



(12) 发明专利申请

(10) 申请公布号 CN 115551880 A

(43) 申请公布日 2022.12.30

(21) 申请号 202180028122.X

舒纳克·麦金德尔

(22) 申请日 2021.05.04

(74) 专利代理机构 成都超凡明远知识产权代理

(30) 优先权数据

63/019,751 2020.05.04 US

有限公司 51258

专利代理师 张大皓

(85) PCT国际申请进入国家阶段日

2022.10.12

(51) Int. Cl.

C07K 14/435 (2006.01)

C07K 14/78 (2006.01)

(86) PCT国际申请的申请数据

PCT/US2021/030635 2021.05.04

C12N 5/071 (2006.01)

C12N 5/09 (2006.01)

(87) PCT国际申请的公布数据

W02021/226071 EN 2021.11.11

G01N 33/574 (2006.01)

(71) 申请人 梅约医学教育与研究基金会

地址 美国明尼苏达州

(72) 发明人 大卫·A·阿尔奎斯特

约翰·B·基谢尔 威廉·R·泰勒

道格拉斯·W·马奥尼

权利要求书10页 说明书61页

序列表14页

(54) 发明名称

检测胰腺神经内分泌肿瘤

(57) 摘要

本文提供了如下技术：用于胰腺神经内分泌肿瘤筛查，并且特别地但不排他地涉及用于检测胰腺神经内分泌肿瘤的存在的方法、组合物和相关用途。

1. 一种方法,所述方法包括:
通过以下方式测量人类个体生物样品中一种或多种基因的甲基化水平:
将所述生物样品中的基因组DNA用以甲基化特异性方式修饰DNA的试剂处理;
使用一组针对所选择的一种或多种基因的引物扩增经过处理的基因组DNA;以及
通过聚合酶链式反应、核酸测序、质谱法、甲基化特异性核酸酶、基于质量的分离和靶标捕获确定所述一种或多种基因的甲基化水平;
其中所述一种或多种基因选自表1A和表2A。
2. 如权利要求1所述的方法,其中所述一种或多种基因选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B。
3. 如权利要求1所述的方法,其中所述一种或多种基因选自SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBKL2A。
4. 如权利要求1所述的方法,其中所述一种或多种基因选自SRRM3、HCN2、SPTBN4和TMC6_A。
5. 如权利要求1所述的方法,其中将所述DNA用以甲基化特异性方式修饰DNA的试剂处理。
6. 如权利要求5所述的方法,其中所述试剂包含甲基化敏感性限制酶、甲基化依赖性限制酶和亚硫酸氢盐试剂中的一种或多种。
7. 如权利要求6所述的方法,其中将所述DNA用亚硫酸氢盐试剂处理以产生经过亚硫酸氢盐处理的DNA。
8. 如权利要求1所述的方法,其中所述测量包括多重扩增。
9. 如权利要求1所述的方法,其中测量至少一种甲基化标志基因的量包括使用一种或多种选自由以下组成的组的方法:甲基化特异性PCR、定量甲基化特异性PCR、甲基化特异性DNA限制酶分析、定量亚硫酸氢盐焦磷酸测序、皮瓣内切核酸酶测定、PCR-皮瓣测定和亚硫酸氢盐基因组测序PCR。
10. 如权利要求1所述的方法,其中所述样品包含血浆样品、血液样品或组织样品(例如胰腺组织)中的一种或多种。
11. 如权利要求1所述的方法,其中所述一组针对所选择的一种或多种基因的引物在表3中叙述。
12. 一种表征样品的方法,所述方法包括:
 - a) 测量来自所述样品的DNA中的至少一种甲基化标志基因的量,其中所述至少一种甲基化标志基因是选自表1A和表2A的一种或多种基因;
 - b) 测量所述DNA中的至少一种参考标志物的量;以及
 - c) 将在所述DNA中测量的所述至少一种甲基化标志基因的量的值计算为在所述DNA中测量的所述参考标志基因的量的百分比,其中所述值指示在所述样品中测量的所述至少一种甲基化标志DNA的量。

13. 如权利要求12所述的方法,其中所述一种或多种基因选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FBNP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B。

14. 如权利要求12所述的方法,其中所述一种或多种基因选自SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FBNP1、FAM78A、IER2、PNMAL2和MOBKL2A。

15. 如权利要求12所述的方法,其中所述一种或多种基因选自SRRM3、HCN2、SPTBN4和TMC6_A。

16. 如权利要求12所述的方法,其中所述至少一种参考标志物包含选自B3GALT6 DNA、ZDHHC1 DNA、 β -肌动蛋白DNA和非癌DNA的一种或多种参考标志物。

17. 如权利要求12所述的方法,其中所述样品包含血浆样品、血液样品或组织样品(例如胰腺组织)中的一种或多种。

18. 如权利要求12所述的方法,其中从所述样品提取所述DNA。

19. 如权利要求12所述的方法,其中将所述DNA用以甲基化特异性方式修饰DNA的试剂处理。

20. 如权利要求19所述的方法,其中所述试剂包含甲基化敏感性限制酶、甲基化依赖性限制酶和亚硫酸氢盐试剂中的一种或多种。

21. 如权利要求20所述的方法,其中将所述DNA用亚硫酸氢盐试剂处理以产生经过亚硫酸氢盐处理的DNA。

22. 如权利要求20所述的方法,其中使用一组针对所选择的一种或多种基因的引物扩增经过修饰的DNA。

23. 如权利要求22所述的方法,其中所述一组针对所选择的一种或多种基因的引物在表3中叙述。

24. 如权利要求12所述的方法,其中测量甲基化标志基因的量包括使用聚合酶链式反应、核酸测序、质谱法、甲基化特异性核酸酶、基于质量的分离和靶标捕获中的一种或多种。

25. 如权利要求24所述的方法,其中所述测量包括多重扩增。

26. 如权利要求24所述的方法,其中测量至少一种甲基化标志基因的量包括使用一种或多种选自由以下组成的组的方法:甲基化特异性PCR、定量甲基化特异性PCR、甲基化特异性DNA限制酶分析、定量亚硫酸氢盐焦磷酸测序、皮瓣内切核酸酶测定、PCR-皮瓣测定和亚硫酸氢盐基因组测序PCR。

27. 一种用于表征生物样品的方法,所述方法包括:

(a) 通过以下方式测量人类个体生物样品中选自表1A和表2A的一种或多种基因的CpG位点的甲基化水平:

将所述生物样品中的基因组DNA用亚硫酸氢盐处理;

使用一组针对所选择的一种或多种基因的引物扩增经过亚硫酸氢盐处理的基因组DNA;以及

通过甲基化特异性PCR、定量甲基化特异性PCR、甲基化敏感性DNA限制酶分析、定量亚硫酸氢盐焦磷酸测序或亚硫酸氢盐基因组测序PCR确定所述CpG位点的所述甲基化水平；

(b) 将所述甲基化水平与无PNET的对照样品中的一组相应基因的甲基化水平进行比较；以及

(c) 当在所述一种或多种基因中测量的所述甲基化水平高于在所述相应对照样品中测量的所述甲基化水平时确定所述个体患有PNET。

28. 如权利要求27所述的方法，其中所述一种或多种基因选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B。

29. 如权利要求27所述的方法，其中所述一种或多种基因选自SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBKL2A。

30. 如权利要求27所述的方法，其中所述一种或多种基因选自SRRM3、HCN2、SPTBN4和TMC6_A。

31. 如权利要求27所述的方法，其中所述一组针对所选择的一种或多种基因的引物在表3中叙述。

32. 如权利要求27所述的方法，其中所述生物样品为血浆样品、血液样品或组织样品（例如胰腺组织）。

33. 如权利要求27所述的方法，其中所述一种或多种基因由表1A和表2A中所示的基因组坐标描述。

34. 如权利要求27所述的方法，其中所述CpG位点存在于编码区或调控区中。

35. 如权利要求27所述的方法，其中所述测量一种或多种基因的CpG位点的甲基化水平包括选自由以下组成的组的确定：确定所述CpG位点的甲基化评分和确定所述CpG位点的甲基化频率。

36. 一种方法，所述方法包括：

(a) 测量人类个体生物样品中选自表1A和表2A的一种或多种基因的CpG位点的甲基化水平，

所述测量通过以下方式进行：

将所述生物样品中的基因组DNA用亚硫酸氢盐处理；

使用一组针对所选择的一种或多种基因的引物扩增经过亚硫酸氢盐处理的基因组DNA；以及

通过甲基化特异性PCR、定量甲基化特异性PCR、甲基化敏感性DNA限制酶分析、定量亚硫酸氢盐焦磷酸测序或亚硫酸氢盐基因组测序PCR确定所述CpG位点的所述甲基化水平。

37. 如权利要求36所述的方法，其中所述一种或多种基因选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、

TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B。

38. 如权利要求36所述的方法,其中所述一种或多种基因选自SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBKL2A。

39. 如权利要求36所述的方法,其中所述一种或多种基因选自SRRM3、HCN2、SPTBN4和TMC6_A。

40. 如权利要求36所述的方法,其中所述一组针对所选择的一种或多种基因的引物在表1A和表2A中叙述。

41. 如权利要求36所述的方法,其中所述生物样品为血浆样品、血液样品或组织样品(例如胰腺组织)。

42. 如权利要求36所述的方法,其中所述一种或多种基因由表3中所示的基因组坐标描述。

43. 如权利要求36所述的方法,

其中若所述生物样品为组织样品,则所述一种或多种基因选自:

●ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B;

●SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBKL2A;以及

●SRRM3、HCN2、SPTBN4和TMC6_A;

其中若所述生物样品为血浆样品,则所述一种或多种基因选自:

●ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B;以及

●SRRM3、HCN2、SPTBN4和TMC6_A。

44. 如权利要求36所述的方法,其中所述测量一种或多种基因的CpG位点的所述甲基化水平包括选自由以下组成的组的确定:确定所述CpG位点的甲基化评分和确定所述CpG位点的甲基化频率。

45. 一种在从受试者获得的样品中筛查PNET的方法,所述方法包括:

1) 测定包含具有表1A和表2A中叙述的注释的染色体区域的DNA甲基化标志物的甲基化状态,并且

2) 当所述标志物的所述甲基化状态不同于在未患PNET的受试者中测定的所述标志物的甲基化状态时将所述受试者鉴定为患有PNET。

46. 如权利要求45所述的方法,其中所述一种或多种基因选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B。

47. 如权利要求45所述的方法,其中所述一种或多种基因选自SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBKL2A。

48. 如权利要求45所述的方法,其中所述一种或多种基因选自SRRM3、HCN2、SPTBN4和TMC6_A。

49. 如权利要求45所述的方法,所述方法包括测定多种标志物。

50. 如权利要求45所述的方法,其中所述标志物在高CpG密度启动子中。

51. 如权利要求45所述的方法,其中所述样品为粪便样品、组织样品、胰腺组织样品、血浆样品或尿液样品。

52. 如权利要求45所述的方法,其中所述测定包括使用甲基化特异性聚合酶链式反应、核酸测序、质谱法、甲基化特异性核酸酶、基于质量的分离或靶标捕获。

53. 如权利要求45所述的方法,其中所述测定包括使用甲基化特异性寡核苷酸。

54. 如权利要求45所述的方法,

其中若所述生物样品是组织样品,则所述一种或多种基因选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B;

其中若所述生物样品为血浆样品,则所述一种或多种基因选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B。

55. 一种用于表征来自人类患者的样品的方法,所述方法包括:

a) 从人类患者的样品获得DNA;

b) 测定包含具有表1A和表2A中叙述的注释的染色体区的DNA甲基化标志物的甲基化状态;

c) 将所述一种或多种DNA甲基化标志物的所测定的甲基化状态与未患PNET的人类患者的所述一种或多种DNA甲基化标志物的甲基化水平参考进行比较。

56. 如权利要求55所述的方法,其中所述一种或多种基因选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B。

57. 如权利要求55所述的方法,其中所述一种或多种基因选自SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、

MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBKL2A。

58.如权利要求55所述的方法,其中所述一种或多种基因选自SRRM3、HCN2、SPTBN4和TMC6_A。

59.如权利要求55所述的方法,其中所述样品为粪便样品、组织样品、胰腺组织样品、血浆样品或尿液样品。

60.如权利要求55所述的方法,所述方法包括测定多种DNA甲基化标志物。

61.如权利要求55所述的方法,

其中若所述生物样品是组织样品,则所述一种或多种基因选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B;

其中若所述生物样品为血浆样品,则所述一种或多种基因选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B。

62.如权利要求55所述的方法,其中所述测定包括使用甲基化特异性聚合酶链式反应、核酸测序、质谱法、甲基化特异性核酸酶、基于质量的分离或靶标捕获。

63.如权利要求55所述的方法,其中所述测定包括使用甲基化特异性寡核苷酸。

64.如权利要求63所述的方法,其中所述甲基化特异性寡核苷酸选自表3中叙述的一组针对所选择的一种或多种基因的引物。

65.一种用于表征从人类受试者获得的样品的方法,所述方法包括使包含DMR的核酸与亚硫酸氢盐试剂反应以产生与亚硫酸氢盐反应过的核酸;对所述与亚硫酸氢盐反应过的核酸进行测序以提供所述与亚硫酸氢盐反应过的核酸的核苷酸序列;将所述与亚硫酸氢盐反应过的核酸的所述核苷酸序列与来自未患PNET的受试者的包含所述DMR的核酸的核苷酸序列进行比较以鉴定所述两个序列的差异。

66.一种用于表征从人类受试者获得的样品的系统,所述系统包括:分析组件,所述分析组件被配置用于确定样品的甲基化状态;软件组件,所述软件组件被配置用于将所述样品的所述甲基化状态与数据库中记录的对照样品或参考样品的甲基化状态进行比较;以及警报组件,所述警报组件被配置用于基于甲基化状态的组合确定单个值并向用户发出关于PNET相关的甲基化状态的警报。

67.如权利要求66所述的系统,其中所述样品包含含有DMR的核酸。

68.如权利要求66所述的系统,所述系统还包括用于分离核酸的组件。

69.如权利要求66所述的系统,所述系统还包括用于收集样品的组件。

70.如权利要求66所述的系统,其中所述样品为粪便样品、组织样品、胰腺组织样品、血浆样品或尿液样品。

71.如权利要求66所述的系统,其中所述数据库包含含有DMR的核酸序列。

72.如权利要求66所述的系统,其中所述数据库包含来自未患PNET的受试者的核酸序

列。

73. 一种试剂盒,所述试剂盒包括:

1) 亚硫酸氢盐试剂;以及

2) 包含来自选自由表1A和表2A的DMR 1-198组成的组的DMR的序列并具有与未患PNET的受试者相关的甲基化状态的对照核酸。

74. 一种试剂盒,所述试剂盒包括亚硫酸氢盐试剂和根据SEQ ID NO 1-66的寡核苷酸。

75. 一种试剂盒,所述试剂盒包括用于从受试者获得样品的样品收集器;用于从所述样品分离核酸的试剂;亚硫酸氢盐试剂;以及根据SEQ ID NO 1-66的寡核苷酸。

76. 根据权利要求75所述的试剂盒,其中所述样品为粪便样品、组织样品、胰腺组织样品、血浆样品或尿液样品。

77. 一种组合物,所述组合物包含含有DMR的核酸和亚硫酸氢盐试剂。

78. 一种组合物,所述组合物包含含有DMR的核酸和根据SEQ ID NO 1-66的寡核苷酸。

79. 一种组合物,所述组合物包含含有DMR的核酸和甲基化敏感性限制酶。

80. 一种组合物,所述组合物包含含有DMR的核酸和聚合酶。

81. 一种用于在从受试者获得的样品中筛查PNET的方法,所述方法包括使包含DMR的核酸与亚硫酸氢盐试剂反应以产生与亚硫酸氢盐反应过的核酸;对所述与亚硫酸氢盐反应过的核酸进行测序以提供所述与亚硫酸氢盐反应过的核酸的核苷酸序列;将所述与亚硫酸氢盐反应过的核酸的所述核苷酸序列与来自未患PNET的受试者的包含所述DMR的核酸的核苷酸序列进行比较以鉴定所述两个序列的差异;以及当存在差异时将所述受试者鉴定为患有PNET。

82. 一种用于在从受试者获得的样品中筛查PNET的系统,所述系统包括:分析组件,所述分析组件被配置用于确定样品的甲基化状态;软件组件,所述软件组件被配置用于将所述样品的所述甲基化状态与数据库中记录的对照样品或参考样品的甲基化状态进行比较;以及警报组件,所述警报组件被配置用于基于甲基化状态的组合确定单个值并向用户发出关于PNET相关的甲基化状态的警报。

83. 如权利要求82所述的系统,其中所述样品包含含有DNA甲基化标志物的核酸,所述DNA甲基化标志物包含差异甲基化区域(DMR)中的碱基,所述DMR选自由表1A和表2A的DMR 1-198组成的组。

84. 如权利要求82所述的系统,所述系统还包括用于分离核酸的组件。

85. 如权利要求82所述的系统,所述系统还包括用于收集样品的组件。

86. 如权利要求82所述的系统,所述系统还包括用于收集粪便样品、胰腺组织样品和/或血浆样品的组件。

87. 如权利要求82所述的系统,其中所述数据库包含来自未患PNET的受试者的核酸序列。

88. 一种用于表征生物样品的方法,所述方法包括:

通过以下方式测量人类个体生物样品中SRRM3、HCN2、SPTBN4和TMC6_A中的一种或多种的CpG位点的甲基化水平:

将所述生物样品中的基因组DNA用亚硫酸氢盐处理;

使用对SRRM3的CpG位点具有特异性的引物、对HCN2的CpG位点具有特异性的引物、对

SPTBN4的CpG位点具有特异性的引物和对TMC6_A的CpG位点具有特异性的引物扩增经过亚硫酸氢盐处理的基因组DNA，

其中对SRRM3具有特异性的所述引物能够结合由SEQ ID No:39和40结合的扩增子，其中由SEQ ID No:39和40结合的所述扩增子是包含7号染色体坐标75896582-75896785的遗传区域的至少一部分，

其中对HCN2具有特异性的所述引物能够结合由SEQ ID No:13和14结合的扩增子，其中由SEQ ID No:13和14结合的所述扩增子是包含19号染色体坐标591692-591781的遗传区域的至少一部分，

其中对SPTBN4具有特异性的所述引物能够结合由SEQ ID No:37和38结合的扩增子，其中由SEQ ID No:37和38结合的所述扩增子是包含19号染色体坐标41060185-41060270的遗传区域的至少一部分；并且

其中对TMC6_A具有特异性的所述引物能够结合由SEQ ID No:43和44结合的扩增子，其中由SEQ ID No:43和44结合的所述扩增子是包含17号染色体坐标76123640-76123768的遗传区域的至少一部分；

通过甲基化特异性PCR、定量甲基化特异性PCR、甲基化敏感性DNA限制酶分析、定量亚硫酸氢盐焦磷酸测序或亚硫酸氢盐基因组测序PCR确定所述SRRM3、HCN2、SPTBN4和TMC6_A中的一种或多种的所述CpG位点的所述甲基化水平。

89. 如权利要求88所述的方法，其中所述生物样品为血液样品或胰腺组织样品。

90. 如权利要求88所述的方法，其中所述CpG位点存在于编码区或调控区中。

91. 如权利要求88所述的方法，其中所述测量所述SRRM3、HCN2、SPTBN4和TMC6_A中的一种或多种的所述CpG位点的所述甲基化水平包括选自由以下组成的组的确定：确定所述CpG位点的甲基化评分和确定所述CpG位点的甲基化频率。

92. 一种用于表征生物样品的方法，所述方法包括：

测量人类个体生物样品中选自以下的一种或多种标志物的CpG位点的甲基化水平：ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBK2A、PNMAL2、S1PR4_A、LGALS3和MYO15B，所述测量通过以下方式进行：

将所述生物样品中的基因组DNA用亚硫酸氢盐处理；

使用对所述一种或多种标志物的CpG位点具有特异性的引物扩增经过亚硫酸氢盐处理的基因组DNA，

其中对每种标志物具有特异性的所述引物能够结合由表3中叙述的相应引物序列结合的扩增子，其中由相应引物序列结合的所述扩增子是包含表1A或表2A中叙述的相应染色体区域的遗传区域的至少一部分；

通过甲基化特异性PCR、定量甲基化特异性PCR、甲基化敏感性DNA限制酶分析、定量亚硫酸氢盐焦磷酸测序或亚硫酸氢盐基因组测序PCR确定所述一种或多种标志物的所述CpG位点的所述甲基化水平。

93. 如权利要求92所述的方法，其中所述生物样品为血液样品或胰腺组织样品。

94. 如权利要求92所述的方法，其中所述CpG位点存在于编码区或调控区中。

95. 如权利要求92所述的方法,其中所述测量所述一种或多种标志物的所述CpG位点的所述甲基化水平包括选自以下组成的组的确定:确定所述CpG位点的甲基化评分和确定所述CpG位点的甲基化频率。

96. 一种用于从人类个体的生物样品制备脱氧核糖核酸(DNA)级分的方法,所述脱氧核糖核酸(DNA)级分可用于分析参与一种或多种染色体畸变的一个或多个遗传基因座,所述方法包括:

(a) 从人类个体的生物样品中提取基因组DNA;

(b) 通过以下方式产生所提取的基因组DNA的级分:

(i) 将所述所提取的基因组DNA用亚硫酸氢盐处理;

(ii) 使用对表1A和表2A中叙述的一种或多种标志物的CpG位点具有特异性的单独引物扩增经过亚硫酸氢盐处理的基因组DNA;

(c) 通过测量所述一种或多种标志物中的每一种标志物的所述CpG位点的甲基化水平来分析所述所提取的基因组DNA所产生的级分中的一个或多个遗传基因座。

97. 如权利要求96所述的方法,其中测量所述一种或多种标志物中的每一种标志物的所述CpG位点的甲基化水平通过甲基化特异性PCR、定量甲基化特异性PCR、甲基化敏感性DNA限制酶分析或亚硫酸氢盐基因组测序PCR确定。

98. 如权利要求96所述的方法,其中使用对所述一种或多种标志物中的每一种标志物的CpG位点具有特异性的引物扩增所述经过亚硫酸氢盐处理的基因组DNA是特异性结合如表1A和/或表2A所示的所述标志物的遗传区域的至少一部分的一组引物。

99. 如权利要求96所述的方法,其中所述生物样品为粪便样品、组织样品、器官分泌样品、CSF样品、唾液样品、血液样品、血浆样品或尿液样品。

100. 如权利要求96所述的方法,其中所分析的一个或多个遗传基因座中的每一个遗传基因座均与PNET相关。

101. 如权利要求96所述的方法,其中所述一种或多种标志物选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSP0、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B。

102. 如权利要求96所述的方法,其中所述一种或多种标志物选自SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSP0、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBKL2A1。

103. 如权利要求96所述的方法,其中所述一种或多种标志物选自SRRM3、HCN2、SPTBN4和TMC6_A。

104. 一种用于从人类个体的生物样品制备脱氧核糖核酸(DNA)级分的方法,所述脱氧核糖核酸(DNA)级分可用于分析参与一种或多种染色体畸变的一个或多个DNA片段,所述方法包括:

(a) 从人类个体的生物样品中提取基因组DNA;

(b) 通过以下方式产生所提取的基因组DNA的级分:

(i) 将所述所提取的基因组DNA用亚硫酸氢盐处理；

(ii) 使用对表1A和表2A中叙述的一种或多种标志物的CpG位点具有特异性的单独引物扩增经过亚硫酸氢盐处理的基因组DNA；

(c) 通过测量所述一种或多种标志物中的每一种标志物的所述CpG位点的甲基化水平来分析所述所提取的基因组DNA所产生的级分中的一个或多个DNA片段。

105. 如权利要求104所述的方法, 其中测量所述一种或多种标志物中的每一种标志物的所述CpG位点的甲基化水平通过甲基化特异性PCR、定量甲基化特异性PCR、甲基化敏感性DNA限制酶分析或亚硫酸氢盐基因组测序PCR确定。

106. 如权利要求104所述的方法, 其中使用对所述一种或多种标志物中的每一种标志物的CpG位点具有特异性的引物扩增所述经过亚硫酸氢盐处理的基因组DNA是特异性结合如表1A和/或表2A所示的所述标志物的遗传区域的至少一部分的一组引物。

107. 如权利要求104所述的方法, 其中所述生物样品为粪便样品、组织样品、器官分泌样品、CSF样品、唾液样品、血液样品、血浆样品或尿液样品。

108. 如权利要求104所述的方法, 其中所分析的DNA片段中的每一个DNA片段均与PNET相关。

109. 如权利要求104所述的方法, 其中所述一种或多种标志物选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B。

110. 如权利要求104所述的方法, 其中所述一种或多种标志物选自SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBKL2A1。

111. 如权利要求104所述的方法, 其中所述一种或多种标志物选自SRRM3、HCN2、SPTBN4和TMC6_A。

检测胰腺神经内分泌肿瘤

[0001] 相关申请的交叉引用

[0002] 本申请要求2020年5月4日提交的美国临时专利申请号63/019,751的优先权,该临时专利申请特此通过引用整体并入。

技术领域

[0003] 本文提供了如下技术:用于胰腺神经内分泌肿瘤筛查,并且特别地但不排他地涉及用于检测胰腺神经内分泌肿瘤的存在的方法、组合物和相关用途。

背景技术

[0004] 胰腺神经内分泌肿瘤(PNET)的怀疑是基于其增强胰腺实性病变的特征性放射学表现,并且通常通过EUS引导的活检证实诊断。PNET有时可能是囊性的并与其它胰腺囊性病变更酷似,导致诊断不确定。目前没有用于诊断PNET的基于血液的或囊液生物标志物。偶然诊断出PNET会导致治疗困境,因为胰腺切除术具有重大风险,尽管PNET通常生长缓慢,允许在选择的病例中进行观察等待和定期监测成像而不进行治疗,但生物学行为可能是不可预知的并且与大小无关。

[0005] 目前,世界卫生组织(WHO)根据胰腺组织中评估的有丝分裂计数和增殖指数(Ki-67)将所有PNET分为低等级(G1)、中等级(G2)和高等级(G3)类别。没有用于确定等级的非侵入性标志物,因此对于哪些患者群体可以安全观察缺乏明确的共识。在接受胰腺切除的患者中,复发并不少见,并且可在手术后数年发生。此外,在患有转移性疾病的患者中,当前药物治疗法是抑制肿瘤的,并且在治疗期间没有监测疾病活动的生物标志物。

[0006] 因此,临床需要可应用于囊液和血液进行诊断、分期和监护的准确的PNET生物标志物。

[0007] 本发明满足了此类需要。实际上,本发明提供了新型甲基化DNA标志物,其区分各种生物样品(例如,组织、血液)中的PNET病例。

发明内容

[0008] 已经将甲基化DNA作为大多数肿瘤类型组织中潜在的一类生物标志物进行研究。在许多情况下,DNA甲基转移酶向DNA中的胞嘧啶-磷酸-鸟嘌呤(CpG)岛位点添加甲基,作为基因表达的表观遗传控制。在生物学上有吸引力的机制中,认为肿瘤抑制基因启动子区域中的获得性甲基化事件会使表达沉默,从而促进肿瘤发生。DNA甲基化可能是比RNA或蛋白质表达更具有化学和生物学稳定性的诊断工具(Laird(2010) Nat Rev Genet 11:191-203)。此外,在散发性结肠癌等其它癌症中,甲基化标志物提供了优良的特异性并且比单个DNA突变提供更广泛的信息且更灵敏(Zou等人(2007) Cancer Epidemiol Biomarkers Prev 16:2686-96)。

[0009] 当应用于动物模型和人类细胞系时,对CpG岛的分析产生了重要的发现。例如,Zhang和同事发现来自同一CpG岛的不同部分的扩增子可具有不同的甲基化水平(Zhang等

人(2009) *PLoS Genet* 5:e1000438)。此外,甲基化水平在高度甲基化和未甲基化的序列之间呈双峰分布,进一步支持DNA甲基转移酶活性的二进制开关样模式(Zhang等人(2009) *PLoS Genet* 5:e1000438)。对体内小鼠组织和体外细胞系的分析证明,只有约0.3%的高CpG密度启动子(HCP,定义为在300个碱基对区域内具有>7%的CpG序列)被甲基化,而低CpG密度的区域(LCP,定义为在300个碱基对区域内具有<5%的CpG序列)倾向于以动态组织特异性模式频繁甲基化(Meissner等人(2008) *Nature* 454:766-70)。HCP包括普遍存在的看家基因和高度调控的发育基因的启动子。在>50%甲基化的HCP位点中有几个已确立的标志物,诸如Wnt 2、NDRG2、SFRP2和BMP3(Meissner等人(2008) *Nature*454:766-70)。

