



(12) 发明专利申请

(10) 申请公布号 CN 116113708 A

(43) 申请公布日 2023.05.12

(21) 申请号 202180054264.3

(22) 申请日 2021.07.08

(30) 优先权数据

63/049,538 2020.07.08 US

(85) PCT国际申请进入国家阶段日

2023.03.02

(86) PCT国际申请的申请数据

PCT/US2021/040971 2021.07.08

(87) PCT国际申请的公布数据

W02022/011197 EN 2022.01.13

(71) 申请人 仁慈生物分析公司

地址 美国马萨诸塞州

(72) 发明人 J·C·塞德拉克 L·T·博尔托林

D·P·塞勒姆 E·S·温-德恩

D·古森莱特内尔 A·D·库维朗

(74) 专利代理机构 北京律盟知识产权代理有限公司

责任公司 11287

专利代理师 沈锦华

(51) Int.Cl.

C12Q 1/6804 (2006.01)

权利要求书10页 说明书146页 附图34页

(54) 发明名称

用于检测肺癌的组合物和方法

(57) 摘要

本公开一方面提供了用于检测肺癌例如,早期检测肺癌的技术。在另一方面,本文提供的技术可用于选择和/或监测和/或评估施用于确定患有或易患肺癌的受试者的治疗的功效。在一些实施方案中,本文提供的技术可用于开发伴随诊断,例如,通过结合治疗测量肿瘤负荷和肿瘤负荷的变化。

1. 一种方法,所述方法包括以下步骤:

(a) 提供或获得来自受试者的血液来源样品;

(b) 在所述血液来源样品中检测表达第一靶生物标志物特征的细胞外囊泡(“第一靶生物标志物特征表达细胞外囊泡”),所述第一靶生物标志物特征包含:

至少一个细胞外囊泡相关膜结合多肽和

选自由以下组成的组的至少一个靶生物标志物:表面蛋白生物标志物、囊泡内蛋白生物标志物和囊泡内RNA生物标志物,其中:

所述表面蛋白生物标志物选自ABCC3、ALCAM、ARSL、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274(PD-L1)、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、DMBT1、DSG2、EGFR、EPCAM、EPHX3、EVA1A、FOLR1、GJB1、GJB2、GPC4、HS6ST2、IG1FR、KDEL3、KRTCAP3、LAMB3、LFNG、LSR、MANEAL、MET、MSLN、MUC1、MUC21、PIGT、PODXL2、PRRG4、ROS1、SDC1、SERINC2、SEZ6L2、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、TNFRSF10B及其组合;

所述囊泡内蛋白生物标志物选自AOC1、C12orf45、CRABP2、CST1、ETV4、FAM83A、FOXA2、HMGB3、LGALS3BP、MIF、NAPSA、PPP1R14D、S100A14、SBK1、SCGB3A2、SFTA2、SFTPA1、SFTPA2、SFTPB、SPINK1、TGFA、ZC3H11A及其组合;

所述囊泡内RNA生物标志物选自ABCC3、AOC1、ARSL、B3GNT3、C12orf45、CDCP1、CDH1、CDH3、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CRABP2、CST1、DMBT1、DSG2、EPCAM、EPHX3、ETV4、EVA1A、FAM83A、FOLR1、FOXA2、GJB1、GJB2、GPC4、HMGB3、HS6ST2、KDEL3、KRTCAP3、LAMB3、LFNG、LGALS3BP、LSR、MANEAL、MIF、MSLN、MUC1、MUC21、NAPSA、PIGT、PODXL2、PPP1R14D、PRRG4、ROS1、S100A14、SBK1、SCGB3A2、SDC1、SERINC2、SEZ6L2、SFTA2、SFTPA1、SFTPA2、SFTPB、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、SPINK1、ST14、TGFA、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、ZC3H11A及其组合;

(c) 将指示所述血液来源样品中表达所述第一靶生物标志物特征的细胞外囊泡的水平样品信息与包括第一参考阈值水平的参考信息进行比较;

(d) 当所述血液来源样品示出相对于参考所述第一参考阈值水平的分类截止值升高的第一靶生物标志物特征表达细胞外囊泡水平时,将所述受试者分类为患有或易患肺癌。

2. 如权利要求1所述的方法,其中当所述至少一个靶生物标志物选自所述表面蛋白生物标志物中的一个或多个时,所述选定的一个或多个表面蛋白生物标志物和所述至少一个细胞外囊泡相关膜结合多肽是不同的。

3. 如权利要求1或2所述的方法,其中所述步骤(b)和(c)对至少第二靶生物标志物特征是重复的,并且其中所述分类截止值参考所述第一参考阈值水平和对应于所述至少第二靶生物标志物特征的至少第二参考阈值水平。

4. 如权利要求1至3中任一项所述的方法,其中所述细胞外囊泡相关膜结合多肽是或包含ALCAM、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274(PD-L1)、CEACAM5、CEACAM6、CLDN3、CLDN4、DSG2、EGFR、EPCAM、FOLR1、GJB1、GJB2、IG1FR、LAMB3、MET、MSLN、MUC1、PIGT、PODXL2、ROS1、SDC1、SLC34A2、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMPRSS4、TSPAN8、TNFRSF10B或其组合。

5. 如权利要求1至4中任一项所述的方法,其中所述第一和/或第二靶生物标志物特征包含至少一个细胞外囊泡相关膜结合多肽和至少两个选自由以下组成的组的靶生物标志物:表面蛋白生物标志物、囊泡内蛋白生物标志物和囊泡内RNA生物标志物。

6. 如权利要求1至5中任一项所述的方法,其中所述至少两个靶生物标志物包含以下组合之一:

- 至少两个不同的表面蛋白生物标志物;
- 至少两个不同的囊泡内蛋白生物标志物;
- 至少两个不同的囊泡内RNA生物标志物;
- 表面蛋白生物标志物和囊泡内蛋白生物标志物;
- 表面蛋白生物标志物和囊泡内RNA生物标志物;
- 囊泡内蛋白生物标志物和囊泡内RNA生物标志物。

7. 一种方法,所述方法包括以下步骤:

(a) 提供或获得来自受试者的血液来源样品;

(b) 在所述血液来源样品中检测表达第一靶生物标志物特征的细胞外囊泡(“第一靶生物标志物特征表达细胞外囊泡”),所述第一靶生物标志物特征包含:

至少一个细胞外囊泡相关膜结合多肽和

至少一个靶生物标志物,其中所述靶生物标志物是或包含表面蛋白生物标志物,其中所述靶生物标志物特征包含如表4中所列的生物标志物组合;

(c) 将指示所述血液来源样品中表达所述第一靶生物标志物特征的细胞外囊泡的水平的样品信息与包括第一参考阈值水平的参考信息进行比较;

(d) 当所述血液来源样品示出相对于参考所述第一参考阈值水平的分类截止值升高的第一靶生物标志物特征表达细胞外囊泡水平时,将所述受试者分类为患有或易患肺癌。

8. 如权利要求1至6中任一项所述的方法,其中所述细胞外囊泡相关膜结合多肽是或包含CEACAM5多肽和/或SLC34A2多肽。

9. 如权利要求1至8中任一项所述的方法,其中所述第一和/或第二靶生物标志物特征包含(i) 至少一个细胞外囊泡相关膜结合多肽生物标志物,其是或包含CEACAM5多肽和/或SLC34A2多肽;和(ii) 至少一个靶生物标志物CEACAM6,其可以是表面蛋白生物标志物或囊泡内RNA生物标志物。

10. 如权利要求1至8中任一项所述的方法,其中所述第一和/或第二靶生物标志物特征包含(i) 至少一个细胞外囊泡相关膜结合多肽生物标志物,其是或包含CEACAM5多肽和/或SLC34A2多肽;和(ii) 至少两个靶生物标志物CEACAM6和EPCAM,其各自可以是表面蛋白生物标志物或囊泡内RNA生物标志物。

11. 如权利要求1至8中任一项所述的方法,其中所述第一和/或第二靶生物标志物特征包含(i) 至少一个细胞外囊泡相关膜结合多肽生物标志物,其是或包含CEACAM5多肽;和(ii) 至少两个靶生物标志物CEACAM6和SLC34A2,其各自可以是表面蛋白生物标志物或囊泡内RNA生物标志物。

12. 如权利要求1至11中任一项所述的方法,其中所述第一和/或第二参考阈值水平通过在来自非癌症受试者群体的可比样品中观察到的靶生物标志物特征表达细胞外囊泡的水平来确定。

13. 如权利要求11所述的方法,其中所述非癌症受试者群体包括以下受试者群体中的一个或多个:健康受试者、诊断为患有良性肿瘤的受试者以及患有非肺相关疾病、病症和/或病状的受试者。

14. 如权利要求1至13中任一项所述的方法,其中所述血液来源样品已经经过尺寸排阻色谱以分离(例如,直接从所述血液来源样品)具有包括细胞外囊泡的关注的大小范围的纳米颗粒。

15. 如权利要求1至14中任一项所述的方法,其中所述检测步骤包括捕获测定。

16. 如权利要求15所述的方法,其中所述捕获测定涉及使所述血液来源样品与包含结合到所述至少一个细胞外囊泡相关膜结合多肽的靶捕获部分的捕获剂接触。

17. 如权利要求15所述的方法,其中所述捕获剂是或包含固体基底,所述固体基底包含与其缀合的靶捕获部分。

18. 如权利要求17所述的方法,其中所述固体基底包含磁珠。

19. 如权利要求16至18中任一项所述的方法,其中所述靶捕获部分是或包含抗体剂。

20. 如权利要求1至19中任一项所述的方法,其中所述检测步骤包括检测测定。

21. 如权利要求1至19中任一项所述的方法,其中所述检测步骤包括捕获测定和检测测定,所述捕获测定在所述检测测定之前进行。

22. 如权利要求20至21中任一项所述的方法,其中当所述第一和/或第二靶生物标志物特征包含至少一个囊泡内RNA生物标志物时,所述检测测定涉及逆转录qPCR。

23. 如权利要求20至22中任一项所述的方法,其中当所述第一和/或第二靶生物标志物特征包含至少一个囊泡内蛋白生物标志物时,在所述检测测定之前对所述靶生物标志物特征表达细胞外囊泡进行涉及固定和/或透化的处理。

24. 如权利要求20至23中任一项所述的方法,其中当所述第一和/或第二靶生物标志物特征包含至少一个表面蛋白生物标志物和/或囊泡内蛋白生物标志物时,所述检测测定涉及免疫测定(包括例如,免疫-PCR和/或邻近连接测定)。

25. 如权利要求24所述的方法,其中所述检测测定涉及邻近连接测定。

26. 如权利要求25所述的方法,其中所述邻近连接测定包括以下步骤:

(a) 将表达所述至少一个胞外囊泡相关膜结合多肽的所述靶生物标志物特征表达细胞外囊泡(“细胞外囊泡相关膜结合多肽生物表达细胞外囊泡”)与一组检测探针接触,每个检测探针都针对所述靶生物标志物特征的靶生物标志物,所述组包含至少两个检测探针,从而产生包含所述细胞外囊泡和所述检测探针组的组合。

其中所述检测探针各自包含:

(i) 针对所述靶生物标志物特征的所述靶生物标志物的靶结合部分;和

(ii) 偶联至所述靶结合部分的寡核苷酸结构域,所述寡核苷酸结构域包含双链部分和从所述寡核苷酸结构域的一端延伸的单链悬突部分,

其中所述检测探针的所述单链悬突部分的特征在于它们在所述检测探针结合至相同的细胞外囊泡时能彼此杂交;

(b) 将所述组合保持在允许所述检测探针组与其在所述细胞外囊泡上的各自靶标结合的条件下,使得所述至少两个检测探针可结合至表达所述靶生物标志物特征的相同的细胞外囊泡以形成双链复合物;

(c)使所述双链复合物与核酸连接酶接触以生成连接模板;以及

(d)检测所述连接模板,其中所述连接模板的存在指示所述靶生物标志物特征表达细胞外囊泡在血液来源样品中的存在;以及

(e)使用正交靶生物标志物特征任选地重复步骤a至d至少额外一次。

27.如权利要求26所述的方法,其中所述至少两个检测探针的靶结合部分针对相同靶生物标志物(例如,CEACAM6)。

28.如权利要求27所述的方法,其中所述至少两个检测探针的寡核苷酸结构域不同。

29.如权利要求19至28中任一项所述的方法,其中所述捕获测定的靶捕获部分是或包含至少一个针对至少一个细胞外囊泡相关膜结合多肽(例如,CEACAM5多肽、EpCAM多肽和/或SLC34A2多肽)的抗体剂。

30.如权利要求1至29中任一项所述的方法,其中进行所述方法以筛查所述受试者的早期肺癌、晚期肺癌或复发肺癌。

31.如权利要求1至30中任一项所述的方法,其中所述受试者具有以下特征中的至少一个或多个:

(i)易患肺癌(例如,处于肺癌平均群体风险(即,没有遗传风险)或具有遗传风险的)的无症状受试者;

(ii)具有肺癌家族病史的受试者(例如,具有一个或多个有肺癌史的一级亲属);

(iii)是吸烟者的受试者;

(iv)暴露于二手烟、氡气、石棉、沥青“烟煤”和/或其他致癌物(例如,砷、铬、镍、电离辐射、多环芳烃、一氧化氮、 $<2.5\mu\text{m}$ 的高水平颗粒物)的受试者;

(v)年龄在40岁以上的受试者;

(vi)具有一种或多种肺癌非特异性症状的受试者,任选地至少一种非特异性症状与一种或多种常见的呼吸道症状类似,所述呼吸道症状诸如咳嗽、咯血、气道阻塞、呼吸短促等,与非癌症疾病、病症或状况相关;

(vii)推荐进行胸部成像诸如X射线、CT扫描或低剂量CT扫描;

(viii)被诊断为具有影像学证实的肺部肿块的受试者;

(ix)具有肺部良性肿瘤的受试者;

(x)先前接受过肺癌治疗的受试者;

(xi)确定患有COPD的受试者;

(xii)确定患有肺纤维化的受试者;

(xiii)具有慢性支气管炎、结核病和/或肺炎病史的受试者;

(xiv)确定患有HIV和/或AIDS的受试者;

(xv)当前或曾经饮酒量高的受试者;

(xvi)具有EGFR、细胞色素p450酶和/或DNA修复基因遗传突变的受试者;和

(xvii)暴露于放射疗法和/或化学疗法的受试者。

32.如权利要求1至31中任一项所述的方法,其中所述方法与以下诊断测定中的一种或多种组合使用:

(i)受试者每年的体检;

(ii)胸部成像(例如,X射线、CT扫描或低剂量CT扫描);

(iii) 痰细胞学;

(iv) 筛选血浆中与癌症相关的循环肿瘤DNA和/或蛋白质生物标志物中基因突变的基因测定;和

(v) 涉及鉴定细胞表型和标志物表达的免疫荧光染色的测定,然后通过下一代测序进行扩增和分析。

33. 如权利要求1至32中任一项所述的方法,其中所述肺癌是小细胞肺癌或非小细胞肺癌。

34. 如权利要求1至33中任一项所述的方法,其中所述肺癌是非小细胞肺癌。

35. 如权利要求34所述的方法,其中所述非小细胞肺癌是肺腺癌。

36. 如权利要求1至35中任一项所述的方法,其中进行所述方法以监测肺癌患者对抗肺癌疗法(例如,手术、放疗、化学疗法、放射手术、靶向药物治疗、免疫治疗)的治疗的反应和/或癌症复发/转移。

37. 如权利要求1至35中任一项所述用于检测癌症的方法,其包括以下步骤:检测来自人血液样品中至少两个生物标志物在完整细胞外囊泡表面的共定位,所述生物标志物组合表达水平已被确定与癌症相关;将检测到的共定位水平与确定的水平进行比较;以及当检测到的共定位水平处于或高于确定的水平时检测癌症。

38. 如权利要求1至35中任一项所述用于检测癌症的方法,其包括以下步骤:将包含外泌体的样品与一组特异性结合所述外泌体上的表面生物标志物的检测探针接触,以检测样品中与癌症相关的外泌体,特异性在95%至100%范围内,灵敏度在30%至100%范围内。

39. 如权利要求1至35中任一项所述的方法,其包括以下步骤:用捕获剂从生物样品中捕获外泌体,所述捕获剂与所述外泌体上的癌症特异性表面蛋白生物标志物选择性相互作用;并将捕获的外泌体与至少一组至少两个检测探针接触,每个探针选择性地与所述外泌体上的表面蛋白生物标志物相互作用;以及当所述组中至少两个检测探针足够紧密邻近时,检测到形成的产物,此类检测指示了表面蛋白生物标志物的共定位。

