLA3蛋白,并且具有编码功能性HSD17B13蛋白的核酸和/或具有功能性HSD17B13蛋白时,则所述方法还包括向患者施用HSD17B13的抑制剂并向患者施用肝病治疗剂。在一些实施方案中,当患者具有编码包含I148M变异的PNPLA3蛋白的核酸和/或具有包含I148M变异的PNPLA3蛋白,但不具有编码功能性HSD17B13蛋白的核酸和/或具有功能性HSD17B13蛋白时,则所述方法还包括向患者施用肝病治疗剂。

[0311] 肝病治疗剂的实例包括但不限于双硫仑、纳曲酮、阿坎酸、泼尼松、泼尼松、硫唑嘌呤、青霉胺、曲恩汀、去铁胺、环丙沙星、诺氟沙星、头孢曲松、氧氟沙星、阿莫西林-克拉维酸、植物甲萘醌、布美他尼、呋塞米、氢氯噻嗪、氯噻嗪、阿米洛利、氨苯蝶啶、螺内酯、奥曲肽、阿替洛尔、美托洛尔、纳多洛尔、普萘洛尔、噻吗洛尔和卡维地洛。

[0312] 肝病治疗剂的其他实例(例如,用于慢性丙型肝炎治疗)包括但不限于利巴韦林、帕利瑞韦、西咪匹韦(Olysio)、依巴司韦、雷迪帕韦、奥比他韦、艾尔巴韦、达卡他韦(Daklinza)、达萨布韦、利托那韦、索非布韦、维帕他韦、伏西瑞韦、格卡瑞韦、哌仑他韦、聚乙二醇干扰素 α -2a、聚乙二醇干扰素 α -2b和干扰素 α -2b。

[0313] 肝病治疗剂的其他实例(例如,用于非酒精性脂肪性肝病)包括但不限于减重诱导剂,如奥利司他或西布曲明;胰岛素增敏剂,如噻唑烷二酮类(TZD)、二甲双胍和氯茴苯酸类;降脂药物,如他汀类、贝特类和 ω -3脂肪酸;抗氧化剂,如维生素E、甜菜碱、N-乙酰半胱氨酸、卵磷脂、水飞蓟素和 β -胡萝卜素;抗TNF药物,如己酮可可碱;益生菌,如VSL#3;和细胞保护剂,如熊去氧胆酸(UDCA)。其他合适的治疗包括ACE抑制剂/ARB、低聚果糖和肠降血糖素类似物。

[0314] 肝病治疗剂的其他实例(例如,用于NASH)包括但不限于奥贝胆酸(Ocaliva[®])、司隆色替(Selonseritib)、Elafibranor、塞尼克韦罗(Cenicriviroc)、GR_MD_02、MGL_3196、IMM124E、花生四烯酰胺基胆酸(AramcholTM)、GS0976、Emricasan、Volixibat、NGM282、GS9674、Tropifexor、MN_001、LMB763、BI_1467335、MSDC_0602、PF_05221304、DF102、Saroglitasar、BMS986036、Lanifibranor、索马鲁肽、硝唑尼特、GRI_0621、EYP001、VK2809、纳美芬、LIK066、MT_3995、Elobixibat、Namodenoson、佛拉鲁单抗(Foralumab)、SAR425899、Sotagliflozin、EDP_305、Isosabutate、吉卡宾(Gemcabene)、TERN_101、KBP_042、PF_06865571、DUR928、PF_06835919、NGM313、BMS_986171、那麦珠单抗(Namacizumab)、CER_209、ND_L02_s0201、RTU_1096、DRX_065、IONIS_DGAT2Rx、INT_767、NC_001、Seladepar、PXL770、TERN_201、NV556、AZD2693、SP_1373、VK0214、Hepastem、TGFTX4、RLBN1127、GKT_137831、RYI_018、CB4209-CB4211和JH_0920。

[0315] 本公开还提供了用于制备用于在人受试者中治疗肝病的药物的HSD17B13的抑制剂,所述人受试者是PNPLA3 Ile148Met阳性或PNPLA3 Ile144Met阳性,并且还对于功能性HSD17B13是纯合的或杂合的。在一些实施方案中,受试者对于PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体是纯合的。在一些实施方案中,受试者对于PNPLA3 Ile148Met变体或PNPLA3 Ile144Met变体是杂合的。在一些实施方案中,受试者对于功能性HSD17B13是纯合的。在一些实施方案中,受试者对于功能性HSD17B13是杂合的。

[0316] 在一些实施方案中,HSD17B13的抑制剂用于在人受试者中治疗肝病,所述人受试者具有在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸的PNPLA3蛋白,或编码在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸或在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸的PNPLA3蛋白的核酸分子,或者编码在对应于根据SEQ ID NO:43的位置148的位置处包含甲硫氨酸的PNPLA3蛋白的核酸分子。在一些实施方案中,变体PNPLA3蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3蛋白包含根据SEQ ID NO:42的氨基酸序列。在一些实施方案中,变体PNPLA3蛋白在对应于根据SEQ ID NO:43的位置144的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3蛋白包含根据SEQ ID NO:43的氨基酸序列。

[0317] 在一些实施方案中,编码变体PNPLA3蛋白的核酸分子是基因组DNA。在一些实施方案中,基因组DNA在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子。在一些实施方案中,基因组DNA包含根据SEQ ID NO:31的核苷酸序列。

[0318] 在一些实施方案中,编码变体PNPLA3蛋白的核酸分子是mRNA。在一些实施方案中,mRNA在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:34的核苷酸序列。在一些实施方案中,mRNA在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:35的核苷酸序列。

[0319] 在一些实施方案中,编码变体PNPLA3蛋白的核酸分子是cDNA。在一些实施方案中,cDNA在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:38的核苷酸序列。在一些实施方案中,cDNA在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:39的核苷酸序列。

[0320] 在一些实施方案中,所述肝病是慢性肝病。在一些实施方案中,慢性肝病是非酒精性脂肪性肝病 (NAFLD)、酒精性肝病 (ALD)、非酒精性脂肪性肝炎 (NASH)、肝硬化、脂肪变性或肝细胞癌。在一些实施方案中,所述肝病是酒精性肝病。在一些实施方案中,酒精性肝病包括由饮酒引起的肝硬化、脂肪变性或肝细胞癌中的一种或多种。在一些实施方案中,所述肝病是非酒精性肝病。在一些实施方案中,非酒精性肝病包括非酒精性脂肪性肝病 (NAFLD) 或非酒精性脂肪性肝炎 (NASH)。在一些实施方案中,非酒精性肝病包括并非由饮酒导致的肝硬化、脂肪变性或肝细胞癌中的一种或多种。

[0321] 在一些实施方案中,人受试者对于功能性HSD17B13是纯合的或杂合的。在一些实施方案中,受试者对于功能性HSD17B13是纯合的。在一些实施方案中,受试者对于功能性HSD17B13是杂合的。

[0322] 在本文所述的方法中的任一种中,探针或引物或改变特异性探针或改变特异性引物可以与单一核酸种类特异性互补或特异性杂交。例如,与HSD17B13转录物A、转录物B、转录物E或转录物I的核酸分子(例如,本文所述的功能性HSD17B13的mRNA、cDNA、RNA转录物或cDNA转录物中的任一种)特异性互补或特异性杂交的探针或引物或改变特异性探针或改变特异性引物不与变体HSD17B13的任何核酸分子(例如,HSD17B13的变体C、D、F、G、H的mRNA、cDNA、RNA转录物或cDNA转录物中的任一种)互补或杂交。

[0323] 本公开还提供了一种用于在人受试者中治疗肝病的HSD17B13的抑制剂,所述人受试者具有包含I148M变异的PNPLA3蛋白并具有功能性HSD17B13蛋白。在一些实施方案中,人受试者已被测试为对于包含I148M变异的PNPLA3蛋白和对于功能性HSD17B13蛋白是阳性的。在一些实施方案中,所述治疗包括确定人受试者是否具有包含I148M变异的PNPLA3蛋白和功能性HSD17B13蛋白。在一些实施方案中,已经通过使用由如本文定义的方法中的任一种将人受试者鉴定为用于通过抑制HSD17B13治疗或抑制肝病的候选者。

[0324] 在一些实施方案中,变体PNPLA3蛋白在对应于根据SEQ ID NO:42的位置148的位置处包含甲硫氨酸。在一些实施方案中,变体PNPLA3蛋白包含根据SEQ ID NO:42的氨基酸序列,或与SEQ ID NO:42具有至少90%序列同一性并包含I148M变异的氨基酸序列。在一些实施方案中,编码变体PNPLA3蛋白的核酸分子是基因组DNA。在一些实施方案中,基因组DNA在对应于根据SEQ ID NO:31的位置5107至5109的位置处包含ATG密码子。在一些实施方案中,基因组DNA包含根据SEQ ID NO:31的核苷酸序列,或与SEQ ID NO:31具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列。在一些实施方案中,编码变体

PNPLA3蛋白的核酸分子是mRNA。在一些实施方案中,mRNA在对应于根据SEQ ID NO:34的位置442至444的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:34的核苷酸序列,或与SEQ ID NO:34具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列。在一些实施方案中,mRNA在对应于根据SEQ ID NO:35的位置430至432的位置处包含AUG密码子。在一些实施方案中,mRNA包含根据SEQ ID NO:35的核苷酸序列,或与SEQ ID NO:35具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列。在一些实施方案中,编码变体PNPLA3蛋白的核酸分子是cDNA。在一些实施方案中,cDNA在对应于根据SEQ ID NO:38的位置442至444的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:38的核苷酸序列,或与SEQ ID NO:38具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列。在一些实施方案中,cDNA在对应于根据SEQ ID NO:39的位置430至432的位置处包含ATG密码子。在一些实施方案中,cDNA包含根据SEQ ID NO:39的核苷酸序列,或与SEQ ID NO:39具有至少90%序列同一性并编码包含I148M变异的PNPLA3蛋白的核苷酸序列。

[0325] 在一些实施方案中,所述肝病是慢性肝病。在一些实施方案中,慢性肝病是非酒精性脂肪性肝病 (NAFLD)、酒精性肝病 (ALD)、非酒精性脂肪性肝炎 (NASH)、肝硬化、脂肪变性或肝细胞癌。在一些实施方案中,所述肝病是酒精性肝病。在一些实施方案中,酒精性肝病包括由饮酒引起的肝硬化、脂肪变性或肝细胞癌中的一种或多种。在一些实施方案中,所述肝病是非酒精性肝病。在一些实施方案中,非酒精性肝病包括非酒精性脂肪性肝病 (NAFLD) 或非酒精性脂肪性肝炎 (NASH)。在一些实施方案中,非酒精性肝病包括并非由饮酒导致的肝硬化、脂肪变性或肝细胞癌中的一种或多种。

[0326] 在一些实施方案中,人受试者对于功能性HSD17B13是纯合的或杂合的。

[0327] 上文或下文引用的所有专利文件、网站、其他出版物、登记号等出于所有目的通过引用以其整体并入,并入程度如同明确且单独地指明每个单独的项目是通过引用如此并入一般。如果一个序列的不同版本在不同时间与一个登记号相关联,则意味着在本申请的有效提交日期与所述登记号相关联的版本。有效提交日期意指实际提交日期或优先权申请的提交日期中较早的一个,如果适用的话,参考登记号。同样,如果出版物、网站等的不同版本在不同时间发布,除非另有说明,否则意指在申请的有效提交日期最近发布的版本。除非另外特别指出,否则本公开的任何特征、步骤、元件、实施方案或方面均可以与任何其他特征、步骤、元件、实施方案或方面组合使用。尽管已出于清楚和理解的目的借助说明和实例详细描述了本公开,但将显而易见的是,可以在所附权利要求的范围内实践某些变化和修改。

[0328] 本文所述的核苷酸和氨基酸序列是使用核苷酸碱基的标准字母缩写和氨基酸的单字母代码示出的。核苷酸序列遵循标准惯例,从序列的5'端开始,并且向前行进(即每行从左到右)到3'端。仅示出每个核苷酸序列的一条链,但应理解在任何提到所展示链时包括互补链。氨基酸序列遵循标准惯例,从序列的氨基末端开始,并且向前行进(即每行从左到右)到羧基末端。

[0329] 提供以下实施例以更详细地描述实施方案。这些实施例旨在说明而非限制要求保护的实施方案。

[0330] 实施例

[0331] 提出以下实施例以便为本领域普通技术人员提供对如何制备和评价本文要求保

护的化合物、组合物、制品、装置和/或方法的完整公开和描述,并且仅意图是示例性的并且不意图限制本发明人所认定的主题范围。已努力确保关于数字(例如量、温度等)的准确性,但应考虑到一些误差和偏差。除非另外指示,否则份数是重量份,温度是以℃计或者是环境温度,并且压力等于或接近大气压。

[0332] 实施例1:PNPLA3 rs738409 (p.I148M) 与HSD17B13 rs72613567之间的遗传相互作用-研究设计

[0333] 在这项研究中,在DiscovEHR人遗传学研究(将外显子组序列数据与电子健康记录(EHR)相关联的群组)中以及在三个另外的研究中,使用外显子组测序鉴定与作为肝细胞损伤标记物的血清丙氨酸转氨酶(ALT)和天冬氨酸转氨酶(AST)水平相关的变体。还在DiscovEHR和两个独立群组中研究所牵涉的遗传变体与慢性肝病的临床诊断之间的关联。还在进行肝脏活检的肥胖手术患者的独立群组中研究这些变体之一与肝病的组织病理学严重程度之间的关联。

[0334] 研究设计和参与者

[0335] 人遗传学研究是使用基因组DNA样品和来自六个群组的数据进行的。这些研究包括两个Regeneron Genetics Center和Geisinger Health System (GHS) DiscovEHR研究群体,所述DiscovEHR研究群体源自来自GHS20的MyCode® Community Health Initiative的首批50,726名成年同意参与者。GHS发现群组由在2007年至2016年间从基层医疗门诊和专科诊所招募的46,544名欧洲个体组成,不包括招募到肥胖手术群组的所有个体。GHS肥胖手术群组由已接受肥胖手术的2,644名欧洲个体组成。在达拉斯心脏研究和宾夕法尼亚大学医学生物库进行了与肝转氨酶的关联的复制研究,所述复制研究分别包括1,357名和8,527名欧洲血统的个体。与慢性肝病的关联的复制研究包括来自达拉斯肝脏研究(DLS)的517名个体和来自达拉斯儿科肝脏研究(DPLS)的439名个体。补充附录中的方法部分描述了完整的研究描述和临床表型和疾病定义。

[0336] 来自达拉斯肝脏和儿科肝脏研究的多种族基因分型病例和对照的基线特征在图5中示出。

[0337] 样品制备、测序和基因分型

[0338] 如前所述,在Regeneron Genetics进行DiscovEHR研究、达拉斯心脏研究和宾夕法尼亚大学医学生物库的参与者的DNA样品制备和全外显子组测序(Dewey等人,Science, 2016, In Press)。在达拉斯肝脏研究和达拉斯儿科肝脏研究中,通过Taqman测定对HSD17B13rs72613567进行基因分型(并通过对每种基因型的5名个体进行Sanger测序进行了验证)。

[0339] 发现群组中的临床测量和慢性肝病定义

[0340] 丙氨酸转氨酶(ALT)和天冬氨酸转氨酶(AST)的临床实验室测量是从来自GHS发现群组和肥胖手术群组的参与者的EHR提取的。用两次或更多次测量计算所有参与者的ALT和AST中值,并在关联分析之前对其进行log₁₀转换,以对分布进行标准化。

[0341] 从EHR提取了国际疾病分类第九次修订版(ICD-9)疾病诊断代码,并将其分类为非病毒性、非酒精性(ICD-9 571.40、571.41、571.49、571.5、571.8、571.9)或酒精性(ICD-9 571.0、571.1、571.2、571.3)肝病病例定义的临床疾病类别。基于单一诊断代码的其他病例定义包括:酒精性肝硬化(ICD-9 571.2)、非酒精性肝硬化(ICD-9 571.5)和HCC(ICD-9

155.0)。对于这些病例定义,将无肝病的普通对照组(“无肝病”)定义为没有病例标准或指示任何类型的肝病的单一征兆(single-encounter)或问题列表诊断代码的参与者。

[0342] GHS发现群组中HSD17B13周围区域的丙氨酸转氨酶(ALT;A)和天冬氨酸转氨酶(AST;B)水平的区域关联图在图6(图A和图B)中示出。紫色菱形指示剪接变体rs72613567。每个圆圈指示单核苷酸变体,其中圆圈的颜色指示该变体与s72613567之间的连锁不平衡(在DiscovEHR群组中计算的 r^2)。蓝线指示HapMap中估计的重组率。图的底部示出了基因座中每个基因的相对位置和转录链。AST或ALT与邻近基因HSD17B11的编码区或剪接区变体之间没有显著关联(ALT和AST的最显著P值分别为 1.4×10^{-1} 和 4.3×10^{-2})。

[0343] 肥胖手术群组中的肝脏组织病理学表型定义

[0344] GHS肥胖手术群组由欧洲血统的2,644名个体组成,其中可以从这些个体中的2,391名获得术中肝脏活检标本。如前所述,将肝脏活检标本用福尔马林固定,并用苏木精和伊红染色以进行组织学分析,并且用马森氏三色染色(Masson's trichrome stain)以评估纤维化(Gerhard等人, *Patient Saf. Surg.*, 2011, 5, 1)。由肝病理学家使用先前建立的标准来确定组织学诊断(Brunt等人, *Am. J. Gastroenterol.*, 1999, 94, 2467-74)。将组织学诊断用于定义以下表型:1) 正常:无脂肪变性、NASH或纤维化的证据;2) 单纯性脂肪变性:脂肪变性(不考虑等级),无NASH或纤维化的证据;3) NASH:小叶炎症或肝细胞气球样变的任何存在(不考虑等级),或纤维化的任何存在(不考虑阶段)。

[0345] 来自发现和复制群组的所测序欧洲血统个体的基线特征在图1中示出。发现群组中以 $P < 1.0 \times 10^{-7}$ 与血清转氨酶水平相关的单核苷酸变体在图2中示出。

[0346] DNA样品制备和测序

[0347] 简而言之,根据制造商推荐的方案(Roche NimbleGen),使用NimbleGen探针进行外显子组捕获。对捕获的DNA进行PCR扩增,并通过qRT-PCR(Kapa Biosystems)定量。在Illumina v4 HiSeq 2500上使用75bp成对末端测序对多重化样品进行测序,其覆盖深度足以在96%的样品中提供超过85%的所靶向碱基的大于20x的单倍体读取深度(约80x的所靶向碱基的平均单倍体读取深度)。将来自每次Illumina HiSeq 2500运行的原始序列数据上传到DNAnexus平台(Reid等人, *BMC Bioinformatics*, 2014, 15, 30),以用于序列读段比对和变体鉴定。简而言之,将原始序列数据从BCL文件转换为样品特异性FASTQ文件,将其用BWA-mem与人参考构件GRCh37.p13进行比对(Li等人, *Bioinformatics*, 2009, 25, 1754-60)。使用基因组分析工具包鉴定单核苷酸变体(SNV)和插入/缺失(indel)序列变体(McKenna等人, *Genome Res.*, 2010, 20, 1297-303)。

[0348] 肝酶和慢性肝病表型的全外显子组关联分析

[0349] 对于与转氨酶水平的关联,检查了502,219个双等位基因变体,它们的缺失数据率 $< 1\%$,Hardy-Weinberg平衡P值 $> 1.0 \times 10^{-6}$,并且次要等位基因频率 $> 0.1\%$ 。针对年龄、年龄²、性别、BMI和血统的前四个主要组分,对Log₁₀转换的中值ALT和AST进行了调整。为了解释研究参与者之间的相关性,将遗传相关性矩阵拟合为随机效应协变量。主要组分和遗传相关性矩阵二者都是从呈近似连锁平衡且次要等位基因频率 $> 0.1\%$ 的39,858个非MHC标记物构建的。如在GCTA包中所实行(Yang等人, *Am. J. Hum. Genet.*, 2011, 88, 76-82),使用线性混合模型测试性状残留与单核苷酸变体之间的关联。正文中报告的所有P值均对应于等位基因模型。

[0350] 在三个独立的欧洲血统群组中尝试复制GHS发现群组中的关联：GHS肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库(如上所述)。对来自GHS肥胖手术群组和来自宾夕法尼亚大学医学生物库的ALT和AST测量值进行了 $\log_{10}$ 转换,并针对年龄、年龄²、性别、BMI和血统的前四个主要组分进行了调整。包括遗传相关性矩阵列作为随机效应协变量,并使用GCTA中的线性混合模型进行分析。在达拉斯心脏研究中,针对年龄、年龄²、性别、BMI和血统的前10个主要组分调整 $\log_{10}$ 转换的ALT和AST测量值,并使用在PLINK中实行的线性回归进行分析。使用METAL对三个复制群组的汇总统计进行综合分析(Willer等人, Bioinformatics, 2010, 26, 2190-1) (复制综合分析)。类似地对发现群组和三个复制群组的汇总统计进行综合分析(联合综合分析)。

[0351] 在三个单独的欧洲血统群组中,来自发现群组的35个全外显子组重要单核苷酸变体的复制和联合综合分析在图3中示出。

[0352] 对于GHS发现群组中与转氨酶具有全外显子组显著关联($p < 1 \times 10^{-7}$)的变体,如本文所述,在本文所述的欧洲血统复制研究中进行关联分析和综合分析。将通过所测试的变体数量确定的Bonferroni显著性阈值用于定义复制的关联。还进行发现和复制研究的综合分析。正文中报告的所有P值均对应于等位基因模型。

[0353] 还检查了转氨酶相关单核苷酸变体与慢性肝病表型的关联(如本文所述定义和分析)。将依据所测试的变体数量和慢性肝病大类确定的Bonferroni显著性阈值用于确定关联的显著性。还检查了复制的新型变体与来自GHS肥胖手术群组的组织病理学定义的肝脏表型的关联。

[0354] 与慢性肝病表型的关联分析

[0355] 分析了来自肝酶ExWAS的13种重要的复制的单核苷酸变体与从GHS发现群组定义的慢性肝病表型的关联,如上所述。将 $P < 0.05/26$ ($P < 1.92 \times 10^{-3}$)的Bonferroni显著性阈值用于解释所测试的13种变体和两个慢性肝病大类(酒精性和非酒精性)。还测试了HSD17B13 rs72613567变体与来自GHS肥胖手术群组的组织病理学定义的肝脏表型的关联,如上所述。在针对年龄、年龄²、性别、BMI和血统的前四个主要组分加以调整后,使用费斯惩罚似然逻辑回归法估计优势比。使用相同的协变量估计HSD17B13 rs72613567的基因型优势比。

[0356] 通过逻辑回归估计DLS中肝病的优势比,针对年龄、年龄²、性别、BMI和自我报告的种族进行调整。将来自达拉斯心脏研究的具有可用的rs72613567基因型的参与者用作正常对照($n=4,279$)。通过逻辑回归估计DPLS中的优势比。

[0357] 发现群组中13种全外显子组重要复制性单核苷酸变体与肝病表型的关联在图4中示出。

[0358] PNPLA3 rs738409(p.I148M)与HSD17B13 rs72613567之间的遗传相互作用-分析

[0359] 为了评估PNPLA3 rs738409和HSD17B13 rs72613567的组合效应,分别使用线性回归和逻辑回归进行定量(ALT和AST)和二元(非酒精性肝病和酒精性肝病)性状的关联分析,假设加性遗传模型,对两种遗传变体以及交互作用项的主要效应进行建模。针对年龄、年龄²、性别、BMI和血统的前四个主要组分,对所有模型进行了调整。使用基础R中的glm函数进行了统计分析。

[0360] 软件

[0361] 使用GCTA软件1.25.0版(Yang等人, Am. J. Hum. Genet., 2011, 88, 76-82)和

PLINK1.9.0版进行了遗传关联分析。使用R软件3.2.1版(用于统计计算的R项目)产生分位数-分位数和曼哈顿图。使用LocusZoom产生区域关联图(Pruim等人,Bioinformatics, 2010,26,2336-7)。

[0362] RNA测序研究

[0363] 通过在Agilent RNA纳米生物分析仪芯片上运行总RNA来评价RNA质量和浓度;所有样品均具有大于8的RNA完整指数(RIN)。通过使用寡(dT)25珠(Thermo Fisher Scientific)的两轮富集来分离聚腺苷酸化的RNA转录物。用RNAClean XP珠(Beckman Coulter)纯化样品并浓缩,并且热碎片化为大约140个碱基对。使用随机六聚体,用SuperScript III逆转录酶(Thermo Fisher Scientific)完成了第一链合成;在第二链合成过程中,用dUTP替代了dTTP。根据上文针对外显子组提到的标准DNA文库制备方法处理样品,其中加入尿嘧啶DNA-糖基化酶步骤以产生链特异性测序文库。合并样品并在Illumina v4 HiSeq 2500上使用75bp成对末端测序进行测序。

[0364] 新型HSD17B13转录物的鉴定和验证

[0365] 使用ArrayStudio[®]软件(OmicSoft[®],凯瑞市,北卡罗来纳州)将读数映射到人B38,允许两次错配。采用两种方法来鉴定新型HSD17B13转录物。使用ArrayStudio基于Gencode v24发现了新型外显子连接。使用Trinity(v2.2.0)在默认设置下进行从头转录装配。构建定制基因模型以并入HSD17B13的新型转录物,并通过与所述定制基因模型进行读段比对来估计转录物定量。确定了所有已鉴定的HSD17B13同种型的蛋白质序列比对。使用SuperScript[™]一步式RT-PCR系统用Platinum[™] Taq DNA聚合酶(ThermoFisher)对来自人肝脏样品的总RNA进行RT-PCR。每50μL RT-PCR反应含有1X反应混合物、各500nM的正向和反向引物(PST516:ATGAACATCATCCTAGAAATCCTTC;SEQ ID NO:62;和PST517:ATCATGCATACATCTCTGGCTGGAG;SEQ ID NO:63)、1μL RT/Platinum Taq和75ng RNA。循环条件为:45℃持续30分钟的一个循环;94℃持续2分钟的一个循环;94℃持续20秒,53℃持续30秒,以及72℃持续90秒的40个循环;72℃持续5分钟的一个循环;然后保持10℃。使用QIAquick PCR纯化试剂盒(Qiagen)纯化产物,并且提交使用引物DE002(ATCAGAACTTCAGGCCTTGG;SEQ ID NO:64)进行直接Sanger测序。为了鉴定B和C转录物,将RT-PCR产物在用SYBR Gold[®]SYBR[®]金核酸凝胶染色剂(ThermoFisher)染色的2%琼脂糖凝胶上运行电泳,并且切下预期分子量的条带并使用QIAquick凝胶提取试剂盒(Qiagen)纯化,然后用TOPO[®]TA克隆试剂盒(ThermoFisher)进行克隆。使用M13F和M13R测序引物对TOPO克隆进行测序。使用Sequencher DNA分析软件(Gene Codes Corporation)进行序列分析。

[0366] 使用第一个外显子(GCAAAGCCATGAACATCATCC;SEQ ID NO:65)和最后一个外显子(TCTTGATGTAGTGGGAGTCGGATT;SEQ ID NO:66)中的基因特异性引物,用SuperScript III一步式RT-PCR系统用Platinum Taq High Fidelity(Thermo Fisher Scientific)从50ng总RNA直接扩增全长HSD17B13转录物,以产生约2.2kb的扩增子(最大预测大小的转录物)。在Agilent生物分析仪上验证扩增子。将PacBio相容的条形码衔接子连接到扩增子,并用PacBio PB珠(Pacific Biosciences)清洁。将文库以等量合并,并在PacBio RSII平台的一个SMRT单元上测序180分钟。使用PacBio软件smrtanalysis v2.3工具labelzmv对数据进行多路解编,并且然后用ConsensusTools AmpliconAnalysis进行分析。将所得扩增子与

HSD17B13 RefSeq基因进行比较,以确定同种型和基因型状态。

[0367] HSD17B13同种型的亚细胞定位

[0368] 用携带HSD17B13 A和D转录物的慢病毒感染HepG2细胞,选择稳定的细胞系,并使用免疫荧光法使HSD17B13同种型、脂滴和内质网可视化。简而言之,在补充有10%胎牛血清的伊格爾最低基础培养基中培养HepG2细胞。将HSD17B13转录物A和D亚克隆到Myc-DDK骨架慢病毒构建体中,并产生慢病毒。用携带HSD17B13转录物的慢病毒感染HepG2细胞。用1-3mg/ml硫酸遗传霉素G-418在完整培养基中选择表达每个HSD17B13转录物的稳定细胞系持续两周。在固定后,用小鼠抗Myc抗体检测HSD17B13同种型。用BODIPY FL染料(Sigma)标记脂滴。用兔抗PLIN抗体(Sigma)和兔抗钙联蛋白抗体(Cell Signaling Technology)标记脂质外壳蛋白和内质网。免疫荧光的二级抗体是Alexa Fluor 488驴抗兔IgG和Alexa Fluor 594驴抗小鼠IgG(Jackson ImmunoResearch)。

[0369] 人肝活检组织中HSD17B13蛋白表达的定量

[0370] 在蛋白酶和磷酸酶抑制剂混合物(Thermo-Fisher)存在下,将人肝脏和细胞沉淀样品在冰冷的1x RIPA裂解缓冲液(EMD Millipore)中匀浆化。收集上清液并且用于使用BCA蛋白测定(Thermo-Fisher)进行蛋白浓缩。将人组织裂解物以30μg/孔加载,并将稳定的细胞系以9μg/孔加载,并且在SDS/PAGE凝胶(Bio-Rad)上分离并转移到PVDF膜(Bio-Rad)。用补充有0.1%Tween20(Bio-Rad)的1x TBS中的5%(重量/体积)牛奶将膜封闭1小时。将膜与针对HSD17B13(1:200, Thermo-Fisher)和β-肌动蛋白(1:500, Cell Signaling Technology)的抗体在4℃下一起孵育过夜。使用HRP缀合的抗兔抗体(1:10,000, Jackson ImmunoResearch)检测结合的抗体,并使用化学发光试剂(Thermo-Fisher)增强。使用Image J软件对条带强度进行定量。

[0371] HSD17B13酶活性的体外和细胞表征

[0372] 从用含有HSD17B13转录物A或转录物D的质粒DNA转化的大肠杆菌(Genscript)纯化重组人HSD17B13蛋白。HSD17B13变体在C末端处含有10xHis标签,并使用Ni²⁺亲和纯化从可溶性级份进行纯化。通过使用NAD(P)H-Glo检测系统(Promega)测量NADH产量来确定酶活性。将反应在25℃下在0.2M tris-HCl(pH 7.5)、0.5mM NAD⁺、75μM底物(Sigma)和500ng纯化酶中以100μL的最终体积进行3小时。在孵育后,将20μL反应物与20μL萤光素酶试剂(Promega)组合,在室温下孵育1小时,并在Envision读板仪(Perkin Elmer)上读取。

[0373] 在基于细胞的测定中,将过表达HSD17B13转录物A、转录物D或绿色荧光蛋白(GFP,对照)的HEK293细胞用于研究HSD17B13对雌二醇的活性。将雌二醇喂给每种细胞类型。在48小时后,收集培养基,并通过LC-MS对雌二醇及其转化产物雌酮的浓度进行鉴定和定量。通过LC-MS鉴定了羟雌二醇(来自雌二醇的代谢产物)和羟雌酮(来自雌酮的代谢产物)。

[0374] 实施例2:66个人肝脏样品中HSD17B13和PNPLA3的基因表达分析

[0375] 用66个人肝脏样品分析了HSD17B13和PNPLA3的基因表达。所有样品均来自无脂肪变性、小叶炎症或纤维化的对照供体。HSD17B13 rs72613567(T/T、T/TA和TA/TA)和PNPLA3 rs738409(C/C、C/G和G/G)基因型的分布在表1中示出。

[0376]

基因型	C/C	C/G	G/G	ND
T/T	12	8	1	0
T/TA	15	12	0	2

TA/TA	12	4	0	0
-------	----	---	---	---

[0377] 在HSD17B13 rs72613567剪接变体的纯合交替载体中,PNPLA3的表达显著降低(参见图7)。mRNA表达以FPKM单位展示。观察到与具有FDR 0.0071的T/T相比下降了1.6倍。当与HSD17B13 T/T载体相比时,具有HSD17B13 TA/TA基因型的变体PNPLA3 C/C载体的表达显著降低:降低1.7倍(FDR 0.017)。具有TA/TA基因型的变体PNPLA3 C/G载体显示出表达下降,但没有统计学显著性(1.4倍,FDR 1)。图8示出了63个PNPLA3 rs738409载体(C/C和C/G,参见表1)在三种HSD17B13 rs72613567基因型(T/T、T/TA、TA/TA)中的表达差异。

[0378] 变体PNPLA3 p.I148M变体是最充分验证的NAFLD遗传风险因子,并且148M等位基因以纯合状态存在于5%-25%的个体中,这取决于血统。为了了解HSD17B13 rs72613567:TA是否改变了与PNPLA3 p.I148M相关的肝损伤风险,在DiscovEHR中对与ALT、AST和慢性肝病表型相关的两个变体之间的相互作用进行分析。这些分析在所有参与者以及肥胖(体重指数(BMI)>30kg/m²)和非肥胖(BMI≤30kg/m²)亚群体中进行。在ALT(对于相互作用,P=1.8x 10⁻³)和AST(对于相互作用,P=4.5x 10⁻³)水平的关联分析中,HSD17B13 rs72613567:TA与PNPLA3 p.I148M之间存在名义上显著的相互作用;这些关联主要是由肥胖个体中的关联驱动的(参见图9)。在这些分析中,rs72613567:TA等位基因降低PNPLA3 148M等位基因与增加的ALT和AST的等位基因剂量依赖性关联(参见图10)。参考图10,图A示出了在具有每个PNPLA3 p.I148M基因型的个体中HSD17B13 rs72613567与ALT的关联,并且图B示出了在具有每个PNPLA3 p.I148M基因型的个体中HSD17B13 rs72613567与AST的关联。使用线性回归计算了效应估计值(β和95%CI),并针对年龄、年龄²、性别、BMI和血统的四个主要组分进行了调整。图11(图A至图F)示出了依据PNPLA3 rs738409(p.I148M)和HSD17B13rs72613567基因型的原始和残差化ALT水平。通过针对年龄、年龄²、性别、BMI和四个主要组分1-4进行调整的线性回归计算了残差。图12(图A至图F)示出了依据PNPLA3 rs738409(p.I148M)和HSD17B13rs72613567基因型的原始和残差化AST水平。通过针对年龄、年龄²、性别、BMI和四个主要组分1-4进行调整的线性回归计算了残差。图13(图A至图F)示出了HSD17B13剪接变体的纯合参考(T/T)、杂合(T/TA)和纯合交替(TA/TA)载体中四个另外新型HSD17B13转录物(E-H)的mRNA表达。基因模型中的编码区以红色指示,并且非翻译区以黑色指示。转录物E和H在外显子3与4之间含有另外的外显子。转录物F包括从外显子6到内含子6的通读。蓝色箭头指示来自rs72613567的A插入。转录物G缺少外显子2。转录物G和H中的星号表明在外显子6的3'端插入G,这导致蛋白质的过早截短(类似于转录物D)。转录物根据HSD17B13基因型差异地表达,如方框图所示。mRNA表达以FPKM单位展示。

[0379] 这些数据表明,HSD17B13 rs72613567:TA变体通过变体PNPLA3p.I148M变体减小遗传易患脂肪性肝病的个体中的肝损伤风险。这一发现可能提示了对HSD17B13的治疗性调节重要的亚群体,即对于变体PNPLA3 148M等位基因为杂合或纯合的个体。

[0380] 实施例3:66个人肝脏样品中HSD17B13和PNPLA3的基因表达分析

[0381] 外显子变体与天冬氨酸和丙氨酸转氨酶的关联

[0382] 在来自DiscovEHR研究(“GHS发现群组”;图1中的基本人口统计)的欧洲血统的46,544名个体中,检查了502,219个双等位基因单遗传变体与血清ALT或AST水平的关联。发现19个基因中的总共35个变体以P<1.0x 10⁻⁷与ALT或AST相关(参见图14和图2)。参见图14,示出了GHS发现群组中与血清转氨酶水平的单核苷酸变体关联的曼哈顿图(左)和分位数-分

位数图(右)。在16个基因中有31个变体以 $P < 1.0 \times 10^{-7}$ 与丙氨酸转氨酶(ALT)水平显著相关(参见图A)。在10个基因中有12个变体以 $P < 1.0 \times 10^{-7}$ 与天冬氨酸转氨酶(AST)水平显著相关(参见图B)。所有显著关联均在图2中示出。在三个单独的欧洲血统群组的复制综合分析中,在9个基因(此处用它们的基因名称指示)中有13个变体(包括HSD17B13)仍然与ALT或AST显著相关(参见图3)。充分校准关联测试,如通过全外显子组分位数-分位数图和基因组对照 λ 值所示。

[0383] 在欧洲血统个体的三个群组中进行了复制研究:1)来自DiscovEHR的肥胖手术患者($n=2,644$) (“GHS肥胖手术群组”);2)来自达拉斯心脏研究的1,357名个体;和3)来自宾夕法尼亚大学医学生物库的8,526名个体。在复制群组的综合分析中,9个基因中的13个变体与ALT或AST的血清水平显著相关(对于所测试的35个变体,Bonferroni显著性阈值为 $P < 1.43 \times 10^{-3}$) (参见图3)。这些包括先前报道的与升高的转氨酶水平相关的变体,如PNPLA3 (Romeo等人,Nat.Genet.,2008,40,1461-5)、TM6SF2 (Kozlitina等人,Nat.Genet.2014,46,352-6)、SERPINA1 (Brantly等人,Am.J.Med.,1988,84,13-31)、SAMD5 (Kitamoto等人,Hum.Genet.,2013,132,783-92)和ERLIN1 (Feitosa等人,Atherosclerosis,2013,228,175-80)。SERPINA1编码 α -1-抗胰蛋白酶,所述酶的功能缺陷会导致肝病;与SAMD5的关联是通过与PNPLA3中变异的连锁不平衡介导的,并且ERLIN1与肝脏脂肪沉积有关。还鉴定了先前没有报道与肝病相关的变体。这些包括GPT和GOT1中的几个变体,所述基因分别编码ALT和AST,以及编码溶质载体家族39成员12的SLC39A12。

[0384] 还鉴定了HSD17B13(所述基因编码17- β 羟基类固醇脱氢酶家族的未表征成员羟基类固醇17- β 脱氢酶13)中的变体与降低的ALT(发现 $P=4.2 \times 10^{-12}$,复制 $P=1.7 \times 10^{-4}$)和AST(发现 $P=6.2 \times 10^{-10}$,复制 $P=1.7 \times 10^{-4}$,参见图3)水平之间的可复现关联。相关变体rs72613567是在外显子6的供体剪接位点(TA等位基因)附近插入腺嘌呤,并且在GHS发现群组中等位基因频率为26.0%。先前,Chambers等人鉴定了与ALT水平相关的4q22(rs6834314)的邻近基因座(Chambers等人,Nat.Genet.,2011,43,1131-8);迄今为止,还没有报道rs72613567与转氨酶水平相关。HSD17B13位于HSD17B11(同一基因家族的另一成员)上游30kb。在发现群组中(参见图6)或在发现群组和三个复制群组的联合综合分析中,在HSD17B11的编码或剪接变体与转氨酶水平之间没有观察到全外显子组显著关联。此外,在所有血统组中,rs72613567与HSD17B11中变体的连锁不平衡较低(在所有血统组中,HSD17B11中所有已确定变体的 $r^2 < 0.4$;数据未示出)。总的来说,这些发现表明HSD17B13是基因组区域中最有可能与转氨酶水平功能性相关的基因。

[0385] 外显子变体与慢性肝病临床诊断的关联

[0386] 还分析了在发现和复制群组中发现的9个基因中的13个转氨酶相关变体与慢性肝病(包括酒精性和非酒精性(非病毒性)肝病,以及慢性肝病的最晚期形式:酒精性肝硬化、非酒精性肝硬化和肝细胞癌(HCC))之间的关系。使用所测试的13个变体的 $P < 1.92 \times 10^{-3}$ 的Bonferroni显著性阈值,发现5个基因(HSD17B13、SERPINA1、TM6SF2、PNPLA3和SAMD5)中的6个变体与慢性肝病表型之间的显著关联(参见图4)。SERPINA1、TM6SF2、PNPLA3和SAMD5关联证实了先前报道的关联。在发现群组中,HSD17B13rs72613567:TA以等位基因剂量依赖性方式与酒精性和非酒精性肝病二者的所有EHR衍生类别的较低优势相关(参见图15,图A):酒精性肝病的所有类别,杂合优势比(OR_{het}) (95%置信区间)0.58(0.42-0.80)、纯合OR

(OR_{hom}) 0.47 (0.23-0.97)、等位基因OR (OR_{等位基因}) 0.62 (0.48-0.81), $P=1.8 \times 10^{-4}$; 非酒精性肝病的所有类别, OR_{het} 0.83 (0.75-0.92)、OR_{hom} 0.70 (0.57-0.87)、OR_{等位基因} 0.84 (0.78-0.91), $P=1.3 \times 10^{-5}$ 。HSD17B13 rs72613567:TA还与较低优势酒精性肝硬化和非酒精性肝硬化相关, 杂合子和纯合子的酒精性肝硬化的较低优势分别为42%和73% (OR_{het} 0.58 (0.39-0.86)、OR_{hom} 0.27 (0.09-0.85)、OR_{等位基因} 0.56 (0.41-0.78), $P=3.4 \times 10^{-4}$), 并且杂合子和纯合子的非酒精性肝硬化的较低优势分别为26%和49% (OR_{het} 0.74 (0.60-0.93), OR_{hom} 0.51 (0.31-0.85), OR_{等位基因} 0.74 (0.62-0.88), $P=4.5 \times 10^{-4}$)。HSD17B13 rs72613567:TA还与HCC的较低优势名义上相关。

[0387] 这些发现在多种族达拉斯肝脏研究 (DLS) 和达拉斯儿科肝脏研究 (DPLS) 中得到证实和扩展 (参见图5)。在DLS中, TA等位基因以等位基因剂量依赖性方式与任何肝病的较低优势相关 (OR_{het} 0.74 (0.57-0.97), OR_{hom} 0.41 (0.21-0.83), OR_{等位基因} 0.70 (0.5-0.88), $P=1.8 \times 10^{-3}$, 参见图15, 图B)。在EHR衍生的肝病亚型中观察到类似的效应, 包括与晚期肝硬化形式的酒精性 (OR_{等位基因} 0.72 (0.53-0.99), $P=4.4 \times 10^{-2}$) 和非酒精性 (OR_{等位基因} 0.65 (0.40-1.07), $P=9.0 \times 10^{-2}$) 肝病的保护性关联。在依据自我报告的种族分组的个体的子集分析中, 西班牙裔美国人中与肝病的所述关联是显著的 ($n=326$ 名病例和722 名对照, OR_{等位基因} 0.51 (0.35-0.74), $P=4.0 \times 10^{-4}$); 在DLS的非洲裔美国人 ($n=33$ 名病例和2,291 名对照, OR_{等位基因} 0.74 (0.25-2.47), $P=0.67$) 和欧洲裔美国人 ($n=158$ 名病例和1,266 名对照, OR_{等位基因} 0.87 (0.65-1.15), $P=0.32$) 子集中也发现了类似的数字趋势, 但未实现统计学显著性。在DPLS (西班牙裔美国人儿科肝病和肥胖对照的单独研究) 中, TA等位基因还与肝病的较低优势相关 (OR_{等位基因} 0.61 (0.37-0.99), $P=4.6 \times 10^{-2}$)。因此, 在三个独立群体的成人和儿童中, HSD17B13 rs72613567:TA与多种形式的慢性肝病 (包括肝硬化) 的降低的优势相关。

[0388] 参见图15, 示出了HSD17B13 rs72613567:TA与酒精性和非酒精性肝病表型的降低的风险相关。在GHS发现群组中, HSD17B13 rs72613567以等位基因剂量依赖性方式与非酒精性和酒精性肝病、肝硬化和肝细胞癌的较低优势相关 (参见图A)。使用逻辑回归计算了优势比, 并针对年龄、年龄²、性别、BMI和血统的主要组分进行了调整。还示出了杂合 (Het OR) 和纯合 (Hom OR) 载体的基因型优势比。在达拉斯肝脏研究中, HSD17B13 rs72613567以等位基因剂量依赖性方式与任何肝病的较低优势相关 (参见图B)。在肝病亚型中观察到了相似的等位基因剂量依赖性效应。使用逻辑回归计算了优势比, 并针对年龄、年龄²、性别、BMI和自我报告的种族进行了调整。

[0389] PNPLA3 rs738409 (p.I148M) 与HSD17B13 rs72613567之间的遗传相互作用

[0390] 变体PNPLA3 p.I148M变体是最充分验证的NAFLD的遗传风险因子, 并且148M等位基因以纯合状态存在于5%-25%的个体中, 这取决于血统。为了了解HSD17B13 rs72613567:TA是否改变了与PNPLA3 p.I148M相关的肝损伤风险, 在DiscovEHR中对与ALT、AST和慢性肝病表型相关的两个变体之间的相互作用进行分析。这些分析在所有参与者以及肥胖 (体重指数 [BMI] $\geq 30 \text{ kg/m}^2$) 和非肥胖 (BMI $< 30 \text{ kg/m}^2$) 亚群体中进行。在ALT (对于相互作用, $P=1.8 \times 10^{-3}$) 和AST (对于相互作用, $P=4.5 \times 10^{-3}$) 水平的关联分析中, HSD17B13 rs72613567:TA与PNPLA3 p.I148M之间存在名义上显著的相互作用; 这些关联主要是由肥胖个体中的关联驱动的 (参见图9)。在这些分析中, rs72613567:TA等位基因降低PNPLA3 148M等位基因与增加的ALT和AST的等位基因剂量依赖性关联 (参见图16、图11和图12)。基

于RNA测序的表达分析显示,HSD17B13rs72613567:TA以等位基因剂量依赖性方式与减少的PNPLA3 mRNA表达相关(参见图7)。这些数据表明,HSD17B13 rs72613567:TA变体通过变体PNPLA3 p.I148M变体减小遗传易患脂肪性肝病的个体中的肝损伤风险。

[0391] 参见图16,示出了HSD17B13 rs72613567:TA减小与PNPLA3p.I148M相关的肝损伤风险。在具有每个PNPLA3 p.I148M基因型的个体中,HSD17B13 rs72613567与ALT的关联(参见图A)。在具有每个PNPLA3 p.I148M基因型的个体中,HSD17B13 rs72613567与AST的关联(参见图B)。使用线性回归计算了效应估计值(B和95%CI),并针对年龄、年龄²、性别、BMI和血统的四个主要组分进行了调整。在ALT和AST水平的关联分析中,HSD17B13 rs72613567:TA与PNPLA3 p.I148M之间相互作用的P值分别为 $P=1.8 \times 10^{-3}$ 和 $P=4.5 \times 10^{-3}$ 。

[0392] HSD17B13 rs72613567:TA与肝脏病理的关联

[0393] NAFLD描述了范围从没有明显炎症证据的肝脏脂肪堆积(单纯性脂肪变性)到更具临床影响的NASH的疾病谱。为了确认HSD17B13rs72613567:TA与EHR衍生的肝病诊断代码之间的关联,并为了进一步了解其与脂肪变性到NASH的组织病理学进展的关联,在GHS肥胖手术群组中进行了关联测试。在肥胖手术时通过肝脏活检评估的2,391名全外显子组测序个体的该群组中,总共555名(23%)个体没有脂肪变性、脂肪性肝炎或纤维化的证据(“正常”),830名(35%)患有单纯性脂肪变性,并且1006名(42%)患有NASH。当依据基因型比较正常肝脏、单纯性脂肪变性和NASH的患病率时,观察到对于每个TA等位基因,正常肝脏的患病率依据基因型似乎没有不同(对于T/T、T/TA和TA/TA载体,分别为23%、24%和23%,通过比例趋势的卡方检验, $P=0.5$),但NASH的患病率下降(对于T/T、T/TA和TA/TA载体,分别为45%、40%和31%, $P=1.6 \times 10^{-4}$),并且单纯性脂肪变性的患病率升高(对于T/T、T/TA和TA/TA载体,分别为33%、35%和47%, $P=1.1 \times 10^{-3}$)(参见图17,图A)。在患有脂肪变性的个体中,与单纯性脂肪变性相比,TA等位基因以等位基因剂量依赖性方式与NASH的统计学上显著较低的优势相关($OR_{het} 0.87 (0.71-1.06)$, $OR_{hom} 0.48 (0.33-0.70)$, $OR_{等位基因} 0.77 (0.66-0.90)$, $P=6.5 \times 10^{-4}$,参见图17,图B)。总之,这些数据表明了HSD17B13在介导NAFLD从单纯性脂肪变性进展到更晚期的NASH和纤维化中的作用。

[0394] 参见图17,示出了HSD17B13 rs72613567:TA与降低的从单纯性脂肪变性进展到脂肪性肝炎和纤维化的风险相关。在来自GHS肥胖手术群组的2,391名肝脏活检个体中,根据HSD17B13 rs72613567基因型的组织病理学表征的肝病的患病率(参见图A)。对于每个TA等位基因,正常肝脏的患病率依据基因型似乎没有不同(通过比例趋势的卡方检验, $P=0.5$),但NASH的患病率降低($P=1.6 \times 10^{-4}$),并且单纯性脂肪变性的患病率升高($P=1.1 \times 10^{-3}$)。在GHS肥胖手术群组中,HSD17B13 rs72613567分别与杂合和纯合TA载体中NASH的13%和52%的较低优势相关(参见图B)。使用逻辑回归计算了优势比,并针对年龄、年龄²、性别、BMI和血统的主要组分进行了调整。还示出了杂合(Het OR)和纯合(Hom OR)载体的基因型优势比。

[0395] rs72613567:TA对HSD17B13 mRNA和HSD17B13蛋白表达的效应

[0396] 检查了HSD17B13 rs72613567:TA等位基因对所述基因的已知和新型转录物表达的效应。将RNA测序用于评估组织学上正常的肝样品中的HSD17B13 mRNA表达,所述样品来自HSD17B13 rs72613567剪接变体的22个T/T纯合、30个T/TA杂合和17个TA/TA纯合载体。除了两种HSD17B13转录物A和B,还鉴定了两种新型转录物:转录物C,其缺少外显子6;和转录

物D,其含有在外显子6的3'端处插入的鸟嘌呤核苷酸,这将被预测导致蛋白质的过早截短。另外四种转录物(E-H)以非常低的水平表达(参见图13)。通过RT-PCR和Sanger测序验证了转录物(数据未示出)。还使用长读段cDNA测序验证了D转录物。这些转录物的表达水平根据HSD17B13rs72613567基因型而改变;对于每个TA等位基因,转录物A和转录物B的水平下降,而转录物C和转录物D的水平以等位基因剂量依赖性方式升高(参见图18,图A和图B)。编码全长300个氨基酸蛋白质的转录物A是T/T纯合子中的主要转录物,而编码过早截短的蛋白质的转录物D是TA/TA纯合子中的主要转录物。在人肝脏活检组织中,截短的同种型D蛋白在杂合子和TA/TA纯合子中以最低限度存在,并且同种型A蛋白的丰度以等位基因剂量依赖性方式减小(参见图18,图C和图D)。这些数据与改变mRNA剪接的HSD17B13rs72613567一致,导致在人肝脏中合成表达显著降低的截短形式的蛋白质。

[0397] 参见图18,示出了新型HSD17B13转录物的表达、亚细胞定位和酶活性。HSD17B13rs72613567剪接变体的纯合参考(T/T)、杂合(T/TA)和纯合交替(TA/TA)载体中HSD17B13转录物A和转录物D的表达(参见图A)。HSD17B13基因中的编码区以红色指示,非翻译区以粗黑线指示,并且内含子以细黑线指示。转录物D中的星号指示来自rs72613567的A插入。mRNA表达以FPKM单位(每百万个映射读段每千碱基转录物的片段)展示。来自过表达HSD17B13转录物A和D的HepG2细胞的蛋白质印迹。与HSD17B13转录物A相比,HSD17B13转录物D被翻译成具有较低分子量的截短蛋白(参见图B)。来自新鲜冷冻人肝脏和HEK293细胞样品的HSD17B13蛋白质印迹(参见图C)。人肝脏样品来自HSD17B13rs72613567剪接变体的纯合参考(T/T)、杂合(T/TA)和纯合交替(TA/TA)载体。细胞样品来自过表达未标记的HSD17B13转录物A和转录物D的HEK293细胞。HSD17B13转录物D被翻译成分子量低于HSD17B13 IsoA的截短蛋白IsoD。在人肝脏(左)和细胞(右)样品二者中,HSD17B13 IsoD蛋白水平低于IsoA蛋白水平(参见图D)。针对肌动蛋白标准化的蛋白质水平在条形柱中示出;** $P < 0.001$,* $P < 0.05$ 。HSD17B13同种型A和D二者均定位于脂滴膜上(参见图E)。将稳定地过表达HSD17B13转录物A或D的HepG2用BODIPY标记以显示脂滴,并用抗Myc标记以显示HSD17B13定位。所有图都被放大到相同的程度。比例尺指示10 μ m。插图表示原始图像的4倍放大。HSD17B13同种型A和D对17- β 雌二醇(雌二醇)、白三烯B₄(LTB₄)和13-羟基十八碳二烯酸(13(S)-HODE)的酶活性(参见表F)。HSD17B13同种型D显示同种型A对应值的<10%酶活性。G,HSD17B13同种型D在HEK293细胞中过表达时,在培养基中测量时,未显示将雌二醇(底物)大量转化为雌酮(产物),而过表达的HSD17B13同种型A显示稳健转化。

[0398] HSD17B13在人肝细胞中的表达

[0399] HSD17B13主要在肝脏中表达(Liu等人,Acta Biochim.Pol.,2007,54,213-8),其中它定位于脂滴(Su等人,Proc.Natl.Acad.Sci.USA,2014,111,11437-42),与在脂肪性肝病的发病机制中的作用一致。在用表达HSD17B13转录物A和D的慢病毒稳定转导的永生化人肝细胞系中评价HSD17B13的表达及其定位。HSD17B13同种型A主要在BODIPY标记的脂滴周围的膜上检测到(参见图18,图E)。在脂滴表面处观察到HSD17B13同种型D的类似亚细胞定位(参见图18,图F)。

[0400] 在体外和细胞模型中rs72613567:TA对HSD17B13活性的效应

[0401] 为了理解由于rs72613567:TA导致的HSD17B13蛋白过早截短的功能性后果,使用重组蛋白在体外评价同种型A和D的酶活性。检查了多于300种假定的底物,其中雌二醇、白

三烯B4和13-羟基十八碳二烯酸被HSD17B13酶促转化,从而导致羟基氧化成酮基。HSD17B13同种型D显示对3种底物显著降低的活性(参见图18,图F)。

[0402] 与GFP对照相比,过表达HSD17B13转录物A的细胞在细胞培养基中具有较低的雌二醇浓度和较高的雌酮浓度,从而表明针对雌二醇的酶活性(参见图18,图G)。过表达HSD17B13转录物D的细胞具有与GFP对照细胞类似的雌酮/雌二醇比率,从而表明HSD17B13转录物D具有显著的功能丧失。质谱分析显示雌酮快速转化为羟基雌酮和其他产物,从而解释了与消耗的雌二醇相比雌酮的低积累。

[0403] 通过大规模外显子组测序,鉴定出在HSD17B13中的剪接变体与降低的血清转氨酶水平以及降低的非酒精性和酒精性形式的肝病(包括肝病的晚期肝硬化形式和HCC)风险之间的新型关联。据我们所知,这是第一次报告改变蛋白质的变体与肝病具有保护性关联。HSD17B13 rs72613567:TA等位基因与单纯性脂肪变性无关,但降低了进展到NASH的风险。在几个不同的肝病类别和种族中,四个独立群组(DiscovEHR、DiscovEHR中的独立肥胖手术群组、DLS和DPLS)中剂量依赖性保护性关联的一致性支持以下观点:所报告的HSD17B13变体防止进展到临床上更晚期的慢性肝病。观察到的等位基因剂量依赖性也证明,对HSD17B13功能的更显著调节可能导致对疾病风险和进展的更显著效应。HSD17B13 rs72613567:TA等位基因还通过变体PNPLA3 p.I148M变体减小遗传易患脂肪性肝病的个体中的肝损伤风险。这一发现可能提示了对HSD17B13的治疗性调节重要的亚群体,即对于变体PNPLA3 I148M等位基因为杂合或纯合的个体。

[0404] 本文所述的关联发现主要基于对具有升高BMI的欧洲和西班牙裔美国人的观察。HSD17B13与HSD17B11非常接近,HSD17B11是同一个基因家族的成员,与HSD17B13具有高序列相似度,但组织分布更广。总之,本文提供的数据支持以下观点:HSD17B13是预防和治疗人脂肪性肝病的潜在治疗靶标。本文提供的数据表明,靶向HSD17B13可减少肝病从脂肪变性进展到更晚期的NASH、纤维化和肝硬化,它们与显著的发病率和死亡率相关,并且目前没有针对它们的有效治疗。

[0405] 实施例4:HSD17B13 rs72613567:TA减小与PNPLA3 I148M相关的酒精性和非酒精性肝病的风险

[0406] 通过将来自无脂肪变性、小叶炎症或纤维化的对照供体的29,928个人肝脏样品的HSD17B13和PNPLA3基因型与来自患有酒精性肝病的患者的190个样品或来自患有非酒精性肝病的患者的1857样品进行比较,分析了HSD17B13和PNPLA3基因型与肝病的关联。对于HSD17B13和PNPLA3基因型的每种组合,通过(患病组的发病率)/(对照组的发病率)的方程计算优势比,其中置信区间为95%。参考图19,图A示出了在具有每个PNPLA3 p.I148M基因型的个体中HSD17B13 rs72613567与酒精性肝病的关联,并且图B示出了在具有每个PNPLA3 p.I148M基因型的个体中HSD17B13 rs72613567与非酒精性肝病的关联。数据表明,PNPLA3 p.I148M以剂量依赖性方式与酒精性和非酒精性肝病二者的较高发病率相关。HSD17B13rs72613567:TA基因型以等位基因剂量依赖性方式与降低的酒精性和非酒精性肝病二者的风险相关。

序列表

<110> 雷杰纳荣制药公司 (Regeneron Pharmaceuticals, Inc.)