[0010] 已经将DNA甲基转移酶在胞嘧啶-磷酸-鸟嘌呤(CpG) 岛位点处对DNA的表观遗传甲基化作为大多数肿瘤类型组织中潜在的一类生物标志物进行研究。

[0011] 有几种方法可用于搜索新的甲基化标志物。虽然基于微阵列的CpG甲基化询问是一种合理的高通量方法,但这种策略偏向于已知的感兴趣区域,主要是已确立的肿瘤抑制启动子。在过去十年中,已经开发了用于全基因组DNA甲基化分析的替代方法。有四种基本方法。第一种采用通过识别特定甲基化位点的限制酶消化DNA,然后是几种可能的分析技术,这些技术提供的甲基化数据仅限于酶识别位点或用于在定量步骤中扩增DNA的引物(例如甲基化特异性PCR;MSP)。第二种方法使用针对甲基胞嘧啶或其它甲基化特异性结合结构域的抗体富集基因组DNA的甲基化部分,然后进行微阵列分析或测序以将片段映射至参考基因组。这种方法不提供片段内所有甲基化位点的单核苷酸分辨率。第三种方法首先是对DNA进行亚硫酸氢盐处理,将所有未甲基化的胞嘧啶转化为尿嘧啶,然后是限制酶消化,并在与接头配体偶联后对所有片段进行完整测序。限制酶的选择可以富集片段的CpG密集区域,减少分析期间可映射到多个基因位置的冗余序列的数量。第四种方法涉及对DNA进行无亚硫酸氢盐处理,该处理描述一种无亚硫酸盐的碱基分辨率测序方法,TET辅助的吡啶硼烷测序(TAPS),用于无损直接检测5-甲基胞嘧啶和5-羟甲基胞嘧啶,而不影响未修饰的胞嘧啶(参见Liu等人,2019, *Nat Biotechnol.* 37,第424-429页)。在一些实施方案中,不管具体的酶促转化方法如何,仅转化甲基化胞嘧啶。

[0012] 简化代表性亚硫酸氢盐测序(RRBS)在中高读取覆盖率下在单核苷酸分辨率下产生所有CpG岛的80-90%和大多数肿瘤抑制启动子的CpG甲基化状况数据。在癌症病例对照研究中,对这些读数的分析可鉴定差异甲基化区域(DMR)。在先前对胰腺癌样本的RRBS分析中,显现出数百个DMR,其中许多从未与致癌作用相关联,并且许多未注释。对独立组织样品集的进一步验证研究证实了在性能方面具有100%灵敏度和特异性的标志物CpG。

[0013] PNET占胰腺肿瘤的一小部分但很重要,可以实性或囊性胰腺肿块的形式存在。过去十年中在美国PNET的患病率增加,这主要是由于广泛诊断使用高清晰度腹部成像的偶然检出(参见Dasari A等人, *JAMA oncology* 2017;3(10):1335-42;Hallet J等人, *Cancer*. 2015;121(4):589-97)。绝大多数PNET是无功能性的,并且不呈现激素过量产生的临床综合征。尽管在临床上沉默的,但PNET在组织学上可能是高等级的,并且在初始检测时偶而会表现出转移性疾病,而与原发病变的大小无关。目前没有用于准确检测PNET的非侵入性生物标志物,并且诊断依赖于组织取样,组织取样由于诊断组织收率低和相关的胰腺炎风险,在小病变中通常具有挑战性。此外,由于NET可出现在胰腺以外的多个其它器官(肺、小肠)中,因此基于血液的分子诊断测试定位原发性癌症的部位将是有价值的。

[0014] 本发明解决了PNET诊断和管理中的一个重要空白,即缺乏准确的生物标志物。先前完成了新型甲基化DNA标志物(MDM)的全甲基化组发现和验证以检测组织中的胰腺导管腺癌(PDAC),从而鉴定了胰腺囊液、胰液和血液中的MDM组,该MDM组可以准确地区分PDAC与健康对照(参见Kisiel JB等人,Clin Cancer Res.2015;21(19):4473-81;Majumder S,Gastroenterology.150(4):S120-S1;Majumder S等人,Gastroenterology.152(5):S148)。

[0015] 实际上,如实施例I所述,在鉴定本发明实施方案的过程中进行的实验鉴定了一组新型差异甲基化区域(DMR),用于区分PNET来源的DNA与非肿瘤对照DNA。

[0016] 此类实验列出并描述了198种将PNET组织与良性组织区别开来的新型DNA甲基化标志物(参见表1A、表1B、表2A、表2B、表4、表5A和表5C,以及实施例I)。

[0017] 从这198种新型DNA甲基化标志物中,进一步的实验鉴定了以下能够区别PNET组织(例如,囊性PNET组织、实性PNET组织、转移性PNET组织)与良性组织的标志物和/或标志物组:

[0018] • ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B(参见表4、表5A和表5C、实施例I);

[0019] • SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBKL2A(参见表5C、实施例I);以及

[0020] • SRRM3、HCN2、SPTBN4和TMC6_A(参见表5C、实施例I)。

[0021] 从这198种新型DNA甲基化标志物中,进一步的实验鉴定了以下用于在血液样品(例如,血浆样品、全血样品、白细胞样品、血清样品)中检测PNET的标志物和/或标志物组:

[0022] • ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B(参见表4、表5A和表5C、实施例I);以及

[0023] • SRRM3、HCN2、SPTBN4和TMC6_A(参见表5C、实施例I)。

[0024] 从这198种新型DNA甲基化标志物中,进一步的实验鉴定了以下用于在血液样品(例如,血浆样品、全血样品、白细胞样品、血清样品)中检测转移性PNET的标志物和/或标志物组:

[0025] • SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBKL2A(参见表5C、实施例I)。

[0026] 从这198种新型DNA甲基化标志物中,进一步的实验鉴定了以下用于在血液样品(例如,血浆样品、全血样品、白细胞样品、血清样品)中检测肺神经内分泌肿瘤(NET)的标志物和/或标志物组:

[0027] • SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、

PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBK2A(参见表5C、实施例I)。

[0028] 从这198种新型DNA甲基化标志物中,进一步的实验鉴定了以下用于在血液样品(例如,血浆样品、全血样品、白细胞样品、血清样品)中检测小肠神经内分泌肿瘤(NET)的标志物和/或标志物组:

[0029] • SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBK2A(参见表5C、实施例I)。

[0030] 如本文所述,这项技术提供了对PNET总体和各种相关NET类型(例如肺NET、小肠NET)具有高区分力的许多甲基化DNA标志物和其子集(例如,2种、3种、4种、5种、6种、7种或8种标志物的集合)。实验将选择过滤器应用于候选标志物以鉴定提供高信噪比和低背景水平,从而为PNET筛查或诊断提供高特异性的标志物。

[0031] 在一些实施方案中,这项技术涉及评估生物样品(例如胰腺组织样品、血液样品)中本文鉴定的一种或多种标志物的存在和甲基化状态。这些标志物包含如本文所讨论,例如,如表1A和表2A中提供的一个或多个差异甲基化区域(DMR)。在这项技术的实施方案中评估甲基化状态。因而,本文提供的技术在测量基因的甲基化状态所采用的方法方面不受限制。例如,在一些实施方案中,甲基化状态通过基因组扫描方法测量。例如,一种方法涉及限制性标记基因组扫描(Kawai等人(1994) Mol. Cell. Biol. 14:7421-7427),另一个实例涉及甲基化敏感性任意引物PCR(Gonzalzo等人(1997) Cancer Res. 57:594-599)。在一些实施方案中,特定CpG位点处甲基化模式的变化通过用甲基化敏感性限制酶消化基因组DNA,随后对感兴趣区域进行Southern分析(消化-Southern法)来监测。在一些实施方案中,分析甲基化模式的变化涉及基于PCR的过程,该过程涉及在PCR扩增之前用甲基化敏感性限制酶或甲基化依赖性限制酶消化基因组DNA(Singer-Sam等人(1990) Nucl. Acids Res. 18:687)。另外,已经报道了利用对DNA的亚硫酸氢盐处理作为甲基化分析起点的其它技术。这些技术包括甲基化特异性PCR(MSP)(Herman等人(1992) Proc. Natl. Acad. Sci. USA 93:9821-9826)和限制酶消化从亚硫酸氢盐转化的DNA扩增的PCR产物(Sadri和Hornsby(1996) Nucl. Acids Res. 24:5058-5059;以及Xiong和Laird(1997) Nucl. Acids Res. 25:2532-2534)。已开发PCR技术用于基因突变的检测(Kuppuswamy等人(1991) Proc. Natl. Acad. Sci. USA 88:1143-1147)和等位基因特异性表达的定量(Szabo和Mann(1995) Genes Dev. 9:3097-3108;以及Singer-Sam等人(1992) PCR Methods Appl. 1:160-163)。此类技术使用退火于PCR生成的模板,并且紧靠待测定的单核苷酸的5'终止的内部引物。使用如美国专利号7,037,650中所述的“定量Ms-SNuPE测定”的方法在一些实施方案中使用。

[0032] 在评估甲基化状态时,常常将甲基化状态表示为在特定位点处(例如在单核苷酸处,在特定区域或位置处,在较长感兴趣序列例如DNA的多达约100bp、200bp、500bp、1000bp的子序列或更长序列处)被甲基化的单个DNA链相对于样品中包含该特定位点的DNA总群体的分数或百分比。传统上,未甲基化核酸的量通过使用校准物进行PCR来确定。然后,对已知量的DNA进行亚硫酸氢盐处理(或非亚硫酸氢盐处理(参见Liu等人,2019, Nat

Biotechnol. 37, 第424-429页)), 并使用实时PCR或其它指数扩增, 例如QuARTS测定(例如, 如美国专利号8,361,720、8,715,937、8,916,344和9,212,392所提供的) 确定所产生的甲基化特异性序列。

[0033] 例如, 在一些实施方案中, 方法包括通过使用外部标准来生成未甲基化靶标的标准曲线。标准曲线由至少两个点构建, 并将未甲基化DNA的实时Ct值与已知的定量标准相关联。然后, 由至少两个点和外部标准构建甲基化靶标的第二标准曲线。该第二标准曲线将甲基化DNA的Ct与已知的定量标准相关联。接着, 确定甲基化和未甲基化群体的测试样品Ct值, 并根据前两个步骤产生的标准曲线计算DNA的基因组当量。感兴趣位点处的甲基化百分比由甲基化DNA的量相对于群体中DNA的总量计算, 例如(甲基化DNA的数量)/(甲基化DNA的数量+未甲基化DNA的数量) × 100。

[0034] 在一些实施方案中, 多个不同的靶区域包含参考靶区域, 并且在某些优选的实施方案中, 参考靶区域包含β-肌动蛋白和/或ZDHHC1, 和/或B3GALT6。

[0035] 本文还提供了用于实施所述方法的组合物和试剂盒。例如, 在一些实施方案中, 对一种或多种MDM具有特异性的试剂(例如, 引物、探针) 单独或成组提供(例如, 用于扩增多个标志物的引物对组)。还可以提供用于进行检测测定的其它试剂(例如, 酶、缓冲液、用于进行QuARTS、PCR、测序的阳性和阴性对照、亚硫酸氢盐、十-十一易位(TET)酶(例如, 人TET1、人TET2、人TET3、鼠TET1、鼠TET2、鼠TET3、纳氏虫TET(NgTET)、灰盖鬼伞(CcTET)或其变体)、有机硼烷或其它测定)。在一些实施方案中, 试剂盒含有能够以甲基化特异性方式修饰DNA的试剂(例如, 甲基化敏感性限制酶、甲基化依赖性限制酶、十-十一易位(TET)酶(例如, 人TET1、人TET2、人TET3、鼠TET1、鼠TET2、鼠TET3、纳氏虫TET(NgTET)、灰盖鬼伞(CcTET)或其变体)、有机硼烷) 和/或能够检测升高水平的本文所述的蛋白质标志物的试剂。在一些实施方案中, 提供了含有一种或多种对于执行方法所必需、足够或有用的试剂的试剂盒。还提供了含有试剂的反应混合物。还提供了含有多种试剂的预混试剂组, 这些试剂可以相互添加和/或添加至测试样品中以完善反应混合物。

[0036] 在一些实施方案中, 本文所述的技术与可编程机器相关联, 该可编程机器被设计用于执行由本文所述的方法提供的一系列算术或逻辑运算。例如, 这项技术的一些实施方案与计算机软件和/或计算机硬件相关联(例如, 在其中执行)。在一个方面, 这项技术涉及一种计算机, 该计算机包括一种形式的存储器、用于执行算术和逻辑运算的元件以及用于执行一系列指令(例如, 本文提供的方法)以读取、操作和存储数据的处理元件(例如, 微处理器)。在一些实施方案中, 微处理器是系统的一部分, 用于: 确定甲基化状态(例如, 一种或多种DMR, 例如表1A和表2A中提供的DMR 1-198); 比较甲基化状态(例如, 一种或多种DMR, 例如表1A和表2A中提供的DMR 1-198); 生成标准曲线; 确定Ct值; 计算甲基化的分数、频率或百分比(例如, 一种或多种DMR, 例如表1A和表2A中提供的DMR 1-198); 鉴定CpG岛; 确定测定或标志物的特异性和/或灵敏度; 计算ROC曲线和相关的AUC; 序列分析; 皆如本文所述或本领域已知。

[0037] 在一些实施方案中, 微处理器或计算机在算法中使用甲基化状态数据来预测癌症部位。

[0038] 在一些实施方案中, 软件或硬件组件接收多个测定的结果, 并且基于多个测定(例如确定例如如表1B和表2B中提供的多个DMR的甲基化状态的测定)的结果来确定用以向用

户报告的指示癌症风险的单值结果。相关实施方案基于来自例如确定多种标志物(诸如例如如表1A和表2A中提供的多个DMR)的甲基化状态的多个测定的结果的数学组合(例如加权组合、线性组合)计算风险因数。在一些实施方案中,DMR的甲基化状态确定维度并可在多维空间中具有数值,并且由多个DMR的甲基化状态确定的坐标是例如用以向用户报告的例如与癌症风险相关的结果。

[0039] 一些实施方案包括存储介质和存储器组件。存储器组件(例如易失性和/或非易失性存储器)可用于存储指令(例如如本文提供的方法的实施方案)和/或数据(例如工作事项,诸如甲基化测量结果、序列以及与之相关的统计描述)。一些实施方案涉及还包括CPU、图形卡和用户界面中的一者或多者(例如包括输出装置诸如显示器和输入装置诸如键盘)的系统。

[0040] 与技术相关的可编程机器包括常规现存技术性机器和处于开发中或尚有待于开发的技术性机器(例如量子计算机、化学计算机、DNA计算机、光学计算机、基于自旋电子学的计算机等)。

[0041] 在一些实施方案中,技术包括用于传输数据的有线(例如金属电缆、光纤)或无线传输介质。举例来说,一些实施方案涉及经网络(例如局域网(LAN)、广域网(WAN)、自组网络、因特网等)进行的数据传输。在一些实施方案中,可编程机器作为对等机存在于这种网络上,并且在一些实施方案中,可编程机器具有客户端/服务器关系。

[0042] 在一些实施方案中,数据存储于计算机可读存储介质诸如硬盘、闪存存储器、光学介质、软盘等上。

[0043] 在一些实施方案中,本文提供的技术与协同操作以执行如本文所述的方法的多个可编程装置相关联。例如,在一些实施方案中,多台计算机(例如,通过网络连接)可以并行工作以收集和处理数据,例如,在集群计算或网格计算或一些其它依赖于完整计算机(具有板载CPU、存储器、电源、网络接口等)通过常规网络接口(例如以太网、光纤)或无线网络技术连接到网络(私有、公共或因特网)的分布式计算机架构的执行中。

[0044] 例如,一些实施方案提供了包括计算机可读介质的计算机。该实施方案包括耦合到处理器的随机存取存储器(RAM)。处理器执行被存储在存储器中的计算机可执行程序指令。此类处理器可以包括微处理器、ASIC、状态机或其它处理器,并且可以是多种计算机处理器中的任一种,例如来自Santa Clara (California)的Intel公司和Schaumburg (Illinois)的Motorola公司的处理器。此类处理器包括介质(例如计算机可读介质)或者可以与介质通信,所述介质存储指令,当由处理器执行时,该指令使处理器进行本文所述的步骤。

[0045] 计算机可读介质的实施方案包括但不限于能够为处理器提供计算机可读指令的电子、光学、磁性或其它存储或传输装置。合适介质的其它实例包括但不限于软盘、CD-ROM、DVD、磁盘、存储芯片、ROM、RAM、ASIC、配置的处理器、所有光学介质、所有磁带或其它磁性介质,或计算机处理器可以从中读取指令的任何其它介质。此外,各种其它形式的计算机可读介质也可以向计算机传输或传送指令,包括有线和无线的路由器、私有或公共网络或其它传输装置或信道。指令可以包括来自任何合适的计算机编程语言的代码,计算机编程语言包括例如C、C++、C#、Visual Basic、Java、Python、Perl和JavaScript。

[0046] 在一些实施方案中,计算机连接到网络。计算机还可以包括许多外部或内部装置,

诸如鼠标、CD-ROM、DVD、键盘、显示器或其它输入或输出装置。计算机的实例是个人计算机、数字助理、个人数字助理、蜂窝电话、移动电话、智能电话、寻呼机、数字平板电脑、膝上型计算机、因特网设备和其它基于处理器的装置。一般来说，与本文提供的技术的方面有关的计算机可为在能够支持一种或多种包含本文提供的技术的程序的任何操作系统上运行的任何类型的基于处理器的平台，所述操作系统诸如Microsoft Windows、Linux、UNIX、Mac OS X等。一些实施方案包括执行其它应用程序(例如，应用程序)的个人计算机。应用程序可以含在存储器中并且可以包括例如文字处理应用程序、电子表格应用程序、电子邮件应用程序、即时消息应用程序、演示应用程序、因特网浏览器应用程序、日历/组织器应用程序和任何其它能够由客户端装置执行的应用程序。

[0047] 本文所述的与这项技术相关联的所有此类组件、计算机和系统可以是逻辑的或虚拟的。

[0048] 因此，本文提供了与在从受试者获得的样品中筛查PNET的方法相关的技术，所述方法包括测定从受试者获得的样品(例如胰腺组织)(例如血液样品)中标志物的甲基化状态，并且当标志物的甲基化状态不同于在未患PNET的受试者中测定的标志物的甲基化状态时，将所述受试者鉴定为患有PNET，其中标志物包含差异甲基化区域(DMR)中的碱基，所述DMR选自由表1A和表2A中提供的DMR1-198组成的组。

[0049] 在一些实施方案中，其中从受试者获得的样品是组织(例如，胰腺组织)并且以下标志物中的一种或多种的甲基化状态不同于在未患PNET的受试者中测定的所述一种或多种标志物的甲基化状态指示受试者患有PNET: ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B(参见表4、表5A和表5C、实施例I)。

[0050] 在一些实施方案中，其中从受试者获得的样品是血液样品(例如，血浆样品、全血样品、白细胞样品、血清样品)并且以下标志物中的一种或多种的甲基化状态不同于在未患PNET的受试者中测定的所述一种或多种标志物的甲基化状态指示受试者患有PNET: ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B(参见表4、表5A和表5C、实施例I)。

[0051] 在一些实施方案中，其中从受试者获得的样品是血液样品(例如，血浆样品、全血样品、白细胞样品、血清样品)并且以下标志物中的一种或多种的甲基化状态不同于在未患转移性PNET的受试者中测定的所述一种或多种标志物的甲基化状态指示受试者患有转移性PNET: SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBKL2A(参见表5C、实施例I)。

[0052] 在一些实施方案中，其中从受试者获得的样品是血液样品(例如，血浆样品、全血样品、白细胞样品、血清样品)并且以下标志物中的一种或多种的甲基化状态不同于在未患

肺NET的受试者中测定的所述一种或多种标志物的甲基化状态指示受试者患有肺NET：SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBK2A(参见表5C、实施例I)。

[0053] 在一些实施方案中,其中从受试者获得的样品是血液样品(例如,血浆样品、全血样品、白细胞样品、血清样品)并且以下标志物中的一种或多种的甲基化状态不同于在未患小肠NET的受试者中测定的所述一种或多种标志物的甲基化状态指示受试者患有小肠NET：SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBK2A(参见表5C、实施例I)。

[0054] 这项技术涉及鉴定和区分PNET和/或各种形式的NET(例如,肺NET、小肠NET)。一些实施方案提供了包括测定多种标志物的方法,例如,包括测定2至11种到100或120或198种标志物(例如,1-4、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-13、1-14、1-15、1-16、1-17、1-18、1-19、1-20、1-25、1-50、1-75、1-100、1-150、1-198种)(例如,2-4、2-6、2-7、2-8、2-9、2-10、2-11、2-12、2-13、2-14、2-15、2-16、2-17、2-18、2-19、2-20、2-25、2-50、2-75、2-100、2-198种)(例如,3-4、3-6、3-7、3-8、3-9、3-10、3-11、3-12、3-13、3-14、3-15、3-16、3-17、3-18、3-19、3-20、3-25、3-50、3-75、3-100、3-198种)(例如,4-5、4-6、4-7、4-8、4-9、4-10、4-11、4-12、4-13、4-14、4-15、4-16、4-17、4-18、4-19、4-20、4-25、4-50、4-75、4-100、4-198种)(例如,5-6、5-7、5-8、5-9、5-10、5-11、5-12、5-13、5-14、5-15、5-16、5-17、5-18、5-19、5-20、5-25、5-50、5-75、5-100、5-198种)。

[0055] 这项技术在评估的甲基化状态方面不受限制。在一些实施方案中,评估样品中标志物的甲基化状态包括确定一种碱基的甲基化状态。在一些实施方案中,测定样品中标志物的甲基化状态包括确定多种碱基的甲基化程度。此外,在一些实施方案中,标志物的甲基化状态包含相对于标志物的正常甲基化状态增加的标志物甲基化。在一些实施方案中,标志物的甲基化状态包含相对于标志物的正常甲基化状态减少的标志物甲基化。在一些实施方案中,标志物的甲基化状态包含相对于标志物的正常甲基化状态不同的标志物甲基化模式。

[0056] 此外,在一些实施方案中,标志物是100个或更少碱基的区域,标志物是500个或更少碱基的区域,标志物是1000个或更少碱基的区域,标志物是5000个或更少碱基的区域,或在一些实施方案中,标志物是一个碱基。在一些实施方案中,标志物在高CpG密度启动子中。

[0057] 这项技术不受样品类型的限制。例如,在一些实施方案中,样品是粪便样品、组织样品(例如,胰腺组织样品)、血液样品(例如,血浆、白细胞、血清、全血)、排泄物或尿液样品。

[0058] 此外,这项技术在用于确定甲基化状态的方法方面不受限制。在一些实施方案中,测定包括使用甲基化特异性聚合酶链式反应、核酸测序、质谱法、甲基化特异性核酸酶、基于质量的分离或靶标捕获。在一些实施方案中,测定包括使用甲基化特异性寡核苷酸。在一些实施方案中,这项技术使用大规模平行测序(例如,下一代测序)来确定甲基化状态,例

如,边合成边测序、实时(例如,单分子)测序、珠粒乳液测序(bead emulsion sequencing)、纳米孔测序等。

[0059] 这项技术提供了用于检测DMR的试剂,例如,在一些实施方案中提供了一组寡核苷酸,其包含由SEQ ID NO:1-66提供的序列(参见表3)。在一些实施方案中,提供了包含与具有DMR中的碱基的染色体区域互补的序列的寡核苷酸,例如对DMR的甲基化状态敏感的寡核苷酸。

[0060] 这项技术提供了用于鉴定PNET的各种标志物组,例如在一些实施方案中,标志物包含具有以下注释的染色体区域:ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBK2A、PNMAL2、S1PR4_A、LGALS3和MYO15B(参见表4、表5A和表5C、实施例I)。

[0061] 提供了试剂盒实施方案,例如包含以下的试剂盒:能够以甲基化特异性方式修饰DNA的试剂(例如,甲基化敏感性限制酶、甲基化依赖性限制酶、十-十一易位(TET)酶(例如,人TET1、人TET2、人TET3、鼠TET1、鼠TET2、鼠TET3、纳氏虫TET(NgTET)、灰盖鬼伞(CcTET)或其变体)、有机硼烷);和包含来自DMR 1-198(来自表1A和表2A)的一个或多个序列并且具有与未患癌症的受试者相关的甲基化状态的对照核酸。在一些实施方案中,试剂盒包含亚硫酸氢盐试剂和如本文所述的寡核苷酸。在一些实施方案中,试剂盒包含能够以甲基化特异性方式修饰DNA的试剂(例如,甲基化敏感性限制酶、甲基化依赖性限制酶、十-十一易位(TET)酶(例如,人TET1、人TET2、人TET3、鼠TET1、鼠TET2、鼠TET3、纳氏虫TET(NgTET)、灰盖鬼伞(CcTET)或其变体)、有机硼烷);和包含来自DMR 1-198(来自表1A和表2A)的一个或多个序列并且具有与患有特定类型的癌症的受试者相关的甲基化状态的对照核酸。一些试剂盒实施方案包括用于从受试者获得样品(例如,粪便样品;组织样品;血浆样品;血清样品;全血样品)的样品收集器;能够以甲基化特异性方式修饰DNA的试剂(例如,甲基化敏感性限制酶、甲基化依赖性限制酶、十-十一易位(TET)酶(例如,人TET1、人TET2、人TET3、鼠TET1、鼠TET2、鼠TET3、纳氏虫TET(NgTET)、灰盖鬼伞(CcTET)或其变体)、有机硼烷);以及如本文所述的寡核苷酸。

[0062] 这项技术涉及组合物(例如,反应混合物)的实施方案。在一些实施方案中,提供了一种组合物,其包含:包含DMR的核酸和能够以甲基化特异性方式修饰DNA的试剂(例如,甲基化敏感性限制酶、甲基化依赖性限制酶、十-十一易位(TET)酶(例如,人TET1、人TET2、人TET3、鼠TET1、鼠TET2、鼠TET3、纳氏虫TET(NgTET)、灰盖鬼伞(CcTET)或其变体)、有机硼烷)。一些实施方案提供了一种组合物,其包含:包含DMR的核酸和如本文所述的寡核苷酸。一些实施方案提供了一种组合物,其包含:包含DMR的核酸和甲基化敏感性限制酶。一些实施方案提供了一种组合物,其包含:包含DMR的核酸和聚合酶。

[0063] 提供了用于在从受试者获得的样品(例如,胰腺组织样品;血液样品;粪便样品)中筛查PNET的额外相关方法实施方案,例如包括以下的方法:确定样品中的标志物的甲基化状态,所述标志物包含作为DMR 1-198(来自表1A和表2A)中的一个或多个的DMR中的碱基;将来自受试者样品的标志物的甲基化状态与来自未患PNET的受试者的正常对照样品的标志物的甲基化状态进行比较;以及确定受试者样品和正常对照样品的甲基化状态差异的置信区间和/或p值。在一些实施方案中,置信区间为90%、95%、97.5%、98%、99%、99.5%、

99.9%或99.99%，并且p值为0.1、0.05、0.025、0.02、0.01、0.005、0.001或0.0001。方法的一些实施方案提供了以下步骤：使包含DMR的核酸与能够以甲基化特异性方式修饰核酸的试剂（例如，甲基化敏感性限制酶、甲基化依赖性限制酶、十-十一易位（TET）酶（例如，人TET1、人TET2、人TET3、鼠TET1、鼠TET2、鼠TET3、纳氏虫TET（NgTET）、灰盖鬼伞（CcTET）或其变体）、有机硼烷）反应，以产生例如以甲基化特异性方式修饰的核酸；对以甲基化特异性方式修饰的核酸进行测序，以提供以甲基化特异性方式修饰的核酸的核苷酸序列；将以甲基化特异性方式修饰的核酸的核苷酸序列与来自未患特定类型的癌症的受试者的包含DMR的核酸的核苷酸序列进行比较，以鉴定这两个序列的差异；以及当存在差异时，将受试者鉴定为患有PNET（例如，PNET和/或一种形式的NET：肺NET、小肠NET）。

[0064] 这项技术提供了用于在从受试者获得的样品中筛查PNET的系统。系统的示例性实施方案包括例如用于在从受试者获得的样品（例如胰腺组织样品；血浆样品；粪便样品）中筛查PNET和/或相关NET类型（例如，肺NET、小肠NET）的系统，该系统包括：分析组件，所述分析组件被配置用于确定样品的甲基化状态；软件组件，所述软件组件被配置用于将样品的甲基化状态与数据库中记录的对照样品或参考样品的甲基化状态进行比较；以及警报组件，所述警报组件被配置用于向用户发出关于PNET相关的甲基化状态的警报。在一些实施方案中，警报由软件组件确定，所述软件组件接收多个测定的结果（例如，确定多种标志物，如例如表1A和表2A中提供的DMR的甲基化状态的测定）并基于多个结果计算值或结果以进行报告。一些实施方案提供与本文提供的每个DMR相关联的加权参数的数据库，用于计算值或结果和/或警报以向用户（例如，如医师、护士、临床医生等）报告。在一些实施方案中，报告了来自多个测定的所有结果，并且在一些实施方案中，一个或多个结果用于提供评分、值或结果，所述评分、值或结果基于来自多个测定的一个或多个结果的复合物，其指示受试者中的癌症风险。

[0065] 在系统的一些实施方案中，样品包含含有DMR的核酸。在一些实施方案中，该系统还包括用于分离核酸的组件、用于收集样品的组件，例如用于收集粪便样品的组件。在一些实施方案中，系统包含含有DMR的核酸序列。在一些实施方案中，数据库包含来自未患PNET和/或相关NET类型（例如，肺NET、小肠NET）的受试者的核酸序列。还提供了核酸，例如一组核酸，每个核酸具有包含DMR的序列。在一些实施方案中，该组核酸中每个核酸具有来自未患PNET和/或相关NET类型（例如，肺NET、小肠NET）的受试者的序列。相关系统实施方案包括如所述的一组核酸和与该组核酸相关的核酸序列数据库。一些实施方案还包括能够以甲基化特异性方式修饰DNA的试剂（例如，甲基化敏感性限制酶、甲基化依赖性限制酶、十-十一易位（TET）酶（例如，人TET1、人TET2、人TET3、鼠TET1、鼠TET2、鼠TET3、纳氏虫TET（NgTET）、灰盖鬼伞（CcTET）或其变体）、有机硼烷）。一些实施方案还包括核酸测序仪。