40. 如权利要求1至35中任一项所述的方法,其包括以下步骤:将包含外泌体的样品与一组特异性结合所述外泌体上的表面生物标志物的探针接触,以检测所述样品中的癌症相关外泌体,其中:(i)所述组中的每个探针包含针对所述外泌体上的靶结合部分;和(ii)所述组包含至少一个捕获探针和至少两个检测探针,其中每个检测探针还包含一个检测单元。

41. 如权利要求1至35中任一项所述的方法,其包括以下步骤:进行检测来自人受试者的外泌体上的表面生物标志物特征的邻近性测定,此步骤在进行事先的检测来自所述人受试者外泌体上的表面生物标志物特征的测定进行一段时间后进行;以及

将所进行的测定结果与之前的测定结果进行比较。

42. 如权利要求1至35中任一项所述的方法,其包括以下步骤:将外泌体与至少两个检测探针接触,其中每个检测探针包含(i)结合部分;和(ii)寡核苷酸实体,其中所述结合部分相同并且所述寡核苷酸实体彼此互补。

43. 如权利要求1至35中任一项所述的方法,其包括检测外泌体表面的标志物邻近性,包括将所述外泌体与至少一对结合剂接触的改进,每个结合剂包含结合部分和邻近部分,其中所述结合部分相同,并且所述邻近部分彼此互补;以及检测邻近部分之间的相互作用。

44. 一种用于检测肺癌的试剂盒,所述试剂盒包含:

(a) 捕获剂,其包含针对细胞外囊泡相关膜结合多肽的靶捕获部分;和

(b) 至少一组检测探针,所述组包含至少两个检测探针,每个检测探针针对肺癌的靶生物标志物特征的靶生物标志物,其中所述检测探针各自包含:

(i) 针对所述肺癌靶生物标志物特征的靶生物标志物的靶结合部分;以及

(ii) 偶联至所述靶结合部分的寡核苷酸结构域,所述寡核苷酸结构域包含双链部分和从所述寡核苷酸结构域的一端延伸的单链悬突部分,

其中所述至少两个检测探针的单链悬突部分的特征在于它们在所述至少两个检测探针结合至相同的细胞外囊泡时能彼此杂交;

其中所述肺癌的靶生物标志物特征包含:

至少一个细胞外囊泡相关膜结合多肽和

选自自由以下组成的组的至少一个靶生物标志物:表面蛋白生物标志物、囊泡内蛋白生物标志物和囊泡内RNA生物标志物,其中:

所述表面蛋白生物标志物选自ALCAM、ABCC3、ARSL、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、DMBT1、DSG2、EGFR、EPCAM、EPHX3、EVA1A、FOLR1、GJB1、GJB2、GPC4、HS6ST2、KDEL3、KRTCAP3、LAMB3、LFNG、LSR、MANEAL、MET、MSLN、MUC1、MUC21、PIGT、PODXL2、PRRG4、ROS1、SDC1、SERINC2、SEZ6L2、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、TNFRSF10B及其组合;

所述囊泡内蛋白生物标志物选自AOC1、C12orf45、CRABP2、CST1、ETV4、FAM83A、FOXA2、HMGB3、LGALS3BP、MIF、NAPSA、PPP1R14D、S100A14、SBK1、SCGB3A2、SFTA2、SFTPA1、SFTPA2、SFTPB、SPINK1、TGFA、ZC3H11A及其组合;

所述囊泡内RNA生物标志物选自ABCC3、AOC1、ARSL、B3GNT3、C12orf45、CDCP1、CDH1、CDH3、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CRABP2、CST1、DMBT1、DSG2、EPCAM、EPHX3、ETV4、EVA1A、FAM83A、FOLR1、FOXA2、GJB1、GJB2、GPC4、HMGB3、HS6ST2、KDEL3、KRTCAP3、LAMB3、LFNG、LGALS3BP、LSR、MANEAL、MIF、MSLN、MUC1、MUC21、NAPSA、PIGT、PODXL2、PPP1R14D、PRRG4、ROS1、S100A14、SBK1、SCGB3A2、SDC1、SERINC2、SEZ6L2、SFTA2、SFTPA1、SFTPA2、SFTPB、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、SPINK1、ST14、TGFA、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、ZC3H11A及其组合。

45. 如权利要求44所述的试剂盒,其中当所述至少一个靶生物标志物选自所述表面蛋白生物标志物中的一个或多个时,所述选定的一个或多个表面蛋白生物标志物和所述至少一个细胞外囊泡相关膜结合多肽不同。

46. 如权利要求44或45所述的试剂盒,其中所述细胞外囊泡相关膜结合多肽是或包含ALCAM、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CLDN3、CLDN4、DSG2、EGFR、EPCAM、FOLR1、GJB1、GJB2、IG1FR、LAMB3、MET、MSLN、MUC1、PIGT、PODXL2、ROS1、SDC1、SLC34A2、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMPRSS4、TSPAN8、TNFRSF10B或其组合。

47. 一种用于检测肺癌的试剂盒,所述试剂盒包含:

(a) 捕获剂,其包含针对细胞外囊泡相关膜结合多肽的靶捕获部分;和

(b) 至少一组检测探针, 所述组包含至少两个检测探针, 每个检测探针针对肺癌的靶生物标志物特征的靶生物标志物, 其中所述检测探针各自包含:

(i) 针对所述肺癌靶生物标志物特征的靶生物标志物的靶结合部分; 以及

(ii) 偶联至所述靶结合部分的寡核苷酸结构域, 所述寡核苷酸结构域包含双链部分和从所述寡核苷酸结构域的一端延伸的单链悬突部分,

其中所述至少两个检测探针的单链悬突部分的特征在于它们在所述至少两个检测探针结合至相同的细胞外囊泡时能彼此杂交;

其中所述肺癌的靶生物标志物特征包含:

至少一个细胞外囊泡相关膜结合多肽和

至少一个靶生物标志物, 其中所述靶生物标志物是或包含表面蛋白生物标志物, 其中所述靶生物标志物特征包含如表4中所列的生物标志物组合。

48. 如权利要求44至46中任一项所述的试剂盒, 其中所述细胞外囊泡相关膜结合多肽是或包含CEACAM5多肽和/或SLC34A2多肽。

49. 如权利要求44至48中任一项所述的试剂盒, 其中所述至少两个检测探针的靶结合部分各自针对所述靶生物标志物特征的相同的靶生物标志物。

50. 如权利要求49所述的试剂盒, 其中所述相同的靶生物标志物是或包含CEACAM6。

51. 如权利要求50所述的试剂盒, 其中所述至少两个检测探针的寡核苷酸结构域不同。

52. 如权利要求44至48中任一项所述的试剂盒, 其中所述至少两个检测探针的靶结合部分各自针对所述靶生物标志物特征的不同的靶生物标志物。

53. 如权利要求52所述的试剂盒, 其中所述细胞外囊泡相关膜结合多肽是或包含CEACAM5多肽和/或SLC34A2多肽; 并且所述至少两个检测探针分别针对CEACAM6和EPCAM。

54. 如权利要求52所述的试剂盒, 其中所述细胞外囊泡相关膜结合多肽是或包含CEACAM5多肽; 并且所述至少两个检测探针分别针对CEACAM6和SLC34A2。

55. 如权利要求44至54中任一项所述的试剂盒, 其还包含至少一种另外的试剂(例如, 连接酶、固定剂和/或透化剂)。

56. 如权利要求44至55中任一项所述的试剂盒, 其包含至少两组(包括例如, 至少三组)检测探针, 每组包含至少两个各自针对肺癌的不同靶生物标志物特征的靶生物标志物的检测探针。

57. 如权利要求44至46中任一项所述的试剂盒, 其包含:

(a) 包含靶捕获部分的第一捕获剂;

(b) 包含靶捕获部分的第二捕获剂;

(c) 至少两组检测探针, 其中所述检测探针各自包含:

(i) 针对靶表面蛋白生物标志物的靶结合部分; 和

(ii) 偶联至所述靶结合部分的寡核苷酸结构域, 所述寡核苷酸结构域包含双链部分和从所述寡核苷酸结构域的一端延伸的单链悬突部分,

其中所述至少两个检测探针的单链悬突部分的特征在于它们在所述至少两个检测探针结合至相同的细胞外囊泡时能彼此杂交。

58. 如权利要求44至46中任一项所述的试剂盒, 其包含:

(a) 包含靶捕获部分的第一捕获剂;

(b) 包含靶捕获部分的第二捕获剂;

(c) 包含靶捕获部分的第三捕获剂;

(d) 至少三组检测探针,其中所述检测探针各自包含:

(i) 针对靶表面蛋白生物标志物的靶结合部分;和

(ii) 偶联至所述靶结合部分的寡核苷酸结构域,所述寡核苷酸结构域包含双链部分和从所述寡核苷酸结构域的一端延伸的单链悬突部分,

其中所述至少两个检测探针的单链悬突部分的特征在于它们在所述至少两个检测探针结合至相同的细胞外囊泡时能彼此杂交。

59. 一种复合物,所述复合物包含:

(a) 表达肺癌的靶生物标志物特征的细胞外囊泡,其中所述靶生物标志物特征包含:

至少一个细胞外囊泡相关膜结合多肽和

选自以下组成的组的至少一个靶生物标志物:表面蛋白生物标志物、囊泡内蛋白生物标志物和囊泡内RNA生物标志物,其中:

所述表面蛋白生物标志物选自ABCC3、ALCAM、ARSL、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、DMBT1、DSG2、EGFR、EPCAM、EPHX3、EVA1A、FOLR1、GJB1、GJB2、GPC4、IG1FR、HS6ST2、KDEL3、KRTCAP3、LAMB3、LFNG、LSR、MANEAL、MET、MSLN、MUC1、MUC21、PIGT、PODXL2、PRRG4、ROS1、SDC1、SERINC2、SEZ6L2、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、TNFRSF10B及其组合;

所述囊泡内蛋白生物标志物选自AOC1、C12orf45、CRABP2、CST1、ETV4、FAM83A、FOXA2、HMGB3、LGALS3BP、MIF、NAPSA、PPP1R14D、S100A14、SBK1、SCGB3A2、SFTA2、SFTPA1、SFTPA2、SFTPB、SPINK1、TGFA、ZC3H11A及其组合;

所述囊泡内RNA生物标志物选自ABCC3、AOC1、ARSL、B3GNT3、C12orf45、CDCP1、CDH1、CDH3、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CRABP2、CST1、DMBT1、DSG2、EPCAM、EPHX3、ETV4、EVA1A、FAM83A、FOLR1、FOXA2、GJB1、GJB2、GPC4、HMGB3、HS6ST2、KDEL3、KRTCAP3、LAMB3、LFNG、LGALS3BP、LSR、MANEAL、MIF、MSLN、MUC1、MUC21、NAPSA、PIGT、PODXL2、PPP1R14D、PRRG4、ROS1、S100A14、SBK1、SCGB3A2、SDC1、SERINC2、SEZ6L2、SFTA2、SFTPA1、SFTPA2、SFTPB、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、SPINK1、ST14、TGFA、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、ZC3H11A及其组合;

其中所述细胞外囊泡固定在固体基底上,所述固体基底包含针对所述细胞外囊泡相关膜结合多肽的靶捕获部分;

(b) 各自结合到所述细胞外囊泡的第一检测探针和第二检测探针,其中每个检测探针包含:

(i) 针对所述肿瘤靶标生物标志物特征的靶标生物标志物之一的靶标结合部分;和

(ii) 偶联至所述靶结合部分的寡核苷酸结构域,所述寡核苷酸结构域包含双链部分和从所述寡核苷酸结构域的一端延伸的单链悬突部分,

其中所述第一检测探针和所述第二检测探针的单链悬突部分彼此杂交。

60. 如权利要求59所述的复合物,其中当所述至少一个靶生物标志物选自所述表面蛋

白生物标志物中的一个或多个时,所述选定的一个或多个表面蛋白生物标志物和所述至少一个细胞外囊泡相关膜结合蛋白生物标志物不同;

61. 如权利要求59或60所述的试剂盒,其中所述细胞外囊泡相关膜结合多肽是或包含ALCAM、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CLDN3、CLDN4、DSG2、EGFR、EPCAM、FOLR1、GJB1、GJB2、IG1FR、LAMB3、MET、MSLN、MUC1、PIGT、PODXL2、ROS1、SDC1、SLC34A2、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMPRSS4、TSPAN8、TNFRSF10B或其组合。

62. 一种复合物,所述复合物包含:

(a) 表达肺癌的靶生物标志物特征的细胞外囊泡,其中所述靶生物标志物特征包含:

至少一个细胞外囊泡相关膜结合多肽和

至少一个靶生物标志物,其中所述靶生物标志物是或包含表面蛋白生物标志物,其中所述靶生物标志物特征是或包含如表4中所列的生物标志物组合;

其中所述细胞外囊泡固定在固体基底上,所述固体基底包含针对所述细胞外囊泡相关膜结合多肽的靶捕获部分;

(b) 各自结合到所述细胞外囊泡的第一检测探针和第二检测探针,其中每个检测探针包含:

(i) 针对所述肿瘤靶标生物标志物特征的靶标生物标志物之一的靶标结合部分;和

(ii) 偶联至所述靶结合部分的寡核苷酸结构域,所述寡核苷酸结构域包含双链部分和从所述寡核苷酸结构域的一端延伸的单链悬突部分,

其中所述第一检测探针和所述第二检测探针的单链悬突部分彼此杂交。

63. 如权利要求59至61中任一项所述的复合物,其中所述细胞外囊泡相关膜结合多肽是或包含CEACAM5多肽和/或SLC34A2多肽和/或CLDN6多肽。

64. 如权利要求59至62中任一项所述的复合物,其中所述至少两个检测探针的靶结合部分各自针对所述靶生物标志物特征的相同的靶生物标志物。

65. 如权利要求64所述的复合物,其中所述相同的靶生物标志物是或包含CEACAM6。

66. 如权利要求64所述的复合物,其中所述至少两个检测探针的寡核苷酸结构域不同。

67. 如权利要求59至62中任一项所述的复合物,其中所述至少两个检测探针的靶结合部分各自针对所述靶生物标志物特征的不同靶生物标志物。

68. 如权利要求67所述的复合物,其中所述细胞外囊泡相关膜结合多肽是或包含CEACAM5多肽和/或SLC34A2多肽;并且所述至少两个检测探针分别针对CEACAM6和EPCAM。

69. 如权利要求67所述的复合物,其中所述细胞外囊泡相关膜结合多肽是或包含CEACAM5多肽;并且所述至少两个检测探针分别针对CEACAM6和SLC34A2。

70. 如权利要求59至69中任一项所述的复合物,其中所述固体基底包含磁珠。

71. 如权利要求59至70中任一项所述的复合物,其中所述靶捕获部分是或包含抗体剂。

72. 如权利要求59至71中任一项所述的复合物,其包含:(a) 其表面具有至少一个靶生物标志物的外泌体;和(b) 每个都结合到所述外泌体的第一检测探针和第二检测探针,其中所述第一检测探针和第二检测探针的每一个包含:(i) 针对所述外泌体表达的靶生物标志物的靶结合部分;和(ii) 与靶结合部分偶联的寡核苷酸域,所述寡核苷酸域包含双链部分和从所述寡核苷酸域一端延伸的单链悬突部分,其中所述第一和第二检测探针的单链悬垂

部分彼此杂交。

73. 如权利要求59至71中任一项所述的复合物,其包含由来自人血液样品的结合一组至少两个探针的细胞外囊泡,每个探针包含生物标志物结合部分和寡核苷酸结构域,其中两个或更多个结合的探针相互靠近,使得其寡核苷酸结构域相互杂交以形成可连接的杂合体。

74. 如权利要求59至71中任一项所述的复合物,其包含:(a) 包含癌症相关靶生物标志物的外泌体;和(b) 每个都结合到所述外泌体的至少第一检测探针和第二检测探针,其中所述检测探针中的每一个包含:(i) 针对所述靶生物标志物特征的靶结合部分;和(ii) 偶联至所述结合部分的寡核苷酸结构域,所述寡核苷酸结构域包含双链部分和从所述寡核苷酸结构域的一端延伸的单链悬突部分,其中所述检测探针的单链悬突部分至少部分地互补。