<120> 在表达PNPLA3 I148M变异的患者的肝病治疗中对HSD17B13的抑制

<130> 189238.01002 (10390W001)

<150> 62/570,985

<151> 2017-10-11

<160> 66

<170> PatentIn版本3.5

<210> 1

<211> 19118

<212> DNA

<213> 智人 (Homo sapiens)

<400> 1

```

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60
aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120
tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180
ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240
gacagagcat attggttctg tgggatatta ataagtaat gtatacatct tccaactttt 300
taaagtcaca gagtaagata tgtattttaa gaattatttg acttaccatc tacttatctt 360
tgtatttttg tttttcaaag tttgataaat tccctgggcc cttagtctgt atatgtgtca 420
ggttagttag atgaagggaag tgtaattaag aactaagcag cgatttttat gacatgggtg 480
gcaggttgat agaaagactc aggagccagt ctcttccaa gctgctaaat gaggcaagtc 540
acatattatc tctcagcctg ttttcttggc tctgaagtgg ggataataac ttaggggatg 600
ggcaagaacg ggatctgaaa attacagcta caaacaaaag tcaaacgaag aacttgcaac 660
agaaaccttt agtgcctccc ctcatgcaca agcaacacag ttctaaaata tttactgtct 720
gaccttttac agaaaatggt tgccagtccg tagtcaaaag gattaaataa gtaatatattt 780
cagcacttag catatgataa acgatacgtg gcacatgata aacaataact gtgttaaata 840
aaatatgtgc gcagtgagtc aggcttttcc ttggacatta gtatttttcc tgtgttctta 900
cttgtaaaca ctacattaac aaccccaaat aaaactgaag gaactgaaat cttgtatcat 960
tttctctaaa cttgtaaatt ctggtgaagg catgaaaata tatgcagaga agtggtttaca 1020
ggattttagg atttgaaaaa ttgtgaagta ctcttgaga atcacatttt ctgcaaatta 1080
cagtggtttt aattaccatt atattattac tttctcatgt tctttgctgt catgttttagt 1140
tgaaacctaa aatgtctctt acacttagag aactaattct tttctgtttt ttttctgaat 1200
agtgaagaat actatacaaa aaagctacta cttttttatt taacagatat gagcatttat 1260
ataatagagg agttgatgta tataaaaatg atttgccatc tttttggtct ttgaagaaat 1320
tcgaatgaac tttctggaag atagcaagaa ttacaaaata gagaaaattg ttgcctgctg 1380
ttctcaggca tttgtccaaa aatataaata agtataaatc tatgaaaagg gcttgatgaa 1440
atctaacctt caaatctctt tccagatgtg tatttttggg gaaagggcta tatttattaa 1500

```

```

gttttttttta aatttttaaaa tttccagaga caagagaaaa gtaaattaga aggaagtcgt 1560
attaaaaaatg acttaagggc ggggtgcagt gctcacacct gtaatcccag cactttggga 1620
gacggagggtg ggcagattgc tggagcccag gagttcaaga ccagcctggg cagcacagca 1680
aaaccccccaa ctctacaaaa aatacaaaaa ttagctgggt gcgggggtgc acaccgtag 1740
tcccagctac tcgggaggct gaggtgggag gatcgtttca gttcaggaag ccaaggctgc 1800
aatgagctat gatggcatca ttgcactcca agctgggcaa tagagccagg ctctgtctca 1860
aaaaaaaataa aaaaagactt aagaaaaata ggtaacccaa cctcaaaaat tctctttgaa 1920
tcattaaatt tcatggttaa acatttaagc tactgaatga ttcactctaa ggctgtaatg 1980
taactcagat ctcttttagg cgaggaagat gctggctgag ttttcatcat aactggctcc 2040
ttttgccctg tgagatgaga gacacagtag cagtttggct cttatgcaat ctaaactgtt 2100
gcgttgggaa tacggttcaa aaaacacatt ggagttaaag ctaaagcaag tgttttgcta 2160
acaaaaagac aaggcatcac attttgcaat tgtctagctc agttataaaa cagaagaata 2220
ggccggacgc ggtggctcac gctgtaatc ccagcacttt gggaggccga gacgggcgga 2280
tcacgaggtc aggagatcga gaccatcctg gataacacag tgaaaccccg tctctactaa 2340
aaatacaaaa aaattagcca ggcgtagtgg cgggcgcctg tagtcccagc tactcgggag 2400
gctgaggcag gagaatggtg tgaacccggg aggcggagct tgcagtgagc cgagatgacg 2460
ccactgcact ccagcctggg cgacagagcg agactccgtc tcaaaaaaaaa aaaaaaaaaa 2520
aaaactgaag aataattaat tcttcaatca aaacatctga tgaatgctct ggtaacttat 2580
gctctctact gacctagaaa caaatgagag agtatggtgt ggtttgtgca atctggcagt 2640
gagcaagcta ccaactaaat cagtgaaga ctctcctatt ctttttttac tcttctgcaa 2700
tcccacaaaa ggctatttga ggggatactg actttgagac tgggtcctaa catccatgtt 2760
tggggagttc aggctgctgc tccagggttt agcctacagt agcgaaatac aaaggacca 2820
gagaccactc attcaaggtt tgccctaaat agcagcaaca ccactgtcat ctcaatacac 2880
gaagaatagg gcttttcagg tatccttgcc tctttgtcac agagaagagt ttacagattg 2940
tgagacggaa aagtataatt tttaaaacct tataatattt tctataaaag tcacctgagg 3000
tgaaaacttg aaaagaatta taattttcca gaatgtgagt caagaaacat tagagcaatt 3060
ttatcttagg aaagaggtct ttgaatttag gctgaaagta aattgctctg tctccatgtc 3120
ctatggttat gggcaagttt ggtacataaa tgagaaatcc atccagtggc cttgcccac 3180
tcaactccaa acacctgaag aatgtaatgt tatactcct agagtagcag catggtctcc 3240
ctatgaaagt cttcttctt taaggagact tctttccctt cctcctagg aggatgagtc 3300
agaatcatca agaaaaatat gatgggcaga ggcatacagt ttaccattac cactagttta 3360
gaattactac ttagcacttt actgcctatt acatagttgg tgctcaacaa atgtatgata 3420
aattaatggt tgagtttttc tttcttctcc atattcatct tccatgacac cacgaagagc 3480
aatgtttttc aagaatgttc ttcaaggttt gaaagtagcc tgcttttagag aaactgccta 3540
ctgtacagcc tccaaccaag aggaaaagct gaaaaaagca tgaagggatt ttgttttggt 3600
ttgtttggtt tggttttaat atgagcattc cctggcagaa aagccagggg taatctcatt 3660
gcaactaggc aatcactctc aagaaatfff ctaacaaata aggaggccaa tttttatfff 3720
atfttgagac gaagtccac tctgtcacc aggttgagat gcaatggaat gatttcagct 3780
cactgcaacc tccgcctccc gggttcaagt gattctctg tctaaacttc ccgagtagct 3840

```

```

gggattacag gctcccacca ccacgcccag ctaatttttt gtatttttag tagagatggg 3900
gtttcaccat tttggccaga ctggtctcaa actcctgacc tcaagtgatc caccctcctc 3960
ggcctcctaa agtgctggga ttacaggcgt gagccaccac acctgaccca ggaggccaat 4020
ttttaaaagg ttaactaatc ttcattgtcca aaatgaatgt taattgttca ttttgacat 4080
gaatgttaat tttttttttt tttttttttg agacagagtc tcactctgtt gccaggctg 4140
gagtcagtg gcactatctc cactcactgc aacttcctcc tcccaggttc aagcaattat 4200
cctgcctcag cctcccaagt agctgggatt acaggcccac accatcaggc ctggctaatt 4260
tttgtatttt tagtagagac ggggtttcac catgttggcc aggctggtct tgaactcctg 4320
acctcgtgat ccgcccctct cgccaacca aagtgtctggg attacaggcg tgagccaccg 4380
cgcctagccg aatgttaatt gtctaaaaat ttttcttctc caatgtcttc tcctccactt 4440
ttttcggaat ttgtttcttc ctaattacag cgcggtgtgg aggaaactgc agctgagtgc 4500
cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 4560
tatcgtcttc taaatcaggt gagactgcag gttcacaat ttcttcagat tattttgttt 4620
cctaggacgc tgacgtggaa aatgagaaaag gtctttatga ctgcctgatt taaattggat 4680
tttagctgct aactgaagta gttatgtcac caaggaagga tatatacttt ttttcttgta 4740
tgtaatccac tcagctctgc ccattattat tgttcatatt attaataat ttcatctga 4800
tcagaagtgt gagcagtggc acagagtgc tgacaaaaga tttatcatca gggaatatgg 4860
atcacttcct agttttgttt tagtcctatt aactttgcag taattccagc ttctctttaa 4920
ttatttccct tgtgagattt tattttggtg ttaatgtagt cttctgtaga aaatgtaata 4980
ttaataatta ttatcacaat tattttaaaa gagtaaatac caaataatca caatgaacta 5040
agcactctaa caaactttac attttttaat tcaatcccta caataactct gtaaacttca 5100
ttttacagat aagcaaatta tgactcagag aggttaagcc agaccaggt catgtagtta 5160
ttaggttatg aaaccaggat ttctcaacca gcactttaga ccagggtcgg tggttcacac 5220
atgtaatccc agcactttgt gaggccaagg tggaaggatc acatgagacc aagagttcaa 5280
gaccagccca ggcaacatag tgagacccta tctctaaaaa aaaaaaaaaa aaaaaaaaaa 5340
aaagttttaa gaaaaacaca tttttaaaaa atgaacactt taaaaatatt tggtcagaat 5400
ttatatagga atttatcaac ataaatgta atttcacttt actgataaac ttgcaaaaca 5460
tgatgtgctg ggtactgaaa tttagatgtt aaaagaacag tttatccac ctttatgaca 5520
gtgttccctt ggctccacg atttgagctc aacagtctgt cttgcctgaa ctctgagaga 5580
cctcatacaa tagaagaaaag actctcatct ttggattata ttgggtccaa aactttgagt 5640
ttgaataata caccagtgaa aagtgttctt tcaatttcaa aaggtgaaga aagaagtggg 5700
tgatgtaaca atcgtggtga ataattgtgg gacagtatat ccagccgatc ttctcagcac 5760
caaggatgaa gagattacca agacatttga ggtcaacatc ctaggacatt tttgggtgag 5820
tgtgagtcag aaacatttct gatttgtgca cttctctta agatacatga aacttataac 5880
ggagttcaca tacttctgga caggaaactg gccagatctt tgccttaatc aagaatcatt 5940
aaatttgttt gagtagaaga gccacagagt ctctgacaca aggacacaga attcaagtgg 6000
acacaacaca ccagaatgta agctacttgg tctgtcttgt ccaccagtat ctgacacaaa 6060
gcttggcatg taccaggagc tcaacaaatg tttgtggagg tttgttaagg gttgtcagtg 6120
tacatctttt caatgctgtc acttgtgact tcattttttt ccctccacac catgattttg 6180

```

```

taatgtgtcc tcattttgtg gaattttaga atggaaagga catcagaagt aattacttgg 6240
atgtatatag gatcgaggac acttttggac gagactctga ggcaagtgtt ctagatccat 6300
ggggtgctgg aactgagaaa tgcagctata cagacctcat ataattgggtt agttttgtgg 6360
gagatggaaa tatcaacttc aactgccttt gtatagaaat ttttatgatt aatcttccag 6420
tgcctcaata ttagtgtaga atctagggca gatctggatt ctagaagaaa gaagaaaaaa 6480
aagagatgtg tcccccttac ctttaccagc tcttcacata tgtgaattgg ctcccatgcc 6540
caccaaaacta cacggagacc tcatacatta gctacctata gctgcataac aaattataca 6600
aaacttagtg gtttaaagca acaatgtatg ttcactatcc tctcacagtt tctatgggtt 6660
gggaatttgg aggtagcttg ggttgggagt tctagttcta tgaatttgca taggattttat 6720
taaattctta taaaatttta ttgatgtttc tcacaaaaga ggtttttgga aaaaaagaaa 6780
gacttgTTTT ctgtaacatc aacatataat atacaatatt acaaataggg agatagtga 6840
ttcaatcatg attcattagt gtggtgtaga actctcagct tacactactc aactgtctta 6900
atacagttac acaagatttc actcttttaa ttagaatgat aaagcccaa accaaaaaat 6960
tatatgacac caaattatca taaggaataa ttttagttct gaaaactctg aatttttccc 7020
ttaatattgt ttagatgaca tatccaaaaa aggatctatt tgattccttc tgaagggaag 7080
gagggggagt actgagatta gtgttggcat ggggcttacc ataccaataa atttgtatct 7140
ttatttctat catttgtaaa gaattaatca tggaatgctt ggaagtattt tatttcattg 7200
tataagttct ctcaaagcc tttctgtctt aacaaaaata aaactacctg atttggaaac 7260
ctaacgtcta tgtcattgtc tttcttcttt ctgcaatgat ccttaagatc acaaaagcac 7320
ttcttccatc gatgatggag agaaatcatg gccacatcgt cacagtggct tcagtgtgcg 7380
gccacgaagg gattccttac ctcatcccat attggtaagt atcacatgcc agccatgtta 7440
tatattttta tactttgaag ggagcattac acttcaaatt gttaccactg gagagtcctg 7500
gttcttggca tcttgaacaa agaattggac aaaactcacc acaaaagcca ggaaagaatg 7560
aagcaacaaa agcagagatt tattgaaaat gaaagtacgc ttacagggtt gggagtgggc 7620
ccaagcacag gggctcaaga gccaattaca gaattttctg gggtttaa atccccctaga 7680
ggtttccact ggttacttgg tgtacgcctt atgtaaatga agaggatgaa ttaaagttac 7740
agagtcgttt actcagtgtg caccatatgt aaatggagag gatatttcct gtcatagctg 7800
gagtgtttcc atttgattta gttctaggaa gtcagcatga atcggcctta tgttcctgc 7860
ctccagaccc tgttctcctg cctcaagatt acaatgctga gagcagagtg atttggattt 7920
acagaattta aatttatagt agtttagaat gatTTTTTaa atgactTTTT ctaaaacaat 7980
gaaaccaggt tgtaattata ttttaagatat ttttagattt ctgcaaactc ctctgtagaa 8040
caatgagaga aaacagtaat gccaaagcatg tttccattgt ttcttggaat aagaaacaga 8100
aaccacacag actgagaagc aaaacctaca gaagctaaaa tgaacacatg tctatgtcat 8160
ggccttgggtg cccaagataa gacaatcaga gtggtccctg gatcaaaaca ttttacagtg 8220
tgcttgtgcc atgaaagtgt gtgtgtgtgt gtgtgtgtgt gtgtgagaga gagagagaga 8280
gagaaaacga ctctacctga ctaaaagttg cagataccac actccatgca ccaccaaaaga 8340
cataaaggga aggaggtgag aggcgttaag gatgtactgc tgtatttgcc aaatatcctt 8400
tctgtaaac tcttctccag atcctcataa taaaattaag aggccaaagt ggcaaccatt 8460
gtcaagagaa aaactatcaa ccattgtcaa gagaataact cagttattga gagagagagg 8520

```

```

agaaatgagc agagtcctac agaagtctgt caacacagat accagttttg tagaatttct 8580
aaatgtatttt ttcctgattc atatttttca aaataaaaagc agcaataaaa actgattaga 8640
aaacagtttg aagattcaat ggaaaaacct tacatgtagg atggaaaact gaacattaag 8700
ccaatcaata gagttatttt tgttcttttg ttatcattgt tgtttaagaa atgagatacg 8760
ttcacaaattc tgcttaatca tgtaagaaaa tgaaaatgaa ttgccattta tactctcaga 8820
aaaatcacaa gtggctgatt tttggcttcc acttgttctt aagccaaatg atacgcctt 8880
ctcacagaaa gctgaggatt ggtttcactc tcccttagct aacaatgctt aataattctc 8940
ttacagttcc agcaaatttg ccgctgttg ctttcacaga ggtctgacat cagaacttca 9000
ggccttggga aaaactggta tcaaaacctc atgtctctgc ccagtttttg tgaatactgg 9060
gttcacacaaa aatccaagca caaggtaagg tcaaaatcaa gttagaatgg gtatgtggta 9120
tgataaattg atatgaaaac taatgagaaa tgttttaggca ggccaactaa tagaagaaaa 9180
tgaagaagga aaaataattt ttcttattat tattattatc ttgaaattaa aggaataaag 9240
ggggaaaaaca cattagggac tagcaggaat gatcagccac cgatgaactg ggatatttat 9300
ttgtgtccgg gagaaagcac atacatttga tcaccgttac caccctgtct ttaaaatgca 9360
aatgttccaa ggaccagcaa ataaattgag tatctagctc cttagtcaag gtgaatttct 9420
gcaagaactc ttgtctctgg tgagacagga tttgagacca caagagaaga aaaattagtc 9480
ctgaaaggag aagaaaaaag caggaaggtg tggataagaa cccgaaaatt aagccatctg 9540
cttaacaaat ttttctaata ctagtatata ttctgctgca ggttaacaaa atatactaag 9600
cttaatgatt cgaaaccaat tttttactgg aagggaatta atcctaaata tattcattca 9660
aaagaactaa acaattctct gttgagtgcc gcctcatttg aggatactga ctcttacagc 9720
ctgagttagc tatgtggctc ctgcagctgg aatcactccc tgccactgga gtccttcattg 9780
gtgttagacc ataggfactg ttgactaaag aaaaaaaaaa gtttttgttt ttatttttgt 9840
tttttttgag acagagtctc actctgtcac ccaggctgga gtacagtggc gcgatctcag 9900
ctcaccgcaa cctccgcctt tctgggttca agcaattctc ctctctcagc ctctgagta 9960
tttggtattac aggcgccac caccagcct ggctaatttt tgtattttta gtagagacgg 10020
ggtttcacca tgttggccag gctggtctca aactcctgac ctgaggtgtc ctacctgcct 10080
tggcctccta aaatgctggg attacaggag tgagccacca tgcccggcca aaaaaataag 10140
tttttaaaga attaaagtc atcctggcta acacagtga acccgtctc tactaaaaaa 10200
cacaaaaaaa ttagccgggc gtggtggcgg gcgcctgtag tcccagctgc gcgggaggct 10260
gaggcaggag aatggcgtga acccgggagg cggagcttgc agtgagccga gatcgcgcca 10320
ctgcactcca gcctgggcga cagagcgaga ctccgtctca aaaaaaaaaa aaaaaaaaaa 10380
aaaaaaaaaaga attaaagtg ttaattttat ttagaagcct tactgaagac tacagtcgga 10440
ggcctatagc ctgagagcag cccttttagag aggttcagtt gaactgttct gatagtgggg 10500
gccatgtgct ctatcctgta ttgtcttcaa agcatcttcc cagagagctg cacattgtca 10560
cagagtcagg gactttgtga aattatgctg acaaccagaa gtgagtaaac gtggcttctt 10620
acatttgcta cgttgtctca cagtacttaa taagtatgca atatgtaagt aaatactata 10680
gtactattgc aactcctgat tgttttctta gacaaggaat tgggcccaat aaaaaccctc 10740
ttggtaggca ttcaggcttc gtgtaccatg agctttccta aggttatcct gccactcttg 10800
gggaaggcat gatagatgag gggagtaagg ataatggaac tctgggtaca gggttccttg 10860

```

```

gggctaactt agaggtagac acaggcaatg ctaaataattt gggattgatt ttatagaggt 10920
tgctagattg tgaatttcct tagtaagggc taaggcattg atatgtaatg tcacacttgg 10980
ctccgaggct gggttgttgg atccatgtag atgaaatcag ggagagaaaag ggcagaacgg 11040
agtaatttag aaatgtattg atttgtatta ctctctgttg gcttgctatt caaggcagtg 11100
gagaactcaa tcacataata atctgcagca aaccacagat catcccaggg aatgaagttt 11160
taacattcgc tggctcccta actcctcacc cagcctttac attcactggc tgttcagtcc 11220
atgcctggac atcttaattt gaatacaaca ttttaaattcc atttttctgt catcatcttg 11280
cactaacaga caattctaca ctaagcctat gtttatgaat atttctcaag agtacatgta 11340
cacagccttc agtataagga aaactggaag tatgacatac ctccagttgt catactcctt 11400
gggccccctc taaattctca ttaaactgca ggataggcaa gtcagagggtg aatctcaaat 11460
acgaaattct taccgaaaag gggttccaat ccagacccca agagagggtt cttagatttc 11520
tcgcaagaaa taattcgggg caaggccaca gtgcaaagca aaagcaagtt tattaggaaa 11580
gtaaaggagt agagaacagc tactccatgg agaagaatgg cttgagctgc tccaccaagg 11640
gtatttagag ttacttcttg attatatgct aaacaagggg tggattattc atgagttttc 11700
cgggaaaagg gtgagcaatt cccagaactg agatttcctc cccttttttag gccatatagg 11760
gtaacttcct gccattgcca tggatattgt aaactgtcat agtgctggtg gaagtgtctc 11820
ttagcttgct aatgtattat agttagctta taatgagcag tgaggacaac cagaggtcac 11880
tttcatcacc atcttggttt tgggtgggttt tggccggctt ctttactgca ccctatttta 11940
tcaacaaggt ctttatgacc tgaatcttgt gccaacctcc tatctcatcc tgtgacaaaag 12000
aatgccttaa ctctctggga atgcagccca gtaggtgtca gccttatttt acccagaccc 12060
tattcaagat ggagttgctc tgatttaaac gcctctgaca aaatgacgac ctcaaaaaca 12120
tccagcttta tggaataacct ccacaagaaa gaaagtatac ttagctatag aattttctcc 12180
ttgcatccaa caggctttga gatgtcagat gtttccttc tgtcccatga ttaatcctag 12240
ccattcctct ttcttgctg gctccactac tccttaccat ctaatgcctc gccaccattt 12300
tgatatattt actaagtgag ctatgaaaca cacctactgg atatgaaagt ataagtttct 12360
gataacaaaa catcaacatg ggatgtggag gaagtgggta ggggtggcatt aatgcagcaa 12420
atcctggaat attttaaatc ttcattctaa atttagtaaa aatataggat aattttcctg 12480
ccatcattta ctataaaaat taaaatttta gaaaataaaa ataataattt cctcttttta 12540
atcacagatt atggcctgta ttggagacag atgaagtcgt aagaagtctg atagatggaa 12600
tacttaccaa taagaaaatg atttttgttc catcgatat caatatcttt ctgagactac 12660
agaagtaagt acagcacaga acacccaaat actaaaacac caatagagct tttttttttg 12720
cttttttttt ttttagacag agtctcactc tgtcaccttg gctggattgc ggtgggttgca 12780
gtggcatgat cttggctcac tgcaacctcc gcctcctggg ttcaagcaat tctcatgcct 12840
cagaccccca agtaactggg attataggtg tgtgtgccca cactacaccc agctaatttt 12900
tgtatttttt gatagagaca ggtttcccca tgttgccag gctggactcg aactcctgac 12960
ctcaagttat cctcctgtct cggcctccca aagtgtgagg attacagtca tgagccacca 13020
tgcttgccc aatagagcta ttattatgga gcatctttca gttgtgaaaa ttggcatgga 13080
aactctccat ccctggggag aacagttatt tcctctgtta ttttctacc cagtctataa 13140
aaagagagtg attcattttc tctaccaaat ctactgtctc tgcccaaact ttgctgaaga 13200

```

```

ctattctaac taaaggaaac acagtttaaa aagaatgcaa tatagtgaag tagttaataa 13260
taaagactcc atttttaaaa gtctgctgga agtttggttg ggattgcact gaatctatag 13320
agcaattggg gagtattgac atatcaacaa tattgagttt tctaatacaa gaacataata 13380
tctattttta aaatcttctt caaaatcttt aaatctttaa attgtatttt gtagtttttg 13440
gtgtttaagt cttgcacata ttttgtcaga tttattccaa agtatttcac gggttctttt 13500
tttttttttt tttttttttt ttgagacaga gtttcaccct tgttgcccag gctggagtgc 13560
agtggcgtga tcttggtcga ctgcagcttc tgccctctgg cttcaagtga ttctcctgcc 13620
tcagcctccc aagtagctgg gattacaggc acctgcccc tcgcccact aactttttgt 13680
gtttgtagta gagacagggt ttcaccatgt tggccaggct ggtctcgaac tcctgacctc 13740
atgtgatcca cctgcctcag cctcccaaag tgctgggatt acaggcatga gccatcatgc 13800
ccagccctat ttgacggttt ttgacgctaa tgcaagtggc attttaaaaa attttatatt 13860
tcccattgtt tgtgtcagt atatattgga tttttgtaat ttgatctcat attttgcagt 13920
cttgctaaat tgctaaacct ctttttgcata aactcgataa gctttttttt ttttggtaga 13980
ttcctgggcc tctaattttc tttatgggaa agtttttaaat tacaatttta atttctttaa 14040
tagctacatg gctattcaat ttacttatta attcttggtta atgtgtgtct ttcaaggaat 14100
ttgtccatth catctaagtt gtagaatttc tttggcataa atttgtacat aacattccct 14160
tattatcctt ttaatgtctt tagaatgtct tattttattta tttattttatt tttattatat 14220
ttttttgaga cagagtctcg ctctgttgcc caggctggag tgcagtggca caatcttggc 14280
tcaactgaag ctccgccttc tgggttcatt ccattctcct gcctcagcct ccctagttgc 14340
tgggactaca ggcgccctgca accatgcccc gcttattttt tttttttttt tttttttttt 14400
tttttttttt tttttttttt tagtagagac ggggtttcac cctgttagcc aggatggtct 14460
cgatctcctg acctggtgat ccgcccgcct cagcctccca aagtgtctggg attacaggcg 14520
tgagccacca agcccagcct atttatttat ttagtagaga cagtctcact ttgctgcccc 14580
ggcaacaaag gttttgaatg cctggcctca agcagtcctc ctgccttggc ctcccaaagt 14640
gctgggatta caggcatgag ccactgcacc tggccaaatg aatatgctga taatatcttc 14700
tttataagga tgacataaga ataaaataat gtaatacaaa caaagcccct gtcactgaaa 14760
atgtatagac ttcaaatgtt aaagtcttag agaacagaat ttatatgaaa tagcaacagc 14820
aacaatttcc cagaggaaat actctctcag ctttcttctg aggagcagtt tctaaattga 14880
aattgtatca gtgagaagat aactatacta acttcataag ccttgggcct ttttgaaaca 14940
aatccatata aactatgaac aaacttgaat tagaacaatt tgagaacagg gtacaaactg 15000
cattggtgta tcaatttcag tatttggttt tagcttaaat agactgactt gagataacat 15060
aaggagaacc ttgaccccca agcaacatca tctcgcgagt tgactaggcc ggggtgtggtg 15120
tctcacgcct gtaattccag cactttggga ggccacagca ggcagatcac ttgaggtcag 15180
gcattcgaga ccagcctggc caacatggtg aaacctcagc tctactaaag atacgaaaat 15240
tagcaggcat agtggcctgc acctgtaata ccaggcactc gcaggagaat cccttgaacc 15300
cggaaggcgg agattgcagt aaacctgatg tgtgccactg cactccagcc tgggcaacag 15360
gagactctgt ctcggaaaaa taaatttttt aaaaaaatga aaaaaataa aagttgacta 15420
aattagtgtc ttggtactaa gcactgtagg aagtgagttt catggaaccc caactctctt 15480
ggggcccaaa gcaagtcata ttaattattga aaattacatg catatacatg catatgacca 15540

```

```

aggtgataaa aacaattatt ctgcctgagt tggagaatag tatcccagta aaataaacia 15600
gagtctcaaa gtcttttgta tcctttgaag ctgtcatggt ggtttgtaac taggcaacag 15660
gtatatattg ttaatcttct ttgcatttaa ttccttttat agagagacac aattttacga 15720
gcagatgcaa ttactagcat gaaggtttct ttgtgagggt agttaaagg cccacatgag 15780
ctctcttctt atccttgtcc ttctttcagc cagatcttcc ctgccccttt gctcattcca 15840
tctttcaccc acctaccccc aaaacaagga agtaaactct gcattagtca acaataccaa 15900
agtgattttc aatatgactt tctctgcaga atgttattat ttctgcctct ttacattcac 15960
atactgtctt cttttttttt tttttttttt tttttttttt tagattgggt ctcactctgt 16020
tgcccaggct ggagtgagct ggcttgatct cagctcactg taacctccac ctctgagtt 16080
caagcaattc tcctgcctca gcctcctgag tagctgggat tacaggcatg tgccaccaca 16140
cctggctagt ttttttgat ttttagtaga gacagggttt caccatgttg gtcaagctgg 16200
tctcgaactc ctgacctcat gatctgacca cctgtgctc tcaaagtgt gggattacag 16260
gcgtgagcca cggggccagc cactctcttc ctttcagttg cctactcatc tcttatgcat 16320
tcctggacat cagttgtcct tttgaagctt tcctccacta tcccagccca tgtgaatcct 16380
ccttcagtt atagccctta attctagatg gctgatattt ttcaataatt gttttaagat 16440
gaccatttta gcctatcagc taaacaatat caaagacaat agctattttt caagtacttt 16500
agtttacctt attatagagt gcataataga tattcagtaa atagtaaagg agaggtgaag 16560
gcttgcatag aatggattct ggtggtgtct cttggtgagc ttttagcatc aagattaatc 16620
agcagtttca gcaatgagct cagacctca gttttagatc tttactcata tcagataaga 16680
gagtgagaag agtggatgt atcagtgtt tatttatatt tgcattcaat ttgaactatg 16740
aatattacaa aggtgcacac ataggttcag acagattgat ttaaaatgac caaagatgac 16800
ctgtcgtaag caacctgggt atcttaagat gcactccttg gagagggaat gttcctaaaa 16860
acattttcag agggacgaac tgtatgaaat tcagtaaaac ataaatcatg agggaaactg 16920
attactctct ttttgacatg aaatgagagt tttaatgcat ggttacgatt attaacgtac 16980
tccgtgcaa gacgttaata aagttactgt tttgcaggct agaatgtctt gatgctgtaa 17040
tcagaacaca ctttttcccc tttcttcag cttcaaatgc agattcataa ttgggctgac 17100
ttctaataac tgcaatgttt tctgccttgg gcttgagca gaagcctgac aaaatagtgt 17160
ttgttttaggc aataatttat ttatttattt attgagatgg agtttcattc ttgtcgccca 17220
ggctggagtg caatggcgtg atctcggtc actgcaacct ctgtgttcag gcaataattt 17280
agactttacc ttacttgtga ttactatagc aattactata gccacaaggc ataattttac 17340
tgtctcattt caattttatg aatttgaatg tttttacact tttcctaagc aagtccacta 17400
tgaagtattg tcaaaaaaaaa aaaagaaaaa gaaagatgca cacgtaaaag agaggtggtt 17460
gcaagagaag aaaagaacgg aggaaagtta aacgcaaacc agataactct cagcgtattc 17520
taaagacca aaacagaaac tctgttgta aagattttta atggaaaatt tttcaatttt 17580
tttttctttt ttgtacaggt ttcttcctga acgcgcctca gcgattttta atcgtatgca 17640
gaatattcaa tttgaagcag tggttggcca caaaatcaaa atgaaatgaa taaataagct 17700
ccagccagag atgtatgcat gataatgata tgaatagttt cgaatcaatg ctgcaaagct 17760
ttatttcaca ttttttcagt cctgataata ttaaaaacat tggtttgga ctagcagcag 17820
tcaaacgaac aagattaatt acctgtcttc ctgtttctca agaataattt cgtagttttt 17880

```

cataggtctg tttttccttt catgcctctt aaaaacttct gtgcttacat aaacatactt 17940
 aaaaggtttt ctttaagata ttttatTTTT ccatttaaag gtggacaaaa gctacctccc 18000
 taaaagtaaa tacaaagaga acttatttac acagggaagg tttaagactg ttcaagtagc 18060
 attccaatct gtagccatgc cacagaatat caacaagaac acagaatgag tgcacagcta 18120
 agagatcaag tttcagcagg cagctttatc tcaacctgga catatTTTaa gattcagcat 18180
 ttgaaagatt tccctagcct cttcctTTTT cattagccca aaacggtgca actctattct 18240
 ggacttttatt acttgattct gtcttctgta taactctgaa gtccaccaa agtggaccct 18300
 ctatatTTcc tccctTTTTa tagtcttata agatacata tgaaagggtga ccgactctat 18360
 tttaaatctc agaattTTaa gttctagccc catgataacc tttttctttg taattttatgc 18420
 tttcatatat ccttggTccc agagatgttt agacaatttt aggctcaaaa attaaagcta 18480
 acacaggaaa aggaactgta ctggctatta cataagaaac aatggaccca agagaagaaa 18540
 aggaagaaaag aaaggTTTTt tggTTTTgt tttgTTTTgt tttgTTTTtt gTTTTTTga 18600
 gatggagtct cactctttcg cccaggtggt agtgcagtgg tatgatctca gctcactgca 18660
 agctccacct cccgggttca cgccattctc ctgcctcagc ctctgagta gctgggacta 18720
 caggcgcccc ccaccacacc cggctaattt tttgtatttt ttgtagagac ggggtttcac 18780
 catgttagcc aagatggtct cgatctcctg acctcgtgat ccacctgcct cggcctccca 18840
 aagtgtctggg attacgggtg tgagccaccg tgcccagcct tttttTTTTt aatagaaaaa 18900
 ataatccgac tcccactaca tcaagactaa tcttgttttg tgtgtttttc acatgtatta 18960
 tagaatgctt ttgcatggac tatcctcttg tttttattaa aaacaaatga tttttttaaa 19020
 agtcacaaaa acaattcact aaaaataaat atgtcattgt gctttaaaaa aataacctct 19080
 tgtagttata aaataaaacg tttgacttct aaactctg 19118

<210> 2

<211> 19119

<212> DNA

<213> 智人 (Homo sapiens)

<400> 2

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60
 aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120
 tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180
 ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240
 gacagagcat attggttctg tgggatatta ataaggtaat gtatacatct tccaactttt 300
 taaagtcaca gagtaagata tgtatTTTaa gaattatttg acttaccatc tacttatctt 360
 tgtatTTTTg tttttcaaag tttgataaat tccttggtcc cttagtctgt atatgtgtca 420
 ggttagttag atgaaggga tgtaattaag aactaagcag cgattttttat gacatggtgt 480
 gcaggttgat agaaagactc aggagccagt ctcttccaa gctgctaaat gaggcaagtc 540
 acatattatc tctcagcctg ttttcttggc tctgaagtgg ggataataac ttaggggatg 600
 ggcaagaacg ggatctgaaa attacagcta caaacaaaag tcaaacgaag aacttgcaac 660
 agaaaccttt agtgctccc ctcatgcaca agcaacacag ttctaaaata tttactgtct 720
 gaccctttac agaaaaatgt tgccagtccg tagtcaaaag gattaaataa gtaatatttt 780

cagcacttag catatgataa acgatacgtg gcacatgata aacaataact gtgttaaata 840
 aaatatgtgc gcagtgagtc aggcttttcc ttggacatta gtatttttcc tgtgttctta 900
 cttgtaaaca ctacattaac aaccccaaat aaaactgaag gaactgaaat cttgtatcat 960
 tttctctaaa cttgtaaatt ctggtaaggc catgaaaata tatgcagaga agtgtttaca 1020
 ggatttttagg attggaaaaa ttgtgaagta ctcttgaga atcacatttt ctgcaaatta 1080
 cagtggtttt aattaccatt atattattac tttctcatgt tctttgctgt catgttttagt 1140
 tgaaacctaa aatgtctctt acacttagag aactaattct tttctgtttt ttttctgaat 1200
 agtgaagaat actatacaaa aaagctacta cttttttatt taacagatat gagcatttat 1260
 ataatagagg agttgatgta tataaaaaatg atttgccatc tttttggtct ttgaagaaat 1320
 tcgaatgaac tttctggaag atagcaagaa ttacaaaata gagaaaattg ttgcctgctg 1380
 ttctcaggca tttgtccaaa aatataaata agtataaatc tatgaaaagg gcttgatgaa 1440
 atctaacctt caaatctctt tccagatgtg ttttttggg gaaagggcta tatttattaa 1500
 gtttttttta aattttaaaa tttccagaga caagagaaaa gtaaattaga aggaagtcgt 1560
 attaaaaatg acttaagggc gggtagcgtg gctcacacct gtaatcccag cactttggga 1620
 gacggaggtg ggcagattgc tggagcccag gagttcaaga ccagcctggg cagcacagca 1680
 aaacccccaa ctctacaaaa aatacaaaaa ttagctgggt gcgggggtgc acaccgtag 1740
 tcccagctac tcgggaggct gaggtgggag gatcgtttca gttcaggaag ccaaggctgc 1800
 aatgagctat gatggcatca ttgcactcca agctgggcaa tagagccagg ctctgtctca 1860
 aaaaaataa aaaaagactt aagaaaaata ggtaacccaa cctcaaaaat tctctttgaa 1920
 tcattaaatt tcatggttaa acatttaagc tactgaatga ttcactctaa ggctgtaatg 1980
 taactcagat ctcttttagg cgaggaagat gctggctgag ttttcatcat aactggctcc 2040
 ttttgccctg tgagatgaga gacacagtag cagtttggt cttatgcaat ctaaactgtt 2100
 gcgttgggaa tacggttcaa aaaacacatt ggagttaaag ctaaagcaag tgttttgcta 2160
 acaaaaagac aaggcatcac attttgcaat tgtctagctc agttataaaa cagaagaata 2220
 ggccggacgc ggtggctcac gcctgtaatc ccagcacttt gggaggccga gacgggcgga 2280
 tcacgaggtc aggagatcga gaccatcctg gataacacag tgaaaccccg tctctactaa 2340
 aaatacaaaa aaattagcca ggcgtagtgg cgggcgcctg tagtcccagc tactcgggag 2400
 gctgaggcag gagaatgggt tgaacccggg aggcggagct tgcagtgagc cgagatgacg 2460
 cactgcact ccagcctggg cgacagagcg agactccgtc tcaaaaaaaaa aaaaaaaaaa 2520
 aaaactgaag aataattaat tcttcaatca aaacatctga tgaatgctct ggtaacttat 2580
 gctctctact gacctagaaa caaatgagag agtatggtgt ggtttgtgca atctggcagt 2640
 gagcaagcta ccaactaaat cagtgaaga ctctctatt ctttttttac tcttctgcaa 2700
 tcccacaaaa ggctatttga ggggatactg actttgagac tgggtcctaa catccatgtt 2760
 tggggagttc aggctgctgc tccagggttt agcctacagt agcgaaatac aaaggaccca 2820
 gagaccactc attcaagggt tgcctaaat agcagcaaca ccactgtcat ctcaatacac 2880
 gaagaatagg gcttttcagg tatccttgcc tctttgtcac agagaagagt ttacagattg 2940
 tgagacggaa aagtataatt tttaaaacct tataatattt tctataaaag tcacctgagg 3000
 tgaaaacttg aaaagaatta taattttcca gaatgtgagt caagaaacat tagagcaatt 3060
 ttatcttagg aaagaggctt ttgaatttag gctgaaagta aattgctctg tctccatgtc 3120

ctatggttat gggcaagttt ggtacataaa tgagaaatcc atccagtggc cttgcccac 3180
 tcactcccaa acacctgaag aatgtaatgt tatatctcct agagtagcag catggctctcc 3240
 ctatgaaagt ccttcttctt taaggagact tctttccctt ccctcctagg aggatgagtc 3300
 agaatcatca agaaaaatat gatgggcaga ggcatacagt ttaccattac cactagttta 3360
 gaattactac ttagcacttt actgcctatt acatagttagg tgctcaacaa atgtatgata 3420
 aattaatggt tgagtttttc tttcttctcc atattcatct tccatgacac cacgaagagc 3480
 aatgtttttc aagaatgttc ttcaagggtt gaaagtagcc tgcttttagag aaactgccta 3540
 ctgtacagcc tccaaccaag aggaaaagct gaaaaaagca tgaagggatt ttgttttggt 3600
 ttgttttggt tggttttaat atgagcattc cctggcagaa aagccagggg taatctcatt 3660
 gcaactaggc aatcactctc aagaaatttt ctaacaaata aggaggccaa tttttatttt 3720
 attttgagac gaagtcccac tctgtcacc aggttggagt gcaatggaat gatttcagct 3780
 cactgcaacc tccgcctccc gggttcaagt gattctctg tctaaacttc ccgagtagct 3840
 gggattacag gctcccacca ccacgccag ctaatttttt gtatttttag tagagatggg 3900
 gtttcacat tttggccaga ctggtctcaa actcctgacc tcaagtgate caccctctc 3960
 ggctcctaa agtgctggga ttacaggcgt gagccaccac acctgacca ggaggccaat 4020
 ttttaaagg ttaactaatc ttcatgtcca aatgaatgt taattgttca ttttgacat 4080
 gaatgttaat tttttttttt tttttttttg agacagagtc tctctgtgt gccaggctg 4140
 gagtccagt gcactatctc cactcactgc aacttctcc tcccagggtc aagcaattat 4200
 cctgcctcag cctcccaagt agctgggatt acaggccac accatcaggc ctggctaatt 4260
 tttgtatttt tagtagagac ggggtttcac catgttggcc aggctggtct tgaactcctg 4320
 acctcgtgat ccgcccctc cggccaacca aagtgtctggg attacaggcg tgagccaccg 4380
 cgcctagccg aatgttaatt gtctaaaaat tttcttctc caatgtcttc tcctccactt 4440
 ttttcggaat ttgtttcttc ctaattacag cgcggtgtgg aggaaactgc agctgagtgc 4500
 cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 4560
 tatcgctctc taaatcagggt gagactgcag gttcaciaat ttcttcagat tattttgttt 4620
 cctaggacgc tgacgtggaa aatgagaaag gtctttatga ctgcctgatt taaattggat 4680
 tttagctgct aactgaagta gttatgtcac caaggaagga tatatacttt ttttcttgta 4740
 tgtaatccac tcagctctgc ccattattat tgttcattat attaatcaat ttcatctga 4800
 tcagaagtgt gagcagtggc acagagtgc tgacaaaaga tttatcatca gggaatatgg 4860
 atcacttcct agttttgttt tagtcttatt aactttgcag taattccagc ttctctttaa 4920
 ttatttccct tgtgagattt tattttggtg ttaatgtagt cttctgtaga aaatgtaata 4980
 ttaataatta ttatcacaat tattttaaaa gagtaaatac caaataatca caatgaacta 5040
 agcactctaa caaactttac attttttaat tcaatcccta caataactct gtaaacttca 5100
 ttttacagat aagcaaatta tgactcagag aggttaagcc agaccaggt catgtagtta 5160
 ttaggttatg aaaccaggat ttctcaacca gcactttaga ccagggtcgg tggttcacac 5220
 atgtaatccc agcactttgt gaggccaagg tggaaggatc acatgagacc aagagttcaa 5280
 gaccagccca ggcaacatag tgagacccta tctctaaaaa aaaaaaaaaa aaaaaaaaaa 5340
 aaagtttaaa gaaaaacaca tttttaaaaa atgaacactt taaaaatatt tggtcagaat 5400
 ttatatagga atttatcaac ataaatgtta atttcacttt actgataaac ttgcaaaaca 5460

```

tgatgtgctg ggtactgaaa tttagatgtt aaaagaacag tttatccac ctttatgaca 5520
gtgttccctt ggctccacg atttgagctc aacagtctgt cttgcctgaa ctctgagaga 5580
cctcatacaa tagaagaaag actctcatct ttggattata ttggtcccaa aactttgagt 5640
ttgaataata caccagtgga aagtgttctt tcaatttcaa aaggtgaaga aagaagtggg 5700
tgatgtaaca atcgtgggtga ataatgctgg gacagtatat ccagccgatc ttctcagcac 5760
caaggatgaa gagattacca agacatttga ggtcaacatc ctaggacatt tttgggtgag 5820
tgtgagtcag aaacatttct gatttgtgca ctttctctta agatacatga aacttataac 5880
ggagttcaca tacttctgga caggaaactg gccagatctt tgccttaatc aagaatcatt 5940
aaattttgtt gagtagaaga gccacagagt ctctgacaca aggacacaga attcaagtgg 6000
acacaacaca ccagaatgta agctacttgg tctgtcttgt ccaccagtat ctgacacaaa 6060
gcttggcatg taccaggagc tcaacaaatg tttgtggagg tttgttaagg gttgtcagtg 6120
tacatctttt caatgctgtc acttgtgact tcattttttt cctccacac catgattttg 6180
taatgtgtcc tcattttgtg gaattttaga atggaaagga catcagaagt aattacttgg 6240
atgtatatag gatcgaggac acttttggac gagactctga ggcaagtgtt ctagatccat 6300
ggggtgctgg aactgagaaa tgcagctata cagacctcat ataattgggt agttttgtgg 6360
gagatggaaa tatcaacttc aactgccttt gtatagaaat ttttatgatt aatcttccag 6420
tgcctcaata ttagttaga atctagggca gatctggatt ctagaagaaa gaagaaaaaa 6480
aagagatgtg tcccccttac ctttaccagc tcttcacata tgtgaattgg ctcccatgcc 6540
caccaacta cacggagacc tcatacatta gctacctata gctgcataac aaattataca 6600
aaacttagtg gtttaaagca acaatgtatg ttcactatcc tctcacagtt tctatgggtt 6660
gggaatttgg aggtagcttg ggttgggagt tctagttcta tgaatttgca taggatttat 6720
taaattctta taaaatttta ttgatgtttc tcacaaaaga ggtttttgga aaaaaagaaa 6780
gacttgtttt ctgtaacatc aacatataat atacaatatt acaaataggg agatagttaa 6840
ttcaatcatg attcattagt gtggtgtaga actctcagct tacactactc aactgtctta 6900
atacagttac acaagatttc actcttttaa ttagaatgat aaagcccaa accaaaaaat 6960
tatatgacac caaattatca taaggaataa ttttagttct gaaaactctg aatttttccc 7020
ttaatattgt ttagatgaca tatccaaaaa aggatctatt tgattccttc tgaagggaag 7080
gagggggagt actgagatta gtgttggcat ggggcttacc ataccaataa atttgtatct 7140
ttatttctat catttgtaaa gaattaatca tggaatgctt ggaagtattt tatttcattg 7200
tataagttct ctcaaatgcc tttctgtctt aacaaaaata aaactacctg atttggaac 7260
ctaacgtcta tgtcattgtc tttcttctt ctgcaatgat ccttaagatc acaaaagcac 7320
ttcttccatc gatgatggag agaaatcatg gccacatcgt cacagtggct tcagtgtgcg 7380
gccacgaagg gattccttac ctcatcccat attggttaagt atcacatgcc agccatgtta 7440
tatattttta tactttgaag ggagcattac acttcaaatt gttaccactg gagagtcttg 7500
gttcttggca tcttgaacaa agaattggac aaaactcacc acaaagcca ggaaagaatg 7560
aagcaacaaa agcagagatt tattgaaaat gaaagtacgc ttacagggt gggagtgggc 7620
ccaagcacag gggctcaaga gccaattaca gaattttctg gggtttaaat accccctaga 7680
ggtttccact gggtacttgg tgtacgcct atgtaaata agaggatgaa ttaaagttac 7740
agagtcgttt actcagtga caccatatgt aaatggagag gatatttcct gtcatagctg 7800

```

gagtgtttcc atttgattta gttctaggaa gtcagcatga atcggcctta tgttccctgc 7860
 ctccagaccc tgttctcctg cctcaagatt acaatgctga gagcagagtg atttgatttt 7920
 acagaattta aatttatagt agtttagaat gattttttta atgacttttt ctaaaacaat 7980
 gaaaccaggt tgtaattata ttttaagatat ttttagattt ctgcaaactc ctctgtagaa 8040
 caatgagaga aaacagtaat gccaagcatg tttccattgt ttcctggaat aagaaacaga 8100
 aacccacacag actgagaagc aaaacctaca gaagctaaaa tgaacacatg tctatgtcat 8160
 ggccttggtg cccaagataa gacaatcaga gtggtccttg gatcaaaaca ttttacagt 8220
 tgcttggtgc atgaaaagtgt gtgtgtgtgt gtgtgtgtgt gtgtgagaga gagagagaga 8280
 gagaaaaacga ctctacctga ctaaaagttg cagataccac actccatgca ccaccaaaaga 8340
 cataaaggga aggaggtgag aggcgttaag gatgtactgc tgtatttgcc aaatatcctt 8400
 tcctgtaaac tcttctccag atcctcataa taaaattaag aggccaaagt ggcaaccatt 8460
 gtcaagagaa aaactatcaa ccattgtcaa gagaataact cagttattga gagagagagg 8520
 agaaatgagc agagtcctac agaagtctgt caacacagat accagttttg tagaatttct 8580
 aaatgtattt ttcctgattc atatttttca aaataaaagc agcaataaaa actgattaga 8640
 aaacagtttg aagattcaat ggaaaaacct tacatgtagg atggaaaact gaacattaag 8700
 ccaatcaata gagttatttt tgttcttttg ttatcattgt tgtttaagaa atgagatacg 8760
 ttcacaattc tgcttaatca tgtaagaaaa tgaaaatgaa ttgccattta tactctcaga 8820
 aaaatcacia gtggctgatt tttggcttcc acttgttctt aagccaaatg ataccgcctt 8880
 ctcacagaaa gctgaggatt ggtttcactc tcccttagct aacaatgctt aataattctc 8940
 ttacagttcc agcaaatttg ccgctgttg ctttcacaga ggtctgacat cagaacttca 9000
 ggccttggga aaaactggta tcaaaacctc atgtctctgc ccagtttttg tgaatactgg 9060
 gttcaccaaa aatccaagca caaggtaagg tcaaaatcaa gttagaatgg gtatgtggta 9120
 tgataaattg atatgaaaac taatgagaaa tgtttaggca ggccaactaa tagaagaaaa 9180
 tgaagaagga aaaataattt ttcttattat tattattatc ttgaaattaa aggaataaag 9240
 ggggaaaaca cattagggac tagcaggaat gatcagccac cgatgaactg ggatatttat 9300
 ttgtgtccgg gagaaagcac atacatttga tcaccgttac caccctgtct ttaaaatgca 9360
 aatgttccaa ggaccagcaa ataaattgag tatctagctc cttagtcaag gtgaatttct 9420
 gcaagaactc ttgtctctgg tgagacagga tttgagacca caagagaaga aaaattagtc 9480
 ctgaaaggag aagaaaaaag caggaaggtg tggataagaa cccgaaaatt aagccatctg 9540
 cttaacaaat ttttctaate ctagtatata ttctgctgca ggtaacaaa atatactaag 9600
 cttaatgatt cgaaaccaat tttttactgg aagggaatta atcctaaata tattcattca 9660
 aaagaactaa acaattctct gttgagtgcc gcctcatttg aggatactga ctcttacagc 9720
 ctgagttagc tatgtggtct ctgcagctgg aatcactccc tgccactgga gtccttcatg 9780
 gtgttagacc ataggtactg ttgactaaag aaaaaaaaaa gtttttgttt ttatttttgt 9840
 tttttttgag acagagtctc actctgtcac ccaggctgga gtacagtggc gcgatctcag 9900
 ctcaccgcaa cctccgcctt tctgggttca agcaattctc ctctctcagc ctcttagta 9960
 tttggattac aggcgcccac caccacgcct ggctaatttt tgtattttta gtagagacgg 10020
 ggtttcacca tgttgccag gctggtctca aactcctgac ctcaggtgtc ctacctgcct 10080
 tggcctccta aaatgctggg attacaggag tgagccacca tgcccggcca aaaaaataag 10140

```

tttttaaaga attaaaggtc atcctggcta acacagtga accccgtctc tactaaaaaa 10200
cacaaaaaaaa ttagccgggc gtggtggcgg gcgcctgtag tcccagctgc gcgggaggct 10260
gaggcaggag aatggcgtga acccgggagg cggagcttgc agtgagccga gatcgcgcca 10320
ctgcactcca gcctgggcga cagagcgaga ctccgtctca aaaaaaaaaa aaaaaaaaaa 10380
aaaaaaaaaga attaaaggtg ttaattttat ttagaagcct tactgaagac tacagtcgga 10440
ggcctatagc ctgagagcag ccctttagag aggttcagtt gaactgttct gatagtgggg 10500
gccatgtgct ctatcctgta ttgtcttcaa agcatcttct cagagagctg cacattgtca 10560
cagagtcagg gactttgtga aattatgctg acaaccagaa gtgagtaaac gtggcttctt 10620
acatttgcta cgttgtctca cagtacttaa taagtatgca atatgtaagt aaatactata 10680
gtactattgc aactcctgat tgttttctta gacaaggaat tgggccaat aaaaaccctc 10740
ttggtaggca ttcaggttc gtgtacatg agctttccta aggttatcct gccactcttg 10800
gggaaggcat gatagatgag gggagtaagg ataatggaac tctgggtaca gggttcctgg 10860
gggctaactt agaggtagac acaggcaatg ctaaataattt gggattgatt ttatagaggt 10920
tgctagattg tgaatttcct tagtaagggc taaggcattg atatgtaatg tcacacttg 10980
ctccgaggct gggttgttgg atccatgtag atgaaatcag ggagagaaag ggcagaacgg 11040
agtaatttag aaatgtattg atttgtatta ctctctgttg gcttgctatt caaggcagt 11100
gagaactcaa tcacataata atctgcagca aaccacagat catcccaggg aatgaagttt 11160
taacattcgc tggctcccta actcctcacc cagcctttac attcactggc tgttcagtcc 11220
atgcctggac atcttaattt gaataacaaca ttttaaattc atttttctgt catcatcttg 11280
cactaacaga caattctaca ctaagcctat gtttatgaat atttctcaag agtacatgta 11340
cacagccttc agtataagga aaactggaag tatgacatac ctccagttgt catactcctt 11400
gggccccctc taaattctca ttaaactgca ggataggcaa gtcagaggtg aatctcaaat 11460
acgaaattct taccggaag gggttccaat ccagacccca agagagggtt ctagatttc 11520
tcgcaagaaa taattcgggg caaggccaca gtgcaaagca aaagcaagtt tattaggaaa 11580
gtaaaggagt agagaacagc tactccatgg agaagaatgg cttgagctgc tccaccaagg 11640
gtatttagag ttacttcttg attatatgct aaacaagggg tggattattc atgagttttc 11700
cgggaaaagg gtgagcaatt ccagaaactg agatttcctc cccttttttag gccatatagg 11760
gtaacttcct gccattgcc tggattttgt aaactgtcat agtgctgggtg gaagtgtctc 11820
ttagcttgct aatgtattat agttagctta taatgagcag tgaggacaac cagaggtcac 11880
tttcatcacc atcttggttt tgggtgggtt tggccggctt ctttactgca ccctatttta 11940
tcaacaaggc ctttatgacc tgaatcttgt gccaacctcc tatctcatcc tgtgacaaag 12000
aatgccttaa cttcctggga atgcagccca gtaggtgtca gccttatttt acccagacc 12060
tattcaagat ggagttgctc tgatttaaac gcctctgaca aatgacgac ctcaaaacaa 12120
tccagcttta tggaatacct ccacaagaaa gaaagtatac ttagctatag aattttctcc 12180
ttgcatccaa caggctttga gatgtcagat gtttccttcc tgtcccatga ttaatectag 12240
ccattcctct ttcttgtctg gctccactac tccttaccat ctaatgcctc gccaccattt 12300
tgatattttg actaagttag ctatgaaaca cacctactgg atatgaaagt ataagtttct 12360
gataacaaaa catcaacatg ggatgtggag gaagtgggta ggggtggcatt aatgcagcaa 12420
atcctggaat attttaaatc ttcattctaa atttagtaaa aatataggat aattttcctg 12480

```

ccatcattta cttataaaat taaaatttta gaaaataaaa ataatatattt cctcttttta 12540
atcacagatt atggcctgta ttggagacag atgaagtcgt aagaagtctg atagatggaa 12600
tacttaccaa taagaaaatg atttttgttc catcgtatat caatatcttt ctgagactac 12660
agaagttaag tacagcacag aacacccaaa tactaaaaca ccaatagagc tttttttttt 12720
gctttttttt tttttagaca gagtctcact ctgtcaccct ggctggattg cggtggttgc 12780
agtggcatga tcttggctca ctgcaacctc cgctcctgg gttcaagcaa ttctcatgcc 12840
tcagaccccc aagtaactgg gattataggt gtgtgctgcc aactacacc cagctaattt 12900
ttgtattttt tgatagagac aggtttcccc atgttggcca ggctggactc gaactcctga 12960
cctcaagtta tcctcctgtc tcggcctccc aaagtgctgg gattacagtc atgagccacc 13020
atgcctggcc caatagagct attattatgg agcatctttc agttgtgaaa attggcatgg 13080
aaactctcca tccctgggga gaacagttat ttctctgtt attttctac ccagtctata 13140
aaaagagagt gattcatttt ctctaccaa tctactgtct ctgcccac tttgctgaag 13200
actattctaa ctaaaggaaa cacagtttaa aaagaatgca atatagtga gtagttaata 13260
ataaagactc catttttaaa agtctgctgg aagtttggtt gggattgcac tgaatctata 13320
gagcaattgg ggagtattga catatcaaca atattgagtt ttctaatacca agaacataat 13380
atctattttt aaaatcttct tcaaaatctt taaatcttta aattgtattt tgtagttttt 13440
gggtgttaag tcttgacat atttgtcag atttattcca aagtatttca cgggttcttt 13500
tttttttttt tttttttttt ttgagacag agtttcaccc ttgttgccca ggctggagtg 13560
cagtggcgtg atcttggctc actgcagctt ctgcctcctg gcttcaagtg attctcctgc 13620
ctcagcctcc caagtagctg ggattacagg cacctgcccc ctgcccac taactttttg 13680
tgttttagt agagacaggg tttcaccatg ttggccaggc tggctcga ctcctgacct 13740
catgtgatcc acctgcctca gcctcccaa gtgctgggat tacaggcatg agccatcatg 13800
cccagcccta ttgacggtt ttgacgcta atgcaagtgg cattttaaaa aattttatat 13860
ttcccatgtt ttgtgtcag tatatattgg atttttgtaa ttgatctca tattttgcag 13920
tcttgctaaa ttgctaaacc tctttttgct aaactcgata agcttttttt ttttggtag 13980
attcctgggc ctctaatttt ctttatggga aagtttttaa ttacaaattt aatttcttta 14040
atagctacat ggctattcaa tttacttatt aattcttggg aatgtgtgtc tttcaaggaa 14100
tttgtccatt tcatctaagt ttagaattt ctttggcata aatttgtaca taacattccc 14160
ttattatcct ttaaatgtct ttagaatgtc ttattttttt atttatttat ttttattata 14220
tttttttgag acagagtctc gctctgttgc ccaggctgga gtgcagtggc acaatcttgg 14280
ctcactgcaa gctccgcctt ctgggttcat gccattctcc tgcctcagcc tccctagtgtg 14340
ctgggactac aggcgcctgc aacctagccc agcttatttt tttttttttt tttttttttt 14400
tttttttttt tttttttttt ttagtagaga cggggtttca ccctgttagc caggatggtc 14460
tcgatctcct gacctggtga tccgccccgc tcagcctccc aaagtgtgtg gattacaggc 14520
gtgagccacc aagcccagcc tatttattta tttagtagag acagtctcac tttgctgccc 14580
aggcaacaaa ggttttgaat gcctggcctc aagcagtcct cctgccttgg cctcccaaag 14640
tgctgggatt acaggcatga gccactgcac ctggccaaat gaatatgtg ataatatctt 14700
ctttataagg atgacataag aataaaaataa tgtaatacaa acaaagcccc tgtcactgaa 14760
aatgtataga cttcaaagt taaagtctta gagaacagaa tttatatgaa atagcaacag 14820

caacaatttc ccagaggaaa tactctctca gctttcttct gaggagcagt ttctaaattg 14880
 aaattgtatc agtgagaaga taactatact aacttcataa gccttgggcc tttttgaaac 14940
 aaatccatat aaactatgaa caaacttgaa atagaacaat ttgagaacag ggtacaaaact 15000
 gcattgggtg atcaatttca gtatttgggt ttagcttaaa tagactgact tgagataaca 15060
 taaggagAAC cttgaccccc aagcaacatc atctcgcgag ttgactaggc cgggtgtggt 15120
 gtctcacgcc tgtaattcca gcactttggg aggccacagc aggcagatca cttgaggtca 15180
 ggcatcgcag accagcctgg ccaacatggt gaaacctcag ctctactaaa gatacgaAAA 15240
 ttagcaggca tagtggcctg cacctgtaat accaggcact cgcaggagaa tcccttgaac 15300
 ccggaaggcg gagattgcag taaacctatga ttgtgccact gcactccagc ctgggcaaca 15360
 ggagactctg tctcggAAAA ataaattttt taaaaaatg aaaaaaata aaagttgact 15420
 aaattagtg tttggtacta agcactgtag gaagtgaatt tcatggaacc ccaactctct 15480
 tggggcccaa agcaagtcatt attaatattg aaaattacat gcataacat gcataatgacc 15540
 aaggtgataa aaacaattat tctgcctgag ttggagaata gtatcccagt aaaataaaca 15600
 agagtctcaa agtcttttgt atcctttgaa gctgtcatgg tggtttgtta ctaggcaaca 15660
 ggtatatatt gttaatcttc tttgcattta attcctttta tagagagaca caattttacg 15720
 agcagatgca attactagca tgaaggtttc tttgtgagg tagttaaaag gccacatga 15780
 gctctcttct tatccttgct cttctttcag ccagatcttc cctgcccctt tgctcattcc 15840
 atctttcacc cacctacccc caaaacaagg aagtaaactt tgcattagtc aacaatacca 15900
 aagtgatttt caatatgact ttctctgcag aatgttatta tttctgcctc tttacattca 15960
 catactgtct tccttttttt tttttttttt tttttttttt ttagattggg tctcactctg 16020
 ttgccaggc tggagtgcag tggcttgatc tcagctcact gtaacctcca cctcctgagt 16080
 tcaagcaatt ctctgcctc agcctcctga gtagctggga ttacaggcat gtgccaccac 16140
 acctggctag tttttttgta ttttttagtag agacagggtt tcaccatggt ggtcaagctg 16200
 gtctcgaact cctgacctca tgatctgacc acctgtgcct ctcaaagtgc tgggattaca 16260
 ggcgtgagcc accgggccag ccactctctt ctttcagtt gcctactcat ctcttatgca 16320
 ttcttgga tcaagtgtcc ttttgaagct ttctccact atcccagccc atgtgaatcc 16380
 tccttccagt tatagccctt aattctagat ggctgatatt tttcaataat tgttttaaga 16440
 tgaccatttt agcctatcag ctaaacaata tcaaagacaa tagctatttt tcaagtactt 16500
 tagtttacct tattatagag tgcataatag atattcagta aatagtaaag gagaggtgaa 16560
 ggcttgcata gaatggattc tgggtgtgtc tcttggtgag ctttttagcat caagattaat 16620
 cagcagtttc agcaatgagc tcagaccttc agtttttagat ctttactcat atcagataag 16680
 agagtgagaa gagggttatg tatcagtgtt tttttatat ttgcatccaa tttgaactat 16740
 gaatattaca aaggtgcaca cataggttca gacagattga tttaaaatga ccaaagatga 16800
 cctgtcgtaa gcaacctggg tatcttaaga tgcactcctt ggagaggga tgttcctaaa 16860
 aacattttca gagggacgaa ctgtatgaaa ttcagtaaaa cataaatcat gaggaact 16920
 gattactctc tttttgacat gaaatgagag ttttaatgca tggttacgat tattaacgta 16980
 ctccgctgca agacgttaat aaagtactg ttttgcaggc tagaatgtct tgatgctgta 17040
 atcagaacac actttttccc ctttcttcca gcttcaaatg cagattcata attgggctga 17100
 cttctaataa ctgcaatggt ttctgccttg ggcttgcagc agaagcctga caaatagtg 17160

```

tttgttttagg caataattta tttattttatt tattgagatg gagtttcatt cttgtcgccc 17220
aggctggagt gcaatggcgt gatctcggct cactgcaacc tctgtgttca ggcaataatt 17280
tagactttac cttacttgtg attactatag caattactat agccacaagg cataatttta 17340
ctgtctcatt tcaattttat gaatttgaat gtttttacac ttttcctaata gaagtccact 17400
atgaagtatt gtcaaaaaaa aaaaagaaaa agaaagatgc acacgtaaaa gagaggtggt 17460
tgcaagagaa gaaaagaacg gaggaagtt aaacgcaaac cagataactc tcagcgtatt 17520
ctaaatgacc aaaaacagaa ctctgttgtc aaagatttta aatggaaaat ttttcaattt 17580
ttttttcttt tttgtacagg tttcttctctg aacgcgcctc agcgatttta aatcgatgc 17640
agaatattca atttgaagca gtggttggtc acaaaatcaa aatgaaatga ataaataagc 17700
tccagccaga gatgtatgca tgataatgat atgaatagtt tcgaatcaat gctgcaaagc 17760
tttattttcac attttttcag tcttgataat attaaaaaca ttggtttggtc actagcagca 17820
gtcaaacgaa caagattaat tacctgtctt cctgtttctc aagaatattt acgtagtttt 17880
tcataggtct gtttttcctt tcatgcctct taaaaacttc tgtgcttaca taaacatact 17940
taaaaggttt tctttaagat attttatttt tccatttaaa ggtggacaaa agctacctcc 18000
ctaaaagtaa atacaaagag aacttattta cacaggaag gttaagact gttcaagtag 18060
cattccaatc tgtagccatg ccacagaata tcaacaagaa cacagaatga gtgcacagct 18120
aagagatcaa gtttcagcag gcagctttat ctcaacctgg acatatttta agattcagca 18180
tttgaaagat ttccctagcc tcttcctttt tcattagccc aaaacggtgc aactctattc 18240
tggtactttat tacttgattc tgtcttctgt ataactctga agtccaccaa aagtggacct 18300
tctatatattc ctcccttttt atagtcttat aagatacatt atgaaaggtg accgactcta 18360
ttttaaatct cagaatttta agttctagcc ccatgataac ctttttcttt gtaatttatg 18420
ctttcatata tccttgggtc cagagatggt tagacaattt taggctcaaa aattaaagct 18480
aacacaggaa aaggaactgt actggctatt acataagaaa caatggacct aagagaagaa 18540
aaggaagaaa gaaaggtttt ttggtttttg ttttgttttg ttttgttttt tgtttttttg 18600
agatggagtc tcaactcttc gccagggtg gagtgagtg gtatgatctc agctcactgc 18660
aagctccacc tcccgggttc acgccattct cctgcctcag cctcctgagt agctgggact 18720
acaggcgccc gccaccacac ccggctaatt ttttgatttt tttgtagaga cggggtttca 18780
ccatgttagc caagatggtc tcgatctcct gacctcgtga tccacctgcc tcggcctccc 18840
aaagtgcagg gattacgggt gtgagccacc gtgcccagcc tttttttttt taatagaaaa 18900
aataatccga ctcccactac atcaagacta atcttgtttt gtgtgttttt cacatgtatt 18960
atagaatgct tttgcatgga ctatcctctt gtttttatta aaaacaaatg atttttttta 19020
aagtcacaaa aacaattcac taaaaataaa tatgtcattg tgcttttaaaa aaataacctc 19080
ttgtagttat aaaataaaac gtttgacttc taaactctg 19119

```

<210> 3

<211> 900

<212> DNA

<213> 智人(Homo sapiens)

<400> 3

```

augaacauca uccuagaaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60

```

ucguugguga aguuuuuau uccucagagg agaaaaucug uggcugggga gauuguucuc 120
 auuacuggag cugggcaugg aaauaggcagg cagacuacuu augaauuugc aaaacgacag 180
 agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240
 cgaaaacuag gcgucacugc gcaugcgauu gugguagacu gcagcaacag agaagagauc 300
 uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaauugu ggugaauau 360
 gcugggacag uauauccagc cgaucuucuc agcaccaagg augaagagau uaccaagaca 420
 uuugagguca acauccuagg acauuuuugg aucacaaaag cacuucuucc aucgaugaug 480
 gagagaaauc auggccacau cgucacagug gcuucagugu gcggccacga agggauuccu 540
 uaccucaucc cauauuguuc cagcaauuuu gccgcuguug gcuuucacag aggucugaca 600
 ucagaacuuc aggccuuggg aaaaacuggu aucaaaaccu caugucucug cccaguuuuu 660
 gugaauacug gguucaccaa aaauccaagc acaagauuau ggccuguaau ggagacagau 720
 gaagucguua gaagucugau agauggaaua cuuaccaaua agaaaugau uuuuguucca 780
 ucguauauca auauuuuuu gagacuacag aaguuuucuc cugaacgcgc cucagcgauu 840
 uuaaaucgua ugcagaauau ucaauuugaa gcagugguug gccacaaaau caaaaugaaa 900

<210> 4

<211> 792

<212> DNA

<213> 智人 (Homo sapiens)

<400> 4

augaacauca uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60
 ucguugguga aguuuuuau uccucagagg agaaaaucug uggcugggga gauuguucuc 120
 auuacuggag cugggcaugg aaauaggcagg cagacuacuu augaauuugc aaaacgacag 180
 agcauauugg uucuguggga uauuaauaag gugaagaaag aaguggguga uguaacaau 240
 guggugaaua augcugggac aguauaucca gccgaucuuc ucagcaccaa ggaugaagag 300
 auuaccaaga cauugaggu caacauccua ggacauuuuu ggaucaaaa agcacuucuu 360
 ccaucgauga uggagagaaa ucauggccac aucgucacag ugguucagu gucggccac 420
 gaagggaauu cuuaccucau cccauauugu uccagcaauu uugccgcugu ugguuucac 480
 agaggucuga caucagaacu ucaggccuug ggaaaaacug guaucaaaac cucaugucuc 540
 ugcccaguuu uugugaauac ugguuacac aaaaauccaa gcacaagauu auggccagua 600
 uuggagacag augaagucgu aagaagucug auagauggaa uacuuaccau uaagaaaau 660
 auuuuuguuc caucguauau caauaucuuu cugagacuac agaaguucu uccugaacgc 720
 gccucagcga uuuuaaaucg uaugcagaau auucauuuug aagcaguggu ugccacaaa 780
 aucaaaauga aa 792

<210> 5

<211> 783

<212> DNA

<213> 智人 (Homo sapiens)