[0066] 在某些实施方案中，提供了用于表征来自人类患者的样品（例如，胰腺组织样品；血液样品；粪便样品）的方法。例如，在一些实施方案中，此类实施方案包括从人类患者的样品获得DNA；测定DNA甲基化标志物的甲基化状态，所述DNA甲基化标志物包含差异甲基化区域（DMR）中的碱基，所述DMR选自由表1A和表2A的DMR 1-198组成的组；以及将所述一种或多种DNA甲基化标志物的所测定的甲基化状态与未患PNET和/或相关NET类型（例如，肺NET、小肠NET）的人类患者的所述一种或多种DNA甲基化标志物的甲基化水平参考进行比较。

[0067] 此类方法不限于来自人类患者的特定类型的样品。在一些实施方案中，样品是胰

腺组织样品。在一些实施方案中,样品是血浆样品。在一些实施方案中,样品是粪便样品、组织样品、胰腺组织样品、血液样品(例如,白细胞样品、血浆样品、全血样品、血清样品)或尿液样品。

[0068] 在一些实施方案中,此类方法包括测定多种DNA甲基化标志物(例如,1-4、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-13、1-14、1-15、1-16、1-17、1-18、1-19、1-20、1-25、1-50、1-75、1-100、1-150、1-198种)(例如,2-4、2-6、2-7、2-8、2-9、2-10、2-11、2-12、2-13、2-14、2-15、2-16、2-17、2-18、2-19、2-20、2-25、2-50、2-75、2-100、2-198种)(例如,3-4、3-6、3-7、3-8、3-9、3-10、3-11、3-12、3-13、3-14、3-15、3-16、3-17、3-18、3-19、3-20、3-25、3-50、3-75、3-100、3-198种)(例如,4-5、4-6、4-7、4-8、4-9、4-10、4-11、4-12、4-13、4-14、4-15、4-16、4-17、4-18、4-19、4-20、4-25、4-50、4-75、4-100、4-198种)(例如,5-6、5-7、5-8、5-9、5-10、5-11、5-12、5-13、5-14、5-15、5-16、5-17、5-18、5-19、5-20、5-25、5-50、5-75、5-100、5-198种)。在一些实施方案中,此类方法包括测定2至11种DNA甲基化标志物。在一些实施方案中,此类方法包括测定12至120种DNA甲基化标志物。在一些实施方案中,此类方法包括测定2至198种DNA甲基化标志物。在一些实施方案中,此类方法包括测定样品中的一种或多种DNA甲基化标志物的甲基化状态,包括确定一种碱基的甲基化状态。在一些实施方案中,此类方法包括测定样品中的一种或多种DNA甲基化标志物的甲基化状态,包括确定多种碱基处的甲基化程度。在一些实施方案中,此类方法包括测定正向链的甲基化状态或测定反向链的甲基化状态。

[0069] 在一些实施方案中,DNA甲基化标志物是100个或更少碱基的区域。在一些实施方案中,DNA甲基化标志物是500个或更少碱基的区域。在一些实施方案中,DNA甲基化标志物是1000个或更少碱基的区域。在一些实施方案中,DNA甲基化标志物是5000个或更少碱基的区域。在一些实施方案中,DNA甲基化标志物是一个碱基。在一些实施方案中,DNA甲基化标志物在高CpG密度启动子中。

[0070] 在一些实施方案中,测定包括使用甲基化特异性聚合酶链式反应、核酸测序、质谱法、甲基化特异性核酸酶、基于质量的分离或靶标捕获。

[0071] 在一些实施方案中,测定包括使用甲基化特异性寡核苷酸。在一些实施方案中,甲基化特异性寡核苷酸选自自由SEQ ID NO:1-66(表3)组成的组。

[0072] 在一些实施方案中,具有选自自由ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSP0、CUX1、FAM78A、FNBP1、IER2、MOBK12A、PNMAL2、S1PR4_A、LGALS3和MYO15B(参见表4、表5A和表5C、实施例I)组成的组的注释的染色体区域构成DNA甲基化标志物。

[0073] 在一些实施方案中,此类方法包括确定两种DNA甲基化标志物的甲基化状态。在一些实施方案中,此类方法包括确定在表1A和表2A的一行中提供的一对DNA甲基化标志物的甲基化状态。

[0074] 在某些实施方案中,这项技术提供了用于表征从人类患者获得的样品(例如,胰腺组织样品;白细胞样品;血浆样品;全血样品;血清样品;粪便样品)的方法。在一些实施方案中,此类方法包括确定样品中的DNA甲基化标志物的甲基化状态,所述DNA甲基化标志物包含选自自由表1A和表2A的DMR 1-198组成的组的DMR中的碱基;将来自患者样品的DNA甲基化

标志物的甲基化状态与来自未患PNET和/或相关NET类型(例如,肺NET、小肠NET)的人类受试者的正常对照样品的DNA甲基化标志物的甲基化状态进行比较;以及确定人类患者和正常对照样品的甲基化状态差异的置信区间和/或p值。在一些实施方案中,置信区间为90%、95%、97.5%、98%、99%、99.5%、99.9%或99.99%,并且p值为0.1、0.05、0.025、0.02、0.01、0.005、0.001或0.0001。

[0075] 在某些实施方案中,这项技术提供了用于表征从人类受试者获得的样品(例如,胰腺组织样品;白细胞样品;血浆样品;全血样品;血清样品;粪便样品)的方法,所述方法包括使包含DMR的核酸与能够以甲基化特异性方式修饰DNA的试剂(例如,甲基化敏感性限制酶、甲基化依赖性限制酶和亚硫酸氢盐试剂)反应,以产生以甲基化特异性方式修饰的核酸;对以甲基化特异性方式修饰的核酸进行测序,以提供以甲基化特异性方式修饰的核酸的核苷酸序列;将以甲基化特异性方式修饰的核酸的核苷酸序列与来自未患PNET的受试者的包含DMR的核酸的核苷酸序列进行比较,以鉴定这两个序列的差异。

[0076] 在某些实施方案中,这项技术提供了用于表征从人类受试者获得的样品(例如,胰腺组织样品;血浆样品;粪便样品)的系统,所述系统包括:分析组件,所述分析组件被配置用于确定样品的甲基化状态;软件组件,所述软件组件被配置用于将样品的甲基化状态与数据库中记录的对照样品或参考样品的甲基化状态进行比较;以及警报组件,所述警报组件被配置用于基于甲基化状态的组合确定单个值并向用户发出关于PNET相关的甲基化状态的警报。在一些实施方案中,样品包含含有DMR的核酸。

[0077] 在一些实施方案中,此类系统还包括用于分离核酸的组件。在一些实施方案中,此类系统还包括用于收集样品的组件。

[0078] 在一些实施方案中,样品是粪便样品、组织样品、胰腺组织样品、血液样品(例如,血浆样品、白细胞样品、全血样品、血清样品)或尿液样品。

[0079] 在一些实施方案中,数据库包含含有DMR的核酸序列。在一些实施方案中,数据库包含来自未患PNET的受试者的核酸序列。

[0080] 基于本文所含的教导内容,额外实施方案对于相关领域的技术人员来说将是显而易见的。

[0081] 定义

[0082] 为促进对本发明技术的理解,以下定义多个术语和短语。另外的定义在整个详细描述中阐述。

[0083] 在整个说明书和权利要求书中,除非上下文另有明确规定,否则以下术语采用与本文明确相关的含义。如本文所用的短语“在一个实施方案中”不一定指同一实施方案,但是可以指同一实施方案。此外,如本文所用的短语“在另一个实施方案中”不一定指不同的实施方案,但是可以指不同的实施方案。因此,如下所述,可以容易地组合本发明的各种实施方案,而不脱离本发明的范围或精神。

[0084] 另外,如本文所用,除非上下文另有明确规定,否则术语“或”是包括性的“或”运算符并且等同于术语“和/或”。除非上下文另有明确规定,否则术语“基于”不是排他性的,并且允许基于未描述的另外的因素。另外,在整个说明书中,“一个(种)(a/an)”和“这个(种)(the)”的含义包括复数个(种)指示物。“在……中”的含义包括“在……中”和“在……上”。

[0085] 如本申请中的权利要求中使用的过渡短语“基本上由……组成”将权利要求的范

围限制为指定的材料或步骤“以及不实质性影响要求保护的发明的基本和新颖特征的那些”，如In re Herz, 537F.2d 549, 551-52, 190USPQ 461, 463 (CCPA 1976)中所讨论的。举例来说，“基本上由所列举要素组成的”组合物可以含有一定水平的未列举的污染物，使得污染物尽管存在，但与纯组合物，即“由所列举组分组成”的组合物相比，不会改变所列举组合物的功能。

[0086] 如本文所用，“核酸”或“核酸分子”一般是指任何核糖核酸或脱氧核糖核酸，其可以是未修饰的或修饰的DNA或RNA。“核酸”包括不限于单链和双链核酸。如本文所用，术语“核酸”还包括含有一个或多个修饰碱基的如上所述的DNA。因此，具有出于稳定性或其它原因而进行修饰的主链的DNA是“核酸”。如本文所用，术语“核酸”涵盖这样的化学、酶促或代谢修饰形式的核酸，以及病毒和细胞（包括例如简单和复杂细胞）特征性的DNA化学形式。

[0087] 术语“寡核苷酸”或“多核苷酸”或“核苷酸”或“核酸”是指具有两个或更多个，优选地超过三个和通常超过十个脱氧核糖核苷酸或核糖核苷酸的分子。确切的大小将取决于许多因素，而这些因素又取决于寡核苷酸的最终功能或用途。寡核苷酸可以通过任何方式产生，包括化学合成、DNA复制、逆转录或它们的组合。DNA的典型脱氧核糖核苷酸是胸腺嘧啶、腺嘌呤、胞嘧啶和鸟嘌呤。RNA的典型核糖核苷酸是尿嘧啶、腺嘌呤、胞嘧啶和鸟嘌呤。

[0088] 如本文所用，术语核酸的“基因座”或“区域”是指核酸的亚区域，例如染色体上的基因、单核苷酸、CpG岛等。

[0089] 术语“互补”和“互补性”是指通过碱基配对规则相关的核苷酸（例如1个核苷酸）或多核苷酸（例如核苷酸序列）。例如，序列5'-A-G-T-3'与序列3'-T-C-A-5'互补。互补性可以是“部分的”，其中仅一些核酸碱基根据碱基配对规则匹配。或者，核酸之间可能存在“完全”或“全部”互补。核酸链之间的互补程度影响核酸链之间杂交的效率和强度。这在依赖于核酸之间的结合的扩增反应和检测方法中特别重要。

[0090] 术语“基因”是指包含产生RNA或多肽或其前体所必需的编码序列的核酸（例如DNA或RNA）序列。功能性多肽可由全长编码序列或由编码序列的任何部分编码，只要保留多肽的所需活性或功能特性（例如酶活性、配体结合、信号转导等）即可。当在提及基因时使用，术语“部分”是指该基因的片段。片段的大小可以在几个核苷酸至整个基因序列减去一个核苷酸的范围内变化。因此，“包含基因至少一部分的核苷酸”可以包含基因片段或整个基因。

[0091] 术语“基因”还涵盖结构基因的编码区，并且包括在5'末端与3'末端两者上邻近于编码区定位的序列，例如在任一末端上持续约1kb的距离，以致基因对应于全长mRNA（例如包含编码序列、调控序列、结构序列和其他序列）的长度。位于编码区5'并且存在于mRNA上的序列称为5'非翻译或未翻译序列。位于编码区3'或下游并且存在于mRNA上的序列称为3'非翻译或3'未翻译序列。术语“基因”涵盖基因的cDNA与基因组形式。在一些生物体（例如真核生物）中，基因的基因组形式或克隆含有被称为“内含子”或“插入区”或“插入序列”的非编码序列中断的编码区。内含子是转录成核RNA（hnRNA）的基因区段；内含子可以含有调控元件，例如增强子。内含子从核转录本或初级转录本中移出或“剪除”；因此，在信使RNA（mRNA）转录本中不存在内含子。mRNA在翻译期间起对初生多肽中的氨基酸序列或顺序进行指定的作用。

[0092] 除含有内含子外，基因的基因组形式还可以包括位于RNA转录本上存在的序列的

5'和3'末端的序列。这些序列被称为“侧接”序列或区域(这些侧接序列位于mRNA转录本上存在的非翻译序列的5'或3')。5'侧接区域可以含有调控序列,例如启动子和强化子,其控制或影响基因的转录。3'侧接区域可以含有指导转录终止、转录后裂解和多腺苷酸化的序列。

[0093] 在提及基因时术语“野生型”是指具有从天然存在的来源分离的基因特征的基因。在提及基因产物时术语“野生型”是指具有从天然存在的来源分离的基因产物特征的基因产物。如应用于物体的术语“天然存在”是指可以在自然界中发现物体的事实。例如,存在于生物体(包括病毒)中的可以从自然界来源分离并且未经实验室人员有意修饰的多肽或多核苷酸序列是天然存在的。野生型基因通常是在群体中最常观察到的基因或等位基因,因此被任意指定为基因的“正常”或“野生型”形式。相比之下,在提及基因或基因产物时术语“修饰的”或“突变的”分别是指与野生型基因或基因产物相比显示序列和/或功能特性修饰(例如特征改变)的基因或基因产物。注意,可以分离天然存在的突变体;这些通过它们与野生型基因或基因产物相比特征改变的事实来鉴定。

[0094] 术语“等位基因”是指基因的变异;所述变异包括但不限于变体和突变体、多态基因座和单核苷酸多态基因座、移码和剪接突变。等位基因可能在群体中天然存在,或者其可能在群体中任何特定个体的生命期间出现。

[0095] 因此,在提及核苷酸序列时使用术语“变体”和“突变体”是指与另一个通常相关的核苷酸序列相差一个或多个核苷酸的核酸序列。“变异”是两个不同核苷酸序列之间的差异;典型地,一个序列是参考序列。

[0096] “扩增”是涉及模板特异性的核酸复制的特定情况。其与非特异性模板复制(例如依赖于模板但不依赖于特定模板的复制)形成对比。在这里,模板特异性不同于复制保真度(例如合成适当多核苷酸序列)和核苷酸(核糖核苷酸或脱氧核糖核苷酸)特异性。模板特异性常根据“靶标”特异性来描述。在设法从其它核酸中分选出来的意义上,靶标序列是“靶标”。扩增技术主要是为这种分选而设计的。

[0097] 在核酸背景下术语“扩增(amplifying或amplification)”是指典型地从少量多核苷酸(例如单个多核苷酸分子)开始,产生多拷贝的多核苷酸或多核苷酸的一部分,其中扩增产物或扩增子一般是可检测的。多核苷酸的扩增涵盖多种化学和酶促过程。在聚合酶链式反应(PCR)或连接酶链式反应(LCR;参见,例如美国专利号5,494,810;通过引用整体并入本文)期间从靶或模板DNA分子的一个或几个拷贝产生多个DNA拷贝是扩增形式。另外的扩增类型包括但不限于等位基因特异性PCR(参见,例如美国专利号5,639,611;通过引用整体并入本文)、组装PCR(参见,例如美国专利号5,965,408;通过引用整体并入本文)、解旋酶依赖性扩增(参见,例如美国专利号7,662,594;通过引用整体并入本文)、热启动PCR(参见,例如美国专利号5,773,258和5,338,671;各自通过引用整体并入本文)、序列间特异性PCR、反向PCR(参见,例如,Triglia等人(1988) *Nucleic Acids Res.*, 16:8186;通过引用整体并入本文)、连接介导的PCR(参见,例如Guilfoyle, R.等人, *Nucleic Acids Research*, 25:1854-1858 (1997);美国专利号5,508,169;各自通过引用整体并入本文)、甲基化特异性PCR(参见,例如Herman等人, (1996) *PNAS* 93 (13) 9821-9826;通过引用整体并入本文)、微小引物PCR、多重连接依赖性探针扩增(参见,例如Schouten等人, (2002) *Nucleic Acids Research* 30 (12):e57;通过引用整体并入本文)、多重PCR(参见,例如,Chamberlain等人, (1988)

Nucleic Acids Research 16 (23) 11141-11156;Ballabio等人,(1990) Human Genetics 84 (6) 571-573;Hayden等人,(2008) BMC Genetics 9:80;各自通过引用整体并入本文)、巢式PCR、重叠延伸PCR(参见,例如Higuchi等人,(1988) Nucleic Acids Research 16 (15) 7351-7367;通过引用整体并入本文)、实时PCR(参见,例如,Higuchi等人,(1992) Biotechnology 10:413-417;Higuchi等人,(1993) Biotechnology 11:1026-1030;各自通过引用整体并入本文)、逆转录PCR(参见,例如Bustin,S.A.(2000) J.Molecular Endocrinology 25:169-193;通过引用整体并入本文)、固相PCR、热不对称交错PCR和降落PCR(参见,例如Don等人,Nucleic Acids Research(1991) 19 (14) 4008;Roux,K.(1994) Biotechniques 16 (5) 812-814;Hecker等人,(1996) Biotechniques 20 (3) 478-485;各自通过引用整体并入本文)。也可以使用数字PCR完成多核苷酸扩增(参见,例如Kalinina等人,Nucleic Acids Research. 25;1999-2004,(1997);Vogelstein和Kinzler,Proc Natl Acad Sci USA. 96; 9236-41,(1999);国际专利公开号W005023091A2;美国专利申请公开号20070202525;各自通过引用整体并入本文)。

[0098] 术语“聚合酶链式反应”(“PCR”)是指K.B.Mullis美国专利号4,683,195、4,683,202和4,965,188的方法,这些专利描述了在无需克隆或纯化的情况下增加基因组或其它DNA或RNA混合物中靶标序列的区段浓度的方法。该扩增靶标序列的过程由以下组成:将大量过量的两种寡核苷酸引物引入到含有所需靶标序列的DNA混合物中,接着在DNA聚合酶的存在下进行一系列精确的热循环。这两个引物与其各自的双链靶标序列链互补。为了实现扩增,将混合物变性,然后将引物退火为靶分子内的互补序列。在退火之后,将引物用聚合酶延伸以便形成一对新的互补链。变性、引物退火和聚合酶延伸的步骤可以重复多次(即,变性、退火和延伸构成一个“循环”;可以存在多个“循环”)以获得高浓度的所需靶标序列的扩增区段。所需靶标序列的扩增区段的长度由引物彼此之间的相对位置决定,因此,该长度是一个可控参数。由于该过程的重复方面,将该方法称为“聚合酶链式反应”(“PCR”)。因为靶标序列的所需扩增区段成为混合物中的主要序列(就浓度而言),所以将它们称为“经过PCR扩增的”并且是“PCR产物”或“扩增子”。本领域的技术人员将理解,术语“PCR”涵盖使用例如实时PCR、巢式PCR、逆转录PCR(RT-PCR)、单引物和随机引物PCR等最初描述的方法的许多变体。

[0099] 大部分扩增技术是通过选择酶来实现模板特异性的。扩增酶是在使用它们的条件下将仅加工核酸的异质混合物中的特定核酸序列的酶。例如,在Q- β 复制酶的情况下,MDV-1RNA是复制酶的特定模板(Kacian等人,Proc.Natl.Acad.Sci.USA,69:3038[1972])。其它核酸不会被这种扩增酶复制。类似地,在T7 RNA聚合酶的情况下,这种扩增酶对其自身的启动子具有严格的特异性(Chamberlin等人,Nature,228:227[1970])。在T4 DNA连接酶的情况下,酶不会连接两种寡核苷酸或多核苷酸,其中寡核苷酸或多核苷酸底物与模板之间在连接点处存在错配(Wu和Wallace(1989) Genomics 4:560)。最终,发现依赖于热稳定模板的DNA聚合酶(例如Taq和Pfu DNA聚合酶)由于它们能够在高温下发挥作用,而对由引物限制并因此由引物界定的序列显示高特异性;高温产生有利于引物与靶标序列杂交而不利于与非靶标序列杂交的热力学条件(H.A.Erlich(编辑),PCR Technology,Stockton Press [1989])。

[0100] 如本文所用,术语“核酸检测测定”是指确定感兴趣核酸的核苷酸组成的任何方

法。核酸检测测定包括但不限于DNA测序法、探针杂交法、结构特异性裂解测定(例如INVADER测定(Hologic, Inc.)并且例如在美国专利号5,846,717、5,985,557、5,994,069、6,001,567、6,090,543和6,872,816;Lyamichev等人, Nat. Biotech., 17:292(1999); Hall等人, PNAS, USA, 97:8272(2000)和美国专利号9,096,893中有描述,其各自通过引用整体并入本文用于所有目的);酶错配裂解法(例如, Variagenics, 美国专利号6,110,684、5,958,692、5,851,770,通过引用整体并入本文);如上所述的聚合酶链式反应(PCR);支化杂交法(例如, Chiron, 美国专利号5,849,481、5,710,264、5,124,246和5,624,802,通过引用整体并入本文);滚环复制(例如, 美国专利号6,210,884、6,183,960和6,235,502,通过引用整体并入本文);NASBA(例如, 美国专利号5,409,818,通过引用整体并入本文);分子信标技术(例如, 美国专利号6,150,097,通过引用整体并入本文);电子传感器技术(Motorola, 美国专利号6,248,229、6,221,583、6,013,170和6,063,573,通过引用整体并入本文);环状探针技术(例如, 美国专利号5,403,711、5,011,769和5,660,988,通过引用整体并入本文);Dade Behring信号扩增法(例如, 美国专利号6,121,001、6,110,677、5,914,230、5,882,867和5,792,614,通过引用整体并入本文);连接酶链式反应(例如, Baranay Proc. Natl. Acad. Sci USA 88,189-93(1991));以及夹心杂交法(例如, 美国专利号5,288,609,通过引用整体并入本文)。

[0101] 术语“可扩增核酸”是指可以通过任何扩增方法扩增的核酸。预期“可扩增核酸”通常包含“样品模板”。

[0102] 术语“样品模板”是指源自针对“靶标”(定义如下)的存在进行分析的样品的核酸。相反,“背景模板”用于指样品中可能存在或可能不存在的样品模板以外的核酸。背景模板通常是无意的。它可能是交叉污染的结果,或者可能是由于存在试图从样品中净化掉的核酸污染物。例如,来自生物体的除待检测核酸之外的核酸可能作为测试样品中的背景存在。

[0103] 术语“引物”是指当置于诱导与核酸模板链互补的引物延伸产物合成的条件下(例如在核苷酸和诱导剂如DNA聚合酶存在下以及在合适温度和pH值下)时能够充当合成起始点的寡核苷酸,无论天然存在,如例如来自限制消化物的核酸片段,还是合成产生。引物优选地是单链的以在扩增中获得最大效率,但可以替代地是双链的。如果是双链的,则首先处理引物以将其链分开,然后再用于制备延伸产物。优选地,引物是寡脱氧核糖核苷酸。引物必须足够长以在诱导剂存在下引发延伸产物的合成。引物的确切长度取决于许多因素,包括温度、引物来源和方法的使用。

[0104] 术语“探针”是指寡核苷酸(例如,核苷酸序列),无论是在纯化的限制消化物中天然存在的还是合成、重组或通过PCR扩增产生的,其都能够与另一感兴趣的寡核苷酸杂交。探针可以是单链的或双链的。探针可用于检测、鉴定和分离特定基因序列(例如,“捕获探针”)。考虑到本发明中使用的任何探针在一些实施方案中都可以用任何“报告分子”标记,使其在任何检测系统中都可检测到,包括但不限于酶(例如ELISA,以及基于酶的组织化学测定)、荧光、放射性和发光系统。本发明并不旨在限于任何特定的检测系统或标记。

[0105] 如本文所用,术语“靶标”是指试图例如通过探针结合、扩增、分离、捕获等从其它核酸中分选出来的核酸。例如,当关于聚合酶链式反应使用时,“靶标”是指用于聚合酶链式反应的引物所结合的核酸区域,而当用于其中靶标DNA未被扩增的测定时,例如在侵入性裂解测定的一些实施方案中,靶标包含探针和侵入性寡核苷酸(例如,INVADER寡核苷酸)结合

形成侵入性裂解结构的位点,使得可以检测到靶标核酸的存在。“区段”定义为靶标序列内的核酸区域。

[0106] 因此,如本文所用,“非靶标”,例如,当它用于描述核酸诸如DNA时,是指可能存在于反应中但不是通过反应检测或表征的主题的核酸。在一些实施方案中,非靶标核酸可以指样品中存在的例如不含靶标序列的核酸,而在一些实施方案中,非靶标可以指外源核酸,即不是来源于含有或怀疑含有靶标核酸的样品,是添加到反应中,例如,使酶(例如聚合酶)的活性归一化以减少反应中酶性能的可变性的核酸。如本文所用,“甲基化”是指胞嘧啶的C5或N4位置的胞嘧啶甲基化、腺嘌呤的N6位置甲基化或其它类型的核酸甲基化。体外扩增的DNA通常是未甲基化的,因为典型的体外DNA扩增方法不保留扩增模板的甲基化模式。然而,“未甲基化DNA”或“甲基化DNA”也可以分别指原始模板未甲基化或甲基化的扩增DNA。

[0107] 如本文所用,术语“扩增试剂”是指除引物、核酸模板和扩增酶以外,扩增所需的那些试剂(脱氧核糖核苷三磷酸、缓冲液等)。通常,扩增试剂连同其它反应组分一起放置并容纳在反应容器中。

[0108] 如本文所用,术语“对照”在关于核酸检测或分析使用时是指具有已知特征(例如,已知序列、已知每个细胞的拷贝数),用于与实验靶标(例如,未知浓度的核酸)比较的核酸。对照可以是内源性的,优选不变的基因,测定中的试验或靶标核酸可以针对该基因进行归一化。此类归一化对照针对可能在例如样品处理、测定效率等中出现的样品间变异,并允许准确的样品间数据比较。可用于归一化对人类样品的核酸检测测定的基因包括例如β-肌动蛋白、ZDHHC1和B3GALT6(参见,例如,美国专利申请序列号14/966,617和62/364,082,各自通过引用并入本文)。

[0109] 对照也可以是外部的。例如,在诸如qPCR、QuARTS等定量测定中,“校准物”或“校准对照”是已知序列,例如具有与实验靶标核酸的一部分相同的序列,并且已知浓度或一系列浓度的核酸(例如,用于在定量PCR中产生校准曲线的连续稀释对照靶标)。通常,使用与实验DNA相同的试剂和反应条件分析校准对照。在某些实施方案中,校准物的测量与实验测定同时进行,例如在同一热循环仪中进行。在优选的实施方案中,可以在单个质粒中包括多个校准物,使得可以容易地以等摩尔量提供不同的校准物序列。在特别优选的实施方案中,例如用一种或多种限制酶消化质粒校准物,以从质粒载体中释放校准物部分。参见,例如WO 2015/066695,其通过引用包括在本文中。

[0110] 如本文所用,“ZDHHC1”是指编码描述为含锌指DHHC型1的蛋白质,位于人DNA中的16号染色体(16q22.1)上并且属于DHHC棕榈酰转移酶家族的基因。如本文所用,“甲基化”是指胞嘧啶的C5或N4位置的胞嘧啶甲基化、腺嘌呤的N6位置甲基化或其它类型的核酸甲基化。体外扩增的DNA通常是未甲基化的,因为典型的体外DNA扩增方法不保留扩增模板的甲基化模式。然而,“未甲基化DNA”或“甲基化DNA”也可以分别指原始模板未甲基化或甲基化的扩增DNA。

[0111] 如本文所用,“甲基化”是指胞嘧啶的C5或N4位置的胞嘧啶甲基化、腺嘌呤的N6位置甲基化或其它类型的核酸甲基化。体外扩增的DNA通常是未甲基化的,因为典型的体外DNA扩增方法不保留扩增模板的甲基化模式。然而,“未甲基化DNA”或“甲基化DNA”也可以分别指原始模板未甲基化或甲基化的扩增DNA。

[0112] 因此,如本文所用,“甲基化核苷酸”或“甲基化核苷酸碱基”是指在核苷酸碱基上

存在甲基部分,其中该甲基部分不存在于公认的典型核苷酸碱基中。例如,胞嘧啶在其嘧啶环上不含甲基部分,但5-甲基胞嘧啶在其嘧啶环的位置5处含有甲基部分。因此,胞嘧啶不是甲基化核苷酸,而5-甲基胞嘧啶是甲基化核苷酸。在另一个实例中,胸腺嘧啶在其嘧啶环的位置5处含有甲基部分;然而,出于本文的目的,当胸腺嘧啶存在于DNA中时,不认为胸腺嘧啶是甲基化核苷酸,因为胸腺嘧啶是DNA的典型核苷酸碱基。

[0113] 如本文所用,“甲基化核酸分子”是指含有一个或多个甲基化核苷酸的核酸分子。

[0114] 如本文所用,核酸分子的“甲基化状态”、“甲基化谱”和“甲基化状况”是指在核酸分子中存在或不存在一个或多个甲基化核苷酸碱基。例如,含有甲基化胞嘧啶的核酸分子被认为是甲基化的(例如,核酸分子的甲基化状态是甲基化的)。不含任何甲基化核苷酸的核酸分子被认为是未甲基化的。

[0115] 特定核酸序列(例如,如本文所述的基因标志物或DNA区域)的甲基化状态可以指示序列中每个碱基的甲基化状态,或者可以指示序列内碱基子集(例如,一个或多个胞嘧啶)的甲基化状态,或者可以指示关于序列内区域甲基化密度的信息,提供或不提供序列内发生甲基化的位置的精确信息。