75. 一组用于如权利要求1至71中任一项的方法、试剂盒或复合物的探针,其中每组探针包含:(a) 生物标志物结合部分,其特异性结合来自癌细胞的细胞外囊泡上的表面生物标志物;和(b) 寡核苷酸结构域,其中所述组内探针的寡核苷酸结构域被排列并构建使得当所述探针结合其靶生物标志物时,只有当所述靶生物标志物彼此邻近时,其寡核苷酸结构域互相杂交以形成可连接的杂合体。

用于检测肺癌的组合物和方法

[0001] 相关申请的交叉引用

[0002] 本申请要求2020年7月8日提交的美国临时申请序列号63/049,538的权益,所述申请的内容全部并入本文。

背景技术

[0003] 癌症的早期检测大大增加了成功治疗的机会。然而,包括肺癌在内的许多癌症仍然缺乏有效的筛查建议。癌症筛查测试的典型挑战包括有限的灵敏度和特异性。假阳性结果率高可能特别令人担忧,因为它会给不想不必要地施用(或接受)可能具有不希望的副作用的抗癌疗法的临床医生和患者带来困难的管理决策。相反,假阴性结果率高无法满足筛查测试的目的,因为错过了需要治疗的患者,从而导致治疗延迟,并因此降低成功的可能性。

发明内容

[0004] 本公开尤其提供了用于实现有效肺癌筛查的见解和技术。在一些实施方案中,本公开尤其提供了对非小细胞肺癌筛查特别有用的见解和技术。在一些实施方案中,提供的技术对于检测早期肺癌(例如,在一些实施方案中为非小细胞肺癌)是有效的。在一些实施方案中,提供的技术即使在应用于包含无症状个体或由无症状个体组成的群体时也是有效的(例如,由于足够高的灵敏度和/或低的假阳性和/或假阴性结果率)。在一些实施方案中,提供的技术在应用于包含没有发展肺癌(例如,在一些实施方案中为非小细胞肺癌)的遗传风险的个体(例如,无症状个体)或由所述个体组成的群体时是有效的。在一些实施方案中,提供的技术在应用于包含有症状个体(例如,遭受一种或多种肺癌症状的个体)或由所述个体组成的群体时是有效的。在一些实施方案中,提供的技术在应用于包含具有肺癌风险的个体(例如,具有肺癌的遗传和/或生活史相关风险因素的个体)或由所述个体组成的群体时是有效的。在一些实施方案中,提供的技术可以是或包括一种或多种组合物(例如,分子实体或复合物、系统、细胞、集合、组合、试剂盒等)和/或方法(例如,制造、使用、评估等),如对于阅读本文提供的公开的本领域技术人员来说将是清楚的。

[0005] 在一些实施方案中,本公开用某些现有技术(包括例如检测和诊断肺癌的某些常规方法)确定问题的根源。例如,本公开认识到许多常规的诊断测定,例如X射线成像、痰液测试、低剂量CT扫描和/或基于游离核酸、血清蛋白(例如,癌胚抗原(CEA)、细胞角蛋白19片段(CYFRA 21-1)、神经元特异性烯醇化酶(NSE)、血清胃泌素释放肽前体(ProGRP)和/或鳞状细胞癌抗原(SCCA))的分子测试,和/或细胞外囊泡的整体分析,可能耗时、昂贵和/或缺乏足以提供可靠和全面的诊断评估的灵敏度和/或特异性。在一些实施方案中,本公开尤其提供通过检测个体细胞外囊泡中肺癌的靶生物标志物特征的共定位来解决此类问题的技术(包括系统、组合物和方法),所述细胞外囊泡包含至少一个细胞外囊泡相关膜结合多肽和选自以下组成的组的至少一个靶生物标志物:表面蛋白生物标志物、内部蛋白生物标志物和RNA生物标志物。在一些实施方案中,本公开尤其提供了通过使用基于个体细胞外囊

泡中至少两个或更多个靶实体(例如,靶生物标志物特征)的相互作用和/或共定位的靶实体检测方法(由申请人开发并描述于均在2020年2月28日提交且名为“Systems, Compositions, and Methods for Target Entity Detection”的美国申请号16/805,637(公布为US2020/0299780)和国际申请PCT/US2020/020529(公布为W02020180741))检测肺癌的此类靶生物标志物特征来解决此类问题的技术(包括系统、组合物和方法)。

[0006] 在一些实施方案中,本公开尤其提供了以下见解:对无症状个体进行筛查(例如,在出现症状之前或以其他方式在没有出现症状的情况下进行定期筛查)可能是有益的,并且甚至对于肺癌(例如,在一些实施方案中为非小细胞肺癌)的有效管理(例如,成功治疗)是重要的。在一些实施方案中,每本公开提供了在一些实施方案中在无症状个体中可以被实施以检测肺癌(例如,在一些实施方案中为非小细胞肺癌),包括早期癌症的肺癌筛查系统。在一些实施方案中,提供的技术被实施以实现无症状个体的定期筛查。本公开提供了例如组合物(例如,试剂、试剂盒、组分等)以及提供和/或使用它们的方法,包括涉及对一个或多个个体(例如,有症状或无症状个体)进行定期测试的策略。本公开定义了此类系统的有用性,并提供了用于实施它们的组合物和方法。

[0007] 在一些实施方案中,提供的技术实现对肺癌的一个或多个特征(例如,发病、进展、对治疗的反应性、复发等)的检测(例如,早期检测,例如,在一个或多个无症状个体和/或一个或多个群体中),其中灵敏度和/或特异性(例如,假阳性和/或假阴性结果的比率)适用于允许将提供的技术有用地应用于单次和/或定期(例如,周期性)评估。在一些实施方案中,提供的技术可用于与常规医疗检查结合,所述常规医疗检查诸如但不限于:体检、全科医生就诊、胆固醇/血脂测试、糖尿病(2型)筛查、结肠镜检查、血压筛查、甲状腺功能测试、前列腺癌筛查、乳腺图像、HPV/Pap涂片和/或预防接种。在一些实施方案中,提供的技术可用于与一种或多种治疗方案结合;在一些实施方案中,提供的技术可以改进此类一种或多种治疗方案的一个或多个特征(例如,根据接受的参数的成功率)。

[0008] 在一些方面,提供了用于将受试者(例如,无症状受试者)分类为患有或易患肺癌(例如,在一些实施方案中为非小细胞肺癌)的技术。在一些实施方案中,本公开提供了用于将受试者(例如,无症状受试者)分类为患有或易患肺癌(例如,在一些实施方案中为非小细胞肺癌)的方法或测定。在一些实施方案中,提供的方法或测定包括(a)在来自有需要的受试者的血液来源样品中检测表达肺癌的靶生物标志物特征的细胞外囊泡,所述靶生物标志物特征包含:至少一个细胞外囊泡相关膜结合多肽和至少一个选自由表面蛋白生物标志物、囊泡内蛋白生物标志物和囊泡内RNA生物标志物组成的组的靶生物标志物,其中所述表面蛋白生物标志物选自ALCAM、ABCC3、ARSL、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、DMBT1、DSG2、EGFR、EPCAM、EPHX3、EVA1A、FOLR1、GJB1、GJB2、GPC4、HS6ST2、KDEL3、KRTCAP3、IG1FR、LAMB3、LFNG、LSR、MANEAL、MET、MSLN、MUC1、MUC21、PIGT、PODXL2、PRRG4、ROS1、SDC1、SERINC2、SEZ6L2、SLC34A2、sTn抗原、Tn抗原、T抗原、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、ST14、TACSTD2、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TNFRSF10B、TSPAN1、TSPAN8及其组合;所述囊泡内蛋白生物标志物选自AOC1、C12orf45、CRABP2、CST1、ETV4、FAM83A、FOXA2、HMGB3、LGALS3BP、MIF、NAPSA、PPP1R14D、S100A14、SBK1、SCGB3A2、SFTA2、SFTPA1、SFTPA2、SFTPB、SPINK1、TGFA、ZC3H11A及其组合;并且所述囊泡内RNA(例如,mRNA)生物标志物选自

ABCC3、AOC1、ARSL、B3GNT3、C12orf45、CDCP1、CDH1、CDH3、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CRABP2、CST1、DMBT1、DSG2、EPCAM、EPHX3、ETV4、EVA1A、FAM83A、FOLR1、FOXA2、GJB1、GJB2、GPC4、HMGB3、HS6ST2、KDELR3、KRTCAP3、LAMB3、LFNG、LGALS3BP、LSR、MANEAL、MIF、MSLN、MUC1、MUC21、NAPSA、PIGT、PODXL2、PPP1R14D、PRRG4、ROS1、S100A14、SBK1、SCGB3A2、SDC1、SERINC2、SEZ6L2、SFTA2、SFTPA1、SFTPA2、SFTPB、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、SPINK1、ST14、TGFA、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、ZC3H11A及其组合；(b) 将指示血液来源样品中靶生物标志物特征表达细胞外囊泡水平的样品信息与包括参考阈值水平在内的参考信息进行比较；以及(c) 当血液来源样品相对于参考参考阈值水平的分类截止值显示出靶生物标志物特征表达细胞外囊泡水平升高时，将受试者分类为患有或易患肺癌。

[0009] 在一些方面，提供了用于将受试者(例如，无症状受试者)分类为患有或易患肺癌(例如，在一些实施方案中为非小细胞肺癌)的技术。在一些实施方案中，本公开提供了用于将受试者(例如，无症状受试者)分类为患有或易患肺癌(例如，在一些实施方案中为非小细胞肺癌)的方法或测定。在一些实施方案中，提供的方法或测定包括(a) 在来自有需要的受试者的血液来源样品中检测表达肺癌(例如，在一些实施方案中为非小细胞肺癌)的靶生物标志物特征的细胞外囊泡，所述靶生物标志物特征包含：至少一个细胞外囊泡相关膜结合多肽和至少一个选自由表面蛋白生物标志物(如本文所述)、囊泡内蛋白生物标志物(如本文所述)和囊泡内RNA生物标志物(如本文所述)组成的组的靶生物标志物；(b) 将指示血液来源样品中靶生物标志物特征表达细胞外囊泡水平的样品信息与包括参考阈值水平在内的参考信息进行比较；以及(c) 当血液来源样品相对于参考参考阈值水平的分类截止值显示出靶生物标志物特征表达细胞外囊泡水平升高时，将受试者分类为患有或易患肺癌。在一些实施方案中，至少一个此类靶生物标志物是或包含选自以下的表面蛋白生物标志物：ABCA3、ABCC1、ABCC3、ACBD3、ACSL5、AGER、ALCAM、AP1M2、APH1A、APOO、ATP11A、ATP11B、ATP1B1、ATP6AP2、B4GALT4、BCAP31、BSPRY、CD109、CD55、CD9、CDC42、CDH1、CDH3、CDKAL1、CEACAM5、CEACAM6、CELSR1、CIP2A、CISD2、CKAP4、CLCA2、CLDN1、CLIC6、CLPTM1L、CLSTN1、CNTN1、CPD、CYP2S1、CYP4F11、CYP4F3、DPY19L1、DSC2、DSC3、DSG2、DSG3、EGFR、EPCAM、EPHB3、FAT2、FBXO45、FERMT1、FOLR1、FZD6、GALNT1、GALNT3、GALNT5、GALNT6、GGCX、GOLM1、GOLPH3L、GRHL2、HACD3、IER3IP1、IGSF3、IL1RAP、ITGA2、ITGB6、KLRG2、KPNA2、KRTCAP3、LAD1、LAMB3、LAMC2、LAMP3、LAMTOR2、LCLAT1、LPCAT1、LSR、MAGT1、MARCKSL1、MET、MGAT1、MSLN、MUC1、MUC4、NCSTN、NECTIN1、NECTIN4、NRAS、NT5E、NUP210、PARL、PEX13、PIGN、PIGT、PLA2G4A、PLCH1、PLEC、PSMD2、PTDSS1、PTGFRN、PTPRF、QSOX1、RAB25、RAB38、RAB6B、RAP2B、RCC2、RIT1、SCAMP3、SDC1、SEL1L3、SHROOM2、SLC2A1、SLC34A2、SLC35B2、SLC39A11、SMPDL3B、SOAT1、SPAST、SSR1、SSR4、SURF4、SYNGR2、TACSTD2、TESC、TFRC、TMC5、TMC01、TMED2、TMED3、TMEM132A、TMEM33、TMPRSS4、TMTC3、TOMM22、TOR1AIP2、TRAM1、TRPV4、TTC33、UGT1A6、UPK1B、VAMP8、VMA21、VRK2、VWA1、XPR1、XXYLT1、ADAM28、AXL、BSG、CD274、CD47、CLU、DKK1、ERBB3、FLT4、GM3、HGF、IGF1R、IL6、KDR、LAG3、Lewis Y/B抗原、LY6E、NOTCH2、NOTCH3、磷脂酰丝氨酸、TIGIT、TNFRSF10A、TNFRSF10B、TNFSF18、TPBG、VEGFA、Tn抗原、Lewis Y/CD174、唾液酸化Lewis X(sLex)(也称为唾液酸化SSEA-1(SLX))、NeuGcGM3及其组合。在一些实施方案中，至少一个此类靶生物标志物是或包含选自以下的囊泡内蛋白生物标志物：ABRACL、ACP5、

ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、AOC1、AP1M2、APOBEC3B、APOBEC3C、ARNTL2、ASF1B、AURKB、BAIAP2L1、BIRC5、C12orf45、C15orf48、C19orf33、C1S、C8orf4、CA9、CALML3、CAPNS2、CBLC、CCL19、CCNB2、CDC20、CDC45、CDCA4、CDCA5、CDK1、CDKN2A、CDKN2B、CENPW、CEP55、CES1、CHMP4C、CNN2、CPA3、CRABP2、CST1、CSTA、CTSC、CTSE、CYP2S1、DPYSL3、EFS、EGLN3、EHF、ELF3、ELF4、ENAH、ESRP1、ETV4、EVPL、FAM129B、FAM60A、FAM83A、FAM83D、FAM83H、FBP1、FERMT1、FOXA2、FOX E1、FOX M1、GBP6、GNA15、GPX2、GRHL2、GSTA1、HCK、HMGB3、HOXB7、ID1、IGF2BP2、IMPA2、IRF6、IVL、JUP、KIAA1522、KIF2C、KIFC1、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、LGALS3BP、LGALS7B、LSP1、MAGEA4、MAGEA6、MCM2、MDFI、MIF、MYBL2、MYH14、MZB1、NAPSA、NCF2、NNMT、NRARP、NUP210、NUSAP1、OSGIN1、PALLD、PITX1、PKP1、PKP3、PLEK、PLEK2、POSTN、PPP1R14C、PPP1R14D、PRAME、PTPN6、RBP1、RIN2、RIPK4、RPS4Y1、RRM2、S100A11、S100A14、S100A16、S100A2、S100P、SBK1、SCGB3A2、SERPINB13、SERPINB3、SERPINB5、SFTA2、SFTPA1、SFTPA2、SFTPB、SH3BP4、SNAI2、SOX2、SPI1、SPINK1、SPINT1、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、SULF1、SYK、SYTL1、TBC1D2、TEAD2、TEAD3、TFAP2C、TGFA、THBS2、TK1、TOP2A、TP63、TPD52、TPX2、TRIM29、TRIP13、UBE2C、YAP1、ZC3H11A、ZNF217、ZNF750及其组合。在一些实施方案中,至少一个此类靶生物标志物是或包含选自以下的囊泡内RNA(例如,mRNA):ABCA3、ABCC1、ABCC3、ABRACL、ACP5、ADAM23、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、ANTXR1、AOC1、AP1M2、APOBEC3B、APOBEC3C、AQP3、AREG、ARNTL2、ARSL、ASF1B、ATP8B1、AURKB、B3GNT3、B3GNT5、BAIAP2L1、BCAM、BIK、BIRC5、C12orf45、C15orf48、C19orf33、C1S、C8orf4、CA12、CA9、CALML3、CAPNS2、CBLC、CCL19、CCL5、CCNB2、CD109、CD24、CD53、CD74、CD9、CDC20、CDC42EP1、CDC45、CDCA4、CDCA5、CDCP1、CDH1、CDH3、CDK1、CDKN2A、CDKN2B、CEACAM5、CEACAM6、CELSR1、CENPW、CEP55、CES1、CHMP4C、CLCA2、CLDN1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CNN2、COL17A1、CPA3、CRABP2、CST1、CSTA、CTSC、CTSE、CX3CL1、CXADR、CXCR4、CYBB、CYP2S1、CYP4F11、DAPL1、DMBT1、DPYSL3、DSC2、DSC3、DSG2、DSG3、DSP、EFNA1、EFS、EGFR、EGLN3、EHD2、EHF、ELF3、ELF4、EMP1、EMP2、ENAH、EPCAM、EPHA2、EPHB3、EPHX1、EPHX3、ESRP1、ETV4、EVA1A、EVPL、F11R、F2R、F2RL1、F3、FAM129B、FAM60A、FAM83A、FAM83D、FAM83H、FAT1、FAT2、FBLIM1、FBP1、FCER1G、FERMT1、FGFR2、FGFR3、FOLR1、FOXA2、FOX E1、FOX M1、FX YD3、GALNT3、GBP6、GJA1、GJB1、GJB2、GJB3、GJB5、GJB6、GNA15、GPC1、GPC3、GPC4、GPNMB、GPR87、GPC5A、GPX2、GRHL2、GSTA1、HAS3、HCK、HMGB3、HOXB7、HS6ST2、ID1、IGF2BP2、IGSF9、IL2RG、IMPA2、IRF6、ITGA2、ITGA6、ITGB4、ITGB6、IVL、JAG2、JUP、KCNS3、KDEL R3、KIAA1522、KIF2C、KIFC1、KITLG、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、KRTCAP3、LAMB3、LAMP3、LAPTM5、LFNG、LGALS3BP、LGALS7B、LRP11、LRRC4、LSP1、LSR、LYPD3、MAGEA4、MAGEA6、MAL2、MANEAL、MAOA、MARCO、MCM2、MDFI、MET、MIF、MMP14、MPZL2、MSLN、MUC1、MUC21、MYBL2、MYH14、MYOF、MZB1、NAPSA、NCF2、NKG7、NNMT、NOTCH3、NRARP、NTRK2、NUP210、NUSAP1、OSGIN1、OSMR、PALLD、PDPN、PDZK1IP1、PECAM1、PERP、PIGR、PIGT、PITX1、PKP1、PKP3、PLEK、PLEK2、PLVAP、PMP22、PODXL2、POSTN、PPL、PPP1R14C、PPP1R14D、PRAME、PROM2、PRRG4、PRSS8、PTGES、PTGFRN、PTPN6、PTPRF、PTPRZ1、RAB25、RAB38、RAET1L、RARRES1、RBP1、RGS1、RHCG、RHOV、RIN2、RIPK4、ROS1、RPS4Y1、RRM2、S100A10、S100A11、