<400> 5

augaacauca uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60

ucguugguga aguuuuucau uccucagagg agaaaaucug uggcugggga gauuguucuc 120
 auuacuggag cugggcaugg aaauaggcagg cagacuacuu augaauuugc aaaacgacag 180
 agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240
 cgaaaacuag gcgucacugc gcaugcguau gugguagacu gcagcaacag agaagagauc 300
 uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaauugu ggugaauaau 360
 gcugggacag uauauccagc cgaucuucuc agcaccaagg augaagagau uaccaagaca 420
 uuugagguca acauccuagg acauuuuugg aucacaaaag cacuucuucc aucgaugaug 480
 gagagaaauc auggccacau cgucacagug gcuucagugu gcggccacga agggauuccu 540
 uaccucaucc cauauuguuc cagcaaaauu gccgcuguug gcuuucacag aggucugaca 600
 ucagaacuuc aggccuuggg aaaaacuggu aucaaaaccu caugucucug cccaguuuuu 660
 gugaauacug gguucaccaa aaauccaagc acaagguuuc uuccugaacg cgccucagcg 720
 auuuuaaauc guaugcagaa uauucaauuu gaagcagugg uuggccacaa aaucaaaug 780
 aaa 783

<210> 6

<211> 822

<212> DNA

<213> 智人 (Homo sapiens)

<400> 6

augaacauca uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60
 ucguugguga aguuuuucau uccucagagg agaaaaucug uggcugggga gauuguucuc 120
 auuacuggag cugggcaugg aaauaggcagg cagacuacuu augaauuugc aaaacgacag 180
 agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240
 cgaaaacuag gcgucacugc gcaugcguau gugguagacu gcagcaacag agaagagauc 300
 uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaauugu ggugaauaau 360
 gcugggacag uauauccagc cgaucuucuc agcaccaagg augaagagau uaccaagaca 420
 uuugagguca acauccuagg acauuuuugg aucacaaaag cacuucuucc aucgaugaug 480
 gagagaaauc auggccacau cgucacagug gcuucagugu gcggccacga agggauuccu 540
 uaccucaucc cauauuguuc cagcaaaauu gccgcuguug gcuuucacag aggucugaca 600
 ucagaacuuc aggccuuggg aaaaacuggu aucaaaaccu caugucucug cccaguuuuu 660
 gugaauacug gguucaccaa aaauccaagc acaagauuau ggccuguauu ggagacagau 720
 gaagucguua gaagucugau agauggaaua cuuaccaaua agaaaauugau uuuguuucca 780
 ucguauauca auauuuuuu gagacuacag aagguuuuuu cc 822

<210> 7

<211> 972

<212> DNA

<213> 智人 (Homo sapiens)

<400> 7

augaacauca uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60
 ucguugguga aguuuuucau uccucagagg agaaaaucug uggcugggga gauuguucuc 120

auuacuggag cugggcaugg aaauaggcagg cagacuacuu augaauuugc aaaacgacag 180
 agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240
 cgaaaacuag gcgucacugc gcaugcguau gugguagacu gcagcaacag agaagagauc 300
 uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaauugu ggugaauau 360
 gcugggacag uauauccagc cgaucuucuc agcaccaagg augaagagau uaccaagaca 420
 uuugagguca acauccuagg acauuuuugg aauggaaagg acaucagaag uaauuacuug 480
 gauguauaua ggaucgagga cacuuuugga cgagacucug agaucacaaa agcacuucuu 540
 ccaucgauga uggagagaaa ucauggccac aucgucacag uggcuucagu gugcgccac 600
 gaagggaauuc cuuaccucau ccuauauugu uccagcaauu uggccgcugu uggcuuucac 660
 agaggucuga caucagaacu ucaggccuug ggaaaaacug guaucaaaac cucaugucuc 720
 ugcccaguuu uugugaauac uggguucacc aaaaauccaa gcacaagauu auggccugua 780
 uuggagacag augaagucgu aagaagucug auagauggaa uacuuaccaa uaagaaaug 840
 auuuuuguuc caucguauau caauaucuuu cugagacuac agaaguucu uccugaacgc 900
 gccucagcga uuuuaaaucg uaugcagaau auucaauuug aagcaguggu uggccacaaa 960
 aucaaaauga aa 972

<210> 8

<211> 852

<212> DNA

<213> 智人(Homo sapiens)

<400> 8

augaacauc uccuagaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60
 ucguugguga aguuuuucau uccucagagg agaaaucug uggcugggga gauuguucuc 120
 auuacuggag cugggcaugg aaauaggcagg cagacuacuu augaauuugc aaaacgacag 180
 agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240
 cgaaaacuag gcgucacugc gcaugcguau gugguagacu gcagcaacag agaagagauc 300
 uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaauugu ggugaauau 360
 gcugggacag uauauccagc cgaucuucuc agcaccaagg augaagagau uaccaagaca 420
 uuugagguca acauccuagg acauuuuugg aucacaaaag cacuucucc aucgaugaug 480
 gagagaaauc auggccacau cgucacagug gcuucagugu gcggccacga agggauuccu 540
 uaccucaucc cauauuguuc cagcaaaauu gccgcuguug gcuuucacag aggucugaca 600
 ucagaacuuc aggccuuggg aaaaacuggu aucaaaaccu caugucucug cccaguuuuu 660
 gugaauacug gguucaccaa aaauccaagc acaagauuau ggccuguauu ggagacagau 720
 gaagucguaa gaagucugau agauggaaua cuuaccaaua agaaaugau uuuuguucca 780
 ucguauauca auaucuuucu gagacuacag aaguuaagua cagcacagaa caccacaaau 840
 cuaaaacacc aa 852

<210> 9

<211> 714

<212> DNA

<213> 智人(Homo sapiens)

<400> 9

augaacauca uccuagaaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60
 ucguugguga aguuuuucau uccucagagg agaaaucug uggcugggga gauuguucuc 120
 auuacuggag cugggcaugg aaauaggcagg cagacuacuu augaauuugc aaaacgacag 180
 agcauauugg uucuguggga uauuaauaag gugaagaaag aaguggguga uguaacaau 240
 guggugaaua augcugggac aguauaucca gccgaucuuc ucagcaccaa ggaugaagag 300
 auuaccaaga cauugaggu caacauccua ggacauuuuu ggaucacaaa agcacuucuu 360
 ccaucgauga uggagagaaa ucauggccac aucgucacag uggcuucagu gugcgccac 420
 gaagggaauuc cuuaccucau cccauauugu uccagcaauu uugccgcugu uggcuuucac 480
 agaggucuga caucagaacu ucaggccuug ggaaaaacug guaucaaaac cucaugucuc 540
 ugcccaguuu uugugaauac uggguucacc aaaaauccaa gcacaagauu auggccugua 600
 uuggagacag augaagucgu aagaagucug auagauggaa uacuuacca uagaaaaug 660
 auuuuuguuc caucguauau caauaucuuu cugagacuac agaagguuuc uucc 714

<210> 10

<211> 894

<212> DNA

<213> 智人(Homo sapiens)

<400> 10

augaacauca uccuagaaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60
 ucguugguga aguuuuucau uccucagagg agaaaucug uggcugggga gauuguucuc 120
 auuacuggag cugggcaugg aaauaggcagg cagacuacuu augaauuugc aaaacgacag 180
 agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240
 cgaaaacuag gcgucacugc gcaugcguau gugguagacu gcagcaacag agaagagauc 300
 uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaauugu ggugaauau 360
 gcugggacag uauauccagc cgaucuucuc agcaccaagg augaagagau uaccaagaca 420
 uuugagguca acauccuagg acauuuugg aauggaaagg acaucagaag uaauuacuug 480
 gauguauaua ggaucgagga cacuuuugga cgagacucug agaucacaaa agcacuucuu 540
 ccaucgauga uggagagaaa ucauggccac aucgucacag uggcuucagu gugcgccac 600
 gaagggaauuc cuuaccucau cccauauugu uccagcaauu uugccgcugu uggcuuucac 660
 agaggucuga caucagaacu ucaggccuug ggaaaaacug guaucaaaac cucaugucuc 720
 ugcccaguuu uugugaauac uggguucacc aaaaauccaa gcacaagauu auggccugua 780
 uuggagacag augaagucgu aagaagucug auagauggaa uacuuacca uagaaaaug 840
 auuuuuguuc caucguauau caauaucuuu cugagacuac agaagguuuc uucc 894

<210> 11

<211> 813

<212> DNA

<213> 智人(Homo sapiens)

<400> 11

augaacauca uccuagaaaau ccuucugcuu cugaucacca ucaucuacuc cuacuuggag 60

ucguugguga aguuuuucau uccucagagg agaaaaucug uggcugggga gauuguucuc 120
 auuacuggag cugggcaugg aaauaggcagg cagacuacuu augaauuugc aaaacgacag 180
 agcauauugg uucuguggga uauuaauaag cgcggugugg aggaaacugc agcugagugc 240
 cgaaaacuag gcgucacugc gcaugcguau gugguagacu gcagcaacag agaagagauc 300
 uaucgcucuc uaaaucaggu gaagaaagaa gugggugaug uaacaauugu ggugaauaau 360
 gcugggacag uauauccagc cgaucuuuc agcaccaagg augaagagau uaccaagaca 420
 uuugagguca acauccuagg acauuuuugg aucacaaaag cacuucuucc aucgaugaug 480
 gagagaaauc auggccacau cgucacagug gcuucagugu gcggccacga agggauuccu 540
 uaccucaucc cauauuguuc cagcaaaauu gccgcuguug gcuuucacag aggucugaca 600
 ucagaacuuc aggccuuggg aaaaacuggu aucaaaaccu caugucucug cccaguuuuu 660
 gugaauacug gguucaccaa aaauccaagc acaagauuau ggccuguauu ggagacagau 720
 gaagucguaa gaagucugau agauggaaua cuuaccaaua agaaaugau uuuuguucca 780
 ucguauauca auaucuuucu gagacuacag aag 813

<210> 12

<211> 900

<212> DNA

<213> 智人 (Homo sapiens)

<400> 12

atgaacatca tcctagaaat ctttctgctt ctgatcacca tcattctactc ctacttggag 60
 tcgttgggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120
 attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180
 agcatattgg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240
 cgaaaactag gcgtcactgc gcatgcgtat gtggttagact gcagcaacag agaagagatc 300
 tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360
 gctgggacag tatatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420
 tttgaggtca acatcctagg acatttttgg atcacaaaag cacttcttcc atcgatgatg 480
 gagagaaatc atggccacat cgtcacagtg gcttcagtgt gcggccacga agggattcct 540
 tacctcatcc catattgttc cagcaaattt gccgctgttg gctttcacag aggtctgaca 600
 tcagaacttc aggccttggg aaaaactggt atcaaaacct catgtctctg cccagttttt 660
 gtgaatactg ggttcaccaa aaatccaagc acaagattat ggctgtatt ggagacagat 720
 gaagtcgtaa gaagtctgat agatggaata cttaccaata agaaaatgat ttttgttcca 780
 tcgtatatca atatctttct gagactacag aagtttcttc ctgaacgcgc ctacgcgatt 840
 ttaaatcgta tgcagaatat tcaatttgaa gcagtgggtg gccacaaaat caaaatgaaa 900

<210> 13

<211> 792

<212> DNA

<213> 智人 (Homo sapiens)

<400> 13

atgaacatca tcctagaaat ctttctgctt ctgatcacca tcattctactc ctacttggag 60

tcgttgggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120
 attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180
 agcatattgg ttctgtggga tattaataag gtgaagaaag aagtgggtga tgtaacaatc 240
 gtgggtgaata atgctgggac agtatatcca gccgatcttc tcagcaccaa ggatgaagag 300
 attaccaaga catttgaggt caacatccta ggacattttt ggatcacaaa agcacttctt 360
 ccatcgatga tggagagaaa tcatggccac atcgtcacag tggcttcagt gtgcggccac 420
 gaagggattc cttacctcat cccatattgt tccagcaaat ttgccgctgt tggctttcac 480
 agaggtctga catcagaact tcaggccttg ggaaaaactg gtatcaaaac ctcatgtctc 540
 tgcccagttt ttgtgaatac tgggttcacc aaaaatccaa gcacaagatt atggcctgta 600
 ttggagacag atgaagtcgt aagaagtctg atagatggaa tacttaccaa taagaaaatg 660
 atttttgttc catcgtatat caatatcttt ctgagactac agaagtttct tcctgaacgc 720
 gcctcagcga ttttaaactg tatgcagaat attcaatttg aagcagtggg ttggccacaaa 780
 atcaaaatga aa 792

<210> 14

<211> 783

<212> DNA

<213> 智人(Homo sapiens)

<400> 14

atgaacatca tcctagaaat ctttctgctt ctgatcacca tcatctactc ctacttggag 60
 tcgttgggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120
 attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180
 agcatattgg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240
 cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 300
 tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360
 gctgggacag tataatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420
 tttagaggtca acatcctagg acatttttgg atcacaaaag cacttcttcc atcgatgatg 480
 gagagaaaatc atggccacat cgtcacagtg gcttcagtgt gcggccacga agggattcct 540
 tacctcatcc catattgttc cagcaaattt gccgctgttg gctttcacag aggtctgaca 600
 tcagaacttc aggccttggg aaaaactggg atcaaaacct catgtctctg cccagttttt 660
 gtgaataactg gggtcaccaa aaatccaagc acaaggtttc ttctgaacg cgcctcagcg 720
 attttaaatc gtatgcagaa tattcaattt gaagcagtgg ttggccacaa aatcaaaatg 780
 aaa 783

<210> 15

<211> 822

<212> DNA

<213> 智人(Homo sapiens)

<400> 15

atgaacatca tcctagaaat ctttctgctt ctgatcacca tcatctactc ctacttggag 60
 tcgttgggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120

attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180
 agcatatttg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240
 cgaaaactag gcgtcactgc gcatgcgtat gtggttagact gcagcaacag agaagagatc 300
 tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360
 gctgggacag tatatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420
 tttgagggtca acatcctagg acatttttgg atcacaaaag cacttcttcc atcgatgatg 480
 gagagaaatc atggccacat cgtcacagtg gcttcagtgt gcggccacga agggattcct 540
 tacctcatcc catattgttc cagcaaattt gccgctgttg gctttcacag aggtctgaca 600
 tcagaacttc aggccttggg aaaaactggg atcaaaacct catgtctctg cccagttttt 660
 gtgaatactg ggttcaccaa aaatccaagc acaagattat ggctgtatt ggagacagat 720
 gaagtcgtaa gaagtctgat agatggaata cttaccaata agaaaatgat ttttgttcca 780
 tcgtatatca atatctttct gagactacag aaggtttctt cc 822

<210> 16

<211> 972

<212> DNA

<213> 智人 (Homo sapiens)

<400> 16

atgaacatca tcctagaaat ctttctgctt ctgatcacca tcatctactc ctacttggag 60
 tcgttggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120
 attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180
 agcatatttg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240
 cgaaaactag gcgtcactgc gcatgcgtat gtggttagact gcagcaacag agaagagatc 300
 tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360
 gctgggacag tatatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420
 tttgagggtca acatcctagg acatttttgg aatggaaagg acatcagaag taattacttg 480
 gatgtatata ggatcgagga cacttttggg cgagactctg agatcacaaa agcacttctt 540
 ccatcgatga tggagagaaa tcatggccac atcgtcacag tggcttcagt gtgcggccac 600
 gaagggattc cttacctcat cccatattgt tccagcaaat ttgccgctgt tggctttcac 660
 agagggtctga catcagaact tcaggccttg ggaaaaactg gtatcaaaac ctcatgtctc 720
 tgcccagttt ttgtgaatac tgggttcacc aaaaatccaa gcacaagatt atggcctgta 780
 ttggagacag atgaagtcgt aagaagtctg atagatggaa tacttaccaa taagaaaatg 840
 atttttgttc catcgatat caatatcttt ctgagactac agaagtttct tcctgaacgc 900
 gcctcagcga ttttaaactg tatgcagaat attcaatttg aagcagtggg tggccacaaa 960
 atcaaaatga aa 972

<210> 17

<211> 852

<212> DNA

<213> 智人 (Homo sapiens)

<400> 17

atgaacatca tcctagaaat ctttctgctt ctgataccca tcattctactc ctacttggag 60
 tcgttgggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120
 attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180
 agcatatttg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240
 cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 300
 tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360
 gctgggacag tatatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420
 tttgagggtca acatcctagg acatttttgg atcacaaaag cacttcttcc atcgatgatg 480
 gagagaaatc atggccacat cgtcacagtg gcttcagtgt gcggccacga agggattcct 540
 tacctcatcc catattgttc cagcaaattt gccgctgttg gctttcacag aggtctgaca 600
 tcagaacttc aggccttggg aaaaactggg atcaaaacct catgtctctg cccagttttt 660
 gtgaatactg ggttcaccaa aaatccaagc acaagattat ggctgtatt ggagacagat 720
 gaagtcgtaa gaagtctgat agatggaata cttaccaata agaaaatgat ttttgttcca 780
 tcgtatatca atatctttct gagactacag aagttaagta cagcacagaa cacccaaata 840
 ctaaaacacc aa 852

<210> 18

<211> 714

<212> DNA

<213> 智人 (Homo sapiens)

<400> 18

atgaacatca tcctagaaat ctttctgctt ctgataccca tcattctactc ctacttggag 60
 tcgttgggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120
 attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180
 agcatatttg ttctgtggga tattaataag gtgaagaaag aagtgggtga tgtaacaatc 240
 gtgggtgaata atgctgggac agtatatcca gccgatcttc tcagcaccaa ggatgaagag 300
 attaccaaga catttgagggt caacatccta ggacattttt ggatcacaaa agcacttctt 360
 ccatcgatga tggagagaaa tcattggccac atcgtcacag tggcttcagt gtgcggccac 420
 gaagggattc cttacctcat cccatattgt tccagcaaat ttgccgctgt tggctttcac 480
 agagggtctga catcagaact tcaggccttg ggaaaaactg gtatcaaaac ctcatgtctc 540
 tgcccagttt ttgtgaatac tgggttcacc aaaaatccaa gcacaagatt atggcctgta 600
 ttggagacag atgaagtcgt aagaagtctg atagatggaa tacttaccaa taagaaaatg 660
 atttttgttc catcgatat caatatcttt ctgagactac agaaggtttc ttcc 714

<210> 19

<211> 894

<212> DNA

<213> 智人 (Homo sapiens)

<400> 19

atgaacatca tcctagaaat ctttctgctt ctgataccca tcattctactc ctacttggag 60
 tcgttgggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120

attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180
 agcatatttg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240
 cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 300
 tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360
 gctgggacag tatatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420
 tttgaggtca acatcctagg acatttttgg aatggaaagg acatcagaag taattacttg 480
 gatgtatata ggatcgagga cacttttggg cgagactctg agatcacaaa agcacttctt 540
 ccatcgatga tggagagaaa tcatggccac atcgtcacag tggcttcagt gtgcggccac 600
 gaagggattc cttacctcat cccatattgt tccagcaaat ttgccgctgt tggctttcac 660
 agaggtctga catcagaact tcaggccttg ggaaaaactg gtatcaaaac ctcatgtctc 720
 tgcccagttt ttgtgaatac tgggttcacc aaaaatccaa gcacaagatt atggcctgta 780
 ttggagacag atgaagtcgt aagaagtctg atagatggaa tacttaccaa taagaaaatg 840
 atttttgttc catcgtatat caatatcttt ctgagactac agaaggtttc ttcc 894

<210> 20

<211> 813

<212> DNA

<213> 智人(Homo sapiens)

<400> 20

atgaacatca tcctagaaat ctttctgctt ctgatcacca tcatctactc ctacttggag 60
 tcgttgggtga agtttttcat tcctcagagg agaaaatctg tggctgggga gattgttctc 120
 attactggag ctgggcatgg aataggcagg cagactactt atgaatttgc aaaacgacag 180
 agcatatttg ttctgtggga tattaataag cgcggtgtgg aggaaactgc agctgagtgc 240
 cgaaaactag gcgtcactgc gcatgcgtat gtggtagact gcagcaacag agaagagatc 300
 tatcgctctc taaatcaggt gaagaaagaa gtgggtgatg taacaatcgt ggtgaataat 360
 gctgggacag tatatccagc cgatcttctc agcaccaagg atgaagagat taccaagaca 420
 tttgaggtca acatcctagg acatttttgg atcacaaaag cacttcttcc atcgatgatg 480
 gagagaaaatc atggccacat cgtcacagtg gcttcagtgt gcggccacga agggattcct 540
 tacctcatcc catattgttc cagcaaattt gccgctgttg gctttcacag aggtctgaca 600
 tcagaacttc aggccttggg aaaaactggt atcaaaacct catgtctctg cccagttttt 660
 gtgaataactg ggttcaccaa aaatccaagc acaagattat ggctgtatt ggagacagat 720
 gaagtcgtaa gaagtctgat agatggaata cttaccaata agaaaatgat ttttgttcca 780
 tcgtatatca atatctttct gagactacag aag 813

<210> 21

<211> 300

<212> PRT

<213> 智人(Homo sapiens)

<400> 21

Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Ile Thr Ile Ile Tyr

1

5

10

15

Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys		
20	25	30
Ser Val Ala Gly Glu Ile Val Leu Ile Thr Gly Ala Gly His Gly Ile		
35	40	45
Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val		
50	55	60
Leu Trp Asp Ile Asn Lys Arg Gly Val Glu Glu Thr Ala Ala Glu Cys		
65	70	75
Arg Lys Leu Gly Val Thr Ala His Ala Tyr Val Val Asp Cys Ser Asn		
85	90	95
Arg Glu Glu Ile Tyr Arg Ser Leu Asn Gln Val Lys Lys Glu Val Gly		
100	105	110
Asp Val Thr Ile Val Val Asn Asn Ala Gly Thr Val Tyr Pro Ala Asp		
115	120	125
Leu Leu Ser Thr Lys Asp Glu Glu Ile Thr Lys Thr Phe Glu Val Asn		
130	135	140
Ile Leu Gly His Phe Trp Ile Thr Lys Ala Leu Leu Pro Ser Met Met		
145	150	155
Glu Arg Asn His Gly His Ile Val Thr Val Ala Ser Val Cys Gly His		
165	170	175
Glu Gly Ile Pro Tyr Leu Ile Pro Tyr Cys Ser Ser Lys Phe Ala Ala		
180	185	190
Val Gly Phe His Arg Gly Leu Thr Ser Glu Leu Gln Ala Leu Gly Lys		
195	200	205
Thr Gly Ile Lys Thr Ser Cys Leu Cys Pro Val Phe Val Asn Thr Gly		
210	215	220
Phe Thr Lys Asn Pro Ser Thr Arg Leu Trp Pro Val Leu Glu Thr Asp		
225	230	235
Glu Val Val Arg Ser Leu Ile Asp Gly Ile Leu Thr Asn Lys Lys Met		
245	250	255
Ile Phe Val Pro Ser Tyr Ile Asn Ile Phe Leu Arg Leu Gln Lys Phe		
260	265	270
Leu Pro Glu Arg Ala Ser Ala Ile Leu Asn Arg Met Gln Asn Ile Gln		
275	280	285
Phe Glu Ala Val Val Gly His Lys Ile Lys Met Lys		
290	295	300
<210> 22		
<211> 264		
<212> PRT		

<213> 智人 (Homo sapiens)

<400> 22

```

Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Ile Thr Ile Ile Tyr
1           5           10           15
Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys
          20           25           30
Ser Val Ala Gly Glu Ile Val Leu Ile Thr Gly Ala Gly His Gly Ile
          35           40           45
Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val
          50           55           60
Leu Trp Asp Ile Asn Lys Val Lys Lys Glu Val Gly Asp Val Thr Ile
65           70           75           80
Val Val Asn Asn Ala Gly Thr Val Tyr Pro Ala Asp Leu Leu Ser Thr
          85           90           95
Lys Asp Glu Glu Ile Thr Lys Thr Phe Glu Val Asn Ile Leu Gly His
          100          105          110
Phe Trp Ile Thr Lys Ala Leu Leu Pro Ser Met Met Glu Arg Asn His
          115          120          125
Gly His Ile Val Thr Val Ala Ser Val Cys Gly His Glu Gly Ile Pro
          130          135          140
Tyr Leu Ile Pro Tyr Cys Ser Ser Lys Phe Ala Ala Val Gly Phe His
145          150          155          160
Arg Gly Leu Thr Ser Glu Leu Gln Ala Leu Gly Lys Thr Gly Ile Lys
          165          170          175
Thr Ser Cys Leu Cys Pro Val Phe Val Asn Thr Gly Phe Thr Lys Asn
          180          185          190
Pro Ser Thr Arg Leu Trp Pro Val Leu Glu Thr Asp Glu Val Val Arg
          195          200          205
Ser Leu Ile Asp Gly Ile Leu Thr Asn Lys Lys Met Ile Phe Val Pro
          210          215          220
Ser Tyr Ile Asn Ile Phe Leu Arg Leu Gln Lys Phe Leu Pro Glu Arg
225          230          235          240
Ala Ser Ala Ile Leu Asn Arg Met Gln Asn Ile Gln Phe Glu Ala Val
          245          250          255
Val Gly His Lys Ile Lys Met Lys
          260

```

<210> 23

<211> 261

<212> PRT

<213> 智人 (Homo sapiens)

<400> 23

```

Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Ile Thr Ile Ile Tyr
1           5           10           15
Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys
           20           25           30
Ser Val Ala Gly Glu Ile Val Leu Ile Thr Gly Ala Gly His Gly Ile
           35           40           45
Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val
           50           55           60
Leu Trp Asp Ile Asn Lys Arg Gly Val Glu Glu Thr Ala Ala Glu Cys
65           70           75           80
Arg Lys Leu Gly Val Thr Ala His Ala Tyr Val Val Asp Cys Ser Asn
           85           90           95
Arg Glu Glu Ile Tyr Arg Ser Leu Asn Gln Val Lys Lys Glu Val Gly
           100          105          110
Asp Val Thr Ile Val Val Asn Asn Ala Gly Thr Val Tyr Pro Ala Asp
           115          120          125
Leu Leu Ser Thr Lys Asp Glu Glu Ile Thr Lys Thr Phe Glu Val Asn
           130          135          140
Ile Leu Gly His Phe Trp Ile Thr Lys Ala Leu Leu Pro Ser Met Met
145          150          155          160
Glu Arg Asn His Gly His Ile Val Thr Val Ala Ser Val Cys Gly His
           165          170          175
Glu Gly Ile Pro Tyr Leu Ile Pro Tyr Cys Ser Ser Lys Phe Ala Ala
           180          185          190
Val Gly Phe His Arg Gly Leu Thr Ser Glu Leu Gln Ala Leu Gly Lys
           195          200          205
Thr Gly Ile Lys Thr Ser Cys Leu Cys Pro Val Phe Val Asn Thr Gly
           210          215          220
Phe Thr Lys Asn Pro Ser Thr Arg Phe Leu Pro Glu Arg Ala Ser Ala
225          230          235          240
Ile Leu Asn Arg Met Gln Asn Ile Gln Phe Glu Ala Val Val Gly His
           245          250          255
Lys Ile Lys Met Lys
           260

```

<210> 24

<211> 274

<212> PRT

<213> 智人 (Homo sapiens)

<400> 24

```

Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Ile Thr Ile Ile Tyr
1           5           10           15
Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys
          20           25           30
Ser Val Ala Gly Glu Ile Val Leu Ile Thr Gly Ala Gly His Gly Ile
          35           40           45
Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val
          50           55           60
Leu Trp Asp Ile Asn Lys Arg Gly Val Glu Glu Thr Ala Ala Glu Cys
65           70           75           80
Arg Lys Leu Gly Val Thr Ala His Ala Tyr Val Val Asp Cys Ser Asn
          85           90           95
Arg Glu Glu Ile Tyr Arg Ser Leu Asn Gln Val Lys Lys Glu Val Gly
          100          105          110
Asp Val Thr Ile Val Val Asn Asn Ala Gly Thr Val Tyr Pro Ala Asp
          115          120          125
Leu Leu Ser Thr Lys Asp Glu Glu Ile Thr Lys Thr Phe Glu Val Asn
          130          135          140
Ile Leu Gly His Phe Trp Ile Thr Lys Ala Leu Leu Pro Ser Met Met
145          150          155          160
Glu Arg Asn His Gly His Ile Val Thr Val Ala Ser Val Cys Gly His
          165          170          175
Glu Gly Ile Pro Tyr Leu Ile Pro Tyr Cys Ser Ser Lys Phe Ala Ala
          180          185          190
Val Gly Phe His Arg Gly Leu Thr Ser Glu Leu Gln Ala Leu Gly Lys
          195          200          205
Thr Gly Ile Lys Thr Ser Cys Leu Cys Pro Val Phe Val Asn Thr Gly
          210          215          220
Phe Thr Lys Asn Pro Ser Thr Arg Leu Trp Pro Val Leu Glu Thr Asp
225          230          235          240
Glu Val Val Arg Ser Leu Ile Asp Gly Ile Leu Thr Asn Lys Lys Met
          245          250          255
Ile Phe Val Pro Ser Tyr Ile Asn Ile Phe Leu Arg Leu Gln Lys Val
          260          265          270

```

Ser Ser

<210> 25

<211> 324

<212> PRT

<213> 智人(Homo sapiens)

<400> 25

```

Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Leu Ile Thr Ile Ile Tyr
1           5           10           15
Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys
          20           25           30
Ser Val Ala Gly Glu Ile Val Leu Ile Thr Gly Ala Gly His Gly Ile
          35           40           45
Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val
          50           55           60
Leu Trp Asp Ile Asn Lys Arg Gly Val Glu Glu Thr Ala Ala Glu Cys
65           70           75           80
Arg Lys Leu Gly Val Thr Ala His Ala Tyr Val Val Asp Cys Ser Asn
          85           90           95
Arg Glu Glu Ile Tyr Arg Ser Leu Asn Gln Val Lys Lys Glu Val Gly
          100          105          110
Asp Val Thr Ile Val Val Asn Asn Ala Gly Thr Val Tyr Pro Ala Asp
          115          120          125
Leu Leu Ser Thr Lys Asp Glu Glu Ile Thr Lys Thr Phe Glu Val Asn
          130          135          140
Ile Leu Gly His Phe Trp Asn Gly Lys Asp Ile Arg Ser Asn Tyr Leu
145          150          155          160
Asp Val Tyr Arg Ile Glu Asp Thr Phe Gly Arg Asp Ser Glu Ile Thr
          165          170          175
Lys Ala Leu Leu Pro Ser Met Met Glu Arg Asn His Gly His Ile Val
          180          185          190
Thr Val Ala Ser Val Cys Gly His Glu Gly Ile Pro Tyr Leu Ile Pro
          195          200          205
Tyr Cys Ser Ser Lys Phe Ala Ala Val Gly Phe His Arg Gly Leu Thr
          210          215          220
Ser Glu Leu Gln Ala Leu Gly Lys Thr Gly Ile Lys Thr Ser Cys Leu
225          230          235          240
Cys Pro Val Phe Val Asn Thr Gly Phe Thr Lys Asn Pro Ser Thr Arg
          245          250          255
Leu Trp Pro Val Leu Glu Thr Asp Glu Val Val Arg Ser Leu Ile Asp
          260          265          270
Gly Ile Leu Thr Asn Lys Lys Met Ile Phe Val Pro Ser Tyr Ile Asn
          275          280          285

```

Ile Phe Leu Arg Leu Gln Lys Phe Leu Pro Glu Arg Ala Ser Ala Ile
 290 295 300
 Leu Asn Arg Met Gln Asn Ile Gln Phe Glu Ala Val Val Gly His Lys
 305 310 315 320
 Ile Lys Met Lys
 <210> 26
 <211> 284
 <212> PRT
 <213> 智人 (Homo sapiens)
 <400> 26
 Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Ile Thr Ile Ile Tyr
 1 5 10 15
 Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys
 20 25 30
 Ser Val Ala Gly Glu Ile Val Leu Ile Thr Gly Ala Gly His Gly Ile
 35 40 45
 Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val
 50 55 60
 Leu Trp Asp Ile Asn Lys Arg Gly Val Glu Glu Thr Ala Ala Glu Cys
 65 70 75 80
 Arg Lys Leu Gly Val Thr Ala His Ala Tyr Val Val Asp Cys Ser Asn
 85 90 95
 Arg Glu Glu Ile Tyr Arg Ser Leu Asn Gln Val Lys Lys Glu Val Gly
 100 105 110
 Asp Val Thr Ile Val Val Asn Asn Ala Gly Thr Val Tyr Pro Ala Asp
 115 120 125
 Leu Leu Ser Thr Lys Asp Glu Glu Ile Thr Lys Thr Phe Glu Val Asn
 130 135 140
 Ile Leu Gly His Phe Trp Ile Thr Lys Ala Leu Leu Pro Ser Met Met
 145 150 155 160
 Glu Arg Asn His Gly His Ile Val Thr Val Ala Ser Val Cys Gly His
 165 170 175
 Glu Gly Ile Pro Tyr Leu Ile Pro Tyr Cys Ser Ser Lys Phe Ala Ala
 180 185 190
 Val Gly Phe His Arg Gly Leu Thr Ser Glu Leu Gln Ala Leu Gly Lys
 195 200 205
 Thr Gly Ile Lys Thr Ser Cys Leu Cys Pro Val Phe Val Asn Thr Gly
 210 215 220
 Phe Thr Lys Asn Pro Ser Thr Arg Leu Trp Pro Val Leu Glu Thr Asp

225	230	235	240
Glu Val Val Arg Ser Leu Ile Asp Gly Ile Leu Thr Asn Lys Lys Met			
	245	250	255
Ile Phe Val Pro Ser Tyr Ile Asn Ile Phe Leu Arg Leu Gln Lys Leu			
	260	265	270
Ser Thr Ala Gln Asn Thr Gln Ile Leu Lys His Gln			
	275	280	
<210> 27			
<211> 238			
<212> PRT			
<213> 智人 (Homo sapiens)			
<400> 27			
Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Leu Ile Thr Ile Ile Tyr			
1	5	10	15
Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys			
	20	25	30
Ser Val Ala Gly Glu Ile Val Leu Ile Thr Gly Ala Gly His Gly Ile			
	35	40	45
Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val			
	50	55	60
Leu Trp Asp Ile Asn Lys Val Lys Lys Glu Val Gly Asp Val Thr Ile			
65	70	75	80
Val Val Asn Asn Ala Gly Thr Val Tyr Pro Ala Asp Leu Leu Ser Thr			
	85	90	95
Lys Asp Glu Glu Ile Thr Lys Thr Phe Glu Val Asn Ile Leu Gly His			
	100	105	110
Phe Trp Ile Thr Lys Ala Leu Leu Pro Ser Met Met Glu Arg Asn His			
	115	120	125
Gly His Ile Val Thr Val Ala Ser Val Cys Gly His Glu Gly Ile Pro			
	130	135	140
Tyr Leu Ile Pro Tyr Cys Ser Ser Lys Phe Ala Ala Val Gly Phe His			
145	150	155	160
Arg Gly Leu Thr Ser Glu Leu Gln Ala Leu Gly Lys Thr Gly Ile Lys			
	165	170	175
Thr Ser Cys Leu Cys Pro Val Phe Val Asn Thr Gly Phe Thr Lys Asn			
	180	185	190
Pro Ser Thr Arg Leu Trp Pro Val Leu Glu Thr Asp Glu Val Val Arg			
	195	200	205
Ser Leu Ile Asp Gly Ile Leu Thr Asn Lys Lys Met Ile Phe Val Pro			

210	215	220
Ser Tyr Ile Asn Ile Phe Leu Arg Leu Gln Lys Val Ser Ser		
225	230	235
<210> 28		
<211> 298		
<212> PRT		
<213> 智人 (Homo sapiens)		
<400> 28		
Met Asn Ile Ile Leu Glu Ile Leu Leu Leu Leu Ile Thr Ile Ile Tyr		
1	5	10
Ser Tyr Leu Glu Ser Leu Val Lys Phe Phe Ile Pro Gln Arg Arg Lys		
20	25	30
Ser Val Ala Gly Glu Ile Val Leu Ile Thr Gly Ala Gly His Gly Ile		
35	40	45
Gly Arg Gln Thr Thr Tyr Glu Phe Ala Lys Arg Gln Ser Ile Leu Val		
50	55	60
Leu Trp Asp Ile Asn Lys Arg Gly Val Glu Glu Thr Ala Ala Glu Cys		
65	70	75
Arg Lys Leu Gly Val Thr Ala His Ala Tyr Val Val Asp Cys Ser Asn		
85	90	95
Arg Glu Glu Ile Tyr Arg Ser Leu Asn Gln Val Lys Lys Glu Val Gly		
100	105	110
Asp Val Thr Ile Val Val Asn Asn Ala Gly Thr Val Tyr Pro Ala Asp		
115	120	125
Leu Leu Ser Thr Lys Asp Glu Glu Ile Thr Lys Thr Phe Glu Val Asn		
130	135	140
Ile Leu Gly His Phe Trp Asn Gly Lys Asp Ile Arg Ser Asn Tyr Leu		
145	150	155
Asp Val Tyr Arg Ile Glu Asp Thr Phe Gly Arg Asp Ser Glu Ile Thr		
165	170	175
Lys Ala Leu Leu Pro Ser Met Met Glu Arg Asn His Gly His Ile Val		
180	185	190
Thr Val Ala Ser Val Cys Gly His Glu Gly Ile Pro Tyr Leu Ile Pro		
195	200	205
Tyr Cys Ser Ser Lys Phe Ala Ala Val Gly Phe His Arg Gly Leu Thr		
210	215	220
Ser Glu Leu Gln Ala Leu Gly Lys Thr Gly Ile Lys Thr Ser Cys Leu		
225	230	235
Cys Pro Val Phe Val Asn Thr Gly Phe Thr Lys Asn Pro Ser Thr Arg		

	245		250		255
Leu Trp Pro Val	Leu Glu Thr Asp	Glu Val Val Arg	Ser Leu Ile Asp		
	260		265		270
Gly Ile Leu Thr	Asn Lys Lys Met	Ile Phe Val Pro	Ser Tyr Ile Asn		
	275		280		285
Ile Phe Leu Arg	Leu Gln Lys Val	Ser Ser			
	290		295		
<210>	29				
<211>	271				
<212>	PRT				
<213>	智人 (Homo sapiens)				
<400>	29				
Met Asn Ile Ile	Leu Glu Ile Leu	Leu Leu Ile Thr	Ile Ile Tyr		
1	5	10	15		
Ser Tyr Leu Glu	Ser Leu Val Lys	Phe Phe Ile Pro	Gln Arg Arg Lys		
	20	25	30		
Ser Val Ala Gly	Glu Ile Val Leu	Ile Thr Gly Ala	Gly His Gly Ile		
	35	40	45		
Gly Arg Gln Thr	Thr Tyr Glu Phe	Ala Lys Arg Gln	Ser Ile Leu Val		
	50	55	60		
Leu Trp Asp Ile	Asn Lys Arg Gly	Val Glu Glu Thr	Ala Ala Glu Cys		
65	70	75	80		
Arg Lys Leu Gly	Val Thr Ala His	Ala Tyr Val Val	Asp Cys Ser Asn		
	85	90	95		
Arg Glu Glu Ile	Tyr Arg Ser Leu	Asn Gln Val Lys	Lys Glu Val Gly		
	100	105	110		
Asp Val Thr Ile	Val Val Asn Asn	Ala Gly Thr Val	Tyr Pro Ala Asp		
	115	120	125		
Leu Leu Ser Thr	Lys Asp Glu Glu	Ile Thr Lys Thr	Phe Glu Val Asn		
	130	135	140		
Ile Leu Gly His	Phe Trp Ile Thr	Lys Ala Leu Leu	Pro Ser Met Met		
145	150	155	160		
Glu Arg Asn His	Gly His Ile Val	Thr Val Ala Ser	Val Cys Gly His		
	165	170	175		
Glu Gly Ile Pro	Tyr Leu Ile Pro	Tyr Cys Ser Ser	Lys Phe Ala Ala		
	180	185	190		
Val Gly Phe His	Arg Gly Leu Thr	Ser Glu Leu Gln	Ala Leu Gly Lys		
	195	200	205		
Thr Gly Ile Lys	Thr Ser Cys Leu	Cys Pro Val Phe	Val Asn Thr Gly		

210	215	220
Phe Thr Lys Asn Pro Ser Thr Arg Leu Trp Pro Val Leu Glu Thr Asp		
225	230	235
Glu Val Val Arg Ser Leu Ile Asp Gly Ile Leu Thr Asn Lys Lys Met		
245	250	255
Ile Phe Val Pro Ser Tyr Ile Asn Ile Phe Leu Arg Leu Gln Lys		
260	265	270

<210> 30

<211> 23830

<212> DNA

<213> 智人 (Homo sapiens)

<400> 30

```

atgggtccgag gggggcgggg ctgacgtcgc gctgggaatg ccctggccga gacactgagg 60
cagggtagag agcgcttgcg ggcgccgggc ggagctgctg cggatcagga cccgagccga 120
ttcccgatcc cgaccagat cctaaccgcg gccccgcgcc cgccgccgcc gccatgtacg 180
acgcagagcg cggctggagc ttgtccttcg cgggctgcgg cttcctgggc ttctaccacg 240
tcggggcgac ccgtgcctg agcgagcacg cccgcacct cctccgcgac gcgcgcatgt 300
tgttcggcgc ttcggccggg gcgttgcaact gcgtcggcgt cctctccggt atcccgttg 360
gtgcgtctgg ggacgtgcc cgggctccac gtgcggagtg ggtgccccct aggccgggga 420
gcgggggatc cccaggggtc gcggggccct ggaggagcgg gcatcggacg cggacacggc 480
gggggtgcac ccgagggccc cctccgaggc agatgcttcc tgcgggggcg ctgttctctg 540
gccccggaag ggggcgttgg aaccccgagc ggtccgggcc gaagcctggg actctcgtgc 600
gtccccaccc ctacccccat caggcgcccg tgcatgaagg gagaccctca cctccggact 660
gagagtcgga gcgtctcgga gcgacgggga gtagggagcg ggaccggggg cggagggtag 720
tgctggcccc tgcggactcc ggggtcccctg tgtcctctcg ggaggggctg gacgggctga 780
gctgccgagg ggccgatttg ccctgggccc gacaaagagt ggggctttgg ccggtcccc 840
acgggtgggct ccttccctct ggggattgag ggactcaaga caccgcgcgc ctgcgctttt 900
cttttctttt tttctttttt tttttttgag acggagtttc gctcagtcgc ccaggctgga 960
gtgcagtggc gtgatctcaa ctcaactgaa gctccacetc ccaggttcac gccattctcc 1020
tgctcagcc tcccagtag ctgggactac aggcgccagc caccaagccc ggctaatttt 1080
ttgtattttt tagtagagac ggggtttcac cgtgttagcc aggatggtct cgatctcctg 1140
acctcgtgat ctgccacct cggcctccca gaatgctggg gttacaggcg tgagccactg 1200
ctccctgctg cctacgctct ctgggtcgca gccagcctt ctgggggctg ggtagcctcc 1260
cagaagggca accctgggca tcctccaggg caggctaact ggagtctagt ggggaggggt 1320
acctgaaaag aggaaagtgt tttctctctc ctctctctcc tccagtgttt gggacccttc 1380
ctgggggctg gagtgcaccc ctggacaccc cccaatccca tcctcttctc tagtttccac 1440
tgacctaggc ccaccctccc ctctccggtc cagtactcct ggaaatgaga ttccgtacat 1500
ttgaatcttg tcctaataaa atatttgtcc atgtgggtac ctgtgtgtgt gtggtggggg 1560
tgcagacgga gggtttgttt ctactagct ggaactactg ggggtgtgta tgcttctctg 1620

```

gaatttgtgt gccacagtcc tggaggcgag gagggggttg tgagccagta ggcaggggct 1680
 ggggcaagta gcattgtgaa gctattgaca cccagacgtc cccaggcagg agattatgcc 1740
 cccattagcc cccttttatc tgggcttcct taacaatgga ctctttgccc tgcctgccag 1800
 agccagcagg gagtgactgt tcagtgggtga ggaagcgggc agaggaagcc ctgccattgg 1860
 gtaggagcag tgggcagccc ctgggctgac tgggaggtgg ggattaggga ttagacagtc 1920
 ctggctgtct gccttccccct aagccagggg gagaggagca aagggcacga aatgtggcct 1980
 ccaggaggat tagaccgcca catgatcatt tgcacaccct ggggttttagc aacaataaaa 2040
 gtcagctttt ttgtatccca aggtggcctg tggacacca catggacaaa tgtttacact 2100
 gggacagaat tcaaatgcag aggtcccagg agcctaaagt aactcactc tggatataga 2160
 aggattcctt actgggcaga ggacaggtgc agcctggggc tttcccaggc aggacacagg 2220
 gaggtcagg aaccaccaag tccttgggaag gtggatctgg aggcgttggc aggagccact 2280
 ccctgggttc cagggtcca ggttcttct ttaacccct gtctcacaga gggctgtgca 2340
 cttgggggct gctgagcatg tcccagaggc tgcctcctgg acacagcacc tcagtgcac 2400
 tgagctgagg ctaacttggc aggagggaca ggcagaacct gccagccacg tgcaattcca 2460
 cccctctggc cactcaggga aggagagctg tgagtcaaga tcagatttgg gtcaggacag 2520
 gctggggcct gcctgtccct gtgcatcca agatttatgg ctggccaggg gttgggctgg 2580
 gaggggtggt cttgcatgcc aggagagtgc agatcagcct gagaggccag gccagtaagt 2640
 gaggtcagat ctctgcacc tgatagcatt aaggccatct acaccaaagc tctaattctg 2700
 atatgttcct ggctctatg tggggcatgg aggtggggca tggaggtgag gcctgtctgc 2760
 ctgggcttct ggaagtggga gactcattcc tgtggctgag gcctacagca gtgctgtgtg 2820
 gtaggaatac actggaagcc atgatgtcat tgtgcatttt ctagaagcca cattgaataa 2880
 agtaaaagac acaggtagaa ttaatttcat tgagcccaat atatccaaa taatatcatt 2940
 ttcacatcta ttcaatataa aaatttacta atgagatatt tcatactaag ccactgaaat 3000
 ccagtttgta tcttacacat ctgagttttg acgagccaca tttcaagggc gtgatagcca 3060
 catgtggctc ccatagtaga cagtactggg ctagagaaat gttggtggca tccttgctgt 3120
 ctggtttctg gccttgccaa aagtattacc atcccagtg ggtacattct ttcattgtatt 3180
 tgtctcctgt cccagagca gactctgcag gtctctcag atcttgtgcg gaaggccagg 3240
 agtcggaaca ttggcatctt ccatccatcc ttcaacttaa gcaagttcct ccgacagggt 3300
 ctctgcaaat gcctcccggc caatgtccac cagctcatct ccggcaaaat aggcatctct 3360
 cttaccagag tgtctgatgg ggaaaacgtt ctgggtgtctg actttcggtc caaagacgaa 3420
 gtcgtggatg taagcagttt gcttatctgg acgttgtcaa gttagaaaag ctgttttggg 3480
 atgggtgtgg tggctcatgc ctgtcatccc ggcactttgg gaggccgaag cgggtgggtt 3540
 gcttgagccc aggagctcga gaccaacatg atgaaacca gtctctacaa aaattacaga 3600
 aaaattagct aggcattggt ttgtgggccc atagtccag ctactaggga ggctgaggca 3660
 ggagaattgc ttgagcctgg gaggtggagg ttgcagtaag tcatgatcat gccactgtac 3720
 tccagcccgg gtgacagtga gatgtgtct ggaaaaaaa aaaaaagaaa gactgttttg 3780
 ttttggaagc aacacaggca gttgtaggcc ccctgtgcca gagtacata aactctgtac 3840
 acctccagtg atttggtcca tgtttgtaaa ccctgaatgt tccagggcag tttcttttct 3900
 tcaactttta tctctttttt ttgggtgggg gggcggggta cagagtcttg ctctgtctcc 3960

caggctggag tgcagtggcg caatctcaac ctcccaggga gctgggacta caggcacagg 4020
 ccatcacacc ttgctaattgt ttgtactttt tgtagagacg gggttttgcc ctgttgccca 4080
 ggctgggtccc aaactcctgc acccaagtaa tctgcccacc tctgcctggc agttacaatt 4140
 tcaataaatt cctccctttc cttcaacact tggctcatga ccgtccagtc caaggaacct 4200
 gtcctgcagg tgtgcctctc ccgagcttcc tctatgcatc ttccataatg aagatgcctt 4260
 ctcactggaa accctacaag ggtgggaacg tgccttattt gcctgtatcc tcagggtcta 4320
 gcagagagaa gataatttgt aataccaaaa caccattaaa ttcagctgat gctttcataa 4380
 gcgctccttg gaggaaggac tccatttact tgacagatct gtgcaagaca gcagcctggc 4440
 gcgtctaacc tgcagccagt tgcatectct gtttaacctt gtttgcgga gctttctcta 4500
 aacagccagc acttgtctgt tcccacatgg gtccgttctc ccagtgaatc accgtggtgc 4560
 ctgctgactg ctctgtagca cagtgttcg caaagtgtga tcctgggacc agcagagcag 4620
 cagctccttt gagcttattg gaatggcaga cctcaggtc ccacctctga cctgctgcat 4680
 gggaattctg gggagggacg cagaatctct ggttccacag gctctccggt gatgctaatt 4740
 aataccggca tttgaacagc accgatctag cccctttcag tccatgagcc aacaacctt 4800
 ggtcctgtct gtggtgaccc agtgtgactc tcatggggag caaggagagg aagttgaagt 4860
 tcaactgacag ggttgtaag gggattatgc aatagatgag acccatgggc ctgaagtccg 4920
 aggggtgatg ttagttcccc gttcttttga cccatggatt aacctactct gtgcaaagg 4980
 cattttcaag tttgttgccc tgctcacttg gagaaagctt atgaaggatc aggaaaatta 5040
 aaagggtgct ctgcctata acttctctct ctttgcctt cacaggcctt ggtatgttcc 5100
 tgcttcatcc ccttctacag tggccttacc cctccttct tcagaggcgt ggtaagtcgg 5160
 ctttctctgc tagcgtgag tcctgggggc ctctgaagtg tgctcacaca tctcctgcct 5220
 gcagggcact ggtgtcgggc acctcagggt ctgtcccatg gtggagcccc atgcctcact 5280
 gcctttcaga cagagtagcc acagctggcc ctatttccag gctaccggg cagcaaaact 5340
 tactgcatgt gtaattaatt atttggtat ctgtaaggta aactggctgg ttactttaat 5400
 ctgcacctta agcatcagat agcttctcag tgatctagtt aaactatatg atgttgcca 5460
 ggcgcggtgg ctcatgtctg taatcccagc actttgggag cctgaagcag gcagatcact 5520
 tgaggtcagg agttcgagac cagcctggcc aacagtgtga aactctgtct ctctaaaaa 5580
 tacaaaaatt agctgggcat ggtggtgtgc acctgtaatc ccagctgctc gggaggctga 5640
 ggcaggagaa ttgcttgaac ttgggaggcg gaagttgcag tgagccaaga tcgcaccact 5700
 gcactccatc ctgggtgaca gagcgagact ctatctcaaa aagaaaaaaa aaaaaagg 5760
 aaataaagta tatgacactg aagaatctgt taccctgga aggtggagct ttactcttag 5820
 ggggaactat aacagtcata tatatatatt ttttctttt cttttttttt ttttttgaga 5880
 tggagtctgg ctctgtcgcc caggtggag tgcagtgggt caatctcggc tcaactgcaac 5940
 ctccacctca caggttcagg caattctct gcctcaacct cccgagtagc tgggattaca 6000
 ggtgcctgcc gtcacgcaa gctaattttt gtatttttag tagagacagg gtttcatcat 6060
 attggccagg ctggtctcca actcctgacc tcaggtgatc cggccgctt ggctcccaa 6120
 agtgctgaga ttacaggcgt gagccatggt gcccgccaa caatcacatg tgttgtaaac 6180
 aacaacaaaa atctgtcagc ctggtctaac ctagatttgt gctttgtttt gttttgccac 6240
 tttgtgatgc acaggaggaa gtttaggctg taaaatacta gccttttagg gtaatttttg 6300

aactcacaag agcagcagcg gaacctttga tgcaatcctg tatgtagcac cagcagagcc 6360
acgtggcaga gggactcgca ttaggagcct cccattacag actacgtgct cctgtgcgtt 6420
atcttatagg gtccccacaa ccaaggggag atgtgattat tcatcctgtg tggctgtggg 6480
gaacttgaga gtcatacttg cccaaagagc acggccagcg agcttgacc caggtcactc 6540
tctgtcctc tgtcagaaca gggcatgtct tggttcactg cagggcggct cttctcattc 6600
tctgtagttt ggggtccagg atagtggctc acggagccac tggagtggcc agctactgag 6660
tgaccaaagc atatttttga tttccgacat tgccacagca tggttgggca tcagcaggac 6720
cccaacccct tgttatgtcg gtggctttat gtggttattt gatcttcccc agaactcagc 6780
aggagtgcac ccagcagcac cgtagtgatg ctctctggct cccagtgca cggttctggc 6840
tttccttcct ggtcgagagt ttcaagccct ctgggtccta ctctgtcctt ttcagcccat 6900
agctttgttc aaaagctgct ggtagtgct agatttggt gagttcagtg aatatgtgca 6960
ttggctgatt tctgagccat gccaggggga tggagaagcc gaagcaggag tgtttgttct 7020
gcaggctctg gagtaggcat tgggtctgtg ccggtcact tgctagtctt gcatccttc 7080
ccaacccct ctggggatgt ctggccacat cagaagacag tttgggttgt cagaactggg 7140
ggagtaccag gccgaggtgg gtggatcatg aggtcaggag atcgagacca tcctggctaa 7200
cacagtgaac cctcatctct actaaacata cgaaaaaat tagctgggcg tgggtggcggg 7260
cgctgtagt cccagctact cgggaggtg aggcaggaga atggtgtgaa cccggggggc 7320
ggagcttgca gtgagctgag atcctgccac tgcactccag cctgggcaac aaagcgagac 7380
tcctgtctac aaacaaaaca aaacaaaaca aaacaaaatc tgggggagtg ccactggcat 7440
ctgatgtata gaggcccgag atgctgtgtc atcacccgtt gagtgcgtc ataggcatct 7500
tcctgacaat tagaaccat tattcttcaa attcaatgca agcaaattca aagcattact 7560
gtgtacatac cgcatgctaa tcaattgcac cactggagct cctaaattca aaacattact 7620
ataaaaaagt tcaaaatgca tggaaaagt gtacatggca ggagaatatt tgggcttctg 7680
actacccctt gaatgaagat gatccaccag ccgccttct ccttggtctt cactccagat 7740
tcctagcatt tcattctgtg tctctttatg cagtgaggtt tttgtttgtt ttttgagaca 7800
gagtctcact gtatcaccta ggcctggagt gcagtggcg gatctcagct cactgcaacc 7860
ctcggctcct gggtttaagc gattctcctg cctcagcctc ccgagcagct gagattacaa 7920
gcacacatcc ccatgcccag ctaatttttg tatttttagc agagacaggg tttcaccatg 7980
ttgcccaggc tggctctgaa ctctggcct caagtgatcc atgtgcctca gccttccaaa 8040
gtgctgggat tacaggcgtg agccaccatg cccagctcct agtgaggtt ttgatgcctt 8100
gctacatctg ccctagaaat tgtgtgacta cgattttgga aatgttgctg tgtaaacttg 8160
tgatcatttc tggactccag gcaagaatct tgatggctaa ggtgtggctg aacatgtctg 8220
attctctcct ggacctgtt taggccaac tctgctctga aattcctccg tgtggaaggg 8280
cgggtgggg agagcctccc agctggaatc ttttgatgc ctttctctgt gggatatctga 8340
tggctggctc tgatggctgg ctgtgatggc tgtggctgga aatcattgtt gacatgagtt 8400
tcacagatgc aggtctgtc caaattgtag caaaagctgc ctgccccagc cgagctatgg 8460
gcaataaggt ggtttaagga tatagatgaa ggaaaactca cccttagaat aatttatcca 8520
aaatgctgct gtgttggtgg ttagaggaca ttttctgagg tcccaggttc attgtttcat 8580
ttaagtctca aaagtccctc caggtgttgg ttctaattgt caaagcatgg ggggagatgg 8640

```

gctcatgggt taaaggctctt atcccagatt tctgtatcct ccttgcaagc agcaaagggg 8700
tctggatttg aatccatgac catgtttctc ctttgggttt ccatcacact ctgtccccgt 8760
gcactgagca ccctttagtt catatgaccc ccttaggcat gttacatggg cactcctata 8820
gggtgcccac tggccctagg acttggccaa cacaacatgg actccagttt ccatctgcct 8880
ctttgccagg cacttttgtg cagtgcacac actgtacaac agtagacggc aaccctgaga 8940
gccagagtag agcctgtcct agcaccggaa tgctcggtaa ggatttgtcg caggagtgat 9000
tccaaagcca atgtcctccc tccatatcag cctgtttgtg gctctgagaa gctctgcccc 9060
catgtgaaaag cttgttaagc acttaagcac taaccagag cttcagacag tgccagtcct 9120
ttttccccctt ctttaaaaagc gatatgtgga tggaggagtg agtgacaacg tacccttcat 9180
tgatgccaaa acaaccatca ccgtgtcccc cttctatggg gagtacgaca tctgccctaa 9240
agtcaagtcc acgaactttc ttcatgtgga catcaccaag ctcagtctac gcctctgcac 9300
agggaacctc taccttctct cgagagcttt tgtcccccg gatctcaagg tgagttggtg 9360
gtgagggggc aggtgttctg ggggtgcagt cttctttgcc tccctgattg ccaggagcta 9420
ccagttactg tctgcacaat caaacagaaa tagacctgtc cttgatggtt aacggaaata 9480
aaaggcgctt gtcccagaag ctcaggtgag gcaccacct gattatggga atcacctggg 9540
aacatatacc cagacctaaa actcagatcc acttcccagg ctgtggttat atagtcaggg 9600
gggtgcagta tgggtattag gattttttat tttttagtta taaagatttt tttttggttt 9660
gtttttgaga cagggtcttg ctctgccgct taggctggag tgcagtggg caatcatagc 9720
tactgaagc ctcagactcc tgggttcaag cagtcctccc acctcagcct cctaaggagc 9780
tgggaccac aggcatgcag caccacacct ggctaatttt taaaaatttt gtggagtgtt 9840
gccaggctg gtctcacact cctggcctca agcgatctc ccaccccagc ctcccaatgt 9900
gttgggatta caggcatgag ccattgtacc cagccactaa gatgattctt atttgaaac 9960
acggctcaaga acaactgcgt tcggtagttt aacctttttt gattgtggtg gtttttagtat 10020
gccttaccac tctaccatag taagaaattt gcagaccatg tacaccaacc tttggtgctc 10080
ctggggagaa agaaagaagg ctatgcaatg caatgcatgc tcacagtcca aggagaggg 10140
aaagctgtct aacaggattg gttttccgt gtgctttata agcagatgag tagaggagac 10200
agctcttatt gtcctagtgg caattgggat aggctgcaaa gtttgttagg gtggaggctt 10260
attccgggac caagggagcc caaagaaaca agctcctgcc aggcgcggtg gctcacgcct 10320
gtaatcccag cactttggga ggctgaggca ggtggatcac ctgaggtcag gagtttgaga 10380
ccagcctggc caacatggtg aaacccctc tccatgaaaa atacaaaaat taccgggca 10440
tggtggcggg cacctgtaat ccagctact agggaggctg aggcaggaaa atggcttgaa 10500
cctcggaagt ggaggtggcc gttagccgag atcacgccac tgcactccag cctgggcaac 10560
agagcaagac tctgccttaa aaaaaaaaaa aaaaaaaga aaagtaaaag gaaaaaaaaa 10620
aggctctggc ctgctggggt gcctgcaaag tctccgtgga agggtgacat tcaagccgag 10680
acctccaggg aactgtctcc tgggagcaca gagccctttg ctcagcccc aggtggtcct 10740
gtgccccag ccagcagact cagagcttgc atgattcttt ggtgctctct gcggtcttcc 10800
aatgatgctg aaataaatgg tgcttggtgt ctccctgctg tagtccccct gcttgctttg 10860
ctcacagggt ctgggagaga tatgccttcg aggatatttg gatgcattca gtttcttgga 10920
agagaagggt atgtatgggc tgggaggatc agccatgccc ttttgacaag catttactag 10980

```

cggctcttgggt aaagacttga gatttgcctt agttctaaca cttagtgtccc aacgccttcc 11040
 ttgtgttgct caacctactc atgagcccag gagataggaa atctccgtcc cattgtacag 11100
 atggggaaac agaatttttg aaaggagagc caagcagcac acaccctcc ctgaggggca 11160
 gagccgagat ttgaactggg atgtcatgac tccagggccc tctccctccc cagggtcccc 11220
 ttatctgaag gcggtttttc tttccagctc gacctcttgt gacccttagt ttaacaagg 11280
 ccgaagttaa agagtttctg cgcctggacc ccaaataag caatcagatt tctcatctcc 11340
 agtcagggtg ggggtccaag ccactagaca agtttgctct tcccagagca catttctgcc 11400
 ttcaagtcac cctggcttgt cagggctggg ggagttctgc tgtagaaata ttagagtga 11460
 aggaaaaaga tgtgttgga gctatttttc tttataacta aaagttgggt gatgaatttg 11520
 tcgttgcca agaccaagga gactgcattt ttaaggacat atgtgtattt atctgctcag 11580
 aaaatgttca ttgctgtgtg ctagggatac tgcagtgaac acagaggtgt gacccttgcc 11640
 agccttgtga gagaagtga cagataagta agcagaagg tgatgctgtg tcgatggga 11700
 agtacagggt ccaatgagaa ggccaggtg tcaaggagaa gacacaggat gctggaggct 11760
 catgcaggat ggatctccaa ggcccagggg aagaagggcc tctcggagga cgtgaatcca 11820
 cattaagact ttggggataa gtaggagcgc ctaggcatg gggacccatg gatgcaggc 11880
 ctgtaggaca cagacaggat ggcatgaagg cctgtgcaac tggaggggtg gggatggga 11940
 cactaagaga tggctggaag tgtgggggtg gggacactaa gagatgactg gagaagagg 12000
 ggtcaggagt ggtgaaaaat gggagaggag ggcaggctgg gccttttgga tacagggga 12060
 ttgcatcctg cagtggtagg gagccactga gggctgctgc agtaggagtg aggggatcag 12120
 aggagagctt tggaagcccc ctggatgcgg gacaggaagc gagataccag tgtctaggag 12180
 gccagtgagg cagccacagg ctccaccagg atcagggtc cgagggtcat gaggaggaaa 12240
 ccaatttgaa ggagtccagg ggaataggac ttggaaatga ccgatgggac atttggaag 12300
 aggaagacag aagagcgcag tcccggcttc tggttttagc agttgggcaa ggggagatgg 12360
 ggagatgtgc ccatgggttg agggttgagg acattaggag ggagccggtg tggcaggaag 12420
 agctgggtgt ccagagatgc tggaagcagc atctgcctga gaacagatac ctggcaatat 12480
 tcctaaggga aagtacatc tcggagggtg aggaggcat ctgatagggc ctggaaagag 12540
 ccggggcaag catgaatgtg aggttatctt ggggggcaag gctcaggcgt tgaggagcag 12600
 cccctggtct cttcagcctg aagttggaag ccagagttgg gccaggtgca gctgtggtt 12660
 tctgaagtcc cctcccccga gccagtggt ccaatgctgt aagagcaagg gccgctcact 12720
 ggtgctggtg gctgagtcac agcaccagc acagggcctg gcacatactg gtgccaatc 12780
 ctcccttctg ggtgcttctt ccaaggcctt gtgatggaag tgagtaccct cttcgacatc 12840
 agaccagct tcaaatcctg gctctgctat gtattggctg cgtggcttta gacaagtctt 12900
 ttaaccttgc tgtgcttctg atttctcagc tgaaaaatgg agatgatgat agtggtttct 12960
 gtaaggcctt atggtgaagc acctagctca gggcctggaa ggcaggtgta accagtgggt 13020
 cagttgttat aaaccaacac taacctcgc ctttgcacct catgaaacca gatatgtaga 13080
 tggagccac aaagctagca ggagccaagc tcacgtgtgt cctgctttta agccccatac 13140
 cctttctcc gggtgacaaa cacctgtgct cgttctcttc cttccctc tcccccttgc 13200
 atttggttaa taacaggcca gctgcctgcc tccctgcagt ttggtagatg ggtgggtaac 13260
 gaccaccact cccacgttcg cctgatgggc ttgttttccg tgcccttcac aggcattctgc 13320

```

aacaggcccc agccaggcct gaagtcattc tcagaaggga tggatcctga ggtcgccatg 13380
cccagctggg caaacatgag tctggattct tccccggagt cggctgcctt ggctgtgagg 13440
ctggagggag atgagctgct agaccacctg cgtctcagca tcctgccctg ggatgagagc 13500
atcctggaca ccctctcgcc caggctcgct acaggtacct actcctcggg gtgagcacgg 13560
gcagcacctt gttttctttt ttgtgcatta tggaggaaga tggtagtgcc acatgggagc 13620
gatagggatg ggcaacctat acaggtgggt gggaacatct ctttccatgt gtacagcctg 13680
ggctgctgcc atcactccca gcacagcccc caaccccccc aatcctggaa ctttgccaag 13740
tctcccttcc catgggggtca tgaccaggag gaaaacaaac tccagctgag ccccttgggg 13800
ttccccatat aggtcctgct ctgtggcagc tgggcctctt gtaccccttt ccaactctgt 13860
ctccctaaca tggcacctga gctcctgcca tcttggattt catggacccc aaggatgggg 13920
gtcctgcata tgggacttgg cctattactc ggagctcctt ttcagccgcc tccctccacc 13980
tgtccaccca cctcaagget cttttcttga gacctctctt aatttctccc ttcccctaaa 14040
cccacaattt tgaacctcca tcgaatgggt ctgtatttta taatgtcatc aaatatcaaa 14100
tggagacagt gctatgggtc aaatgattgt gtacccccca gaatttgtct tttgaaatcc 14160
taacccccaa catgatgggt ttaggaggtg gggccttttg gaggagatta ggtcatgagg 14220
aaagggtgt catgaatggg attgggtgcc ttattaaaca gaccaagag aggtcccttg 14280
tcccttctac tgtgtgagga ctcagaaggt ggtgtctatg aagaagcagg ccctcaccag 14340
acaccaacat gtctgctgcc ccttgatctg ggaccttgca gcctctagaa ctctgaaaaa 14400
tcgatgtttg ttgttttata agccactcag ttgggtggcat tttgttagag tagcctgaac 14460
acggactaag tcaaacagaa gaaccacaaa accagctaca gagttgggca tttggagaaa 14520
ttcaaaaatg agtcagacat aactccttat tcttgaggtg ccctaagaga tgggacacag 14580
cagctgcccc ggtgcattag tttgttctca cattgctata aagaaatacc tgagactggg 14640
taactcataa agaaagaggt tgaattgggt cacagttgca caggctggac aggaagcatg 14700
gtgctggcat ctgctcagct tctggggagg cctcaggaaa cttacaatca tggcagaagg 14760
tgaacgggaa gcatgcacat cccatgactg gagcaggagt gagagagaga gggaaataga 14820
gggaagggtc catacacttt taaacaacca gatctcatga gaacacattc actatcaaga 14880
gaacagcacc agtggggaaa tccgccccca tgatccaatc acctcccatc aggtccgcc 14940
tccaacactg ggaattacaa tttgacatga gatgtgggca gggacacaga tccaaacct 15000
atgaccagat taatacgatt tgaggcatca cgaggtcatt aaagagaggg aataaaagac 15060
tggggctcca ggaagaaggc tctggaatcc agcagagggt caaggaccag cttgtaaagc 15120
tggtgggtgcc tgagaagtac ctaggagAAC atagatgctg tgacgtttga ttagctgtgt 15180
ttttgttttg tgttttggtt tttgagacag agtctcactc tgttgcccag gctggagtgt 15240
gcagtggcgt gatcttgggt cactggagcc tccatctccc aggttcaaata gatcctcatg 15300
cctcagcctc ctgagttgct gggattacag gtgcacacca ccacgcctgg ctaatttttg 15360
tgttttcagt agagacaggg tttcaccatg ttggccaggc tggctcttgaa ctctgacct 15420
caagtgatcc aacaacttca gcctcccaaa gtgctgggat gacaggcatg agccaccatg 15480
cccagcctga ttagctgtt tctgtgcaca ttatttgctg tgggtatat ttagatttct 15540
taatacaaga tgattctttg cctcatgact tacacacat tttctattta atttcagcta 15600
tgatattgga aatggacatg tcttttcaag gaaaataaaa gcaggctttc tggaatggcg 15660