[0116] 核酸分子中核苷酸基因座的甲基化状态是指核酸分子中特定基因座处甲基化核苷酸的存在或不存在。例如,当核酸分子中第7个核苷酸处存在的核苷酸为5-甲基胞嘧啶时,核酸分子中第7个核苷酸处的胞嘧啶的甲基化状态是甲基化的。类似地,当核酸分子中第7个核苷酸处存在的核苷酸为胞嘧啶(而不是5-甲基胞嘧啶)时,核酸分子中第7个核苷酸处的胞嘧啶的甲基化状态是未甲基化的。

[0117] 甲基化状况可以任选地由“甲基化值”(例如,代表甲基化频率、分数、比率、百分比等)表示或指示。例如,可以通过量化在用甲基化依赖性限制酶进行限制消化后完整核酸的量,或通过比较亚硫酸氢盐反应后的扩增概况,或通过比较经过亚硫酸氢盐处理和未处理的核酸的序列来产生甲基化值。因此,值,例如甲基化值,代表甲基化状况并且因此可以用作跨基因座的多个拷贝的甲基化状况的定量指标。这在期望将样品中序列的甲基化状态与阈值或参考值进行比较时特别有用。

[0118] 在一些实施方案中,样品是粪便样品、组织样品、血液样品(例如,血浆样品、全血样品、血清样品)或尿液样品。在一些实施方案中,样品包含血液、血清、血浆、胃分泌物、胰液、脑脊髓液(CSF)样品、胃肠活检样品和/或从粪便中回收的细胞。在一些实施方案中,受试者是人类。样品可以包括来自淋巴腺、乳房、肝脏、胆管、胰腺、胃、结肠、直肠、食道、小肠、阑尾、十二指肠、息肉、胆囊、肛门和/或腹膜的细胞、分泌物或组织。在一些实施方案中,样品包含细胞液、腹水、尿液、粪便、胃分泌物、胰液、在内窥镜检查期间获得的流体、血液。

[0119] 如本文所用,“甲基化频率”或“甲基化百分比(%)”是指相对于分子或基因座未甲基化的示例数,分子或基因座被甲基化的示例数。

[0120] 因此,甲基化状态描述了核酸(例如,基因组序列)的甲基化状态。另外,甲基化状态是指在特定基因组基因座处的核酸区段与甲基化相关的特征。此类特征包括但不限于,该DNA序列中的任何胞嘧啶(C)残基是否被甲基化、甲基化C残基的位置、甲基化C在核酸任何特定区域中的频率或百分比,以及由于例如等位基因来源的差异而导致的等位基因甲基化差异。术语“甲基化状态”、“甲基化谱”和“甲基化状况”还指生物样品中核酸的任何特定区域整个中甲基化C或未甲基化C的相对浓度、绝对浓度或模式。例如,如果核酸序列内的胞

嘧啶(C)残基被甲基化,则它可以称为“高甲基化”或“甲基化增加”,而如果DNA序列内的胞嘧啶(C)残基未被甲基化,则它可以称为“低甲基化”或“甲基化减少”。同样,如果与另一核酸序列(例如,来自不同区域或来自不同个体等)相比,核酸序列内的胞嘧啶(C)残基被甲基化,那么认为该序列与其它核酸序列相比是高甲基化的或甲基化增加。可替代地,如果与另一核酸序列(例如,来自不同区域或来自不同个体等)相比,DNA序列内的胞嘧啶(C)残基未被甲基化,那么认为该序列与其它核酸序列相比是低甲基化的或甲基化降低。此外,如本文所用的术语“甲基化模式”是指核酸区域上甲基化和未甲基化核苷酸的集合位点。当整个区域中甲基化和未甲基化核苷酸的数量相同或相似但甲基化和未甲基化核苷酸的位置不同时,两个核酸可能具有相同或相似的甲基化频率或甲基化百分比但具有不同的甲基化模式。当序列的甲基化程度(例如一个的甲基化相对于另一个增加或减少)、频率或模式有差异时,称其为“差异甲基化”或具有“甲基化差异”或具有“不同甲基化状态”。术语“差异甲基化”是指与癌症阴性样品中的核酸甲基化水平或模式相比,癌症阳性样品中的核酸甲基化水平或模式的差异。它还可能指手术后癌症复发的患者与未复发的患者之间水平或模式的差异。差异甲基化和DNA甲基化的特定水平或模式是预后和预测性生物标志物,例如,一旦确定了正确的截止或预测特征。

[0121] 甲基化状态频率可用于描述个体群体或来自单个个体的样品。例如,具有50%的甲基化状态频率的核苷酸基因座在50%的情况下是甲基化的并且在50%的情况下是未甲基化的。例如,此类频率可用于描述个体群体或核酸集合中核苷酸基因座或核酸区域被甲基化的程度。因此,当第一群体或核酸分子库中的甲基化不同于第二群体或核酸分子库中的甲基化时,第一群体或库的甲基化状态频率将不同于第二群体或库的甲基化状态频率。此类频率也可以用于例如描述单个个体中核苷酸基因座或核酸区域被甲基化的程度。例如,此类频率可用于描述来自组织样品的一组细胞在核苷酸基因座或核酸区域甲基化或非甲基化的程度。

[0122] 如本文所用,“核苷酸基因座”是指核苷酸在核酸分子中的位置。甲基化核苷酸的核苷酸基因座是指甲基化核苷酸在核酸分子中的位置。

[0123] 典型地,人类DNA的甲基化发生在包括相邻鸟嘌呤和胞嘧啶的二核苷酸序列上,其中胞嘧啶位于鸟嘌呤的5' (也称为CpG二核苷酸序列)。在人类基因组中CpG二核苷酸内的大多数胞嘧啶是甲基化的,但是一些在称为CpG岛的特定的富含CpG二核苷酸的基因组区域中保持未甲基化(参见,例如,Antequera等人(1990) *Cell* 62:503-514)。

[0124] 如本文所用,“CpG岛”是指基因组DNA的富含G:C的区域,其含有相对于总基因组DNA数量增加的CpG二核苷酸。CpG岛的长度可以为至少100、200个或更多个碱基对,其中该区域的G:C含量为至少50%并且观察到的CpG频率与期望频率的比率为0.6;在一些情况下,CpG岛的长度可以为至少500个碱基对,其中该区域的G:C含量为至少55%)并且观察到的CpG频率与期望频率的比率为0.65。可以根据Gardiner-Garden等人(1987) *J. Mol. Biol.* 196:261-281提供的方法计算观察到的CpG频率与期望频率的比率。例如,观察到的CpG频率与期望频率的比率可以根据公式 $R = (A \times B) / (C \times D)$ 计算,其中R是观察到的CpG频率与期望频率的比率,A是分析序列中CpG二核苷酸的数量,B是分析序列中的核苷酸总数,C是分析序列中的C核苷酸总数,且D是分析序列中的G核苷酸总数。典型地在CpG岛中,例如在启动子区域确定甲基化状态。尽管如此,也应理解人类基因组中的其它序列易于发

生DNA甲基化,诸如CpA和CpT(参见Ramsahoye(2000) Proc.Natl.Acad.Sci.USA 97:5237-5242;Salmon和Kaye(1970) Biochim.Biophys.Acta.204:340-351;Grafstrom(1985) Nucleic Acids Res.13:2827-2842;Nyce(1986) Nucleic Acids Res.14:4353-4367;Woodcock(1987) Biochem.Biophys.Res.Commun.145:888-894)。

[0125] 如本文所用,“甲基化特异性试剂”是指根据核酸分子的甲基化状态修饰核酸分子的核苷酸的试剂,或甲基化特异性试剂是指可以以反映核酸分子甲基化状态的方式改变核酸分子的核苷酸序列的化合物或组合物或其它试剂。用此类试剂处理核酸分子的方法可以包括使核酸分子与试剂接触,如果需要,结合附加步骤,以实现核苷酸序列的所需改变。可以以将未甲基化的核苷酸(例如,每个未甲基化的胞嘧啶)修饰成不同的核苷酸的方式应用此类方法。例如,在一些实施方案中,这种试剂可以使未甲基化的胞嘧啶核苷酸脱氨基,产生脱氧尿嘧啶残基。此类试剂的实例包括但不限于甲基化敏感性限制酶、甲基化依赖性限制酶和亚硫酸氢盐试剂。

[0126] 甲基化特异性试剂对核酸核苷酸序列的改变也可以产生其中每个甲基化核苷酸被修饰成不同的核苷酸的核酸分子。

[0127] 如本文所用,术语“用化学选择性基团修饰的UDP葡萄糖”是指已经官能化,特别是在6-羟基位置,具有能够经由点击化学与亲和标签反应的官能团的尿苷二磷酸葡萄糖分子。

[0128] 术语“氧化的5-甲基胞嘧啶”是指已经在5-位被氧化的氧化5-甲基胞嘧啶残基。氧化5-甲基胞嘧啶残基因此包括5-羟甲基胞嘧啶、5-甲酰基胞嘧啶和5-羧甲基胞嘧啶。根据本发明的一个实施方案,与有机硼烷进行反应的氧化5-甲基胞嘧啶残基是5-甲酰基胞嘧啶和5-羧甲基胞嘧啶。

[0129] 术语“甲基化测定”是指用于确定核酸序列内的一个或多个CpG二核苷酸序列的甲基化状态的任何测定。

[0130] 术语“MS AP-PCR”(甲基化敏感性任意引物聚合酶链式反应)是指本领域公认的技术,该技术允许使用富含CG的引物对基因组进行全局扫描,以关注最有可能含有CpG二核苷酸的区域,并由Gonzalzo等人(1997) Cancer Research 57:594-599进行描述。

[0131] 术语“MethyLight™”是指Eads等人(1999) Cancer Res.59:2302-2306描述的本领域公认的基于荧光的实时PCR技术。

[0132] 术语“HeavyMethyl™”是指其中覆盖介于扩增引物之间或被扩增引物覆盖的CpG位置的甲基化特异性封闭探针(本文也称为封闭剂)能够实现核酸样品的甲基化特异性选择性扩增的测定。

[0133] 术语“HeavyMethyl™MethyLight™”测定是指HeavyMethyl™MethyLight™测定,它是MethyLight™测定的变型,其中MethyLight™测定与覆盖扩增引物之间的CpG位置的甲基化特异性封闭探针组合。

[0134] 术语“Ms-SNuPE”(甲基化敏感性单核苷酸引物延伸)是指由Gonzalzo和Jones(1997) Nucleic Acids Res.25:2529-2531描述的本领域公认的测定。

[0135] 术语“MSP”(甲基化特异性PCR)是指Herman等人(1996) Proc.Natl.Acad.Sci.USA 93:9821-9826和美国专利号5,786,146描述的本领域公认的甲基化测定。

[0136] 术语“COBRA”(联合亚硫酸氢盐限制分析)是指Xiong和Laird(1997) Nucleic

Acids Res.25:2532-2534描述的本领域公认的甲基化测定。

[0137] 术语“MCA”(甲基化CpG岛扩增)是指Toyota等人(1999)Cancer Res.59:2307-12和WO 00/26401A1描述的甲基化测定。

[0138] 如本文所用,“选定核苷酸”是指核酸分子中四种典型存在的核苷酸中的一种核苷酸(DNA为C、G、T和A, RNA为C、G、U和A),并且可以包括典型存在的核苷酸的甲基化衍生物(例如,当C是选定核苷酸时,甲基化和未甲基化的C都包括在选定核苷酸的含义内),而甲基化的选定核苷酸特指甲基化的典型存在的核苷酸并且未甲基化的选定核苷酸特指未甲基化的典型存在的核苷酸。

[0139] 术语“甲基化特异性限制酶”是指根据核酸识别位点的甲基化状态选择性消化核酸的限制酶。在识别位点未甲基化或半甲基化下特异性切割的限制酶的情况下(甲基化敏感性酶),如果在一条或两条链上识别位点甲基化,那么切割不会发生(或会发生,但效率显著降低)。在只有识别位点甲基化时才特异性切割的限制酶的情况下(甲基化依赖性酶),如果识别位点未甲基化,那么切割不会发生(或会发生,但效率显著降低)。优选的是甲基化特异性限制酶,其识别序列含有CG二核苷酸(例如识别序列诸如CGCG或CCCGG)。对于一些实施方案进一步优选的是如果该二核苷酸中的胞嘧啶在碳原子C5处被甲基化则不会切割的限制酶。

[0140] 如本文所用,“不同核苷酸”是指化学上与选定核苷酸不同的核苷酸,通常使得不同核苷酸具有不同于选定核苷酸的沃森-克里克碱基配对(Watson-Crick base-pairing)特性,其中与选定核苷酸互补的典型存在的核苷酸不同于与不同核苷酸互补的典型存在的核苷酸。例如,当C是选定核苷酸时,U或T可以是不同核苷酸,此由C与G的互补性和U或T与A的互补性例示。如本文所用,与选定核苷酸互补或与不同核苷酸互补的核苷酸是指在高度严格条件下以比互补核苷酸与四种典型存在的核苷酸中的三种进行碱基配对高的亲和力与选定核苷酸或不同核苷酸进行碱基配对的核苷酸。互补性的一个实例是DNA(例如A-T和C-G)和RNA(例如A-U和C-G)中的沃森-克里克碱基配对。因此,例如,在高度严格条件下,G与C碱基配对的亲和力高于G与G、A或T碱基配对,因此,当C是选定核苷酸时,G是与选定核苷酸互补的核苷酸。

[0141] 如本文所用,给定标志物(或一起使用的一组标志物)的“灵敏度”是指报告DNA甲基化值高于区别肿瘤和非肿瘤样品的阈值的样品的百分比。在一些实施方案中,阳性定义为报告DNA甲基化值高于阈值(例如,与疾病相关的范围)的经组织学证实的瘤形成,并且假阴性定义为报告DNA甲基化值低于阈值(例如,与无疾病相关的范围)的经组织学证实的瘤形成。因此,灵敏度值反映了从已知患病样品中获得的给定标志物的DNA甲基化测量值在疾病相关测量值范围内的概率。如这里所定义,计算出的灵敏度值的临床相关性表示对当将给定标志物应用于患有该临床疾患的受试者时将检测到该疾患存在的概率的估计。

[0142] 如本文所用,给定标志物(或一起使用的一组标志物)的“特异性”是指报告DNA甲基化值低于区别肿瘤和非肿瘤样品的阈值的非肿瘤样品的百分比。在一些实施方案中,阴性定义为报告DNA甲基化值低于阈值(例如,与无疾病相关的范围)的经组织学证实的非肿瘤样品,并且假阳性定义为报告DNA甲基化值高于阈值(例如,与疾病相关的范围)的经组织学证实的非肿瘤样品。因此,特异性值反映了从已知非肿瘤样品中获得的给定标志物的DNA甲基化测量值在非疾病相关测量值范围内的概率。如这里所定义,计算出的特异性值的临

床相关性表示对当将给定标志物应用于未患临床疾患的患者时将检测到该疾患不存在的概率的估计。

[0143] 如本文所用,术语“AUC”是“曲线下面积”的缩写。其尤其是指接受者操作特征(ROC)曲线下面积。ROC曲线是针对诊断测试的不同可能切点的真阳性率与假阳性率的关系图。它显示了灵敏度与特异性之间的权衡,取决于所选的切点(灵敏度的任何增加都将伴随着特异性的降低)。ROC曲线下面积(AUC)是对诊断测试准确性的度量(面积越大越好;最佳值为1;随机测试的ROC曲线位于对角线上,面积为0.5;供参考:J.P.Egan.(1975) Signal Detection Theory and ROC Analysis,Academic Press,New York)。

[0144] 如本文所用,术语“赘生物”是指组织的任何新的异常生长。因此,赘生物可以是癌前赘生物或恶性赘生物。

[0145] 如本文所用,术语“赘生物特异性标志物”是指可用于指示赘生物的存在任何生物材料或要素。生物材料的实例包括但不限于核酸、多肽、碳水化合物、脂肪酸、细胞组分(例如,细胞膜和线粒体)和全细胞。在一些情况下,标志物是特定核酸区域(例如,基因、基因内区域、特定基因座等)。作为标志物的核酸区域可以称为例如“标志基因”、“标志区域”、“标志序列”、“标志基因座”等。

[0146] 如本文所用,术语“腺瘤”是指腺来源的良性肿瘤。尽管这些生长是良性的,但随着时间的推移,它们可能会进展为恶性。

[0147] 术语“癌前”或“肿瘤前”及其等同词是指任何正在经历恶性转化的细胞增殖性病征。

[0148] 赘生物、腺瘤、癌症等的“部位”是受试者体内赘生物、腺瘤、癌症等所在的组织、器官、细胞类型、解剖区域、身体部位等。

[0149] 如本文所用,“诊断”测试应用包括检测或鉴定受试者的疾病状态或疾患,确定受试者感染给定疾病或疾患的可能性,确定患有疾病或疾患的受试者将对疗法作出反应的可能性,确定患有疾病或疾患的受试者的预后(或其可能的进展或消退),以及确定治疗对患有疾病或疾患的受试者的影响。例如,诊断可用于检测受试者感染赘生物的存在或可能性,或此类受试者将对化合物(例如药品,例如药物)或其它治疗产生有利反应的可能性。

[0150] 术语“分离的”当如在“分离的寡核苷酸”中与核酸相关使用时是指从至少一种在天然来源中通常相关联的污染核酸中鉴定并分离的核酸序列。分离的核酸以不同于在自然界中发现的形式或设置存在。相比之下,未分离的核酸,如DNA和RNA,在它们存在于自然界中的状态下被发现。未分离的核酸的实例包括:在宿主细胞染色体上与邻近基因相邻的给定DNA序列(例如,基因);RNA序列,例如编码特定蛋白质的特定mRNA序列,在细胞中发现其呈与许多其它编码多种蛋白质的mRNA的混合物。然而,编码特定蛋白质的分离的核酸包括例如在通常表达蛋白质的细胞中的这种核酸,其中该核酸在与天然细胞不同的染色体位置中,或者另外由不同于自然界中发现的核酸侧接。分离的核酸或寡核苷酸可以单链或双链形式存在。当分离的核酸或寡核苷酸用于表达蛋白质时,该寡核苷酸将至少含有有义链或编码链(即寡核苷酸可以是单链的),但可以同时含有有义链和反义链(即,寡核苷酸可以是双链的)。分离的核酸在从其天然或典型环境分离后,可以与其它核酸或分子组合。例如,分离的核酸可以存在于其所处的宿主细胞中,例如用于异源表达。

[0151] 术语“纯化的”是指作为核酸或氨基酸序列的分子从其天然环境中移出、分离或分

离出来。因此，“分离的核酸序列”可以是纯化的核酸序列。“基本上纯化的”分子至少60%不含，优选至少75%不含，更优选至少90%不含与它们天然相关的其它组分。如本文所用，术语“纯化的”或“进行纯化”还指从样品中去除污染物。污染蛋白质的去除导致样品中感兴趣的多肽或核酸的百分比增加。在另一个实例中，重组多肽在植物、细菌、酵母或哺乳动物宿主细胞中表达，并且通过去除宿主细胞蛋白质来纯化多肽；从而增加了样品中重组多肽的百分比。

[0152] 术语“包含给定多核苷酸序列或多肽的组合物”泛指含有给定多核苷酸序列或多肽的任何组合物。该组合物可包含含有盐（例如NaCl）、去污剂（例如SDS）和其它组分（例如登哈特溶液（Denhardt's solution）、奶粉、鲑鱼精DNA等）的水溶液。

[0153] 术语“样品”以其最广泛的含义使用。在某种意义上，它可以指动物细胞或组织。在另一种意义上，它是指从任何来源获得的样本或培养物，以及生物和环境样品。生物样品可以从植物或动物（包括人类）获得，并且涵盖流体、固体、组织和气体。环境样品包括环境材料，例如表面物质、土壤、水和工业样品。这些实例不应被解释为限制适用于本发明的样品类型。

[0154] 如本文所用，在一些情况下使用的“远程样品”涉及从不是样品的细胞、组织或器官来源的部位收集的样品。

[0155] 如本文所用，术语“患者”或“受试者”是指要经受这项技术提供的各种测试的生物体。术语“受试者”包括动物，优选哺乳动物，包括人类。在一个优选的实施方案中，受试者是灵长类动物。在一个甚至更优选的实施方案中，受试者是人类。进一步关于诊断方法，优选的受试者是脊椎动物受试者。优选的脊椎动物是温血动物；优选的温血脊椎动物是哺乳动物。优选的哺乳动物最优选是人类。如本文所用，术语“受试者”包括人类和动物受试者。因此，本文提供了兽医治疗用途。因此，本发明的技术提供了对哺乳动物（例如人类）以及以下动物的诊断：那些因濒危而重要的哺乳动物（例如东北虎）；具有经济重要性的动物，例如在农场饲养供人类食用的动物；和/或对人类具有社会重要性的动物，例如作为宠物或在动物园饲养的动物。此类动物的实例包括但不限于：食肉动物，例如猫和狗；猪科动物，包括猪、肉猪和野猪；反刍动物和/或有蹄类动物，例如牛、公牛、绵羊、长颈鹿、鹿、山羊、野牛和骆驼；鳍足类动物；以及马。因此，还提供了家畜的诊断和治疗，包括但不限于家养猪、反刍动物、有蹄类动物、马（包括赛马）等。当前公开的主题还包括用于诊断受试者的肺癌的系统。例如，该系统可作为市售试剂盒提供，该试剂盒可用于在已从其收集生物样品的受试者中筛查肺癌的风险或诊断其患肺癌。根据本发明的技术提供的示例性系统包括评估本文所述的标志物的甲基化状态。

[0156] 如本文所用，术语“试剂盒”是指用于递送材料的任何递送系统。在反应测定的情况下，此类递送系统包括允许从一个位置至另一个位置存储、运输或递送反应试剂（例如适当容器中的寡核苷酸、酶等）和/或支持材料（例如，缓冲液、用于执行测定的书面说明书等）的系统。例如，试剂盒包括含有相关反应试剂和/或支持材料的一个或多个壳体（例如，盒）。如本文使用，术语“零散试剂盒（fragmented kit）”是指包括两个或更多个单独容器的递送系统，每个容器含有全部试剂盒组分的子部分。这些容器可共同或单独递送至预定接受者。例如，第一个容器可以含有用于测定的酶，而第二个容器含有寡核苷酸。术语“零散试剂盒”旨在涵盖含有受联邦食品、药品和化妆品法案第520(e)节监管的分析物特异性试剂（ASR）

的试剂盒,但不限于此。事实上,包含各自含有全部试剂盒组分的子部分的两个或更多个单独容器的任何递送系统都包括在术语“零散试剂盒”中。相比之下,“组合试剂盒”是指在单个容器(例如,在容纳每种所需组分的单个盒)中含有反应测定中的所有组分的递送系统。术语“试剂盒”包括零散试剂盒和组合试剂盒。

[0157] 如本文所用,术语“信息”是指任何事实或数据的集合。关于使用计算机系统(包括但不限于因特网)存储或处理的信息,该术语是指以任何格式(例如,模拟、数字、光学等格式)存储的任何数据。如本文所用,术语“与受试者相关的信息”是指与受试者(例如,人类、植物或动物)有关的事实或数据。术语“基因组信息”是指与基因组有关的信息,包括但不限于核酸序列、基因、甲基化百分比、等位基因频率、RNA表达水平、蛋白质表达、与基因型相关的表型等。“等位基因频率信息”是指与等位基因频率有关的事实或数据,包括但不限于等位基因同一性、等位基因的存在与受试者(例如人类受试者)的特征之间的统计相关性、个体或群体中等位基因的存在或不存在、在具有一种或多种特定特征的个体中存在等位基因的可能性百分比等。

具体实施方式

[0158] 在对各种实施方案的这个详细描述中,出于解释的目的,阐述众多特定细节以提供对所公开实施方案的彻底理解。然而,本领域技术人员将了解,这些各种实施方案可在有或无这些特定细节下实施。在其他情况下,结构和装置以框图形式显示。此外,本领域技术人员可易于了解呈现和进行方法所采用的特定顺序是说明性的,并且预期顺序可变化,并且仍然保持在本文公开的各种实施方案的精神和范围内。

[0159] 本文提供了如下技术:用于PNET筛查,并且特别地但不排他地涉及用于检测PNET和/或相关PNET类型(例如肺NET、小肠NET)的存在的方法、组合物和相关用途。如这项技术在本文中所述,使用的章节标题仅仅是出于组织的目的,而不应视为以任何方式限制主题。

[0160] 实际上,如实施例I所述,在鉴定本发明实施方案的过程中进行的实验鉴定了一组198种新型差异甲基化区域(DMR),用于区分PNET来源的DNA与非肿瘤对照DNA。从这198种新型DNA甲基化标志物中,进一步的实验鉴定出能够区别PNET与正常胰腺组织并在血液中检测PNET的标志物。

[0161] 尽管本文的公开内容涉及某些例示的实施方案,但是应当理解,这些实施方案是作为实例而不是作为限制来呈现的。

[0162] 在特定方面,本发明的技术提供了用于鉴定、确定和/或分类诸如PNET的癌症的组合物和方法。所述方法包括确定从受试者分离的生物样品(例如,粪便样品、胰腺组织样品、血浆样品)中至少一种甲基化标志物的甲基化状况,其中该标志物的甲基化状态的变化指示PNET的存在、类别或部位。特定实施方案涉及包含差异甲基化区域(DMR,例如DMR 1-198,参见表1A和表2A)的标志物,所述标志物用于诊断(例如筛查)PNET和各种NET类型(例如,肺NET、小肠NET)。

[0163] 除了分析包含本文所提供和表1A和表2A中所列出的DMR(例如DMR,例如DMR 1-198)的至少一种标志物、标志物区域或标志物的碱基的甲基化分析的实施方案以外,这项技术还提供了包含含有DMR的至少一种标志物、标志物区域或标志物的碱基的标志物组,用

于检测癌症,尤其是PNET。

[0164] 这项技术的一些实施方案是基于对包含DMR的至少一种标志物、标志物区域或标志物的碱基的CpG甲基化状况的分析。

[0165] 在一些实施方案中,本发明的技术提供了以甲基化特异性方式修饰DNA的试剂(例如甲基化敏感性限制酶、甲基化依赖性限制酶和亚硫酸氢盐试剂)的用途,所述试剂与一种或多种甲基化测定组合,用于确定包含DMR(例如DMR 1-198,参见表1A和表2A)的至少一种标志物内的CpG二核苷酸序列的甲基化状况。基因组CpG二核苷酸可以是甲基化或未甲基化的(可替代地分别称为甲基化上调和甲基化下调)。然而,本发明的方法适合于分析在远程样品(例如血液、器官流出物或粪便)的背景内的非均质性的生物样品,例如低浓度的肿瘤细胞或从中获得的生物材料。因此,当分析此类样品内CpG位置的甲基化状况时,可使用定量测定来确定特定CpG位置的甲基化水平(例如百分比、分数、比率、比例或程度)。

[0166] 根据本发明的技术,对包含DMR的标志物中CpG二核苷酸序列的甲基化状况的确定可用于诊断和表征诸如PNET的癌症。

[0167] 标志物的组合

[0168] 用于分析核酸中是否存在5-甲基胞嘧啶的常用方法基于Frommer等人描述的用于检测DNA中的5-甲基胞嘧啶的亚硫酸氢盐方法(Frommer等人(1992) Proc.Natl.Acad.Sci.USA 89:1827-31,出于所有目的,明确地通过引用整体并入本文)或其变型。映射5-甲基胞嘧啶的亚硫酸氢盐方法基于观察到胞嘧啶而不是5-甲基胞嘧啶与酸式亚硫酸根离子(也称为亚硫酸氢根)反应。反应通常根据以下步骤进行:首先,胞嘧啶与酸式亚硫酸盐反应,形成磺化胞嘧啶。接下来,磺化反应中间体的自发脱氨基产生磺化尿嘧啶。最后,磺化尿嘧啶在碱性条件下脱磺酸形成尿嘧啶。检测是可能的,因为尿嘧啶与腺嘌呤进行碱基配对(从而表现得像胸腺嘧啶),而5-甲基胞嘧啶与鸟嘌呤进行碱基配对(从而表现得像胞嘧啶)。这使得区分甲基化胞嘧啶与非甲基化胞嘧啶成为可能,通过例如亚硫酸氢盐基因组测序(Grigg G和Clark S,Bioessays(1994)16:431-36;Grigg G,DNA Seq.(1996)6:189-98),例如美国专利号5,786,146中公开的甲基化特异性PCR(MSP),或使用包括序列特异性探针裂解的测定,例如QuARTS皮瓣内切核酸酶测定(参见,例如Zou等人(2010)“Sensitive quantification of methylated markers with a novel methylation specific technology”Clin Chem 56:A199;以及美国专利号8,361,720、8,715,937、8,916,344和9,212,392)。

[0169] 一些常规技术与如下方法有关,所述方法包括将待分析的DNA封装在琼脂糖基质中,从而防止DNA的扩散和复性(亚硫酸氢盐仅与单链DNA反应),并用快速透析代替沉淀和纯化步骤(Olek A等人,(1996)“A modified and improved method for bisulfite based cytosine methylation analysis”Nucleic Acids Res.24:5064-6)。因此可以分析单个细胞的甲基化状况,说明该方法的实用性和敏感性。Rein,T.等人,(1998)Nucleic Acids Res.26:2255概述了检测5-甲基胞嘧啶的常规方法。