S100A14、S100A16、S100A2、S100P、SBK1、SCGB3A2、SCNN1A、SDC1、SERINC2、SERPINB13、SERPINB3、SERPINB5、SEZ6L2、SFTA2、SFTPA1、SFTPA2、SFTPB、SH3BP4、SHISA2、SLC1A5、SLC2A1、SLC34A2、SLC40A1、SLC44A4、SLC6A14、SLC6A8、SLC7A7、SLC7A8、SMIM22、SMPDL3B、SNAI2、SOX2、SPI1、SPINK1、SPINT1、SPINT2、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、ST14、STEAP1、SULF1、SYK、SYTL1、TACSTD2、TBC1D2、TEAD2、TEAD3、TFAP2C、TGFA、THBD、THBS2、TK1、TM4SF1、TMC4、TMC5、TMEM30B、TMEM45B、TMEM54、TMPRSS11D、TMPRSS11E、TMPRSS2、TMPRSS4、TNFRSF18、TNS4、TOP2A、TP53I11、TP63、TPD52、TPX2、TREM2、TRIM29、TRIP13、TSPAN1、TSPAN13、TSPAN6、TSPAN7、TSPAN8、TUSC3、TYROBP、UBE2C、UPK1B、VAMP8、VANGL2、WLS、YAP1、ZC3H11A、ZNF217、ZNF750及其组合。

[0010] 在一些方面,提供了用于将受试者(例如,无症状受试者)分类为患有或易患肺癌(例如,在一些实施方案中为非小细胞肺癌例如像肺腺癌(LUAD)和/或肺鳞状细胞癌(LUSC))的技术。在一些实施方案中,本公开提供了用于将受试者(例如,无症状受试者)分类为患有或易患肺癌(例如,在一些实施方案中为非小细胞肺癌例如像LUAD和/或LUSC)的方法或测定。在一些实施方案中,提供的方法或测定包括(a)在来自有需要的受试者的血液来源样品中检测表达肺癌(例如,在一些实施方案中为非小细胞肺癌例如像LUAD和/或LUSC)的靶生物标志物特征的细胞外囊泡,所述靶生物标志物特征包含:至少一个细胞外囊泡相关膜结合多肽和至少一个选自由表面蛋白生物标志物(如本文所述)、囊泡内蛋白生物标志物(如本文所述)和囊泡内RNA生物标志物(如本文所述)组成的组的靶生物标志物;(b)将指示血液来源样品中靶生物标志物特征表达细胞外囊泡水平的样品信息与包括参考阈值水平在内的参考信息进行比较;以及(c)当血液来源样品相对于参考参考阈值水平的分类截止值显示出靶生物标志物特征表达细胞外囊泡水平升高时,将受试者分类为患有或易患肺癌(例如,在一些实施方案中为非小细胞肺癌例如像LUAD和/或LUSC)。在一些实施方案中,至少一个此类靶生物标志物是或包含选自以下的表面蛋白生物标志物:ABCA3、ACBD3、AGER、ALCAM、AP1M2、APH1A、APOO、ATP1B1、ATP6AP2、BCAP31、BSPRY、CDC42、CDH1、CDH3、CDKAL1、CELSR1、CIP2A、CISD2、CKAP4、CLPTM1L、CLSTN1、CPD、DPY19L1、DSG2、EGFR、EPCAM、FZD6、GALNT1、GALNT3、GALNT5、GALNT6、GGCX、GOLPH3L、GRHL2、HACD3、IER3IP1、IL1RAP、ITGA2、ITGB6、KPNA2、KRTCAP3、LAD1、LAMB3、LAMP3、LAMTOR2、LCLAT1、LSR、MAGT1、MARCKSL1、MET、MGAT1、MUC4、NCSTN、NECTIN4、NRAS、NUP210、PEX13、PIGN、PIGT、PLEC、PTPRF、QSOX1、RAB25、RCC2、RIT1、SCAMP3、SDC1、SEL1L3、SHROOM2、SLC35B2、SLC39A11、SOAT1、SPAST、SSR1、SSR4、SURF4、SYNGR2、TACSTD2、TMC01、TMED2、TMED3、TMEM132A、TMEM33、TMPRSS4、TOMM22、TOR1AIP2、TRAM1、TTC33、VAMP8、VMA21、VRK2、VWA1、XPR1、ADAM28、AXL、BSG、CD274、CD47、CLU、DKK1、FLT4、GM3、HGF、IGF1R、IL6、LAG3、Lewis Y/B抗原、LY6E、NOTCH2、NOTCH3、磷脂酰丝氨酸、TIGIT、TNFRSF10A、TNFRSF10B、TPBG、VEGFA、Tn抗原、Lewis Y/CD174、唾液酸化Lewis X(sLex)(也称为唾液酸化SSEA-1(SLX))、NeuGcGM3及其组合。在一些实施方案中,至少一个此类靶生物标志物是或包含选自以下的囊泡内蛋白生物标志物:ABRACL、ACP5、AIF1、ALDH1A1、ALG1L、AP1M2、APOBEC3C、ASF1B、AURKB、BAIAP2L1、BIRC5、C12orf45、C15orf48、C19orf33、C1S、C8orf4、CBLC、CCL19、CCNB2、CDC20、CDCA5、CDK1、CDKN2A、CDKN2B、CEP55、CHMP4C、CNN2、CPA3、CRABP2、CST1、CTSC、DPYSL3、EGLN3、EHF、ELF3、ELF4、ENAH、ESRP1、ETV4、EVPL、FAM129B、FAM60A、FAM83A、FAM83H、GNA15、GRHL2、HCK、HMGB3、HOXB7、ID1、

IMPA2、KIAA1522、KIF2C、KIFC1、KLF4、KLF5、KRT18、KRT19、KRT8、LGALS3BP、LSP1、MDFI、MIF、MYBL2、MYH14、MZB1、NCF2、NNMT、NUP210、NUSAP1、PALLD、PKP3、PLEK、PLEK2、POSTN、PTPN6、RIN2、RIPK4、RPS4Y1、RRM2、S100A11、S100A14、S100A16、SBK1、SH3BP4、SPI1、SPINT1、SULF1、SYK、SYTL1、TBC1D2、TEAD2、TEAD3、TFAP2C、TGFA、THBS2、TK1、TOP2A、TPD52、TPX2、TRIP13、UBE2C、YAP1、ZC3H11A、ZNF217及其组合。在一些实施方案中,至少一个此类靶生物标志物是或包含选自以下的囊泡内RNA(例如,mRNA):ABCA3、ABRACL、ACP5、AIF1、ALDH1A1、ALG1L、ANTXR1、AP1M2、APOBEC3C、AQP3、AREG、ARSL、ASF1B、ATP8B1、AURKB、BAIAP2L1、BCAM、BIK、BIRC5、C12orf45、C15orf48、C19orf33、C1S、C8orf4、CBLC、CCL19、CCL5、CCNB2、CD24、CD53、CD74、CDC20、CDC42EP1、CDCA5、CDCP1、CDH1、CDH3、CDK1、CDKN2A、CDKN2B、CELSR1、CEP55、CHMP4C、CLDN4、CLDN7、CNN2、CPA3、CRABP2、CST1、CTSC、CX3CL1、CXADR、CXCR4、CYBB、DMBT1、DPYSL3、DSG2、EFNA1、EGFR、EGLN3、EHD2、EHF、ELF3、ELF4、EMP1、EMP2、ENAH、EPCAM、EPHA2、EPHX1、EPHX3、ESRP1、ETV4、EVA1A、EVPL、F11R、F2R、F2RL1、F3、FAM129B、FAM60A、FAM83A、FAM83H、FAT1、FBLIM1、FCER1G、GALNT3、GNA15、GPC4、GRHL2、HCK、HMGB3、HOXB7、HS6ST2、ID1、IL2RG、IMPA2、ITGA2、ITGB4、ITGB6、JAG2、KCNS3、KDELRL3、KIAA1522、KIF2C、KIFC1、KITLG、KLF4、KLF5、KRT18、KRT19、KRT8、KRTCAP3、LAMB3、LAMP3、LAPTM5、LFNG、LGALS3BP、LRP11、LSP1、LSR、MAL2、MANEAL、MAOA、MARCO、MDFI、MET、MIF、MMP14、MPZL2、MYBL2、MYH14、MYOF、MZB1、NCF2、NKG7、NNMT、NOTCH3、NUP210、NUSAP1、OSMR、PALLD、PDPN、PDZK1IP1、PECAM1、PIGT、PKP3、PLEK、PLEK2、PLVAP、PMP22、PODXL2、POSTN、PPL、PROM2、PRSS8、PTGES、PTPN6、PTPRF、RAB25、RARRES1、RGS1、RHOF、RIN2、RIPK4、RPS4Y1、RRM2、S100A10、S100A11、S100A14、S100A16、SBK1、SCNN1A、SDC1、SERINC2、SEZ6L2、SH3BP4、SHISA2、SLC1A5、SLC40A1、SLC7A7、SPI1、SPINT1、SPINT2、ST14、STEAP1、SULF1、SYK、SYTL1、TACSTD2、TBC1D2、TEAD2、TEAD3、TFAP2C、TGFA、THBS2、TK1、TM4SF1、TMC4、TMEM30B、TMEM54、TMPRSS11E、TMPRSS4、TNFRSF18、TOP2A、TP53I11、TPD52、TPX2、TREM2、TRIP13、TSPAN1、TSPAN13、TSPAN6、TSPAN7、TUSC3、TYROBP、UBE2C、VAMP8、WLS、YAP1、ZC3H11A、ZNF217及其组合。

[0011] 在一些方面,提供了用于将受试者(例如,无症状受试者)分类为患有或易患某种类型的肺癌(例如,在特定实施方案中为肺腺癌(LUAD))的技术。在一些实施方案中,本公开提供了用于将受试者(例如,无症状受试者)分类为患有或易患某种类型的肺癌(例如,在特定实施方案中为LUAD)的方法或测定。在一些实施方案中,提供的方法或测定包括(a)在来自有需要的受试者的血液来源样品中检测表达某种类型的肺癌(例如,在特定实施方案中为LUAD)的靶生物标志物特征的细胞外囊泡,所述靶生物标志物特征包含:至少一个细胞外囊泡相关膜结合多肽和至少一个选自由表面蛋白生物标志物(如本文所述)、囊泡内蛋白生物标志物(如本文所述)和囊泡内RNA生物标志物(如本文所述)组成的组的靶生物标志物;(b)将指示血液来源样品中靶生物标志物特征表达细胞外囊泡水平的样品信息与包括参考阈值水平在内的参考信息进行比较;以及(c)当血液来源样品相对于参考参考阈值水平的分类截止值显示出靶生物标志物特征表达细胞外囊泡水平升高时,将受试者分类为患有或易患某种类型的肺癌(例如,在特定实施方案中为LUAD)。在一些实施方案中,至少一个对将受试者分类为患有或易患LUAD特别有用的此类靶生物标志物是或包含选自以下的表面蛋白生物标志物:ABCC3、ACSL5、ATP11A、CD55、CEACAM5、CEACAM6、CLIC6、FOLR1、GOLM1、LPCAT1、MSLN、MUC1、NT5E、PLA2G4A、PLCH1、SLC34A2、SMPDL3B、TESC、TMC5、ERBB3、KDR及其组

合。在一些实施方案中,至少一个对将受试者分类为患有或易患LUAD特别有用的此类靶生物标志物是或包含选自以下的囊泡内蛋白生物标志物:AGR2、AOC1、CTSE、FBP1、FOXA2、KRT7、NAPSA、PPP1R14D、S100P、SCGB3A2、SFTA2、SFTPA1、SFTPA2、SFTPB、SPINK1及其组合。在一些实施方案中,至少一个对将受试者分类为患有或易患LUAD特别有用的此类靶生物标志物是或包含选自以下的囊泡内RNA(例如,mRNA):ABCC3、AGR2、AOC1、B3GNT3、CEACAM5、CEACAM6、CLDN18、CLDN3、CLIC6、CTSE、FBP1、FOLR1、FOXA2、GJB1、GPRC5A、KRT7、MSLN、MUC1、MUC21、NAPSA、PIGR、PPP1R14D、ROS1、S100P、SCGB3A2、SFTA2、SFTPA1、SFTPA2、SFTPB、SLC34A2、SLC44A4、SLC6A14、SMIM22、SMPDL3B、SPINK1、TMC5、TMEM45B、TMPRSS2、TSPAN8及其组合。在一些实施方案中,某些前述靶生物标志物在区分LUAD和LUSC方面特别有用。