```

acttccaaac atatttgtca atttaaagga gctgggagtg gggaccctat gtcctgtaag 15720
cactctctta gctgttcttg gctgtgctcc ccgcttcagc ttcacactgc ctttgctgtg 15780
aagggagcag cctgggccgg gcgcggtggc ttacacctgt aatcctagca ctttgggagg 15840
ccgaggtggg tggatcacct gaggtcagga gttcaagacc agcctggcca acatggtgaa 15900
actccatctc tactaaaaat acaaaaaatt agctgggcat ggtggcaggt gcctgtaatc 15960
ccagctactt gggaggctga ggcagaagaa tcgcttgaac ccaggaggcg gaggttgcag 16020
tgagccgaga ttgcgccatt gcactccagc ctggggggcaa caagagcaaa actctgtctg 16080
gaaaaaaaaaag aaaggagcag cttggcaaac cccaccttgt cgcttttctg agtgcctctg 16140
accctttggc tgccaggacg ggcgtatttt atggaaatgc taagcaccaa cagagtaaag 16200
tggtttggtt tttcacagtg gtgggagata atagctcaa attgtctttt tcagcactga 16260
gtgaagaaat gaaagacaaa ggtggataca tgagcaagat ttgcaacttg ctaccatta 16320
ggataatgtc ttatgtaatg ctgcctgta cctgcctgt ggaatctgcc attgcgattg 16380
tccagagtg agcattttag gtggtccgt gtcttctca cagggttgat atgaggatga 16440
aacaagatga tagatcatgg tggcatgtag tctgggacct ggattgtcgt gccacagatc 16500
acagctcaca gtctatgtgc aatgccctg aatgttgccc acctgtctc aagccacaca 16560
tgcacctgta actcagtga agcccagaaa ctccccgtgg ggactcctag agctgtcagt 16620
ggcctcacat agcagctggt ccagtctctt gtgattgccc aaggaaactg aggcctggag 16680
agcttggggc cactgctctg aggccataga gatgcctagt agaaggcca ggcctagaag 16740
caggatcctt gctgccccct tgagctgttt ccatttaaaa tcacatgaag gccggcgccg 16800
tggctcacgg ctgtaatccc agcatttttg gaggccaagg tgggtggatc atgtgaggtc 16860
aggagtttga gaccagcttg gccaacatgg tgaaatgcca tctgtactaa aaatacaaaa 16920
attagtggag catggtggca cgtgcctgta ctcccagcta cttggaaggc tggggcagaa 16980
gaatcgcttg agcctgggag gcagaggttg tagtgagcca agattgtacc actgcactcc 17040
agcctgggtg acaggagaga aaccctatct caaaataaaa tgaaagtaa tgaaatgaat 17100
aaaataataa atcaagtcac ggccgggcac ggtggctcac acctgtaatc ccagcgcttt 17160
gggaggccga ggtgggtgga taatgaggtc aggagttaa gaccagcctg gccaacatgg 17220
tgaaaccatg tctctactaa aaatacaaaa attagctggg catggtggtg catgcctgta 17280
atcccagcta ctccggaggc taaggcagga gaattgcttg aagcaggacc taggaggcgg 17340
aggttggttg cagtgaagcc agatcatgcc actgcactct agcctgggct acagagcgaa 17400
actccgactc aaaaaaaaaa aaaaaaaaaa atcaaatcac atgaaagtag aacataggga 17460
attccatctt tcgttctagg catagtttgt taatatgatt cagagccagc agttaggaga 17520
acacagtgtg actctcctag aacttcttga ttgggcttcc tctgattggg tttcctctga 17580
ttgggcttcc tctgaaagtg ggggggatgg ggggtgggga gcagaatggt cagagcttgg 17640
ctcagcagtc agactgctct tcttcaaate ctggctgcat tgcttactac agctgtgtga 17700
ctccagatga ctgaatccac ctctctgtgc tgcagcttcc cgtctagaga gatcacctgg 17760
agcagagggt ggtcaggaga ctcaatctgg ttactgactc acagtgcagg agtactcact 17820
ccatagtaag catccagcta gagatgttga tttctatttt caggtaataa tgatgatcgt 17880
aaaattagag acagataaaa ggtatgggca ttaggccagg gcactgcaat ttctaagctg 17940
tgtgacctca ggcaagttac tcgacttctc tgagcctcag cggtttctac cgcaatatat 18000

```

ggataggaaa accgacctca gtgggttgtc tgacagtgga gggcacttga ttaaaaaaaaa 18060
aaaaattacc ctggtctgaa tattaccctg gactgaaaga aaaatattga gctaatacag 18120
gcatcaggaa tggggctgca gggagtccag ggaagggaga acgaagagcc tgaaggtgtg 18180
aggaggtgcg agtgctgac tgtctgctac aaagaggctg ctgagcctcc tgtggatgtg 18240
gccctggact tggcagttta atacctgagc tgtaaaaata acctcagatg ctgtgttctt 18300
taaggggtag gattcagatt cctgctgaaa tgcttctgaa agggagggaa tgagccagcc 18360
catccccagt tgctttttaa gatcattggg aagtcttggt cttgccattt gtccctggac 18420
cactcttagg tcctcctgcc ccacttccat ctgggtgtgt gccctgggct gtccaccaca 18480
cagctacatc ctgccatctt ccctcctgga gccactgtgc catgcatgga tctgtagctt 18540
catTTTTtctt ggcttttccc tggTTTTtct ggagcagagt ctctagtaaa ctcccaagga 18600
agaaaacgtt tgactttatg tgtgttgga aacgtgcttt ttttctatta catctcagt 18660
ataggttggc catgtctaga attgcagggt gaaaatcatt tcctctcagt atattggtta 18720
gtgagaagcc tgggactgag acagtcacat tctcattctt ttgcagggtga gtgctcttag 18780
gactgtcttt ttatccctta tactctgaaa tgtcatatgt cttgggtgtaa gtccttattt 18840
cagttattga gctggacaag tactggagac cccttcagtc aaagccttct gtcattctcc 18900
agctctagga aattatcttc tattgttatt tctgttattc cttcccttcc attttctttt 18960
ttcttttttt tttttttttt ttgagacagg gtcttactct ggtgcccagg ctggaatgca 19020
gtgacctgat catggtacac tgcagcctga acctcccaga ctcaagtgat cctcccacct 19080
caacctccta agtagctggg actgcaagca cacatcacca cacccaacaa atatttttta 19140
aaaattttgt aagatgggat ctactatgt tgcccagact ttttcttctt cttcctgggg 19200
ctcttattag gaagatgttt gacttcctgg gttggattcc tgtctccgtg tctgactttc 19260
tctctttgtc atatttttca tctctggtt tctttttgcg tctgctctga cagatttctt 19320
caaattttgt ctctagctcc tctctacag ttttacttt cagcaaatat aatttaactt 19380
ccaagagtac tctctgttc tttttctta gcattctgtt cttgttttat ggatgtaaca 19440
ttctcttgga atatttgctg tcctctagat catcccttct ccatttcttc ttgggctagt 19500
ttttctgttt ctctatcttt ctcttttatg ctacttattc tgggcgtgtt cttggtgggt 19560
tttttcccat atagcaacag aggacttgga gctcaggag aaaagggtag gtgcatcacc 19620
tggcagagct ccagacagt gacaggcagg ctgcgggaag gatgtctact tggcgggtgct 19680
accgctttcc tagaaaccct ttccctggag ctggttgaac tgttgggttt tgccctgggtg 19740
gtgaacgctg gctccccgtg ctctgctgtt tcatcacca gccccctccc cttctgcctg 19800
gggtccagta atctgttgaa atatatatct tgctcattgg tgagctcctg ctccttcctc 19860
gttgctcttg cagatttatt acttctcgta aggtgcgct tgtacttcgg gattttctct 19920
gtgccacact gggaaacata gggtggttgc atgctgcagt cctgagcact tatttcactc 19980
acatctttac acgaagattt ggtgggtgtt tactttgttt ttagtaagtt agtctgtcat 20040
gtcctttgat cttttttttt tgTTTTttga gatggagtct ctctgtgtcc tccaggtgg 20100
agtgcaatgt cgcatctca gctcactgca acctccacct cctgggctca agagattctc 20160
ctgcttcagt ctctgagta gctgggatta caggcatgtg ccaccacacc tggctaattt 20220
ttgtattttt agtagaggtg gggtttggca tgttgccag cctggtctca aactcctgac 20280
ctcctgacct gcctgccttg gcctcccaa gtgctgggat tacaggtgtg agccaccaca 20340

```

cctggccctg attaatcttt taatgcccag tctctccttc aaaagccggc tcctttctct 20400
ccctcgcctt cctagattcc ttctccactc cccaggatca gcctcctcct cccacacca 20460
ccactgccgg ggggatgtct gtggtcaggc atttatcaga gaccctgagg tgggggtcct 20520
ttatgtgtct gggggatgga gagtctagag gaggtagcgt tcagacctct ccatggtgcc 20580
tctgctgggc tcacatgtga ccaagcacag caaacatga ggcaggggat ggtcttgacc 20640
atgagagccc ttgcagcagc tgccatgggc ctgagctcct ctccaagctg ggaagagccc 20700
tgaaaagcca aggtgttttt tttccctct ttatttcagt gtaagtcctt tgagctttct 20760
tgaaccagaa gtgggctcat tttgctttag agatttcagg tgggcttgct cttgtcctag 20820
catcccagat ccaccttctg ggaagtcac agattggagg tgatgttggc agcttttgta 20880
aacaaaggg agtgttgtaa gctgttgtgt ctgcctatgt gtgtgtttgt gtacttggtc 20940
tcactctctg agactggtga catggcttcc agatatgcc gacgatgtcc tgtggttgca 21000
gtgggtgacc tcacaggtgt tcaactcagt gctgatgtgt ctgctccccg cctccaggta 21060
aatactttgg ctgtgggtgt gtgggccgga cgggcacctc tctcatctga tgaggcctca 21120
cacgacattc tagaaacagc tggctgaaca ccaagcaagg agcttgccct tgggtgtggg 21180
gacctgtct catgggaggc agctgagtca gtcagaggtc ctggcacacc tgctgagagc 21240
tgccaccag gccaacctga accggagcct gggaagactt cccgtcggat gagtctcttt 21300
gagtgcagca ttgatggtg aagagcagag agggcccaga taagcaggga aagtgcttc 21360
agacagagt gctgggatga ggactgggga gtgtcagata gcgctggcgt gtctgagcga 21420
aggagctctg gcacccatgg cacaggaagg aggtgggacc ctggaggggc agggctagca 21480
gagctcctcg gagcgtgtgg ctaggtgcct ggtaatgcaa gcccctgtc ctccaccctc 21540
tgttgtactg agtcacagtc tccgggtga agcctagcag tctgcgttga caggccccag 21600
gggatgccgc tacttctga attctgaatt ctggaaactg agccggagt cagggcctgg 21660
ctcccattac cagggttgg cgttatcctg aaaatcatag gccttggtt cctcacttgg 21720
ctaacagggg tgatcccat cccctcaatg ggtttccgtg agctcctgag agcccgtagc 21780
atggtacttg gcacatgctg ggcatcagga ggtatggcct ctcttgctat tgttgttatt 21840
ggtagacaca gaaggattta aaagtagggg aatgcaaaga tccgatttgc tagggaagag 21900
ggcagtagtg gccaaagtag ggtggatcc tgggccctgg ctggcagcag gcagcaaggg 21960
gggctgccag ggcccaggca gggacgatct gtagaccgag aggttccta aggtcttgg 22020
acaggaggag gtgtcggttc caagcctaag gagtggggca gccctggtga ctggtggtca 22080
gtggtgccag gcggtgggtg gtaggacacc ctggcaggca agtaggtttg tgtgggggaa 22140
actgataggc ccctccaggg attcgttggg ggacaacacc tgtgatgtcc agtgggaggt 22200
gtccaggtag ctgggagggc cacaggcttg gaagacctag gtggtgacat cagcccagca 22260
ctgagggcta gaagaagctg tgtctctggc tgtgacggca ccctagagtg tgtgtggtgc 22320
cctctactgg ccggcaatgt ggggtccacc tagctcagac tgcacactgc agcagcggga 22380
acggcctcta agccaacttc ctccatgtgt ttcaggtccc aatgccagt gagcagccaa 22440
caggcctccc catgcacacc tgagcaggac tggcctgct ggactccctg ctccccaaag 22500
ggctgtccag cagagaccaa agcagaggcc accccgcgt ccatcctcag gtccagcctg 22560
aacttcttct tgggcaataa agtacctgct ggtgctgagg ggctctccac ctttcccagt 22620
ttttcactag agaagagtct gtgagtcact tgaggaggcg agtctagcag attctttcag 22680

aggtgctaaa gtttcccatc tttgtgcagc tacctccgca ttgctgtgta gtgaccctg 22740
 cctgtgacgt ggaggatccc agcctctgag ctgagttggt tttatgaaaa gctaggaagc 22800
 aacctttcgc ctgtgcagcg gtccagcact taactctaata acatcagcat gcgttaattc 22860
 agctggttgg gaaatgacac caggaagccc agtgcagagg gtcccttact gactgtttcg 22920
 tggccctatt aatggtcaga ctgttccagc atgaggttct tagaatgaca ggtgtttgga 22980
 tgggtggggg ccttgtgatg gggggtaggc tggcccatgt gtgatcttgt ggggtggagg 23040
 gaagagaata gcatgatccc acttcccat gctgtgggaa ggggtgcagt tcgtcccaa 23100
 gaacgacact gcctgtcagg tggctctgca agatgataac cttgactact aaaaacgtct 23160
 ccatggcggg ggtaacaaga tgataatcta cttaatttta gaacacctt ttcacctaac 23220
 taaaataatg tttaaagagt tttgtataaa aatgtaagga agcgttggtta cctgttgaat 23280
 tttgtattat gtgaatcagt gagatgttag tagaataagc cttaaaaaaa aaaaaatcgg 23340
 ttgggtgcag tggcacacgg ctgtaatccc agcactttgg gaggccaagg ttggcagatc 23400
 acctgaggtc aggagttcaa gaccagtctg gccacatag caaaacctg tctctactaa 23460
 aaatacaaaa attatctggg catggtggtg catgcctgta atcccagcta ttcggaaggc 23520
 tgaggcagga gaatcacttg aaccaggag gcggaggttg cggtagctg agattgcacc 23580
 atttcattcc agcctgggca acatgagtga aagtctgact caaaaaaaaa aaatttaaaa 23640
 aacaaaataa tctagtgtgc agggcattca cctcagcccc ccaggcagga gccaaagcaca 23700
 gcaggagctt ccgcctcctc tccactggag cacacaactt gaacctggct tattttctgc 23760
 agggaccagc cccacatggt cagtgagttt ctccccatgt gtggcgatga gagagtgtag 23820
 aaataaagac 23830

<210> 31

<211> 23830

<212> DNA

<213> 智人(Homo sapiens)

<400> 31

atggtccgag gggggcgggg ctgacgtcgc gctgggaatg ccctggccga gacactgagg 60
 cagggtagag agcgcttgcg ggcgccgggc ggagctgctg cggatcagga cccgagccga 120
 ttcccgatcc cgaccagat cctaaccgc gccccgccc cgccgcccgc gccatgtacg 180
 acgcagagcg cggtggagc ttgtccttcg cgggtgctg cttcctgggc ttctaccagc 240
 tcggggcgac ccgtgcctg agcgagcacg cccgcacct cctccgcgac gcgcgcatgt 300
 tgttcggcgc ttcggccggg gcgttgcaact gcgtcggcgt cctctccggt atcccgtgg 360
 gtgcgtctgg ggacgtgcc cgggtccac gtgcggagt ggtgccccct aggccgggga 420
 gcgggggatc cccaggggtc gcggggccct ggaggagcgg gcatcgagc cggaacggc 480
 ggggtgcac ccgagggccc cctccgagc agatgcttc tgcgggggcg ctgttctctg 540
 gcccgggaag ggggcgttg aaccccgagc ggtccgggccc gaagcctggg actctcgtgc 600
 gtccccaccc ctaccccat caggcgcgcg tgcatgaagg gagacctca cctccgact 660
 gagagtcgga gcgtctcgga gcgacgggga gtagggagcg ggaccgggg cgagggttag 720
 tgctggcccc tcgggactcc ggggtccctg tgctctctcg ggaggggctg gacgggctga 780
 gctgccgagg ggccgattg ccctgggccc gacaaagagt ggggctttgg ccggtcccc 840

```

acgggtgggct ccttccctct ggggattgag ggactcaaga caccgccgcg ctgcgctttt 900
cttttctttt tttctttttt tttttttgag acggagtttc gctcagtcgc ccaggctgga 960
gtgcagtggc gtgatctcaa ctactgcaa gctccacctc ccaggttcac gccattctcc 1020
tgcctcagcc tcccagtag ctgggactac aggcgccagc caccaagccc ggctaatttt 1080
ttgtattttt tagtagagac ggggtttcac cgtgttagcc aggatggtct cgatctcctg 1140
acctcgtgat ctgcccacct cggcctccca gaatgctggg gttacaggcg tgagccactg 1200
ctccctgctg cctacgctct ctgggtcgca gccagcctt ctgggggctg ggtagcctcc 1260
cagaagggca accctgggca tcctccaggg caggctaact ggagtctagt ggggaggggt 1320
accttgaaaag aggaaagtgt tttctctctc ctctctctcc tccagtgttt gggacccttc 1380
ctgggggctg gagtgcattc ctggacaccc cccaatccca tcctcttctc tagtttccac 1440
tgacctaggc ccaccctccc ctctccggtc cagtactctt ggaaatgaga ttccgtacat 1500
ttgaatcttg tcctaataaa atatttgtcc atgtgggtac ctgtgtgtgt gtggtggggg 1560
tgcagacgga ggggtttgtt ctactagct ggaactactg ggggtgtggt tgcttctctg 1620
gaatttgtgt gccacagtcc tggaggcgag gagggggttg tgagccagta ggcaggggct 1680
ggggcaagta gcatttgtaa gctattgaca ccagacgtc ccagggcagg agattatgcc 1740
cccattagcc cccttttatt tgggttctc taacaatgga ctctttgcc tgctgcccag 1800
agccagcagg gactgactgt tcagtgggtg ggaagcgggc agaggaagcc ctgccattgg 1860
gtaggagcag tgggcagccc ctgggctgac tgggaggtgg ggattaggga ttagacagtc 1920
ctggctgtct gccttcccct aagccagggg gagaggagca aagggcacga aatgtggcct 1980
ccaggaggat tagaccgcca catgatcatt tgcacacctt ggggttttag aacaataaaa 2040
gtcagctttt ttgtatccca aggtggcctg tggacaccca catggacaaa tgtttact 2100
gggacagaat tcaaatgcag aggtcccagg agcctaaagt aactcactc tggtatagaa 2160
aggattcctt actgggcaga ggacaggtgc agcctggggc tttcccaggc aggacacagg 2220
gaggctcagg aaccaccaag tccctggaag gtggatctgg aggcgttggc aggagccact 2280
ccctgggttc cagggtccca ggttctctg ttaacccct gtctcacaga gggctgtgca 2340
cttgggggct gctgagcatg tcccagaggc tgcctcctgg acacagcacc tcagtgcatt 2400
tgagctgagg ctaacttggc aggagggaca ggcagaacct gccagccacg tgcaattcca 2460
cccctctggc cactcaggga aggagagctg tgagtcaaga tcagatttgg gtcaggacag 2520
gctggggcct gcctgtccct gtgcattcca agatttatgg ctggccaggg gttgggctgg 2580
gaggggttgt cttgcatgcc aggagagtgc agatcagcct gagaggccag gccagtaagt 2640
gaggtcagat ctctgcacc tgatagcatt aaggccatct acaccaaagc tctaattgct 2700
atatgttcct ggctctatg tggggcatgg aggtggggca tggaggtgag gcctgctcgc 2760
ctgggcttct ggaagtggga gactcattcc tgtggctgag gcctacagca gtgctgtgtg 2820
gtaggaatac actggaagcc atgatgtcat tgtgcatttt ctagaagcca cattgaataa 2880
agtaaaagac acaggtagaa ttaatttcat tgagcccaat atatccaaa taatatcatt 2940
ttcacatcta ttcaataaa aaatttacta atgagatatt tcatactaag ccaactgaaat 3000
ccagtttgta tcttacacat ctgagttttg acgagccaca tttcaagggc gtgatagcca 3060
catgtggctc ccatagtaga cagtactggt ctagagaaat gttggtggca tccttgctgt 3120
ctggtttctg gccttgccaa aagtattacc atcccagtgt ggtacattct ttcattgtatt 3180

```

```

tgtctcctgt cccagagca gactctgcag gtcctctcag atcttgtgcg gaaggccagg 3240
agtcggaaca ttggcatctt ccatccatcc ttcaacttaa gcaagttcct ccgacagggt 3300
ctctgcaaat gcctcccggc caatgtccac cagctcatct ccggcaaaat aggcatctct 3360
cttaccagag tgtctgatgg ggaaaacgtt ctgggtgtctg actttcggtc caaagacgaa 3420
gtcgtggatg taagcagttt gcttatctgg acgttgtcaa gttagaaaag ctgttttggg 3480
atgggtgtgg tggctcatgc ctgtcatccc ggcactttgg gaggccgaag cgggtgggtt 3540
gcttgagccc aggagctcga gaccaacatg atgaaaccca gtctctacaa aaattacaga 3600
aaaattagct aggcatggtg ttgtgggccc atagtcccag ctactaggga ggctgaggca 3660
ggagaattgc ttgagcctgg gaggtggagg ttgcagtaag tcatgatcat gccactgtac 3720
tccagcccggt gtgacagtga gatgctgtct ggaaaaaaa aaaaaagaaa gactgttttg 3780
ttttggaagc aacacaggca gttgtaggcc ccctgtgcca gagtacata aactctgtac 3840
acctccagtg atttggtcca tgtttgtaaa ccctgaatgt tccagggcag tttcttttct 3900
tcaacttttta tctctttttt ttgggtgggg gggcggggta cagagtcttg ctctgtctcc 3960
caggctggag tgcagtggcg caatctcaac ctcccagga gctgggacta caggcacagg 4020
ccatcacacc ttgctaattg ttgtactttt tgtagagacg gggttttgcc ctgttgccca 4080
ggctgggtccc aaactcctgc acccaagtaa tctgccacc tctgcctggc agttacaatt 4140
tcaaataatt cctccccttc cttcaacact tggctcatga ccgtccagtc caaggaacct 4200
gtcctgcagg tgtgcctctc ccgagcttcc tctatgcatc ttccataatg aagatgcctt 4260
ctcactggaa accctacaag ggtgggaacg tgccttattt gcctgtatcc tcagggtcta 4320
gcagagagaa gataatttgt aatacaaaaa caccattaaa ttcagctgat gctttcataa 4380
gcgctccttg gaggaaggac tccatttact tgacagatct gtgcaagaca gcagcctggc 4440
gcgtctaacc tgcagccagt tgcattctct gttaaacctt gtttgcgga gctttctcta 4500
aacagccagc acttgtctgt tcccacatgg gtccgttctc ccagtgaatc accgtggtgc 4560
ctgctgactg ctctgtagca cagtgtctcg caaagtgtga tcctgggacc agcagagcag 4620
cagctccttt gagcttattg gaatggcaga ccctcaggtc ccacctctga cctgctgcat 4680
gggaattctg gggagggacg cagaatctct ggttccacag gctctccggt gatgctaatt 4740
aataccggca tttgaacagc accgatctag cccctttcag tccatgagcc aacaaccctt 4800
ggtcctgtct gtggtgacct agtgtgactc tcatggggag caaggagagg aagttgaagt 4860
tcactgacag ggttgtaag gggattatgc aatagatgag acccatgggc ctgaagtccg 4920
agggtgtatg ttagttcccc gttcttttga ccatggatt aacctactct gtgcaaaggg 4980
cattttcaag tttgttgccc tgctcacttg gagaaagctt atgaaggatc agggaaatta 5040
aaagggtgct ctgcctata acttctctct ctttgcctt cacaggcctt ggtatgttcc 5100
tgcttcatgc cttctacag tggccttate cctccttctc tcagaggcgt ggtaagtcgg 5160
ctttctctgc tagcgctgag tcctgggggc ctctgaagtg tgctcacaca tctcctgcct 5220
gcagggcact ggtgtcgggc acctcagggt ctgtcccatg gtggagcccc atgcctcact 5280
gcctttcaga cagagtagcc acagctggcc ctatttccag gctaccggg cagcaaaact 5340
tactgcatgt gtaattaatt atttggtat ctgtaaggta aactggctgg ttacttaat 5400
ctgcacctta agcatcagat agcttctcag tgatctagtt aaactatatg atgttgcca 5460
ggcgcggtgg ctcatgtctg taatcccagc actttgggag cctgaagcag gcagatcact 5520

```

```

tgaggtcagg agttcgagac cagcctggcc aacagtgtga aactctgtct ctcctaaaaa 5580
tacaaaaatt agctgggcat ggtggtgtgc acctgtaatc ccagctgctc gggaggctga 5640
ggcaggagaa ttgcttgaac ttgggaggcg gaagttgcag tgagccaaga tcgcaccact 5700
gcactccatc ctgggtgaca gagcgagact ctatctcaaa aagaaaaaaa aaaaaaaggt 5760
aaataaagta tatgacactg aagaatctgt tacccttgga aggtggagct ttactcttag 5820
ggggaactat aacagtcata tatatatatt tttttctttt cttttttttt ttttttgaga 5880
tggagtctgg ctctgtcgcc caggctggag tgcagtgggt caatctcggc tcaactgcaac 5940
ctccacctca caggttcagg caattctcct gcctcaacct cccgagtagc tgggattaca 6000
ggtgcctgcc gtcacgccaa gctaattttt gtatttttag tagagacagg gtttcatcat 6060
attggccagg ctggtctcca actcctgacc tcaggtgatc cgcccgctt ggccctccaa 6120
agtgtgaga ttacaggcgt gagccatggt gcccgccaa caatcacatg tgttgtaaac 6180
aacaacaaaa atctgtcagc ctggtctaac ctagatttgt gctttgtttt gttttgccac 6240
tttgtgatgc acaggaggaa gtttagctg taaaatacta gccttttagg gtaatttttg 6300
aactcacaag agcagcagcg gaacctttga tgcaatctg tatgtagcac cagcagagcc 6360
acgtggcaga gggactcgca ttaggagcct ccattacag actacgtgt cctgtgcgtt 6420
atcttatagg gtccccacaa ccaaggggag atgtgattat tcattcctgt tggctgtggg 6480
gaacttgaga gtcatacttg cccaaagagc acggccagcg agcttgacc caggtcactc 6540
tctgtcctc tgtcagaaca gggcatgtct tggttcactg cagggcggct cttctcattc 6600
tctgtagttt ggggtccagg atagtggctc acggagccac tggagtggc agctactgag 6660
tgaccaaagc atattttgga tttccgacat tgccacagca tggttgggca tcagcaggac 6720
cccaaccctt tgttatgctg gtggctttat gtggttattt gatcttcccc agaactcagc 6780
aggagtgcac ccagcagcac cgtagtgatg ctctctggct cccagtgca cggttctggc 6840
tttcttcctt ggtcgagagt ttcaagcctt ctgggtccta ctctgtcctt ttcagcccat 6900
agctttgttc aaaagctgct ggcagtgttc agatttggct gagttcagt aatatgtgca 6960
ttggctgatt tctgagccat gccaggggga tggagaagcc gaagcaggag tgtttgttct 7020
gcaggctctg gagtaggcat tgggtctgtg ccggctcact tgctagtctt gcattccttc 7080
ccaacccctt ctggggatgt ctggccacat cagaagacag tttgggttgt cagaactggg 7140
ggagtaccag gccgaggtgg gtggatcatg aggtcaggag atcgagacca tcctggctaa 7200
cacagtgaat cctcatctct actaaacata cgaaaaaat tagctgggcg tgggtggcggg 7260
cgctgttagt cccagctact cgggaggctg aggaggaga atggtgtgaa cccggggggc 7320
ggagcttgca gtgagctgag atcctgccac tgcactccag cctgggcaac aaagcgagac 7380
tccgtctcac aaacaaaaca aaacaaaaca aaacaaaatc tgggggagtg ccactggcat 7440
ctgatgtata gaggcccgag atgtgtgtc atcacccgtt gagtgcgctc ataggcatct 7500
tcctgacaat tagaaccat tattcttcaa attcaatga agcaaattca aagcattact 7560
gtgtacatac cgcatgctaa tcaattgcac cactggagct cctaaattca aaacattact 7620
ataaaaaagt tcaaaatgca tggaaaagt gtacatggca ggagaatatt tgggtctctg 7680
actaccctt gaatgaagat gatccaccag ccgccttctt ccttggctt cactccagat 7740
tcctagcatt tcattctgtg tctctttatg cagtgaagtt tttgtttgtt ttttgagaca 7800
gagtctcact gtatcaccta ggcctggagt gcagtggcgc gatctcagct cactgcaacc 7860

```

ctccggctcct ggggtttaagc gattctcctg cctcagcctc ccgagcagct gagattacaa 7920
 gcacacatcc ccatgcccag ctaatttttg tattttttagc agagacaggg ttccaccatg 7980
 ttgcccaggc tgggtctcgaa ctccctggcct caagtgatcc atgtgcctca gccttccaaa 8040
 gtgctgggat tacaggcgtg agccaccatg cccagctcct agtgaggttt ttgatgcctt 8100
 gctacatctg ccctagaaat tgtgtgacta cgattttgga aatgttgctg tgtaaacttg 8160
 tgatcatttc tggactccag gcaagaatct tgatggctaa ggtgtggctg aacatgtctg 8220
 attctctcct ggacctgttt taggccaaac tctgctctga aattcctccg tgtggaaggg 8280
 cgggctgggg agagcctccc agctggaatc ttttggatgc ctttctctgt gggatatctga 8340
 tggctggctc tgatggctgg ctgtgatggc tgtggctgga aatcattgtt gacatgagtt 8400
 tcacagatgc aggctctgtc caaattgtag caaaagctgc ctgccccagc cgagctatgg 8460
 gcaataaggt ggtttaagga tatagatgaa ggaaaactca cccttagaat aatttatcca 8520
 aaatgctgct gtgttggtgg ttagaggaca ttttctgagg tcccaggttc attgtttcat 8580
 ttaagtctca aaagtccctc caggtgttgg ttctaattgt caaagcatgg ggggagatgg 8640
 gctcatgggt taaaggctct atcccagatt tctgtatcct ccttgcaagc agcaaagggg 8700
 tctggatttg aatccatgac catgtttctc ctttgggttt ccatcacact ctgtccccgt 8760
 gcactgagca ccctttagtt catatgacct ccttaggcat gttacatggg cactcctata 8820
 ggtgcccatac tggccctagg acttggccaa cacaacatgg actccagttt ccatctgcct 8880
 ctttgccagg cacttttggt cagtgcacac actgtacaac agtagacggc aaccctgaga 8940
 gccagagtag agcctgtcct agcaccggaa tgctcggtaa ggatttgctg caggagtgat 9000
 tccaaagcca atgtcctccc tccatatcag cctgtttgtg gctctgagaa gctctgccc 9060
 catgtgaaag ctgtttaagc acttaagcac taaccagag cttcagacag tgccagtcct 9120
 ttttccccct ctttaaaagc gatatgtgga tggaggagtg agtgacaacg tacccttcat 9180
 tgatgcaaaa acaaccatca ccgtgtcccc cttctatggg gactacgaca tctgcctaa 9240
 agtcaagtcc acgaactttc ttcatgtgga catcaccaag ctcagtctac gcctctgcac 9300
 agggaaacct taccttctct cgagagcttt tgtcccccg gatctcaagg tgagttggtg 9360
 gtgagggggc aggtgttctg ggggtgcagct cttctttgcc tccctgattg ccaggagcta 9420
 ccagttactg tctgcacaat caaacagaaa tagacctgtc cttgatgggt aacggaaata 9480
 aaaggcgctt gtcccagaag ctcaggtgag gcaccacct gattatggga atcacctggg 9540
 aacatatacc cagacctaaa actcagatcc acttcccagg ctgtgggttat atagtcaggg 9600
 ggggtgcagta tgggtattag gattttttat ttttagtta taaagatttt ttttgggtt 9660
 gtttttgaga cagggtcttg ctctgccgct taggctggag tgcagtgggt caatcatagc 9720
 tctactgaagc ctcagactcc tgggttcaag cagtctccc acctcagcct cctaaggagc 9780
 tgggacccac aggcatgcag caccacacct ggctaatttt taaaaatttt gtggagtgtt 9840
 gccaggtg gtctcacact cctggcctca agcgatctc ccaccccagc ctcccaatgt 9900
 gttgggatta caggcatgag ccattgtacc cagccactaa gatgattctt atttggaac 9960
 acggtcaaga acaactgcgt tcggtagttt aacctttttt gattgtggtg gtttttagtat 10020
 gccttaccac tctaccatag taagaaattt gcagaccatg tacaccaacc tttggtgctc 10080
 ctggggagaa agaaagaagg ctatgcaatg caatgcatgc tcacagtcca agggagaggg 10140
 aaagctgtct aacaggattg gttttcccggt gtgctttata agcagatgag tagaggagac 10200

```

agctcttatt gtcctagtgg caattgggat aggctgcaaa gtttgttagg gtggaggctt 10260
attccgggac caagggagcc caaagaaaca agctcctgcc aggcgcggtg gtcacgcct 10320
gtaatcccag cactttggga ggctgaggca ggtggatcac ctgaggtcag gagtttgaga 10380
ccagcctggc caacatgggtg aaaccccgctc tccatgaaaa atacaaaaat taccgggca 10440
tggtggcggg cacctgtaat ccagctact agggaggctg aggcaggaaa atggcttgaa 10500
cctcggaagt ggaggtggcc gttagccgag atcacgccac tgcactccag cctgggcaac 10560
agagcaagac tctgccttaa aaaaaaaaaa aaaaaaaaga aaagtaaaag gaaaaaaaaa 10620
aggctctggc ctgctggggt gcctgcaaag tctccgtgga agggtgacat tcaagccgag 10680
acctccaggg aactgtctcc tgggagcaca gagccctttg ctgagccccc aggtggctca 10740
gtgccccag ccagcagact cagagcttgc atgattcttt ggtgctctct gcggtcttcc 10800
aatgatgctg aaataaatgg tgcttggtgt ctccctgctg tagtccccct gcttgctttg 10860
ctcacagggt ctgggagaga tatgccttcg aggatatttg gatgcattca ggttcttgga 10920
agagaagggt atgtatgggc tgggaggatc agccatgccc ttttgacaag catttactag 10980
cggctcttgg aaagacttga gatttgcctt agttctaaca cttagtgcc aacgccttcc 11040
ttgtgttgct caacctactc atgagcccag gagataggaa atctccgtcc cattgtacag 11100
atggggaaac agaatttttg aaaggagagc caagcagcac acaccctcc ctgaggggca 11160
gagccgagat ttgaactggg atgtcatgac tccagggccc tctccctccc cagggtcccc 11220
ttatctgaag gcggtttttc tttccagctc gacctcttgt gacccttagt ttaacaagg 11280
ccgaagttaa agagtttctg cgcttgacc ccaaatgaag caatcagatt tctcatctcc 11340
agtcagggtg gggccaagc ccactagaca agtttgctct tcccagagca catttctgcc 11400
ttcaagtcat cctggcttgt cagggtggg ggagttctgc tgtagaaata ttagagtgga 11460
aggaaaaaga tgtgttgga gctatttttc ttaataacta aaagttggtt gatgaatttg 11520
tcgttgcca agaccaagga gactgcattt ttaaggacat atgtgtattt atctgctcag 11580
aaaatgttca ttgctgtgtg ctagggatac tgcagtgaac acagaggtgt gacccttgcc 11640
agccttgta gagaagtgag cagataagta agcagaaggg tgatgctgtg tcgatgggaa 11700
agtacagggt ccaatgagaa ggcacagggt tcaaggagaa gacacaggat gctggaggct 11760
catgcaggat ggatctccaa ggcccagggg aagaagggcc tctcgaggga cgtgaatcca 11820
cattaagact ttggggataa gtaggagcgc ctaggcatg gggacccatg gatgcaggc 11880
ctgtaggaca cagacaggat ggcataagg cctgtgcaac tggaggggtg gggatgggga 11940
cactaagaga tggctggaag tgtgggggtg gggacactaa gagatgactg gagaagagg 12000
ggtcaggagt ggtgaaaaat gggagaggag ggcaggctgg gccttttgga tacaggggga 12060
ttgcatcctg cagtggtagg gagccactga gggctgctgc agtaggagtg aggggatcag 12120
aggagagctt tggaagcccc ctggatgcgg gacaggaagc gagataccag tgtctaggag 12180
gccagtgagg cagccacagg ctccaccagg atcagggtcg cgagggtcat gaggaggaaa 12240
ccaatttgaa ggagtccagg ggaataggac ttggaaatga ccgatgggac atttgggaag 12300
aggaagacag aagagcgag tcccggcttc tggttttagc agttgggcaa ggggagatgg 12360
ggagatgtgc ccatgggttg agggttgagg acattaggag ggagccggtg tggcaggaag 12420
agctgggtgt ccagagatgc tggaagcagc atctgcctga gaacagatac ctggcaatat 12480
tcctaaggga aagtgacatc tcggagggtg aggaggcat ctgatagggc ctggaaagag 12540

```

```

ccggggcaag catgaatgtg aggttatctt ggggggcaag gctcaggcgt tgaggagcag 12600
cccctggtct cttcagcctg aagttggaag ccagagttgg gccagggtgca gctgtggttg 12660
tctgaagtcc ccctccccc gcccagtggtg ccaatgctgt aagagcaagg gccgctcact 12720
ggtgctggtg gctgagtgcc agcaccagg acagggcctg gcacatactg gtgcccgaatc 12780
ctcccttctg ggtgcttctt ccaaggcctt gtgatggaag tgagtaccct cttcgacatc 12840
agaccagct tcaaactcctg gctctgctat gtattggctg cgtggcttta gacaagtctt 12900
ttaaccttgc tgtgcttctg atttctcagc tgaaaaatgg agatgatgat agtggtttct 12960
gtaaggcctt atggtgaagc acctagctca gggcctggaa ggcagggtgta accagtgggtt 13020
cagttgttat aaaccaacac taacctcgc ctttgacct catgaaacca gatatgtaga 13080
tgagagccac aaagctagca ggagccaagc tcacgtgtgt cctgctttaa agccccatac 13140
ccctttctcc gggtgacaaa cacctgtgct cgttctcttc ctttccctc tcccccttgc 13200
atttggttaa taacaggcca gctgcctgcc tcctgcagt ttggtagatg ggtgggtaac 13260
gaccaccact cccacgttcg cctgatgggc ttgttttccg tgcccttcac aggcattctgc 13320
aacaggcccc agccaggcct gaagtcatcc tcagaaggga tggatcctga ggtcgccatg 13380
cccagctggg caaacatgag tctggattct tccccggagt cggctgcctt ggctgtgagg 13440
ctggagggag atgagctgct agaccacctg cgtctcagca tcctgccctg ggatgagagc 13500
atcctggaca ccctctcgcc caggctcgct acaggtacct actcctcggg gtgagcacgg 13560
gcagcacctt gttttctttc ttgtgcatta tggaggaaga tggtagtgcc acatgggagc 13620
gatagggtga ggcaacctg acaggtggtt gggaacatct cttccatgt gtacagcctg 13680
ggctgctgcc atcactccca gcacagcccc caaccccc aatcctggaa ccttgccaag 13740
tctcccttcc catgggggtca tgaccaggag gaaaacaaac tccagctgag ccccttgggg 13800
ttccccatat aggcctcctg ctgtggcagc tgggcctct gtaccctttt ccaactctgt 13860
ctccctaaca tggcacctga gctcctgcca tcctggattt catggacccc aaggatgggg 13920
gtcctgcac tgggacttgg cctattactc ggagctcctt ttcagccgcc tccctccacc 13980
tgtccacca cctcaaggct ctttcttga gacctctct aatttctccc tccccctaaa 14040
cccacaattt tgaacctcca tcgaatggtg ctgtatttta taatgtcatc aaatatcaaa 14100
tgagagacagt gctatggtcc aaatgattgt gtacccccca gaatttgtct tttgaaatcc 14160
taacccccaa catgatggtc ttaggaggtg gggcctttgg gaggagatta ggtcatgagg 14220
aaagggtgt catgaatggg attggtgccc ttattaaaca gaccaagag aggtcccttg 14280
tcccttctac tgtgtgagga ctcagaaggt ggtgtctatg aagaagcagg ccctcaccag 14340
acaccaacat gtctgctgcc ccttgatctg ggacctgca gcctctagaa ctctgaaaaa 14400
tcgatgtttg ttgttttata agccactcag ttggtggcat tttgttagag tagcctgaac 14460
acggactaag tcaaacagaa gaaccacaa accagctaca gagttgggca tttggagaaa 14520
ttcaaaaatg agtcagacat aactccttat tcttgaggtg ccctaagaga tgggacacag 14580
cagctgcccc ggtgcattag tttgttctca cattgctata aagaaatacc tgagactggg 14640
taactcataa agaaagaggt tgaattggct cacagttgca caggctggac aggaagcatg 14700
gtgctggcat ctgctcagct tctggggagg cctcaggaaa cttacaatca tggcagaagg 14760
tgaacgggaa gcatgcacat cccatgactg gagcaggagt gagagagaga gggaaataga 14820
gggaagggtgc catacacttt taaacaacca gatctcatga gaacacattc actatcaaga 14880

```

gaacagcacc agtggggaaa tccgccccca tgatccaatc acctcccatc aggctccgcc 14940
tccaacactg ggaattacaa tttgacatga gatgtgggca gggacacaga tccaaacat 15000
atgaccagat taatacgatt tgaggcatca cgaggtcatt aaagagaggg aataaaagac 15060
tggggctcca ggaagaaggc tctggaatcc agcagagggt caaggaccag cttgtaaagc 15120
tggtggtgcc tgagaagtac ctaggagAAC atagatgctg tgacgtttga tntagctgtt 15180
ttttgttttg tgttttggtt tttgagacag agtctcactc tgttgcccag gctggagtgt 15240
gcagtggcgt gatcttggct cactggagcc tccatctccc aggttcaaAT gatcctcatg 15300
cctcagcctc ctgagttgct gggattacag gtgcacacca ccacgcctgg ctaatttttg 15360
tgttttcagt agagacaggg tttcaccatg ttggccaggc tggctttgaa ctctgacct 15420
caagtgatcc aacaacttca gcctcccaaa gtgctgggat gacaggcatg agccaccatg 15480
cccagcctga tntagctgtt tctgtgcaca ttatttgctg tggggtatat tcagatttct 15540
taatacaaga tgattctttg cctcatgact tacacacat tttctattta atttcagcta 15600
tgatattgga aatggacatg tcttttcaag gaaaataaaa gcaggctttc tggaatggcg 15660
acttccaaac atatttgtca atttaaagga gctgggagtg gggaccctat gctccgtaag 15720
cactctctta gctgttcttg gctgtgctcc ccgcttcagc ttcacactgc cttgctgtg 15780
aaggagcag cctgggccgg gcgcggtggc ttacacctgt aatcctagca ctttgggagg 15840
ccgaggtggg tggatcacct gaggtcagga gttcaagacc agcctggcca acatggtgaa 15900
actccatctc tactaaaaat acaaaaaatt agctgggcat ggtggcaggT gcctgtaatc 15960
ccagctactt gggaggctga ggcagaagaa tcgcttgaac ccaggaggcg gaggttcag 16020
tgagccgaga ttgcgccatt gcactccagc ctgggggcaa caagagcaaa actctgtctg 16080
gaaaaaaaaa aaaggagcag cttggcaaac ccaccttgt cgcttttgtg agtgcctctg 16140
accctttggc tgccaggacg ggcgtatttt atggaaatgc taagcaccaa cagagtaaag 16200
tggtttggtt tttcacagtg gtgggagata atagctccaa attgtctttt tcagcactga 16260
gtgaagaaat gaaagacaaa ggtggataca tgagcaagat ttgcaacttg ctaccatta 16320
ggataatgtc ttatgtaatg ctgccctgta ccctgcctgt ggaatctgcc attgcgattg 16380
tccagagggt agcattttag gtggctccgt gtcttctca cagggttgat atgaggatga 16440
aacaagatga tagatcatgg tggcatgtag tctgggacct ggattgtcgt gccacagatc 16500
acagctcaca gtctatgtgc aatgcccctg aatgttgccc acctgtctc aagccacaca 16560
tgcacctgta actcagtgcA agcccagaaa ctccccgtgg ggactcctag agctgtcagt 16620
ggcctcacat agcagctggt ccagtctctt gtgattgccc aaggaaactg aggctggag 16680
agcttggggT cactgctctg aggccataga gatgcctagt agaagggcca ggcctagaag 16740
caggatcctt gctgcccctc tgagctgttt ccatttaaaa tcacatgaag gccggcgccg 16800
tggtcacagg ctgtaatccc agcatttttg gaggccaagg tgggtggatc atgtgaggtc 16860
aggagtttga gaccagcttg gccaacatgg tgaaatgccA tctgtactaa aaatacaaaa 16920
attagtggag catggtggca cgtgcctgta ctcccagcta cttggaaggc tggggcagaa 16980
gaatcgcttg agcctgggag gcagaggttg tagtgagcca agattgtacc actgcactcc 17040
agcctgggtg acaggagaga aaccctatct caaaataaaa tgaaaggtaa tgaaatgaat 17100
aaaataataa atcaagtcac ggccgggcac ggtggctcac acctgtaatc ccagcgcttt 17160
gggaggccga ggtgggtgga taatgaggtc aggagttaa gaccagcctg gccaacatgg 17220

```

tgaaaccatg tctctactaa aaatacaaaa attagctggg catggtggtg catgcctgta 17280
atcccagcta ctccggaggc taaggcagga gaattgcttg aagcaggacc taggaggcgg 17340
aggttggttg cagtga gccg agatcatgcc actgcactct agcctgggct acagagcgaa 17400
actccgactc aaaaaaaaaa aaaaaaaaaa atcaaatcac atgaaagtag aacataggga 17460
attccatctt tcgttctagg catagtittgt taatatgatt cagagccagc agttaggaga 17520
acacagtgtg actctcctag aacttcttga ttgggcttcc tctgattggg tttcctctga 17580
ttgggcttcc tctgaaagtg ggggggatgg ggggtgggga gcagaatggt cagagcttgg 17640
ctcagcagtc agactgctct tcttcaaate ctggctgcat tgcttactac agctgtgtga 17700
ctccagatga ctgaatccac ctctctgtgc tgcagcttcc cgtctagaga gatcacctgg 17760
agcagagggt ggtcaggaga ctcaatctgg ttactgactc acagtgcagg agtactcatc 17820
ccatagtaag catccagcta gagatgttga tttctatitt caggtaataa tgatgatcgt 17880
aaaattagag acagataaaa ggtatgggca ttaggccagg gcactgcaat ttctaagctg 17940
tgtgacctca ggcaagttac tcgacttctc tgagcctcag cggtttctac cgcaatatat 18000
ggataggaaa accgacctca gtgggttgtc tgacagtgga gggcacttga ttaaaaaaaaa 18060
aaaaattacc ctggtctgaa tattaccctg gactgaaaga aaaatattga gctaatacag 18120
gcatcaggaa tggggctgca gggagtccag ggaagggaga acgaagagcc tgaaggtgtg 18180
aggaggtgag agtgctgacg tgtctgctac aaagaggctg ctgagcctcc tgtggatgtg 18240
gccctggact tggcagttta atacctgagc tgtaaaaata acctcagatg ctgtgttctt 18300
taaggggtag gattcagatt cctgctgaaa tgcttctgaa agggagggaa tgagccagcc 18360
catccccagt tgctttttta gatcattggg aagtctggtt cttgccattt gtccctggac 18420
cactcttagg tcctcctgcc ccacttccat ctgggtgtgt gccctgggct gtccaccaca 18480
cagctacatc ctgccatctt ccctcctgga gccactgtgc catgcatgga tctgtagctt 18540
catttttctt ggcttttccc tggtttttct ggagcagagt ctctagtaaa ctcccaagga 18600
agaaaacggt tgactttatg tgtgttggga aacgtgcttt ttttctatta catctcagt 18660
atagggtggc catgtctaga attgcagggt gaaaatcatt tcctctcagt atattggtta 18720
gtgagaagcc tgggactgag acagtcacat tctcattctt ttgcagggtga gtgctcttag 18780
gactgtcttt ttatccctta tactctgaaa tgtcatatgt cttgggtgtaa gtccttattt 18840
cagttattga gctggacaag tactggagac cccttcagtc aaagccttct gtcattctcc 18900
agctctagga aattatcttc tattgttatt tctgttatte cttcccttcc attttctttt 18960
ttcttttttt tttttttttt ttgagacagg gtcttactct ggtgcccagg ctggaatgca 19020
gtgacctgat catggtacac tgcagcctga acctcccaga ctcaagtgat cctcccacct 19080
caacctccta agtagctggg actgcaagca cacatcacca cacccaacaa atatttttta 19140
aaaattttgt aagatgggat ctactatgtt tgcccagact ttttcttctt cttcctgggg 19200
ctcttattag gaagatgttt gacttctggt gttggattcc tgtctccgtg tctgactttc 19260
tctctttgtc atatttttca tcaactggtt tctttttgag tctgctctga cagatttctt 19320
caaattttgt cttctagtcc tatcctacag tttttacttt cagcaaatat aatttaattt 19380
ccaagagtac tctctgttc tttttcttta gcattctgtt cttgttttat ggatgtaaca 19440
ttctcttgga atatttgctg tcctctagat catcccttct ccatttcttc ttgggctagt 19500
ttttctgttt cttcatcttt ctcttttatg ctacttattc tgggcgtgtt cttggtgggt 19560

```

```

tttttcccat atagcaacag aggacttgga gctcaggag aaaaggtag gtgcatcacc 19620
tggcagagct cccagacagt gacaggcagg ctgcgggaag gatgtctact tggcgggtgct 19680
accgctttcc tagaaaccct ttccctggag ctggttgaac tgttgggttt tgccctggtg 19740
gtgaacgctg gctccccgtg ctctgcctgt ttcatcacca gccccctccc cttctgcctg 19800
gggtccagta atctgttgaa atatatatct tgctcattgg tgagctcctg ctccttcctc 19860
gttgctcttg cagatttatc acttctcgta aggtgcgcct tgtacttcgg gattttctct 19920
gtgccacact gggaaacata ggggtggtgc atgctgcagt cctgagcact tatttcactc 19980
acatctttac acgaagattt ggtgggtggt tactttgttt ttagtaagtt agtctgtcat 20040
gtcctttgat cctttttttt tgttttttga gatggagtct ctctgtgtcc tccaggctgg 20100
agtgcaatgt cgcatctca gctcactgca acctccacct cctgggctca agagattctc 20160
ctgcttcagt ctctgagta gctgggatta caggcatgtg ccaccacacc tggctaattt 20220
ttgtattttt agtagagggt gggtttggca tgttggccag cctggctcct aactcctgac 20280
ctcctgacct gcctgccttg gcctcccaaa gtgctgggat tacagggtgtg agccaccaca 20340
cctggccctg attaatcttt taatgccag tctctccttc aaaagccggc tcctttctct 20400
ccctgcctt cctagattcc ttctccactc cccaggatca gcctcctcct cccacccca 20460
ccactgccgg ggggatgtct gtggtcaggc atttatcaga gaccctgagg tgggggtcct 20520
ttatgtgtct gggggatgga gagtctagag gaggtagcgt tcagacctct ccatggtgcc 20580
tctgctgggc tcacatgtga ccaagcacag caaacatga ggcaggggat ggtcttgacc 20640
atgagagccc ttgcagcagc tgccatgggc ctgagctcct ctccaagctg ggaagagccc 20700
tgaaaagcca aggtgttttt tttccctct ttatttcagt gtaagtcctt tgagctttct 20760
tgaaccagaa gtgggctcat ttgtcttag agatttcagg tgggcttgct cttgtcctag 20820
catcccagat ccacctctg ggaagtcac agattggagg tgatgttggc agcttttgta 20880
aacaaagggg agtgttgtaa gctgttgtgt ctgcctatgt gtgtgtttgt gtacttggtc 20940
tcactctctg agactgggtga catggcttcc agatatgcc gacgatgtcc tgtggttgca 21000
gtgggtgacc tcacagggtg tctactcagt gctgatgtgt ctgctccccg cctccaggta 21060
aatactttgg ctgtgggtgt gtgggccgga cgggcacctc tctcatctga tgaggcctca 21120
cacgacattc tagaaacagc tggtgaaca ccaagcaagg agcttgccct tgggtgtggg 21180
gacctgtct catgggaggc agctgagtc gtcagaggtc ctggcacacc tgctgagagc 21240
tgccaccag gccaacctga accggagcct gggaagactt cccgtcggat gagtctcttt 21300
gagtgcagca ttgatggtg aagagcagag aggccccaga taagcaggga aaggtgcttc 21360
agacagagt gctgggatga ggactgggga gtgtcagata gcgctggcgt gtctgagcga 21420
aggagctctg gcacccatgg cacaggaagg aggtgggacc ctggaggggc agggctagca 21480
gagctcctcg gagcgtgtgg ctaggtgcct ggtaatgcaa gccccctgtc ctccaccctc 21540
tgttgtagct agtcacagtc tccggggtga agcctagcag tctgcgttga caggccccag 21600
gggatgccgc tacttcctga attctgaatt ctggaaactg agccggagt cagggcctgg 21660
ctcccattac cagggttggg cgttatcctg aaaatcatag gccttggttt cctcacttgg 21720
ctaacagggg tgatcccat cccctcaatg ggtttccgtg agctcctgag agcccgtagc 21780
atggtagctg gcacatgct ggcatcagga ggtatggcct ctcttgctat tgttggttatt 21840
ggtagacaca gaaggattta aaagtagggg aatgcaaaga tccgatttgc tagggaagag 21900

```

ggcagtagtg gccaaagtaga ggggtggatcc tgggcccttg ctggcagcag gcagcaaggg 21960
 gggctgccag ggcccaggca gggacgatct gtagaccgag aggcttccta aggctcttgg 22020
 acaggaggag gtgtcggttc caagcctaag gagtggggca gccctggtga ctggtggtca 22080
 gtggtgccag gcggtgggtg gtaggacacc ctggcaggca agtaggtttg tgtgggggaa 22140
 actgataggc ccctccaggg attcgttggg ggacaacacc tgtgatgtcc agtgggaggt 22200
 gtccaggtag ctgggagggc cacaggcttg gaagacctag gtggtgacat cagcccagca 22260
 ctgagggcta gaagaagctg tgtctctggc tgtgacggca ccctagagtg tgtgtggtgc 22320
 cctctactgg ccggcaatgt ggggtccacc tagctcagac tgcacactgc agcagcggga 22380
 acggcctcta agccaacttc ctccatgtgt ttcaggtccc aaatgccagt gagcagccaa 22440
 caggcctccc catgcacacc tgagcaggac tggccctgct ggactccctg ctcccccaag 22500
 ggctgtccag cagagaccaa agcagaggcc accccgcggt ccatectcag gtccagcctg 22560
 aacttcttct tgggcaataa agtacctgct ggtgctgagg ggctctccac ctttcccagt 22620
 ttttactag agaagagtct gtgagtcact tgaggaggcg agtctagcag attctttcag 22680
 aggtgctaaa gtttcccatc tttgtgcagc tacctccgca ttgctgtgta gtgaccctg 22740
 cctgtgacgt ggaggatccc agcctctgag ctgagttggt tttatgaaaa gctaggaagc 22800
 aacctttcgc ctgtgcagcg gtccagcact taactctaata acatcagcat gcgttaattc 22860
 agctggttgg gaaatgacac caggaagccc agtgcagagg gtcccttact gactgtttcg 22920
 tggccctatt aatggtcaga ctgttccagc atgaggttct tagaatgaca ggtgtttgga 22980
 tgggtggggg ccttgtgatg gggggtaggc tggcccatgt gtgatcttgt ggggtggagg 23040
 gaagagaata gcatgatccc acttccccat gctgtgggaa ggggtgcagt tcgtcccaa 23100
 gaacgacact gcctgtcagg tggctctgaa agatgataac cttgactact aaaaacgtct 23160
 ccatggcggg ggtaacaaga tgataatcta cttaatTTTA gaacaccttt ttcacctaac 23220
 taaaataatg tttaaagagt tttgtataaa aatgtaagga agcgttgTTA cctgttgaat 23280
 tttgtattat gtgaatcagt gagatgttag tagaataagc cttaaaaaaa aaaaaatcgg 23340
 ttgggtgcag tggcacacgg ctgtaatccc agcactttgg gaggccaagg ttggcagatc 23400
 acctgaggtc aggagttaa gaccagtctg gccaacatag caaaaccctg tctctactaa 23460
 aaatacaaaa attatctggg catggtggtg catgcctgta atcccagcta ttcggaaggc 23520
 tgaggcagga gaatcacttg aaccaggag gcggagggtg cgggtgagctg agattgcacc 23580
 atttcattcc agcctgggca acatgagtga aagtctgact caaaaaaaaa aaatttaaaa 23640
 aacaaaaataa tctagtgtgc agggcattca ctcagcccc ccaggcagga gccaaagcaca 23700
 gcaggagctt ccgcctcctc tccactggag cacacaactt gaacctggct tattttctgc 23760
 agggaccagc cccacatggt cagtgagttt ctccccatgt gtggcgatga gagagtgtag 23820
 aaataaagac 23830
 <210> 32
 <211> 1443
 <212> DNA
 <213> 智人(Homo sapiens)
 <400> 32
 auguacgacg cagagcgcgg cuggagcuug uccuucgchg gcugcgguu ccugggcuuc 60

uaccacgucg gggcgacccg cugccugagc gagcacgccc cgcaccuccu ccgcgacgcg 120
 cgcauguugu ucggcgcuuc ggccggggcg uugcacugcg ucggcguccu cuccgguauc 180
 ccgcuggagc agacucugca gguccucuca gaucuugugc ggaaggccag gagucggaac 240
 auuggcaucu uccaaccauc cuucaacuua agcaaguucc uccgacaggg ucucugcaaa 300
 ugccucccg ccaaugucca ccagcucauc uccggcaaaa uaggcaucuc ucuuaccaga 360
 gugucugaug gggaaaacgu ucuggugucu gacuuucggu ccaaagacga agucguggau 420
 gccuugguau guuccugcuu cauccccuuc uacaguggcc uuauccucc uuccuucaga 480
 ggcgugcgau auguggaugg aggagugagu gacaacguac ccuucauuga ugccaaaaca 540
 accaucaccg ugucucccuu cuauggggag uacgacauuc gcccuaaagu caaguccacg 600
 aacuuucuuuc auguggacau caccaagcuc agucuaagcc ucugcacagg gaaccucuac 660
 cuucucucga gagcuuuugu cccccggau cucaaggugc ugggagagau augccuucga 720
 ggauauuugg augcauucag guucuuggaa gagaaggga ucugcaacag gccccagcca 780
 ggccugaagu cauccucaga agggauggau ccugaggucg ccaugcccag cugggcaaac 840
 augagucugg auucuucccc ggagucggcu gccuuggcug ugaggcugga gggagaugag 900
 cugcuagacc accugcgucu cagcauccug ccugggagug agagcauccu ggacaccuc 960
 ucgccaggc ucgcuacagc acugagugaa gaaaugaaag acaaaggugg auacaugagc 1020
 aagauuugca acuugcuacc cauuaaggau augucuuaug uaaugcugcc cuguaccug 1080
 ccuguggaau cugccauugc gauuguccag agacugguga cauggcuucc agauaugccc 1140
 gacgaugucc ugugguugca gugggugacc ucacaggugu ucacucgagu gcugaugugu 1200
 cugcuccccg ccuccagguc ccaaaucca gugagcagcc aacaggccuc ccaugcaca 1260
 ccugagcagg acuggcccug cuggacuccc ugcucccca agggcugucc agcagagacc 1320
 aaagcagagg ccaccccgcg guccauccuc agguccagcc ugaacuucu cuugggcaau 1380
 aaaguaccug cuggugcuga ggggcucucc accuuucca guuuuucacu agagaagagu 1440
 cug 1443

<210> 33

<211> 1431

<212> DNA

<213> 智人(Homo sapiens)