[0170] 亚硫酸氢盐技术通常包括在亚硫酸氢盐处理后扩增已知核酸的短的特异性片段,然后通过测序(Olek和Walter(1997)Nat.Genet.17:275-6)或引物延伸反应(Gonzalvo和Jones(1997)Nucleic Acids Res.25:2529-31;WO 95/00669;美国专利号6,251,594)测定产物以分析单个胞嘧啶位置。一些方法使用酶消化(Xiong和Laird(1997)Nucleic Acids

Res.25:2532-4)。本领域还描述了通过杂交进行检测(Olek等人,WO 99/28498)。另外,已经描述了使用亚硫酸氢盐技术检测单个基因的甲基化(Grigg和Clark(1994)Bioessays 16:431-6;Zeschnigk等人(1997)Hum Mol Genet.6:387-95;Feil等人(1994)Nucleic Acids Res.22:695;Martin等人(1995)Gene 157:261-4;WO 9746705;WO 9515373)。

[0171] 根据本技术,各种甲基化测定程序可以连同亚硫酸氢盐处理一起使用。这些测定允许确定核酸序列内一个或多个CpG二核苷酸(例如,CpG岛)的甲基化状态。除其它技术外,此类测定包括经过亚硫酸氢盐处理的核酸的测序、PCR(用于序列特异性扩增)、Southern印迹分析和使用甲基化特异性限制酶,例如甲基化敏感性酶或甲基化依赖性酶。

[0172] 例如,通过使用亚硫酸氢盐处理,基因组测序已简化用于分析甲基化模式和5-甲基胞嘧啶分布(Frommer等人(1992)Proc.Natl.Acad.Sci.USA 89:1827-1831)。此外,对从经过亚硫酸氢盐转化的DNA扩增的PCR产物的限制酶消化可用于评估甲基化状态,例如,如Sadri和Hornsby(1997)Nucl.Acids Res.24:5058-5059所述或如在称为COBRA(联合亚硫酸氢盐限制分析)的方法中所体现(Xiong和Laird(1997)Nucleic Acids Res.25:2532-2534)。

[0173] COBRA™分析是一种定量甲基化分析,可用于确定少量基因组DNA中特定基因座处的DNA甲基化水平(Xiong和Laird,Nucleic Acids Res.25:2532-2534,1997)。简言之,限制酶消化用于揭示经过亚硫酸氢钠处理的DNA的PCR产物中的甲基化依赖性序列差异。根据Frommer等人描述的程序,首先通过标准亚硫酸氢盐处理将甲基化依赖性序列差异引入基因组DNA中(Proc.Natl.Acad.Sci.USA89:1827-1831,1992)。然后使用对感兴趣的CpG岛具有特异性的引物对经过亚硫酸氢盐转化的DNA进行PCR扩增,接着进行限制性内切核酸酶消化、凝胶电泳,并使用特定的标记杂交探针进行检测。原DNA样品中的甲基化水平由消化和未消化的PCR产物的相对量在广泛的DNA甲基化水平上以线性定量方式表示。此外,这项技术可以可靠地应用于从显微解剖的石蜡包埋组织样品中获得的DNA。

[0174] 用于COBRA™分析的典型试剂(例如,可在典型的基于COBRA™的试剂盒中找到)可能包括但不限于:针对特定基因座(例如,特定基因、标志物、DMR、基因区域、标志物区域、经过亚硫酸氢盐处理的DNA序列、CpG岛等)的PCR引物;限制酶和适当的缓冲液;基因杂交寡核苷酸;对照杂交寡核苷酸;寡核苷酸探针的激酶标记试剂盒;和标记的核苷酸。另外,亚硫酸氢盐转化试剂可包括:DNA变性缓冲液;磺化缓冲液;DNA回收试剂或试剂盒(例如沉淀、超滤、亲和柱);脱磺酸缓冲液;和DNA回收组分。

[0175] 诸如“MethyLight™”(一种基于荧光的实时PCR技术)(Eads等人,Cancer Res.59:2302-2306,1999)、Ms-SNuPE™(甲基化敏感性单核苷酸引物延伸)反应(Gonzalvo和Jones,Nucleic Acids Res.25:2529-2531,1997)、甲基化特异性PCR(“MSP”;Herman等人,Proc.Natl.Acad.Sci.USA 93:9821-9826,1996;美国专利号5,786,146)和甲基化CpG岛扩增(“MCA”;Toyota等人,Cancer Res.59:2307-12,1999)等测定单独使用或与这些方法中的一种或多种组合使用。

[0176] “HeavyMethyl™”测定技术是一种基于经过亚硫酸氢盐处理的DNA的甲基化特异性扩增来评估甲基化差异的定量方法。覆盖介于扩增引物之间或被扩增引物覆盖的CpG位置的甲基化特异性封闭探针(“封闭剂”)能够实现核酸样品的甲基化特异性选择性扩增。

[0177] 术语“HeavyMethyl™MethyLight™”测定是指HeavyMethyl™MethyLight™测定,它

是MethyLight™测定的变型,其中MethyLight™测定与覆盖扩增引物之间的CpG位置的甲基化特异性封闭探针组合。HeavyMethyl™测定也可与甲基化特异性扩增引物组合使用。

[0178] 用于HeavyMethyl™分析的典型试剂(例如,可在典型的基于MethyLight™的试剂盒中找到)可以包括但不限于:针对特定基因座(例如,特定基因、标志物、基因区域、标志物区域、经过亚硫酸氢盐处理的DNA序列、CpG岛或经过亚硫酸氢盐处理的DNA序列或CpG岛等)的PCR引物;封闭寡核苷酸;优化的PCR缓冲液和脱氧核苷酸;以及Taq聚合酶。

[0179] MSP(甲基化特异性PCR)允许评估CpG岛内几乎任何一组CpG位点的甲基化状况,不依赖于甲基化敏感性限制酶的使用(Herman等人,Proc.Natl.Acad.Sci.USA 93:9821-9826,1996;美国专利号5,786,146)。简单地说,亚硫酸氢钠对DNA进行修饰,将未甲基化而非甲基化的胞嘧啶转化为尿嘧啶,随后使用相对于未甲基化DNA,对甲基化DNA具有特异性的引物扩增产物。MSP只需要少量的DNA,对给定CpG岛基因座的0.1%甲基化等位基因灵敏,并且可以对从石蜡包埋样品中提取的DNA进行。用于MSP分析的典型试剂(例如,可在典型的基于MSP的试剂盒中找到)可以包括但不限于:针对特定基因座(例如,特定基因、标志物、基因区域、标志物区域、经过亚硫酸氢盐处理的DNA序列、CpG岛等)的甲基化或未甲基化PCR引物;优化的PCR缓冲液和脱氧核苷酸;以及特定探针。

[0180] MethyLight™测定是一种高通量定量甲基化测定,该测定利用基于荧光的实时PCR(例如,TaqMan®),在PCR步骤后无需进一步操纵(Eads等人,Cancer Res.59:2302-2306,1999)。简言之,MethyLight™过程以基因组DNA的混合样品开始,该样品在亚硫酸氢钠反应中根据标准程序转化为具有甲基化依赖性序列差异的混合库(该亚硫酸氢盐过程将未甲基化的胞嘧啶残基转化为尿嘧啶)。然后在“偏向”反应中进行基于荧光的PCR,例如,使用与已知CpG二核苷酸重叠的PCR引物。序列区分发生在扩增过程的水平和荧光检测过程的水平。

[0181] MethyLight™测定用作核酸(例如基因组DNA样品)中甲基化模式的定量测试,其中序列区分发生在探针杂交水平。在定量型式中,PCR反应在存在与特定推定甲基化位点重叠的荧光探针的情况下提供甲基化特异性扩增。关于输入DNA的量的无偏对照由其中引物和探针均不与任何CpG二核苷酸重叠的反应提供。可替代地,基因组甲基化的定性测试是通过用不覆盖已知甲基化位点的对照寡核苷酸(例如,基于荧光型式的HeavyMethyl™和MSP技术)或用覆盖潜在甲基化位点的寡核苷酸探测有偏PCR库来实现的。

[0182] MethyLight™方法与任何合适的探针(例如“TaqMan®”探针、Lightcycler®探针等)一起使用。例如,在一些应用中,将双链基因组DNA用亚硫酸氢钠处理并使用TaqMan®探针,例如用MSP引物和/或HeavyMethyl封闭寡核苷酸和TaqMan®探针进行两组PCR反应中的一组。TaqMan®探针采用荧光“报告”和“淬灭”分子进行双重标记并且设计为对相对高GC含量的区域具有特异性,使其在PCR循环中在比正向或反向引物高约10°C的温度下解链。这允许TaqMan®探针在PCR退火/延伸步骤期间保持完全杂交。由于Taq聚合酶在PCR期间酶促合成一条新链,它最终会达到退火的TaqMan®探针。Taq聚合酶5'至3'内切核酸酶活性然后将通过消化TaqMan®探针而取代它,以释放荧光报告分子,以使用实时荧光检测系统定量检测其现在未淬灭的信号。

[0183] 用于MethyLight™分析的典型试剂(例如,可在典型的基于MethyLight™的试剂盒

中找到)可以包括但不限于:针对特定基因座(例如,特定基因、标志物、基因区域、标志物区域、经过亚硫酸氢盐处理的DNA序列、CpG岛等)的PCR引物; **TaqMan®**或**Lightcycler®**探针;优化的PCR缓冲液和脱氧核苷酸;以及Taq聚合酶。

[0184] QMTM(定量甲基化)测定是针对基因组DNA样品中的甲基化模式的可替代定量测试,其中序列区分发生在探针杂交水平。在这种定量型式中,PCR反应在存在与特定推定甲基化位点重叠的荧光探针的情况下提供无偏扩增。关于输入DNA的量的无偏对照由其中引物和探针均不与任何CpG二核苷酸重叠的反应提供。可替代地,基因组甲基化的定性测试是通过用不覆盖已知甲基化位点的对照寡核苷酸(基于荧光型式的HeavyMethylTM和MSP技术)或用覆盖潜在甲基化位点的寡核苷酸探测有偏PCR库来实现的。

[0185] QMTM方法可以在扩增过程中与任何合适的探针,例如“**TaqMan®**”探针、**Lightcycler®**探针一起使用。例如,将双链基因组DNA用亚硫酸氢钠处理并经受无偏引物和**TaqMan®**探针。**TaqMan®**探针采用荧光“报告”和“淬灭”分子进行双重标记并且设计为对相对高GC含量的区域具有特异性,使其在PCR循环中在比正向或反向引物高约10°C的温度下解链。这允许**TaqMan®**探针在PCR退火/延伸步骤期间保持完全杂交。由于Taq聚合酶在PCR期间酶促合成一条新链,它最终会达到退火的**TaqMan®**探针。Taq聚合酶5'至3'内切核酸酶活性然后将通过消化**TaqMan®**探针而取代它,以释放荧光报告分子,以使用实时荧光检测系统定量检测其现在未淬灭的信号。用于QMTM分析的典型试剂(例如,可在典型的基于QMTM的试剂盒中找到)可以包括但不限于:针对特定基因座(例如,特定基因、标志物、基因区域、标志物区域、经过亚硫酸氢盐处理的DNA序列、CpG岛等)的PCR引物; **TaqMan®**或**Lightcycler®**探针;优化的PCR缓冲液和脱氧核苷酸;以及Taq聚合酶。

[0186] Ms-SNuPETM技术是一种基于DNA的亚硫酸氢盐处理,接着进行单核苷酸引物延伸来评估特定CpG位点处的甲基化差异的定量方法(Gonzalzo和Jones, *Nucleic Acids Res.* 25: 2529-2531, 1997)。简言之,使基因组DNA与亚硫酸氢钠反应以将未甲基化的胞嘧啶转化为尿嘧啶,而保持5-甲基胞嘧啶不变。然后使用对经过亚硫酸氢盐转化的DNA具有特异性的PCR引物对所需靶标序列进行扩增,并且分离所得产物并用作感兴趣的CpG位点处的甲基化分析的模板。可以分析少量DNA(例如,显微解剖的病理切片),并且避免利用限制酶来确定CpG位点处的甲基化状况。

[0187] 用于Ms-SNuPETM分析的典型试剂(例如,可在典型的基于Ms-SNuPETM的试剂盒中找到)可以包括但不限于:针对特定基因座(例如,特定基因、标志物、基因区域、标志物区域、经过亚硫酸氢盐处理的DNA序列、CpG岛等)的PCR引物;优化的PCR缓冲液和脱氧核苷酸;凝胶提取试剂盒;阳性对照引物;针对特定基因座的Ms-SNuPETM引物;反应缓冲液(用于Ms-SNuPETM反应);以及标记的核苷酸。另外,亚硫酸氢盐转化试剂可包括:DNA变性缓冲液;磺化缓冲液;DNA回收试剂或试剂盒(例如沉淀、超滤、亲和柱);脱磺酸缓冲液;和DNA回收组分。

[0188] 简化代表性亚硫酸氢盐测序(RRBS)以亚硫酸氢盐处理核酸开始,将所有未甲基化的胞嘧啶转化为尿嘧啶,然后是限制酶消化(例如,通过识别包括CG序列的位点的酶,诸如MspI)并在与接头配体偶联后对片段进行完整测序。限制酶的选择富集片段的CpG密集区域,减少分析期间可映射到多个基因位置的冗余序列的数量。因此,RRBS通过选择用于测序

的限制性片段的子集(例如,通过使用制备凝胶电泳的大小选择)来降低核酸样品的复杂性。与全基因组亚硫酸氢盐测序不同,限制酶消化所产生的每个片段都含有至少一个CpG二核苷酸的DNA甲基化信息。因此,RRBS富集样品的启动子、CpG岛和其它基因组特征,这些区域中的限制酶切割位点频率很高,从而提供了一种评估一个或多个基因组基因座的甲基化状态的测定。

[0189] RRBS的典型方案包括以下步骤:用限制酶(诸如MspI)消化核酸样品、填充突出端和A加尾、连接接头、亚硫酸氢盐转化和PCR。参见,例如,等人(2005)“Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution”*Nat Methods* 7:133-6;Meissner等人(2005)“Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis”*Nucleic Acids Res.*33:5868-77。

[0190] 在一些实施方案中,定量等位基因特异性实时靶标和信号扩增(QuARTS)测定用于评估甲基化状态。在每个QuARTS测定中依序发生三个反应,包括初级反应中的扩增(反应1)和靶探针裂解(反应2);以及次级反应中的FRET裂解和荧光信号生成(反应3)。当用特定引物扩增靶标核酸时,带有皮瓣序列的特定检测探针与扩增子松散结合。靶结合位点处特定侵入性寡核苷酸的存在导致5'核酸酶,例如FEN-1内切核酸酶,通过在检测探针和皮瓣序列之间切割而释放皮瓣序列。皮瓣序列与相应FRET盒的非发夹部分互补。因此,皮瓣序列充当FRET盒上的侵入性寡核苷酸并使FRET盒荧光团与淬灭剂之间发生裂解,从而产生荧光信号。裂解反应可以切割每个靶标的多个探针,从而每个皮瓣释放多个荧光团,提供指数信号扩增。QuARTS可以通过使用具有不同染料的FRET盒检测单个反应孔中的多个靶标。参见,例如,Zou等人(2010)“Sensitive quantification of methylated markers with a novel methylation specific technology”*Clin Chem* 56:A199)以及美国专利号8,361,720、8,715,937、8,916,344和9,212,392,其各自通过引用并入本文用于所有目的。

[0191] 术语“亚硫酸氢盐试剂”是指包含亚硫酸氢盐(bisulfite)、焦亚硫酸盐(disulfite)、酸式亚硫酸盐(hydrogen sulfite)或其组合的试剂,如本文所公开,其用于区别甲基化与未甲基化的CpG二核苷酸序列。所述处理的方法在本领域中是已知的(例如,PCT/EP2004/011715和W02013/116375,其各自通过引用整体并入)。在一些实施方案中,亚硫酸氢盐处理在变性溶剂诸如但不限于正亚烷基二醇或二乙二醇二甲醚(DME)的存在下,或在二噁烷或二噁烷衍生物的存在下进行。在一些实施方案中,变性溶剂以介于1%与35%(v/v)之间的浓度使用。在一些实施方案中,亚硫酸氢盐反应在清除剂的存在下进行,清除剂诸如但不限于色烷衍生物,例如6-羟基-2,5,7,8,-四甲基色烷2-甲酸或三羟基苯甲酸及其衍生物,例如没食子酸(参见:PCT/EP2004/011715,其通过引用整体并入)。在某些优选的实施方案中,亚硫酸氢盐反应包括用酸式亚硫酸铵处理,例如,如W0 2013/116375中所述。

[0192] 在一些实施方案中,使用根据本发明的引物寡核苷酸的集合(例如,参见表V)和扩增酶来扩增经过处理的DNA的片段。几个DNA区段的扩增可以在同一个反应容器中同时进行。通常,使用聚合酶链式反应(PCR)进行扩增。扩增子的长度通常为100至2000个碱基对。

[0193] 在一些实施方案中,这项技术涉及评估包含来自表1A和表2A的DMR(例如,DMR编号1-198)的标志物的组合的甲基化状态。在一些实施方案中,评估超过一种标志物的甲基化状态增加用于鉴定受试者中的赘生物(例如,如本文所述的PNET)的筛查或诊断法的特异性

和/或灵敏度。

[0194] 在另一个实施方案中,本发明提供了一种用于将无细胞DNA中的氧化5-甲基胞嘧啶残基转化为二氢尿嘧啶残基的方法(参见Liu等人,2019,Nat Biotechnol.37,第424-429页;美国专利申请公开号202000370114)。该方法包括使选自5-甲酰基胞嘧啶(5fC)、5-羧甲基胞嘧啶(5caC)及其组合的氧化5mC残基与有机硼烷反应。氧化5mC残基可以是天然存在的,或者更通常地是先前氧化5mC或5hmC残基的结果,例如用TET家族的酶(例如TET1、TET2或TET3)氧化5mC或5hmC,或者例如用过钌酸钾(KRuO_4)或无机过氧化合物或组合物诸如过钨酸盐(参见例如Okamoto等人(2011)Chem. Commun.47:11231-33)和高氯酸铜(II)/2,2,6,6-四甲基哌啶-1-氧基(TEMPO)的组合(参见Matsushita等人(2017)Chem. Commun.53:5756-59)化学氧化5mC或5hmC的结果。

[0195] 有机硼烷可以描述为硼烷和选自氮杂环和叔胺的含氮化合物的复合物。氮杂环可以是单环的、双环的或多环的,但通常是单环的,呈含有氮杂原子和任选一个或多个选自N、O和S的附加杂原子的5元或6元环的形式。氮杂环可以是芳族或脂环族的。本文优选的氮杂环包括2-吡咯啉、2H-吡咯、1H-吡咯、吡啶烷、咪唑烷、2-吡啶啉、2-咪唑啉、吡啶、咪唑、1,2,4-三唑、1,2,4-三唑、哒嗪、嘧啶、吡嗪、1,2,4-三嗪和1,3,5-三嗪,它们中的任何一种可以是未取代的或者被一个或多个非氢取代基取代。典型的非氢取代基是烷基,特别是低级烷基,诸如甲基、乙基、正丙基、异丙基、正丁基、异丁基、叔丁基等。示例性化合物包括吡啶硼烷、2-甲基吡啶硼烷(2-methylpyridine borane,也称为2-picoline borane)和5-乙基-2-吡啶。

[0196] 在可以使用无毒试剂和温和的反应条件的程度上来说,有机硼烷与无细胞DNA中的氧化5mC残基的反应是有利的;不需要任何亚硫酸氢盐,也不需要任何其它可能降解DNA的试剂。此外,用有机硼烷将氧化5mC残基转化为二氢尿嘧啶可以在“一锅”或“一管”反应中进行,而无需分离任何中间体。这是相当重要的,因为转化涉及多个步骤,即(1)连接氧化5mC中的C-4和C-5的烯烃键的还原,(2)脱氨基,和(3)如果氧化5mC是5caC,则脱羧基,或者如果氧化5mC是5fC,则脱甲酰基。

[0197] 除了将无细胞DNA中的氧化5-甲基胞嘧啶残基转化为二氢尿嘧啶残基的方法外,本发明还提供了与上述方法相关的反应混合物。该反应混合物包含含有至少一种选自5caC、5fC及其组合的氧化5-甲基胞嘧啶残基的无细胞DNA样品,以及有效地使所述至少一种氧化5-甲基胞嘧啶残基还原、脱氨基以及脱羧基或脱甲酰基的有机硼烷。有机硼烷是硼烷和选自氮杂环和叔胺的含氮化合物的复合物,如以上所解释的那样。在一个优选的实施方案中,反应混合物基本上不含亚硫酸氢盐,意味着基本上不含亚硫酸氢根离子和亚硫酸氢盐。理想地,反应混合物不含亚硫酸氢盐。

[0198] 在本发明的一个相关方面,提供了一种用于将无细胞DNA中的5mC残基转化为二氢尿嘧啶残基的试剂盒,其中该试剂盒包括用于封闭5hmC残基的试剂、用于将5mC残基氧化(除羟甲基化以外)以提供氧化5mC残基的试剂,和有效地使氧化5mC残基还原、脱氨基、脱羧基或脱甲酰基的有机硼烷。该试剂盒还可以包括使用所述组分进行上述方法的说明书。

[0199] 在另一个实施方案中,提供了一种利用上述氧化反应的方法。该方法使得能够检测无细胞DNA中5-甲基胞嘧啶残基的存在和位置,并且包括以下步骤:

[0200] (a) 修饰片段化的、接头连接的无细胞DNA中的5hmC残基以在其上提供亲和标签,

其中亲和标签使得能够从无细胞DNA中去除含经修饰的5hmC的DNA;

[0201] (b) 从无细胞DNA中去除含经修饰的5hmC的DNA,留下含有未修饰的5mC残基的DNA;

[0202] (c) 氧化未修饰的5mC残基以得到含有选自5caC、5fC及其组合的氧化5mC残基的DNA;

[0203] (d) 使含有氧化5mC残基的DNA与有机硼烷接触,该有机硼烷有效地使氧化5mC残基还原、脱氨基和脱羧基或脱甲酰基,从而提供含有二氢尿嘧啶残基代替氧化5mC残基的DNA;

[0204] (e) 对含有二氢尿嘧啶残基的DNA进行扩增和测序;

[0205] (f) 从(e)中的测序结果确定5-甲基化模式。

[0206] 从来自受试者的身体样品中提取无细胞DNA,其中身体样品通常是全血、血浆或血清,最通常是血浆,但样品也可以是尿液、唾液、粘膜排泄物、痰液、粪便、或眼泪。在一些实施方案中,无细胞DNA来源于肿瘤。在其它实施方案中,无细胞DNA来自患有疾病或其它致病状况的患者。无细胞DNA可能源自或者可能不是源自肿瘤。在步骤(a)中,应注意其中要修饰5hmC残基的无细胞DNA是纯化的、片段化形式并且是接头连接的。可以使用本领域普通技术人员已知的和/或相关文献中描述的任何合适的方法进行该上下文中的DNA纯化,并且,虽然无细胞DNA本身可以高度片段化,但有时可能期望进一步片段化,如例如美国专利公开号2017/0253924中所述。无细胞DNA片段通常在约20个核苷酸至约500个核苷酸的大小范围内,更通常地在约20个核苷酸至约250个核苷酸的范围内。在步骤(a)中修饰的纯化的无细胞DNA片段已经使用常规手段(例如限制酶)进行末端修复,使得片段在每个3'末端和5'末端具有平端。在优选的方法中,如WO 2017/176630中所述,还已经使用聚合酶诸如Taq聚合酶为平端片段提供了包含单个腺嘌呤残基的3'突出端。这有利于后续连接选定的通用接头,即诸如Y-接头或发夹接头的接头,其连接到无细胞DNA片段的两端并含有至少一个分子条码。使用接头还使得能够对接头连接的DNA片段进行选择PCR富集。

[0207] 然后,在步骤(a)中,“纯化的、片段化无细胞DNA”包括接头连接的DNA片段。如步骤(a)中所指定的,用亲和标签修饰这些无细胞DNA片段中的5hmC残基,以便后续从无细胞DNA中去除含经修饰的5hmC的DNA。在一个实施方案中,亲和标签包含生物素部分,诸如生物素、脱硫生物素、氧代生物素、2-亚氨基生物素、二氨基生物素、生物素亚砷、生物胞素等。使用生物素部分作为亲和标签允许用链霉亲和素,例如链霉亲和素珠粒、磁性链霉亲和素珠粒等轻松去除。

[0208] 用生物素部分或其它亲和标签标记5hmC残基是通过将化学选择性基团共价附接到DNA片段中的5hmC残基来实现的,其中化学选择性基团能够与官能化亲和标签发生反应,从而将亲和标签连接到5hmC残基。在一个实施方案中,化学选择性基团是UDP葡萄糖-6-叠氮化物,其与炔烃官能化的生物素部分发生自发的1,3-环加成反应,如Robertson等人(2011) *Biochem. Biophys. Res. Comm.* 411 (1): 40-3, 美国专利号8,741,567和WO 2017/176630中所述。因此添加炔烃官能化的生物素部分会导致生物素部分共价附接到每个5hmC残基。

[0209] 然后可以在步骤(b)中将亲和标记的DNA片段沉降(pull down),在一个实施方案中,使用链霉亲和素珠粒、磁性链霉亲和素珠粒等形式的链霉亲和素沉降,如果需要的话,留作以后分析。去除亲和标记的片段后剩余的上清液含有具有未修饰的5mC残基而无5hmC残基的DNA。

[0210] 在步骤(c)中,使用任何合适的手段将未修饰的5mC残基氧化以提供5caC残基和/或5fC残基。选择氧化剂以氧化5mC残基(除羟甲基化以外),即提供5caC和/或5fC残基。可以使用催化活性TET家族的酶以酶促方式进行氧化。如本文所用的那些术语“TET家族酶”或“TET酶”是指如美国专利号9,115,386中定义的催化活性“TET家族蛋白”或“TET催化活性片段”,该专利的公开内容通过引用并入本文。在该上下文中,优选的TET酶是TET2;参见Ito等人(2011) *Science* 333 (6047):1300-1303。如前一部分中所述,也可以使用化学氧化剂以化学方式进行氧化。合适的氧化剂的实例包括但不限于:无机或有机过钌酸盐形式的过钌酸盐阴离子,包括金属过钌酸盐诸如过钌酸钾(KRuO_4)、四烷基过钌酸铵诸如四丙基过钌酸铵(TPAP)和四丁基过钌酸铵(TBAP)和聚合物负载的过钌酸盐(PSP);以及无机过氧化物和组合物,诸如过钨酸盐或高氯酸铜(II)/TEMPO组合。此时没有必要将含5fC的片段与含5caC的片段分开,只要在该过程的下一步中,步骤(e)将5fC残基和5caC残基都转化为二氢尿嘧啶(DHU)。

[0211] 在一些实施方案中,5-羟甲基胞嘧啶残基用 β -葡糖基转移酶(B3GT)封闭,而5-甲基胞嘧啶残基用TET酶氧化,有效地提供5-甲酰基胞嘧啶和5-羧甲基胞嘧啶的混合物。含有这两种氧化物质的混合物可以与2-甲基吡啶硼烷或另一种有机硼烷反应得到二氢尿嘧啶。在该实施方案的变型中,在步骤(b)中未去除含5hmC的片段。相反,“TET辅助的甲基吡啶硼烷测序(TAPS)”,含5mC的片段和含5hmC的片段一起被酶促氧化以提供含5fC和5caC的片段。在最初存在5mC和5hmC残基的任何地方,与2-甲基吡啶硼烷反应会产生DHU残基。“化学辅助的甲基吡啶硼烷测序(CAPS)”涉及用过钌酸钾选择性氧化含5hmC的片段,保持5mC残基不变。

[0212] 该实施方案的方法有许多优点:不需要亚硫酸氢盐,使用无毒的试剂和反应物;并且该过程在温和的条件下进行。另外,整个过程可以在单管中进行,无需分离任何中间体。

[0213] 在相关实施方案中,上述方法包括进一步的步骤:(g)鉴定在步骤(b)中从无细胞DNA中去除的含5hmC的DNA中的羟甲基化模式。这可以使用WO 2017/176630中详细描述的技术来执行。该过程可以在不去除或分离中间体的情况下以一管法进行。例如,最初,用BGT催化的尿苷二磷酸葡萄糖6-叠氮化物对无细胞DNA片段,优选接头连接的DNA片段进行官能化,然后经由化学选择性叠氮化物基团进行生物素化。该过程导致在每个5hmC位点共价衔接生物素。在下一步中,将生物素化的链和含有未修饰(天然)的5mC的链同时沉降以进行进一步处理。如本领域已知的,使用抗5mC抗体或甲基-CpG-结合结构域(MBD)蛋白将含有天然5mC的链沉降。然后,在5hmC残基被封闭的情况下,使用用于将5mC转化为5fC和/或5caC的任何合适的技术选择性地氧化未修饰的5mC残基,如本文别处所述。

[0214] 通过扩增获得的片段可以带有直接或间接可检测的标记。在一些实施方案中,标记是荧光标记、放射性核素或具有可在质谱仪中检测到的典型质量的可分离分子片段。在所述标记是质量标记的情况下,一些实施方案规定标记的扩增子具有单个正或负净电荷,从而允许在质谱仪中具有更好的可检测性。可以通过例如基质辅助激光解吸/电离质谱法(MALDI)或使用电喷雾质谱法(ESI)来进行检测和可视化。

[0215] 适于这些测定技术的分离DNA的方法是本领域已知的。具体而言,一些实施方案包括分离核酸,如美国专利申请序列号13/470,251(“Isolation of Nucleic Acids”)中所述,其通过引用整体并入本文。