[0012] 在一些方面,提供了用于将受试者(例如,无症状受试者)分类为患有或易患某种类型的肺癌(例如,在特定实施方案中为肺鳞状细胞癌(LUSC))的技术。在一些实施方案中,本公开提供了用于将受试者(例如,无症状受试者)分类为患有或易患某种类型的肺癌(例如,在特定实施方案中为LUSC)的方法或测定。在一些实施方案中,提供的方法或测定包括(a)在来自有需要的受试者的血液来源样品中检测表达某种类型的肺癌(例如,在特定实施方案中为LUSC)的靶生物标志物特征的细胞外囊泡,所述靶生物标志物特征包含:至少一个细胞外囊泡相关膜结合多肽和至少一个选自由表面蛋白生物标志物(如本文所述)、囊泡内蛋白生物标志物(如本文所述)和囊泡内RNA生物标志物(如本文所述)组成的组的靶生物标志物;(b)将指示血液来源样品中靶生物标志物特征表达细胞外囊泡水平的样品信息与包括参考阈值水平在内的参考信息进行比较;以及(c)当血液来源样品相对于参考参考阈值水平的分类截止值显示出靶生物标志物特征表达细胞外囊泡水平升高时,将受试者分类为患有或易患某种类型的肺癌(例如,在特定实施方案中为LUSC)。在一些实施方案中,至少一个对将受试者分类为患有或易患LUSC特别有用的此类靶生物标志物是或包含选自以下的表面蛋白生物标志物:ABCC1、ATP11B、B4GALT4、CD109、CD9、CLCA2、CLDN1、CNTN1、CYP2S1、CYP4F11、CYP4F3、DSC2、DSC3、DSG3、EPHB3、FAT2、FBXO45、FERMT1、IGSF3、KLRG2、LAMC2、NECTIN1、PARL、PSMD2、PTDSS1、PTGFRN、RAB38、RAB6B、RAP2B、SLC2A1、TFRC、TMTC3、TRPV4、UGT1A6、UPK1B、XXYL1、TNFSF18及其组合。在一些实施方案中,至少一个对将受试者分类为患有或易患LUSC特别有用的此类靶生物标志物是或包含选自以下的囊泡内蛋白生物标志物:ADH7、AKR1C1、AKR1C2、AKR1C3、ALDH3A1、ALDH3B2、APOBEC3B、ARNTL2、CA9、CALML3、CAPNS2、CDC45、CDCA4、CENPW、CES1、CSTA、CYP2S1、EFS、FAM83D、FERMT1、FOX E1、FOX M1、GBP6、GPX2、GSTA1、IGF2BP2、IRF6、IVL、JUP、KRT13、KRT14、KRT15、KRT16、KRT17、KRT5、KRT6A、KRT6B、KRT6C、LGALS7B、MAGEA4、MAGEA6、MCM2、NRARP、OSGIN1、PITX1、PKP1、PPP1R14C、PRAME、RBP1、S100A2、SERPINB13、SERPINB3、SERPINB5、SNAI2、SOX2、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、TP63、TRIM29、ZNF750及其组合。在一些实施方案中,至少一个对将受试者分类为患有或易患LUSC特别有用的此类靶生物标志物是或包含选自以下的囊泡内RNA(例如,mRNA):ABCC1、ADAM23、ADH7、AKR1C1、AKR1C2、AKR1C3、ALDH3A1、ALDH3B2、APOBEC3B、ARNTL2、B3GNT5、CA12、CA9、CALML3、CAPNS2、CD109、CD9、CDC45、CDCA4、CENPW、CES1、CLCA2、CLDN1、COL17A1、CSTA、CYP2S1、CYP4F11、DAPL1、DSC2、DSC3、DSG3、DSP、EFS、EPHB3、FAM83D、FAT2、FERMT1、FGFR2、FGFR3、FOX E1、FOX M1、FX YD3、GBP6、GJA1、GJB2、GJB3、GJB5、GJB6、GPC1、GPC3、GPNMB、GPR87、GPX2、GSTA1、HAS3、IGF2BP2、IGSF9、IRF6、ITGA6、IVL、

JUP、KRT13、KRT14、KRT15、KRT16、KRT17、KRT5、KRT6A、KRT6B、KRT6C、LGALS7B、LRRC4、LYPD3、MAGEA4、MAGEA6、MCM2、NRARP、NTRK2、OSGIN1、PERP、PITX1、PKP1、PPP1R14C、PRAME、PRRG4、PTGFRN、PTPRZ1、RAB38、RAET1L、RBP1、RHCG、S100A2、SERPINB13、SERPINB3、SERPINB5、SLC2A1、SLC6A8、SLC7A8、SNAI2、SOX2、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、THBD、TMPRSS11D、TNS4、TP63、TRIM29、UPK1B、VANG2、ZNF750及其组合。在一些实施方案中，某些前述靶生物标志物在区分LUSC和LUAD方面特别有用。

[0013] 在一些实施方案中，可以对至少一个另外的靶生物标志物特征(包括，例如至少一个、至少两个、至少三个或更多个另外的靶生物标志物特征)进行本文所述的方法和测定。在一些此类实施方案中，分类截止值可以参考对应于每个另外的靶生物标志物特征的另外的参考阈值水平。

[0014] 在一些实施方案中，用于本文使用的和/或描述的肺癌的靶生物标志物特征的细胞外囊泡相关膜结合多肽可以是或包含肿瘤特异性生物标志物和/或组织特异性生物标志物(例如，肺组织特异性生物标志物)。在一些实施方案中，此类细胞外囊泡相关膜结合多肽生物标志物可以是或包含非特异性标志物，例如，其存在于一种或多种非靶肿瘤和/或一种或多种非靶组织中。在一些实施方案中，此类细胞外囊泡相关膜结合多肽生物标志物可以是或包含一个或多个本文所述的表面蛋白生物标志物。在一些实施方案中，此类细胞外囊泡相关膜结合多肽生物标志物可以是或包含一个或多个选自以下的多肽生物标志物：ALCAM、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CLDN3、CLDN4、DSG2、EGFR、EPCAM、FOLR1、GJB1、GJB2、IG1FR、LAMB3、MET、MSLN、MUC1、PIGT、PODXL2、ROS1、SDC1、SLC34A2、sTn抗原、Tn抗原、T抗原、SMPDL3B、ST14、TACSTD2、TMPRSS4、TNFRSF10B、TSPAN8及其组合。在一些实施方案中，此类细胞外囊泡相关膜结合多肽生物标志物可以是或包含SLC34A2、CEACAM5、CEACAM6和/或EPCAM。在一些实施方案中，此类细胞外囊泡相关膜结合多肽生物标志物可以是或包含ALCAM、CD55、CDH1、CDH3、CD274 (PD-L1)、CEACAM5、CEACAM6、DSG2、EGFR、EPCAM、FOLR1、IG1FR、MET、MSLN、MUC1、SLC34A2、sTn抗原、Tn抗原、T抗原、TACSTD2、TNFRSF10B或其组合。在一些实施方案中，此类细胞外囊泡相关膜结合多肽可以是或包含SLC34A2多肽。在一些实施方案中，此类细胞外囊泡相关膜结合多肽可以是或包含CEACAM5多肽。在一些实施方案中，此类细胞外囊泡相关膜结合多肽可以是或包含CEACAM6多肽。在一些实施方案中，此类细胞外囊泡相关膜结合多肽可以是或包含EPCAM多肽。

[0015] 在一些实施方案中，肺癌的靶生物标志物特征包含细胞外囊泡相关膜结合多肽(例如，本文所述的那些)和至少一个(包括例如，1个、2个、3个或更多个)另外的靶表面蛋白生物标志物，其在一些实施方案中可以是或包含ALCAM、ABCC3、ARSL、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、DMBT1、DSG2、EGFR、EPCAM、EPHX3、EVA1A、FOLR1、GJB1、GJB2、GPC4、HS6ST2、IG1FR、KDEL3、KRTCAP3、LAMB3、LFNG、LSR、MANEAL、MET、MSLN、MUC1、MUC21、PIGT、PODXL2、PRRG4、ROS1、SDC1、SERINC2、SEZ6L2、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TNFRSF10B、TSPAN1、TSPAN8或其组合。

[0016] 在一些实施方案中，用于本文使用的和/或描述的肺癌的靶生物标志物特征的细

胞外囊泡相关膜结合多肽可以是或包含肿瘤特异性生物标志物和/或组织特异性生物标志物(例如,肺组织特异性生物标志物)。在一些实施方案中,此类细胞外囊泡相关膜结合多肽生物标志物可以是或包含非特异性标志物,例如,其存在于一种或多种非靶肿瘤和/或一种或多种非靶组织中。在一些实施方案中,此类细胞外囊泡相关膜结合多肽生物标志物可以是或包含一个或多个本文所述的表面蛋白生物标志物。在一些实施方案中,此类细胞外囊泡相关膜结合多肽生物标志物可以是或包含一个或多个选自以下的多肽生物标志物: ABCA3、ABCC1、ABCC3、ACBD3、ACSL5、AGER、ALCAM、AP1M2、APH1A、APOO、ATP11A、ATP11B、ATP1B1、ATP6AP2、B4GALT4、BCAP31、BSPRY、CD109、CD55、CD9、CDC42、CDH1、CDH3、CDKAL1、CEACAM5、CEACAM6、CELSR1、CIP2A、CISD2、CKAP4、CLCA2、CLDN1、CLIC6、CLPTM1L、CLSTN1、CNTN1、CPD、CYP2S1、CYP4F11、CYP4F3、DPY19L1、DSC2、DSC3、DSG2、DSG3、EGFR、EPCAM、EPHB3、FAT2、FBX045、FERMT1、FOLR1、FZD6、GALNT1、GALNT3、GALNT5、GALNT6、GGCX、GOLM1、GOLPH3L、GRHL2、HACD3、IER3IP1、IGSF3、IL1RAP、ITGA2、ITGB6、KLRG2、KPNA2、KRTCAP3、LAD1、LAMB3、LAMC2、LAMP3、LAMTOR2、LCLAT1、LPCAT1、LSR、MAGT1、MARCKSL1、MET、MGAT1、MSLN、MUC1、MUC4、NCSTN、NECTIN1、NECTIN4、NRAS、NT5E、NUP210、PARL、PEX13、PIGN、PIGT、PLA2G4A、PLCH1、PLEC、PSMD2、PTDSS1、PTGFRN、PTPRF、QSOX1、RAB25、RAB38、RAB6B、RAP2B、RCC2、RIT1、SCAMP3、SDC1、SEL1L3、SHROOM2、SLC2A1、SLC34A2、SLC35B2、SLC39A11、SMPDL3B、SOAT1、SPAST、SSR1、SSR4、SURF4、SYNGR2、TACSTD2、TESC、TFRC、TMC5、TMC01、TMED2、TMED3、TMEM132A、TMEM33、TMPRSS4、TMTC3、TOMM22、TOR1AIP2、TRAM1、TRPV4、TTC33、UGT1A6、UPK1B、VAMP8、VMA21、VRK2、VWA1、XPR1、XXYLT1、ADAM28、AXL、BSG、CD274、CD47、CLU、DKK1、ERBB3、FLT4、GM3、HGF、IGF1R、IL6、KDR、LAG3、Lewis Y/B抗原、LY6E、NOTCH2、NOTCH3、磷脂酰丝氨酸、TIGIT、TNFRSF10A、TNFRSF10B、TNFSF18、TPBG、VEGFA、Tn抗原、Lewis Y/CD174、唾液酸化Lewis X(sLex)(也称为唾液酸化SSEA-1(SLX))、NeuGcGM3及其组合。

[0017] 在一些实施方案中,肺癌的靶生物标志物特征包含细胞外囊泡相关膜结合多肽(例如,本文所述的那些)和至少一个(包括例如,1个、2个、3个或更多个)另外的靶表面蛋白生物标志物,其在一些实施方案中可以是或包含ABCA3、ABCC1、ABCC3、ACBD3、ACSL5、AGER、ALCAM、AP1M2、APH1A、APOO、ATP11A、ATP11B、ATP1B1、ATP6AP2、B4GALT4、BCAP31、BSPRY、CD109、CD55、CD9、CDC42、CDH1、CDH3、CDKAL1、CEACAM5、CEACAM6、CELSR1、CIP2A、CISD2、CKAP4、CLCA2、CLDN1、CLIC6、CLPTM1L、CLSTN1、CNTN1、CPD、CYP2S1、CYP4F11、CYP4F3、DPY19L1、DSC2、DSC3、DSG2、DSG3、EGFR、EPCAM、EPHB3、FAT2、FBX045、FERMT1、FOLR1、FZD6、GALNT1、GALNT3、GALNT5、GALNT6、GGCX、GOLM1、GOLPH3L、GRHL2、HACD3、IER3IP1、IGSF3、IL1RAP、ITGA2、ITGB6、KLRG2、KPNA2、KRTCAP3、LAD1、LAMB3、LAMC2、LAMP3、LAMTOR2、LCLAT1、LPCAT1、LSR、MAGT1、MARCKSL1、MET、MGAT1、MSLN、MUC1、MUC4、NCSTN、NECTIN1、NECTIN4、NRAS、NT5E、NUP210、PARL、PEX13、PIGN、PIGT、PLA2G4A、PLCH1、PLEC、PSMD2、PTDSS1、PTGFRN、PTPRF、QSOX1、RAB25、RAB38、RAB6B、RAP2B、RCC2、RIT1、SCAMP3、SDC1、SEL1L3、SHROOM2、SLC2A1、SLC34A2、SLC35B2、SLC39A11、SMPDL3B、SOAT1、SPAST、SSR1、SSR4、SURF4、SYNGR2、TACSTD2、TESC、TFRC、TMC5、TMC01、TMED2、TMED3、TMEM132A、TMEM33、TMPRSS4、TMTC3、TOMM22、TOR1AIP2、TRAM1、TRPV4、TTC33、UGT1A6、UPK1B、VAMP8、VMA21、VRK2、VWA1、XPR1、XXYLT1、ADAM28、AXL、BSG、CD274、CD47、CLU、DKK1、ERBB3、FLT4、GM3、HGF、IGF1R、IL6、KDR、LAG3、Lewis Y/B抗原、LY6E、NOTCH2、NOTCH3、磷脂酰丝氨酸、TIGIT、TNFRSF10A、TNFRSF10B、

TNFSF18、TPBG、VEGFA、Tn抗原、Lewis Y/CD174、唾液酸化Lewis X(sLex) (也称为唾液酸化SSEA-1 (SLX))、NeuGcGM3及其组合。

[0018] 在一些实施方案中,肺癌的靶生物标志物特征包含细胞外囊泡相关膜结合多肽(例如,本文所述的那些)和至少一个靶囊泡外RNA(例如,mRNA)生物标志物,其在一些实施方案中可以是或包含ABCC3、AOC1、ARSL、B3GNT3、C12orf45、CDCP1、CDH1、CDH3、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CRABP2、CST1、DMBT1、DSG2、EPCAM、EPHX3、ETV4、EVA1A、FAM83A、FOLR1、FOXA2、GJB1、GJB2、GPC4、HMGB3、HS6ST2、KDELRL3、KRTCAP3、LAMB3、LFNG、LGALS3BP、LSR、MANEAL、MIF、MSLN、MUC1、MUC21、NAPSA、PIGT、PODXL2、PPP1R14D、PRRG4、ROS1、S100A14、SBK1、SCGB3A2、SDC1、SERINC2、SEZ6L2、SFTA2、SFTPA1、SFTPA2、SFTPB、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、SPINK1、ST14、TGFA、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、ZC3H11A或其组合。

[0019] 在一些实施方案中,肺癌的靶生物标志物特征包含细胞外囊泡相关膜结合多肽(例如,本文所述的那些)和至少一个靶囊泡外RNA(例如,mRNA)生物标志物,其在一些实施方案中可以是或包含ABCA3、ABCC1、ABRACL、ACP5、ADAM23、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、ANTXR1、AP1M2、APOBEC3B、APOBEC3C、AQP3、AREG、ARNTL2、ASF1B、ATP8B1、AURKB、B3GNT5、BAIAP2L1、BCAM、BIK、BIRC5、C15orf48、C19orf33、C1S、C8orf4、CA12、CA9、CALML3、CAPNS2、CBLC、CCL19、CCL5、CCNB2、CD109、CD24、CD53、CD74、CD9、CDC20、CDC42EP1、CDC45、CDCA4、CDCA5、CDCP1、CDH1、CDH3、CDK1、CDKN2A、CDKN2B、CEACAM5、CEACAM6、CELSR1、CENPW、CEP55、CES1、CHMP4C、CLCA2、CLDN1、CLDN4、CLDN7、CNN2、COL17A1、CPA3、CRABP2、CSTA、CTSC、CTSE、CX3CL1、CXADR、CXCR4、CYBB、CYP2S1、CYP4F11、DAPL1、DPYSL3、DSC2、DSC3、DSG2、DSG3、DSP、EFNA1、EFS、EGFR、EGLN3、EHD2、EHF、ELF3、ELF4、EMP1、EMP2、ENAH、EPCAM、EPHA2、EPHB3、EPHX1、ESRP1、EVPL、F11R、F2R、F2RL1、F3、FAM129B、FAM60A、FAM83D、FAM83H、FAT1、FAT2、FBLIM1、FBP1、FCER1G、FERMT1、FGFR2、FGFR3、FOX E1、FOX M1、FX YD3、GALNT3、GBP6、GJA1、GJB2、GJB3、GJB5、GJB6、GNA15、GPC1、GPC3、GPNMB、GPR87、GPRC5A、GPX2、GRHL2、GSTA1、HAS3、HCK、HOXB7、ID1、IGF2BP2、IGSF9、IL2RG、IMPA2、IRF6、ITGA2、ITGA6、ITGB4、ITGB6、IVL、JAG2、JUP、KCNS3、KIAA1522、KIF2C、KIFC1、KITLG、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、KRTCAP3、LAMP3、LAPTM5、LGALS7B、LRP11、LRRC4、LSP1、LSR、LYPD3、MAGEA4、MAGEA6、MAL2、MAOA、MARCO、MCM2、MDFI、MET、MMP14、MPZL2、MUC1、MYBL2、MYH14、MYOF、MZB1、NCF2、NKG7、NNMT、NOTCH3、NRARP、NTRK2、NUP210、NUSAP1、OSGIN1、OSMR、PALLD、PDPN、PDZK1IP1、PECAM1、PERP、PIGR、PIGT、PITX1、PKP1、PKP3、PLEK、PLEK2、PLVAP、PMP22、POSTN、PPL、PPP1R14C、PRAME、PROM2、PRRG4、PRSS8、PTGES、PTGFRN、PTPN6、PTPRF、PTPRZ1、RAB25、RAB38、RAET1L、RARRES1、RBP1、RGS1、RHCG、RHOV、RIN2、RIPK4、RPS4Y1、RRM2、S100A10、S100A11、S100A14、S100A16、S100A2、S100P、SCNN1A、SDC1、SERINC2、SERPINB13、SERPINB3、SERPINB5、SEZ6L2、SH3BP4、SHISA2、SLC1A5、SLC2A1、SLC34A2、SLC40A1、SLC6A8、SLC7A8、SNAI2、SOX2、SPI1、SPINT1、SPINT2、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、ST14、STEAP1、SULF1、SYK、SYTL1、TACSTD2、TBC1D2、TEAD2、TEAD3、TFAP2C、THBD、THBS2、TK1、TM4SF1、TMC4、TMEM30B、TMEM54、TMPRSS11D、TMPRSS11E、TMPRSS4、TNFRSF18、TNS4、TOP2A、TP53I11、TP63、TPD52、TPX2、TREM2、TRIM29、TRIP13、TSPAN1、TSPAN13、TSPAN6、