<400> 33

auguacgacg cagagcgcg cuggagcuug uccuucgcgg gcugcggcuu ccugggcuuc 60
 uaccacgucg gggcgacccg cugccugagc gagcacgccc cgcaccuccu ccgcgacgcg 120
 cgcauguugu ucggcgcuuc ggccggggcg uugcacugcg ucggcguccu cuccgagcag 180
 acucugcagg uccucucaga ucuugugcgg aaggccagga gucggacaau uggcaucuuc 240
 cauccauccu ucaacuuaag caaguuccuc cgacaggguc ucugcaaaug ccucccgcc 300
 aauguccacc agcucaucuc cggcaaaaau ggcaucucuc uuaccagagu gucugauggg 360
 gaaaacguuc uggugucuga cuuucggucc aaagacgaag ucguggaugc cuugguau 420
 uccugcuuca ucccuucua caguggccuu auccuccuu ccuucagagg cgugcgauau 480
 guggauggag gagugaguga caacguaccc uucauugaug ccaaaacaac caucaccgug 540
 ucccuucuu auggggagua cgacaucugc ccuaaaguca aguccacgaa cuuucuuc 600

guggacauca ccaagcucag ucuacgccuc ugcacaggga accucuaccu ucucucgaga 660
 gcuuuugucc ccccggauuc caaggugcug ggagagauau gccuucgagg auauuuggau 720
 gcauucaggu ucuuggaaga gaagggcauc ugcaacaggc cccagccagg ccugaaguca 780
 uccucagaag ggauggaucc ugaggucgcc augcccagcu gggcaaacau gagucuggau 840
 ucuuucccgagg agucggcugc cuuggcugug aggcuggagg gagaugagcu gcuagaccac 900
 cugcgucuca gcauccugcc cuggggaugag agcauccugg acaccucuc gcccaggcuc 960
 gcuacagcac ugagugaaga aaugaaagac aaagguggau acaugagcaa gauuugcaac 1020
 uugcuaccca uuaggauaau gucuuaugua augcugcccu guaccucgcc uguggaauuc 1080
 gccauugcga uuguccagag acuggugaca uggcuuccag auaugcccga cgauguccug 1140
 ugguugcagu gggugaccuc acagguguuc acucgagugc ugaugugucu gcuccccgcc 1200
 uccagguccc aaaugccagu gagcagccaa caggccuccc caugcacacc ugagcaggac 1260
 uggcccugcu ggacucccug cucccccaag ggcuguccag cagagaccaa agcagaggcc 1320
 accccgcggg ccauccucag guccagccug aacuucuuuc ugggcaauaa aguaccugcu 1380
 ggugcugagg ggcucuccac cuuucccagu uuucacuaag agaagagucu g 1431

<210> 34

<211> 1443

<212> DNA

<213> 智人(Homo sapiens)

<400> 34

auguacgacg cagagcgagg cuggagcuug uccuucgcgg gcugcggcuu ccugggcuuc 60
 uaccacgucg gggcgacccg cugccugagc gagcacgccc cgcaccuccu ccgcgacgag 120
 cgcauguugu ucggcgcuuc ggccggggcg uugcacugcg ucggcguccu cuccgguauc 180
 ccgcuggagc agacucugca gguccucuca gaucuuugugc ggaaggccag gagucggaac 240
 auuggcaucu uccauccauc cuucaacuua agcaaguucc uccgacaggg ucucugcaaa 300
 ugccucccgg ccaaugucca ccagcucauc uccggcaaaa uaggcaucuc ucuuaccaga 360
 gugucugaug gggaaaacgu ucuggugucu gacuuucggu ccaaagacga agucguggau 420
 gccuugguau guuccugcuu caugcccuuc uacaguggcc uuauccucc uuccuucaga 480
 ggcgugcgau auguggaugg aggagugagu gacaacguac ccuucuuuga ugccaaaaca 540
 accaucaccg ugucucccuu cuauggggag uacgacauuc gcccuaaaagu caaguccacg 600
 aacuucuuuc auguggacau caccaagcuc agucuaagcc ucugcacagg gaaccucuaac 660
 cuucucucga gagcuuuugu cccccggau cucaaggugc ugggagagau augccuucga 720
 ggauauuugg augcauucag guucuuugga gagaagggca ucugcaacag gccccagcca 780
 ggccugaagu cauccucaga agggauuggau ccugaggucg ccaugcccag cugggcaaac 840
 augagucugg auucuuuccc ggagucggcu gccuuggcug ugaggcugga gggagaugag 900
 cugcuagacc accugcgucu cagcauccug ccuugggaug agagcauccu ggacaccuc 960
 ucgccaggc ucgcuaacag acugagugaa gaaaugaaag acaaaggugg auacaugagc 1020
 aagauuugca acuugcuacc cauuaaggau augucuuuug uaaugcugcc cuguaccug 1080
 ccuguggaau cugccauugc gauuguccag agacugguga cauggcuucc agauaugccc 1140
 gacgaugucc ugugguugca gugggugacc ucacaggugu ucacucgagu gcugaugugu 1200

cugcucucccg ccuccagguc ccaaaugcca gugagcagcc aacaggccuc cccaugcaca 1260
 ccugagcagg acuggccug cuggacucc ugcucuccca agggcugucc agcagagacc 1320
 aaagcagagg ccaccccgcg guccauccuc agguccagcc ugaacuucuu cuugggcaau 1380
 aaaguaccug cuggugcuga ggggcucucc accuuucca guuuuucacu agagaagagu 1440
 cug 1443

<210> 35

<211> 1431

<212> DNA

<213> 智人 (Homo sapiens)

<400> 35

auguacgacg cagagcgcg cuggagcuug uccuucgcgg gcugcggcuu ccugggcuuc 60
 uaccacgucg gggcgacccg cugccugagc gagcacgccc cgcaccuccu ccgcgacgcg 120
 cgcauguugu ucggcgcuuc ggccggggcg uugcacugcg ucggcguccu cuccgagcag 180
 acucugcagg uccucucaga ucuugugcgg aaggccagga gucggaacau uggcaucuuc 240
 cauccauccu ucaacuuaag caaguuccuc cgacaggguc ucugcaaaug ccucccgcc 300
 aauguccacc agcucaucuc cggcaaaaau ggcaucucuc uuaccagagu gucugauggg 360
 gaaaacguuc uggugucuga cuuucggucc aaagacgaag ucguggaugc cuugguangu 420
 uccugcuuca ugcccuucua caguggccuu auccuccuu ccuucagagg cgugcgauau 480
 guggauggag gagugaguga caacguaccc uucauugaug ccaaaacaac caucaccgug 540
 ucccccuuuc auggggagua cgacaucugc ccuaaaguca aguccacgaa cuuucuucau 600
 guggacauca ccaagcucag ucuacgccuc ugcacaggga accucuaccu ucucucgaga 660
 gcuuuugucc ccccggaucu caaggugcug ggagagauau gccuucgagg auauuuggau 720
 gcauucaggu ucuuggaaga gaagggauc ugcaacaggc cccagccagg ccugaaguca 780
 uccucagaag ggauggaucc ugaggucgcc augcccagcu gggcaaacu gagucuggau 840
 ucuuccccgg agucggcugc cuuggcugug aggcuggagg gagaugagcu gcuagaccac 900
 cugcgucuca gcauccugcc cugggaugag agcauccugg acaccucuc gcccaggcuc 960
 gcuacagcac ugagugaaga aaugaaagac aaagguggau acaugagcaa gauuugcaac 1020
 uugcuaccca uuaggauaau gucuuaugua augcugcccu guaccugcc uguggaauuc 1080
 gccauugcga uuguccagag acuggugaca uggcuuccag auaugcccga cgauguccug 1140
 ugguugcagu gggugaccuc acagguguuc acucgagugc ugaugugucu gcucccgcc 1200
 uccaggucucc aaaugccagu gagcagccaa caggccucc caugcacacc ugagcaggac 1260
 uggccucgu ggacuccug cuccccaag ggcuguccag cagagaccaa agcagaggcc 1320
 accccgcggu ccauccucag guccagccug aacuucuuu ugggcaauaa aguaccugcu 1380
 ggugcugagg ggcucuccac cuuucccagu uuucacuag agaagagucu g 1431

<210> 36

<211> 1443

<212> DNA

<213> 智人 (Homo sapiens)

<400> 36

atgtacgacg cagagcgcggt ctggagcttg tccttcgcgg gctgcggctt cctgggcttc 60
 taccacgtcg gggcgacccg ctgcctgagc gagcacgccc cgcacctcct ccgcgacgcg 120
 cgcatgttgt tcggcgcttc ggccggggcg ttgcaactgc tcggcgctcct ctccggtatc 180
 ccgctggagc agactctgca ggtcctctca gatcttgtgc ggaaggccag gactcggaac 240
 attggcatct tccatccatc cttcaactta agcaagttcc tccgacaggg tctctgcaaa 300
 tgcctccccg ccaatgtcca ccagctcatc tccggcaaaa taggcatctc tcttaccaga 360
 gtgtctgatg gggaaaacgt tctggtgtct gactttcggt ccaaagacga agtcgtggat 420
 gccttggtat gttcctgctt catcccttc tacagtggcc ttatccctcc ttccttcaga 480
 ggcgtgcgat atgtggatgg aggagtgagt gacaacgtac ccttcattga tgccaaaaca 540
 accatcacccg tgtccccctt ctatggggag tacgacatct gccctaaagt caagtccacg 600
 aactttcttc atgtggacat caccaagctc agtctacgcc tctgcacagg gaacctctac 660
 cttctctcga gagcttttgt cccccggat ctcaaggtgc tgggagagat atgccttcga 720
 ggatatttgg atgcattcag gttcttggaa gagaaggga tctgcaacag gccccagcca 780
 ggctgaagt catcctcaga agggatggat cctgaggtcg ccatgcccag ctgggcaaac 840
 atgagtctgg attcttcccc ggagtcggct gccttggtg tgaggctgga gggagatgag 900
 ctgctagacc acctgcgtct cagcatcctg ccctgggatg agagcatcct ggacaccctc 960
 tcgccaggc tcgctacagc actgagtga gaaatgaaag acaaaggtgg atacatgagc 1020
 aagatttgca acttgctacc cattaggata atgtcttatg taatgctgcc ctgtaccctg 1080
 cctgtggaat ctgccattgc gattgtccag agactggtga catggcttcc agatatgccc 1140
 gacgatgtcc tgtggttgca gtgggtgacc tcacaggtgt tcaactcagt gctgatgtgt 1200
 ctgctccccg cctccagtc ccaaagcca gtgagcagcc aacaggcctc cccatgcaca 1260
 cctgagcagg actggccctg ctggactccc tgctccccca agggctgtcc agcagagacc 1320
 aaagcagagg ccaccccgcg gtccatcctc aggtccagcc tgaacttctt cttgggcaat 1380
 aaagtacctg ctggtgctga ggggctctcc acctttccca gtttttact agagaagagt 1440
 ctg 1443

<210> 37

<211> 1431

<212> DNA

<213> 智人(Homo sapiens)

<400> 37

atgtacgacg cagagcgcggt ctggagcttg tccttcgcgg gctgcggctt cctgggcttc 60
 taccacgtcg gggcgacccg ctgcctgagc gagcacgccc cgcacctcct ccgcgacgcg 120
 cgcatgttgt tcggcgcttc ggccggggcg ttgcaactgc tcggcgctcct ctccgagcag 180
 actctgcagg tcctctcaga tcttgtgcgg aaggccagga gtcggaacat tggcatcttc 240
 catccatcct tcaacttaag caagtctcct cgacagggtc tctgcaaatg cctcccggcc 300
 aatgtccacc agctcatctc cggcaaaaata ggcattctct ttaccagagt gtctgatggg 360
 gaaaacgttc tgggtgtctga ctttcggtcc aaagacgaag tcgtggatgc cttggtatgt 420
 tcctgcttca tcccccttca cagtggcctt atccctcctt ccttcagagg cgtgcgatat 480
 gtggatggag gactgagtga caacgtaccc ttcatgatg ccaaaacaac catcacctg 540

tcccccttct atggggagta cgacatctgc cctaaagtca agtccacgaa ctttcttcat 600
 gtggacatca ccaagctcag tctacgcctc tgcacaggga acctctacct tctctcgaga 660
 gcttttgtcc ccccggatct caaggtgctg ggagagatat gccttcgagg atatttggat 720
 gcattcaggt tcttgaaga gaagggcatc tgcaacaggc cccagccagg cctgaagtca 780
 tcctcagaag ggatggatcc tgaggtcgcc atgccagct gggcaaacaat gagtctggat 840
 tcttccccgg agtcggctgc cttggctgtg aggtctggagg gagatgagct gctagaccac 900
 ctgcgtctca gcatcctgcc ctgggatgag agcatcctgg acaccctctc gccaggctc 960
 gctacagcac tgagtgaaga aatgaaagac aaaggtggat acatgagcaa gatttgcaac 1020
 ttgctaccca ttaggataat gtcttatgta atgctgccct gtaccctgcc tgtggaatct 1080
 gccattgcga ttgtccagag actggtgaca tggcttcag atatgccga cgatgtcctg 1140
 tggttgcagt gggtagacct acaggtgttc actcgagtgc tgatgtgtct gctccccgcc 1200
 tccaggtccc aaatgccagt gagcagccaa caggcctccc catgcacacc tgagcaggac 1260
 tggccctgct ggactccctg ctcccccaag ggctgtccag cagagaccaa agcagaggcc 1320
 accccgcggt ccatcctcag gtccagcctg aacttcttct tgggcaataa agtacctgct 1380
 ggtgtgagg ggctctccac ctttcccagt ttttactag agaagagtct g 1431

<210> 38

<211> 1443

<212> DNA

<213> 智人(Homo sapiens)

<400> 38

atgtacgacg cagagcgcggt ctggagcttg tccttcgcgg gctgcggctt cctgggcttc 60
 taccacgtcg gggcgacccg ctgcctgagc gagcacgcc cgcacctcct ccgcgacgcg 120
 cgcatgttgt tcggcgcttc ggccggggcg ttgcaactgc tcggcgctcct ctccggtatc 180
 ccgctggagc agactctgca ggtcctctca gatcttgtgc ggaaggccag gagtgcgaac 240
 attggcatct tccatccatc cttcaactta agcaagttcc tccgacaggg tctctgcaaa 300
 tgccctcccg ccaatgtcca ccagctcatc tccggcaaaa taggcatctc tcttaccaga 360
 gtgtctgatg gggaaaacgt tctggtgtct gactttcggg ccaaagacga agtcgtggat 420
 gccttggtat gttcctgctt catgcccttc tacagtggcc ttatccctcc ttccttcaga 480
 ggcgtgcgat atgtggatgg aggagtgagt gacaacgtac ctttcattga tgccaaaaca 540
 accatcaccc tgccccctt ctatggggag tacgacatct gccctaaagt caagtccacg 600
 aactttcttc atgtggacat caccaagctc agtctacgcc tctgcacagg gaacctctac 660
 cttctctcga gagcttttgt cccccggat ctcaaggtgc tgggagagat atgccttcga 720
 ggatatattg atgcattcag gttcttggaa gagaagggca tctgcaacag gccccagcca 780
 ggccgaagt catcctcaga agggatggat cctgaggtcg ccatgccag ctgggcaaac 840
 atgagtctgg attcttcccc ggagtcggct gccttggctg tgaggctgga gggagatgag 900
 ctgctagacc acctgcgtct cagcatcctg ccctgggatg agagcatcct ggacaccctc 960
 tcgccaggc tcgctacagc actgagtga gaaatgaaag acaaaggtgg atacatgagc 1020
 aagatttgca acttgctacc cattagata atgtcttatg taatgctgcc ctgtaccctg 1080
 cctgtggaat ctgccattgc gattgtccag agactggtga catggcttcc agatatgcc 1140

gacgatgtcc tgtgggttgca gtgggtgacc tcacaggtgt tcactcgagt gctgatgtgt 1200
 ctgctccccg cctccaggtc ccaaatgcc a gtgagcagcc aacaggcctc cccatgcaca 1260
 cctgagcagg actggccctg ctggactccc tgctccccca agggctgtcc agcagagacc 1320
 aaagcagagg ccaccccgcg gtccatcctc aggtccagcc tgaacttctt cttgggcaat 1380
 aaagtacctg ctggtgctga ggggctctcc acctttccca gtttttctact agagaagagt 1440
 ctg 1443

<210> 39

<211> 1431

<212> DNA

<213> 智人 (Homo sapiens)

<400> 39

atgtacgacg cagagcgcggt ctggagcttg tccttcgcgg gctgcggctt cctgggcttc 60
 taccacgtcg gggcgacccg ctgcctgagc gagcacgccc cgcacctcct ccgcgacgcg 120
 cgcattgtgt tcggcgcttc ggccggggcg ttgcaactgc tcggcgctcct ctccgagcag 180
 actctgcagg tcctctcaga tcttgtgcgg aaggccagga gtcggaacat tggcatcttc 240
 catccatcct tcaacttaag caagtctcct cgacagggtc tctgcaaatg cctccccggc 300
 aatgtccacc agctcatctc cggcaaaata ggcattctct ttaccagagt gtctgatggg 360
 gaaaacgttc tgggtgtctga ctttcggtcc aaagacgaag tcgtggatgc cttggtatgt 420
 tcctgcttca tgcccttcta cagtggcctt atccctcctt ccttcagagg cgtgcgatat 480
 gtggatggag gaggtagtga caacgtaccc ttcatgtatg ccaaaacaac catcaccgtg 540
 tcccccttct atggggagta cgacatctgc cctaaagtca agtccacgaa ctttcttcat 600
 gtggacatca ccaagctcag tctacgcctc tgcacaggga acctctacct tctctcgaga 660
 gcttttgtcc ccccgatct caaggtgctg ggagagatat gccttcgagg atatttggat 720
 gcattcaggt tcttgaaga gaaggcatc tgcaacaggc ccagccagg cctgaagtca 780
 tcctcagaag ggatggatcc tgaggtcgcc atgccagct gggcaaacat gaggctggat 840
 tcttcccccg agtcggctgc cttggctgtg aggctggagg gagatgagct gctagaccac 900
 ctgcgtctca gcatcctgcc ctgggatgag agcatcctgg acacctctc gccaggctc 960
 gctacagcac tgagtgaaga aatgaaagac aaaggtggat acatgagcaa gatttgcaac 1020
 ttgctaccca ttaggataat gtcttatgta atgctgccct gtaccctgcc tgtggaatct 1080
 gccattgcga ttgtccagag actggtgaca tggcttccag atatgcccga cgatgtcctg 1140
 tggttgcagt gggtagacct acaggtgttc actcgagtgc tgatgtgtct gctccccgcc 1200
 tccagggtccc aaatgccagt gagcagccaa caggcctccc catgcacacc tgagcaggac 1260
 tggccctgct ggactccctg ctcccccaag ggctgtccag cagagaccaa agcagaggcc 1320
 accccgcggt ccattcctcag gtccagcctg aactttctt tgggcaataa agtacctgct 1380
 ggtgctgagg ggctctccac ctttcccagt ttttactag agaagagtct g 1431

<210> 40

<211> 481

<212> PRT

<213> 智人 (Homo sapiens)

<400> 40

```

Met Tyr Asp Ala Glu Arg Gly Trp Ser Leu Ser Phe Ala Gly Cys Gly
1           5           10           15
Phe Leu Gly Phe Tyr His Val Gly Ala Thr Arg Cys Leu Ser Glu His
           20           25           30
Ala Pro His Leu Leu Arg Asp Ala Arg Met Leu Phe Gly Ala Ser Ala
           35           40           45
Gly Ala Leu His Cys Val Gly Val Leu Ser Gly Ile Pro Leu Glu Gln
           50           55           60
Thr Leu Gln Val Leu Ser Asp Leu Val Arg Lys Ala Arg Ser Arg Asn
65           70           75           80
Ile Gly Ile Phe His Pro Ser Phe Asn Leu Ser Lys Phe Leu Arg Gln
           85           90           95
Gly Leu Cys Lys Cys Leu Pro Ala Asn Val His Gln Leu Ile Ser Gly
           100          105          110
Lys Ile Gly Ile Ser Leu Thr Arg Val Ser Asp Gly Glu Asn Val Leu
           115          120          125
Val Ser Asp Phe Arg Ser Lys Asp Glu Val Val Asp Ala Leu Val Cys
           130          135          140
Ser Cys Phe Ile Pro Phe Tyr Ser Gly Leu Ile Pro Pro Ser Phe Arg
145          150          155          160
Gly Val Arg Tyr Val Asp Gly Gly Val Ser Asp Asn Val Pro Phe Ile
           165          170          175
Asp Ala Lys Thr Thr Ile Thr Val Ser Pro Phe Tyr Gly Glu Tyr Asp
           180          185          190
Ile Cys Pro Lys Val Lys Ser Thr Asn Phe Leu His Val Asp Ile Thr
           195          200          205
Lys Leu Ser Leu Arg Leu Cys Thr Gly Asn Leu Tyr Leu Leu Ser Arg
           210          215          220
Ala Phe Val Pro Pro Asp Leu Lys Val Leu Gly Glu Ile Cys Leu Arg
225          230          235          240
Gly Tyr Leu Asp Ala Phe Arg Phe Leu Glu Glu Lys Gly Ile Cys Asn
           245          250          255
Arg Pro Gln Pro Gly Leu Lys Ser Ser Ser Glu Gly Met Asp Pro Glu
           260          265          270
Val Ala Met Pro Ser Trp Ala Asn Met Ser Leu Asp Ser Ser Pro Glu
           275          280          285
Ser Ala Ala Leu Ala Val Arg Leu Glu Gly Asp Glu Leu Leu Asp His
           290          295          300

```

Leu Arg Leu Ser Ile Leu Pro Trp Asp Glu Ser Ile Leu Asp Thr Leu
 305 310 315 320
 Ser Pro Arg Leu Ala Thr Ala Leu Ser Glu Glu Met Lys Asp Lys Gly
 325 330 335
 Gly Tyr Met Ser Lys Ile Cys Asn Leu Leu Pro Ile Arg Ile Met Ser
 340 345 350
 Tyr Val Met Leu Pro Cys Thr Leu Pro Val Glu Ser Ala Ile Ala Ile
 355 360 365
 Val Gln Arg Leu Val Thr Trp Leu Pro Asp Met Pro Asp Asp Val Leu
 370 375 380
 Trp Leu Gln Trp Val Thr Ser Gln Val Phe Thr Arg Val Leu Met Cys
 385 390 395 400
 Leu Leu Pro Ala Ser Arg Ser Gln Met Pro Val Ser Ser Gln Gln Ala
 405 410 415
 Ser Pro Cys Thr Pro Glu Gln Asp Trp Pro Cys Trp Thr Pro Cys Ser
 420 425 430
 Pro Lys Gly Cys Pro Ala Glu Thr Lys Ala Glu Ala Thr Pro Arg Ser
 435 440 445
 Ile Leu Arg Ser Ser Leu Asn Phe Phe Leu Gly Asn Lys Val Pro Ala
 450 455 460
 Gly Ala Glu Gly Leu Ser Thr Phe Pro Ser Phe Ser Leu Glu Lys Ser
 465 470 475 480
 Leu
 <210> 41
 <211> 477
 <212> PRT
 <213> 智人(Homo sapiens)
 <400> 41
 Met Tyr Asp Ala Glu Arg Gly Trp Ser Leu Ser Phe Ala Gly Cys Gly
 1 5 10 15
 Phe Leu Gly Phe Tyr His Val Gly Ala Thr Arg Cys Leu Ser Glu His
 20 25 30
 Ala Pro His Leu Leu Arg Asp Ala Arg Met Leu Phe Gly Ala Ser Ala
 35 40 45
 Gly Ala Leu His Cys Val Gly Val Leu Ser Glu Gln Thr Leu Gln Val
 50 55 60
 Leu Ser Asp Leu Val Arg Lys Ala Arg Ser Arg Asn Ile Gly Ile Phe
 65 70 75 80
 His Pro Ser Phe Asn Leu Ser Lys Phe Leu Arg Gln Gly Leu Cys Lys

				85					90					95			
Cys	Leu	Pro	Ala	Asn	Val	His	Gln	Leu	Ile	Ser	Gly	Lys	Ile	Gly	Ile		
				100					105					110			
Ser	Leu	Thr	Arg	Val	Ser	Asp	Gly	Glu	Asn	Val	Leu	Val	Ser	Asp	Phe		
				115					120					125			
Arg	Ser	Lys	Asp	Glu	Val	Val	Asp	Ala	Leu	Val	Cys	Ser	Cys	Phe	Ile		
				130					135					140			
Pro	Phe	Tyr	Ser	Gly	Leu	Ile	Pro	Pro	Ser	Phe	Arg	Gly	Val	Arg	Tyr		
145						150				155					160		
Val	Asp	Gly	Gly	Val	Ser	Asp	Asn	Val	Pro	Phe	Ile	Asp	Ala	Lys	Thr		
				165					170					175			
Thr	Ile	Thr	Val	Ser	Pro	Phe	Tyr	Gly	Glu	Tyr	Asp	Ile	Cys	Pro	Lys		
				180					185					190			
Val	Lys	Ser	Thr	Asn	Phe	Leu	His	Val	Asp	Ile	Thr	Lys	Leu	Ser	Leu		
				195					200					205			
Arg	Leu	Cys	Thr	Gly	Asn	Leu	Tyr	Leu	Leu	Ser	Arg	Ala	Phe	Val	Pro		
				210					215					220			
Pro	Asp	Leu	Lys	Val	Leu	Gly	Glu	Ile	Cys	Leu	Arg	Gly	Tyr	Leu	Asp		
225						230				235					240		
Ala	Phe	Arg	Phe	Leu	Glu	Glu	Lys	Gly	Ile	Cys	Asn	Arg	Pro	Gln	Pro		
				245					250					255			
Gly	Leu	Lys	Ser	Ser	Ser	Glu	Gly	Met	Asp	Pro	Glu	Val	Ala	Met	Pro		
				260					265					270			
Ser	Trp	Ala	Asn	Met	Ser	Leu	Asp	Ser	Ser	Pro	Glu	Ser	Ala	Ala	Leu		
				275					280					285			
Ala	Val	Arg	Leu	Glu	Gly	Asp	Glu	Leu	Leu	Asp	His	Leu	Arg	Leu	Ser		
				290					295					300			
Ile	Leu	Pro	Trp	Asp	Glu	Ser	Ile	Leu	Asp	Thr	Leu	Ser	Pro	Arg	Leu		
305						310				315					320		
Ala	Thr	Ala	Leu	Ser	Glu	Glu	Met	Lys	Asp	Lys	Gly	Gly	Tyr	Met	Ser		
				325					330					335			
Lys	Ile	Cys	Asn	Leu	Leu	Pro	Ile	Arg	Ile	Met	Ser	Tyr	Val	Met	Leu		
				340					345					350			
Pro	Cys	Thr	Leu	Pro	Val	Glu	Ser	Ala	Ile	Ala	Ile	Val	Gln	Arg	Leu		
				355					360					365			
Val	Thr	Trp	Leu	Pro	Asp	Met	Pro	Asp	Asp	Val	Leu	Trp	Leu	Gln	Trp		
				370					375					380			
Val	Thr	Ser	Gln	Val	Phe	Thr	Arg	Val	Leu	Met	Cys	Leu	Leu	Pro	Ala		
385						390				395					400		

Ser	Arg	Ser	Gln	Met	Pro	Val	Ser	Ser	Gln	Gln	Ala	Ser	Pro	Cys	Thr		
				405					410					415			
Pro	Glu	Gln	Asp	Trp	Pro	Cys	Trp	Thr	Pro	Cys	Ser	Pro	Lys	Gly	Cys		
				420					425					430			
Pro	Ala	Glu	Thr	Lys	Ala	Glu	Ala	Thr	Pro	Arg	Ser	Ile	Leu	Arg	Ser		
				435					440					445			
Ser	Leu	Asn	Phe	Phe	Leu	Gly	Asn	Lys	Val	Pro	Ala	Gly	Ala	Glu	Gly		
				450					455					460			
Leu	Ser	Thr	Phe	Pro	Ser	Phe	Ser	Leu	Glu	Lys	Ser	Leu					
				465					470					475			
<210> 42																	
<211> 481																	
<212> PRT																	
<213> 智人(Homo sapiens)																	
<400> 42																	
Met	Tyr	Asp	Ala	Glu	Arg	Gly	Trp	Ser	Leu	Ser	Phe	Ala	Gly	Cys	Gly		
1					5					10					15		
Phe	Leu	Gly	Phe	Tyr	His	Val	Gly	Ala	Thr	Arg	Cys	Leu	Ser	Glu	His		
				20					25					30			
Ala	Pro	His	Leu	Leu	Arg	Asp	Ala	Arg	Met	Leu	Phe	Gly	Ala	Ser	Ala		
				35					40					45			
Gly	Ala	Leu	His	Cys	Val	Gly	Val	Leu	Ser	Gly	Ile	Pro	Leu	Glu	Gln		
				50					55					60			
Thr	Leu	Gln	Val	Leu	Ser	Asp	Leu	Val	Arg	Lys	Ala	Arg	Ser	Arg	Asn		
65					70					75					80		
Ile	Gly	Ile	Phe	His	Pro	Ser	Phe	Asn	Leu	Ser	Lys	Phe	Leu	Arg	Gln		
				85					90					95			
Gly	Leu	Cys	Lys	Cys	Leu	Pro	Ala	Asn	Val	His	Gln	Leu	Ile	Ser	Gly		
				100					105					110			
Lys	Ile	Gly	Ile	Ser	Leu	Thr	Arg	Val	Ser	Asp	Gly	Glu	Asn	Val	Leu		
				115					120					125			
Val	Ser	Asp	Phe	Arg	Ser	Lys	Asp	Glu	Val	Val	Asp	Ala	Leu	Val	Cys		
				130					135					140			
Ser	Cys	Phe	Met	Pro	Phe	Tyr	Ser	Gly	Leu	Ile	Pro	Pro	Ser	Phe	Arg		
145					150					155					160		
Gly	Val	Arg	Tyr	Val	Asp	Gly	Gly	Val	Ser	Asp	Asn	Val	Pro	Phe	Ile		
				165					170					175			
Asp	Ala	Lys	Thr	Thr	Ile	Thr	Val	Ser	Pro	Phe	Tyr	Gly	Glu	Tyr	Asp		
				180					185					190			

Ile Cys Pro Lys Val Lys Ser Thr Asn Phe Leu His Val Asp Ile Thr		
195	200	205
Lys Leu Ser Leu Arg Leu Cys Thr Gly Asn Leu Tyr Leu Leu Ser Arg		
210	215	220
Ala Phe Val Pro Pro Asp Leu Lys Val Leu Gly Glu Ile Cys Leu Arg		
225	230	235 240
Gly Tyr Leu Asp Ala Phe Arg Phe Leu Glu Glu Lys Gly Ile Cys Asn		
	245	250 255
Arg Pro Gln Pro Gly Leu Lys Ser Ser Ser Glu Gly Met Asp Pro Glu		
	260	265 270
Val Ala Met Pro Ser Trp Ala Asn Met Ser Leu Asp Ser Ser Pro Glu		
	275	280 285
Ser Ala Ala Leu Ala Val Arg Leu Glu Gly Asp Glu Leu Leu Asp His		
	290	295 300
Leu Arg Leu Ser Ile Leu Pro Trp Asp Glu Ser Ile Leu Asp Thr Leu		
305	310	315 320
Ser Pro Arg Leu Ala Thr Ala Leu Ser Glu Glu Met Lys Asp Lys Gly		
	325	330 335
Gly Tyr Met Ser Lys Ile Cys Asn Leu Leu Pro Ile Arg Ile Met Ser		
	340	345 350
Tyr Val Met Leu Pro Cys Thr Leu Pro Val Glu Ser Ala Ile Ala Ile		
	355	360 365
Val Gln Arg Leu Val Thr Trp Leu Pro Asp Met Pro Asp Asp Val Leu		
	370	375 380
Trp Leu Gln Trp Val Thr Ser Gln Val Phe Thr Arg Val Leu Met Cys		
385	390	395 400
Leu Leu Pro Ala Ser Arg Ser Gln Met Pro Val Ser Ser Gln Gln Ala		
	405	410 415
Ser Pro Cys Thr Pro Glu Gln Asp Trp Pro Cys Trp Thr Pro Cys Ser		
	420	425 430
Pro Lys Gly Cys Pro Ala Glu Thr Lys Ala Glu Ala Thr Pro Arg Ser		
	435	440 445
Ile Leu Arg Ser Ser Leu Asn Phe Phe Leu Gly Asn Lys Val Pro Ala		
	450	455 460
Gly Ala Glu Gly Leu Ser Thr Phe Pro Ser Phe Ser Leu Glu Lys Ser		
465	470	475 480
Leu		
<210> 43		
<211> 477		

<212> PRT

<213> 智人(Homo sapiens)

<400> 43

```

Met Tyr Asp Ala Glu Arg Gly Trp Ser Leu Ser Phe Ala Gly Cys Gly
1           5           10           15
Phe Leu Gly Phe Tyr His Val Gly Ala Thr Arg Cys Leu Ser Glu His
          20           25           30
Ala Pro His Leu Leu Arg Asp Ala Arg Met Leu Phe Gly Ala Ser Ala
          35           40           45
Gly Ala Leu His Cys Val Gly Val Leu Ser Glu Gln Thr Leu Gln Val
          50           55           60
Leu Ser Asp Leu Val Arg Lys Ala Arg Ser Arg Asn Ile Gly Ile Phe
65           70           75           80
His Pro Ser Phe Asn Leu Ser Lys Phe Leu Arg Gln Gly Leu Cys Lys
          85           90           95
Cys Leu Pro Ala Asn Val His Gln Leu Ile Ser Gly Lys Ile Gly Ile
          100          105          110
Ser Leu Thr Arg Val Ser Asp Gly Glu Asn Val Leu Val Ser Asp Phe
          115          120          125
Arg Ser Lys Asp Glu Val Val Asp Ala Leu Val Cys Ser Cys Phe Met
          130          135          140
Pro Phe Tyr Ser Gly Leu Ile Pro Pro Ser Phe Arg Gly Val Arg Tyr
145          150          155          160
Val Asp Gly Gly Val Ser Asp Asn Val Pro Phe Ile Asp Ala Lys Thr
          165          170          175
Thr Ile Thr Val Ser Pro Phe Tyr Gly Glu Tyr Asp Ile Cys Pro Lys
          180          185          190
Val Lys Ser Thr Asn Phe Leu His Val Asp Ile Thr Lys Leu Ser Leu
          195          200          205
Arg Leu Cys Thr Gly Asn Leu Tyr Leu Leu Ser Arg Ala Phe Val Pro
          210          215          220
Pro Asp Leu Lys Val Leu Gly Glu Ile Cys Leu Arg Gly Tyr Leu Asp
225          230          235          240
Ala Phe Arg Phe Leu Glu Glu Lys Gly Ile Cys Asn Arg Pro Gln Pro
          245          250          255
Gly Leu Lys Ser Ser Ser Glu Gly Met Asp Pro Glu Val Ala Met Pro
          260          265          270
Ser Trp Ala Asn Met Ser Leu Asp Ser Ser Pro Glu Ser Ala Ala Leu
          275          280          285

```

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

<210> 44

<211> 2397

<212> DNA

<213> 智人 (Homo sapiens)

<400> 44

```

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60
aagccaugaa caucauccua gaaauccuuc ugcuuucugau caccaucauc uacuccuacu 120
uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180
uucucauuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240
gacagagcau auugguucug ugggauuuu auaagcgcgg uguggaggaa acugcagcug 300
agugccgaaa acuaggcguc acugcgcaug cguauguggu agacugcagc aacagagaag 360
agaucuaucg cucucuaaaau caggugaaga aagaaguggg ugauguaaca aucgugguga 420
auaaugcugg gacaguauau ccagccgauc uucucagcac caaggaugaa gagauuacca 480
agacauuuga gguaacauc cuaggacauu uuuggaucac aaaagcacuu cuuccaucga 540
ugauggagag aaaucauggc cacaucguca caguggcuuc agugugcggc cacgaaggga 600

```

uuccuuaccu caucccauau uguuccagca aauuugccgc uguuggcuuu cacagagguc 660
 ugacaucaga acuucaggcc uugggaaaaa cugguaucua aaccucaugu cucugcccag 720
 uuuuugugaa uacuggguuc accaaaaaau caagcacaag auuauggccu guauuggaga 780
 cagaugaagu cguaagaagu cugauagau gaauacuua caauaagaaa augauuuug 840
 uuccaucgua uaucaauauc uuucugagac uacagaagu ucuuccugaa cgcgccucag 900
 cgauuuuaaa ucguaugcag aaauuucuuu uugaagcagu gguuggccac aaaaucaaaa 960
 ugaaaugaau aaauaagcuc cagccagaga uguaugcaug auaaugauau gaauaguuuu 1020
 gaaucaaugc ugcaaaagcuu uauuuacauu uuuuucaguc cugauaaauu uaaaaacauu 1080
 gguuuggcac uagcagcagu caaacgaaca agauuaauua ccugucuucc uguuucucua 1140
 gaauauuuac guaguuuuuu auaggucugu uuuuuccuuu augccucua aaaacuucug 1200
 ugcuuacaua aacauacuua aaagguuuu uuuuagauau uuuaauuuu cauuuaaagg 1260
 uggacaaaag cuaccucccu aaaaguauu acaaagagaa cuuauuuaca cagggaaggu 1320
 uuaagacugu ucaaguagca uuccaauucg uagccaugcc acagaauauc aacaagaaca 1380
 cagaugagu gcacagcuua gagaucaagu uucagcaggc agcuuuaucu caaccuggac 1440
 auauuuuaag auucagcauu ugaaagauuu ccuagccuc uuccuuuuu auuagcccaa 1500
 aacggugcaa cucuauucug gacuuuauu cuugauucug ucuucuguau aacucugaag 1560
 uccacaaaaa guggaccuc uauauuuuccu ccuuuuuuu agucuuaaua gauacauau 1620
 gaaaggugac cgacucuauu uuaaaucuca gaauuuuag uucuagcccc augauaaccu 1680
 uuuuucuuug aauuuaugc uucauauauc cuuggucca gagauguuu gacaauuuu 1740
 ggcuaaaaaa uuaaagcuua cacaggaaaa ggaacuguac uggcuauuac auagaaaca 1800
 auggacccaa gagaagaaaa ggaagaaaga aagguuuuu gguuuuuguu uuguuuuguu 1860
 uuguuuuuug uuuuuuugag auggagucuc acucuucgc ccaggcugga gugcaguggu 1920
 augaucucag cucacugcaa gcuccaccuc ccggguucac gccauucucc ugccucagcc 1980
 uccugaguag cugggacuac aggcgcccgc caccacacc ggcuauuuu uuguauuuu 2040
 uguagagacg ggguuucacc auguuagcca agauggucuc gaucuccuga ccucgugauc 2100
 caccugccuc ggccucccaa agugcuggga uuacgggugu gagccaccgu gccagccuu 2160
 uuuuuuuuua auagaaaaaa uauccgacu ccacuacau caagacuaa cuuguuuugu 2220
 guguuuuuca cauguauuau agaauucu ugauggacu auccucuugu uuuuauuaaa 2280
 aacaaugau uuuuuuaaaa gucacaaaa caauucacua aaaauaaaua ugucuuugug 2340
 cuuuuaaaaa auuaccucuu guaguuaaua auuaaaacgu uugacuucua aacucug 2397

<210> 45

<211> 2289

<212> RNA

<213> 智人 (Homo sapiens)

<400> 45

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60
 aagccaugaa caucauccua gaaauccuuc ugcuuucugau caccauauac uacuccuacu 120
 uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180
 uuucuuuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240

gacagagcau auugguucug uggaauauua auuaggugaa gaaagaagug ggugauguaa 300
 caaucguggu gaauaaugcu gggacaguau auccagccga ucuucucagc accaaggau 360
 aagagauuac caagacauuu gaggucaaca uccuaggaca uuuuuggauc acaaaagcac 420
 uucuucauc gaugauggag agaaaucaug gccacaucgu cacaguggcu ucagugugcg 480
 gccacgaagg gauuccuac cucauccau auuguuccag caaauuugcc gcuguuggcu 540
 uucacagagg ucugacauca gaacuucagg ccuugggaaa aacugguaua aaaaccucau 600
 gucucugccc aguuuuugug aaucuggggu ucacaaaaa uccaagcaca agauuauaggc 660
 cuguauugga gacagaugaa gucguuagaa gucugauaga uggaauacuu accaauaaga 720
 aaaugauuuu uguuccaucg uauaucaaua ucuuucugag acuacagaag uuucuuucug 780
 aacgcgccuc agcgauuuua aaucguuagc agaauauuca auuugaagca gugguuggcc 840
 acaaaaauca aaugaaauga auaaaauaagc uccagccaga gauguauca ugauaaugau 900
 augaauaguu ucgaaucaau gcugcaaagc uuuaauucac auuuuuucag uccugauaua 960
 auuaaaaaa uugguuuggc acuagcagca gucaaagaa caagauuaa uaccugucuu 1020
 ccuguuucuc aagaauauuu acguaguuuu ucauaggucu guuuuuccuu ucaugccucu 1080
 uaaaaacuuc ugugcuuaca uaaacauacu uaaaagguuu ucuuaagau auuuuauuuu 1140
 uccauuuaaa gguggacaaa agcuaccucc cuaaaagua auacaaagag aacuuauuuu 1200
 cacagggaag guuuuagacu guucaaguag cauuccaauc uguagccaug ccacagaaua 1260
 ucaacaagaa cacagaaua gugcacagcu aagagaucua guuucagcag gcagcuuuau 1320
 cucaaccugg acauauuuua agauucagca uuugaaagau uuuccuagcc ucuuccuuuu 1380
 ucauuagccc aaaacggugc aacucuauc uggacuuau uacuugauuc ugucuucugu 1440
 auaacucuga aguccacaa aaguggacc ucuauuuuc cuccuuuuu auagucuuau 1500
 aagauacauu augaaaggug accgacucua uuuaaauc cagaauuuu aguucagcc 1560
 ccaugauaac cuuuuucuuu guaaauuau cuuucuuua uccuuggucc cagagauguu 1620
 uagacaauuu uaggcucaaa aaauaaagcu aacacaggaa aaggaacugu acuggcuauu 1680
 acauaagaaa caauggacc aagagaagaa aaggaagaaa gaaagguuuu uugguuuuug 1740
 uuuguuuuug uuuguuuuu uguuuuuuug agauggaguc ucacucuuuc gcccaggcug 1800
 gagucgagug guaugaucuc agcucacugc aagcuccacc ucccggguuc acgccauuc 1860
 ccugccucag ccuccugagu agcugggacu acaggcgccc gccaccacac ccggcuauuu 1920
 uuuguaauuu uuuguagaga cgggguuua ccauguuagc caauggguc ucgaucuccu 1980
 gaccucguga uccaccugcc ucggccucc aaagugcugg gauuacgggu gugagccacc 2040
 gugcccagcc uuuuuuuuu uauuagaaaa auuauuccga cuccacuac aucaagacua 2100
 aucuuguuuu guguguuuuu ccauguauu auagaauuc uugcaugga cuauccucuu 2160
 guuuuuauua aaaacaaaug auuuuuuuua aagucacaaa acaauucac uaaaaauaaa 2220
 uaugucauug ugcuuuuaaa aaauaaccuc uuguaguauu aaauuaaac guuugacuuc 2280
 uaaacucug 2289

<210> 46

<211> 2280

<212> RNA

<213> 智人 (Homo sapiens)

<400> 46

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60
 aagccaugaa caucauccua gaaauccuuc ugcuuugau caccucauc uacuccuacu 120
 uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180
 uucucauuac uggagcuggg cauggauuag gcaggcagac uacuuaugaa uuugcaaaac 240
 gacagagcau auugguucug ugggauuua auuagcgagg uguggaggaa acugcagcug 300
 agugccgaaa acuaggcguc acugcgcaug cguauguggu agacugcagc aacagagaag 360
 agaucuaucg cucucuaaaau caggugaaga aagaaguggg ugauguaaca aucgugguga 420
 auaaugcugg gacaguauau ccagccgauc uucucagcac caaggauaa gagauuacca 480
 agacauuuga gguaacauc cuaggacauu uuuggaucac aaaagcacuu cuuccaucga 540
 ugauggagag aaaucauggc cacaucguca caguggcuuc agugugcggc cacgaaggga 600
 uuuccuaccu caucccauuu uguuccagca aaauugccgc uguuggcuuu cacagagguc 660
 ugacaucaga acuucaggcc uugggaaaaa cugguauca aaccucaugu cucugcccag 720
 uuuuuugugaa uacuggguuu accaaaaaau caagcacaag guuucuuccu gaacgcgccu 780
 cagcgauuuu aaauuguaug cagaauuuu aaauugaagc agugguuggc caaaaauca 840
 aaauugaaug aaauaaauag cuccagccag agauguugc augauaaga uauaauagu 900
 uucgaaucaa ugcugcaaag cuuuuuuua cauuuuuua guccugauaa uauaaaaaac 960
 auugguuugg cacuagcagc agucaaacga acaagauua uuaccugucu uccuguuucu 1020
 caagaauuu uacguaguuu uucauagguc uguuuuuuccu uucaugccuc uaaaaaacuu 1080
 cugugcuuac auaaacauac uuaaaagguu uuuuuuuaga uuuuuuuuu uuccauuuua 1140
 agguaggaaa aagcuaccuc ccuaaaagua aaauacaaaga gaacuuuuu acacaggga 1200
 gguuuuagac uguucaagua gcauucaau cuguagccau gccacagaau aucaacaaga 1260
 acacagaau agugcacagc uaagagauc aguuucagca ggcagcuua ucuaaccug 1320
 gacauuuuu aagaucagc auuugaaaga uuuccuagc cucuuccuuu uucauagcc 1380
 caaacggug caacucuuu cuggacuuu uuacuugauu cugucuucug uauaacucug 1440
 aaguccacca aaaguggacc cucuaauuu ccuccuuuu uauagucua uaagauacu 1500
 uauaaaggu gaccgacucu auuuuaaau ucagaauuu aaguucagc ccaugauaa 1560
 ccuuuuucuu uguauuuuu gcuuucuuu auccuugguc ccagagaugu uuagacauu 1620
 uuaggcucaa aaauuaaagc uaacacagga aaaggaacug uacuggcuau uacauaaga 1680
 acauaggacc caagagaaga aaaggaagaa agaaagguuu uuugguuuuu guuuuguuuu 1740
 guuuuguuuu uuuuuuuuu gagauggagu cucacuuuu cgcccaggcu ggagugcagu 1800
 gguaugauu cagcucacug caagcuccac cuccggguu cagccauuc uccugccua 1860
 gccuccugag uagcugggac uacaggcgcc cgccaccaca cccggcuau uuuuuguauu 1920
 uuuuguagag acggggguu accauguuu ccaagauggu cucgaucucc ugaccucgug 1980
 auccaccugc cucggccucc caaagugcug ggauuacggg ugugagccac cgugcccagc 2040
 cuuuuuuuuu uuaauagaaa aaauaauccg acuccacua caucaagacu aaucuuuuu 2100
 uguguguuuu ucacauuuu uauagaauuc uuugcaugg acuauccucu uguuuuuuu 2160
 aaaaacaaau gauuuuuuu aaagucacaa aaacaauua cuaaaaaua auaugucau 2220
 gugcuuuuu aaaaauaccu cuuguaguua uaaaauaaaa cguuugacu cuaaacucug 2280

<210> 47

<211> 2398

<212> RNA

<213> 智人 (Homo sapiens)

<400> 47

```

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60
aagccaugaa caucauccua gaaauccuuc ugcuuucugau caccaucauc uacuccuacu 120
uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180
uucucauuac uggagcuggg cauggauuag gcaggcagac uacuuaugaa uuugcaaaac 240
gacagagcau auugguucug uggaauuuu auuagcgagg uguggaggaa acugcagcug 300
agugccgaaa acuaggcguc acugcgcaug cguauguggu agacugcagc aacagagaag 360
agaucuaucg cucucuaaaau caggugaaga aagaaguggg ugauguaaca aucgugguga 420
auaauugcug gacaguauau ccagccgauc uucucagcac caaggaugaa gagauuacca 480
agacauuuga gguaacauc cuaggacauu uuuggaucac aaaagcacuu cuuccaucga 540
ugauggagag aaaucauggc cacaucguca caguggcuuc agugugcggc cacgaaggga 600
uuccuuaccu caucccauau uguuccagca aaauugccgc uguuggcuuu cacagagguc 660
ugacaucaga acuucaggcc uugggaaaaa cugguaucaa aaccucaugu cucugcccag 720
uuuuugugaa uacuggguuc accaaaaaau caagcacaag auuauggccu guauuggaga 780
cagaugaagu cguaagaagu cugauagau gaauacuac caauaagaaa augauuuuug 840
uuccaucgua uaucaauauc uuucugagac uacagaaggu uucuuuccga acgcgccuca 900
gcgauuuuua aucguaugca gaauuuucaa uuugaagcag ugguuggcca caaaaucaaa 960
augaaaugaa uaaaauagcu ccagccagag auguauagca gauaaugaua ugaauaguuu 1020
cgaaucaaug cugcaaagcu uuauuuuaca uuuuuucagu ccugauaaua uaaaaacau 1080
ugguuuggca cuagcagcag ucaaacgaac aagaauaaau accugucuuc cuguuucua 1140
agaauauuuu cguaguuuuu cauaggucug uuuuuuccuu caugccucu uaaaaacuuc 1200
gugcuuacau aaacauacu aaaagguuuu cuuuuagaua uuuuuuuuuu ccauuuaaag 1260
guggacaaaa gcuaccuccc uaaaaguaaa uacaaagaga acuuuuuuac acagggaagg 1320
uuuaagacug uucaaguagc auuccaauu guagccaugc cacagaauau caacaagaac 1380
acagaauagag ugcacagcua agagaucag uuucagcagg cagcuuuauc ucaaccugga 1440
cauauuuuuu gauucagcau uugaaagauu ucccuagccu cuuccuuuuu cauagccca 1500
aaacggugca acucuaauuc ggacuuuuuu acuuugauuc gucuucugua uaacucugaa 1560
guccacaaaa aguggacccu cuauuuuuuc uccuuuuuuu uagucuuaa agauacauua 1620
ugaaagguga ccgacucua uuuuuuuuuc agauuuuuu guucuagccc caugauaacc 1680
uuuuuuuuug uauuuuuuagc uuucuuuuu ccuugguccc agagauguuu agacaauuuu 1740
aggcucaaaa auuuuagcua acacaggaaa aggaacugua cuggcuauua cauagaaac 1800
aauggaccca agagaagaaa aggaagaaa aggaaguuuu ugguuuuuuu uuuguuuuuu 1860
uuuguuuuuu guuuuuuuuga gauggagucu cacucuucg cccaggcugg agugcagugg 1920
uaugaucua gcucacugca agcuuccacu cccggguuca cgccauucuc cugccucagc 1980
cuccugagua gcugggacua caggcgcccc ccaccacacc cggcuauuuu uuuguauuuu 2040

```

uuuguagagac gggguuucac cauguuagcc aagauggucu cgaucuccug accucgugau 2100
 ccaccugccu cggccuccca aagugcuggg auuacgggug ugagccaccg ugcccagccu 2160
 uuuuuuuuuu aaugaaaaa auaauccgac ucccacuaca ucaagacuaa ucuuguuuug 2220
 uguguuuuuc acauguauua uagaauugcu uugcauggac uauccucuug uuuuuauuaa 2280
 aaacaaauga uuuuuuuaa agucacaaaa acaauucacu aaaaauaaa augucauugu 2340
 gcuuuuuuuu aaauaaccucu uguaguuaa aaauaaaacg uuugacuucu aaacucug 2398
 <210> 48
 <211> 2469
 <212> RNA
 <213> 智人 (Homo sapiens)
 <400> 48
 agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60
 aagccaugaa caucauccua gaaauccuuc ugcuuucugau caccuauuc uacuccuacu 120
 uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180
 uuucuuuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240
 gacagagcau auugguucug ugggauuuua auuagcgcgg uguggaggaa acugcagcug 300
 agugccgaaa acuaggcguc acugcgcaug cguauguggu agacugcagc aacagagaag 360
 agaucuauucg cucucuuaau caggugaaga aagaaguggg ugauguaaca aucgugguga 420
 auuauugcug gacaguauau ccagccgauc uuucucagc caaggaugaa gagauuacca 480
 agacauuuga ggucaacauc cuaggacauu uuuggaauug aaaggacauc agaaguaauu 540
 acuuggaugu auauaggau caggacacuu uuggacgaga cucugagauc aaaaaagcac 600
 uuucuccauc gaugauggag agaaaucaug gccacaucgu cacaguggcu ucagugugcg 660
 gccacgaagg gauuccuac cucauccau auuguuccag caauuuugcc gcuguuggcu 720
 uuacacagagg ucugacauca gaacuucagg ccuugggaaa aacugguauc aaaaccuacu 780
 gucucugccc aguuuuugug aaucuggggu ucacaaaaa uccaagcaca agauuauggc 840
 cuguauugga gacagaugaa gucguuagaa gucugauaga uggaauacu accaauaaga 900
 aaauaguuuu uguuccaucg uauaucaua ucuuucugag acuacagaag uuucuuucug 960
 aacgcgccuc agcgauuuua aaucguuagc agaauuuca auuugaagca gugguuggcc 1020
 acaaaaauca aaugaaauga auaaaauagc uccagccaga gauguaugca ugauaaugau 1080
 augaaauagu ucgaaucau gcugcaaagc uuuaauucac auuuuuucag uccugauau 1140
 auuuuuuaca uuuguuuugc acuagcagca gucaaacgaa caagauuaau uaccugucuu 1200
 ccuguuuuc uagaauuuu acguaguuuu ucuaaggucu guuuuuuccu ucaugccucu 1260
 uaaaaacuuc ugugcuuaca uaaacauacu uaaaagguuu ucuuuuagau auuuuuuuuu 1320
 ucauuuuuua gguggacaaa agcuaccucc cuaaaaguua auacaaagag aacuuuuuu 1380
 cacagggaag guuuuagacu guucaaguag cauuccauuc uguagccaug ccacagaaua 1440
 ucaacaagaa cacagaaua gugcacagcu aagagaucua guuucagcag gcagcuuuau 1500
 cucaaccugg acuuuuuuu agauucagca uuugaaagau uuuccuagcc ucuuccuuuu 1560
 ucauuagccc aaacgggugc aacucuuuuc uggacuuuau uacuugauuc ugucuuucugu 1620
 auaacucuga aguccaccaa aaguggaccc ucuuuuuuuc cuccuuuuu auagucuuau 1680

aagauacauu augaaaggug accgacucua uuuuaaaucu cagaauuuua aguucuaagcc 1740
 ccaugauaac cuuuuuucuuu guaaauuaug cuuucauuaa uccuuggucc cagagauguu 1800
 uagacaauuu uaggcucaaa aaauaaagcu aacacaggaa aaggaacugu acuggcuauu 1860
 acauaagaaa caauggaccc aagagaagaa aaggaagaaa gaaagguuuu uugguuuuug 1920
 uuuuguuuuug uuuuguuuuu uguuuuuuug agauggaguc ucacucuuuc gcccaggcug 1980
 gagugcagug guaugaucuc agcucacugc aagcuccacc ucccggguuc acgccauucu 2040
 ccugccucag ccuccugagu agcugggacu acaggcgccc gccaccacac ccggcuauuu 2100
 uuuuguauuu uuuuguagaga cgggguuuca ccauguuagc caagaugguc ucgaucuccu 2160
 gaccucguga uccaccugcc ucggccuccc aaagugcugg gauuacgggu gugagccacc 2220
 gugcccagcc uuuuuuuuuu uaaauagaaa aaauauccga cucccacuac aucaagacua 2280
 aucuuguuuu guguguuuuu cacauguauu auagaaugcu uuugcaugga cuauccucuu 2340
 guuuuuauua aaaacaaaug auuuuuuuua aagucacaaa aacaauucac uaaaaauaaa 2400
 uaugucauug ugcuuuaaaa aaauaaccuc uuguaguauu aaaauaaaac guuugacuuc 2460
 uaaacucug 2469

<210> 49

<211> 1715

<212> RNA

<213> 智人(Homo sapiens)

<400> 49

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60
 aagccaugaa caucauccua gaaauccuuc ugcuuugau caccuauuc uacuccuacu 120
 uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180
 uucucauuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240
 gacagagcau auugguucug ugggauuuu auaagcgcgg uguggaggaa acugcagcug 300
 agugccgaaa acuaggcguc acugcgcaug cguauuggu agacugcagc aacagagaag 360
 agaucuaucg cucucuaaa caggugaaga aagaaguggg ugauguaaca aucgugguga 420
 auaauncugg gacaguauu ccagccgauc uucucagcac caaggaugaa gagauuacca 480
 agacauuuga ggucaacauc cuaggacauu uuuggaucac aaaagcacuu cuuccaucga 540
 ugauggagag aaaucauggc cacaucguca caguggcuuc agugugcggc cacgaaggga 600
 uuuccuuaccu caucccauau uguuccagca aaauugccgc uguuggcuuu cacagagguc 660
 ugacaucaga acuuaggcc uugggaaaaa cugguaucaa aaccucaugu cucugcccag 720
 uuuuugugaa uacuggguuc accaaaaaau caagcacaag auuauggccu guauuggaga 780
 cagaugaagu cguaagaagu cugauagau gaauacuua caauaagaaa augauuuuug 840
 uuuccaucgua uaucaauauc uuucugagac uacagaagu aaguacagca cagaacaccc 900
 aaauacuaaa acaccaauag agcuuuuuuu uuugcuuuuu uuuuuuuuag acagagucuc 960
 acucugucac ccuggcugga uugcgguggu ugcaguggca ugaucuuggc ucacugcaac 1020
 cuccgccucc uggguucaag caauucucuu gccucagacc cccaaguaac ugggauuuaa 1080
 ggugugugcu gccacacuac acccagcuua uuuuuguauu uuugauaga gacagguuuc 1140
 cccauguugg ccaggcugga cucgaacucc ugaccucaag uuauccuccu gucucggccu 1200

cccaaagugc ugggauuaca gucaugagcc accaugccug gccc aaauaga gcuauuauua 1260
 uggagcaucu uucaguugug aaaaauuggca uggaaacucu ccaucccugg ggagaacagu 1320
 uauuuccucu guuauuuucc uaccagucu auaaaaagag agugauucau uuucucuacc 1380
 aaauacuacug ucucugccca aacuuugcug aagacuauuc uaacuaaagg aaacacagu 1440
 uaaaaagaau gcaauauagu gaaguaguua auauaaaaga cuccauuuuu aaaagucugc 1500
 uggaaguuug guugggauug cacugaauuc auagagcaau ugaggagauu ugacauauca 1560
 acaauauuga guuuucuaau ccaagaacau aaauucuaau uuuaaaauuc ucuuacaaau 1620
 cuuuuuuauuc uuaaaauugua uuuguaguuu uuugguguuu aagucuugca cauauuuugu 1680
 cagauuuauu ccaaaguauu ucacggguuc uuuuu 1715

<210> 50

<211> 2290

<212> RNA

<213> 智人 (Homo sapiens)

<400> 50

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60
 aagccaugaa caucauccua gaaauccuuc ugcuuugau caccauac uacuccuacu 120
 uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180
 uuucuuuac uggagcuggg cauggauag gcaggcagac uacuuugaa uuugcaaac 240
 gacagagcau auugguucug ugggauuua auuagguu gaaagaagug ggugauguaa 300
 caaucguggu gaauaaugcu gggacaguau auccagccga ucuucucagc accaaggaug 360
 aagagauuac caagacauuu gaggucaaca uccuaggaca uuuuuggauc aaaaagcac 420
 uuuuuccauc gaugauggag agaaaucaug gccacaucgu cacaguggcu ucagugugcg 480
 gccacgaagg gauuccuac cucauccau auuguuccag caauuuugc gcuguuggcu 540
 uuacagagg ucugacauca gaacuucagg ccuugggaaa aacugguauc aaaaccuau 600
 gucucugccc aguuuuugug aaucugggu ucacaaaaa uccaagcaca agauuauugc 660
 cuguauugga gacagaugaa gucguuagaa gucugauaga uggaauacu accaauaaga 720
 aaauuuuuu uguuccaucg uauaucaua ucuuucugag acuacagaag guuuuuccu 780
 gaacgcgccu cagcgauuuu aaucguuag cagaauuuc aaauugaagc agugguugc 840
 caaaaauca aaauuuuug aaauuuuag cuccagccag agauuuuagc augauuuga 900
 uauuuuuu uucgaaucua ugcugcaaag cuuuuuuua cauuuuuuu guccugauaa 960
 uauuuuuu auugguuug cacuagcagc agucuaacga acaaguuua uuaccugucu 1020
 uccuguuuu caagaauuu uacguaguuu uucauagguc uguuuuuuccu uucaugccuc 1080
 uuuuuuuu cuugcuuac auuuuuuu uuuuuuuu uuuuuuuu uuuuuuuu 1140
 uuuuuuuu agguuagaa aagcuaccuc cuuuuuuu aaauuuuu gaacuuuuu 1200
 acacaggga gguuuuuagc uguucaaaga gcauuucaa cuuagccau gccacagaau 1260
 aucaacaaga acacagaau agugcacagc uaagagauc aguuucagc ggcagcuuu 1320
 ucucaaccug gacauuuuu aagauucagc auuuuaga uuuccuagc cuuuccuuu 1380
 uucauuagc caaacggug caacuuuuu cuggacuuuu uuacuugau cuuucucug 1440
 uauaacucug aaguccacca aaaguggacc cuuauuuuu cuuuuuuu uauagucuua 1500

uaagauacau uaugaaaggu gaccgacucu auuuuaaauc ucagaauuuu aaguucuagc 1560
 cccaugauaa ccuuuuucuu uguaaauuuu gcuuucuuu auccuugguc ccagagaugu 1620
 uuagacaauu uuaggcucaa aaauuaaagc uaacacagga aaaggaacug uacuggcuau 1680
 uacauaagaa acaauggacc caagagaaga aaaggaagaa agaaagguu uuugguuuu 1740
 guuuuguuuu guuuuguuuu uuguuuuuu gagauggagu cucacucuuu cgcccaggcu 1800
 ggagugcagu gguaugaucu cagcucacug caagcuccac cucccggguu cagccauuc 1860
 uccugccuca gccuccugag uagcugggac uacaggcgcc cgccaccaca cccggcuau 1920
 uuuuuuguauu uuuuguagag acgggguuuc accauguuag ccaagauggu cucgaucucc 1980
 ugaccucgug auccaccugc cucggccucc caaagugcug ggauuacggg ugugagccac 2040
 cgugcccagc cuuuuuuuuu uuaauagaaa aaauaaucg acucccacua caucaagacu 2100
 aaucuuuguu uguguguuuu ucacauguau uauagaauuc uuugcaugg acuaucucuc 2160
 uguuuuuauu aaaaacaaau gauuuuuuuu aaagucacaa aaacaauuca cuaaaaaua 2220
 auaugucuuu gugcuuuuuu aaaauaaccu cuuguaguua uaaaauaaaa cguuugacuu 2280
 cuaaacucug 2290

<210> 51

<211> 2470

<212> RNA

<213> 智人 (Homo sapiens)