[0216] 在一些实施方案中,本文所述的标志物可用于对粪便样品进行的QUARTS测定。在一些实施方案中,提供用于产生DNA样品的方法,尤其是用于产生包含小体积(例如,小于100微升、小于60微升)的高度纯化的、低丰度核酸并且基本上和/或有效地不含抑制用于测试DNA样品的测定(例如PCR、INVADER、QuARTS测定等)的物质的DNA样品的方法。此类DNA样品可用于定性检测取自患者的样品中存在的基因、基因变体(例如等位基因)或基因修饰(例如甲基化)的存在或定量测量其活性、表达或量的诊断测定。例如,一些癌症与特定突变等位基因或特定甲基化状态的存在相关,因此检测和/或定量此类突变等位基因或甲基化状态在癌症的诊断和治疗中具有预测价值。

[0217] 许多有价值的遗传标志物在样品中以极低的量存在,而且产生此类标志物的许多事件罕见。因此,即使是灵敏的检测方法,诸如PCR,也需要大量DNA来提供足够的低丰度靶标,以满足或替代测定的检测阈值。而且,即使少量的抑制物质的存在也会损害这些针对检测如此少量靶标的测定的准确性和精确度。因此,本文提供了对体积和浓度提供必要管理以产生此类DNA样品的方法。

[0218] 在一些实施方案中,样品包括粪便、组织样品(例如胰腺组织)、器官分泌物、CSF、唾液、血液或尿液。在一些实施方案中,受试者是人类。此类样品可以通过本领域已知的,例如熟练技术人员显而易见的多种方法获得。可以通过使样品经受本领域技术人员已知的各种技术,包括但不限于离心和过滤来获得无细胞或基本上无细胞的样品。尽管通常优选不使用侵入性技术来获得样品,但获得样品诸如组织匀浆、组织切片和活检样本仍可能是优选的。这项技术在用于制备样品以及提供用于测试的核酸的方法方面不受限制。例如,在一些实施方案中,使用直接基因捕获,例如,如美国专利号8,808,990和9,169,511以及WO 2012/155072中所详述的,或者通过相关方法从粪便样品或从血液或从血浆样品中分离DNA。

[0219] 标志物的分析可以与一个测试样品中的其它标志物分开或同时进行。例如,可以将几种标志物组合到一个测试中,以有效处理多个样品并潜在地提供更高的诊断和/或预后准确性。此外,本领域技术人员将认识到测试来自同一受试者的多个样品(例如,在连续时间点)的价值。连续样品的这种测试可以允许鉴定标志物甲基化状态随时间的变化。甲基化状态的变化,以及甲基化状态没有变化,可以提供有关疾病状况的有用信息,包括但不限于鉴定事件发生的大致时间、可挽救组织的存在和量、药物疗法的适当性、各种疗法的有效性以及受试者结果(包括未来事件的风险)的鉴定。

[0220] 生物标志物的分析可以以多种物理格式进行。例如,可以使用微量滴定板或自动化来促进大量测试样品的处理。可替代地,可以开发单一样品格式以促进以及时方式进行即刻治疗和诊断,例如在移动运输(ambulatory transport)或急诊室环境中。

[0221] 预期以试剂盒的形式提供该技术的实施方案。试剂盒包括本文所述的组合物、装置、设备等的实施方案,以及试剂盒的使用说明书。此类说明书描述了用于从样品制备分析物,例如用于收集样品和从样品制备核酸的适当方法。试剂盒的单个组分包装在适当的容器和包装(例如,小瓶、盒子、泡罩包装、安瓿、广口瓶、瓶子、试管等)中,并且所述组分一起包装在适当的容器(例如,一个或多个盒子)中以方便试剂盒的用户储存、运输和/或使用。应当理解,液体组分(例如,缓冲剂)可以以冻干形式提供以由用户重构。试剂盒可包括用于评估、验证和/或确保试剂盒性能的对照或参考。例如,用于测定样品中存在的核酸的量的

试剂盒可以包括对照,该对照包含用于比较的已知浓度的相同核酸或另一种核酸,并且在一些实施方案中,包含对对照核酸具有特异性的检测试剂(例如,引物)。试剂盒适合在临床环境中使用,并且在一些实施方案中,适合在用户家中使用。在一些实施方案中,试剂盒的组分提供用于从样品制备核酸溶液的系统的功能。在一些实施方案中,系统的某些组分由用户提供。

[0222] 利用例如如通过与预测特异性和灵敏度有关的统计技术鉴定的标志物的多种组合预测多种癌症。这项技术提供了鉴定一些癌症的预测组合和验证预测组合的方法。

[0223] 方法

[0224] 在这项技术的一些实施方案中,提供包括以下步骤的方法:

[0225] 1) 使从受试者获得的核酸(如例如从胰腺组织分离的基因组DNA)与区别至少一种标志物内的甲基化与非甲基化CpG二核苷酸的至少一种试剂或一系列试剂接触,所述标志物选自具有表1A叙述的注释的染色体区域,以及

[0226] 2) 检测PNET(例如以大于或等于80%的灵敏度和大于或等于80%的特异性提供)。

[0227] 在这项技术的一些实施方案中,提供包括以下步骤的方法:

[0228] 1) 使从受试者获得的核酸(如例如从胰腺组织分离的基因组DNA)与区别至少一种标志物内的甲基化与非甲基化CpG二核苷酸的至少一种试剂或一系列试剂接触,所述标志物选自具有选自由以下组成的组的注释的染色体区域:ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B,以及

[0229] 2) 检测PNET(例如以大于或等于80%的灵敏度和大于或等于80%的特异性提供)。

[0230] 在这项技术的一些实施方案中,提供包括以下步骤的方法:

[0231] 1) 使从受试者获得的核酸(如例如从血液样品(例如,血浆样品、全血样品、白细胞样品、血清样品)分离的基因组DNA)与区别至少一种标志物内的甲基化与非甲基化CpG二核苷酸的至少一种试剂或一系列试剂接触,所述标志物选自具有表2A叙述的注释的染色体区域,以及

[0232] 2) 检测PNET(例如以大于或等于80%的灵敏度和大于或等于80%的特异性提供)。

[0233] 在这项技术的一些实施方案中,提供包括以下步骤的方法:

[0234] 2) 使从受试者获得的核酸(如例如从血液样品(例如,血浆样品、全血样品、白细胞样品、血清样品)分离的基因组DNA)与区别至少一种标志物内的甲基化与非甲基化CpG二核苷酸的至少一种试剂或一系列试剂接触,所述标志物选自具有选自由以下组成的组的注释的染色体区域:ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B,以及

[0235] 2) 检测PNET(例如以大于或等于80%的灵敏度和大于或等于80%的特异性提供)。

[0236] 在这项技术的一些实施方案中,提供包括以下步骤的方法:

[0237] 3) 使从受试者获得的核酸(如例如从血液样品(例如,血浆样品、全血样品、白细胞样品、血清样品)分离的基因组DNA)与区别至少一种标志物内的甲基化与非甲基化CpG二核

苷酸的至少一种试剂或一系列试剂接触,所述标志物选自具有选自由以下组成的组的注释的染色体区域:SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBK2A,以及

[0238] 2) 检测转移性PNET (例如以大于或等于80%的灵敏度和大于或等于80%的特异性提供)。

[0239] 在这项技术的一些实施方案中,提供包括以下步骤的方法:

[0240] 4) 使从受试者获得的核酸(如例如从血液样品(例如,血浆样品、全血样品、白细胞样品、血清样品)分离的基因组DNA)与区别至少一种标志物内的甲基化与非甲基化CpG二核苷酸的至少一种试剂或一系列试剂接触,所述标志物选自具有选自由以下组成的组的注释的染色体区域:SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBK2A,以及

[0241] 2) 检测肺NET (例如以大于或等于80%的灵敏度和大于或等于80%的特异性提供)。

[0242] 在这项技术的一些实施方案中,提供包括以下步骤的方法:

[0243] 5) 使从受试者获得的核酸(如例如从血液样品(例如,血浆样品、全血样品、白细胞样品、血清样品)分离的基因组DNA)与区别至少一种标志物内的甲基化与非甲基化CpG二核苷酸的至少一种试剂或一系列试剂接触,所述标志物选自具有选自由以下组成的组的注释的染色体区域:SRRM3、HCN2、SPTBN4、TMC6_A、GP1BB_C、GP1BB_A、STX10_B、CACNA1C_A、CDHR2、PTPRN2、MAX.chr17.77788758.77788971、FBXL16_B、RTN2、HPCAL1、RASSF3、TSPO、RUNDC3A、SLC38A2、MAX.chr19.2478419.2478656、PDZD2、LOC100129726、CUX1、ANXA2、RXRA、S1PR4_A、FNBP1、FAM78A、IER2、PNMAL2和MOBK2A,以及

[0244] 2) 检测小肠NET (例如以大于或等于80%的灵敏度和大于或等于80%的特异性提供)。

[0245] 在这项技术的一些实施方案中,提供包括以下步骤的方法:

[0246] 1) 通过用以甲基化特异性方式修饰DNA的试剂(例如其中所述试剂是亚硫酸氢盐试剂、甲基化敏感性限制酶或甲基化依赖性限制酶)处理人类个体的生物样品中的基因组DNA,测量所述生物样品中一种或多种基因的甲基化水平,其中所述一种或多种基因选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBK2A、PNMAL2、S1PR4_A、LGALS3和MYO15B;

[0247]

[0248] 2) 使用一组针对所选择的一种或多种基因的引物扩增经过处理的基因组DNA;以及

[0249] 3) 通过聚合酶链式反应、核酸测序、质谱法、甲基化特异性核酸酶、基于质量的分

离和靶标捕获确定所述一种或多种基因的甲基化水平。

[0250] 在这项技术的一些实施方案中,提供包括以下步骤的方法:

[0251] 1) 测量来自样品的DNA中至少一种甲基化标志基因的量,其中所述一种或多种基因选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B;

[0252] 2) 测量DNA中的至少一种参考标志物的量;以及

[0253] 3) 将在所述DNA中测量的所述至少一种甲基化标志基因的量的值计算为在所述DNA中测量的所述参考标志基因的量的百分比,其中所述值指示在所述样品中测量的所述至少一种甲基化标志DNA的量。

[0254] 在这项技术的一些实施方案中,提供包括以下步骤的方法:

[0255] 1) 通过用亚硫酸氢盐,一种能够以甲基化特异性方式修饰DNA的试剂(例如甲基化敏感性限制酶、甲基化依赖性限制酶和亚硫酸氢盐试剂)处理人类个体的生物样品中的基因组DNA,测量所述生物样品中一种或多种基因的CpG位点的甲基化水平;

[0256] 2) 使用一组针对所选择的一种或多种基因的引物扩增经过修饰的基因组DNA;以及

[0257] 3) 通过甲基化特异性PCR、定量甲基化特异性PCR、甲基化敏感性DNA限制酶分析、定量亚硫酸氢盐焦磷酸测序或亚硫酸氢盐基因组测序PCR确定所述CpG位点的甲基化水平;

[0258] 其中所述一种或多种基因选自ANXA2、CACNA1C_A、CDHR2、FBXL16_B、GP1BB_A、GP1BB_C、HCN2、HPCAL1、LOC100129726、MAX.chr17.77788758-77788971、PDZD2、PTPRN2、RASSF3、RTN2、RUNDC3A、RXRA、SLC38A2、SPTBN4、SRRM3、STX10_B、TMC6_A、TSPO、CUX1、FAM78A、FNBP1、IER2、MOBKL2A、PNMAL2、S1PR4_A、LGALS3和MYO15B。

[0259] 优选地,此类方法的灵敏度为约70%至约100%,或约80%至约90%,或约80%至约85%。优选地,特异性为约70%至约100%,或约80%至约90%,或约80%至约85%。

[0260] 基因组DNA可以通过任何方式分离,包括使用市售试剂盒。简单地说,当感兴趣的DNA被细胞膜包裹时,生物样品必须通过酶促、化学或机械方式进行破坏和溶解。然后可以例如通过用蛋白酶K消化来清除DNA溶液中的蛋白质和其它污染物。然后从溶液中回收基因组DNA。这可以通过多种方法进行,包括盐析、有机萃取或DNA与固相支持物的结合。方法的选择将受到几种因素的影响,包括时间、费用和所需的DNA数量。包含赘生物或肿瘤发生前物质的所有临床样品类型都适用于本发明的方法,例如细胞系、组织切片、活检、石蜡包埋组织、体液、粪便、乳腺组织、子宫内膜组织、白细胞、结肠流出物、尿液、血浆、血清、全血、分离的血细胞、从血液中分离的细胞以及它们的组合。

[0261] 这项技术在用于制备样品以及提供用于测试的核酸的方法方面不受限制。例如,在一些实施方案中,使用直接基因捕获,例如,如美国专利申请序列号61/485386中所详述,或通过相关方法,从粪便样品或从血液或从血浆样品中分离DNA。

[0262] 然后用至少一种试剂或一系列试剂处理基因组DNA样品,所述试剂区别包含DMR(例如DMR 1-198,例如表1A和2A所提供)的至少一种标志物内的甲基化与未甲基化的CpG二核苷酸。

[0263] 在一些实施方案中,试剂将在5'-位置未甲基化的胞嘧啶碱基转化为尿嘧啶、胸腺嘧啶或在杂交行为方面与胞嘧啶不同的另一碱基。然而,在一些实施方案中,试剂可以是甲基化敏感性限制酶。

[0264] 在一些实施方案中,基因组DNA样品通过如下方式进行处理,使得将在5'-位置未甲基化的胞嘧啶碱基转化为尿嘧啶、胸腺嘧啶或在杂交行为方面与胞嘧啶不同的另一碱基。在一些实施方案中,这种处理用亚硫酸氢盐(酸式亚硫酸盐、焦亚硫酸盐)进行,随后进行碱性水解。

[0265] 然后分析处理过的核酸以确定靶基因序列(来自包含DMR,例如至少一种选自DMR 1-198(例如表1A和表2A中所提供)的DMR的标志物的至少一种基因、基因组序列或核苷酸)的甲基化状态。分析方法可选自本领域已知的那些,包括本文所列的那些,例如本文所述的QuARTS和MSP。

[0266] 异常甲基化,更具体地包含DMR(例如,DMR 1-66,例如表3所提供)的标志物的高甲基化与PNET相关联。

[0267] 这项技术涉及与PNET相关的任何样品的分析。例如,在一些实施方案中,样品包含从患者获得的组织和/或生物流体。在一些实施方案中,样品包含分泌物。在一些实施方案中,样品包含血液、血清、血浆、胃分泌物、胰液、胃肠活检样品、来自乳房活检的显微解剖细胞和/或从粪便中回收的细胞。在一些实施方案中,样品包含胰腺组织。在一些实施方案中,受试者是人类。样品可以包括来自子宫内膜、乳房、肝脏、胆管、胰腺、胃、结肠、直肠、食道、小肠、阑尾、十二指肠、息肉、胆囊、肛门和/或腹膜的细胞、分泌物或组织。在一些实施方案中,样品包含细胞液、腹水、尿液、粪便、胰液、在内窥镜镜检查期间获得的流体、血液、粘液或唾液。在一些实施方案中,样品是粪便样品。在一些实施方案中,样品是胰腺组织样品。

[0268] 此类样品可以通过本领域已知的,例如熟练技术人员显而易见的多种方法获得。例如,尿液和粪便样品很容易获得,而血液、腹水、血清或胰液样品可以通过使用例如针头和注射器以肠胃外方式获得。可以通过使样品经受本领域技术人员已知的各种技术,包括但不限于离心和过滤来获得无细胞或基本上无细胞的样品。尽管通常优选不使用侵入性技术来获得样品,但获得样品例如组织匀浆、组织切片和活检样本仍可能是优选的。

[0269] 在一些实施方案中,这项技术涉及一种治疗患者(例如,患有PNET、患有早期PNET或可能患上PNET的患者)的方法,该方法包括确定如本文提供的一种或多种DMR的甲基化状态并基于确定甲基化状态的结果对患者施用治疗。治疗可以是施用药物化合物、疫苗、免疫疗法、进行手术、对患者进行成像、进行另一测试。优选地,所述使用是在临床筛查方法、预后评估方法、监测疗法结果的方法、鉴定最可能对特定治疗性治疗有反应的患者的方法、对患者或受试者进行成像的方法以及药物筛查和开发的方法中。

[0270] 在这项技术的一些实施方案中,提供了一种用于诊断受试者中的PNET的方法。如本文所用,术语“诊断(diagnosing)”和“诊断(diagnosis)”是指熟练技术人员可以估计甚至确定受试者是否患有给定疾病或疾患或将来是否可能患上给定疾病或疾患的方法。熟练技术人员通常基于一种或多种诊断指标做出诊断,所述诊断指标例如生物标志物(例如,如本文公开的DMR),其甲基化状态指示疾患存在、严重性或不存在。

[0271] 连同诊断一起,临床癌症预后涉及确定癌症的侵袭性和肿瘤复发的可能性,以规划最有效的疗法。如果可以做出更准确的预后或者甚至可以评估患上癌症的潜在风险,则

可以为患者选择适当的疗法,在一些情况下,可以为患者选择不那么严重的疗法。癌症生物标志物的评估(例如,确定甲基化状态)可用于将不需要疗法或需要有限疗法的具有良好预后和/或患上癌症风险低的受试者与更可能患上癌症或遭受癌症复发的受益于更深入的治疗的受试者分开。

[0272] 因此,如本文所用,“做出诊断”或“诊断”还包括确定患上癌症的风险或确定预后,这可以基于本文公开的诊断生物标志物(例如,DMR)的测量预测临床结果(有或没有医学治疗)、选择适当的治疗(或治疗是否有效)、或监测当前治疗以及可能改变治疗。此外,在当前公开的主题的一些实施方案中,可以随时间对生物标志物进行多次确定以促进诊断和/或预后。生物标志物的时间变化可用于预测临床结果、监测PNET的进展和/或监测针对癌症的适当疗法的功效。例如,在这样的实施方案中,可能期望在有效疗法过程期间看到生物样品中的一种或多种本文公开的生物标志物(例如DMR)(以及可能的一种或多种额外生物标志物,如果监测)的甲基化状态随时间的变化。

[0273] 在一些实施方案中,本发明公开的主题还提供了一种用于确定是否在受试者中开始或继续癌症的预防或治疗的方法。在一些实施方案中,该方法包括在一段时间内提供来自受试者的一系列生物样品;分析该系列生物样品以确定每个生物样品中至少一种本文公开的生物标志物的甲基化状态;以及比较每个生物样品中一种或多种生物标志物的甲基化状态的任何可测量变化。生物标志物甲基化状态在一段时间内的任何变化都可用于预测患上癌症的风险,预测临床结果,确定是否开始或继续癌症的预防或治疗,以及当前疗法是否有效地治疗癌症。例如,第一时间点可以选择在开始治疗之前,并且第二时间点可以选择在开始治疗之后的某一时间。甲基化状态可以在取自不同时间点的每个样品中测量,并记录定性和/或定量差异。来自不同样品的生物标志物水平的甲基化状态的变化可与受试者中的PNET风险、预后、确定治疗功效和/或癌症进展相关。

[0274] 在优选的实施方案中,本发明的方法和组合物用于在早期阶段,例如在疾病症状出现之前治疗或诊断疾病。在一些实施方案中,本发明的方法和组合物用于在临床阶段治疗或诊断疾病。

[0275] 如所述,在一些实施方案中,可以对一种或多种诊断或预后生物标志物进行多次确定,并且可以使用标志物的时间变化来确定诊断或预后。例如,可以在初始时间确定诊断标志物,并在第二时间再次确定。在此类实施方案中,从初始时间到第二时间标志物的增加可以诊断癌症的特定类型或严重程度,或给定的预后。同样,从初始时间到第二时间标志物的减少可以指示癌症的特定类型或严重程度,或给定的预后。此外,一种或多种标志物的变化程度可与癌症的严重程度和未来的不良事件相关。熟练的技术人员将理解,虽然在某些实施方案中,可以在多个时间点对相同的生物标志物进行比较测量,但也可以在一个时间点测量给定的生物标志物,在第二时间点测量第二生物标志物,并且对这些标志物的比较可以提供诊断信息。

[0276] 如本文所用,短语“确定预后”是指熟练的技术人员可以预测受试者的疾患的过程或结果所采用的方法。术语“预后”并不是指能够以100%的准确度预测疾患的过程或结果,甚至不是基于生物标志物(例如DMR)的甲基化状态大概预测给定的过程或结果发生的可能性。相反,熟练的技术人员将理解,术语“预后”是指某个过程或结果发生的可能性增加;也就是说,与未表现出给定疾患的那些个体相比,过程或结果更有可能在表现出疾患的受试

者发生。例如,在未表现出疾患(例如,具有一种或多种DMR的正常甲基化状态)的个体中,给定结果(例如,患有PNET)的机会可能非常低。

[0277] 在一些实施方案中,统计分析将预后指标与不利结果的倾向相关联。例如,在一些实施方案中,与获自未患癌症的患者的正常对照样品中的甲基化状态不同的甲基化状态可以表明受试者比具有与对照样品中的甲基化状态更相似的水平受试者更有可能患癌症,如统计显著性水平所确定的。此外,甲基化状态相对于基线(例如,“正常”)水平的变化可以反映受试者的预后,并且甲基化状态的变化程度可以与不良事件的严重程度相关。统计显著性常常通过比较两个或更多个群体并确定置信区间和/或p值来确定。参见,例如,Dowdy和Wearden, *Statistics for Research*, John Wiley&Sons, New York, 1983, 通过引用整体并入本文。本发明主题的示例性置信区间为90%、95%、97.5%、98%、99%、99.5%、99.9%和99.99%,而示例性p值为0.1、0.05、0.025、0.02、0.01、0.005、0.001和0.0001。

[0278] 在其它实施方案中,可以建立本文公开的预后或诊断生物标志物(例如,DMR)的甲基化状态变化的阈值程度,并且简单地比较生物样品中生物标志物的甲基化状态变化程度与甲基化状态变化的阈值程度。本文提供的生物标志物的甲基化状态的优选阈值变化为约5%、约10%、约15%、约20%、约25%、约30%、约50%、约75%、约100%和约150%。在还有其它实施方案中,可以建立“列线图”,通过该“列线图”,将预后或诊断指标(生物标志物或生物标志物的组合)的甲基化状态与针对给定结果的相关倾向直接相关联。熟练的技术人员熟悉使用此类列线图来关联两个数值,并理解该测量中的不确定性与标志物浓度的不确定性相同,因为参考的是个体样品测量值,而不是群体平均值。

[0279] 在一些实施方案中,对照样品与生物样品同时分析,以便可以将从生物样品获得的结果与从对照样品获得的结果进行比较。另外,预期可以提供标准曲线,可以将生物样品的测定结果与该标准曲线进行比较。如果使用荧光标记,则此类标准曲线将生物标志物的甲基化状态呈现为测定单位的函数,例如荧光信号强度。使用取自多个供体的样品,可以提供正常组织中一种或多种生物标志物的对照甲基化状态,以及取自具有化生的供体或患有PNET的供体的组织中一种或多种生物标志物的“风险”水平的标准曲线。在该方法的某些实施方案中,在从受试者获得的生物样品中鉴定本文提供的一种或多种DMR的异常甲基化状态后,受试者被鉴定为具有化生。在该方法的其它实施方案中,在从受试者获得的生物样品中检测到此类生物标志物中的一种或多种的异常甲基化状态导致将受试者鉴定为患有癌症。

[0280] 标志物的分析可以与一个测试样品中的其它标志物分开或同时进行。例如,可以将几种标志物组合到一个测试中,以有效处理多个样品并潜在地提供更高的诊断和/或预后准确性。此外,本领域技术人员将认识到测试来自同一受试者的多个样品(例如,在连续时间点)的价值。连续样品的这种测试可以允许鉴定标志物甲基化状态随时间的变化。甲基化状态的变化,以及甲基化状态没有变化,可以提供有关疾病状况的有用信息,包括但不限于鉴定事件发生的大致时间、可挽救组织的存在和量、药物疗法的适当性、各种疗法的有效性以及受试者结果(包括未来事件的风险)的鉴定。

[0281] 生物标志物的分析可以以多种物理格式进行。例如,可以使用微量滴定板或自动化来促进大量测试样品的处理。可替代地,可以开发单一样品格式以促进以及时方式进行即刻治疗和诊断,例如在移动运输或急诊室环境中。

[0282] 在一些实施方案中,如果与对照甲基化状态相比,样品中至少一种生物标志物的甲基化状态存在可测量的差异,则受试者被诊断为患有PNET。相反,当在生物样品中未鉴定出甲基化状态的变化时,可以鉴定受试者为未患PNET、没有患癌症的风险或具有低的患癌症的风险。在这点上,患有癌症或其风险的受试者可以与患有低至基本上没有癌症或其风险的受试者相区分。那些有患上PNET风险的受试者可以进行更密集和/或定期的筛查计划,包括内窥镜监测。另一方面,那些具有低至基本上没有风险的受试者可以避免接受额外的PNET测试(例如,侵入性程序),直到例如未来的筛查,例如根据本发明的技术进行的筛查表明在那些受试者中出现了PNET风险的时间。

[0283] 如上所述,根据本发明的技术方法的实施方案,检测一种或多种生物标志物的甲基化状态的变化可以是定性测定,也可以是定量测定。因此,将受试者诊断为患有PNET或有患上PNET风险的步骤表明进行了某些阈值测量,例如,生物样品中一种或多种生物标志物的甲基化状态不同于预定的对照甲基化状态。在该方法的一些实施方案中,对照甲基化状态是生物标志物的任何可检测的甲基化状态。在对照样品与生物样品同时测试的方法的其它实施方案中,预定的甲基化状态是对照样品中的甲基化状态。在该方法的其它实施方案中,预定的甲基化状态基于标准曲线和/或通过标准曲线鉴定。在该方法的其它实施方案中,预定的甲基化状态是特定状态或状态范围。因此,可以在本领域技术人员显而易见的可接受限度内,部分地基于所实践的方法的实施方案和期望的特异性等来选择预定的甲基化状态。

[0284] 进一步关于诊断方法,优选的受试者是脊椎动物受试者。优选的脊椎动物是温血动物;优选的温血脊椎动物是哺乳动物。优选的哺乳动物最优选是人类。如本文所用,术语“受试者”包括人类和动物受试者。因此,本文提供了兽医治疗用途。因此,本发明的技术提供了对哺乳动物(例如人类)以及以下动物的诊断:那些因濒危而重要的哺乳动物(例如东北虎);具有经济重要性的动物,例如在农场饲养供人类食用的动物;和/或对人类具有社会重要性的动物,例如作为宠物或在动物园饲养的动物。此类动物的实例包括但不限于:食肉动物,例如猫和狗;猪科动物,包括猪、肉猪和野猪;反刍动物和/或有蹄类动物,例如牛、公牛、绵羊、长颈鹿、鹿、山羊、野牛和骆驼;以及马。因此,还提供了家畜的诊断和治疗,包括但不限于家养猪、反刍动物、有蹄类动物、马(包括赛马)等。

[0285] 当前公开的主题还包括用于诊断受试者的PNET的系统。例如,该系统可作为市售试剂盒提供,该试剂盒可用于筛查已收集生物样品的受试者的PNET的风险或诊断其患PNET癌症。根据本发明的技术提供的示例性系统包括评估如表1A和表2A中提供的DMR的甲基化状态。

[0286] 实施例

[0287] 实施例I.