TSPAN7、TUSC3、TYROBP、UBE2C、UPK1B、VAMP8、VANGL2、WLS、YAP1、ZC3H11A、ZNF217、ZNF750或其组合。

[0020] 在一些实施方案中,肺癌的靶生物标志物特征包含细胞外囊泡相关膜结合多肽(例如,本文所述的那些)和至少一个另外的靶囊泡内蛋白生物标志物,其在一些实施方案中可以是或包含AOC1、C12orf45、CRABP2、CST1、ETV4、FAM83A、FOXA2、HMGB3、LGALS3BP、MIF、NAPSA、PPP1R14D、S100A14、SBK1、SCGB3A2、SFTA2、SFTPA1、SFTPA2、SFTPB、SPINK1、TGFA、ZC3H11A或其组合。在一些实施方案中,靶生物标志物特征包含至少一个细胞外囊泡相关膜结合多肽,其是或包含SCL34A2多肽和/或CEACAM5多肽;以及至少一个靶生物标志物SLC34A2、CEACAM5、CEACAM6和/或EPCAM。

[0021] 在一些实施方案中,肺癌的靶生物标志物特征包含细胞外囊泡相关膜结合多肽(例如,本文所述的那些)和至少一个另外的靶囊泡内蛋白生物标志物,其在一些实施方案中可以是或包含ABRACL、ACP5、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、AP1M2、APOBEC3B、APOBEC3C、ARNTL2、ASF1B、AURKB、BAIAP2L1、BIRC5、C15orf48、C19orf33、C1S、C8orf4、CA9、CALML3、CAPNS2、CBLC、CCL19、CCNB2、CDC20、CDC45、CDCA4、CDCA5、CDK1、CDKN2A、CDKN2B、CENPW、CEP55、CES1、CHMP4C、CNN2、CPA3、CRABP2、CSTA、CTSC、CTSE、CYP2S1、DPYSL3、EFS、EGLN3、EHF、ELF3、ELF4、ENAH、ESRP1、EVPL、FAM129B、FAM60A、FAM83D、FAM83H、FBP1、FERMT1、FOX E1、FOX M1、GBP6、GNA15、GPX2、GRHL2、GSTA1、HCK、HOXB7、ID1、IGF2BP2、IMPA2、IRF6、IVL、JUP、KIAA1522、KIF2C、KIFC1、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、LGALS7B、LSP1、MAGEA4、MAGEA6、MCM2、MDFI、MYBL2、MYH14、MZB1、NCF2、NNMT、NRARP、NUP210、NUSAP1、OSGIN1、PALLD、PITX1、PKP1、PKP3、PLEK、PLEK2、POSTN、PPP1R14C、PRAME、PTPN6、RBP1、RIN2、RIPK4、RPS4Y1、RRM2、S100A11、S100A14、S100A16、S100A2、S100P、SERPINB13、SERPINB3、SERPINB5、SH3BP4、SNAI2、SOX2、SPI1、SPINT1、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、SULF1、SYK、SYTL1、TBC1D2、TEAD2、TEAD3、TFAP2C、THBS2、TK1、TOP2A、TP63、TPD52、TPX2、TRIM29、TRIP13、UBE2C、YAP1、ZC3H11A、ZNF217、ZNF750或其组合。

[0022] 在一些实施方案中,用于本文所述的提供的方法或测定中的参考阈值水平通过在来自非肺癌受试者群体的可比样品中观察到的靶生物标志物特征表达细胞外囊泡的水平来确定。

[0023] 在一些实施方案中,可以使用基于抗体的剂检测包含在靶生物标志物特征中的细胞外囊泡相关膜结合多肽。在一些实施方案中,可以使用包含基于抗体的剂的捕获测定检测此类细胞外囊泡相关膜结合多肽。例如,在一些实施方案中,用于检测细胞外囊泡中细胞外囊泡相关膜结合多肽的存在的捕获测定可以涉及使包含细胞外囊泡的血液来源样品与针对此类细胞外囊泡相关膜结合多肽的捕获剂接触。在一些实施方案中,此类捕获剂包含针对细胞外囊泡相关膜结合多肽(例如,本文所述的那些)的结合部分,所述结合部分可以任选地缀合至固体基底。在不受限制的情况下,用于细胞外囊泡相关膜结合多肽的示例性捕获剂可以是或包含固体基底(例如,磁珠)和针对细胞外囊泡相关膜结合多肽的结合部分(例如,抗体剂)。

[0024] 在一些实施方案中,可以使用本领域已知的适当方法检测包括在靶生物标志物特

征中的靶生物标志物,所述适当方法可以随着待检测分析物的类型(例如,表面蛋白、囊泡内蛋白、囊泡内RNA(例如,mRNA))而变化。例如,阅读本公开的本领域技术人员将理解,在一些实施方案中可以使用基于抗体的剂检测表面蛋白生物标志物和/或囊泡内蛋白生物标志物,而在一些实施方案中,可以使用基于核酸的剂(例如,使用定量逆转录PCR)检测囊泡内RNA(例如,mRNA)生物标志物。

[0025] 例如,在靶生物标志物是或包含表面蛋白生物标志物和/或囊泡内蛋白标志物的一些实施方案中,可例如在捕获表达细胞外囊泡相关膜结合多肽的细胞外囊泡(例如,如本文使用和/或描述的那些)的捕获测定(例如,如本文所述的那些)之后涉及邻近连接测定来检测此类靶生物标志物。在一些实施方案中,此类邻近连接测定包括将包含细胞外囊泡的血液来源样品与一组检测探针接触,每个检测探针针对一个靶生物标志物,所述组包括至少两个不同的检测探针,从而生成包含所述细胞外囊泡和所述检测探针组的组合,其中所述两个检测探针各自包含:(i)针对表面蛋白生物标志物和/或囊泡内蛋白生物标志物的结合部分;和(ii)偶联至所述结合部分的寡核苷酸结构域,所述寡核苷酸结构域包含双链部分和从所述寡核苷酸结构域的一端延伸的单链悬突部分。所述检测探针的此类单链悬突部分的特征在于它们在所述检测探针结合至相同的细胞外囊泡时可以彼此杂交。然后将包含所述细胞外囊泡和所述检测探针组的此类组合保持在允许所述检测探针组与其在所述细胞外囊泡上的相应靶标结合的条件下,使得所述检测探针可以结合至相同的细胞外囊泡以形成双链复合物。此类双链复合物可以通过使所述双链复合物与核酸连接酶接触产生连接模板;并检测所述连接模板来检测。此类连接模板的存在指示对肺癌的靶生物标志物特征呈阳性的细胞外囊泡的存在。虽然此类邻近连接测定可以执行得更好,例如具有比其他现有邻近连接测定更高的特异性和/或灵敏度,但阅读本公开的本领域技术人员将理解,可以替代地使用本领域已知的其他形式的邻近连接测定。

[0026] 在靶生物标志物是或包含囊泡内RNA(例如,mRNA)标志物的一些实施方案中,可以涉及核酸检测测定来检测此类靶生物标志物。在一些实施方案中,示例性核酸检测测定可以是或包括逆转录PCR。

[0027] 在靶生物标志物是或包含囊泡内生物标志物(例如,囊泡内蛋白生物标志物和/或囊泡内RNA(例如,mRNA)生物标志物)的一些实施方案中,可在检测测定(例如,如本文所述的邻近连接测定)之前涉及使用于随后检测的一个或多个囊泡内生物标志物暴露的样品处理(例如,固定和/或透化)来检测此类靶生物标志物。

[0028] 本公开尤其认识到,基于批量样品(例如,细胞外囊泡的批量样品)而不是以单个细胞外囊泡的分辨率对单个肺癌相关血清蛋白或多个肺癌相关生物标志物的检测通常在确定从其获得样品的受试者是否可能患有或易患肺癌方面不能提供足够的特异性和/或灵敏度。本公开尤其提供了解决此类问题的技术(包括系统、组合物和/或方法),包括例如通过具体要求用于检测的个体细胞外囊泡的特征在于存在包含至少一个或多个细胞外囊泡相关膜结合多肽和至少一个或多个靶生物标志物的组合的靶生物标志物特征。在特定实施方案中,本公开教导了要求此类个体细胞外囊泡的特征在于存在(例如,通过表达)肺癌的此类靶生物标志物特征的技术,而不包含靶生物标志物特征的细胞外囊泡不产生可检测信号(例如,高于参考水平,例如至少10%或更多,其中在一些实施方案中,参考水平可以是在阴性对样品(诸如其中不存在包含此类靶生物标志物特征的个体细胞外囊泡的样品)中

观察到的水平)。

[0029] 因此,在一些实施方案中,本文提供的技术可以用于检测受试者和/或整个受试者群体中肺癌的发病或复发。在一些实施方案中,可以选择靶生物标志物特征用于检测肺癌。在一些实施方案中,可以选择靶生物标志物特征用于检测特定类别的肺癌,包括但不限于例如肺腺癌、小细胞肺癌、鳞状细胞和移行细胞肺癌、大细胞肺癌、非小细胞癌、其他特定的癌症、肉瘤和其他本领域已知的特定类型的肺癌(参见,例如SEER癌症统计回顾1975-2017)。在一些实施方案中,本文提供的技术可以定期(例如,每年)用于筛查人受试者或整个人受试者群体的早期肺癌或肺癌复发。

[0030] 在一些实施方案中,适用本文提供的用于检测肺癌的发病或复发的技术的受试者可以是对无症状人受试者和/或整个无症状群体。此类无症状受试者可以是对具有肺癌家族病史、先前接受过肺癌治疗、癌症治疗后存在肺癌复发风险、癌症治疗后处于缓解期,和/或先前或定期通过胸部X射线、痰液分析、低剂量CT和/或至少一种肺癌血清生物标志物(例如但不限于CEA、CYFRA 21-1、NSE、ProGRP和/或SCCA血清蛋白)的存在筛查肺癌存在的受试者。在一些实施方案中,此类无症状受试者可以由例如胸部X射线、痰液分析、低剂量CT分析或血清CEA、CYFRA 21-1、NSE、ProGRP和/或SCCA水平确定为具有正常医学诊断结果的受试者。在一些实施方案中,此类无症状受试者可以由例如胸部X射线、痰液分析、低剂量CT分析和/或血清CEA、CYFRA 21-1、NSE、ProGRP和/或SCCA水平在与通常在非肺癌受试者和/或正常健康受试者中观察到的结果相比时确定为具有异常医学诊断结果的受试者。替代地,在一些实施方案中,无症状受试者可以是先前未经肺癌筛查的受试者、未诊断为肺癌的受试者和/或先前未接受肺癌治疗的受试者。

[0031] 在一些实施方案中,可以基于一种或多种特征(诸如年龄、种族、遗传史、病史、个人史和/或病史(例如,吸烟、酒精、药物、致癌剂、饮食、肥胖、糖尿病、体力活动、光暴露、放射暴露、暴露于感染性因子诸如病毒和/或职业危害))来选择受试者或受试者群体。

[0032] 在一些实施方案中,本文提供的技术可以用于选择针对患有或易患肺癌的受试者的疗法。在一些实施方案中,可以根据基于本文提供的技术的发现来选择肺癌疗法和/或辅助疗法。

[0033] 在一些实施方案中,本文提供的技术可以用于监测和/或评估施用于受试者(例如,肺癌受试者)的疗法的功效。

[0034] 在一些实施方案中,本公开提供了用于管理患者护理的技术,以用于例如一个或多个个体受试者和/或整个受试者群体。仅举几个实例,在一些实施方案中,本公开提供了可以用于筛查(例如,临时或偶然动机的筛查和/或非临时或偶然动机的筛查,例如周期性筛查,诸如年度、半年、两年或以某种其他频率)的技术。例如,在一些实施方案中,提供的用于临时动机筛查的技术可以用于筛查超过某个年龄(例如,超过50、55、60、65、70岁或以上)的一个或多个个体受试者或整个受试者(例如,无症状受试者)群体。例如,在一些实施方案中,提供的用于临时动机筛查的技术可以用于筛查具有大于一定数量(例如,5包年、10包年、15包年、20包年、25包年、30包年和/或大于35包年;1包年等于每天抽1包烟持续一年,而每天抽2包烟持续一年等于2包年,或者每天抽1/2包烟持续2年等于1包年等)的香烟包-年史的一个或多个个体受试者或整个受试者(例如,无症状受试者)群体。在一些实施方案中,提供的用于偶然动机筛查的技术可以用于筛查可能经历了促使筛查如本文所述的肺癌的

事故或事件的个体受试者。例如,在一些实施方案中,与确定癌症或其易患性的一个或多个指标有关的偶然动机可以是或包括例如基于他们的家族病史的事故(例如,近亲诸如血缘亲属先前被诊断为肺癌);鉴定与肺癌相关的一个或多个风险因素(例如,生活史风险因素,包括但不限于例如吸烟、酒精、饮食、肥胖、职业危害等)和/或遗传测试(例如,基因组测序)和/或影像诊断测试(例如,X射线、超声、计算机断层扫描(CT)、低剂量CT和/或磁共振成像(MRI)扫描)的先前偶然发现;发展具有肺癌特征的一种或多种体征或症状(例如,可能指示肺癌的异常成像结果和/或症状等)。

[0035] 在一些实施方案中,提供的用于管理患者护理的技术可通知治疗和/或支付(例如,治疗的报销)决定和/或行动。例如,在一些实施方案中,提供的技术可以提供对个体受试者是否具有肺癌的发病或复发的一个或多个指标的确定,从而通知医师和/或患者何时根据此类发现开始治疗。另外或替代地,在一些实施方案中,提供的技术可以例如基于特定反应性生物标志物(例如,肺癌反应性生物标志物)的发现来通知医师和/或患者治疗选择。在一些实施方案中,提供的技术可以例如基于与肺癌相关的一个或多个分子靶标水平的变化的发现来提供对个体受试者是否对当前治疗有反应的确定,从而通知医师和/或患者此类治疗的功效和/或根据此类发现保持或改变治疗的决定。