<400> 51

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60
 aagccaugaa caucauccua gaaauccuuc ugcuuucgau caccauauac uacuccuacu 120
 uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180
 uucucauuac uggagcuggg cauggaauag gcaggcagac uacuuauгаа uuugcaaaac 240
 gacagagcau auugguucug ugggauuuu auaagcgcgg uguggaggaa acugcagcug 300
 agugccgaaa acuaggcguc acugcgcaug cguauguggu agacugcagc aacagagaag 360
 agaucuaucg cucucuaaa caggugaaga aagaaguggg ugauguaaca aucgugguga 420
 auaaagcugg gacaguauu ccagccgauc uucucagcac caaggaugaa gagauuacca 480
 agacauuuga ggucaacau cuaggacauu uuuggaauug aaaggacauc agaaguaau 540
 acuuggaugu auauaggauc gaggacacuu uuggacgaga cucugagauc aaaaagcac 600
 uucuucauc gaugauggag agaaaaucaug gccacaucgu cacaguggcu ucagugugcg 660
 gccacgaagg gauuccuac cucauccau auuguuccag caauuuugcc gcuguuggcu 720
 uucacagagg ucugacauca gaacuucagg ccuugggaaa aacugguauc aaaaccuau 780
 gucucugccc aguuuuugug aaucugggg ucacaaaaa uccaagcaca agauuauugc 840
 cuguauuuga gacagaugaa gucguuagaa gucugauaga uggaauacuu accaauaaga 900
 aaaugauuuu uguuccauc uauaucaau ucuuucugag acuacagaag guuucuuucc 960
 gaacgcgccu cagcgauuu aaucguuag cagaauuuu auuuugaagc agugguuggc 1020
 caaaaauca aaauaaaug aaauaaaag cuccagccag agauguuagc augauaaga 1080
 uaugaauagu uucgaaucaa ugcugcaaag cuuuuuuca cauuuuuca guccugauaa 1140
 uauuuuuuac auugguuugg cacuagcagc agucaaacga acaagauuaa uuaccugucu 1200

uccuguuuucu caagaauauu uacguaguuu uucauagguc uguuuuuccu uucaugccuc 1260
 uaaaaaacuu cugugcuuac auaaacauac uuaaaagguu uucuuuaaga uauuuuauuu 1320
 uuccauuuuaa aggguggacaa aagcuaccuc ccuaaaagua aaauaaaaga gaacuuauuu 1380
 acacagggaa gguuuuagac uguucaagua gcuuuccaau cuguagccau gccacagaau 1440
 aucaacaaga acacagaauug agugcacagc uaagagauca aguuucagca ggcagcuuuu 1500
 ucucaaccug gacauauuuu aagauucagc auuugaaaga uuucccuagc cucuuccuuu 1560
 uucauuagcc caaaacggug caacucuauu cuggacuuaa uuacuugauu cugucuucug 1620
 uauaacucug aaguccacca aaaguggacc cucuauauuu ccucccuuuu uauagucuua 1680
 uaagauacau uaugaaaggu gaccgacucu auuuuaaauc ucagaauuuu aaguucuaagc 1740
 cccaugauaa cccuuuuucu uguauuuuuu gcuuucuuu auccuugguc ccagagaugu 1800
 uuagacaauu uuaggcucaa aaauuaaagc uaacacagga aaaggaacug uacuggcuau 1860
 uacauaagaa acaauaggacc caagagaaga aaaggaagaa agaaagguu uuugguuuuu 1920
 guuuuguuuu guuuuguuuu uuuuuuuuuu gagauggagu cucacucuuu cgcccaggcu 1980
 ggagugcagu gguaugaucu cagcucacug caagcuccac cucccggguu cagccauuc 2040
 uccugccuca gccuccugag uagcugggac uacaggcgcc cgccaccaca cccggcuau 2100
 uuuuuguaau uuuuuguaag acgggguuuc accauguuag ccaagauggu cucgaucucc 2160
 ugaccucgug auccaccugc cucggccucc caaagucug ggauuacggg ugugagccac 2220
 cgugcccagc cccccuuuuu uuaauagaaa aaauaaucg acucccacua caucaagacu 2280
 aaucuuuuu uguguuuuu ucacauuuu uauagaauug uuugcaugg acuaucucu 2340
 uguuuuuauu aaaaacaaa gauuuuuuuu aaagucacaa aaacaauuca cuaaaaaua 2400
 auaugucau gugcuuuuuu aaauuaaccu cuuguaguua uaaaauaaaa cguugacuu 2460
 cuaaacucug 2470

<210> 52

<211> 1714

<212> RNA

<213> 智人(Homo sapiens)

<400> 52

agacaguacc uccucccuag gacuacacaa ggacugaacc agaaggaaga ggacagagca 60
 aagccaugaa caucauccua gaaauccuuc ugcuuucugau caccauauuc uacuccuacu 120
 uggagucguu ggugaaguuu uucauuccuc agaggagaaa aucuguggcu ggggagauug 180
 uucucauuac uggagcuggg cauggaauag gcaggcagac uacuuaugaa uuugcaaaac 240
 gacagagcau auugguucug ugggauuuu auuagcgagg uguggaggaa acugcagcug 300
 agugccgaaa acuaggcguc acugcgcaug cguauguggu agacugcagc aacagagaag 360
 agaucuaucg cucucuaaa caggugaaga aagaaguggg ugauguaaca aucgugguga 420
 auuauugcugg gacaguauuu ccagccgauc uucucagcac caaggaugaa gagauuacca 480
 agacauuuuga ggucaacauc cuaggacauu uuuggaucac aaaagcacuu cuuccaucga 540
 ugauggagag aaaucauggc cacauuguca caguggcuuc agugugcggc cacgaaggga 600
 uuuccuuaccu caucccauuu uguuccagca aaauugccgc uguuggcuuu cacagagguc 660
 ugacaucaga acuuaggcc uuuggaaaaa cugguaucaa aaccucaugu cucugcccag 720

uuuuugugaa uacuggguuc accaaaaaau ccaagcacaag auuauaggccu guauuggaga 780
 cagaugaagu cguaagaagu cugauagaug gaauacuua ccauaagaaa augauuuuug 840
 uucaaucgua uaucaauauc uuucugagac uacagaagua aguacagcac agaacaccca 900
 aaauacuaaaa caccaauaga gcuuuuuuuu uugcuuuuuu uuuuuuuaga cagagucuca 960
 cucugucacc cuggcuggau ugcggugguu gcaguggcau gaucuuggcu cacugcaacc 1020
 uccgccuccu ggguucaagc aaauucuaug ccucagaccc ccaaguaacu gggauuauag 1080
 gugugugcug ccacacuaca cccagcuauu uuuuguauuu uuugauagag acagguuucc 1140
 ccauguuggc caggcuggac ucgaacuccu gaccucaagu uauccuccug ucucggccuc 1200
 ccaaagugcu gggauuacag ucaugagcca ccaugccugg cccaauagag cuuuuuuuau 1260
 ggagcaucuu ucaguuguga aaauuggcau ggaaacucuc caucccuggg gagaacagu 1320
 auuuccucug uuauuuuccu acccagucua uaaaaagaga gugauucau uucucuacca 1380
 aaucuaucug cucugcccaa acuuugcuga agacuauucu aacuaaagga aacacaguuu 1440
 aaaaagaau ccauuauagug aaguaguuaa uaauaaagac uccauuuuuu aaagucugcu 1500
 ggaaguuggg uugggauugc acugaauca uagagcauu ggggaguau gacauaucaa 1560
 cauuauugag uuuuucuauc caagaacaua auaucauuu uuaaaaucuu cuucaaauc 1620
 uuuuuuauuu uaaaauuguau uuuguaguuu uugguguuuu agucuugcac auuuuuuguc 1680
 agauuuauuc caaaguauuu cacggguucu uuuu 1714

<210> 53

<211> 2397

<212> DNA

<213> 智人 (Homo sapiens)

<400> 53

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60
 aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120
 tggagtcgtt ggtgaagttt ttcatcctc agaggagaaa atctgtggct ggggagattg 180
 ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240
 gacagagcat attggttctg tgggatatta ataagcgcgg tgtggaggaa actgcagctg 300
 agtgccgaaa actaggcgtc actgcgcatg cgtatgtggt agactgcagc aacagagaag 360
 agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420
 ataatgctgg gacagtatat ccagccgatc ttctcagcac caaggatgaa gagattacca 480
 agacatttga ggtcaacatc ctaggacatt tttggatcac aaaagcactt cttccatcga 540
 tgatggagag aaatcatggc cacatcgtca cagtggcttc agtgtgcggc cacgaaggga 600
 ttccctacct catcccatat tgttccagca aatttgccgc tgttggtttt cacagaggtc 660
 tgacatcaga acttcaggcc ttgggaaaaa ctggtatcaa aacctcatgt ctctgcccag 720
 tttttgtgaa tactgggttc accaaaaatc caagcacaag attatggcct gtattggaga 780
 cagatgaagt cgtaagaagt ctgatagatg gaatacttac caataagaaa atgatttttg 840
 ttccatcgta tatcaatata tttctgagac tacagaagtt tcttctgaa cgcgcctcag 900
 cgatttttaa tcgtatgcag aatattcaat ttgaagcagt ggttgccac aaaatcaaaa 960
 tgaaatgaat aaataagctc cagccagaga tgtatgcatg ataatgatat gaatagtttc 1020

gaatcaatgc tgcaaagctt tatttcacat tttttcagtc ctgataatat taaaaacatt 1080
 ggtttggcac tagcagcagt caaacgaaca agattaatta cctgtcttcc tgttttctcaa 1140
 gaatatattac gtagtttttc ataggtctgt ttttcctttc atgcctctta aaaacttctg 1200
 tgcttacata aacatactta aaaggttttc ttttaagatat tttatttttc catttaaagg 1260
 tggacaaaag ctacctccct aaaagtaaata acaaagagaa cttattttaca cagggaaggt 1320
 ttaagactgt tcaagtagca ttccaatctg tagccatgcc acagaatatc aacaagaaca 1380
 cagaatgagt gcacagctaa gagatcaagt ttcagcaggc agctttatct caacctggac 1440
 atatttttaag attcagcatt tgaaagattt ccctagcctc ttcctttttc attagcccaa 1500
 aacggtgcaa ctctattctg gactttatta cttgattctg tcttctgtat aactctgaag 1560
 tccacaaaaa gtggaccctc tatatttctc ccctttttat agtcttataa gatacattat 1620
 gaaaggtgac cgactctatt ttaaattctc gaatttttaag ttctagcccc atgataacct 1680
 ttttctttgt aatttatgct ttcatatata cttgggtccca gagatgttta gacaatttta 1740
 ggctcaaaaa ttaaagctaa cacaggaaaa ggaactgtac tggctattac ataagaaaca 1800
 atggacccaa gagaagaaaa ggaagaaaga aaggtttttt ggtttttggt ttgttttggt 1860
 ttgttttttg tttttttgag atggagtctc actctttcgc ccaggctgga gtgcagtggg 1920
 atgatctcag ctactgcaa gctccacctc ccgggttcac gccattctcc tgcctcagcc 1980
 tcctgagtag ctgggactac aggcgcccg caccacaccc ggctaatttt ttgtattttt 2040
 tgtagagacg gggtttcacc atgttagcca agatgggtct gatctcctga cctcgtgatc 2100
 cacctgcctc ggctcccaa agtgctggga ttacgggtgt gagccaccgt gccagcctt 2160
 ttttttttta atagaaaaaa taatccgact ccactacat caagactaat cttgttttgt 2220
 gtgtttttca catgtattat agaattgctt tgcatggact atcctcttgt ttttattaaa 2280
 aacaaatgat ttttttaaaa gtcacaaaaa caattcacta aaaataaata tgtcattgtg 2340
 ctttaaaaaa ataacctctt gtagttataa aataaaacgt ttgacttcta aactctg 2397

<210> 54

<211> 2289

<212> DNA

<213> 智人(Homo sapiens)

<400> 54

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60
 aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120
 tggagtcgtt ggtgaagttt ttcatctctc agaggagaaa atctgtggct ggggagattg 180
 ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa ttgcaaaac 240
 gacagagcat attggttctg tgggatatta ataaggtgaa gaaagaagtg ggtgatgtaa 300
 caatcgtggg gaataatgct gggacagtat atccagccga tcttctcagc accaaggatg 360
 aagagattac caagacattt gaggtcaaca tcctaggaca tttttggatc aaaaagcac 420
 ttcttccatc gatgatggag agaaatcatg gccacatcgt cacagtggct tcagtgtgcg 480
 gccacgaagg gattccttac ctcatcccat attgttccag caaatttgcc gctgttggct 540
 ttcacagagg tctgacatca gaacttcagg ctttgggaaa aactggtatc aaaacctcat 600
 gtctctgccc agtttttgtg aatactgggt tcaccaaaaa tccaagcaca agattatggc 660

ctgtattgga gacagatgaa gtcgtaagaa gtctgataga tggaatactt accaataaga 720
 aaatgatttt tgttccatcg tatatcaata tctttctgag actacagaag tttcttcctg 780
 aacgcgcctc agcgatttta aatcgtatgc agaattattca atttgaagca gtggttggcc 840
 acaaaatcaa aatgaaatga ataaataagc tccagccaga gatgtatgca tgataatgat 900
 atgaatagtt tcgaatcaat gctgcaaagc tttatttcac attttttcag tcctgataat 960
 attaaaaaca ttggtttggc actagcagca gtcaaacgaa caagattaat tacctgtctt 1020
 cctgtttctc aagaatattt acgtagtttt tcataggtct gtttttcctt tcatgcctct 1080
 taaaaacttc tgtgcttaca taaacatact taaaaggttt tctttaagat attttatttt 1140
 tccattttaa ggtggacaaa agctacctcc ctaaaagtaa atacaaagag aacttattta 1200
 cacagggaag gtttaagact gttcaagtag cattccaatc tgtagccatg ccacagaata 1260
 tcaacaagaa cacagaatga gtgcacagct aagagatcaa gtttcagcag gcagctttat 1320
 ctcaacctgg acatatttta agattcagca tttgaaagat ttccctagcc tcttcctttt 1380
 tcattagccc aaaacgggtgc aactctatc tggactttat tacttgattc tgtcttctgt 1440
 ataactctga agtccaccaa aagtggacct tctatatttc ctcccttttt atagtcttat 1500
 aagatacatt atgaaagggt accgactcta ttttaaactc cagaatttta agttctagcc 1560
 ccatgataac ctttttcttt gtaatttatg ctttcatata tccttgggtc cagagatgtt 1620
 tagacaattt taggctcaaa aattaaagct aacacaggaa aaggaactgt actggctatt 1680
 acataagaaa caatggacct aagagaagaa aaggaagaaa gaaaggtttt ttggtttttg 1740
 ttttgttttg ttttgttttt tgtttttttg agatggagtc tctctcttc gccaggtg 1800
 gagtgcagtg gtatgatctc agctcactgc aagctccacc tcccgggttc acgccattct 1860
 cctgcctcag cctcctgagt agctgggact acaggcgccc gccaccacac ccggttaatt 1920
 ttttgatatt tttgtagaga cggggtttca ccatgttagc caagatggtc tcgatctcct 1980
 gacctcgtga tccacctgcc tcggcctccc aaagtgtctg gattacgggt gtgagccacc 2040
 gtgcccagcc tttttttttt taatagaaaa aataatccga ctcccactac atcaagacta 2100
 atcttgtttt gtgtgttttt cacatgtatt atagaatgct tttgcatgga ctatcctctt 2160
 gtttttatta aaaacaaatg atttttttaa aagtcacaaa aacaattcac taaaaataaa 2220
 tatgtcattg tgctttaaaa aaataacctc ttgtagtatt aaaataaaac gtttgacttc 2280
 taaactctg 2289

<210> 55

<211> 2280

<212> DNA

<213> 智人 (Homo sapiens)

<400> 55

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60
 aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120
 tggagtcgtt ggtgaagttt ttcatcctc agaggagaaa atctgtggct ggggagattg 180
 ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaac 240
 gacagagcat attggttctg tgggatatta ataagcgcg tgtggaggaa actgcagctg 300
 agtgccgaaa actaggcgtc actgcgcatg cgtatgtggt agactgcagc aacagagaag 360

agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420
 ataatgctgg gacagtatat ccagccgatac ttctcagcac caaggatgaa gagattacca 480
 agacatttga ggtcaacatc ctaggacatt tttggatcac aaaagcactt cttccatcga 540
 tgatggagag aaatcatggc cacatcgtca cagtggcttc agtgtgcggc cacgaaggga 600
 ttccttacct catcccatat tgttccagca aatttgccgc tgttggcttt cacagaggtc 660
 tgacatcaga acttcaggcc ttgggaaaaa ctggtatcaa aacctcatgt ctctgcccag 720
 tttttgtgaa tactgggttc accaaaaatc caagcacaag gtttcttcct gaacgcgcct 780
 cagcgatttt aaatcgtatg cagaatattc aatttgaagc agtggttggc cacaaaatca 840
 aaatgaaatg aataaataag ctccagccag agatgtatgc atgataatga tatgaatagt 900
 ttcgaatcaa tgctgcaaag ctttatttca cattttttca gtccctgataa tattaaaaac 960
 attggtttgg cactagcagc agtcaaacga acaagattaa ttacctgtct tctgttttct 1020
 caagaatatt tacgtagttt ttcataggtc tgtttttcct ttcattgcctc ttaaaaactt 1080
 ctgtgcttac ataaacatac ttaaaagggt ttctttaaga tatttttattt ttccatttaa 1140
 aggtggacaa aagctacctc cctaaaagta aatacaaaga gaacttattt acacaggga 1200
 ggtttaagac tgttcaagta gcattccaat ctgtagccat gccacagaat atcaacaaga 1260
 acacagaatg agtgcacagc taagagatca agtttcagca ggcagcttta tctcaacctg 1320
 gacatatattt aagattcagc atttgaaaga tttccctagc ctcttccttt ttcattagcc 1380
 caaaacggtg caactctatt ctggacttta ttacttgatt ctgtcttctg tataactctg 1440
 aagtccacca aaagtggacc ctctatatatt cctccctttt tatagtctta taagatacat 1500
 tatgaaagggt gaccgactct attttaaatc tcagaatttt aagttctagc cccatgataa 1560
 cctttttctt tgtaatttat gctttcatat atccttggtc ccagagatgt ttagacaatt 1620
 ttaggctcaa aaattaaagc taacacagga aaaggaactg tactggctat tacataagaa 1680
 acaatggacc caagagaaga aaaggaagaa agaaagggtt tttggttttt gttttgtttt 1740
 gttttgtttt ttgttttttt gagatggagt ctactcttt cgcccaggct ggagtgcagt 1800
 ggtatgatct cagctcactg caagctccac ctcccgggtt cacgccattc tcctgcctca 1860
 gcctcctgag tagctgggac tacaggcgcc cgccaccaca cccggctaatt tttttgtatt 1920
 tttttagtag acgggggttc accatgttag ccaagatggt ctcgatctcc tgacctcgtg 1980
 atccacctgc ctcggcctcc caaagtgtg ggattacggg tgtgagccac cgtgcccagc 2040
 cttttttttt ttaatagaaa aaataatccg actcccacta catcaagact aatcttgttt 2100
 tgtgtgtttt tcacatgtat tatagaatgc ttttgcattg actatcctct tgtttttatt 2160
 aaaaacaaat gattttttta aaagtcacaa aaacaattca ctaaaaataa atatgtcatt 2220
 gtgcttttaa aaaataacct cttgtagtta taaaataaaa cgtttgactt ctaaactctg 2280
 <210> 56
 <211> 2398
 <212> DNA
 <213> 智人(Homo sapiens)
 <400> 56
 agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60
 aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120

```

tggagtcggtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180
ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240
gacagagcat attggttctg tgggatatta ataagcgcgg tgtggaggaa actgcagctg 300
agtgccgaaa actaggcgtc actgcgcatg cgtatgtggt agactgcagc aacagagaag 360
agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420
ataatgctgg gacagtatat ccagccgata ttctcagcac caaggatgaa gagattacca 480
agacatttga ggtcaacatc ctaggacatt tttggatcac aaaagcactt cttccatcga 540
tgatggagag aaatcatggc cacatcgtca cagtggcttc agtgtgcggc cacgaaggga 600
ttccttacct catcccatat tgttccagca aatttgccgc tgttggcttt cacagaggtc 660
tgacatcaga acttcaggcc ttgggaaaaa ctggtatcaa aacctcatgt ctctgcccag 720
tttttgtgaa tactgggttc accaaaaatc caagcacaag attatggcct gtattggaga 780
cagatgaagt cgtaagaagt ctgatagatg gaatacttac caataagaaa atgatttttg 840
ttccatcgta tatcaatata tttctgagac tacagaaggt ttcttctga acgcgcctca 900
gcgattttta atcgtatgca gaatattcaa tttgaagcag tggttggcca caaaatcaaa 960
atgaaatgaa taaataagct ccagccagag atgtatgcat gataatgata tgaatagttt 1020
cgaatcaatg ctgcaaagct ttatttcaca ttttttcagt cctgataata ttaaaaacat 1080
tggtttggca ctagcagcag tcaaacgaac aagattaatt acctgtcttc ctgtttctca 1140
agaatattta cgtagttttt cataggtctg tttttccttt catgcctctt aaaaacttct 1200
gtgcttacat aaacatactt aaaaggtttt ctttaagata ttttattttt ccattttaaag 1260
gtggacaaaa gctacctccc taaaagtaaa tacaaagaga acttatttac acagggaagg 1320
tttaagactg ttcaagtagc attccaatct gtagccatgc cacagaatat caacaagaac 1380
acagaatgag tgcacagcta agagatcaag tttcagcagg cagctttatc tcaacctgga 1440
cataatttta gattcagcat ttgaaagatt tccctagcct cttccttttt cattagccca 1500
aaacgggtgca actctattct ggactttatt acttgattct gtcttctgta taactctgaa 1560
gtccacccaaa agtggaccct ctatatctcc tcccttttta tagtcttata agatacatta 1620
tgaaagggtga ccgactctat tttaaatctc agaattttta gttctagccc catgataacc 1680
tttttctttg taatttatgc tttcatatat ctttggtccc agagatgttt agacaatttt 1740
aggctcaaaa attaaagcta acacaggaaa aggaactgta ctggctatta cataagaaac 1800
aatggaccca agagaagaaa aggaagaaaag aaaggttttt tggtttttgt tttgttttgt 1860
tttgtttttt gtttttttga gatggagtct cactctttcg cccaggctgg agtgcagtgg 1920
tatgatctca gctcactgca agtccacct cccgggttca cgccattctc ctgcctcagc 1980
ctcctgagta gctgggacta caggcgcgccc ccaccacacc cggctaattt tttgtatttt 2040
ttgtagagac ggggtttcac catgttagcc aagatggtct cgatctcctg acctcgtgat 2100
ccacctgcct cggcctccca aagtgetggg attacgggtg tgagccaccg tgcccagcct 2160
tttttttttt aatagaaaaa ataatccgac tcccactaca tcaagactaa tcttgttttg 2220
tgtgtttttc acatgtatta tagaatgctt ttgcatggac tatcctcttg tttttattaa 2280
aaacaaatga tttttttaaa agtcacaaaa acaattcact aaaaataaat atgtcattgt 2340
gcttttaaaaa aataacctct ttagattata aaataaaaac tttgacttct aaactctg 2398

```

<210> 57

<211> 2469

<212> DNA

<213> 智人(Homo sapiens)

<400> 57

```

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60
aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120
tggagtcggt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180
ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240
gacagagcat attggttctg tgggatatta ataagcgagg tgtggaggaa actgcagctg 300
agtgccgaaa actaggcgct actgcgcatg cgtatgtggt agactgcagc aacagagaag 360
agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420
ataatgctgg gacagtatat ccagccgatc ttctcagcac caaggatgaa gagattacca 480
agacatttga ggtcaacatc ctaggacatt tttggaatgg aaaggacatc agaagtaatt 540
acttggatgt atataggatc gaggacactt ttggacgaga ctctgagatc aaaaagcac 600
ttcttccatc gatgatggag agaaatcatg gccacatcgt cacagtggct tcagtgtgcg 660
gccacgaagg gattccttac ctcatcccat attgttccag caaatttgcc gctgttggct 720
ttcacagagg tctgacatca gaacttcagg ccttgggaaa aactggtatc aaaacctcat 780
gtctctgccc agtttttgtg aatactgggt tcaccaaaaa tccaagcaca agattatggc 840
ctgtatttga gacagatgaa gtcgtaagaa gtctgataga tggaatactt accaataaga 900
aaatgatttt tgttccatcg tatatcaata tctttctgag actacagaag tttcttctg 960
aacgcgcctc agcgatttta aatcgtatgc agaattatca atttgaagca gtggttggcc 1020
acaaaatcaa aatgaaatga ataaataagc tccagccaga gatgtatgca tgataatgat 1080
atgaatagtt tcgaatcaat gctgcaaagc tttatttcac attttttcag tcctgataat 1140
attaaaaaca ttggtttggc actagcagca gtcaaagaa caagattaat tacctgtctt 1200
cctgtttctc aagaatattt acgtagtttt tcataggtct gtttttctt tcatgcctct 1260
taaaaacttc tgtgcttaca taaacatact taaaaggttt tctttaagat attttatttt 1320
tccattttaa ggtggacaaa agctacctcc ctaaaagtaa atacaaagag aacttattta 1380
cacaggaag gtttaagact gttcaagtag cattccaatc tgtagccatg ccacagaata 1440
tcaacaagaa cacagaatga gtgcacagct aagagatcaa gtttcagcag gcagctttat 1500
ctcaacctgg acatatttta agattcagca tttgaaagat ttccctagcc tcttcctttt 1560
tcattagccc aaaacggtgc aactctatc tggactttat tacttgattc tgtcttctgt 1620
ataactctga agtccaccaa aagtggacce tctatatttc ctcccttttt atagtcttat 1680
aagatacatt atgaaagggt accgactcta ttttaaactc cagaatttta agttctagcc 1740
ccatgataac ctttttcttt gtaatttatg ctttcatata tccttgggtc cagagatgtt 1800
tagacaattt taggctcaaa aattaaagct aacacaggaa aaggaactgt actggctatt 1860
acataagaaa caatggaccc aagagaagaa aaggaagaaa gaaaggtttt ttggtttttg 1920
ttttgttttg ttttgttttt tgtttttttg agatggagtc tcactcttcc gccaggtg 1980
gagtgcagtg gtatgatctc agctcactgc aagctccacc tcccgggttc acgccattct 2040
cctgcctcag cctcctgagt agctgggact acaggcgccc gccaccacac ccggttaatt 2100

```

ttttgtatatt tttgtagaga cgggggtttca ccatgttagc caagatggtc tcgatctcct 2160
 gacctcgtga tccacctgcc tcggcctccc aaagtgcctg gattacgggt gtgagccacc 2220
 gtgcccagcc tttttttttt taatagaaaa aataatccga ctcccactac atcaagacta 2280
 atcttgtttt gtgtgttttt cacatgtatt atagaatgct tttgcatgga ctatcctctt 2340
 gtttttatta aaaacaaatg atttttttta aagtcacaaa aacaattcac taaaaataaa 2400
 tatgtcattg tgctttaaaa aaataacctc ttgtagttat aaaataaaac gtttgacttc 2460
 taaactctg 2469

<210> 58

<211> 1715

<212> DNA

<213> 智人 (Homo sapiens)

<400> 58

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60
 aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120
 tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180
 ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240
 gacagagcat attggttctg tgggatatta ataagcgcgg tgtggaggaa actgcagctg 300
 agtgccgaaa actaggcgtc actgcgcatg cgtatgtggt agactgcagc aacagagaag 360
 agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420
 ataatgctgg gacagtatat ccagccgatc ttctcagcac caaggatgaa gagattacca 480
 agacatttga ggtcaacatc ctaggacatt tttggatcac aaaagcactt cttccatcga 540
 tgatggagag aaatcatggc cacatcgtca cagtggcttc agtgtgcggc cacgaaggga 600
 ttccctacct catcccatat tgttccagca aatttgccgc tgttggcttt cacagaggtc 660
 tgacatcaga acttcaggcc ttgggaaaaa ctggtatcaa aacctcatgt ctctgcccag 720
 tttttgtgaa tactgggttc accaaaaatc caagcacaag attatggcct gtattggaga 780
 cagatgaagt cgtaagaagt ctgatagatg gaatacttac caataagaaa atgatttttg 840
 ttccatcgta tatcaatatc tttctgagac tacagaagtt aagtacagca cagaacaccc 900
 aaataactaaa acaccaatag agcttttttt tttgcttttt ttttttttag acagagtctc 960
 actctgtcac cctggctgga ttgcggtggg tgcagtggca tgatcttggc tcaactgcaac 1020
 ctccgcctcc tgggttcaag caattctcat gcctcagacc cccaagtaac tgggattata 1080
 ggtgtgtgct gccacactac acccagctaa ttttgtatt ttttgataga gacaggtttc 1140
 cccatgttgg ccaggctgga ctcgaaactc tgacctcaag ttatcctcct gtctcggcct 1200
 cccaaagtgc tgggattaca gtcatgagcc accatgcctg gcccaataga gctattatta 1260
 tggagcatct ttcagttgtg aaaattggca tggaaactct ccatccctgg ggagaacagt 1320
 tatttcctct gtatttttcc taccagctct ataaaaagag agtgattcat tttctctacc 1380
 aaatctactg tctctgcccc aacttttctg aagactatct taactaaagg aaacacagtt 1440
 taaaaagaat gcaatatagt gaagtagtta ataataaaga ctccattttt aaaagtctgc 1500
 tggaagtttg gttgggattg cactgaatct atagagcaat tggggagtat tgacatatca 1560
 acaatattga gtttttcta ccaagaacat aatatctatt tttaaaatct tcttcaaaat 1620

ctttaaatct ttaaattgta tttttagt ttgtgtgtt aagtcttgca catattttgt 1680
cagatttatt ccaaagtatt tcacgggttc ttttt 1715

<210> 59

<211> 2290

<212> DNA

<213> 智人 (Homo sapiens)

<400> 59

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60
aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120
tgtagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180
ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240
gacagagcat attggttctg tgggatatta ataagtgaa gaaagaagtg ggtgatgtaa 300
caatcgtggg gaataatgct gggacagtat atccagccga tcttctcagc accaaggatg 360
aagagattac caagacattt gaggtcaaca tcctaggaca tttttggatc aaaaagcac 420
ttcttccatc gatgatggag agaaatcatg gccacatcgt cacagtggct tcagtgtgcg 480
gccacgaagg gattccttac ctcatcccat attgttccag caaatttgcc gctgttggct 540
ttcacagagg tctgacatca gaacttcagg ccttgggaaa aactggtatc aaaacctcat 600
gtctctgccc agtttttgtg aatactgggt tcaccaaaaa tccaagcaca agattatggc 660
ctgtattgga gacagatgaa gtcgtaagaa gtctgataga tggaatactt accaataaga 720
aaatgatttt tgttccatcg tatatcaata tctttctgag actacagaag gtttcttctt 780
gaacgcgcct cagcgatttt aaatcgtatg cagaatatc aatttgaagc agtggttggc 840
cacaaaatca aaatgaaatg aataaataag ctccagccag agatgtatgc atgataatga 900
tatgaatagt ttcgaatcaa tgctgcaaag ctttatttca ctttttttca gtcctgataa 960
tattaaaaac attggtttgg cactagcagc agtcaaacga acaagattaa ttacctgtct 1020
tcctgtttct caagaatatt tacgtagttt ttcataggtc tgtttttctt ttcatgcctc 1080
ttaaaaactt ctgtgcttac ataaacatac ttaaaagggt ttctttaaga tattttattt 1140
ttccatttaa aggtggacaa aagctacctc ctaaaaagta aatacaaaga gaacttattt 1200
acacagggaa ggtttaagac tgttcaagta gcattccaat ctgtagccat gccacagaat 1260
atcaacaaga acacagaatg agtgcacagc taagagatca agtttcagca ggcagcttta 1320
tctcaacctg gacatatttt aagattcagc atttgaaaga tttccctagc ctcttctctt 1380
ttcattagcc caaaacggtg caactctatt ctggacttta ttacttgatt ctgtcttctg 1440
tataactctg aagtccacca aaagtggacc ctctatattt cctccctttt tatagtctta 1500
taagatacat tatgaaaggt gaccgactct attttaaate tcagaatttt aagttctagc 1560
cccatgataa ctttttctt tgtaatttat gtttcatat atccttggtc ccagagatgt 1620
ttagacaatt ttaggtcaa aaattaaagc taacacagga aaaggaactg tactggctat 1680
tacataagaa acaatggacc caagagaaga aaaggaagaa agaaaggttt tttggttttt 1740
gttttgtttt gttttgtttt ttgttttttt gagatggagt ctactcttt cgcccaggct 1800
ggagtgcagt ggtatgatct cagctcactg caagctccac ctcccgggtt cagccattc 1860
tcctgcctca gcctcctgag tagctgggac tacaggcgcc cgccaccaca cccggcta 1920

tttttgtatt tttttagag acggggtttc accatgtag ccaagatggt ctcatctcc 1980
 tgacctcgtg atccacctgc ctgcgcctcc caaagtgctg ggattacggg tgtgagccac 2040
 cgtgcccagc cttttttttt ttaatagaaa aaataatccg actcccacta catcaagact 2100
 aatcttgttt tgtgtgtttt tcacatgtat tatagaatgc ttttgcatgg actatcctct 2160
 tgttttttatt aaaaacaaat gattttttta aaagtcacaa aaacaattca ctaaaaaataa 2220
 atatgtcatt gtgctttaaa aaaataacct cttgtagtta taaaataaaa cgtttgactt 2280
 ctaaaactctg 2290

<210> 60

<211> 2470

<212> DNA

<213> 智人 (Homo sapiens)

<400> 60

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60
 aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120
 tggagtcgtt ggtgaagttt ttcattcctc agaggagaaa atctgtggct ggggagattg 180
 ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaac 240
 gacagagcat attggttctg tgggatatta ataagcgcgg tgtggaggaa actgcagctg 300
 agtgccgaaa actaggcgtc actgcgcatg cgtatgtggt agactgcagc aacagagaag 360
 agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420
 ataatgctgg gacagtatat ccagccgatc ttctcagcac caaggatgaa gagattacca 480
 agacatttga ggtcaacatc ctaggacatt tttggaatgg aaaggacatc agaagtaatt 540
 acttggatgt atataggatc gaggacactt ttggacgaga ctctgagatc acaaaagcac 600
 ttcttccatc gatgatggag agaaatcatg gccacatcgt cacagtggct tcagtgtgcg 660
 gccacgaagg gattccttac ctcatcccat attgttcag caaatttgcc gctgttggct 720
 ttcacagagg tctgacatca gaacttcagg ctttgggaaa aactggtatc aaaacctcat 780
 gtctctgccc agtttttgtg aatactgggt tcacaaaaa tccaagcaca agattatggc 840
 ctgtatttga gacagatgaa gtcgtaagaa gtctgataga tggaatactt accaataaga 900
 aaatgatttt tgttccatcg tatatcaata tctttctgag actacagaag gtttcttcct 960
 gaacgcgcct cagcgatttt aaatcgtatg cagaatattc aatttgaagc agtggttggc 1020
 cacaaaatca aaatgaaatg aataaataag ctccagccag agatgtatgc atgataatga 1080
 tatgaatagt ttcgaatcaa tgctgcaaag ctttatttca ctttttttca gtcctgataa 1140
 tattaaaaac attggtttgg cactagcagc agtcaaacga acaagattaa ttacctgtct 1200
 tcctgtttct caagaatatt tacgtagttt ttcataggte tgtttttctt ttcatgcctc 1260
 ttaaaaactt ctgtgcttac ataaacatac ttaaaagggt ttctttaaga tatttttatt 1320
 ttccatttaa aggtggacaa aagctacctc cctaaaagta aatacaaaga gaacttattt 1380
 acacagggaa ggtttaagac tgttcaagta gcattccaat ctgtagccat gccacagaat 1440
 atcaacaaga acacagaatg agtgcacagc taagagatca agtttcagca ggcagcttta 1500
 tctcaacctg gacatatatt aagattcagc atttgaaaga tttccctagc ctcttccttt 1560
 ttcattagcc caaaacggtg caactctatt ctggacttta ttacttgatt ctgtcttctg 1620

tataactctg aagtccacca aaagtggacc ctctatatatt cctccctttt tatagtctta 1680
 taagatacat tatgaaaggt gaccgactct attttaaatc tcagaatttt aagttctagc 1740
 cccatgataa cctttttctt tgtaatttat gctttcatat atccttggtc ccagagatgt 1800
 ttagacaatt ttaggctcaa aaattaaagc taacacagga aaaggaactg tactggctat 1860
 tacataagaa acaatggacc caagagaaga aaaggaagaa agaaaggttt tttggttttt 1920
 gttttgtttt gttttgtttt ttgttttttt gagatggagt ctactcttt cgcccaggct 1980
 ggagtgcagt ggtatgatct cagctcactg caagctccac ctcccgggtt cagccattc 2040
 tcctgcctca gcctcctgag tagctgggac tacaggcgcc cgccaccaca cccggctaata 2100
 tttttgtatt tttttagag acgggggttc accatgttag ccaagatggc ctgatctcc 2160
 tgacctcgtg atccacctgc ctcggcctcc caaagtgtg ggattacggg tgtgagccac 2220
 cgtgcccagc cttttttttt ttaatagaaa aaataatccg actcccacta catcaagact 2280
 aatcttgttt tgttgtttt tcacatgtat tatagaatgc ttttgcattg actatcctct 2340
 tgtttttatt aaaaacaaat gattttttta aaagtcacaa aaacaattca ctaaaaataa 2400
 atatgtcatt gtgctttaa aaaataacct cttgtagtta taaaataaaa cgtttgactt 2460
 ctaaactctg 2470

<210> 61

<211> 1714

<212> DNA

<213> 智人(Homo sapiens)

<400> 61

agacagtacc tcctccctag gactacacaa ggactgaacc agaaggaaga ggacagagca 60
 aagccatgaa catcatccta gaaatccttc tgcttctgat caccatcatc tactcctact 120
 tggagtcgtt ggtgaagttt ttcatcctc agaggagaaa atctgtggct ggggagattg 180
 ttctcattac tggagctggg catggaatag gcaggcagac tacttatgaa tttgcaaaac 240
 gacagagcat attggttctg tgggatatta ataagcgcgg tgtggaggaa actgcagctg 300
 agtgccgaaa actaggcgtc actgcgcatg cgtatgtggt agactgcagc aacagagaag 360
 agatctatcg ctctctaaat caggtgaaga aagaagtggg tgatgtaaca atcgtggtga 420
 ataatgctgg gacagtatat ccagccgac ttctcagcac caaggatgaa gagattacca 480
 agacatttga ggtcaacatc ctaggacatt tttggatcac aaaagcactt cttccatcga 540
 tgatggagag aaatcatggc cacatcgtca cagtggcttc agtgtgcggc cacgaaggga 600
 ttcccttacct catcccatat tgttccagca aatttgccgc tgttggcttt cacagaggctc 660
 tgacatcaga acttcaggcc ttgggaaaaa ctggtatcaa aacctcatgt ctctgcccag 720
 tttttgtgaa tactgggttc accaaaaatc caagcacaag attatggcct gtattggaga 780
 cagatgaagt cgtaagaagt ctgatagatg gaatacttac caataagaaa atgatttttg 840
 ttccatcgta tatcaatatc tttctgagac tacagaagta agtacagcac agaacacca 900
 aataactaaaa caccaataga gctttttttt ttgctttttt tttttttaga cagagtctca 960
 ctctgtcacc ctggctggat tgcggtggtt gcagtggcat gatcttggt cactgcaacc 1020
 tccgcctcct gggttcaagc aattctcatg cctcagaccc ccaagtaact gggattatag 1080
 gtgtgtgctg ccacactaca cccagctaata ttttgtattt tttgatagag acaggtttcc 1140

ccatgttggc caggctggac tcgaactcct gacctcaagt ttcctcctg tctcggcctc 1200
ccaaagtgct gggattacag tcatgagcca ccatgcctgg cccaatagag ctattattat 1260
ggagcatcctt tcagttgtga aaattggcat ggaaactctc catccctggg gagaacagtt 1320
atttcctctg ttattttcct acccagtcta taaaaagaga gtgattcatt ttctctacca 1380
aatctactgt ctctgccccaa actttgctga agactattct aactaaagga aacacagttt 1440
aaaaagaatg caatatagtg aagtagttaa taataaagac tccattttta aaagtctgct 1500
ggaagtttgg ttgggattgc actgaatcta tagagcaatt ggggagtatt gacatatcaa 1560
caatattgag ttttctaate caagaacata atatctattt ttaaaatctt cttcaaaatc 1620
tttaaatctt taaattgtat tttgtagttt ttggtgttta agtcttgac atattttgtc 1680
agattttattc caaagtattt cacgggttct tttt 1714

<210> 62

<211> 25

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 正向引物

<400> 62

atgaacatca tcctagaaat ccttc 25

<210> 63

<211> 25

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 反向引物

<400> 63

atcatgcata catctctggc tggag 25

<210> 64

<211> 20

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 引物

<400> 64

atcagaactt caggccttgg 20

<210> 65

<211> 21

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 第一个外显子
<400> 65
gcaaagccat gaacatcatc c 21
<210> 66
<211> 24
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 最后一个外显子
<400> 66
tcttgatgta gtgggagtcg gatt 24

特征	GHS发现群组 (N = 46,544)	GHS肥胖手术群组 (N = 2,644)	达拉斯心脏研究 (N = 1,357)	宾夕法尼亚大学 医学生物库(N = 8,526)
年龄(岁) --中值(IQR)	63 (50 - 74)	53 (44 - 61)	46 (38 - 54)	68 (60 - 76)
女性--数量(%)	26,875 (58)	2,119 (80)	724 (53)	3,242 (38)
体重指数--中值(IQR)	30 (25 - 45)	47 (42 - 54)	28 (25 - 32)	30 (25 - 32)
转氨酶水平(UL) --中值(IQR)				
丙氨酸转氨酶(ALT)	22.0 (17.0 - 29.0)	23.0 (17.5 - 29.5)	20.0 (15.0 - 27.0)	22.0 (17.0 - 30.0)
天冬氨酸转氨酶(AST)	23.0 (20.0 - 27.5)	23.0 (20.0 - 27.0)	21.0 (18.0 - 25.0)	24.0 (20.0 - 30.5)
肝病的存在--N(%)				
酒精性肝病	197 (0.4)	7 (0.3)	-	-
酒精性肝硬化	130 (0.3)	3 (0.1)	-	-
非酒精性(非病毒性)肝病	1,938 (4.2)	1,543 (58.4)	-	-
非酒精性肝硬化	382 (0.8)	24 (0.9)	-	-
肝细胞癌	76 (0.2)	1 (0.04)	-	-
无肝病	30,628 (65.8)	1 (0.04)	-	-

图1

性状	Chr	BP	Ref	Alt	rsID	基因	注释	AA 取代	β (SE)	P	AAF	N
ALT	1	220970028	A	G	rs2642438	MARC1	错义	p.Thr163Ala	0.008 (0.001)	4.67E-08	0.7067	40,834
	4	882331392	T	T	rs72613567	MSD170B3	剪接供体		0.009 (0.001)	4.26E-08	0.2534	40,834
	8	144997604	C	T	rs371119003	PLEC	错义	p.Ala2302Thr	-0.160 (0.026)	1.30E-09	0.0005	40,833
	8	145008502	G	A		PLEC	错义	p.Arg522Cys	-0.268 (0.032)	3.26E-17	0.0003	40,834
	8	145692918	G	A	rs35968570	KIFC2	错义	p.Glu174Lys	0.033 (0.005)	1.40E-11	0.0139	40,834
	8	145730072	G	A	rs143408057	GPT	错义	p.Arg83His	-0.314 (0.036)	3.28E-18	0.0003	40,834
	8	145730161	C	T	rs201815797	GPT	错义	p.Ala87Val	-0.274 (0.014)	6.28E-59	0.0018	40,834
	8	145730221	G	A	rs112574791	GPT	错义	p.Arg107Lys	-0.033 (0.005)	4.25E-11	0.0136	40,834
	8	145731636	T	G	rs145155876	GPT	停止获得	p.Tyr326*	-0.235 (0.031)	1.76E-14	0.0004	40,814
	8	145732114	G	C	rs141505249	GPT	错义	p.Glu430Gln	-0.224 (0.013)	8.84E-64	0.0019	40,795
	8	145732131	G	A	rs143462595	GPT	错义	p.Arg442His	-0.077 (0.013)	1.18E-09	0.0071	40,826
	8	145732180	G	C	rs147998249	GPT	错义	p.Val452Leu	-0.225 (0.013)	8.19E-65	0.0019	40,833
	8	145732305	G	GC		GPT	移码	p.Glu475fs	-0.271 (0.031)	1.00E-18	0.0004	40,834
	8	145748532	A	G	rs567402720	LRRRC24	错义	p.Leu290Ser	-0.185 (0.028)	3.42E-11	0.0004	40,813
	9	117122282	C	T	rs3748177	AKNA	同义	p.Glu755Glu	-0.007 (0.001)	9.51E-09	0.5232	40,834
	9	117124731	G	A	rs3748176	AKNA	错义	p.Pro624Leu	-0.007 (0.001)	4.31E-09	0.5230	40,832
	10	101595996	T	A	rs17227273	ABCC2	错义	p.Val1188Glu	-0.015 (0.003)	2.97E-08	0.0608	40,834
	10	101606861	G	T	rs1137968	ABCC2	同义	p.Val1430Val	-0.015 (0.003)	2.71E-08	0.0608	40,834
	10	101610533	C	T	rs8187707	ABCC2	同义	p.His1496His	-0.015 (0.003)	2.77E-08	0.0608	40,834
	10	101611294	G	A	rs8187710	ABCC2	错义	p.Cys1515Tyr	-0.015 (0.003)	2.15E-08	0.0611	40,834
	10	101612064	T	C	rs2862954	SRH1	错义	p.Leu291Val	-0.012 (0.001)	1.43E-07	0.4755	40,834

灰色阴影指示变体与ALT和AST二者都具有全外显子组显著关联。

缩写: AAF, 交替的等位基因频率; Alt, 交替的等位基因; ALT, 丙氨酸转氨酶;
AST, 天冬氨酸转氨酶; Ref, 参考等位基因; SE, 标准误差;

图2

性状	Chr	BP	Ref	Alt	rsID	基因	注释	AA取代	N				平均ALT或AST水平 (U/L)			
									Ref/Ref	Ref/Alt	Alt/Alt	Alt/Alt	Ref/Ref	Ref/Alt	Alt/Alt	Alt/Alt
ALT	1	2.2E+08	A	G	rs2642438	MARCI	错义	p.Thr165Ala	3,515	17,262	20,637	24.52	23.88	24.52	24.92	24.92
	4	8.8E+07	T	A	rs72813567	rs317913	剪接供体		22,441	16,130	2,843	24.26	25.02	24.26	24.1	24.1
	8	1.4E+08	C	T	rs371119003	PLEC	错义	p.Ala2302Thr	41,373	40	0	18.1	24.67	18.1	NA	NA
	8	1.5E+08	G	A		PLEC	错义	p.Arg322Cys	41,387	27	0	13.8	24.67	13.8	NA	NA
	8	1.5E+08	G	A	rs35968570	KIRC2	错义	p.Glu174Lys	40,271	1,133	10	12.07	24.67	12.07	NA	NA
	8	1.5E+08	G	A	rs143408057	GPI	错义	p.Arg83His	41,383	21	0	12.07	24.67	12.07	NA	NA
	8	1.5E+08	C	T	rs201815297	GPI	错义	p.Ala87Val	41,270	144	0	14.68	24.7	14.68	NA	NA
	8	1.5E+08	G	A	rs112574791	GPI	错义	p.Arg107Lys	40,293	1,111	10	23.09	24.71	23.09	38.35	38.35
	8	1.5E+08	T	G	rs145155876	GPI	停止获得	p.Tyr326*	41,364	30	0	14.07	24.67	14.07	NA	NA
	8	1.5E+08	G	C	rs143505249	GPI	错义	p.Glu430Gln	41,223	150	2	14.48	24.7	14.48	13.75	13.75
	8	1.5E+08	G	A	rs143462595	GPI	错义	p.Arg442His	41,232	174	0	20.87	24.68	20.87	NA	NA
	8	1.5E+08	G	C	rs147998249	GPI	错义	p.Val452Leu	41,254	159	0	14.74	24.7	14.74	NA	NA
	8	1.5E+08	G	GC		GPI	移码	p.Glu475fs	41,385	29	0	14.24	24.67	14.24	NA	NA
	8	1.5E+08	A	G	rs567402720	LHRC24	错义	p.Leu290Ser	41,358	35	0	17.71	24.67	17.71	NA	NA
	9	1.2E+08	C	T	rs3748177	AKNA	同义	p.Glu755Glu	9,414	20,645	11,355	24.72	25.12	24.72	24.18	24.18
	9	1.2E+08	G	A	rs3748176	AKNA	错义	p.Pro624Leu	9,427	20,634	11,351	24.73	25.12	24.73	24.17	24.17
	10	1E+08	T	A	rs17222723	ABCC2	错义	p.Val1188Glu	36,543	4,704	167	23.97	24.77	23.97	22.12	22.12
	10	1E+08	G	T	rs1137968	ABCC2	同义	p.Val1430Val	36,543	4,704	167	23.97	24.77	23.97	22.04	22.04
	10	1E+08	C	T	rs8187707	ABCC2	同义	p.His1496His	36,542	4,706	166	23.97	24.77	23.97	22.03	22.03
	10	1E+08	G	A	rs8187710	ABCC2	错义	p.Cys1515Tyr	36,519	4,726	169	23.97	24.77	23.97	21.99	21.99
	10	1E+08	T	C	rs2882954	FRIN	错义	p.Ile79Val	11,318	20,819	9,277	24.71	25.12	24.71	23.77	23.77

灰色阴影指示变体与ALT和AST二者都具有全外显子组显著的关联。

缩写: AAF, 交替的等位基因频率; Alt, 交替的等位基因; ALT, 丙氨酸转氨酶;

AST, 天冬氨酸转氨酶; Ref, 参考等位基因; SE, 标准误差;

图2(续)

性状	Chr	BP	Ref	Alt	rsID	基因	注释	AA取代	β [SE]	P	AAF	N
	10	101977883	C	T	rs2230804	CHUK	错义	p.Val268Ile	-0.009 (0.001)	1.93E-13	0.5072	40,834
	10	113917085	T	A	rs2254537	GPAM	同义	p.Pro681Pro	-0.008 (0.001)	4.61E-10	0.7073	40,834
	10	113940329	T	C	rs2792751	GPAM	错义	p.Ile43Val	-0.008 (0.001)	2.54E-10	0.7097	40,832
	14	94844947	C	T	rs28929474	SERPINA1	错义	p.Glu366Lys	0.042 (0.003)	9.28E-21	0.0171	40,834
	19	19379549	C	T	rs58542926	TM6SF2	错义	p.Glu167Lys	0.014 (0.002)	4.76E-09	0.0759	40,833
	22	44324727	C	G	rs738409	PNPLA3	错义	p.Ile148Met	0.023 (0.002)	1.34E-50	0.1351	40,834
	22	44324730	C	T	rs738408	PNPLA3	同义	p.Pro149Pro	0.023 (0.002)	1.11E-50	0.2349	40,834
	22	44342116	A	G	rs2294918	PNPLA3	错义	p.Lys434Glu	0.007 (0.001)	8.26E-08	0.5986	40,832
	22	44368122	A	G	rs3761472	SAMM50	错义	p.Asp110Gly	0.019 (0.002)	8.85E-30	0.1682	40,833
	22	44395451	T	C	rs1007863	PARNB	错义	p.Tyr37Arg	0.011 (0.001)	7.98E-16	0.3963	40,834
AST	4	88731392	T	TA	rs73613567	MSD17B13	剪接供体		-0.005 (0.001)	6.24E-10	0.2638	40,193
	10	18242311	A	G	rs10764176	SIC39A12	错义	p.Ser36Gly	-0.006 (0.001)	1.09E-10	0.2881	40,193
	10	101157378	CGTT	C		GOT1	框内插入缺失	p.Asn389del	-0.221 (0.024)	1.96E-20	0.0002	40,193
	10	101165533	G	C	rs374966349	GOT1	错义	p.Gln208Glu	0.271 (0.027)	2.43E-24	0.0302	40,193
	10	101912064	T	C	rs2862954	ERL1A1	错义	p.Ile291Val	-0.005 (0.001)	4.82E-09	0.4354	40,193
	11	22371870	A	T	rs7481951	ANO5	错义	p.Leu322Phe	0.004 (0.001)	9.61E-08	0.5833	40,162
	14	64844947	C	T	rs28929474	SERPINA1	错义	p.Glu366Lys	0.027 (0.003)	2.44E-20	0.0172	40,193
	19	19379549	C	T	rs58542926	TM6SF2	错义	p.Glu167Lys	0.008 (0.002)	6.54E-08	0.0760	40,192
	22	44324727	C	G	rs738409	PNPLA3	错义	p.Ile148Met	0.014 (0.001)	8.31E-46	0.2343	40,193
	22	44324730	C	T	rs738408	PNPLA3	同义	p.Pro149Pro	0.014 (0.001)	8.93E-46	0.2341	40,193
	22	44368122	A	G	rs3761472	SAMM50	错义	p.Asp110Gly	0.011 (0.001)	1.22E-22	0.1680	40,192
	22	44395451	T	C	rs1007863	PARNB	错义	p.Tyr37Arg	0.006 (0.001)	1.31E-13	0.3961	40,193

灰色阴影指示变体与ALT和AST二者都具有全外显子组显著关联。

缩写: AAF, 交替的等位基因频率; Alt, 交替的等位基因; ALT, 丙氨酸转氨酶;

AST, 天冬氨酸转氨酶; Ref, 参考等位基因; SE, 标准误差;

图2(续)

性状	Chr	BP	Ref	Alt	rsID	基因	注释	AA 取代	N			平均ALT或AST水平 (U/L)		
									Ref/Ref	Ref/Alt	Alt/Alt	Ref/Ref	Ref/Alt	Alt/Alt
	10	101977883	C	T	rs22230804	CHUK	错义	p.Val268Ile	10,048	20,733	10,633	25.18	24.75	24.01
	10	113917085	T	A	rs22254537	GPAM	同义	p.Pro681Pro	3,627	16,984	20,803	25	24.97	24.36
	10	113940329	T	C	rs2792751	GPAM	错义	p.Ile43Val	3,567	16,910	20,935	25	24.98	24.35
	14	94844947	C	T	rs28529474	SEPPIN1	错义	p.Glu366Iys	40,006	1,399	9	24.58	26.91	33.89
	19	19379549	C	T	rs5854926	TM6SF2	错义	p.Glu167Iys	35,388	5,280	243	24.33	25.46	26.84
	22	44324727	C	G	rs738439	PNPLA3	错义	p.Ile148Met	24,257	14,837	2,320	24.06	24.99	28.91
	22	44324730	C	T	rs738438	PNPLA3	同义	p.Pro149Pro	24,273	14,824	2,317	24.06	24.98	28.92
	22	44342116	A	G	rs2294918	PNPLA3	错义	p.Lys434Glu	6,691	19,833	14,888	24.15	24.47	25.15
	22	44368122	A	G	rs3761472	SNMM50	错义	p.Asp110Gly	28,626	11,618	1,169	24.23	25.36	28.45
	22	44395451	T	C	rs1007863	PARVB	错义	p.Trp37Arg	15,036	19,920	6,458	24.15	24.6	26.09
AST	4	88231392	T	T	rs72613507	HSN17B13	剪接供体		22,068	15,870	2,815	24.47	24.1	23.96
	10	18242311	A	G	rs10754176	SCC39A12	错义	p.Ser36Gly	20,645	16,738	3,370	24.47	24.15	23.85
	10	101157378	CGTT	C		GOT1	框内插入缺失	p.Asn389del	40,733	20	0	24.29	14.7	NA
	10	101165533	G	C	rs374966349	GOT1	错义	p.Gln208Glu	40,736	17	0	24.28	44.5	NA
	10	101913064	T	C	rs2862354	ERLWI	错义	p.Ile291Val	11,138	20,486	9,129	24.59	24.26	23.99
	11	72271870	A	T	rs7481951	ANC5	错义	p.Leu322Phe	7,123	19,686	13,913	24.03	24.22	24.53
	14	94844947	C	T	rs28529474	SEPPIN1	错义	p.Glu366Iys	39,361	1,384	8	24.24	25.76	30.5
	19	19379549	C	T	rs5854926	TM6SF2	错义	p.Glu167Iys	34,811	5,698	243	24.21	24.74	25.43
	22	44324727	C	G	rs738439	PNPLA3	错义	p.Ile148Met	23,889	14,622	2,242	23.96	24.48	26.62
	22	44324730	C	T	rs738438	PNPLA3	同义	p.Pro149Pro	23,905	14,609	2,239	23.96	24.47	26.63
	22	44368122	A	G	rs3761472	SNMM50	错义	p.Asp110Gly	28,170	11,430	1,132	24.07	24.64	26.24
	22	44395451	T	C	rs1007863	PARVB	错义	p.Trp37Arg	14,761	19,678	6,314	24.02	24.23	25.1

灰色阴影指示变体与ALT和AST二者都具有全外显子组显著关联。

缩写: AAF, 交替的等位基因频率; Alt, 交替的等位基因; ALT, 丙氨酸转氨酶;

AST, 天冬氨酸转氨酶; Ref, 参考等位基因; SE, 标准误差;

图2(续)

性状	Chr	BP	Ref	Alt	RSID	基因	注释	AA 取代	GHS 发现群组	
									B[SE]	P 发现 N
ALT	1	220970028	A	G	rs2642438	MARCI	错义	p.Thr165Ala	0.008 (0.001)	4.67E-08 40834
	4	88231392	T	TA	rs72633567	HS017B13	剪接供体		-0.009 (0.001)	4.16E-12 40834
	8	144997604	C	T	rs371119003	PLEC	错义	p.Ala2302Thr	-0.160 (0.026)	1.30E-09 40833
	8	145008502	G	A		PLEC	错义	p.Arg572Cys	-0.268 (0.032)	3.26E-17 40834
	8	145692918	G	A	rs35968570	KIFC2	错义	p.Glu178Lys	-0.033 (0.005)	1.40E-11 40834
	8	145730072	G	A	rs143408057	GPT	错义	p.Arg83His	-0.314 (0.036)	3.28E-18 40834
	8	145730161	C	T	rs201815297	GPT	错义	p.Ala87Val	-0.224 (0.014)	6.28E-59 40834
	8	145730221	G	A	rs112524791	GPT	错义	p.Arg107Lys	-0.033 (0.005)	4.25E-11 40834
	8	145731636	T	G	rs145355876	GPT	停止获得	p.Lys226*	-0.235 (0.031)	1.76E-14 40814
	8	145732114	G	C	rs141505249	GPT	错义	p.Glu430Gln	-0.224 (0.013)	8.84E-64 40795
	8	145732151	G	A	rs143462595	GPT	错义	p.Arg442His	-0.077 (0.013)	1.18E-09 40826
	8	145732180	G	C	rs147998249	GPT	错义	p.Val452Leu	-0.225 (0.013)	8.19E-65 40833
	8	145732305	G	GC		GPT	移码	p.Glu479fs	-0.271 (0.031)	1.00E-18 40834
	8	145748532	A	G	rs567402720	LARC24	错义	p.Leu290Ser	-0.185 (0.028)	3.42E-11 40813
	9	117122202	C	T	rs3748177	AKNA	同义	p.Glu735Glu	-0.007 (0.001)	9.51E-09 40834
	9	117124731	G	A	rs3748176	AKNA	错义	p.Pro624Leu	-0.007 (0.001)	4.31E-09 40832
	10	101595996	T	A	rs17222723	ABCC2	错义	p.Val1188Glu	-0.015 (0.003)	2.97E-08 40834

灰色阴影指示P值满足Bonferroni显著性阈值 $P < 1.43 \times 10^{-3}$

*复制综合分析包括三个复制群组: GHS肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库

*联合综合分析包括发现群组和三个复制群组: GHS发现群组、GHS肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库

缩写: AAF, 交替的等位基因频率; Alt, 交替的等位基因; Ref, 参考等位基因; SE, 标准误差; 天冬氨酸转氨酶; Ref, 参考等位基因; SE, 标准误差;

图3

性状	Chr	BP	Ref	Alt	RSID	基因	注释	AA 取代	GHS 发现群组		
									β (SE)	P 发现	N
	10	101606861	G	T	rs113793	ABCC2	同义	p.Val1430Val	-0.015 {0.003}	2.71E-08	40834
	10	101610533	C	T	rs818770	ABCC2	同义	p.His1496His	-0.015 {0.003}	2.77E-08	40834
	10	101611294	G	A	rs818771	ABCC2	错义	p.Cys1515Tyr	-0.015 {0.003}	2.15E-08	40834
	10	101912064	T	C	rs286295	ERL1N1	错义	p.Ile291Val	-0.012 {0.001}	2.43E-21	40834
	10	101977883	C	T	rs223980	CHUK	错义	p.Val268Ile	-0.009 {0.001}	1.93E-13	40834
	10	113917085	T	A	rs225453	GPAM	同义	p.Pro681Pro	-0.008 {0.001}	4.61E-10	40834
	10	113940329	T	C	rs279775	GPAM	错义	p.Ile43Val	-0.008 {0.001}	2.54E-10	40832
	14	94844947	C	T	rs289294	SERPINA4	错义	p.Glu366Lys	0.042 {0.005}	9.28E-21	40834
	19	19379549	C	T	rs585429	TNFSF2	错义	p.Glu167Lys	0.014 {0.002}	4.76E-09	40833
	22	44324727	C	G	rs738409	PNPLA3	错义	p.Ile148Met	0.023 {0.002}	1.34E-50	40834
	22	44324730	C	T	rs738408	PNPLA3	同义	p.Pro149Pro	0.023 {0.002}	1.11E-50	40834
	22	44342116	A	G	rs229491	PNPLA3	错义	p.Lys434Glu	0.007 {0.001}	8.26E-08	40832
	22	44368122	A	G	rs376147	SAMM50	错义	p.Asp110Gly	0.019 {0.002}	8.85E-30	40833
	22	44395451	T	C	rs100786	PARVB	错义	p.Trp37Arg	0.011 {0.001}	7.98E-16	40834
灰色阴影指示P值满足Bonferroni显著性阈值 $P < 1.43 \times 10^{-3}$											
* 复制综合分析包括三个复制群组: GHS 肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库											
* 联合综合分析包括发现群组和三个复制群组: GHS 发现群组、GHS 肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库											
缩写: AAF, 交替的等位基因频率; Alt, 交替的等位基因; Ref, 参考等位基因; SE, 标准误差; 天冬氨酸转氨酶; Ref, 参考等位基因; SE, 标准误差;											

图3(续)

性状	Chr	BP	Ref	Alt	RSID	基因	注释	AA 取代	GHS 发现群组		
									β (SE)	P 发现	N
AST	4	88231392	T	TA	rs72613567	HSD17B13	剪接供体		-0.005 (0.001)	6.24E-10	40193
	10	18242311	A	G	rs10764176	SLC39A12	错义	p.Ser36Gly	-0.006 (0.001)	1.09E-10	40193
	10	101157378	CGTT	C		GOT1	框内插入缺失	p.Asn389del	-0.221 (0.024)	1.96E-20	40193
	10	101165533	G	C	rs374966349	GOT1	错义	p.Gln208Glu	0.271 (0.027)	2.43E-24	40193
	10	101912064	T	C	rs2862954	ERLIN1	错义	p.Ile291Val	-0.005 (0.001)	4.82E-09	40193
	11	22271870	A	T	rs7481951	ANC5	错义	p.Leu322Phe	0.004 (0.001)	9.61E-08	40162
	14	94844947	C	T	rs28929474	SERPIN1	错义	p.Glu366Iys	0.027 (0.003)	2.44E-20	40193
	19	19379549	C	T	rs58542926	TM6SF2	错义	p.Glu167Iys	0.008 (0.002)	6.54E-08	40192
	22	44324727	C	G	rs738409	PNPLA3	错义	p.Ile148Met	0.014 (0.001)	8.31E-46	40193
	22	44324730	C	T	rs738408	PNPLA3	同义	p.Pro149Pro	0.014 (0.001)	8.93E-46	40193
	22	44368122	A	G	rs3761472	SAMM50	错义	p.Asp110Gly	0.011 (0.001)	1.22E-22	40192
	22	44395451	T	C	rs1007863	PARVB	错义	p.Trp37Arg	0.006 (0.001)	1.31E-13	40193
灰色阴影指示P值满足Bonferroni显著性阈值 $P < 1.43 \times 10^{-3}$											
*复制综合分析包括三个复制群组;GHS 肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库											
*联合综合分析包括发现群组和三个复制群组:GHS 发现群组、GHS 肥胖手术群组、达拉斯心脏研究											
和宾夕法尼亚大学医学生物库											
缩写:AST,交替的等位基因频率;Alt,交替的等位基因;ALT,丙氨酸转氨酶;AST,天冬氨酸转氨酶;Ref,参考等位基因;SE,标准误差;											

图3(续)