[0288] 材料和方法

[0289] 组织和血液从Mayo Clinic生物样本库获得,并接受机构IRB监督。样本的选择严格遵守主题研究授权和纳入/排除标准(在研究方案中详述)。病例由28个实性和16个囊性胰腺神经内分泌肿瘤(PNET)组成。对照包括来自无癌患者的13个非肿瘤性胰腺组织和18个血沉棕黄层样品。来自有胰腺导管腺癌(PDAC)病史的患者、过去6个月内接受过化疗类药物或腹部进行过治疗性放射的患者的组织样品排除在研究之外。对组织进行宏观解剖并由专

门的病理学家审查组织学。样品经过年龄匹配,随机化并且是盲样。分别使用Qiagen QIAmp FFPE组织试剂盒和QIAamp DNA Blood Mini试剂盒(Qiagen,Valencia CA)纯化来自组织和血液样品的DNA。DNA用AMPure XP珠粒(Beckman-Coulter,Brea CA)重新纯化,并通过PicoGreen(Thermo-Fisher,Waltham MA)定量。使用qPCR评估DNA完整性。

[0290] 使用经过修改的NuGEN Ovation RRBS Methyl-Seq试剂盒(Tecan Genomics, Redwood City,CA)制备RRBS测序文库。样品以4重格式合并,并由Mayo Genomics Facility在Illumina HiSeq 4000仪器(Illumina, San Diego CA)上进行测序。读数由Illumina管道模块处理,用于图像分析和碱基判读。使用Mayo开发的生物信息学套件SAAP-RRBS进行二次分析。简单地说,使用Trim-Galore清理读数,并与使用BMAP构建的GRCh37/hg19参考基因组比对。对于覆盖率 $\geq 10X$ 并且碱基质量得分 ≥ 20 的CpG,甲基化比率通过计算 $C/(C+T)$ 或对于映射到反向链的读数,相反地,计算 $G/(G+A)$ 来确定。

[0291] 单个CpG按高甲基化比率排序,即给定基因座的甲基化胞嘧啶数量占该基因座总胞嘧啶计数的比率。对于病例,要求比率 ≥ 0.20 (20%);对于组织对照,组织对比组织分析, ≤ 0.05 (5%);组织对比血沉棕黄层, ≥ 0.20 (20%);对于血沉棕黄层对照, ≤ 0.01 (1%)。CpG高甲基化定义为在病例中至少20%甲基化,相比之下在组织对照中 $\leq 5\%$ 或;对于血沉棕黄层对照 $\leq 1\%$ 。将不符合这些标准的CpG丢弃。随后,候选CpG通过基因组定位被分入DMR(差异甲基化区域),范围在大约40-220bp,每个区域的最小截止值为5个CpG。CpG密度过高($>30\%$)的DMR被排除在外,以避免在验证阶段出现GC相关的扩增问题。进行了两项分析,比较PNET组织与正常组织对照以及PNET组织与血沉棕黄层对照。回归后,将两组DMR按p值、接受者操作特征曲线下面积(AUC)及病例与所有对照之间的分数甲基化比率排序。由于事先计划了独立验证,因此在此阶段没有对错误发现进行调整。

[0292] 对于每个候选区域,创建一个2D矩阵,对于病例和对照两者而言以样品-样品的方式比较单个CpG。然后将这些CpG矩阵与参考序列进行比较,以评估在初始过滤过程中是否丢弃了基因组连续的甲基化位点。从这个区域子集中,最终选择要求在每个样品水平上跨DMR序列的单个CpG协调和连续的高甲基化(在病例中)。相反,对照样品的甲基化程度必须至少比病例低5倍,并且CpG模式必须更加随机且不那么协调。要求至少10%的癌症样品对DMR内的每个CpG位点具有至少50%的高甲基化比率。

[0293] 在单独但互补的分析中,实验利用专有的DMR鉴定管道和回归包,根据CpG二核苷酸的平均甲基化值推导出DMR。比较PNET病例、组织对照和血沉棕黄层对照之间平均甲基化百分比的差异;使用每个映射的CpG的100个碱基对内的平铺阅读框(tiled reading frame)鉴定对照甲基化 $<5\%$ 的DMR;只有当覆盖的总深度为每个受试者平均10个读数并且亚组间的差异 >0 时才分析DMR。

[0294] 回归后,DMR按p值、接受者操作特征曲线下面积(AUC)和病例与所有对照之间的倍数变化差异排序。由于事先计划了独立验证,因此在此阶段没有对错误发现进行调整。

[0295] 选择DMR的一个子集进行进一步开发。标准主要是逻辑推导的ROC曲线下面积度量,它提供了对该区域区分潜力的性能评估。选择0.90的AUC作为截止值(对于病例与血沉棕黄层比较为0.95)。此外,计算甲基化倍数变化比(平均癌症高甲基化比率/平均对照高甲基化比率),组织与组织比较采用下限5,并且组织与血沉棕黄层比较采用下限100。要求p值小于0.01。DMR必须在平均和单个CpG选择过程中列出。使用MethPrimer(参见Li LC和

Dahiya R. *Bioinformatics* 2002年11月;18(11):1427-31) 为候选区域设计定量甲基化特异性PCR (qMSP) 引物并且QC对20ng (6250当量) 的阳性和阴性基因组甲基化对照进行检查。测试多个退火温度以获得最佳区分。在qMSP的两个阶段中进行验证。第一个阶段由重新测试已测序的DNA样品组成。这样做是为了验证DMR是真正具有区分力的,而不是过度拟合极大的下一代数据集的结果。第二个阶段利用更大的独立样品集:67个原发性PNET (50个实性、17个囊性)、25个转移性PNET、36个肺神经内分泌肿瘤和36个小肠神经内分泌肿瘤、24个正常胰腺对照组织和36个正常血沉棕黄层样品。

[0296] 组织像以前一样通过专家临床和病理审查进行鉴定。使用Qiagen QIAmp FFPE组织试剂盒进行DNA纯化。EZ-96DNA甲基化试剂盒 (Zymo Research, Irvine CA) 用于亚硫酸氢盐转化步骤。在Roche 480LightCyclers (Roche, Basel Switzerland) 上使用SYBR Green检测扩增10ng经过转化的DNA (每个标志物)。连续稀释的通用甲基化基因组DNA (Zymo Research) 用作定量标准。CpG不可知的ACTB (β -肌动蛋白) 测定用作输入参考和归一化对照。结果表示为甲基化拷贝 (特定标志物) / ACTB拷贝。

[0297] 对单个MDM (甲基化DNA标志物) 性能的结果进行逻辑分析。对于标志物的组合,使用两种技术。首先,将rPart技术应用于整个MDM集并仅限于3个MDM的组合,在此基础上计算出rPart预测的癌症概率。第二种方法使用随机森林回归 (rForest), 它生成500个单独的rPart模型,这些模型与原始数据的boot strap样品拟合 (大约2/3数据用于训练) 并用于估计整个MDM组的交叉验证误差 (1/3数据用于测试), 并重复500次,以避免低估或高估真实交叉验证度量的虚假分开。然后将结果在500次迭代中取平均值。

[0298] 结果

[0299] 实验利用样品制备、测序、分析管道和过滤器的专有方法 (在方法中有概述) 来鉴定PNET相关的差异甲基化区域 (DMR) 并使其缩减为可以在临床测试环境中查询和利用的差异甲基化区域。

[0300] 从组织对组织分析,鉴定了72个高甲基化DMR (表1A和表1B)。它们包括PNET特定区域以及针对更普遍的癌症谱的区域。组织对白细胞 (血沉棕黄层) 的分析得到126个高甲基化组织DMR,其中WBC中的噪音小于1% (表2A和表2B)。满足选择标准的区域的单个AUC范围为0.90-1.00,其中许多超过0.95。由于两种细胞类型的特定表观遗传性质和特征,pNET组织和血沉棕黄层的比较得到了甲基化信号方面最显著的差异,而比较正常胰腺和pNET组织的组织分析则不太如此,但有几个MDM在两组均表现出高区分力并且被选择用于后续验证。

[0301] 从组织和血沉棕黄层标志物组中选择33个候选标志物作为初始试点。开发了甲基化特异性PCR测定并在两轮组织样品上进行了测试;那些被测序 (冷冻) 的样品和更大的独立队列 (FFPE)。短扩增子引物 (<150bp) 被设计成靶向DMR中最具区分力的CpG,并在对照上进行测试以确保完全甲基化的片段以线性方式稳健地扩增,而未甲基化和/或未转化的片段不会扩增。表3列出了66个引物序列和退火温度。

[0302] 对第一阶段验证的结果进行逻辑分析以确定AUC和倍数变化。组织和血沉棕黄层对照的分析是分开运行的。结果在表4中突出显示。蓝色阴影指示AUC超过0.90的标志物。许多测定在PNET与血沉棕黄层样品中区分力为100%,而其它测定在PNET与对照组织分析中接近完美。

[0303] 这些结果为进行独立样品测试的高性能候选物提供了丰富的来源。在最初的33个

测定中,选择了31个。由于设想最终PNET测试是基于血液的,因此这些MDM区分正常白细胞来源的cfDNA的能力极为重要。因此,选择权衡主要偏向在比较血沉棕黄层样品时高性能的MDM。将两个标志物排除在外,因为它们的AUC小于0.90。其余31个落在0.90-1.00的AUC范围内。所有这些测定都展示了高分析性能-线性、效率、序列特异性(使用解链曲线分析评估)和强扩增。

[0304] 在第2轮验证中,与上一步一样,整个样品和标志物集合一批运行。每个标记物运行约10ng FFPE来源的样品DNA-总共310。表5列出了单个MDM的PNET对比正常组织和血沉棕黄层的结果。表5A示出了PNET组织、囊性PNET组织、实性PNET组织、转移性PNET组织对比正常胰腺组织以及PNET组织对比转移性PNET组织的AUC。表5B示出了小肠神经内分泌组织(NET)和肺NET对比PNET组织的AUC。表5C示出了转移性PNET组织、肺NET和小肠NET对比血沉棕黄层的AUC。在单个标志物候选物的接受者操作特征分析中,PNET对比对照组织比较的最佳拟合AUC范围为0.51至0.98。对于PNET对比血沉棕黄层比较,AUC范围为0.91-1.0。中位AUC分别为0.88和0.99。四个MDM(SRRM3、HCN2、SPTBN4和TMC6_A)实现了单独的交叉验证AUC ≥ 0.95 (表6)。这些MDM在转移性PNET组织和原发性肺和小肠NET中具有类似的区分力。这4个MDM中有三个完美区别PNET组织与血沉棕黄层,AUC为1并且可能非常适合进一步开发基于血液的测定。

[0305] 总之,全甲基化组测序、严格的过滤标准和生物学验证得出针对胰腺神经内分泌肿瘤的出色候选MDM。而且,这些MDM还可以将转移性PNET与正常胰腺组织区别开来。

[0306] 表1A.

DMR 编号	基因注释	染色体 编号	DMR 起止位置 (GRCh37/hg19)
1	ADCY4	14	24808742-24808909
2	ANKHD1-EIF4EBP3	5	139927684-139927755
3	ANXA2	15	60690826-60690959
4	BTBD17	17	72352858-72353448
5	C7orf23	7	86848467-86848674

[0307]

[0308]

6	CACNA1C_A	12	2692333-2692493
7	CCDC102A	16	57571027-57571105
8	CDHR2	5	175969650-175969749
9	CTBP2	10	126850434-126850576
10	DIDO1	20	61560373-61560657
11	EEF1A2	20	62119620-62119812
12	ENO3	17	4853728-4853800
13	FBXL16_A	16	744986-745216
14	FBXL16_B	16	750628-750715
15	FLJ22184	19	7934721-7934770
16	FLOT1	6	30711524-30711803
17	GP1BB_A	22	19710945-19711074
18	GP1BB_B	22	19711446-19711516
19	GP1BB_C	22	19711681-19711850
20	HCN2	19	591692-591781
21	HPCAL1	2	10444412-10444495
22	IZUMO1	19	49250409-49250467
23	JSRP1_A	19	2252384-2252398
24	JSRP1_B	19	2253163-2253319
25	KCNH6	17	61615733-61615864
26	LGALS3	14	55595729-55595844
27	LIN28A	1	26735086-26735256
28	LMO4	1	87796247-87796350
29	LOC100129726	2	43452130-43452585
30	LTBP4	19	41116418-41116462
31	MAST1	19	12978384-12978558
32	MAX.chr11.518965-519004	11	518965-519004
33	MAX.chr14.69283008-69283105	14	69283008-69283105
34	MAX.chr15.76638902-76638974	15	76638902-76638974
35	MAX.chr17.2627804-2627879	17	2627804-2627879
36	MAX.chr17.2659373-2659452	17	2659373-2659452
37	MAX.chr17.42357220-42357323	17	42357220-42357323
38	MAX.chr17.77788758-77788971	17	77788758-77788971
39	MAX.chr19.13266472-13266581	19	13266472-13266581
40	MAX.chr19.2478419-2478656	19	2478419-2478656
41	MAX.chr22.25678201-25678290	22	25678201-25678290
42	MAX.chr3.127176070-127176139	3	127176070-127176139
43	MAX.chr9.137028591-137028864	9	137028591-137028864
44	MT1G	16	56701913-56702151
45	MYO15B	17	73584898-73585121
46	NCRNA00245	10	77164134-77164642
47	OCA2	15	28339873-28340223
48	PDE2A	11	72301460-72301582
49	PDZD2	5	31855237-31855491
50	PHLDB3	19	43979342-43979600
51	PPARGC1B	5	149111737-149111884
52	PTPRN2	7	157484509-157484663
53	RASSF3	12	65004980-65005539

[0309]

54	RTN2	19	45996433-45996499
55	RUNDC3A	17	42392883-42393032
56	RXRA	9	137217099-137217558
57	SAMD11	1	861287-861369
58	SLC38A2	12	46767122-46767329
59	SLC38A3	3	50243435-50243504
60	SLC8A2	19	47933463-47933678
61	SPTBN4	19	41060185-41060270
62	SRRM3	7	75896582-75896785
63	STX10_A	19	13265725-13265813
64	STX10_B	19	13265939-13266255
65	SYT5	19	55684965-55685063
66	SYT7	11	61322823-61322904
67	TGIF2	20	35203463-35203676
68	TMC6_A	17	76123640-76123768
69	TMC6_B	17	76124136-76124298
70	TMEM145	19	42818005-42818094
71	TNKS1BP1	11	57078772-57078977
72	TSPO	22	43548112-43548218

[0310] 表1B.

[0311]

DMR 编号	基因注释	曲线下面积	倍数变 化	p 值
1	ADCY4	0.9156	6.628	0.0001779
2	ANKHD1-EIF4EBP3	0.9502	7.534	2.84E-06
3	ANXA2	0.9106	6.497	0.0002795
4	BTBD17	0.9356	13.09	1.42E-06
5	C7orf23	0.9606	44.2	0.00953
6	CACNA1C_A	0.9219	26.87	0.0006922
7	CCDC102A	0.9462	7.609	3.41E-05
8	CDHR2	0.9431	25.78	0.00822
9	CTBP2	0.9623	11.93	6.45E-06
10	DIDO1	0.9034	6.459	0.005633
11	EEF1A2	0.9811	23.6	1.08E-05
12	ENO3	0.937	12.57	5.35E-05
13	FBXL16_A	0.9577	10.56	1.54E-06
14	FBXL16_B	0.9528	36.61	1.39E-05
15	FLJ22184	0.927	10.86	3.43E-06
16	FLOT1	0.8974	19.06	0.0009995
17	GP1BB_A	0.9233	9.225	3.04E-05
18	GP1BB_B	0.8981	12.2	5.42E-06
19	GP1BB_C	0.9487	13.51	2.24E-06
20	HCN2	0.9583	48.06	4.15E-05
21	HPCAL1	0.9545	27.51	0.00094
22	IZUMO1	0.9484	13.03	9.94E-08
23	JSRP1_A	0.9122	7.722	8.39E-05

24	JSRP1_B	0.898	10.98	1.58E-05	
25	KCNH6	0.9081	20.38	1.47E-06	
26	LGALS3	0.9557	21.06	0.00266	
27	LIN28A	0.9568	19.93	7.84E-07	
28	LMO4	0.9744	35.78	9.11E-06	
29	LOC100129726	0.9423	68.4	0.007458	
30	LTBP4	0.8958	5.247	0.0002122	
31	MAST1	0.9757	15.32	2.72E-07	
32	MAX.chr11.518965-519004	0.8986	7.527	7.00E-05	
33	MAX.chr14.69283008-69283105	0.9207	8.658	0.004686	
34	MAX.chr15.76638902-76638974	0.9051	7.104	0.0001099	
35	MAX.chr17.2627804-2627879	0.9649	12.9	9.25E-08	
36	MAX.chr17.2659373-2659452	0.9475	28.74	0.000786	
37	MAX.chr17.42357220-42357323	0.9628	15.42	3.36E-07	
38	MAX.chr17.77788758-77788971	0.9286	30.46	0.001833	
39	MAX.chr19.13266472-13266581	0.9235	67.09	0.002876	
40	MAX.chr19.2478419-2478656	0.9306	16.61	5.56E-05	
41	MAX.chr22.25678201-25678290	0.9118	13.14	3.83E-05	
42	MAX.chr3.127176070-127176139	0.9183	6.631	3.78E-05	
43	MAX.chr9.137028591-137028864	0.9514	12.79	4.75E-05	
44	MT1G	0.926	8.244	7.36E-05	
45	MYO15B	0.9069	24.67	0.0006389	
46	NCRNA00245	0.9206	10.96	0.000104	
[0312]	47	OCA2	0.9137	14.03	0.0006256
	48	PDE2A	0.9558	13.12	8.20E-06
	49	PDZD2	0.9139	34.56	0.004131
	50	PHLDB3	0.9494	39.41	1.43E-06
	51	PPARGC1B	0.9028	10.56	0.0002482
	52	PTPRN2	0.9938	20.81	1.32E-06
	53	RASSF3	0.956	20.96	0.0003587
	54	RTN2	0.9919	21.95	2.10E-09
	55	RUNDC3A	0.937	33	0.0001732
	56	RXRA	0.9206	27.38	0.00304
	57	SAMD11	0.9792	13.05	0.0003039
	58	SLC38A2	0.9524	43.57	0.002501
	59	SLC38A3	0.941	15.42	6.02E-06
	60	SLC8A2	0.9156	15.67	0.0002196
	61	SPTBN4	0.9679	22.71	5.82E-07
	62	SRRM3	0.9579	47.41	5.42E-05
	63	STX10_A	0.9208	51.95	0.00188
	64	STX10_B	0.9614	39.5	0.0002594
	65	SYT5	0.9455	15.69	0.0006154
	66	SYT7	0.9011	7.743	0.0001346
	67	TGIF2	0.9425	27.97	2.25E-06
	68	TMC6_A	0.9901	27.34	1.69E-07
	69	TMC6_B	0.9557	14.43	3.50E-05
	70	TMEM145	0.9486	14.52	3.66E-05
	71	TNKS1BP1	0.9226	27.14	6.68E-05
[0313]	72	TSPO	0.9667	9.424	3.38E-07

[0314] 表2A.

[0315]

DMR 编号	基因注释	染 色 体 编 号	DMR 起止位置 (GRCh37/hg19)
73	ABHD8	19	17403232-17403460
74	ACAP1	17	7240013-7240106
75	ARHGAP30	1	161039227-161039440
76	AXIN1	16	375199-375316
77	BCL9	1	147016991-147017127
78	C1orf38	1	28195587-28195817
79	C2orf85	2	242810333-242810434
80	CACNA1C_B	12	2800272-2800464
81	CDK9	9	130545435-130545566
82	CRMP1	4	5868034-5868199
83	CSK	15	75069555-75069761
84	CTU2	16	88769741-88770116
85	CUX1	7	101500189-101500515
86	DEDD2	19	42703469-42703649
87	ELMO1	7	37392953-37393050
88	EPS15L1	19	16482561-16482712
89	FAM129C	19	17634027-17634213
90	FAM78A	9	134151363-134151474
91	FERMT3	11	63974772-63974955
92	FGF18	5	170878125-170878222
93	FHAD1	1	15672512-15672643
94	FNBP1	9	132650764-132651026
95	GMFG	19	39826176-39826273
96	GNG7	19	2561967-2562206
97	GRK6	5	176858662-176858774
98	HIVEP3	1	42204698-42204867
99	HMGA1	6	34203522-34203631
100	HMHA1_A	19	1069295-1069500
101	HMHA1_B	19	1074514-1074737
102	HPCAL1	2	10471174-10471748
103	IER2	19	13264692-13264807
104	IL17C	16	88700993-88701075
105	INPP5D	2	233925165-233925301
106	KIAA0195	17	73483829-73483938
107	KIAA0427	18	46361212-46361352
108	LMTK2	7	97831890-97832023
109	LOC100130872	4	1195670-1196131
110	LOC100507463	6	32813441-32813592
111	LOC285696	5	17130456-17130634
112	MAD1L1	7	1980049-1980132

[0316]

113	MARK2	11	63637233-63637411
114	MAX.chr1.16488894-16489075	1	16488894-16489075
115	MAX.chr1.210426160-210426264	1	210426160-210426264
116	MAX.chr1.225655507-225655620	1	225655507-225655620
117	MAX.chr11.68049738-68049894	11	68049738-68049894
118	MAX.chr12.12163358-12163631	12	12163358-12163631
119	MAX.chr14.102188698-102188818	14	102188698-102188818
120	MAX.chr14.107253099-107253355	14	107253099-107253355
121	MAX.chr15.31727007-31727144	15	31727007-31727144
122	MAX.chr15.70550976-70551130	15	70550976-70551130
123	MAX.chr16.11327016-11327312	16	11327016-11327312
124	MAX.chr16.50300428-50300651	16	50300428-50300651
125	MAX.chr16.50308404-50308570	16	50308404-50308570
126	MAX.chr17.74994454-74994572	17	74994454-74994572
127	MAX.chr17.76339840-76340086	17	76339840-76340086
128	MAX.chr2.10169502-10169736	2	10169502-10169736
129	MAX.chr2.235355101-235355212	2	235355101-235355212
130	MAX.chr20.56008091-56008227	20	56008091-56008227
131	MAX.chr3.187676577-187676668	3	187676577-187676668
132	MAX.chr4.4765181-4765330	4	4765181-4765330
133	MAX.chr5.53942200-53942315	5	53942200-53942315
134	MAX.chr6.159519777-159519949	6	159519777-159519949
135	MAX.chr6.170580966-170581132	6	170580966-170581132
136	MAX.chr6.20024141-20024570	6	20024141-20024570
137	MAX.chr6.24936094-24936246	6	24936094-24936246
138	MAX.chr7.391295-391422	7	391295-391422
139	MAX.chr8.142046288-142046398	8	142046288-142046398
140	MAX.chr8.142216497-142216631	8	142216497-142216631
141	MAX.chr8.144217550-144217700	8	144217550-144217700
142	MAX.chr8.145900710-145901246	8	145900710-145901246
143	MAX.chr8.80804237-80804308	8	80804237-80804308
144	MAX.chr9.87904996-87905372	9	87904996-87905372
145	MBP	18	74818401-74818536
146	MGAT1	5	180230498-180230723
147	MIR200C	12	7068171-7068303
148	MOBKL2A	19	2085442-2085612
149	NBEAL2	3	47029453-47029597
150	NCOR2_A	12	124941846-124941955
151	NCOR2_B	12	124950687-124950803
152	NELF	9	140356296-140356348
153	OSM_A	22	30662000-30662103
154	OSM_B	22	30662697-30662807
155	PARVG	22	44577550-44577908
156	PKN1	19	14551093-14551303
157	PNMAL2	19	46996516-46996606
158	PPP6R1	19	55765917-55766155
159	PRIC285	20	62199539-62199703
160	PRKAR1B	7	644126-644374

[0317]	161	PTK2B	8	27221308-27221453
	162	PTPRE	10	129845667-129845938
	163	RAC2 A	22	37626189-37626295
	164	RAC2 B	22	37637570-37637727
	165	RAP1GAP2	17	2699553-2699729
	166	RASSF1	3	50378492-50378750
	167	RBM38	20	55964848-55965398
	168	RHOF	12	122231058-122231184
	169	S1PR4 A	19	3178378-3178781
	170	S1PR4 B	19	3179828-3180413
	171	SDK2	17	71587461-71587557
	172	SEPTIN9 A	17	75449912-75450101
	173	SEPTIN9 B	17	75461597-75461735
	174	SH3BP2	4	2813867-2814151
	175	SHANK3	22	51110972-51111091
	176	SHISA5	3	48520645-48520772
	177	SHROOM1	5	132161293-132161522
	178	SKI	1	2232144-2232470
	179	SNX20	16	50715181-50715339
	180	STAT5A	17	40440733-40441156
	181	SUCLG2	3	67706348-67706568
	182	SUN2 A	22	39148139-39148300
	183	SUN2 B	22	39152758-39152893
	184	SUSD3	9	95821778-95821978
	185	TCF3	19	1650722-1650865
	186	TMC6 C	17	76127199-76127566
	187	TMEM132E	17	32964651-32964776
	188	TMEM163	2	135464600-135464735
	189	TNFRSF10C	8	22961173-22961268
	190	TNFRSF25	1	6526055-6526198
	191	TRABD	22	50629316-50629609
	192	TRAF3IP3	1	209943070-209943218
	193	UHRF1	19	4916882-4916984
	194	VAV1	19	6772930-6773075
	195	VILL	3	38035405-38035508
	196	ZC3H12D	6	149803439-149803687
	197	ZDHHC18	1	27160118-27160221
	198	ZFYVE28	4	2292779-2292933

[0318] 表2B.

DMR 编号	基因注释	曲线下 面积	倍数变 化	p 值	
[0319]	73	ABHD8	0.9967	257.2	1.39E-08
	74	ACAP1	1	606.1	4.15E-06
	75	ARHGAP30	1	678.3	2.03E-11
	76	AXIN1	1	1146	5.59E-07

[0320]

77	BCL9	1	506.8	7.00E-06
78	C1orf38	1	511.7	8.00E-07
79	C2orf85	1	104.7	8.46E-08
80	CACNA1C_B	1	359	3.44E-05
81	CDK9	1	6861	9.63E-05
82	CRMP1	1	573.6	7.98E-11
83	CSK	1	164.7	1.25E-09
84	CTU2	1	356.6	2.44E-18
85	CUX1	1	1511	0.0007766
86	DEDD2	1	686.5	5.75E-16
87	ELMO1	1	2199	1.61E-08
88	EPS15L1	1	100.3	3.76E-13
89	FAM129C	1	357	6.75E-10
90	FAM78A	1	2394	3.94E-09
91	FERMT3	1	6902	0.002545
92	FGF18	0.9782	103.7	0.0001005
93	FHAD1	1	460.7	7.58E-15
94	FNBP1	1	2139	2.36E-07
95	GMFG	1	445.8	2.66E-09
96	GNG7	1	562.4	1.13E-09
97	GRK6	1	1610	1.96E-10
98	HIVEP3	1	197	5.04E-14
99	HMGA1	1	418.8	0.0005092
100	HMHA1_A	1	619.7	0.0004255
101	HMHA1_B	1	212.2	3.57E-11
102	HPCAL1	1	373.2	5.45E-12
103	IER2	0.9806	508.5	0.009347
104	IL17C	1	828.8	1.33E-08
105	INPP5D	1	1836	8.39E-06
106	KIAA0195	1	260.1	2.58E-12
107	KIAA0427	0.9984	184.7	3.14E-11
108	LMTK2	1	1776	5.52E-07
109	LOC100130872	1	304.9	2.04E-14
110	LOC100507463	1	1204	4.42E-14
111	LOC285696	1	780.5	8.46E-12
112	MAD1L1	1	1757	5.89E-12
113	MARK2	1	246.7	3.90E-10
114	MAX.chr1.16488894-16489075	1	258.5	5.63E-10
115	MAX.chr1.210426160-210426264	0.9769	140.2	1.09E-05
116	MAX.chr1.225655507-225655620	1	714.1	3.91E-06
117	MAX.chr11.68049738-68049894	1	836.5	3.22E-07
118	MAX.chr12.12163358-12163631	1	427.7	1.01E-06
119	MAX.chr14.102188698-102188818	1	550	7.29E-06
120	MAX.chr14.107253099-107253355	1	183.3	1.10E-05
121	MAX.chr15.31727007-31727144	1	143.2	2.62E-13
122	MAX.chr15.70550976-70551130	1	283.2	3.68E-16
123	MAX.chr16.11327016-11327312	1	1448	8.67E-05
124	MAX.chr16.50300428-50300651	1	263.9	1.08E-07

[0321]

125	MAX.chr16.50308404-50308570	1	1026	2.51E-11
126	MAX.chr17.74994454-74994572	1	340.1	5.73E-23
127	MAX.chr17.76339840-76340086	1	242.1	7.49E-18
128	MAX.chr2.10169502-10169736	1	344.5	3.21E-11
129	MAX.chr2.235355101-235355212	1	833.6	9.00E-14
130	MAX.chr20.56008091-56008227	1	1437	8.70E-10
131	MAX.chr3.187676577-187676668	1	605.6	5.73E-13
132	MAX.chr4.4765181-4765330	0.9984	121.1	9.86E-11
133	MAX.chr5.53942200-53942315	1	281.6	1.71E-15
134	MAX.chr6.159519777-159519949	1	140.7	6.66E-11
135	MAX.chr6.170580966-170581132	1	249.1	1.20E-12
136	MAX.chr6.20024141-20024570	1	324.6	9.02E-21
137	MAX.chr6.24936094-24936246	1	450.6	6.59E-07
138	MAX.chr7.391295-391422	0.9783	286.9	9.61E-11
139	MAX.chr8.142046288-142046398	1	1012	5.66E-06
140	MAX.chr8.142216497-142216631	1	197.9	7.42E-07
141	MAX.chr8.144217550-144217700	0.9581	146.4	7.12E-17
142	MAX.chr8.145900710-145901246	1	596.4	1.23E-30
143	MAX.chr8.80804237-80804308	1	296.9	1.15E-05
144	MAX.chr9.87904996-87905372	1	356.7	5.38E-18
145	MBP	1	486.8	3.78E-15
146	MGAT1	1	1066	4.04E-10
147	MIR200C	1	769.4	5.84E-11
148	MOBKL2A	1	3343	0.0004202
149	NBEAL2	1	2617	1.81E-07
150	NCOR2_A	1	304.9	5.69E-11
151	NCOR2_B	1	457.2	4.35E-11
152	NELF	1	905.7	3.88E-05
153	OSM_A	1	1317	1.07E-10
154	OSM_B	1	897.5	9.33E-09
155	PARVG	1	412.3	4.08E-06
156	PKN1	1	354.8	4.63E-14
157	PNMAL2	0.9984	154.4	1.24E-08
158	PPP6R1	1	544.1	1.65E-06
159	PRIC285	1	398.8	3.43E-21
160	PRKAR1B	1	517.4	1.87E-13
161	PTK2B	1	547.3	1.03E-10
162	PTPRE	1	1138	1.45E-05
163	RAC2_A	1	1953	0.0001195
164	RAC2_B	1	184.3	1.45E-11
165	RAP1GAP2	1	1875	3.91E-05
166	RASSF1	1	353.5	5.28E-13
167	RBM38	1	164.3	5.90E-15
168	RHOF	1	456.5	4.76E-07
169	S1PR4_A	1	3268	5.18E-09
170	S1PR4_B	1	1640	9.87E-14
171	SDK2	1	490.5	7.27E-09
172	SEPTIN9_A	1	358.4	2.80E-10

173	SEPTIN9_B	1	197.9	9.73E-12
174	SH3BP2	1	710.9	2.18E-06
175	SHANK3	1	123.4	1.56E-18
176	SHISA5	1	523.2	2.96E-13
177	SHROOM1	1	368.2	3.10E-25
178	SKI	1	198.5	6.16E-18
179	SNX20	1	965.2	2.45E-08
180	STAT5A	1	462.7	4.41E-12
181	SUCLG2	1	2193	7.49E-19
182	SUN2_A	1	4978	8.42E-06
183	SUN2_B	1	285.7	5.54E-09
184	SUSD3	1	646.6	5.09E-21
185	TCF3	1	108.9	8.36E-09
186	TMC6_C	1	1019	3.82E-22
187	TMEM132E	1	185.3	4.91E-09
188	TMEM163	1	584.1	8.19E-06
189	TNFRSF10C	1	579.4	1.64E-08
190	TNFRSF25	1	145.2	4.19E-09
191	TRABD	1	204.3	6.93E-08
192	TRAF3IP3	1	239.7	1.32E-10
193	UHRF1	1	884.8	4.78E-14
194	VAV1	1	610.6	2.89E-10
195	VILL	1	131.6	1.28E-09
196	ZC3H12D	1	165.9	4.91E-13
197	ZDHHC18	1	1722	6.65E-13
198	ZFYVE28	1	388.5	2.58E-20

[0323] 表3.