[0036] 在一些实施方案中,提供的技术可通知关于健康保险提供者是否报销(或不报销)例如以下的决定:(1) 筛查本身(例如,报销仅可用于周期性/定期筛查或仅可用于临时和/或偶然动机的筛查);和/或(2) 根据提供的技术的发现开始、保持和/或改变治疗。例如,在一些实施方案中,本公开提供了与以下有关的方法:(a) 接收如本文所述的筛查结果,并且还接收对筛查和/或特定治疗方案的报销请求;(b) 如果根据适当的时间表或对相关事故的反应对受试者进行筛查,则批准报销筛查;和/或如果治疗方案根据接收的筛查结果代表适当治疗,则批准报销治疗方案;以及任选地(c) 实施报销或提供报销被拒绝的通知。在一些实施方案中,如果接收的筛查结果检测到代表相关治疗方案的批准的生物标志物(例如,如可以在处方信息标签中和/或通过批准的伴随诊断指出的生物标志物)的生物标志物,则根据接收的筛查结果,治疗方案是适当的。替代地或另外,本公开设想允许或促进如本文所述的筛查结果和/或报销决定的报告和/或处理的报告系统(例如,通过一个或多个适当的电子装置和/或一个或多个通信系统实施)。

[0037] 本文提供的一些方面涉及用于提供的技术的系统和试剂盒。在一些实施方案中,系统或试剂盒可以包括针对肺癌的肿瘤生物标志物特征的检测剂(例如,本文所述的那些)。在一些实施方案中,此类系统或试剂盒可以包括针对存在于与肺癌相关的细胞外囊泡(例如,本文使用和/或描述的那些)中的细胞外囊泡相关膜结合多肽的捕获剂;和(b) 针对肺癌的靶生物标志物特征的一个或多个靶生物标志物的至少一种或多种检测剂,所述靶生物标志物可以是或包含另外的一个或多个表面蛋白生物标志物(例如,如本文使用和/或描述的那些)、一个或多个囊泡内蛋白质生物标志物(例如,如本文使用和/或描述的那些),和/或一个或多个囊泡内RNA(例如,mRNA)生物标志物(例如,如本文使用和/或描述的那些)。

[0038] 在一些实施方案中,包括在系统和/或试剂盒中的捕获剂可以包含针对细胞外囊泡相关膜结合多肽(例如,本文所述的那些)的结合部分。在一些实施方案中,此类结合部分可以缀合至固体基底,其在一些实施方案中可以是或包含固体基底。在一些实施方案中,此

类固体基底可以是或包含磁珠。在一些实施方案中,包括在提供的系统和/或试剂盒中的示例性捕获剂可以是或包含固体基底(例如,磁珠)和针对与其缀合的细胞外囊泡相关膜结合多肽的抗体剂。

[0039] 在靶生物标志物包括表面蛋白生物标志物和/或囊泡内蛋白生物标志物的一些实施方案中,系统和/或试剂盒可以包括用于进行邻近连接测定的检测剂(例如,如本文所述的那些)。在一些实施方案中,用于进行邻近连接测定的此类检测剂可以包含一组检测探针,每个检测探针针对靶生物标志物特征的靶生物标志物,所述组包含至少两个检测探针,其中所述两个检测探针各自包含:(i)针对靶生物标志物的多肽结合部分;和(ii)偶联至所述结合部分的寡核苷酸结构域,所述寡核苷酸结构域包含双链部分和从所述寡核苷酸结构域的一端延伸的单链悬突部分,其中所述检测探针的单链悬突部分的特征在于它们在所述检测探针结合至相同的细胞外囊泡时可以彼此杂交。

[0040] 在一些实施方案中,提供的系统和/或试剂盒可以包括多个(例如,2个、3个、4个、5个或更多个)检测探针组,其中每组包含两个或更多个(例如,3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个或更多个)检测探针。在一些实施方案中,至少一组检测探针可以针对肺癌的检测。例如,在一些实施方案中,提供的系统和/或试剂盒可以包括用于检测肺癌的至少一组检测探针和用于检测不同癌症(例如,胰腺癌)的至少一组检测探针。在一些实施方案中,两个或更多个的检测探针可以针对不同类别的肺癌,包括但不限于例如肺腺癌肺癌、小细胞肺癌、鳞状细胞和移行细胞肺癌、大细胞肺癌、非小细胞肺癌、其他特定的肺癌、肉瘤型肺癌和其他本领域已知的特定类型的肺癌(参见,例如SEER癌症统计回顾1975-2017)。在一些实施方案中,两组或更多组可以针对不同类别的肺癌的检测。在一些实施方案中,两组或更多组可以针对相同类别的肺癌的检测。在一些实施方案中,两组或更多组可以针对不同分期的肺癌的检测。在一些实施方案中,两组或更多组可以针对相同分期的肺癌的检测。

[0041] 在一些实施方案中,提供的试剂盒中的检测探针可以作为容器中的单一混合物提供。在一些实施方案中,多组检测探针可以作为独立容器中的单独混合物提供。在一些实施方案中,每个检测探针在独立容器中单独地提供。

[0042] 在靶生物标志物包含囊泡内RNA(例如,mRNA)生物标志物的一些实施方案中,此类系统和/或试剂盒可以包括用于进行核酸检测测定的检测剂。在一些实施方案中,此类系统和/或试剂盒可以包括用于进行定量逆转录PCR的检测剂,例如,所述定量逆转录PCR可以包括针对一个或多个囊泡内RNA(例如,mRNA)靶标的引物。

[0043] 在一些实施方案中,提供的系统和/或试剂盒可以包括至少一种化学试剂,以例如处理样品和/或其中的细胞外囊泡。在一些实施方案中,提供的系统和/或试剂盒可以包括至少一种化学试剂以处理样品中的细胞外囊泡,所述化学试剂包括但不限于例如固定剂、透化剂和/或封闭剂。在一些实施方案中,提供的系统和/或试剂盒可以包括核酸连接酶和/或核酸聚合酶。在一些实施方案中,提供的系统和/或试剂盒可以包括一个或多个引物和/或探针。在一些实施方案中,提供的系统和/或试剂盒可以包括一对或多对引物,例如用于PCR,例如定量PCR(qPCR)反应。在一些实施方案中,提供的系统和/或试剂盒可以包括一个或多个探针,例如像在一些实施方案中可以被设计为增加qPCR的特异性的水解探针(例如,TaqMan探针)。在一些实施方案中,提供的系统和/或试剂盒可以包括一个或多个多路复用

探针,例如当采用同时或并行qPCR反应时可能是有用的(例如,以促进或改进读出)。

[0044] 在一些实施方案中,提供的系统和/或试剂盒可以用于筛查(例如,定期筛查)和/或个体(例如,无症状或有症状受试者)的其他评估以检测(例如,早期检测)肺癌。在一些实施方案中,提供的系统和/或试剂盒可以用于筛查和/或易患肺癌的个体(例如,具有已知遗传、环境或经验风险等的个体)的其他评估。在一些实施方案中,提供的系统和/或试剂盒可以用于监测先前接受过治疗的受试者的肺癌的复发。在一些实施方案中,提供的系统和/或试剂盒可以结合用于患有肺癌的受试者的疗法用作伴随诊断。在一些实施方案中,提供的系统和/或试剂盒可以用于监测或评估施用于患有肺癌的受试者的疗法的功效。在一些实施方案中,提供的系统和/或试剂盒可以用于为患有肺癌的受试者选择疗法。在一些实施方案中,提供的系统和/或试剂盒可以用于为具有与肺癌相关的一种或多种症状(例如,非特异性症状)的受试者做出治疗决定和/或选择疗法。

[0045] 通过进行本文所述的方法和/或使用本文所述的系统和/或试剂盒形成的复合物也在本公开的范围。例如,在一些实施方案中,复合物包含:(a)靶生物标志物特征表达细胞外囊泡,其中至少两个囊泡包含至少一个细胞外囊泡相关膜结合多肽和至少一个选自以下组成的组的靶生物标志物:表面蛋白生物标志物、囊泡内蛋白生物标志物和囊泡内RNA生物标志物,其中所述表面蛋白生物标志物选自ABCC3、ALCAM、ARSL、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274(PD-L1)、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、DMBT1、DSG2、EGFR、EPCAM、EPHX3、EVA1A、FOLR1、GJB1、GJB2、GPC4、HS6ST2、IG1FR、KDEL3、KRTCAP3、LAMB3、LFNG、LSR、MANEAL、MET、MSLN、MUC1、MUC21、PIGT、PODXL2、PRRG4、ROS1、SDC1、SERINC2、SEZ6L2、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TNFRSF10B、TSPAN1、TSPAN8及其组合;囊泡内蛋白生物标志物选自:AOC1、C12orf45、CRABP2、CST1、ETV4、FAM83A、FOXA2、HMGB3、LGALS3BP、MIF、NAPSA、PPP1R14D、S100A14、SBK1、SCGB3A2、SFTA2、SFTPA1、SFTPA2、SFTPB、SPINK1、TGFA、ZC3H11A及其组合;并且所述囊泡内RNA(例如,mRNA)生物标志物选自ABCC3、AOC1、ARSL、B3GNT3、C12orf45、CDCP1、CDH1、CDH3、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CRABP2、CST1、DMBT1、DSG2、EPCAM、EPHX3、ETV4、EVA1A、FAM83A、FOLR1、FOXA2、GJB1、GJB2、GPC4、HMGB3、HS6ST2、KDEL3、KRTCAP3、LAMB3、LFNG、LGALS3BP、LSR、MANEAL、MIF、MSLN、MUC1、MUC21、NAPSA、PIGT、PODXL2、PPP1R14D、PRRG4、ROS1、S100A14、SBK1、SCGB3A2、SDC1、SERINC2、SEZ6L2、SFTA2、SFTPA1、SFTPA2、SFTPB、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、SPINK1、ST14、TGFA、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、ZC3H11A及其组合,其中所述细胞外囊泡固定在包含针对此类细胞外囊泡相关膜结合多肽的结合部分的固体基底上。此类复合物还包含至少两个针对存在于细胞外囊泡中的靶生物标志物的至少一个靶生物标志物的检测探针,其中每个检测探针结合至此类靶生物标志物,并且每个检测探针都包含:(i)针对所述靶生物标志物的结合部分;和(ii)偶联至所述结合部分的寡核苷酸结构域,所述寡核苷酸结构域包含双链部分和从所述寡核苷酸结构域的一端延伸的单链悬突部分,其中所述检测探针的单链悬突部分彼此杂交。

[0046] 在一些实施方案中,存在于形成复合物的细胞外囊泡中的细胞外囊泡相关膜结合多肽生物标志物包含本文所述的一个或多个表面蛋白生物标志物。在一些实施方案中,此

类细胞外囊泡相关膜结合多肽可以是或包含一个或多个选自以下的多肽:ALCAM、B3GNT3、CDCP1、CDHA1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CLDN3、CLDN4、DSG2、EGFR、EPCAM、FOLR1、GJB1、GJB2、IG1FR、LAMB3、MET、MSLN、MUC1、PIGT、PODXL2、ROS1、SDC1、SLC34A2、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMPRSS4、TSPAN8、TNFRSF10B及其组合。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含SLC34A2多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含CEACAM5多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含CEACAM6多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含EPCAM多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含ALCAM多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含CD55多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含CDH1多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含CDH3多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含CD274 (PD-L1) 多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含DSG2多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含EGFR多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含FOLR1多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含IG1FR多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含MET多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含MSLN多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含MUC1多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含sTn抗原糖基化多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含Tn抗原糖基化多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含T抗原糖基化多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含TACSTD2多肽。在一些实施方案中,此类细胞外囊泡相关膜结合多肽可以是或包含TNFRSF10B多肽。

[0047] 下文和权利要求书中更详细地描述本公开所涵盖的这些和其他方面。

附图说明

[0048] 图1是说明剖析个体细胞外囊泡 (EV) 的示例性工作流程的示意图。所述图示出了使用尺寸排阻色谱法 (SEC) 从血浆中纯化EV和免疫亲和捕获显示出特定膜结合蛋白标志物的EV (图A); 使用根据本文所述的一些实施方案的靶实体检测测定检测捕获的EV上的共定位靶标志物 (例如, 囊泡内蛋白质或表面蛋白质) (图B)。

[0049] 图2是说明根据本文所述的一些实施方案的靶实体检测测定的示意图。在一些实施方案中, 靶实体检测测定使用检测探针的组合, 所述组合对癌症的检测具有特异性。在一些实施方案中, 将包括针对靶蛋白1 (例如, 癌症标志物1) 的第一检测探针和针对靶蛋白2 (例如, 癌症标志物2) 的第二检测探针的双链体系统添加到包含生物实体 (例如, 细胞外囊泡) 的样品中。在一些实施方案中, 检测探针各自包含偶联至寡核苷酸结构域的靶结合部分 (例如, 针对靶蛋白的抗体剂), 所述寡核苷酸结构域包含双链部分和从寡核苷酸结构域的一端延伸的单链悬突。当第一检测探针和第二检测探针的不同靶结合部分 (例如, 分别针对靶蛋白1和靶蛋白2的抗体剂) 紧密邻近地定位至同一生物实体 (例如, 细胞外囊泡) 使得相应的单链悬突彼此杂交, 从而允许它们的寡核苷酸结构域发生连接时, 生成检测信号。例

如,对照实体(例如,来自健康受试者样品的生物实体)不表达靶蛋白1(例如,癌症标志物1)和靶蛋白2(例如,癌症标志物2)中的一个或两个,因此不能生成检测信号。然而,当来自癌症(例如,肺癌)样品的生物实体表达靶蛋白1和靶蛋白2,并且靶蛋白以彼此足够短的距离存在于同一生物实体(例如,细胞外囊泡)中时,生成检测信号。

[0050] 图3示出了不同肺癌和支气管肺癌组织学亚型的患病率的图形表示。数据获得自美国国立癌症研究所的监测、流行病学和最终结果(SEER)计划癌症统计回顾1975-2017(National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program Cancer Statistics Review 1975-2017),其全部内容以引用的方式并入本文。类别“其他”包括大细胞癌、非小细胞癌、其他特定癌、癌NOS、肉瘤和其他特定类型。此回顾中记录了组织学分类的确定和分组。

[0051] 图4是试验患者群组的人口统计学的图形表示。对获得自此群组的合计39份患者血浆样品进行分析,示出了患者肺癌诊断分期的年龄、性别和群组规模的概述。随后通过本文所述的示例性测定对群组样品进行评估。

[0052] 图5是如本文所述的示例性肺腺癌(LUAD)诊断测定的图形表示。使用基于SLC34A2抗体的用CEACAM6+CEACAM6检测探针的EV捕获归一化健康对照和LUAD患者群组的信号。水平截止线表示100%特异性阈值。实现了16.7%的I期LUAD灵敏度、60%的II期LUAD灵敏度、50%的III期LUAD灵敏度和100%的IV期LUAD灵敏度。

[0053] 图6是如本文所述的示例性肺腺癌诊断测定的图形表示。使用基于SLC34A2抗体的用CEACAM6+EPCAM检测探针的EV捕获归一化健康对照和LUAD患者群组的信号。水平截止线表示100%特异性阈值。实现了20%的II期LUAD灵敏度、50%的III期LUAD灵敏度和75%的IV期LUAD灵敏度。

[0054] 图7是如本文所述的示例性肺腺癌诊断测定的图形表示。使用基于CEACAM5抗体的用CEACAM6+SLC34A2检测探针的EV捕获归一化健康对照和LUAD患者群组的信号。水平截止线表示100%特异性阈值。实现了16.7%的I期LUAD灵敏度、20%的II期LUAD灵敏度、50%的III期LUAD灵敏度和75%的IV期LUAD灵敏度。

[0055] 图8是如本文所述的示例性肺腺癌诊断测定之间相关性的图形表示。来自基于SLC34A2抗体的用CEACAM6+CEACAM6检测探针的捕获的信号沿x轴进行描绘,而来自基于SLC34A2抗体的用CEACAM6+EPCAM检测探针的捕获的信号沿y轴进行描绘。使用皮尔逊积矩相关系数确定相关性。观察到强相关性,特别是对III期和IV期LUAD样品。

[0056] 图9是如本文所述的示例性肺腺癌诊断测定之间相关性的图形表示。来自基于SLC34A2抗体的用CEACAM6+CEACAM6检测探针的捕获的信号沿x轴进行描绘,而来自基于CEACAM5抗体的用CEACAM6+SLC34A2检测探针的捕获的信号沿y轴进行描绘。使用皮尔逊积矩相关系数确定相关性。观察到强相关性,特别是对III期和IV期LUAD样品。

[0057] 图10是如本文所述的示例性肺腺癌诊断测定之间相关性的图形表示。来自基于SLC34A2抗体的用CEACAM6+EPCAM检测探针的捕获的信号沿x轴进行描绘,而来自基于CAECAM5抗体的用CEACAM6+SLC34A2检测探针的捕获的信号沿y轴进行描绘。使用皮尔逊积矩相关系数确定相关性。观察到强相关性,特别是对III期和IV期LUAD样品。