性状	Chr	复制群组									
		GHS肥胖手术群组					达拉斯心脏研究				
		β (SE)	P	N	β (SE)	P	β (SE)	P	N	β (SE)	P
ALT	1	0.005 (0.005)	3.10E-01	2475	0.009 (0.008)	2.58E-01	0.006 (0.004)	1.81E-01	1356	0.006 (0.004)	1.81E-01
	4	-0.010 (0.005)	5.57E-02	2475	-0.014 (0.008)	9.68E-02	-0.012 (0.004)	4.85E-03	1356	-0.012 (0.004)	4.85E-03
	8	-0.492 (0.165)	2.84E-03	2475	NA (NA)	NA	-0.054 (0.071)	4.46E-01	NA	-0.054 (0.071)	4.46E-01
	8	-0.161 (0.165)	3.29E-01	2475	NA (NA)	NA	-0.259 (0.143)	6.90E-02	NA	-0.259 (0.143)	6.90E-02
	8	-0.009 (0.02)	6.48E-01	2475	0.027 (0.035)	4.48E-01	-0.051 (0.019)	7.52E-03	1355	-0.051 (0.019)	7.52E-03
	8	-0.189 (0.165)	2.50E-01	2475	NA (NA)	NA	-0.305 (0.101)	2.54E-03	NA	-0.305 (0.101)	2.54E-03
	8	-0.341 (0.074)	3.64E-06	2475	NA (NA)	NA	-0.144 (0.054)	7.67E-03	NA	-0.144 (0.054)	7.67E-03
	8	-0.009 (0.02)	6.45E-01	2475	0.024 (0.035)	5.01E-01	-0.059 (0.018)	1.13E-03	1356	-0.059 (0.018)	1.13E-03
	8	-0.314 (0.165)	5.71E-02	2475	-0.334 (0.137)	1.49E-02	-0.151 (0.143)	2.90E-01	1355	-0.151 (0.143)	2.90E-01
	8	-0.273 (0.048)	9.83E-09	2474	-0.244 (0.073)	8.91E-04	-0.188 (0.041)	5.52E-06	1356	-0.188 (0.041)	5.52E-06
	8	-0.115 (0.058)	4.82E-02	2475	-0.092 (0.097)	3.43E-01	-0.042 (0.043)	3.36E-01	1355	-0.042 (0.043)	3.36E-01
	8	-0.273 (0.050)	4.26E-08	2475	-0.198 (0.068)	3.90E-03	-0.188 (0.041)	5.52E-06	1356	-0.188 (0.041)	5.52E-06
	8	-0.161 (0.165)	3.29E-01	2475	NA (NA)	NA	-0.506 (0.202)	1.22E-02	NA	-0.506 (0.202)	1.22E-02
	8	-0.161 (0.165)	3.29E-01	2475	NA (NA)	NA	-0.303 (0.143)	3.37E-02	NA	-0.303 (0.143)	3.37E-02
	9	-0.004 (0.005)	4.09E-01	2475	0.003 (0.008)	6.46E-01	-0.007 (0.004)	6.38E-02	1356	-0.007 (0.004)	6.38E-02
灰色阴影指示P值满足Bonferroni显著性阈值 $P < 1.43 \times 10^{-3}$											
*复制综合分析包括三个复制群组: GHS肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学图书馆											
*联合综合分析包括发现群组三个复制群组: GHS发现群组, GHS肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学图书馆											
缩写: AAT, 交替的等位基因频率; ALT, 交替的等位基因; ALT, 丙氨酸转氨酶; AST, 天冬氨酸转氨酶; Ref, 参考等位基因; SE, 标准误差;											

图3(续)

性状	Chr	复制群组									
		GHS肥胖手术群组					达拉斯心脏研究				
		β (SE)	P	N	β (SE)	P	β (SE)	P	N	β (SE)	P
	9	-0.004 (0.005)	3.90E-01	2475	0.002 (0.008)	7.69E-01	1355	-0.007 (0.004)	5.29E-02	6158	
	10	-0.002 (0.010)	8.01E-01	2475	-0.003 (0.017)	8.37E-01	1356	-0.015 (0.007)	4.49E-02	6158	
	10	-0.003 (0.010)	7.74E-01	2475	-0.005 (0.017)	7.49E-01	1356	-0.014 (0.007)	4.86E-02	6158	
	10	-0.003 (0.010)	7.93E-01	2475	-0.005 (0.017)	7.49E-01	1356	-0.014 (0.007)	5.02E-02	6158	
	10	-0.001 (0.010)	9.11E-01	2475	-0.008 (0.016)	6.41E-01	1356	-0.013 (0.007)	7.46E-02	6158	
	10	-0.01 (0.005)	2.91E-02	2475	-0.006 (0.007)	4.02E-01	1356	-0.009 (0.004)	2.06E-02	6158	
	10	-0.006 (0.005)	2.05E-01	2475	-0.001 (0.007)	9.07E-01	1356	-0.011 (0.004)	5.26E-03	6158	
	10	-0.003 (0.005)	5.80E-01	2475	-0.014 (0.008)	8.25E-02	1356	-0.007 (0.004)	7.45E-02	6158	
	10	-0.003 (0.005)	5.61E-01	2475	-0.014 (0.008)	9.08E-02	1356	-0.008 (0.004)	6.34E-02	6158	
	14	0.035 (0.020)	7.97E-02	2475	0.044 (0.032)	1.63E-01	1356	0.056 (0.013)	1.38E-05	6158	
	19	0.040 (0.010)	2.40E-05	2475	0.025 (0.014)	7.24E-02	1356	0.013 (0.008)	1.07E-01	6158	
	22	0.019 (0.006)	5.54E-04	2475	0.005 (0.009)	5.75E-01	1356	0.018 (0.005)	5.51E-05	6158	
	22	0.019 (0.006)	5.51E-04	2475	0.005 (0.009)	5.75E-01	1356	0.018 (0.005)	5.78E-05	6158	
	22	0.001 (0.005)	7.77E-01	2475	0.004 (0.008)	6.26E-01	1356	0.005 (0.004)	2.00E-01	6158	
	22	0.009 (0.006)	1.66E-01	2475	-0.002 (0.01)	8.80E-01	1356	0.021 (0.005)	5.29E-05	6158	
	22	0.003 (0.005)	5.22E-01	2475	0.007 (0.008)	3.31E-01	1356	0.008 (0.004)	3.82E-02	6158	
灰色阴影指示P值满足Bonferroni显著性阈值 $P < 1.43 \times 10^{-3}$											
*复制综合分析包括三个复制群组:GHS肥胖手术群组、达拉斯心脏研究和											
宾夕法尼亚大学医学图书馆											
*联合综合分析包括发现群组和三个复制群组:GHS发现群组、GHS肥胖手术群组、达拉斯心脏研究											
和宾夕法尼亚大学医学图书馆											
缩写:AAE,交替的等位基因频率;Alt,交替的等位基因;ALT,丙氨酸转氨酶;AST,											
天冬氨酸转氨酶;Ref,参考等位基因;SE,标准误差;											

图3(续)

性状	Chr	复制综合分析*		联合综合分析**	
		β (SE)	P复制	β (SE)	P联合
ALT	1	0.006 (0.003)	4.77E-02	0.007 (0.001)	6.91E-09
	4	-0.012 (0.003)	1.67E-04	-0.010 (0.001)	3.85E-15
	8	-0.124 (0.066)	5.92E-02	-0.155 (0.024)	2.41E-10
	8	-0.259 (0.143)	6.90E-02	-0.264 (0.03)	4.65E-18
	8	-0.024 (0.013)	6.89E-02	-0.032 (0.005)	3.36E-12
	8	-0.305 (0.101)	2.54E-03	-0.308 (0.033)	2.21E-20
	8	-0.213 (0.044)	1.01E-06	-0.223 (0.013)	4.00E-64
	8	-0.029 (0.013)	2.09E-02	-0.032 (0.005)	2.89E-12
	8	-0.264 (0.085)	1.84E-03	-0.238 (0.029)	1.35E-16
	8	-0.227 (0.029)	2.39E-15	-0.224 (0.012)	1.94E-77
	8	-0.070 (0.033)	3.12E-02	-0.076 (0.012)	1.12E-10
	8	-0.218 (0.029)	3.94E-14	-0.224 (0.012)	2.92E-77
	8	-0.506 (0.202)	1.22E-02	-0.272 (0.030)	6.42E-20
	8	-0.303 (0.143)	3.37E-02	-0.189 (0.027)	2.99E-12
	9	-0.005 (0.003)	9.79E-02	-0.007 (0.001)	3.57E-09

灰色阴影指示P值满足Bonferroni显著性阈值 $P < 1.43 \times 10^{-3}$

*复制综合分析包括三个复制群组: GHS肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库

*联合综合分析包括发现群组和三个复制群组: GHS发现群组、GHS肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库

缩写: ALT, 交替的等位基因频率; ALT, 交替的等位基因; ALT, 丙氨酸转氨酶; ALT, 天冬氨酸转氨酶; Ref, 参考等位基因; SE, 标准误差;

图3 (续)

性状	Chr	复制综合分析*		联合综合分析**	
		β{SE}	P复制 (N=3)	β{SE}	P联合 (N=4)
	9	-0.005 {0.003}	7.27E-02	-0.007 {0.001}	1.19E-09
	10	-0.01 {0.006}	8.55E-02	-0.014 {0.002}	9.49E-09
	10	-0.01 {0.006}	7.96E-02	-0.014 {0.002}	8.03E-09
	10	-0.01 {0.006}	8.38E-02	-0.014 {0.002}	8.70E-09
	10	-0.009 {0.006}	1.19E-01	-0.014 {0.002}	1.07E-08
	10	-0.009 {0.003}	1.14E-03	-0.011 {0.001}	1.76E-23
	10	-0.008 {0.003}	5.11E-03	-0.009 {0.001}	4.17E-15
	10	-0.007 {0.003}	2.44E-02	-0.008 {0.001}	3.82E-11
	10	-0.007 {0.003}	2.13E-02	-0.008 {0.001}	1.85E-11
	14	0.049 {0.010}	1.51E-06	0.044 {0.004}	9.55E-26
	19	0.024 {0.006}	1.58E-05	0.016 {0.002}	1.35E-12
	22	0.017 {0.003}	2.60E-07	0.022 {0.001}	7.75E-56
	22	0.017 {0.003}	2.71E-07	0.022 {0.001}	6.77E-56
	22	0.004 {0.003}	2.03E-01	0.006 {0.001}	6.49E-08
	22	0.014 {0.004}	2.57E-04	0.018 {0.002}	2.36E-32
	22	0.006 {0.003}	2.63E-02	0.01 {0.001}	1.83E-16

灰色阴影指示P值满足Bonferroni显著性阈值 $P < 1.4 \times 10^{-3}$

*复制综合分析包括三个复制群组: GHS肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库

**联合综合分析包括发现群组和三个复制群组: GHS发现群组, GHS肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库

缩写: AAF, 交替的等位基因频率; ALT, 交替的等位基因; ALZ, 丙氨酸转氨酶; AST, 天冬氨酸转氨酶; Ref, 参考等位基因; SE, 标准误差;

图3(续)

性状	Chr	复制群组									
		GHS 肥胖手术群组					达拉斯心脏研究				
		β (SE)	P	N	β (SE)	P	β (SE)	P	N	β (SE)	P
AST	4	-0.010 (0.003)	3.12E-03	2469	-0.012 (0.006)	5.87E-02	1356	-0.006 (0.004)	1.02E-01	6166	
	10	-0.010 (0.003)	2.91E-03	2469	-0.003 (0.006)	6.25E-01	1356	-0.010 (0.004)	6.75E-03	6166	
	10	-0.205 (0.062)	8.57E-04	2469	NA (NA)	NA	NA	-0.244 (0.089)	5.90E-03	6165	
	10	NA (NA)	NA	2469	NA (NA)	NA	NA	0.339 (0.079)	1.85E-05	6166	
	10	-0.004 (0.003)	1.54E-01	2469	-0.007 (0.006)	2.18E-01	1356	-0.003 (0.003)	3.13E-01	6166	
	11	-0.001 (0.003)	7.85E-01	2466	0.006 (0.006)	2.80E-01	1356	-0.003 (0.003)	3.54E-01	6165	
	14	0.023 (0.013)	7.79E-02	2469	0.046 (0.024)	6.09E-02	1356	0.052 (0.011)	4.75E-06	6166	
	19	0.023 (0.006)	1.99E-04	2469	0.010 (0.011)	3.42E-01	1356	0.004 (0.007)	5.94E-01	6166	
	22	0.014 (0.004)	1.27E-04	2469	0.004 (0.007)	5.53E-01	1356	0.017 (0.004)	1.16E-05	6166	
	22	0.014 (0.004)	1.32E-04	2469	0.004 (0.007)	5.53E-01	1356	0.017 (0.004)	1.17E-05	6166	
	22	0.008 (0.004)	6.03E-02	2469	-0.001 (0.008)	9.33E-01	1356	0.018 (0.005)	6.47E-05	6166	
	22	0.003 (0.003)	4.12E-01	2469	0.006 (0.006)	3.03E-01	1356	0.009 (0.003)	1.37E-02	6166	
灰色阴影指示P值满足Bonferroni显著性阈值 $P < 1.43 \times 10^{-3}$											
*复制综合分析包括三个复制群组: GHS 肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库											
*联合综合分析包括发现群组三个复制群组: GHS 发现群组、GHS 肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库											
缩写: AAT, 交替的等位基因频率; AAT, 交替的等位基因; AAT, 丙氨酸转氨酶; AAT, 天冬氨酸转氨酶; AAT, 参考等位基因; SE, 标准误差;											

图3(续)

性状	Chr	复制综合分析*			联合综合分析**		
		β (SE)	P复制	(N=3)	β (SE)	P联合	(N=4)
AST	4	-0.009 (0.002)	1.69E-04		-0.006 (0.001)	1.13E-12	
	10	-0.009 (0.002)	8.86E-05		-0.006 (0.001)	9.66E-14	
	10	-0.205 (0.062)	8.57E-04		-0.220 (0.022)	1.66E-24	
	10	0.339 (0.079)	1.85E-05		0.271 (0.027)	2.43E-24	
	10	-0.004 (0.002)	3.92E-02		-0.005 (0.001)	5.52E-10	
	11	-0.001 (0.002)	7.03E-01		0.004 (0.001)	1.48E-06	
	14	0.040 (0.008)	6.56E-07		0.029 (0.003)	2.78E-25	
	19	0.014 (0.004)	1.20E-03		0.009 (0.002)	5.92E-10	
	22	0.014 (0.002)	2.00E-08		0.014 (0.001)	1.12E-52	
	22	0.014 (0.002)	2.10E-08		0.014 (0.001)	1.26E-52	
	22	0.011 (0.003)	1.77E-04		0.011 (0.001)	1.01E-25	
	22	0.005 (0.002)	1.34E-02		0.006 (0.001)	6.61E-15	

灰色阴影指示 P 值满足Bonferroni显著性阈值 $P < 1.43 \times 10^{-3}$

*复制综合分析包括三个复制群组: GHS肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库

*联合综合分析包括发现群组和三个复制群组: GHS发现群组, GHS肥胖手术群组、达拉斯心脏研究和宾夕法尼亚大学医学生物库

缩写: AAs, 交替的等位基因频率; AEs, 交替的等位基因; AEs, 丙氨酸转氨酶; AST, 天冬氨酸转氨酶; Ref, 参考等位基因; SE, 标准误差;

图3(续)

CHR:BP:Ref:Alt	基因	rsID	酒精性肝病		酒精性肝硬化	
			OR (95% CI)	P	OR (95% CI)	P
4:88231392:T:TA	HSD17B13	rs72613567	0.62 (0.48-0.81)	1.82E-04	0.56 (0.41-0.78)	3.35E-04
8:145730161:C:T	GPT	rs201815297	3.83 (1.05-13.94)	8.88E-02	6.33 (1.71-23.43)	2.88E-02
8:145732114:G:C	GPT	rs141505249	0.77 (0.06-10.73)	8.43E-01	1.13 (0.08-15.39)	9.30E-01
8:145732180:G:C	GPT	rs147998249	0.73 (0.05-11.76)	8.17E-01	1.07 (0.07-17.16)	9.60E-01
10:18242311:A:G	SLC39A12	rs10764176	0.85 (0.68-1.07)	1.64E-01	0.92 (0.70-1.22)	5.80E-01
10:101157378:CGTT:C	GOI1		4.60 (0.25-86.41)	3.93E-01	7.11 (0.38-133.19)	3.00E-01
10:101165533:G:C	GOI1	rs374966349	2.20 (0.13-37.68)	6.24E-01	3.47 (0.20 - 59.04)	4.70E-01
10:101912064:T:C	ERLIN1	rs2862954	0.92 (0.75-1.12)	4.05E-01	1.05 (0.82-1.34)	7.13E-01
14:94844947:C:T	SERPINA1	rs28929474	2.49 (1.49-4.17)	2.30E-03	3.35 (1.93-5.83)	3.01E-04
19:19379549:C:T	TM6SF2	rs58542926	1.47 (1.06-2.04)	2.76E-02	1.35 (0.89-2.04)	1.80E-01
22:44324727:C:G	PNPLA3	rs738409	1.76 (1.43-2.18)	4.98E-07	2.07 (1.60-2.67)	1.08E-07
22:44324730:C:T	PNPLA3	rs738408	1.77 (1.43-2.18)	4.70E-07	2.07 (1.61-2.67)	1.03E-07
22:44368122:A:G	SAMM50	rs3761472	1.90 (1.52-2.38)	1.36E-07	2.28 (1.75-2.98)	1.83E-08

灰色阴影指示P值满足Bonferroni显著性阈值 $P < 1.92 \times 10^{-3}$

图4

CHR:BP:Ref:Alt	基因	rsID	非酒精性肝病			非酒精性肝硬化			肝细胞癌		
			OR (95% CI)	P		OR (95% CI)	P		OR (95% CI)	P	
4:88231392:T:TA	HSD17B13	rs72613567	0.84 (0.78-0.91)	1.31E-05		0.74 (0.62-0.88)	4.48E-04		0.67 (0.45-1.00)	4.66E-02	
8:145730161:C:T	GPT	rs201815297	0.23 (0.04-1.14)	1.86E-02		1.25 (0.24-6.38)	7.98E-01		3.66 (0.70-19.02)	2.01E-01	
8:145732114:G:C	GPT	rs141505249	1.02 (0.49-2.11)	9.70E-01		0.36 (0.02-5.37)	3.82E-01		1.84 (0.15-23.25)	6.88E-01	
8:145732180:G:C	GPT	rs147998249	1.03 (0.49-2.17)	9.30E-01		0.34 (0.02-5.59)	3.67E-01		1.74 (0.11-27.05)	7.21E-01	
10:18242311:A:G	SLC39A12	rs10764176	0.92 (0.86-0.99)	3.43E-02		1.03 (0.88-1.21)	7.15E-01		1.29 (0.93-1.79)	1.37E-01	
10:101157378:CGT:C	GOT1		2.37 (0.61-9.27)	2.50E-01		8.27 (1.44-47.49)	5.92E-02		9.81 (0.52-183.54)	2.43E-01	
10:101165533:G:C	GOT1	rs374966349	1.63 (0.53-4.96)	4.20E-01		1.17 (0.07-20.09)	9.13E-01		5.37 (0.37-91.12)	3.55E-01	
10:101912064:T:C	ERUN1	rs2862954	0.98 (0.91-1.04)	4.61E-01		1.13 (0.98-1.31)	9.90E-02		0.94 (0.69-1.28)	6.94E-01	
14:94844947:C:T	SEPPHA1	rs28929474	1.50 (1.21-1.87)	5.29E-04		2.99 (2.11-4.24)	9.08E-08		1.86 (0.74-4.67)	2.40E-01	
19:19379549:C:T	TM6SF2	rs58542926	1.36 (1.21-1.52)	2.42E-07		1.64 (1.31-2.05)	6.04E-05		1.93 (1.22-3.04)	1.08E-02	
22:44324727:C:G	PNPLA3	rs738409	1.65 (1.54-1.78)	1.31E-41		2.05 (1.76-2.38)	1.70E-19		2.20 (1.60-3.02)	5.59E-06	
22:44324730:C:T	PNPLA3	rs738408	1.65 (1.54-1.78)	1.42E-41		2.05 (1.77-2.38)	1.45E-19		2.20 (1.60-3.03)	5.41E-06	
22:44368122:A:G	SAMM50	rs3761472	1.52 (1.41-1.65)	7.33E-24		1.86 (1.58-2.19)	1.81E-12		1.66 (1.16-2.39)	1.05E-02	

灰色阴影指示P值满足Bonferroni显著性阈值 $P < 1.92 \times 10^{-3}$

图4(续)

特征	达拉斯肝脏 研究病例 (N = 517)	达拉斯肝脏 研究对照 (N = 4,279)	达拉斯儿科肝脏 研究病例 (N = 203)	达拉斯儿科肝脏 研究对照 (N = 244)
年龄{岁}~中值{IQR}	35 {48 - 60}	44 {36 - 53}	12 {10 - 15}	12 {11 - 14}
女性~数量{}	277 {54}	2,494 {58}	65 {32}	126 {52}
体重指数~中值{IQR}	30 {27 - 35}	30 {26 - 35}	30 {27 - 34}	31 {28 - 35}
自我报告的种族				
非洲裔美国人	33 {6}	2,291 {54}	-	-
欧洲裔美国人	158 {31}	1,266 {30}	-	-
西班牙裔美国人	326 {63}	722 {17}	203 {100}	244 {100}
肝病的存在~是{}				
酒精性肝病	223 {43}	-	-	-
酒精性肝硬化	215 {42}	-	-	-
非酒精性{非病毒性}肝病	212 {20}	-	-	-
非酒精性肝硬化	100 {19}	-	-	-
肝细胞癌	44 {9}	-	-	-
无肝病	-	4,279 {100}	-	244 {100}

图5

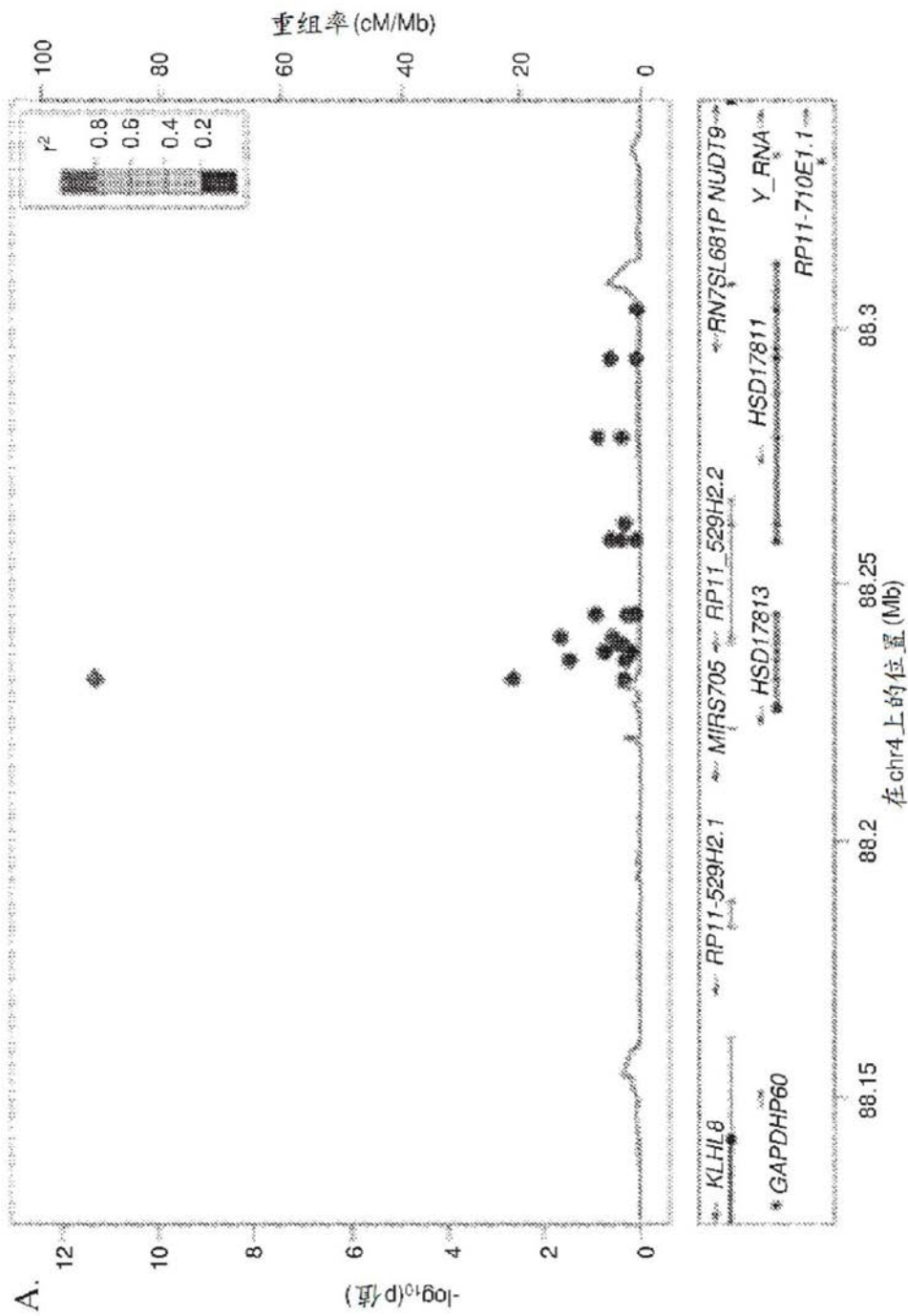


图6

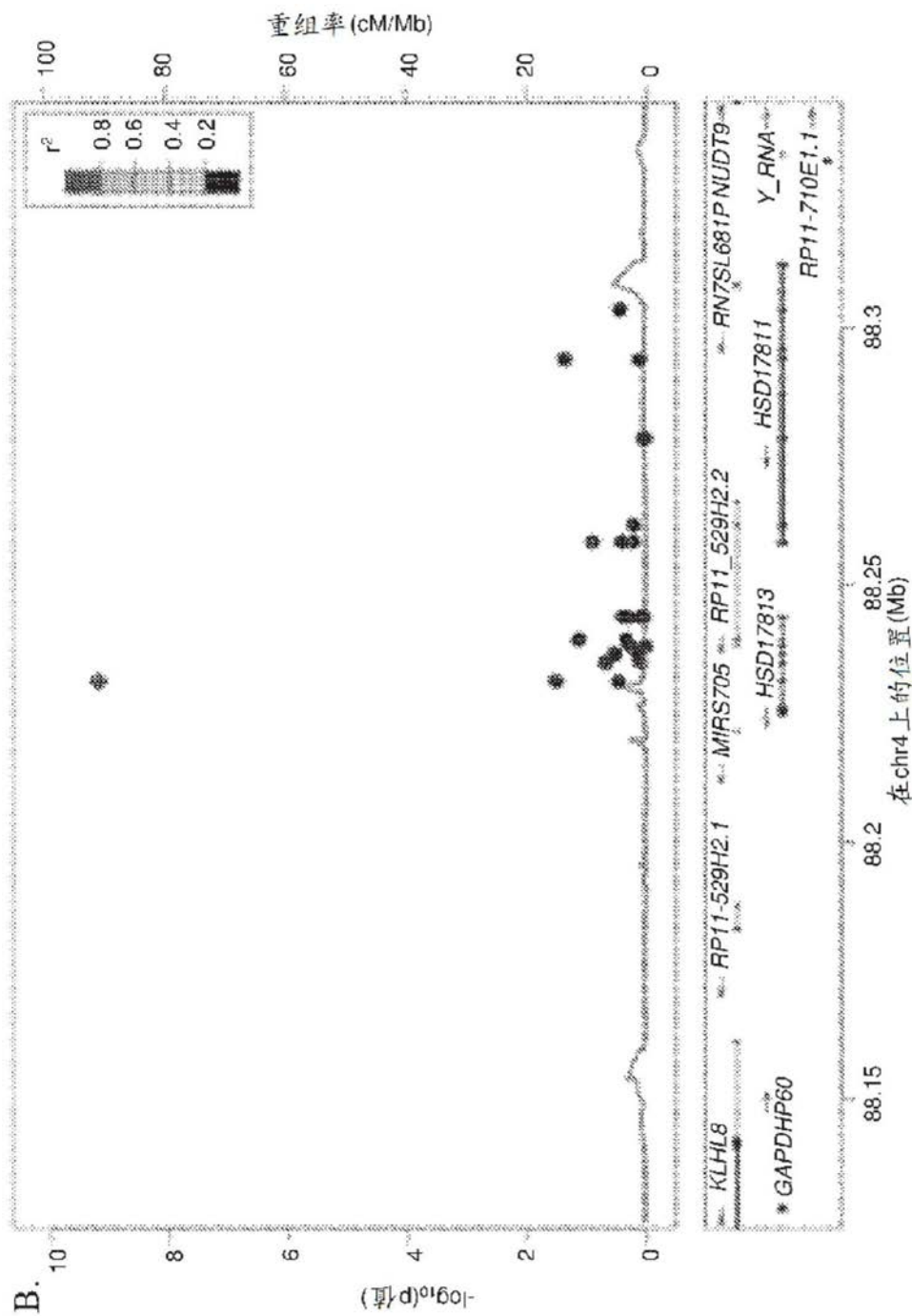


图6(续)

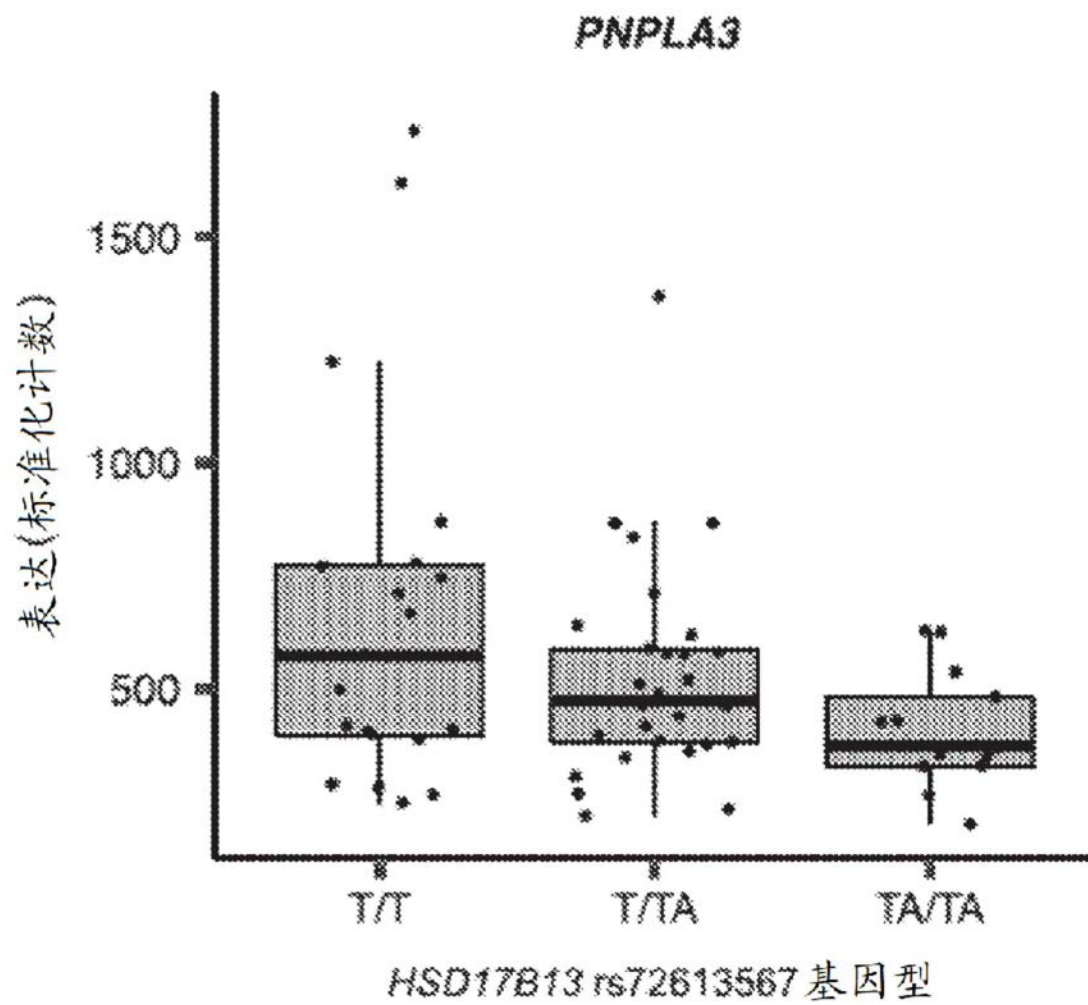


图7

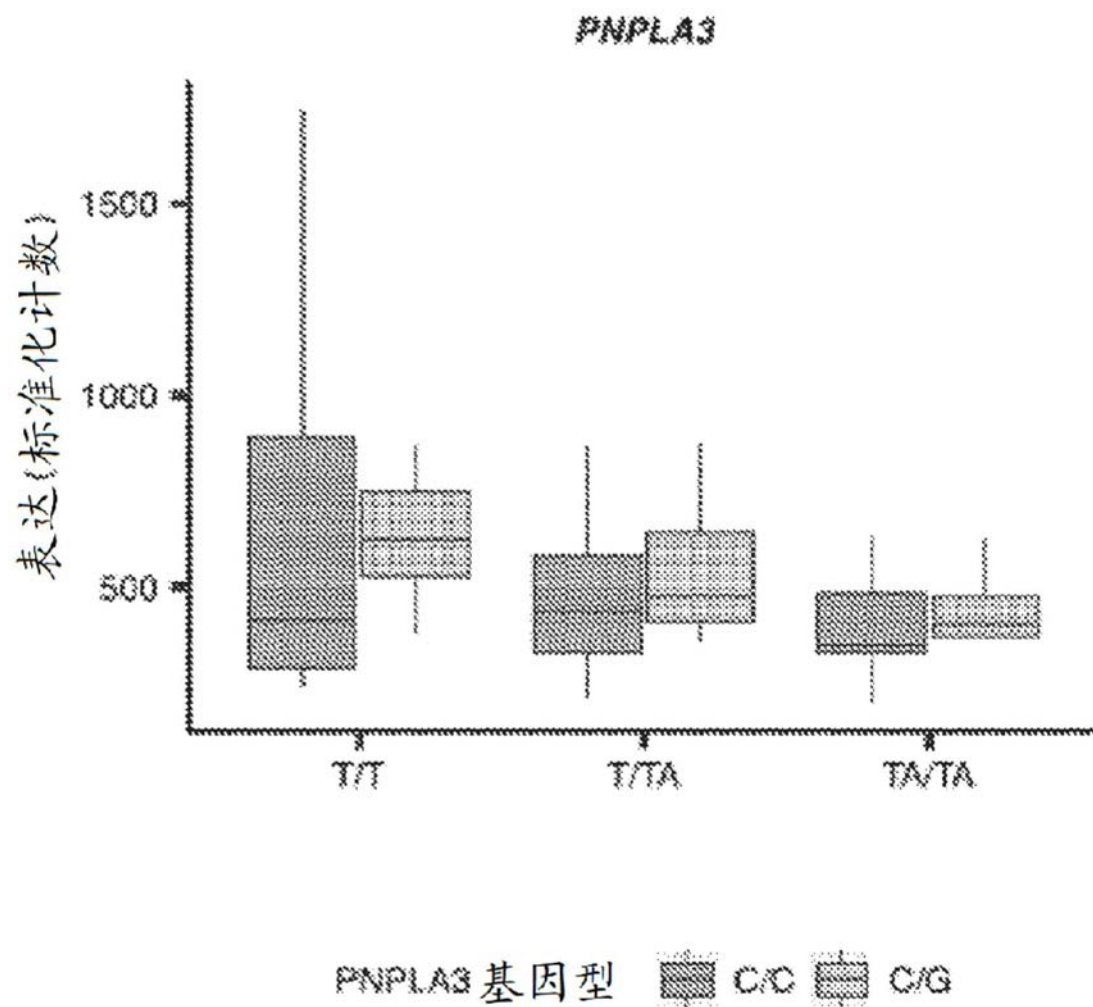


图8

表型/子集	N	HSD17B13 rs72613567 : [相互作用效应]	HSD17B13 rs72613567 [主要效应]	PNPLA3 rs738409 [主要效应]
log₁₀ (AIF)				
子集	N	效应(95% CI)	效应(95% CI)	效应(95% CI)
全部	43,309	1.80E-03 {-0.011, 0.002}	1.88E-04 {-0.006, -0.009, -0.003}	1.88E-04 {-0.022, 0.029}
肥胖	23,051	1.01E-03 {-0.015, -0.004}	6.53E-05 {-0.013, -0.005}	3.13E-55 {-0.032, 0.042}
非肥胖	20,258	1.49E-03 {-0.004, -0.01, 0.002}	3.53E-01 {-0.002, -0.007, 0.002}	1.56E-07 {-0.008, 0.018}
log₁₀ (AIF)				
子集	N	效应(95% CI)	效应(95% CI)	效应(95% CI)
全部	42,662	4.53E-03 {-0.004, -0.007, -0.001}	4.78E-04 {-0.006, -0.002}	9.69E-46 {-0.014, 0.018}
肥胖	22,719	1.04E-03 {-0.006, -0.01, -0.003}	1.42E-04 {-0.009, -0.003}	7.65E-49 {-0.02, 0.027}
非肥胖	19,943	4.97E-03 {-0.001, -0.005, 0.003}	3.42E-01 {-0.004, 0.002}	1.56E-06 {-0.005, 0.011}
非酒精性肝病				
子集	N对照	OR (95% CI)	OR (95% CI)	OR (95% CI)
全部	29,928	1.857 {0.919, 0.812, 1.039}	2.40E-02 {0.88, 0.787, 0.983}	1.76E-32 {1.764, 1.606, 1.938}
肥胖	14,243	1.445 {0.906, 0.786, 1.044}	7.81E-02 {0.894, 0.788, 1.012}	2.06E-72 {1.714, 1.537, 1.91}
非肥胖	15,685	7.71E-01 {0.964, 0.75, 1.235}	1.67E-01 {0.845, 0.662, 1.069}	1.96E-11 {1.887, 1.566, 2.269}
酒精性肝病				
子集	N对照	OR (95% CI)	OR (95% CI)	OR (95% CI)
全部	23,928	1.90 {1.112, 0.749, 1.637}	4.55E-03 {0.578, 0.391, 0.834}	7.80E-05 {1.689, 1.298, 2.185}
肥胖	14,243	97 {1.295, 0.741, 2.224}	1.22E-02 {0.501, 0.283, 0.839}	2.30E-02 {1.533, 1.052, 2.201}
非肥胖	15,685	93 {0.956, 0.544, 1.65}	1.39E-01 {0.666, 0.382, 1.105}	1.00E-03 {1.853, 1.275, 2.664}

图9

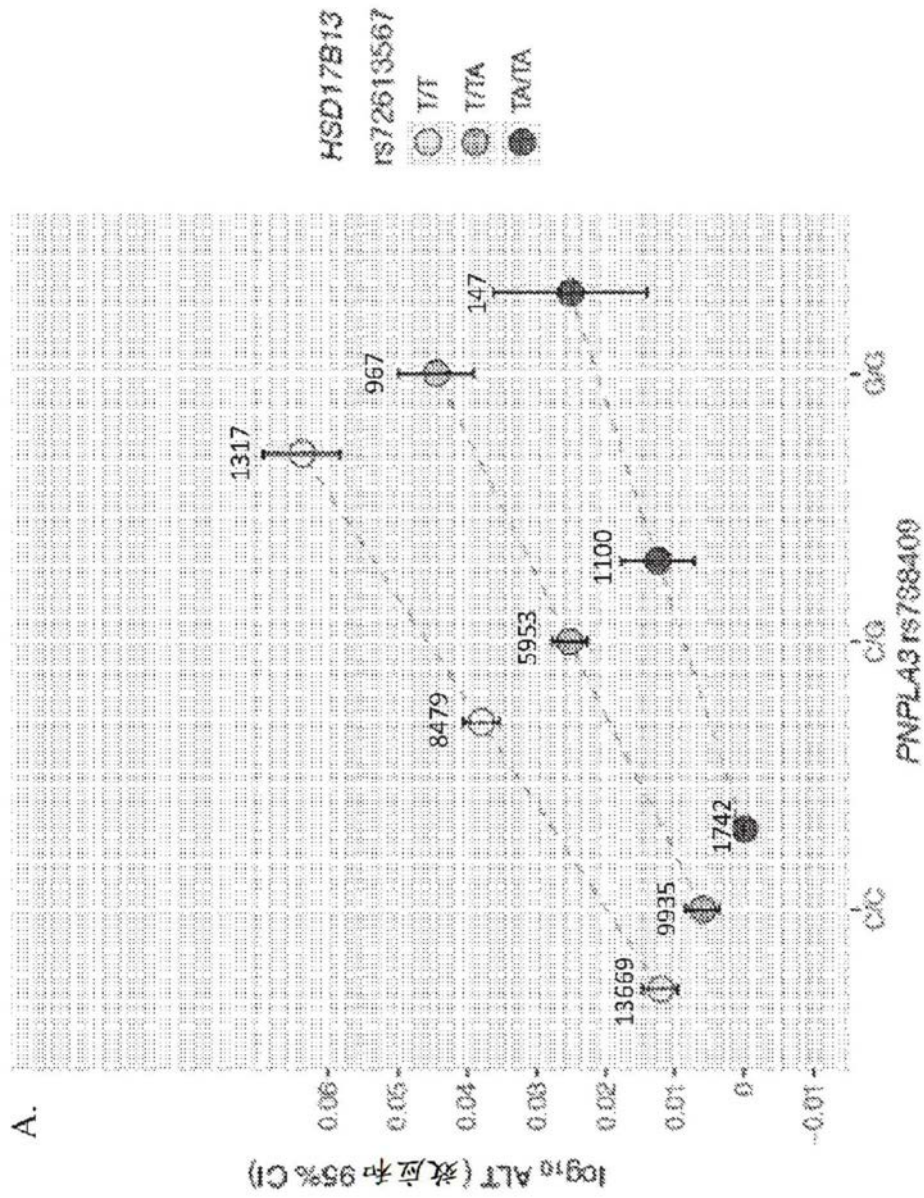


图10

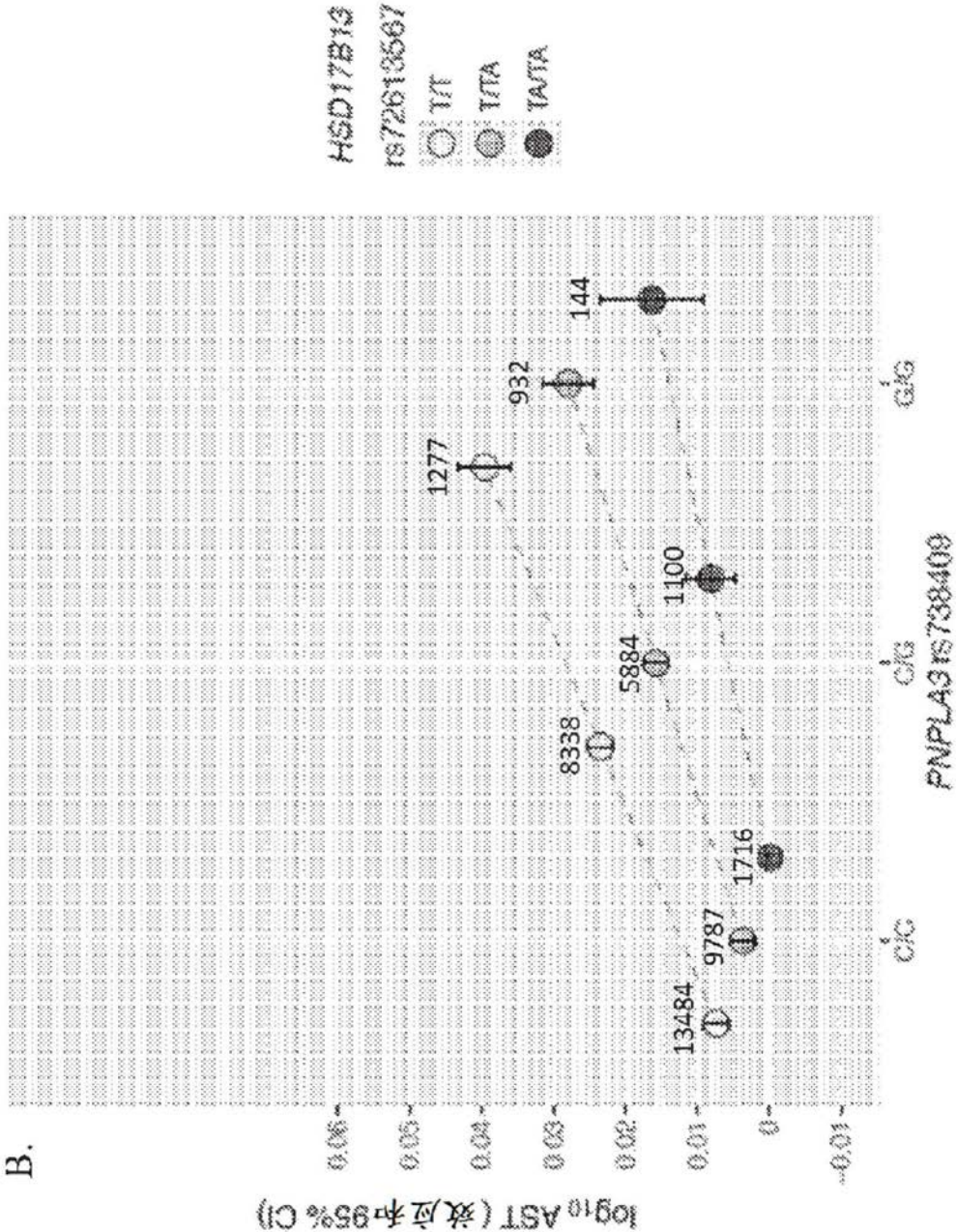


图10(续)

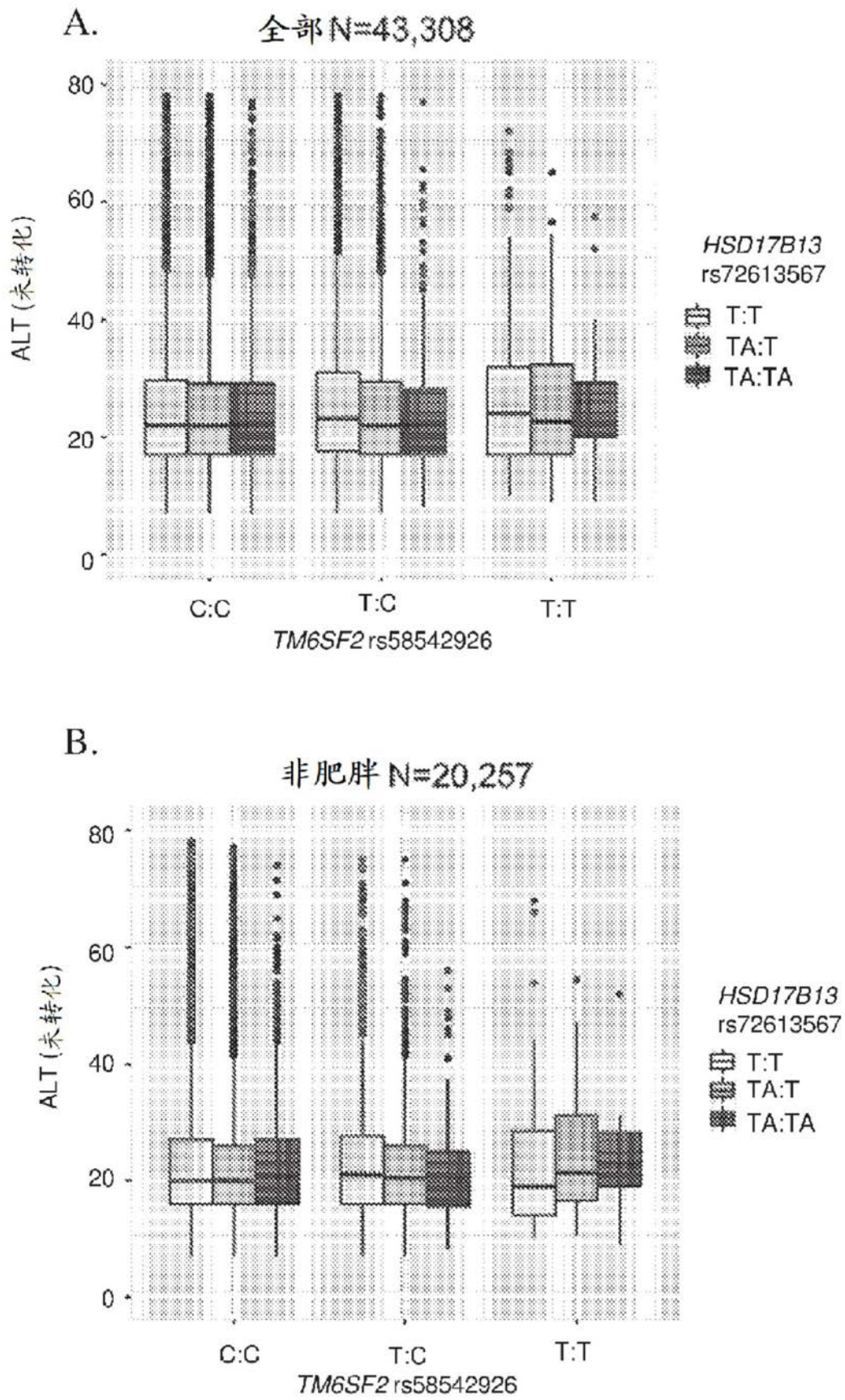


图11

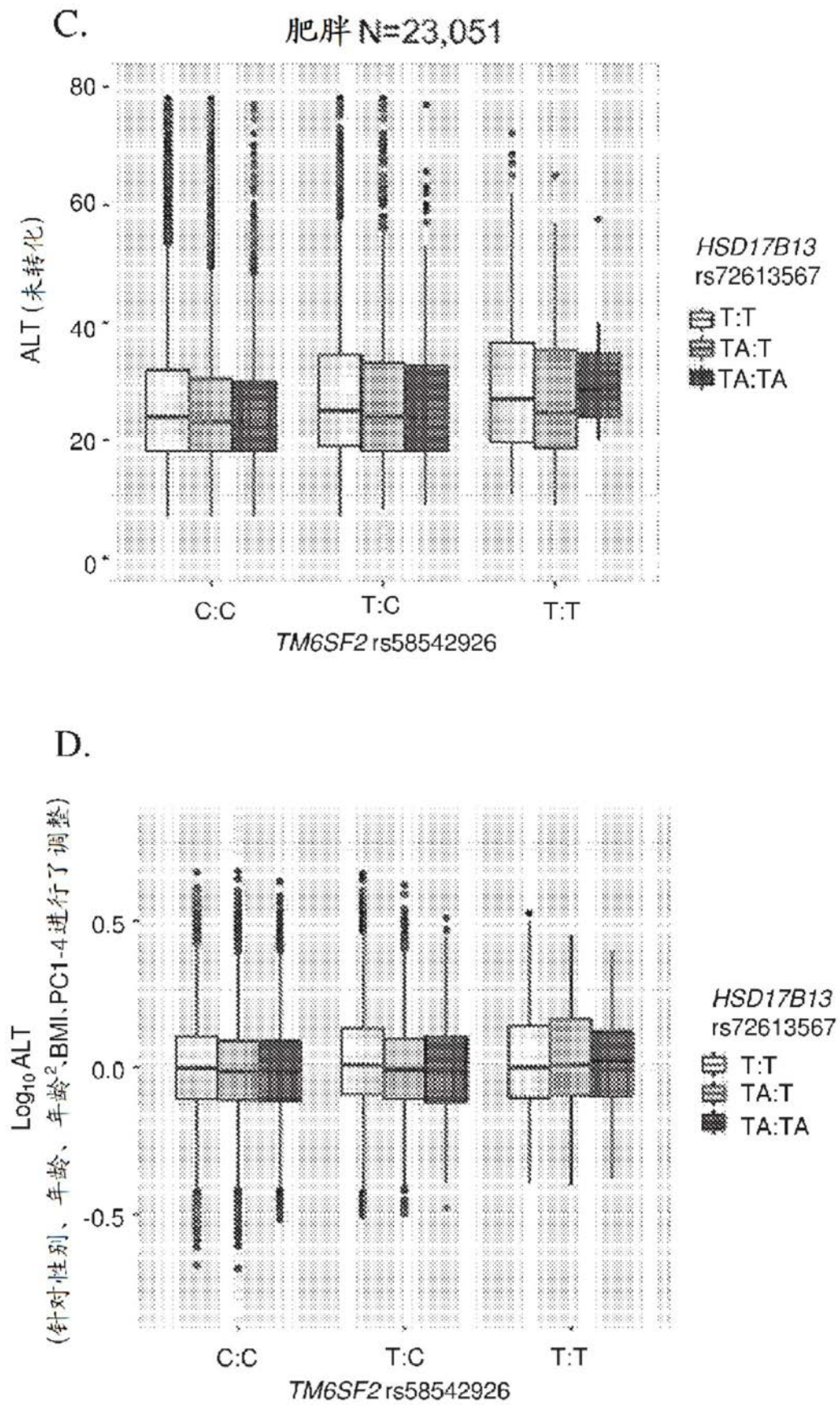


图11 (续)

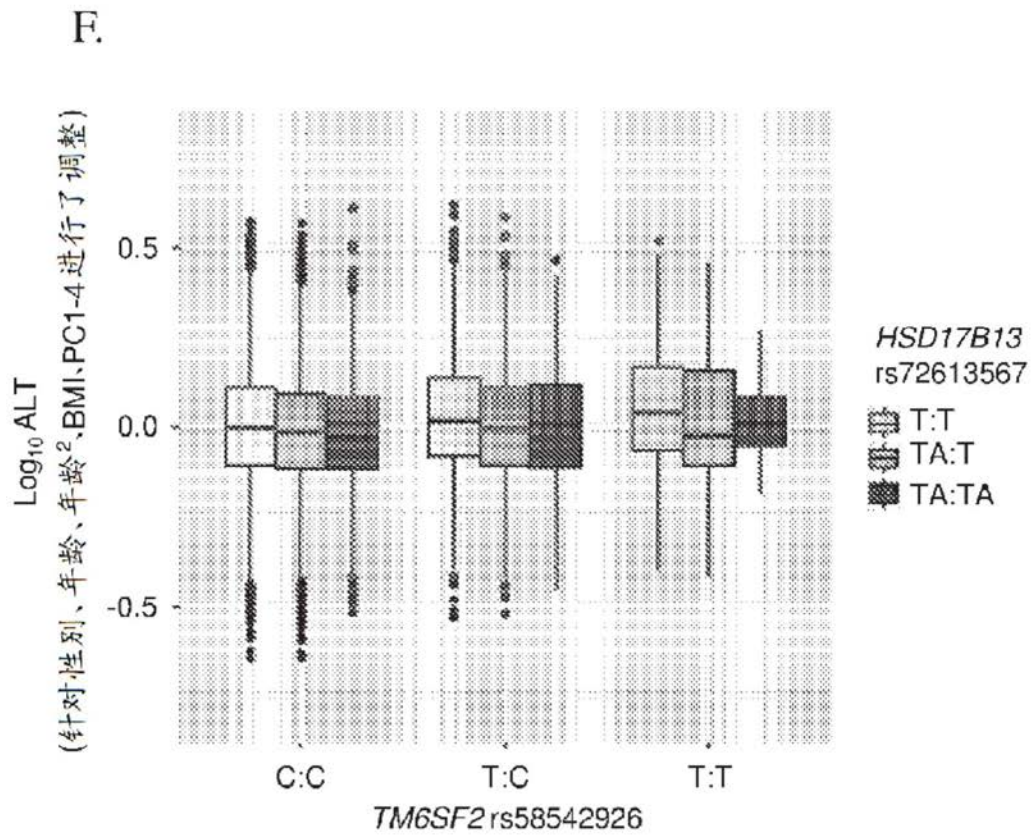
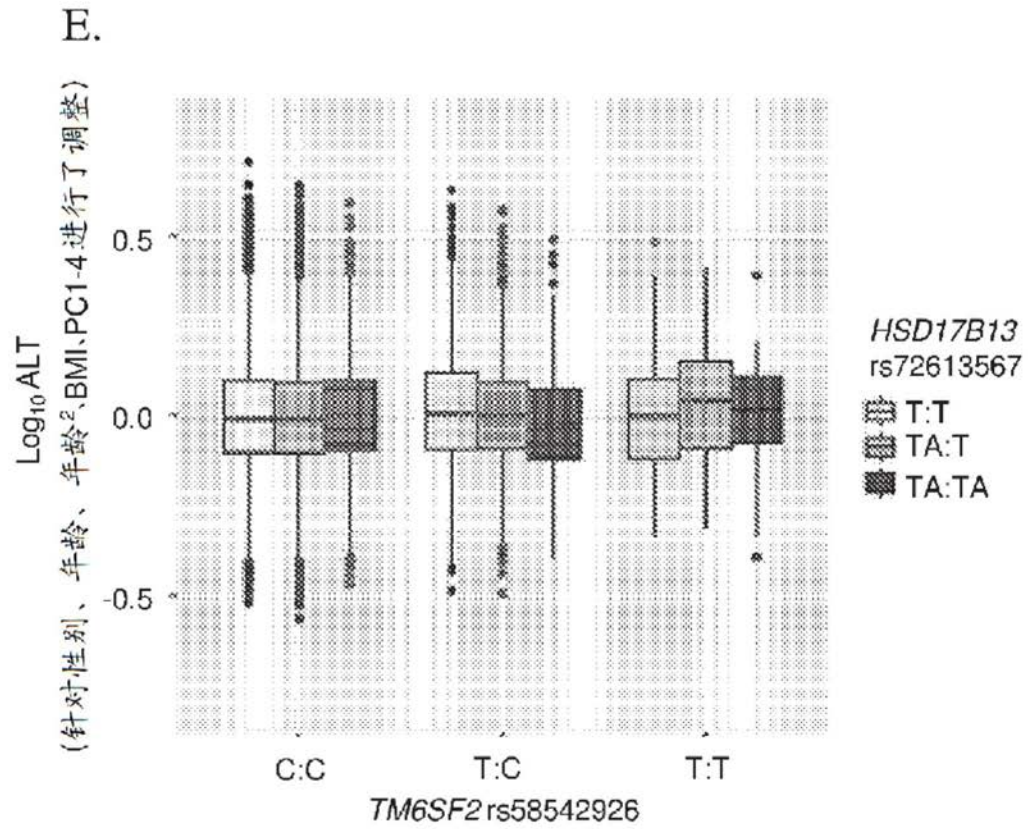


图11 (续)

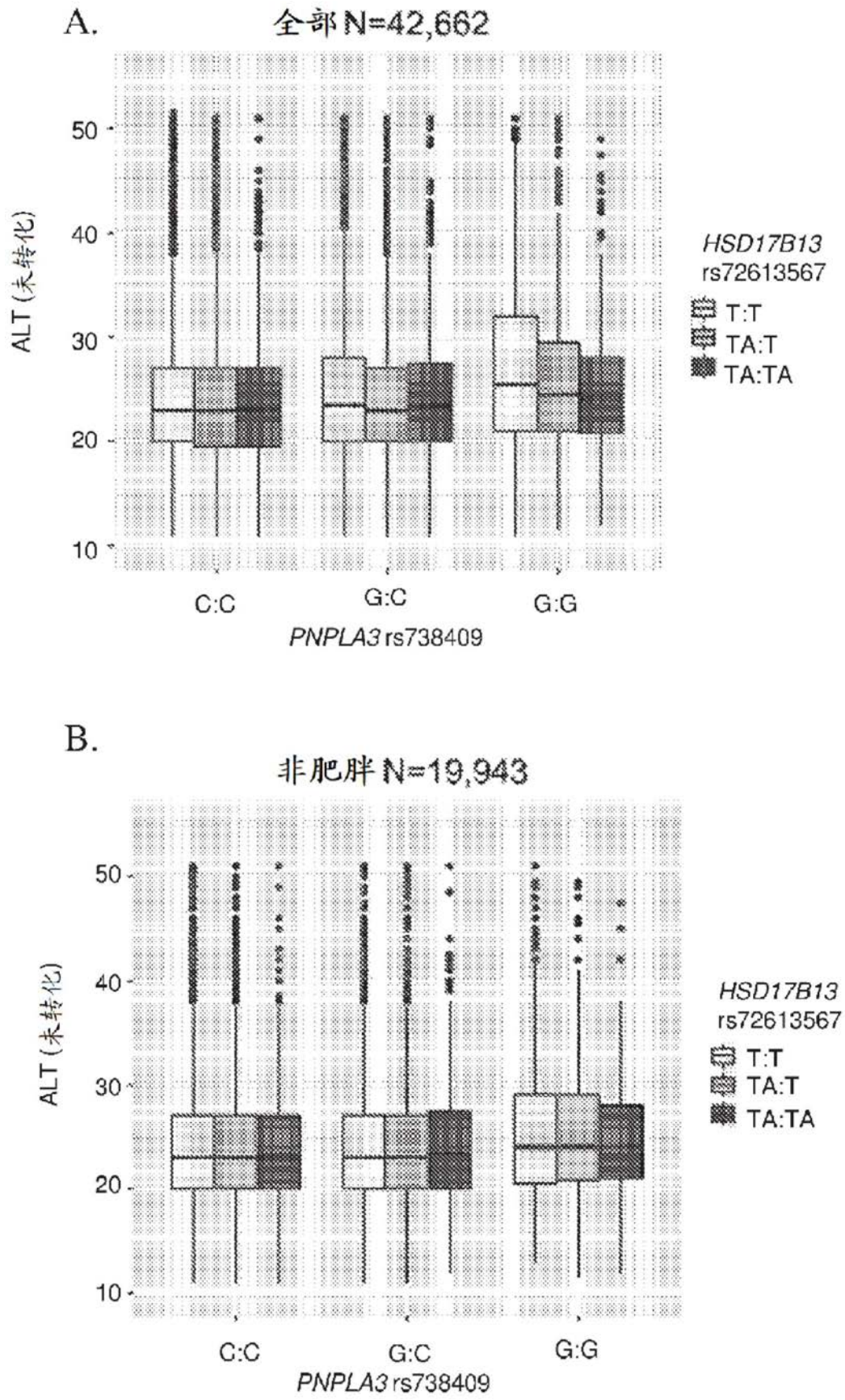


图12

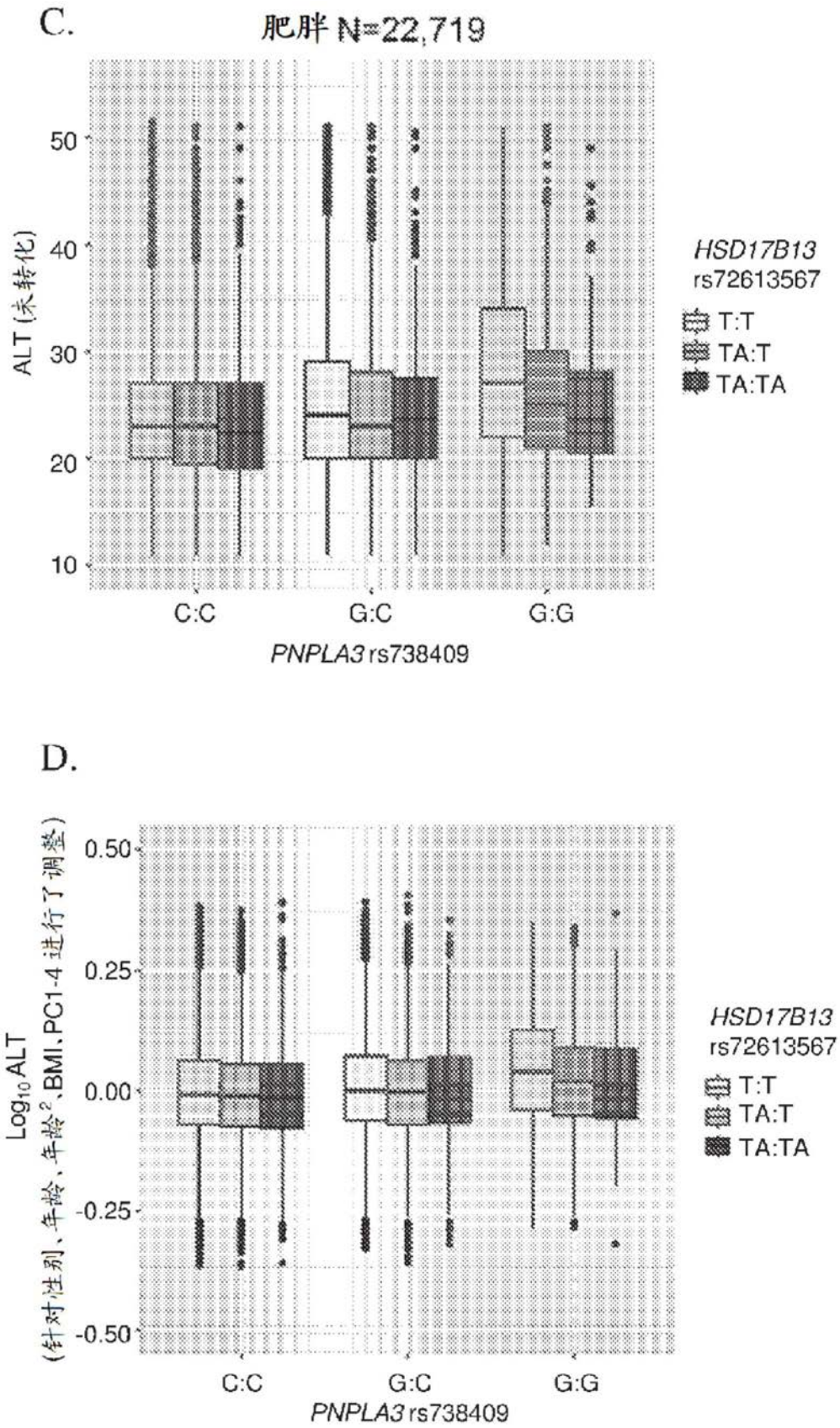


图12(续)