DMR 编号	基因注释	Seq ID No.	正向引物 5'-3'序列	Seq ID No.	反向引物 5'-3'序列
3	ANXA2	1	AGGATTGTTTGTT ACGAGGTCGCGT	2	ACTATACTCCGCT TCTCTCCGCGCC
6	CACNA1 C_A	3	GGTCGCGGCGTT TGTTTAGAGGC	4	AACTCATCCTCCC TCCCGAAACGTC
8	CDHR2	5	CGAGTTTAGTGG TTTTTAGGTAACG G	6	TACGAAATCCGAA AAAAATCCGTA
14	FBXL16_ B	7	TCGAGGCGGTAG TATTAGGTTTACG	8	AATCTAAAAAACG AAAATCCCCGCT
17	GP1BB_ A	9	TTATTTTGTAGC GGGAGGCGTAGG C	10	CTAAACTATTTCTA ACCAAACCGCA
19	GP1BB_ C	11	TAGTAGAGCGGG TCGGGAGCGTAA GC	12	CGCCTACTACCCT ATCTAACCGAAAA CGAAC
20	HCN2	13	TTGGGAAGTTTG	14	AAAATCCGTA AAA

[0325]

			CGGTTTTTTCGTT		ACTATCCTAAAAC GCCC
21	HPCAL1	15	AGGATGTAGTTTA GTTCGTGGAGTT CGAG	16	GAAAAACGCCAA TTTTACGCCGTAA
29	LOC1001 29726	17	GTGTAATTTGGGT CGCGGTTTTTCGC	18	CGTACCTTTAACA CGCGCGATACGTT
38	MAX.chr 17.77788 758-7778 8971	19	TTAGGGTTCGGGA AAGGATTTTTTAT CGT	20	GAAACCGAACTC GAAATCCACGCG
40	MAX.chr 19.24784 19-24786 56	21	GATTTTGGGTTGC GGTGGTCGT	22	AAACACATATAAA AACATTTCAACGA A
49	PDZD2	23	AGTTATAGTTTCG GAGGCGCGGAGC	24	CCGAAAAACGAA AAAAACAAACGC T
52	PTPRN2	25	CGGGTTATAGTTA TAGGTTGGGGTAT TTCGG	26	ACTCTCGCCA ACTCCGCGAA
53	RASSF3	27	GCGGTTTTTTGGT ATTAGGAGTCGT	28	ATAAACGTAACCG AATTAACCCGAC
54	RTN2	29	TTTTATTGAAGTG GGTAAAATTTTCG AG	30	TCCGAATAAAAAA CTAAAAACACCGC TA
55	RUNDC3 A	31	TTAGGGGTTAGG GTAGGTCGTGCG T	32	CGCGAAAAACGA AAACTAAAAAAC GTA
56	RXRA	33	CGTTTGTTTAGGA AGGTTGGGTTTG GC	34	GCCGTCTCGAACG ACTAAAATTCGAA
58	SLC38A2	35	ACGGTTATGGAA ATTGGATTAGCGG	36	CCAAACGACCTTA AAAACGCCGAA
61	SPTBN4	37	TCGGATTGCGGG AGGTTGTC	38	CACGTCGAAATAA TACTACTCCACCT AAAAAACG
62	SRRM3	39	TCGTTTTAGCGGA ACGGCGG	40	GTACGTACATCGA ACGAACTATACGC CGAAC
64	STX10_B	41	TTATAGTATTAGG TGGAGTTGAGCG G	42	AACGATTCCTCGA AAAAAATACGAA
68	TMC6_A	43	GTTTTGTTGGGG TTTTGGGTTTCGG	44	AAACAAAAACCG ACAAAACCTCGCT
72	TSPO	45	TTAGGTGCGTTG TAGTTTAGACGG	46	AAACAAATCCCA AAAACACTCGAC
85	CUX1 (v1)	47	GGAGAGGATTTG AAGGGTTTCGT	48	CTCTAAAATCCTA CCCAACTCCGAT

[0326]	85	CUX1 (v2)	49	CGTAGGTTTAAA AGTGGTTCGCGG C	50	CCGATTCCTATTTT TATTA AACGAA
	90	FAM78A	51	GGGTGTTTCGGTA GCGGAGTATTAC GTT	52	ATAAAAACCTCCA TCGACCCCGTCC
	94	FNBP1	53	GCGTGATTGATG GGTGTATTACGT	54	ATAAACTTCCGAT CCCTACAACGAA
	103	IER2	55	GTCGAATCGTCG GTTCGAGGGC	56	TCTCCACGATTTT CGCGAACGCT
	148	MOBKL 2A	57	ATTTTGTTATTGT TTCGGGGATCGT	58	CACTCAAAACTTA TCTCTCAAACGCC
	157	PNMAL2	59	GAGGAAAGAGA AGTGGGCGTTTCG A	60	CCCAATCCTAATC CACTTAACGCGTC
	169	S1PR4_A	61	GGTTGGAAAGGG GTGGTTTATTTTCG A	62	GAAAACCCGCAA AAAACCCCGAA
	26	LGALS3	63	GCGTTACGGAATT TAACGGTGGTAG CG	64	CGACGAAAAAAAA CGCGAACACTAA AAAACG
	45	MYO15B	65	TTTCGAGGATAGT TCGCGGGTTTTTC	66	ATTATCGCTCGCG TCCTTAACCGAC

[0327] 表4.

DMR 编号	基因注释	AUC PNET 组织对 比良性 胰腺组 织(正 常)	AUC PNET 组织对 比血沉 棕黄层
3	ANXA2	0.81119	1
6	CACNA1C_A	0.86888	0.99621
8	CDHR2	0.96329	0.99621
14	FBXL16_B	0.92657	0.98106
17	GP1BB_A	0.88986	1

[0328]

[0329]

19	GP1BB_C	0.95629	1
20	HCN2	0.96154	0.97917
21	HPCAL1	0.8951	0.97159
26	LGALS3	0.80944	0.87879
29	LOC100129726	0.76049	0.91098
38	MAX.chr17.77788758-77788971	0.87325	0.9536
40	MAX.chr19.2478419-2478656	0.84091	0.99242
45	MYO15B	0.88287	0.73011
49	PDZD2	0.89336	0.94318
52	PTPRN2	0.95105	0.97727
53	RASSF3	0.87762	0.93939
54	RTN2	0.94056	1
55	RUNDC3A	0.90909	0.95265
56	RXRA	0.72552	0.94886
58	SLC38A2	0.88287	0.99811
61	SPTBN4	0.98427	1

62	SRRM3	0.95717	0.98485
64	STX10_B	0.95105	1
68	TMC6_A	0.96503	1
72	TSPO	0.87238	0.97727
85	CUX1 (具有引物 SEQ ID No: 47 和 48)	0.81643	1
85	CUX1 (具有引物 SEQ ID No: 49 和 50)	0.78671	1
90	FAM78A	0.63112	1
94	FNBP1	0.50699	1
103	IER2	0.8007	0.93371
148	MOBKL2A	0.56643	1
157	PNMAL2	0.52273	1
169	S1PR4_A	0.54545	1

[0330]

[0331] 表5A.

DM R 编 号	基因注释	胰腺神经内 分泌肿瘤 (PNET)组织 对比正常胰 腺组织	囊性 PNET 组 织对比正 常组织	实性 PNET 组织 对比 正常 组织	转移性 PNET 组 织对比正 常组织	PNET 组织对 比转移 性 PNET 组织
		AUC (95% CI)	AUC (95% CI)	AUC (95% CI)	AUC (95% CI)	AUC (95% CI)
62	SRRM3	0.96 (0.93-1)	0.96	0.97	0.94	0.49

[0332]

[0333]

			(0.88-1.03)	(0.92-1.01)	(0.85-1.02)	(0.34-0.63)
20	HCN2	0.96 (0.92-1)	0.91 (0.81-1.02)	0.97 (0.93-1.01)	0.89 (0.78-1)	0.4 (0.25-0.54)
61	SPTBN4	0.96 (0.92-1)	0.86 (0.71-1.01)	0.99 (0.97-1)	0.99 (0.97-1.01)	0.39 (0.27-0.52)
68	TMC6_A	0.95 (0.9-1)	0.94 (0.84-1.04)	0.95 (0.9-1)	0.89 (0.77-1)	0.37 (0.23-0.5)
19	GP1BB_C	0.93 (0.88-0.98)	0.93 (0.84-1.03)	0.93 (0.87-0.99)	0.76 (0.61-0.9)	0.43 (0.27-0.6)
17	GP1BB_A	0.93 (0.87-0.98)	0.96 (0.9-1.02)	0.91 (0.85-0.98)	0.78 (0.64-0.91)	0.34 (0.2-0.48)
64	STX10_B	0.92 (0.87-0.98)	0.88 (0.76-1.01)	0.93 (0.88-0.99)	0.86 (0.75-0.97)	0.36 (0.23-0.49)
6	CACNA1C_A	0.92 (0.86-0.98)	0.87 (0.74-1)	0.93 (0.88-0.99)	0.83 (0.7-0.96)	0.42 (0.28-0.56)
8	CDHR2	0.92 (0.86-0.98)	0.88 (0.74-1.02)	0.93 (0.86-0.99)	0.83 (0.7-0.96)	0.44 (0.3-0.58)
52	PTPRN2	0.91 (0.85-0.98)	0.9 (0.76-1.04)	0.92 (0.84-0.99)	0.93 (0.84-1.03)	0.53 (0.39-0.67)
38	MAX.chr17.777887 58.77788971	0.91 (0.85-0.97)	0.92 (0.83-1.01)	0.9 (0.83-0.97)	0.92 (0.85-1)	0.58 (0.44-0.72)
14	FBXL16_B	0.9 (0.84-0.97)	0.89 (0.75-1.04)	0.91 (0.83-0.98)	0.77 (0.63-0.92)	0.38 (0.25-0.51)
54	RTN2	0.9 (0.84-0.96)	0.88 (0.74-1.02)	0.91 (0.84-0.98)	0.92 (0.84-1.01)	0.56 (0.43-0.69)
21	HPCAL1	0.86 (0.79-0.94)	0.95 (0.87-1.02)	0.84 (0.74-0.93)	0.89 (0.8-0.99)	0.49 (0.36-0.63)
53	RASSF3	0.85 (0.77-0.93)	0.9 (0.77-1.03)	0.84 (0.74-0.93)	0.82 (0.68-0.96)	0.42 (0.28-0.55)
72	TSPO	0.85 (0.77-0.93)	0.85 (0.7-0.99)	0.85 (0.76-0.94)	0.84 (0.72-0.97)	0.49 (0.34-0.63)
55	RUNDC3A	0.81 (0.72-0.9)	0.9 (0.76-1.04)	0.78 (0.67-0.89)	0.72 (0.56-0.88)	0.45 (0.3-0.59)
58	SLC38A2	0.78 (0.69-0.88)	0.83 (0.67-0.99)	0.77 (0.66-0.88)	0.79 (0.64-0.93)	0.47 (0.34-0.6)
40	MAX.chr19.247841	0.77	0.75	0.77	0.68	0.43

	9.2478656	(0.67-0.86)	(0.56-0.94)	(0.67-0.88)	(0.51-0.84)	(0.3-0.56)
49	PDZD2	0.76 (0.67-0.86)	0.69 (0.49-0.89)	0.79 (0.69-0.89)	0.66 (0.5-0.81)	0.37 (0.25-0.5)
29	LOC100129726	0.75 (0.66-0.85)	0.83 (0.68-0.99)	0.72 (0.61-0.84)	0.87 (0.76-0.99)	0.58 (0.45-0.71)
85	CUX1 (具有引物 SEQ ID No: 47 和 48)	0.67 (0.56-0.77)	0.61 (0.39-0.83)	0.69 (0.56-0.81)	0.48 (0.31-0.66)	0.38 (0.25-0.51)
3	ANXA2	0.66 (0.55-0.77)	0.57 (0.37-0.77)	0.69 (0.57-0.81)	0.73 (0.57-0.88)	0.56 (0.42-0.69)
56	RXRA	0.62 (0.51-0.73)	0.57 (0.35-0.78)	0.64 (0.52-0.77)	0.57 (0.39-0.74)	0.46 (0.33-0.6)
[0334] 85	CUX1 (具有引物 SEQ ID No: 49 和 50)	0.62 (0.51-0.73)	0.56 (0.33-0.79)	0.64 (0.51-0.76)	0.44 (0.26-0.62)	0.39 (0.26-0.51)
169	S1PR4_A	0.6 (0.49-0.71)	0.73 (0.54-0.92)	0.56 (0.43-0.69)	0.43 (0.25-0.6)	0.38 (0.25-0.51)
94	FNBP1	0.51 (0.39-0.63)	0.68 (0.5-0.87)	0.45 (0.32-0.58)	0.29 (0.13-0.45)	0.36 (0.23-0.49)
90	FAM78A	0.5 (0.36-0.64)	0.64 (0.47-0.82)	0.45 (0.3-0.6)	0.4 (0.24-0.56)	0.39 (0.24-0.54)
103	IER2	0.49 (0.37-0.6)	0.42 (0.21-0.64)	0.51 (0.38-0.64)	0.57 (0.4-0.74)	0.59 (0.47-0.71)
157	PNMAL2	0.37 (0.24-0.51)	0.34 (0.17-0.51)	0.39 (0.25-0.53)	0.36 (0.21-0.52)	0.51 (0.38-0.64)
148	MOBK2A	0.33 (0.21-0.46)	0.42 (0.24-0.61)	0.3 (0.18-0.43)	0.22 (0.08-0.35)	0.39 (0.26-0.51)

[0335] 表5B.

DMR 编号	基因注释	小肠神经内分泌肿瘤(NET)组织对比 PNET 组织	肺 NET 组织对比 PNET 组织
		AUC (95% CI)	AUC (95% CI)
62	SRRM3	0.37 (0.26-0.49)	0.44 (0.32-0.57)
20	HCN2	0.35 (0.23-0.47)	0.71 (0.6-0.81)

[0336]

[0337]

61	SPTBN4	0.45 (0.33-0.56)	0.62 (0.5-0.74)
68	TMC6_A	0.27 (0.17-0.38)	0.67 (0.56-0.78)
19	GP1BB_C	0.32 (0.21-0.43)	0.78 (0.69-0.88)
17	GP1BB_A	0.38 (0.27-0.5)	0.68 (0.57-0.79)
64	STX10_B	0.4 (0.29-0.51)	0.33 (0.22-0.45)
6	CACNA1C_A	0.3 (0.2-0.4)	0.26 (0.15-0.37)
8	CDHR2	0.33 (0.23-0.44)	0.65 (0.54-0.76)
52	PTPRN2	0.38 (0.27-0.49)	0.56 (0.44-0.68)
38	MAX.chr17.77788758.77788971	0.17 (0.08-0.25)	0.49 (0.38-0.6)
14	FBXL16_B	0.61 (0.5-0.72)	0.77 (0.66-0.87)
54	RTN2	0.4 (0.29-0.51)	0.58 (0.46-0.71)
21	HPCAL1	0.14 (0.07-0.22)	0.5 (0.39-0.62)
53	RASSF3	0.25 (0.15-0.35)	0.4 (0.29-0.52)
72	TSPO	0.36 (0.25-0.47)	0.65 (0.54-0.77)
55	RUNDC3A	0.46 (0.34-0.57)	0.43 (0.31-0.55)
58	SLC38A2	0.18 (0.1-0.27)	0.41 (0.3-0.52)
40	MAX.chr19.2478419.2478656	0.2 (0.11-0.29)	0.35 (0.23-0.46)
49	PDZD2	0.17 (0.09-0.25)	0.14 (0.07-0.21)
29	LOC100129726v1	0.22 (0.12-0.31)	0.62 (0.5-0.75)
85	CUX1 (具有引物 SEQ ID No: 47 和 48)	0.39 (0.28-0.5)	0.38 (0.26-0.49)
3	ANXA2	0.31 (0.2-0.42)	0.57 (0.45-0.68)
56	RXRA	0.23 (0.13-0.32)	0.39 (0.27-0.5)
85	CUX1 (具有引物 SEQ ID No: 49 和 50)	0.42 (0.3-0.53)	0.41 (0.29-0.53)
169	S1PR4_A	0.52 (0.4-0.64)	0.61 (0.5-0.73)
94	FBNP1	0.59 (0.48-0.71)	0.49

			(0.37-0.61)
90	FAM78A	0.48 (0.36-0.6)	0.71 (0.6-0.81)
[0338] 103	IER2	0.66 (0.55-0.77)	0.49 (0.38-0.61)
157	PNMAL2	0.26 (0.16-0.36)	0.61 (0.49-0.72)
148	MOBKL2A	0.63 (0.52-0.74)	0.71 (0.6-0.82)

[0339] 表5C.

DM R 编 号	基因注释	PNET 组织对 比血沉 棕黄层 AUC (95% CI)	转移性 PNET 组织 对比血沉棕 黄层 AUC (95% CI)	肺 NET 组 织对比血沉 棕黄层 AUC (95% CI)	小肠 NET 组 织对比 血沉棕 黄层 AUC (95% CI)
62	SRRM3	1 (1-1)	1 (1-1)	0.99 (0.99-1)	1 (0.99-1)
20	HCN2	1 (1-1)	0.99 (0.97-1.01)	1 (1-1)	1 (1-1)
61	SPTBN4	0.99 (0.98-1)	0.99 (0.97-1.01)	0.99 (0.98-1)	0.99 (0.98-1)
[0340] 68	TMC6_A	1 (1-1)	0.99 (0.97-1.01)	1 (1-1)	0.98 (0.93-1.0 2)
19	GP1BB_C	1 (1-1)	0.99 (0.98-1.01)	1 (1-1)	1 (0.99-1)
17	GP1BB_A	1 (1-1)	1 (1-1)	1 (1-1)	1 (0.99-1)
64	STX10_B	1 (1-1)	1 (0.99-1.01)	0.99 (0.98-1.01)	1 (1-1)
6	CACNA1C_A	1 (1-1)	1 (0.99-1)	0.98 (0.94-1.02)	1 (1-1)
8	CDHR2	1 (1-1)	1 (1-1)	1 (1-1)	1 (1-1)
52	PTPRN2	0.93 (0.88-0. 98)	0.94 (0.87-1.01)	0.94 (0.89-1)	0.97 (0.93-1.0 1)
38	MAX.chr17.77788758.77 788971	0.96 (0.93-0. 99)	0.97 (0.93-1)	0.99 (0.97-1.01)	0.84 (0.74-0.9 4)
14	FBXL16_B	1 (1-1)	1 (0.99-1)	1 (1-1)	1 (1-1)
54	RTN2	0.97 (0.95-1)	0.99 (0.96-1.01)	0.96 (0.93-1)	1 (0.99-1)

	21	HPCAL1	0.84 (0.76-0.91)	0.84 (0.75-0.94)	0.87 (0.8-0.95)	0.56 (0.41-0.7)
	53	RASSF3	0.91 (0.85-0.97)	0.9 (0.81-0.99)	0.95 (0.89-1)	0.77 (0.66-0.89)
	72	TSPO	0.92 (0.86-0.97)	0.92 (0.85-0.99)	0.95 (0.9-1.01)	0.93 (0.87-0.98)
	55	RUNDC3A	0.99 (0.98-1)	0.99 (0.98-1)	1 (0.99-1)	1 (0.99-1)
	58	SLC38A2	1 (0.99-1)	1 (0.99-1)	1 (0.99-1)	0.98 (0.96-1.01)
	40	MAX.chr19.2478419.2478656	0.86 (0.79-0.94)	0.8 (0.66-0.94)	0.67 (0.53-0.81)	0.5 (0.36-0.65)
[0341]	49	PDZD2	0.83 (0.75-0.91)	0.75 (0.62-0.88)	0.43 (0.29-0.56)	0.51 (0.37-0.64)
	29	LOC100129726v1	0.99 (0.98-1)	1 (0.99-1)	1 (0.99-1)	1 (0.99-1)
	85	CUX1 (具有引物 SEQ ID No: 47 和 48)	1 (1-1)	1 (1-1)	1 (1-1)	1 (1-1)
	3	ANXA2	1 (1-1)	1 (0.99-1)	1 (1-1)	1 (1-1)
	56	RXRA	0.91 (0.85-0.97)	0.9 (0.82-0.97)	0.86 (0.78-0.95)	0.8 (0.68-0.92)
	85	CUX1 (具有引物 SEQ ID No: 49 和 50)	1 (1-1)	1 (1-1)	1 (1-1)	1 (1-1)
	169	S1PR4 A	1 (1-1)	1 (1-1)	1 (1-1)	1 (1-1)
	94	FNBP1	1 (1-1)	1 (1-1)	1 (1-1)	1 (1-1)
	90	FAM78A	1 (1-1)	1 (1-1)	1 (1-1)	1 (1-1)
	103	IER2	1 (1-1)	1 (1-1)	1 (0.99-1)	1 (1-1)
	157	PNMAL2	0.98 (0.96-1.01)	0.98 (0.96-1.01)	0.99 (0.97-1.01)	0.96 (0.9-1.01)
	148	MOBKL2A	1 (1-1)	1 (1-1)	1 (1-1)	1 (1-1)

[0342] 通过引用并入

[0343] 出于所有目的,将本文所提到的每篇专利文件和科学论文的全部公开内容通过引用并入。

[0344] 等效方案

[0345] 在不脱离本发明的精神或基本特征的情况下,本发明能以其它具体形式来实施。因此,前述实施方案在所有方面都应被视为说明性的,而非是对本文中描述的本发明的限制。因此,本发明的范围是由随附权利要求而非由上述描述所指示,并且属于权利要求的等效性含义和范围内的所有变化都意图包括在其中。

序列表

<110> 梅奥医学教育与研究基金会 (MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH)

<120> 检测胰腺神经内分泌肿瘤

<130> PPI22171784US

<150> US 63/019,751

<151> 2020-05-04

<160> 66

<170> PatentIn version 3.5

<210> 1

<211> 25

<212> DNA

<213> 人工序列 (Artificial Sequence)

<220>

<223> 合成的

<400> 1

aggattgttt gttacgaggt cgcgt 25

<210> 2

<211> 25

<212> DNA

<213> 人工序列 (Artificial Sequence)

<220>

<223> 合成的

<400> 2

actatactcc gcttctctcc gcgcc 25

<210> 3

<211> 23

<212> DNA

<213> 人工序列 (Artificial Sequence)

<220>

<223> 合成的

<400> 3

ggtcgcggcg tttgtttaga ggc 23

<210> 4

<211> 25

<212> DNA

<213> 人工序列 (Artificial Sequence)

<220>

<223> 合成的
<400> 4
aactcatcct ccctcccgaac acgtc 25
<210> 5
<211> 26
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 5
cgagtttagt ggttttagg taacgg 26
<210> 6
<211> 24
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 6
tacgaaatcc gaaaaaatc cgta 24
<210> 7
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 7
tcgaggcggc agtattaggt ttacg 25
<210> 8
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 8
aatctaaaaa acgaaatcc ccgct 25
<210> 9
<211> 26
<212> DNA
<213> 人工序列 (Artificial Sequence)

<220>
<223> 合成的
<400> 9
tttattttgt agcgggaggc gtaggc 26
<210> 10
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 10
ctaaactatt tctaaccaaa ccgca 25
<210> 11
<211> 26
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 11
tagtagagcg ggtcgggagc gtaagc 26
<210> 12
<211> 31
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 12
cgcctactac cctatctaac cgaaaacgaa c 31
<210> 13
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 13
ttgggaagtt tgcggttttt tcggtt 25
<210> 14
<211> 30
<212> DNA

<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 14
aaaatccgta aaaactatcc taaaacgccc 30
<210> 15
<211> 29
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 15
aggatgtagt ttagttcgtg gagttcgag 29
<210> 16
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 16
gaaaaacgcc aattttacgc cgtaa 25
<210> 17
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 17
gtgtaatttg ggtcgcggtt ttcgc 25
<210> 18
<211> 26
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 18
cgtaccttta acacgcgcga tacggt 26
<210> 19
<211> 28

<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 19
ttagggctcgg gaaaggattt tttatcgt 28
<210> 20
<211> 24
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 20
gaaaccgaac tcgaaatcca cgcg 24
<210> 21
<211> 22
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 21
gattttgggt tgcggtggtc gt 22
<210> 22
<211> 27
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 22
aaacacatat aaaaacattt caacgaa 27
<210> 23
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 23
agttatagtt tcggaggcgc ggagc 25
<210> 24

<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 24
ccgaaaaaacg aaaaaaacaac acgct 25
<210> 25
<211> 31
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 25
cgggttatag ttatagggtg ggtatttcg g 31
<210> 26
<211> 20
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 26
actctcgcca actccgcaa 20
<210> 27
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 27
gcggtttttt ggtattagga gtcgt 25
<210> 28
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 28
ataaacgtaa ccgaattaac ccgac 25

<210> 29
<211> 28
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 29
ttttattgaa gtgggtaaaa ttttcgag 28
<210> 30
<211> 28
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 30
tccgaataaa aaactaaaa caccgcta 28
<210> 31
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 31
ttaggggtta gggtaggtcg tgcgt 25
<210> 32
<211> 27
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 32
cgcgaaaaac gaaaactaaa aaacgta 27
<210> 33
<211> 27
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 33

cgtttgttta ggaaggttgg gtttggc 27
<210> 34
<211> 26
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 34
gccgtctcga acgactaaaa ttcgaa 26
<210> 35
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 35
acggttatgg aaattggatt agcgg 25
<210> 36
<211> 24
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 36
ccaaacgacc ttaaaaacgc cgaa 24
<210> 37
<211> 20
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 37
tcgattgcg ggaggttgtc 20
<210> 38
<211> 34
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的

<400> 38
cacgtcgaaa taatactact ccacctaaaa aacg 34
<210> 39
<211> 20
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 39
tcgttttagc ggaacggcgg 20
<210> 40
<211> 31
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 40
gtacgtacat cgaacgaact atacgccgaa c 31
<210> 41
<211> 26
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 41
ttatagtatt aggtggagtt gagcgg 26
<210> 42
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 42
aacgattcct cgaaaaaaat acgaa 25
<210> 43
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>

<223> 合成的
<400> 43
gttttgtttg gggttttggg ttcgg 25
<210> 44
<211> 24
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 44
aaacaaaaac cgacaaaact cgct 24
<210> 45
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 45
tttaggtgcg ttgtagttta gacgg 25
<210> 46
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 46
aaacaaatcc caaaaactac tcgac 25
<210> 47
<211> 23
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 47
ggagaggatt tgaagggttt cgt 23
<210> 48
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)

<220>
<223> 合成的
<400> 48
ctctaaaatc ctacccaact ccgat 25
<210> 49
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 49
cgtaggttta aaagtgggtc gcggc 25
<210> 50
<211> 26
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 50
ccgattccta tttctattaa aacgaa 26
<210> 51
<211> 27
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 51
gggtgttcgg tagcggagta ttacgtt 27
<210> 52
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 52
ataaaaacct ccatcgaccc cgtcc 25
<210> 53
<211> 24
<212> DNA

<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 53
gcgtgattga tgggtgtatt acgt 24
<210> 54
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 54
ataaaacttcc gatccctaca acgaa 25
<210> 55
<211> 22
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 55
gtcgaatcgt cggttcgagg gc 22
<210> 56
<211> 23
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 56
tctccacgat tttcgcgaac gct 23
<210> 57
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 57
atthttgttat tgthttcgggg atcgt 25
<210> 58
<211> 26

<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 58
cactcaaaaac ttatctctca aacgcc 26
<210> 59
<211> 24
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 59
gaggaaagag aagtgggcgt tcga 24
<210> 60
<211> 26
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 60
cccaatccta atccacttaa cgcgtc 26
<210> 61
<211> 26
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 61
ggttgaaaag ggtggttta tttcga 26
<210> 62
<211> 23
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 62
gaaaaccgc aaaaaacccc gaa 23
<210> 63

<211> 27
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 63
gcgttacgga atttaacggt gtagcg 27
<210> 64
<211> 30
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 64
cgacgaaaaa aacgcgaaca ctaaaaaacg 30
<210> 65
<211> 26
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 65
tttcgaggat agttcgcggg ttttc 26
<210> 66
<211> 25
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 合成的
<400> 66
attatcgctc gcgtccttaa ccgac 25