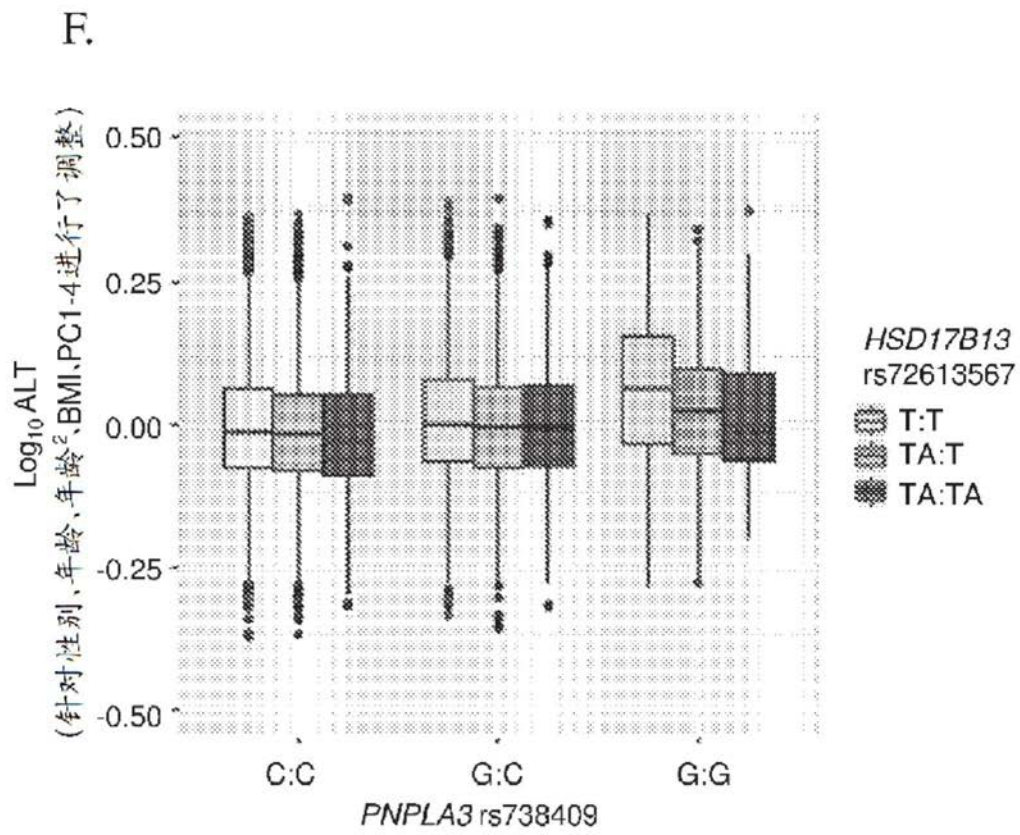
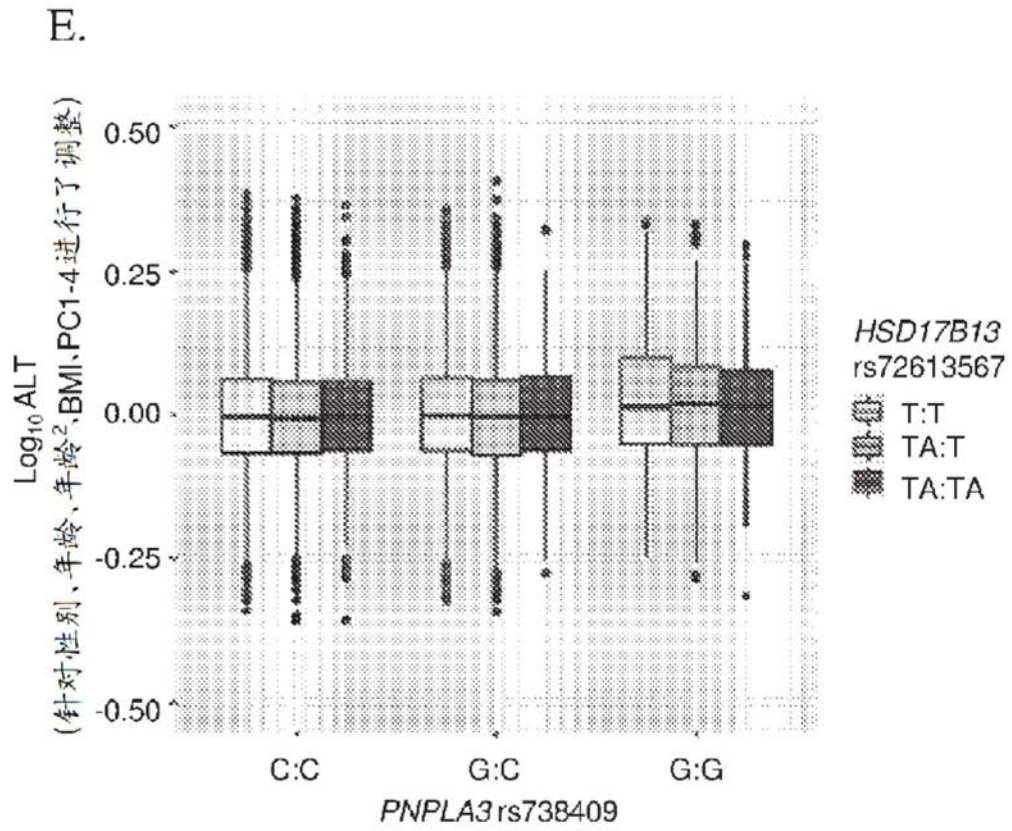


图12(续)

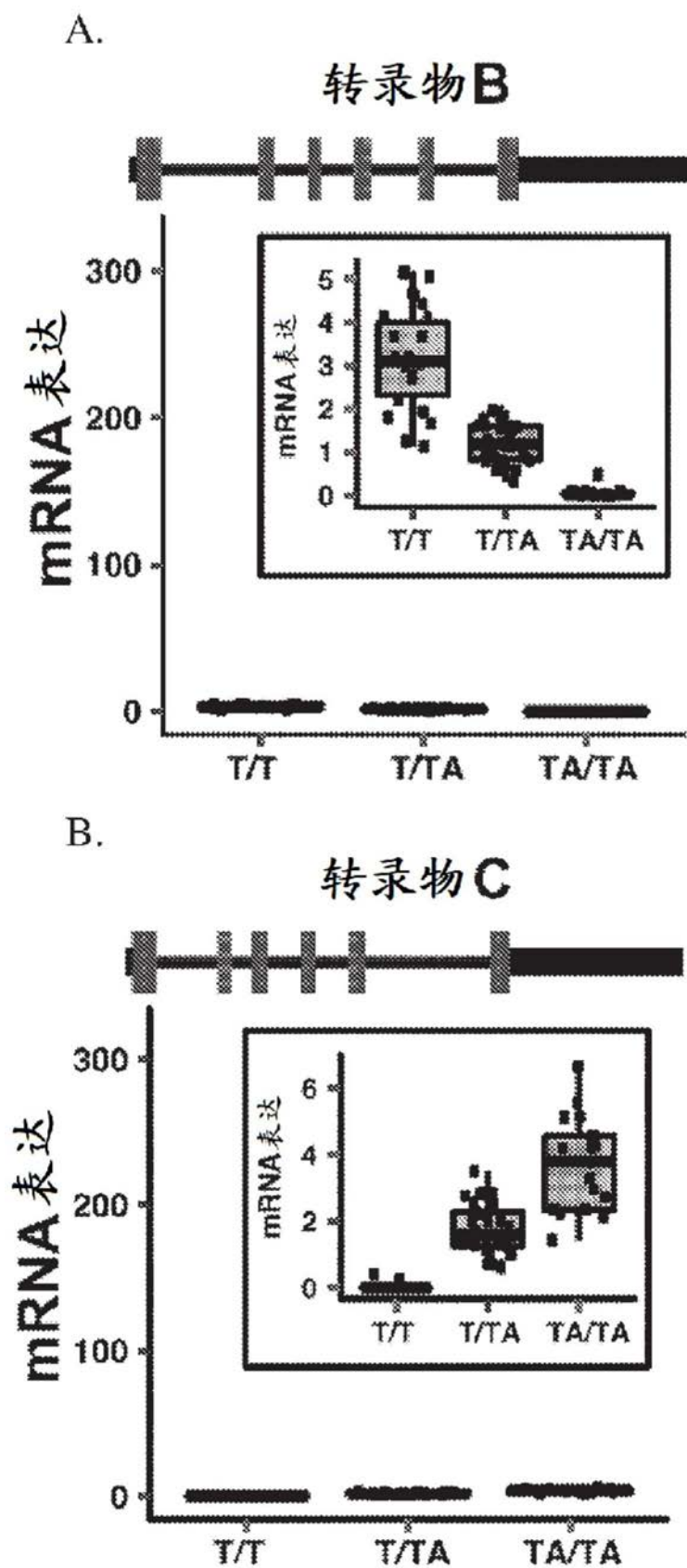


图13

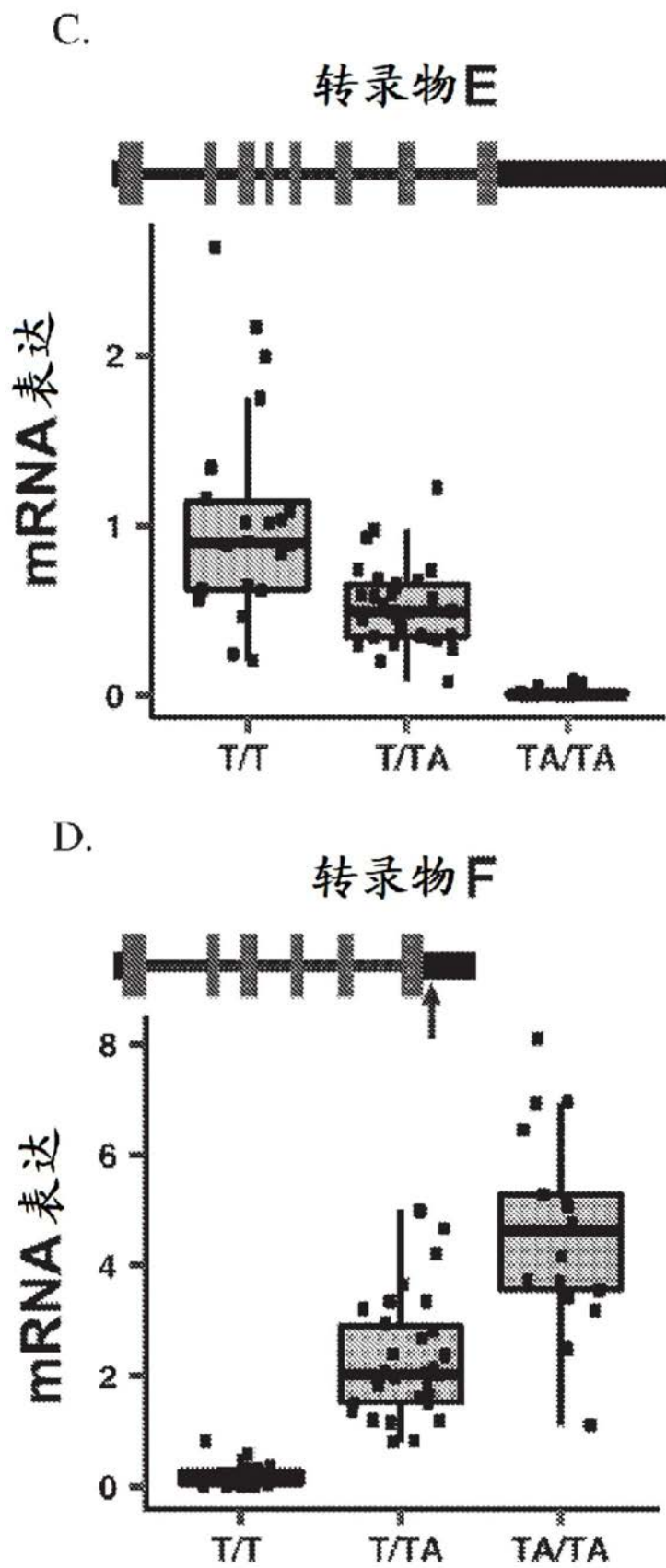


图13(续)

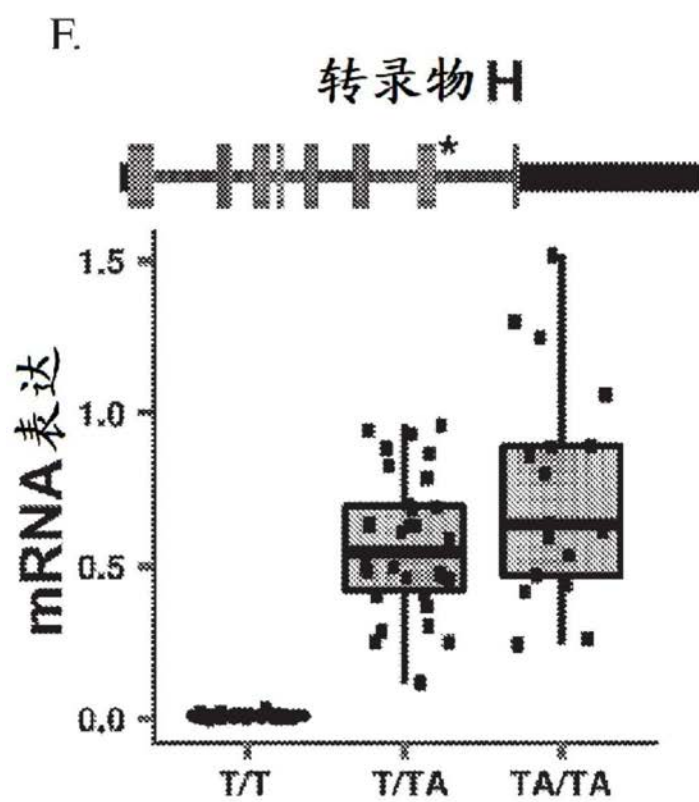
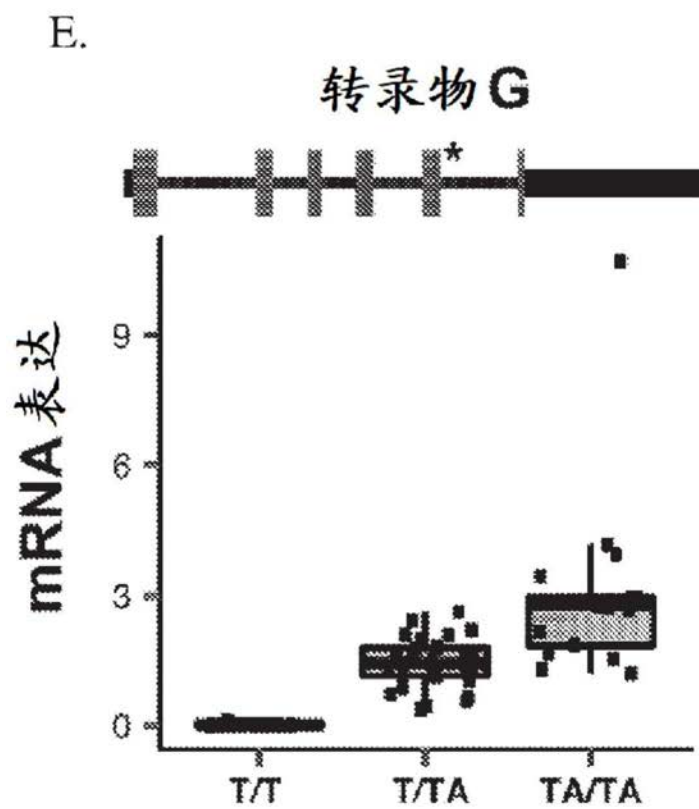


图13(续)

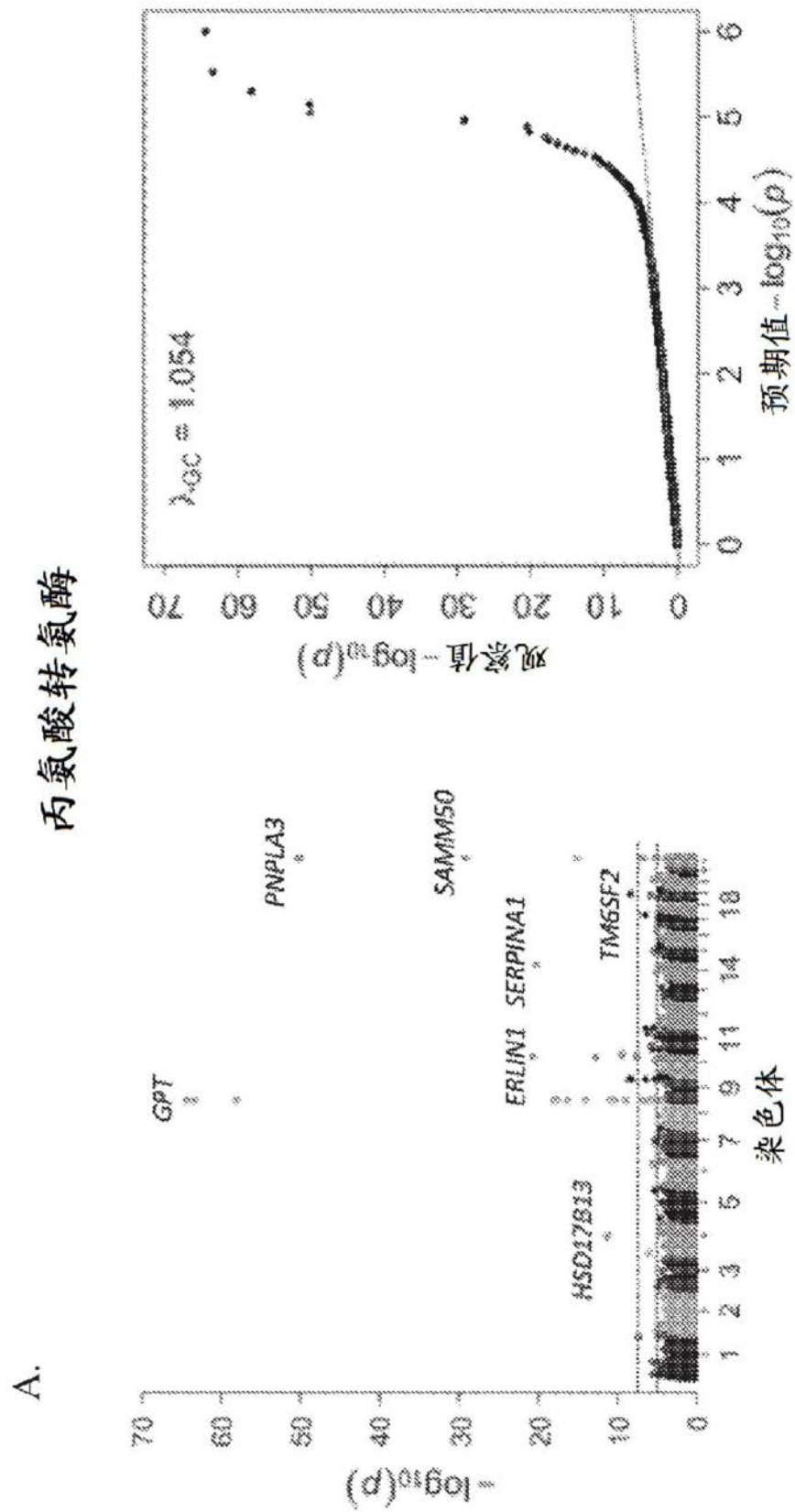


图14

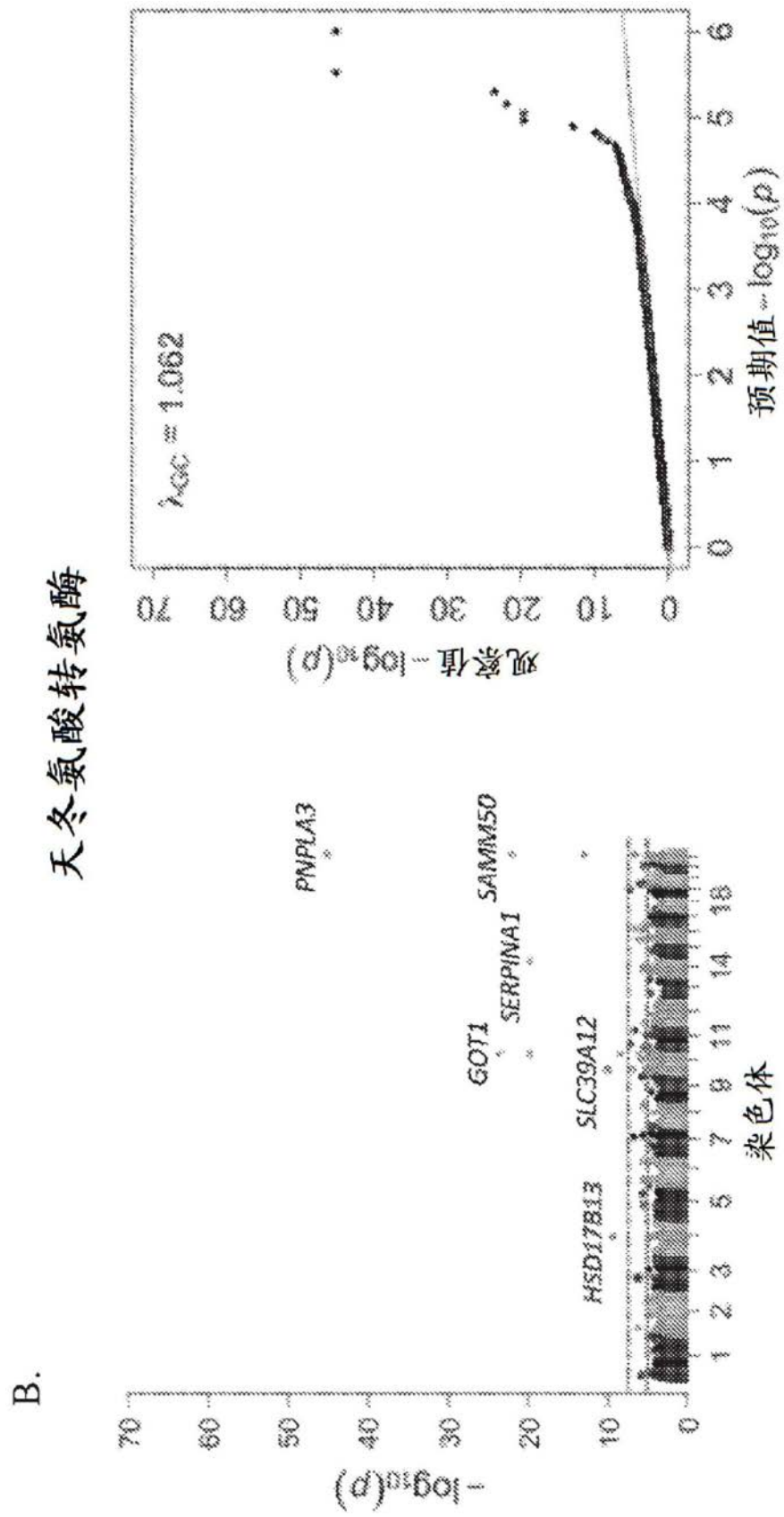


图14 (续)

A.

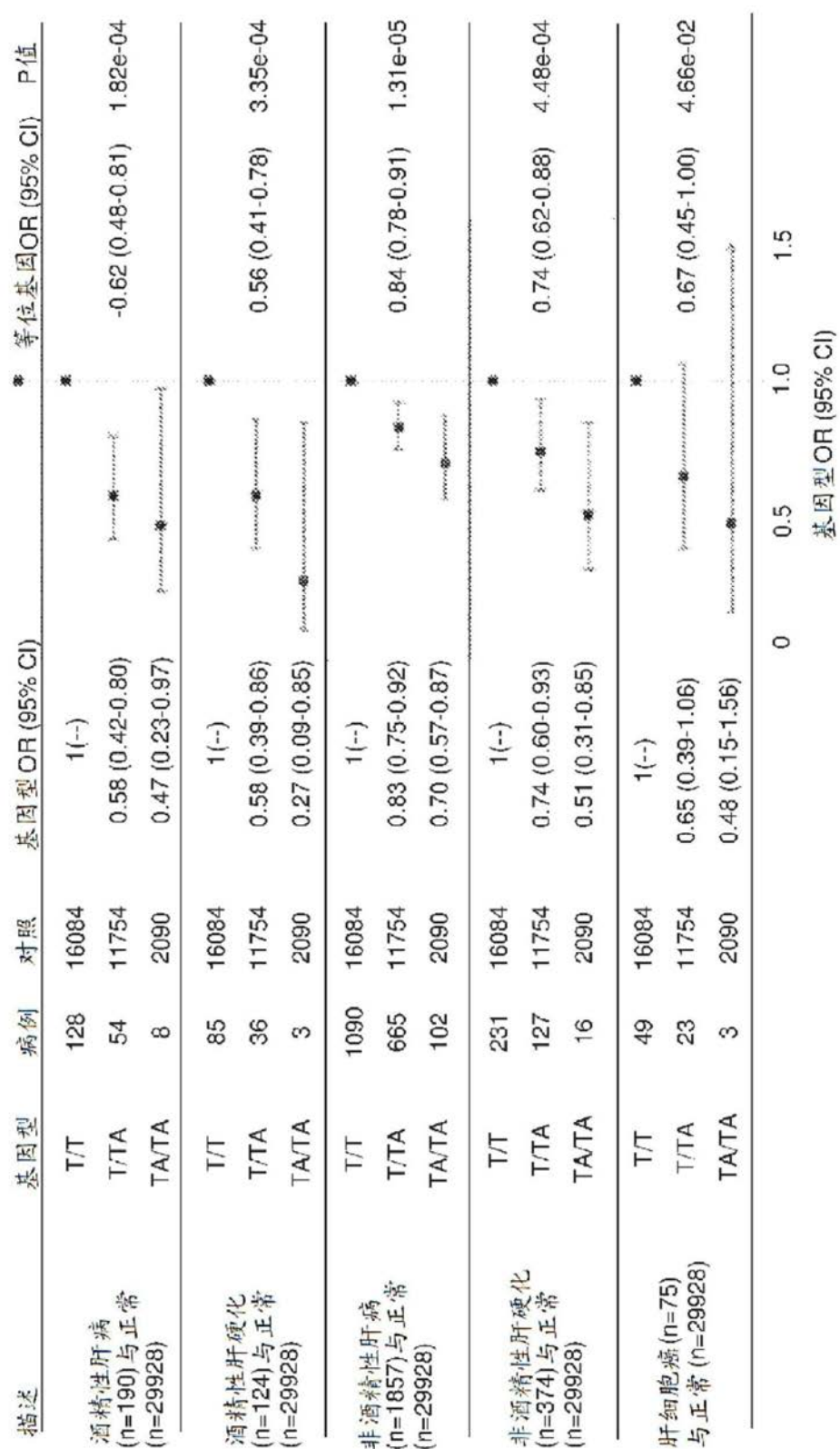


图15

B.

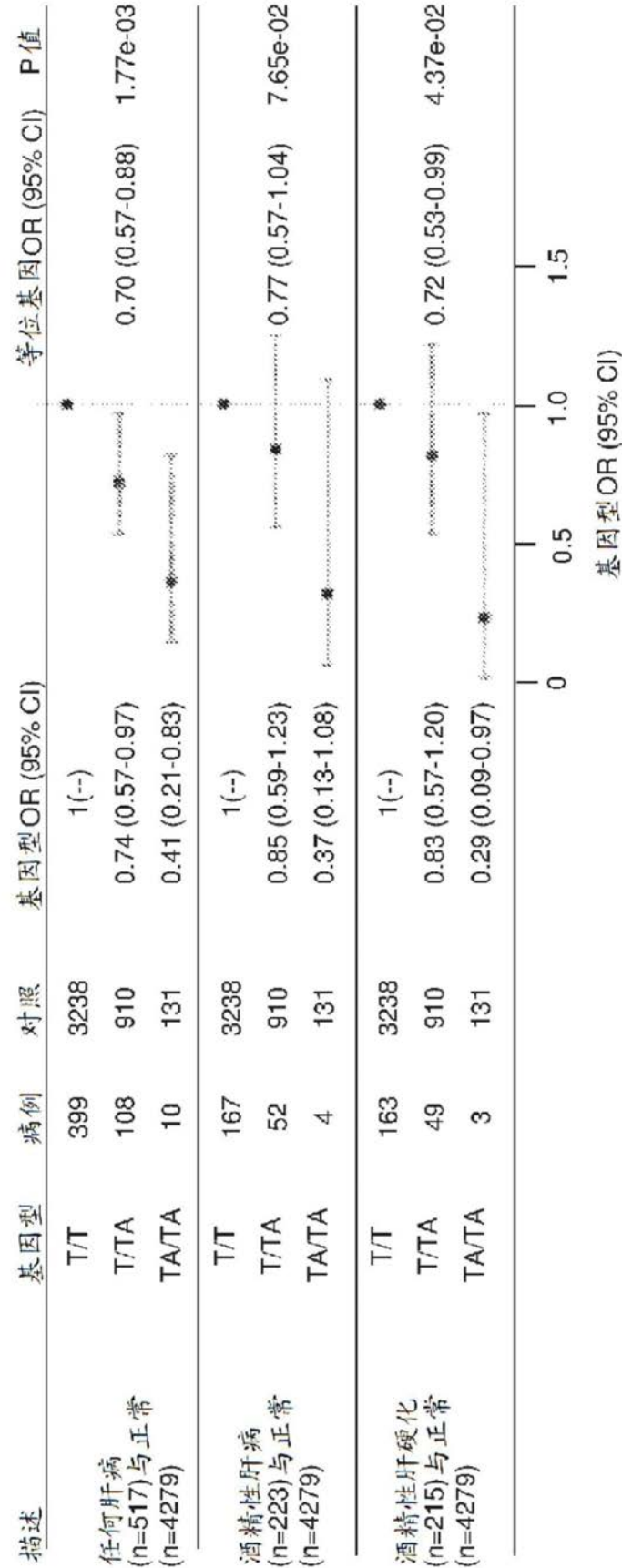


图15(续)

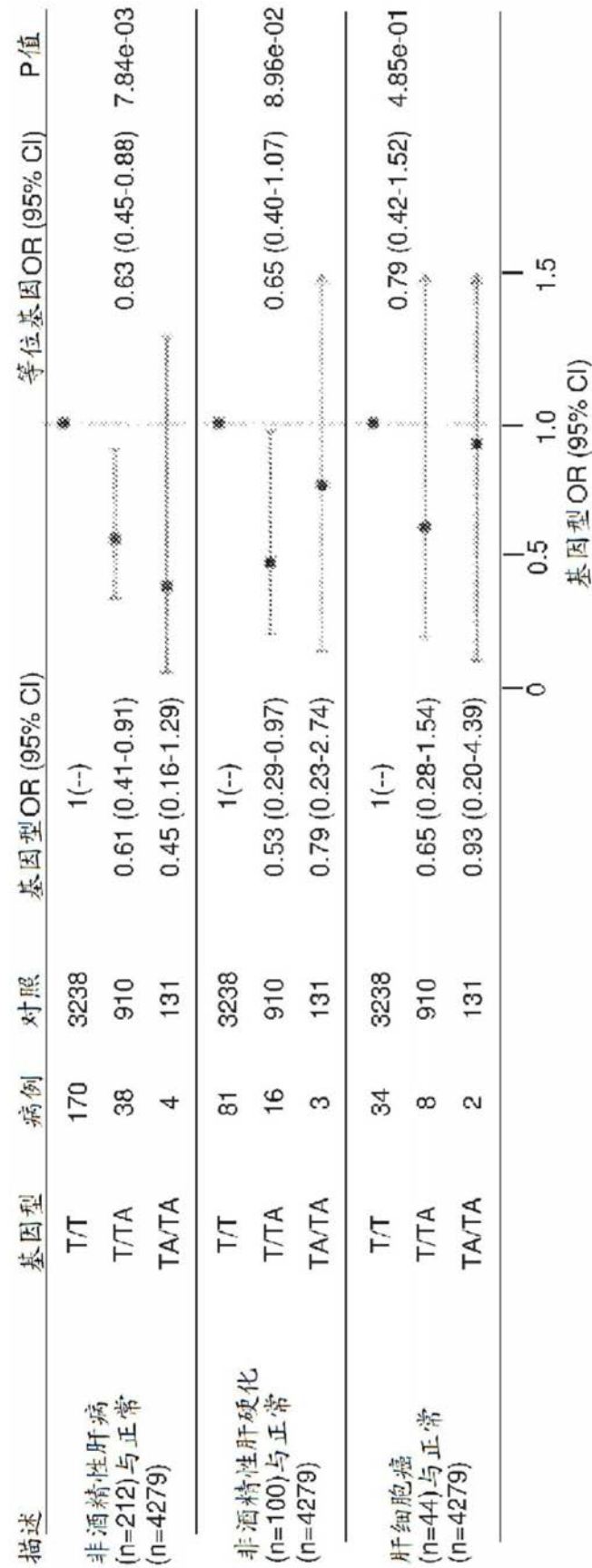


图15(续)

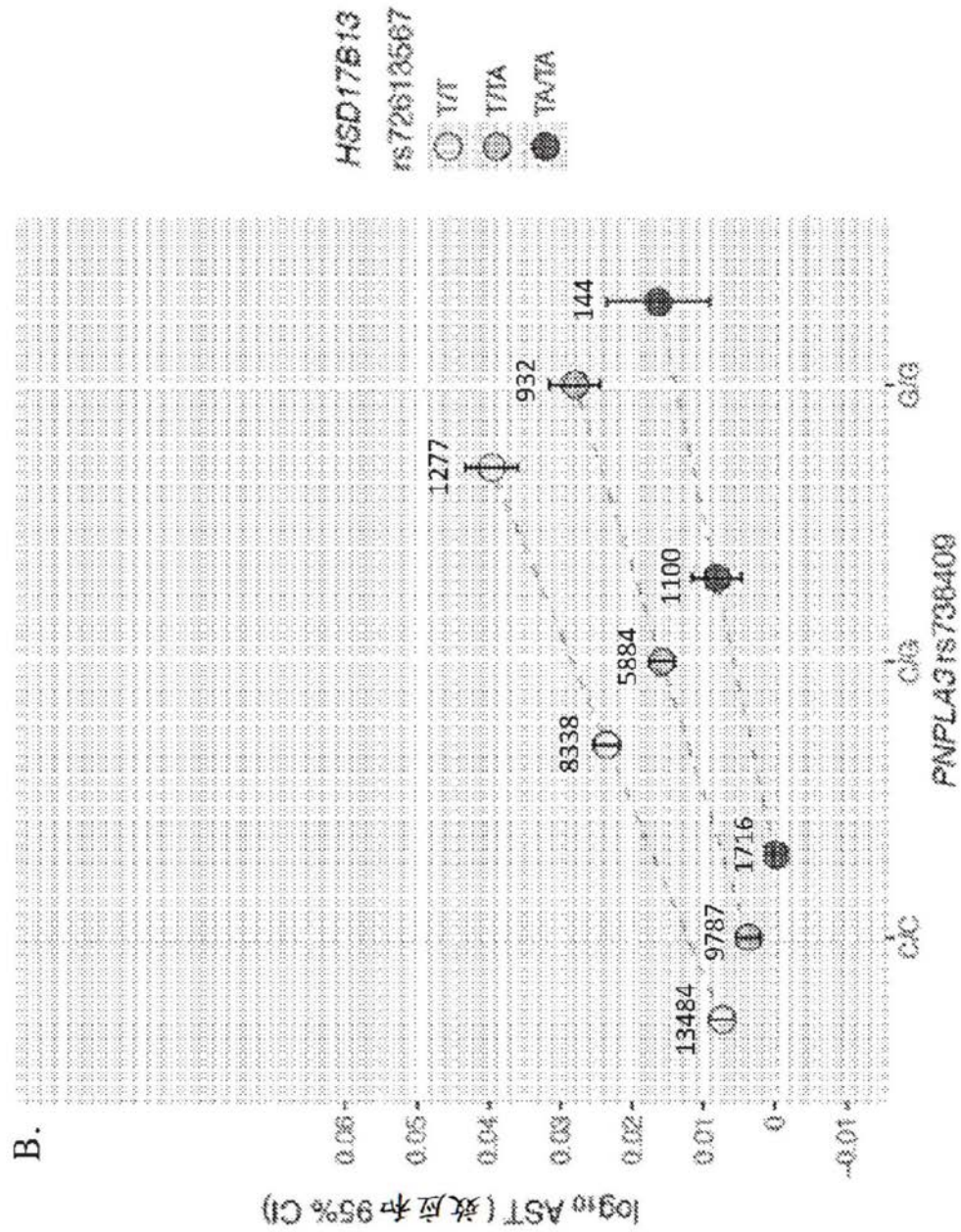


图16(续)

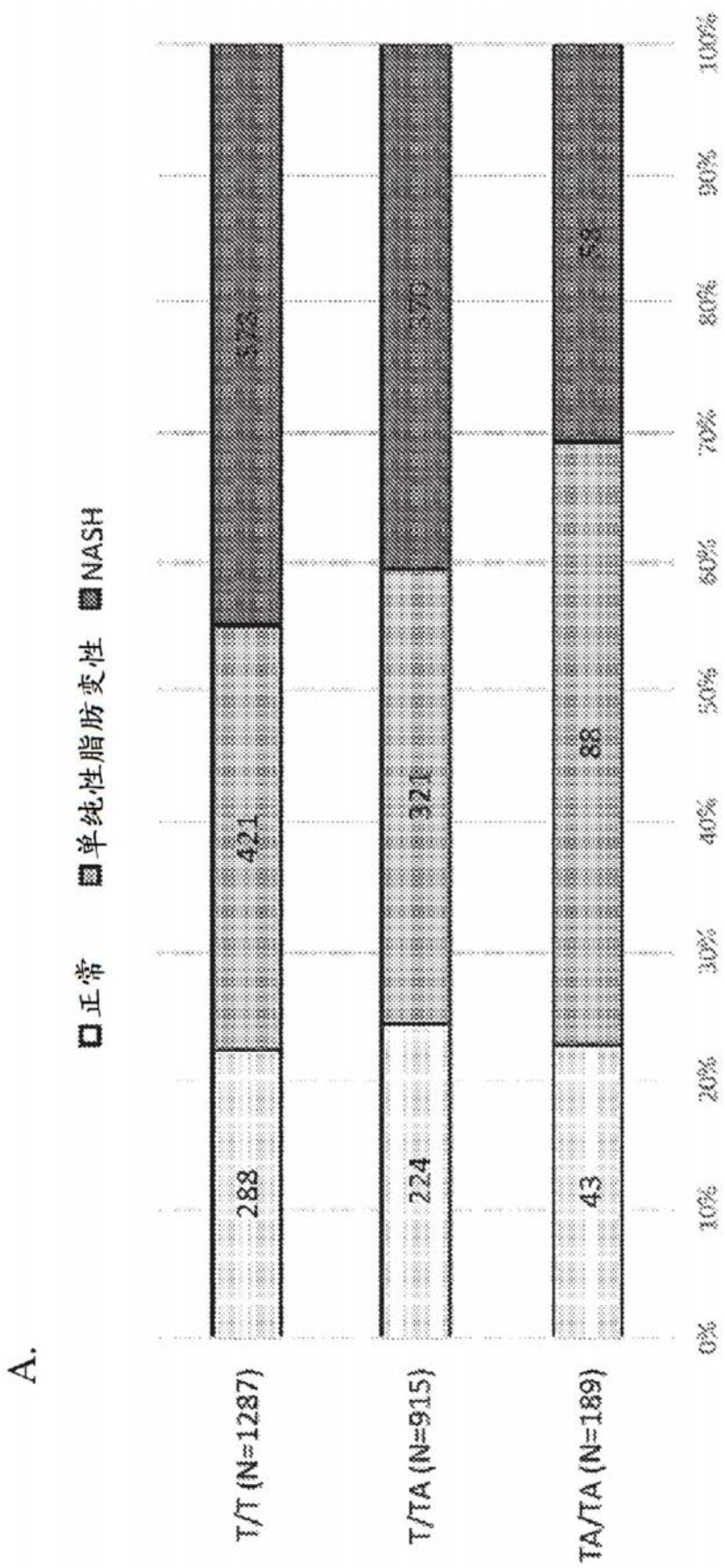


图17

B.

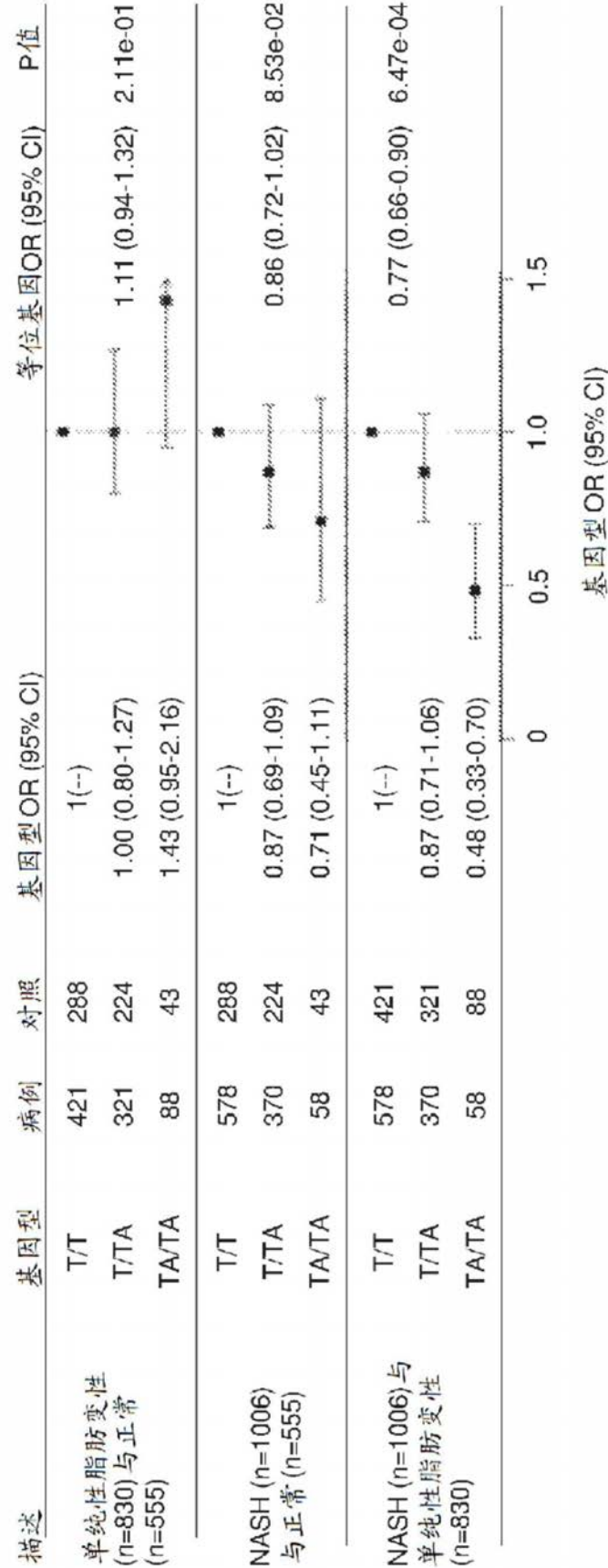


图17 (续)

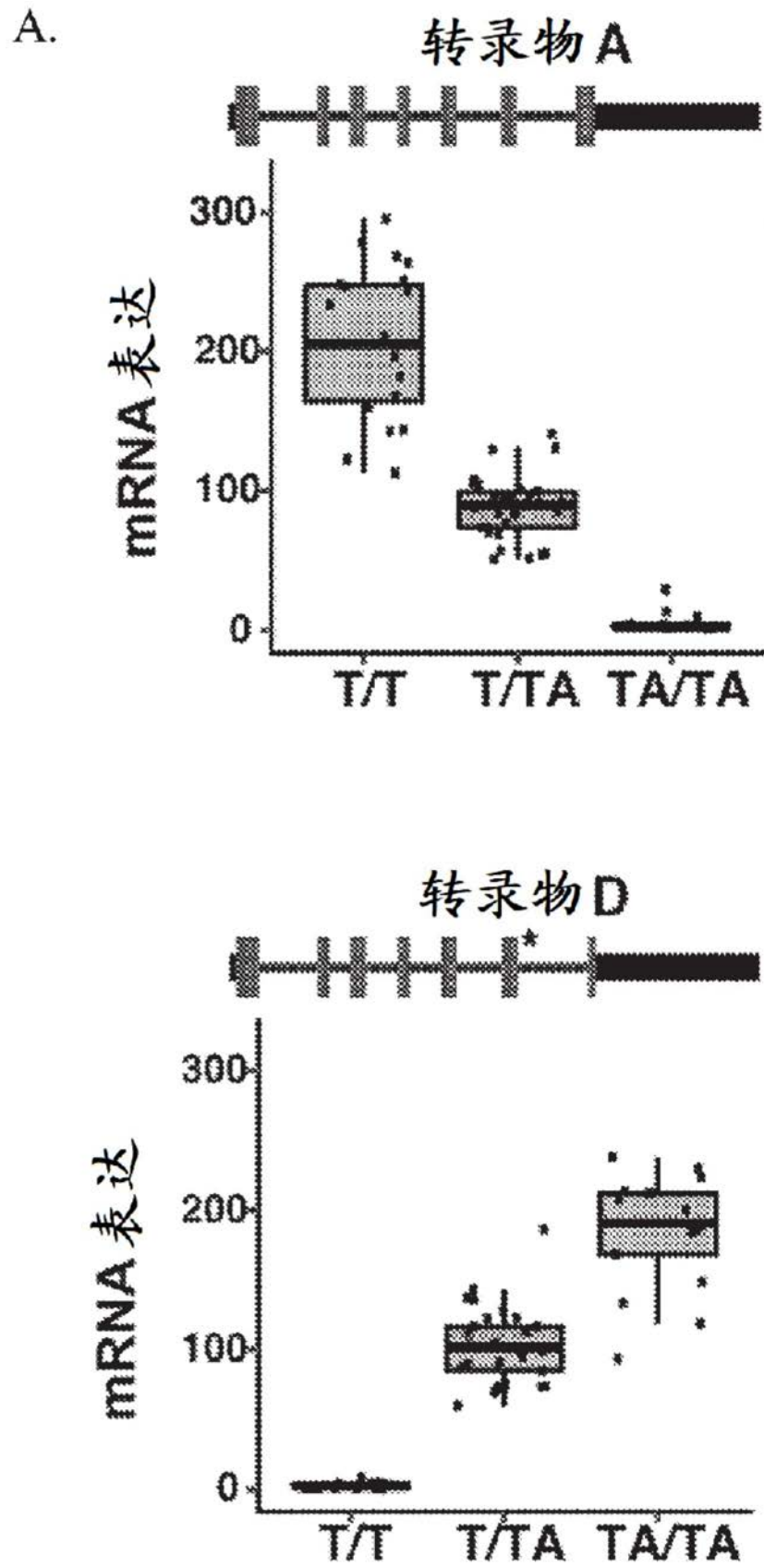


图18

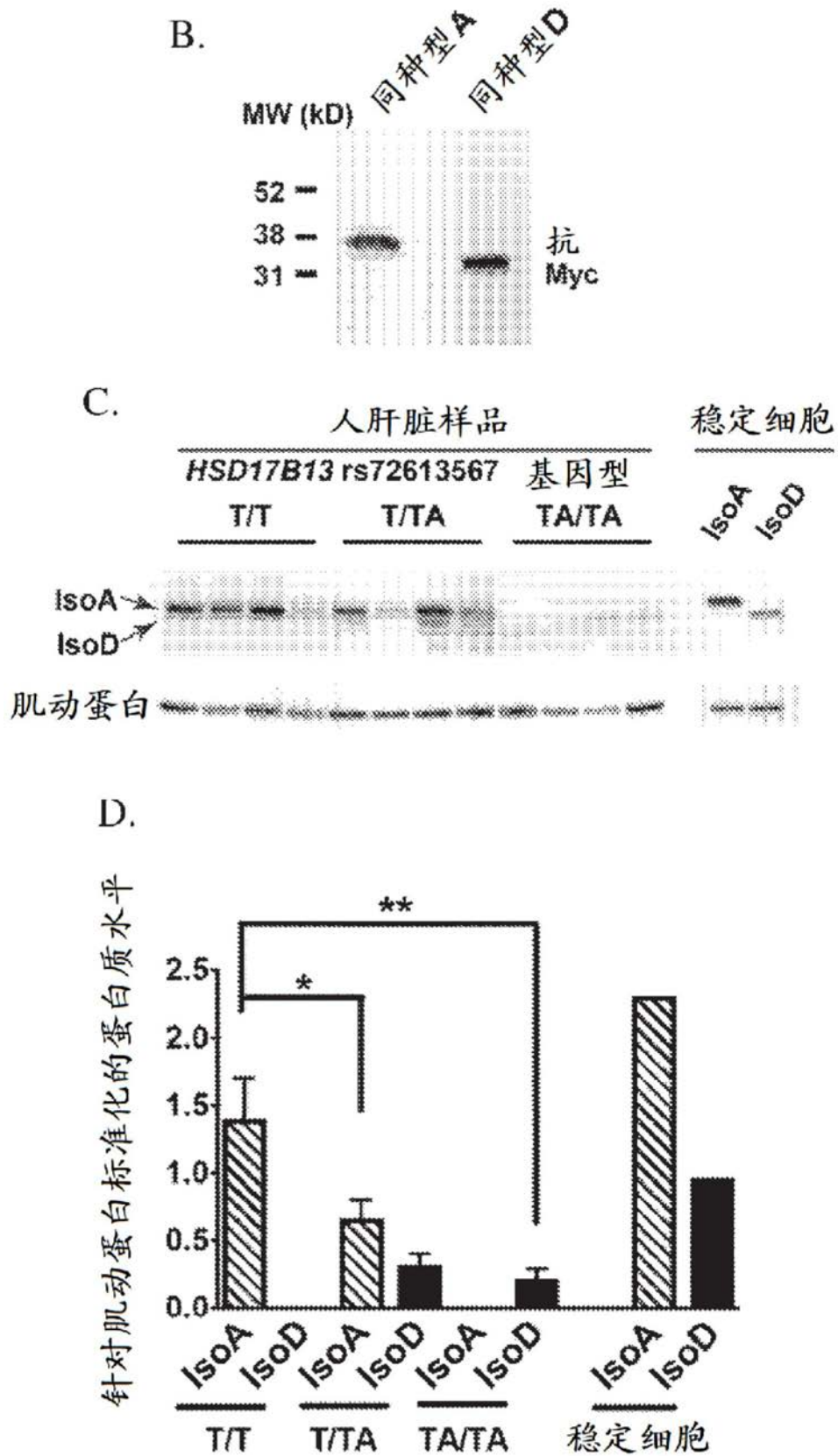


图18(续)

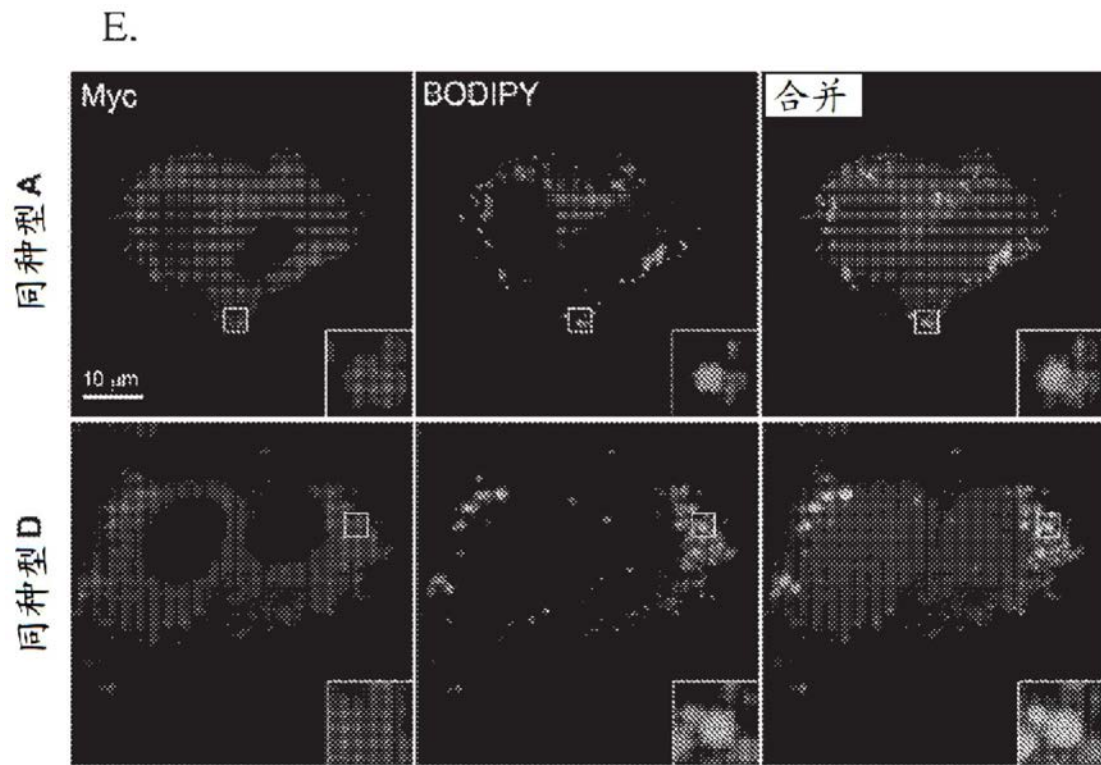


图18(续)

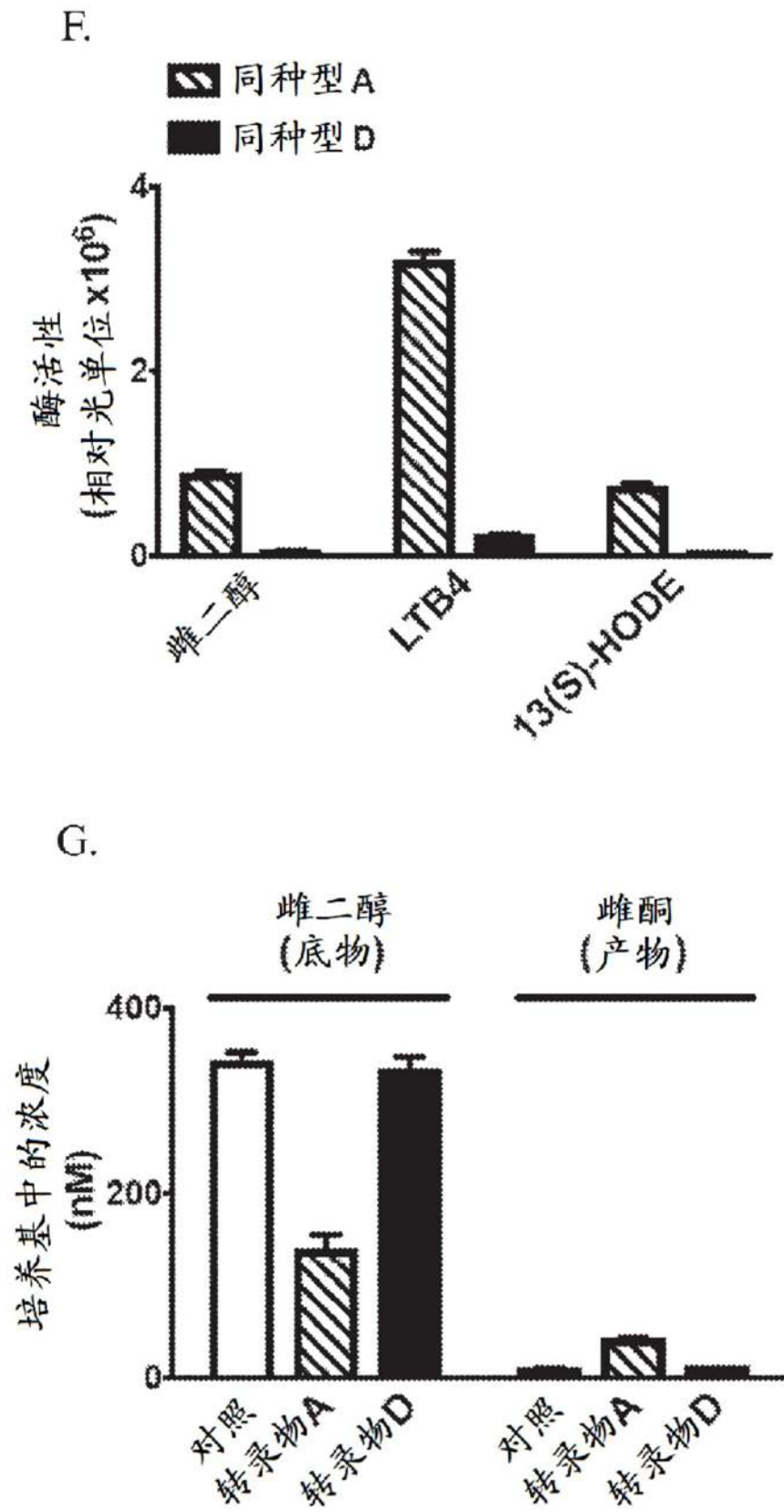


图18 (续)

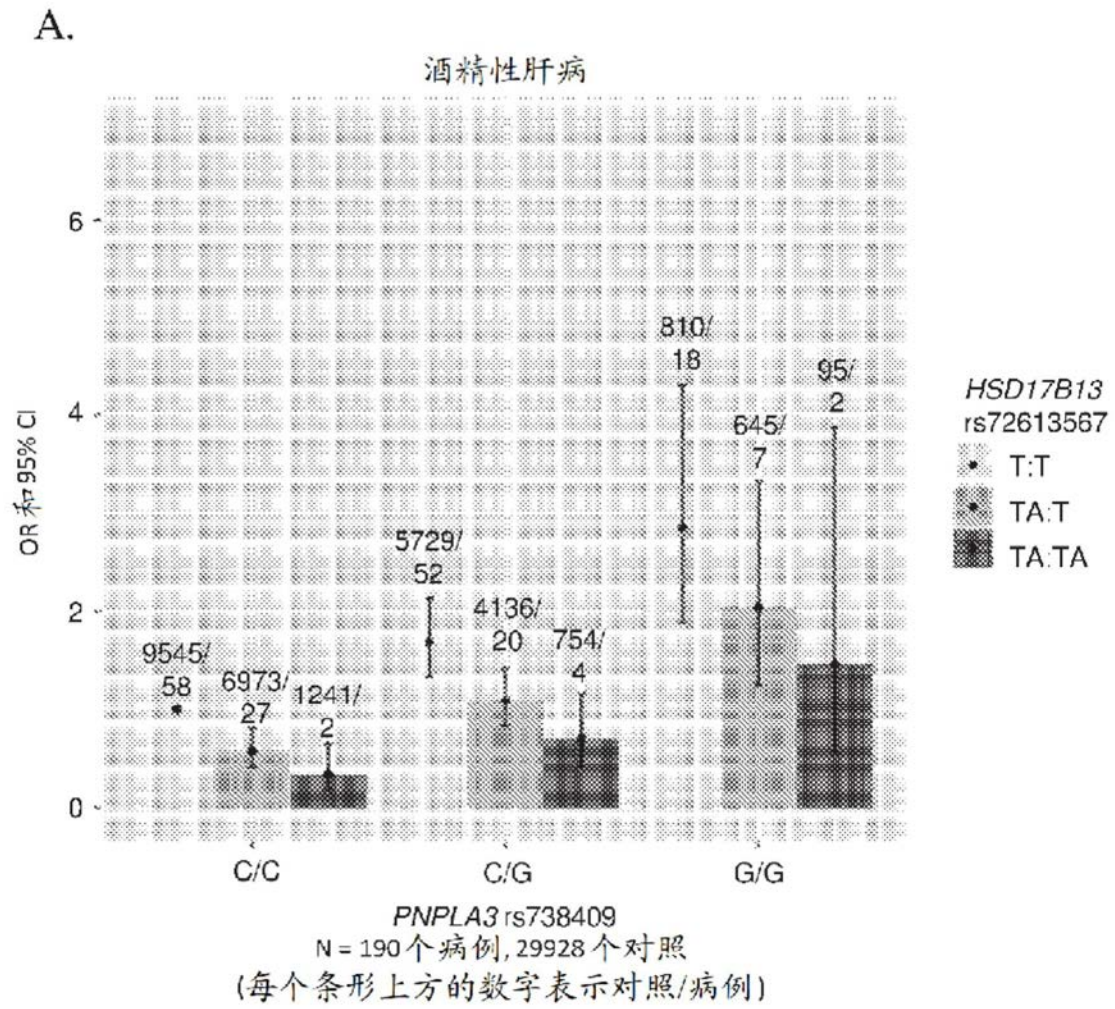


图19

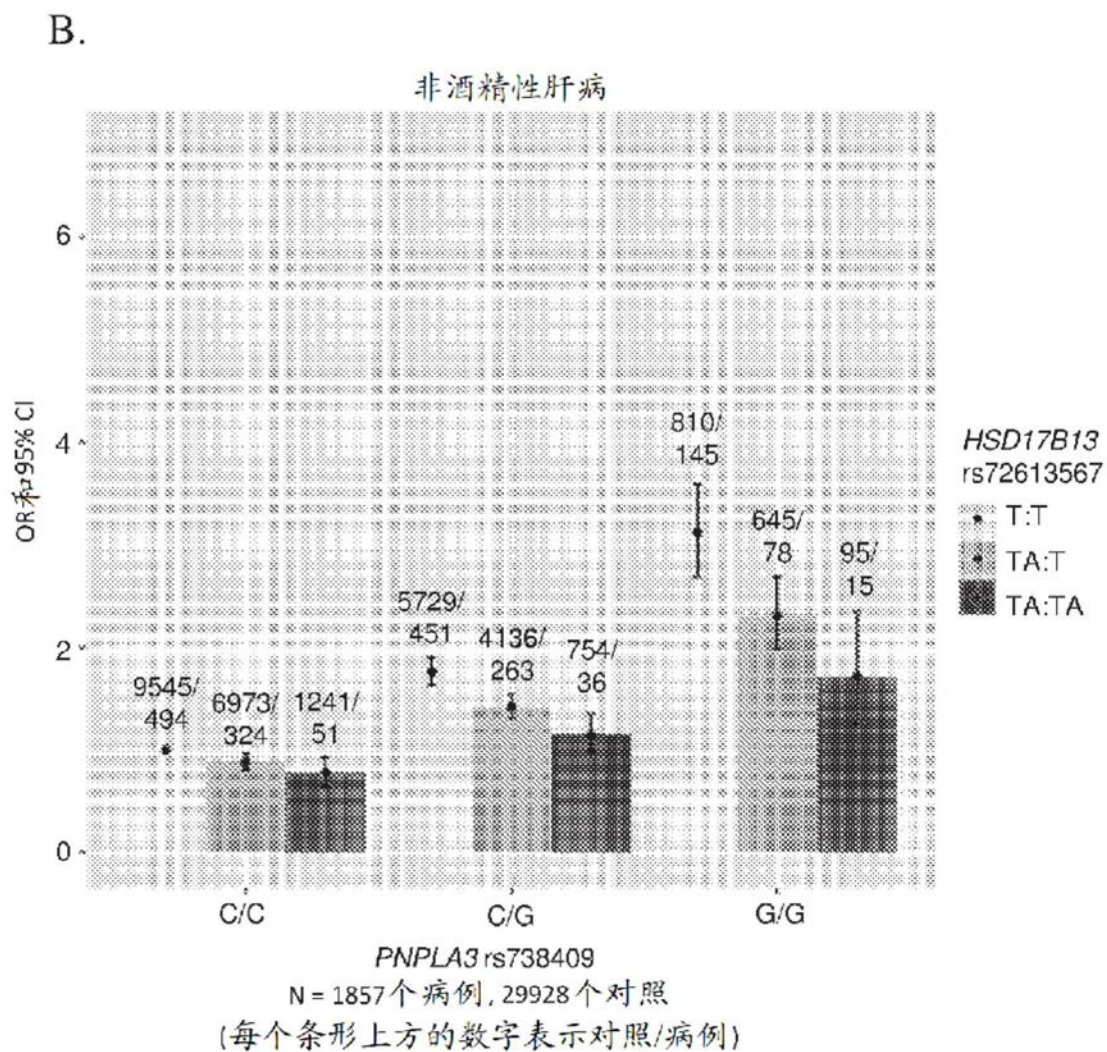


图19(续)