[0058] 图11是扩展患者群组的人口统计学的图形表示。对获得自此群组的合计138份患者血浆样品进行分析,示出了患者肺癌诊断分期的年龄、性别和群组规模的概述。随后通过

本文所述的示例性测定对群组样品进行评估。

[0059] 图12是如本文所述的示例性肺腺癌诊断测定的图形表示。使用基于SLC34A2抗体的用CEACAM6+CEACAM6检测探针的EV捕获归一化健康对照和LUAD患者群组的信号。水平截止线表示99.9%特异性阈值。实现了50%的II期LUAD灵敏度、55.5%的III期LUAD灵敏度和71.4%的IV期LUAD灵敏度。

[0060] 图13是说明根据本文所述的一些实施方案的靶实体检测测定的示意图。所述图示出了示例性三链体靶实体检测系统,其中在一些实施方案中,可以将三个或更多个检测探针(每个针对靶蛋白)添加到包含生物实体(例如,细胞外囊泡)的样品中。在一些实施方案中,检测探针各自包含偶联至寡核苷酸结构域的靶结合部分(例如,针对靶蛋白的抗体剂),所述寡核苷酸结构域包含双链部分和从寡核苷酸结构域的一端延伸的单链悬突。当所有三个或更多个检测探针的相应单链悬突彼此杂交以形成线性双链复合物,并且双链复合物的至少一条链发生连接时,生成检测信号,从而允许检测所得的连接产物。

[0061] 图14是包含四个检测探针的双链复合物的非限制性实例,所述检测探针通过它们各自的单链悬突的杂交以线性排列彼此连接。

[0062] 图15是说明本文所述的示例性实施方案的靶实体检测测定的示意图。在一些实施方案中,将多个检测探针(每个针对不同的靶标)添加到包含生物实体(例如,细胞外囊泡)的样品中。在一些实施方案中,检测探针各自包含偶联至寡核苷酸结构域的靶结合部分(例如,抗体剂),所述寡核苷酸结构域包含双链部分和从寡核苷酸结构域的一端延伸的单链悬突。当所有检测探针都紧密邻近同一生物实体(例如,细胞外囊泡或分析物)定位使得相应的单链悬突杂交以形成线性双链复合物,并且所得的线性双链复合物的至少一条链发生连接时,生成检测信号,从而允许检测连接产物。

[0063] 图16A是通过模拟“健康患者”和“癌症患者”对用于LUAD检测的生物标志物组合的辨别能力的图形表示。使用生物信息学分析,分别从正常和癌症组织数据库中随机选择来自5000个模拟的“健康患者”和5000个模拟的“癌症患者”的血浆样品。模拟的“癌症患者”被建模为具有不同大小的肿瘤,范围为1g至1000g。基于“健康患者”和“癌症患者”两个池,计算在99%特异性下检测LUAD的生物标志物组合的灵敏度。

[0064] 图16B是说明了来自表3中列表的每个可能的生物标志物组合如本文所述基于对100g肿瘤的模拟灵敏度检测LUAD的能力的示例性热图。每行表示一个生物标志物组合,并且每列表示一个LUAD癌症患者。浅灰色表明使用给定的生物标志物组合未检测到LUAD患者的癌症,并且深灰色表示使用给定的生物标志物组合检测到LUAD患者的癌症。热图示出了基于100g肿瘤的灵敏度阈值。

[0065] 图16C是说明了基于表3中的列表针对其他癌症类型的每个可能的LUAD生物标志物组合的AUC值的柱状图。为与其他癌症进行比较,计算癌症分子数据库(例如,癌症基因组图谱)中肺腺癌和除肺鳞状细胞癌外的所有其他肿瘤类型的生物标志物组合的EV评分(其中EV评分是给定生物标志物组合中所有TPM表达值的乘积),然后计算AUC。柱状图示出了对所有其他癌症的AUC值的分布。一般来讲,本文描述的生物标志物组合相对于其他癌症类型示出对LUAD的特异性。

[0066] 图17A是相对于健康吸烟者样品池用于检测早期肺腺癌(LUAD)的某些生物标志物组合的图形表示。生物标志物组合按照它们区分早期LUAD样品池和健康吸烟者样品池的能

力排名(排名最高的在图表的顶部)。x轴表示获得自健康吸烟者混合样品和早期LUAD混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0067] 图17B是相对于健康吸烟者样品池用于检测晚期LUAD的某些生物标志物组合的图形表示。生物标志物组合按照它们区分晚期LUAD样品池和健康吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康吸烟者混合样品和晚期LUAD混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0068] 图18A是相对于健康吸烟者样品池用于检测早期肺鳞状细胞癌(LUSC)的某些生物标志物组合的图形表示。生物标志物组合按照它们区分早期LUSC样品池和健康吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康吸烟者混合样品和早期LUSC混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0069] 图18B是相对于健康吸烟者样品池用于检测晚期LUSC的某些生物标志物组合的图形表示。生物标志物组合按照它们区分晚期LUSC样品池和健康吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康吸烟者混合样品和晚期LUSC混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0070] 图19A是相对于健康非吸烟者样品池用于检测早期LUAD的某些生物标志物组合的图形表示。生物标志物组合按照它们区分早期LUAD样品池和健康非吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康非吸烟者混合样品和早期LUAD混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0071] 图19B是相对于健康非吸烟者样品池用于检测晚期LUAD的某些生物标志物组合的图形表示。生物标志物组合按照它们区分晚期LUAD样品池和健康非吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康非吸烟者混合样品和晚期LUAD混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0072] 图20A是相对于健康非吸烟者样品池用于检测早期LUSC的某些生物标志物组合的图形表示。生物标志物组合按照它们区分早期LUSC样品池和健康非吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康非吸烟者混合样品和早期LUSC混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0073] 图20B是相对于健康非吸烟者样品池用于检测晚期LUSC的某些生物标志物组合的图形表示。生物标志物组合按照它们区分晚期LUSC样品池和健康非吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康非吸烟者混合样品和晚期LUSC混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0074] 图21A是相对于健康吸烟者样品池用于检测早期LUAD的某些生物标志物组合的图形表示。生物标志物组合按照它们区分早期LUAD样品池和健康吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康吸烟者混合样品和早期LUAD混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0075] 图21B是相对于健康吸烟者样品池用于检测晚期LUAD的某些生物标志物组合的图形表示。生物标志物组合按照它们区分晚期LUAD样品池和健康吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康吸烟者混合样品和晚期LUAD混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0076] 图22A是相对于健康吸烟者样品池用于检测早期LUSC的某些生物标志物组合的图形表示。生物标志物组合按照它们区分早期LUSC样品池和健康吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康吸烟者混合样品和早期LUSC混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0077] 图22B是相对于健康吸烟者样品池用于检测晚期LUSC的某些生物标志物组合的图形表示。生物标志物组合按照它们区分晚期LUSC样品池和健康吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康吸烟者混合样品和晚期LUSC混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0078] 图23A是相对于健康非吸烟者样品池用于检测早期LUAD的某些生物标志物组合的图形表示。生物标志物组合按照它们区分早期LUAD样品池和健康非吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康非吸烟者混合样品和早期LUAD混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0079] 图23B是相对于健康非吸烟者样品池用于检测晚期LUAD的某些生物标志物组合的图形表示。生物标志物组合按照它们区分晚期LUAD样品池和健康非吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康非吸烟者混合样品和晚期LUAD混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0080] 图24A是相对于健康非吸烟者样品池用于检测早期LUSC的某些生物标志物组合的图形表示。生物标志物组合按照它们区分早期LUSC样品池和健康非吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康非吸烟者混合样品和早期LUSC混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0081] 图24B是相对于健康非吸烟者样品池用于检测晚期LUSC的某些生物标志物组合的图形表示。生物标志物组合按照它们区分晚期LUSC样品池和健康非吸烟者样品池的能力排名(排名最高的在图表的顶部)。x轴表示获得自健康非吸烟者混合样品和晚期LUSC混合样品的Ct值的差异。y轴表示某些生物标志物组合(捕获探针的靶标、检测探针的靶标)。

[0082] 某些定义

[0083] 施用:如本文所用,术语“施用(administering/administration)”通常是指将组合物施用于受试者,以实现将是组合物或包含在组合物中的剂递送到靶位点或待治疗的位点。本领域普通技术人员将了解在适当的情况下可以用于施用于受试者(例如人)的多种途径。例如,在一些实施方案中,施用可以是胃肠外的。在一些实施方案中,施用可以是口服的。在一些实施方案中,施用可以仅涉及单剂量。在一些实施方案中,施用可以涉及应用固定数量的剂量。在一些实施方案中,施用可以涉及间歇性(例如,在时间上分开的多个剂量)和/或周期性(例如,由常见的时间段分开的单独剂量)给药的给药。在一些实施方案中,施用可以涉及连续给药(例如,灌注)至少一段选定的时间。

[0084] 扩增:术语“扩增(amplification和amplify)”是指导致核酸分子的量和/或水平相对于其初始量和/或水平增加的模板依赖性过程。模板依赖性过程一般是涉及引物分子的模板依赖性延伸的过程,其中新合成的核酸链的序列由众所周知的互补碱基配对规则决定(参见例如,Watson,J.D.等人于:Molecular Biology of the Gene,第4版,W.A.Benjamin,Inc.,Menlo Park,Calif.(1987);出于本文所述的目的,所述文献以引用方式并入本文)。

[0085] 抗体剂:如本文所用,术语“抗体剂”是指与特定抗原特异性结合的剂。在一些实施方案中,术语涵盖包含足以赋予特异性结合的免疫球蛋白结构元件的任何多肽或多肽复合物。示例性抗体剂包括但不限于单克隆抗体或多克隆抗体。在一些实施方案中,抗体剂可以包含作为小鼠、兔、灵长类动物或人抗体特征的一个或多个恒定区域序列。在一些实施方案中,抗体剂可以包含如本领域已知的人源化、灵长类化、嵌合等的一个或多个序列元件。在许多实施方案中,术语“抗体剂”用于指用于在替代呈现中利用抗体结构和功能特征的本领域已知或开发的构建体或形式中的一种或多种。例如,实施方案中,根据本发明使用的抗体剂的形式选自但不限于完整的IgA、IgG、IgE或IgM抗体;双特异性抗体或多特异性抗体(例如,Zybodies®等);抗体片段,诸如Fab片段、Fab'片段、F(ab')₂片段、Fd'片段、Fd片段和分离的互补决定区(CDR)或其组;单链Fv;多肽-Fc融合;单结构域抗体(例如,鲨鱼单结构域抗体,诸如IgNAR或其片段);骆驼抗体;掩蔽抗体(例如,Probodies®);小模块免疫药物(“SMIPsTM”);单链或串联双抗体(TandAb®);VHH; Anticalins®; Nanobodies®微体; BiTE®; 锚蛋白重复蛋白或DARPINs®; Avimers®; DART; TCR样抗体; Adnectins®; Affilins®; Trans-bodies®; Affibodies®; TrimerX®;微蛋白;

Fynomers®、Centyrins®、KALBITOR®和Affimers®。在一些实施方案中,抗体可能缺少在天然产生时具有的共价修饰(例如,聚糖的附接)。在一些实施方案中,抗体可以含有共价修饰(例如,聚糖、有效载荷[例如,可检测部分、治疗部分、催化部分等]或其他侧基[例如,聚乙二醇等]的附着)。在许多实施方案中,抗体剂是或包含其氨基酸序列包含本领域技术人员识别为互补决定区(CDR)的一个或多个结构元件的多肽;在一些实施方案中,抗体剂是或包含其氨基酸序列包含与在参考抗体中发现的CDR基本相同的至少一个CDR(例如,至少一个重链CDR和/或至少一个轻链CDR)的多肽。在一些实施方案中,所包括的CDR与参考CDR基本相同,因为与参考CDR相比,它在序列上相同或含有1-5个之间的氨基酸取代。在一些实施方案中,所包括的CDR与参考CDR基本相同,因为它示出与参考CDR至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的序列同一性。在一些实施方案中,所包括的CDR与参考CDR基本相同,因为它示出与参考CDR至少95%、96%、97%、98%、99%或100%的序列同一性。在一些实施方案中,所包括的CDR与参考CDR基本相同,因为与参考CDR相比,所包括的CDR内的至少一个氨基酸被删除、添加或取代,但是所包括的CDR具有与参考CDR的氨基酸序列以其他方式相同的氨基酸序列。在一些实施方案中,所包括的CDR与参考CDR基本相同,因为与参考CDR相比,所包括的CDR内的1-5个氨基酸被删除、添加或取代,但是所包括的CDR具有与参考CDR以其他方式相同的氨基酸序列。在一些实施方案中,所包括的CDR与参考CDR基本相同,因为与参考CDR相比,所包括的CDR内的至少一个氨基酸被取代,但是所包括的CDR具有与参考CDR的氨基酸序列以其他方式相同的氨基酸序列。在一些实施方案中,所包括的CDR与参考CDR基本相同,因为与参考CDR相比,所包括的CDR内的1-5个氨基酸被删除、添加或取代,但是所包括的CDR具有与参考CDR以其他方式相同的氨基酸序列。在一些实施方案中,抗体剂是或包含其氨基酸序列包含本领域技术人员识别为免疫球蛋白可变结构域的结构元件的多肽。在一些实施方案中,抗体剂是具有与免疫球蛋白结合结构域同源或大体同源的结合结构域的多肽蛋白。

[0086] 技术人员可以使用本领域已知的方法和可商购获得的服务和试剂盒来制备抗体

剂。例如,制备单克隆抗体的方法是本领域众所周知的,并且包括杂交瘤技术和噬菌体展示技术。适合在本公开中使用的另外的抗体描述于例如以下出版物:Antibodies A Laboratory Manual,第二版Edward A.Greenfield.Cold Spring Harbor Laboratory Press (September 30,2013);Making and Using Antibodies:A Practical Handbook,第二版编辑Gary C.Howard and Matthew R.Kaser.CRC Press (July 29,2013);Antibody Engineering:Methods and Protocols,第二版(Methods in Molecular Biology).Patrick Chames.Humana Press (August 21,2012);Monoclonal Antibodies:Methods and Protocols(Methods in Molecular Biology).编辑Vincent Ossipow and Nicolas Fischer.Humana Press (February 12,2014);以及Human Monoclonal Antibodies:Methods and Protocols(Methods in Molecular Biology).Michael Steinitz.Humana Press (September 30,2013)。

[0087] 抗体可以通过标准技术产生,例如通过用适当的多肽或其一个或多个部分进行免疫,或通过使用噬菌体展示文库。如果需要多克隆抗体,则用带有一个或多个所需表位的免疫原性多肽对所选宿主动物(例如,小鼠、兔、山羊、马、鸡等)进行免疫,所述免疫原性多肽任选地半抗原化到另一多肽。根据宿主物种,可以使用各种佐剂来增加免疫反应。此类佐剂包括但不限于弗氏、矿物凝胶(诸如氢氧化铝)和表面活性物质(诸如溶血卵磷脂、普洛尼克多元醇(pluronic polyol)、聚阴离子、肽、油乳剂、匙孔血蓝蛋白和二硝基苯酚)。收集来自免疫动物的血清,并根据已知程序进行处理。如果含有针对所需表位的多克隆抗体的血清含有针对其他抗原的抗体,则可以通过免疫亲和色谱法或本领域已知的任何其他方法来纯化多克隆抗体。用于生产和加工多克隆抗血清的技术是本领域众所周知的。

[0088] 大约或约:如本文所用,当应用于一个或多个关注的值时,术语“大约”或“约”是指与所述参考值类似的值。一般来讲,熟悉上下文的本领域技术人员将理解所述上下文中“约”或“大约”所涵盖的相关差异程度。例如,在一些实施方案中,术语“大约”或“约”可以涵盖在参考值的25%、20%、19%、18%、17%、16%、15%、14%、13%、12%、11%、10%、9%、8%、7%、6%、5%、4%、3%、2%、1%或更小范围内的值范围。

[0089] 适体:如本文所用,术语“适体”通常是指结合至特定靶分子(例如,表位)的核酸分子或肽分子。在一些实施方案中,核酸适体可以由核苷酸序列描述,并且长度通常为约15-60个核苷酸。核酸适体可以是或包含单链和/或双链结构。在一些实施方案中,核酸适体可以是或包含DNA。在一些实施方案中,核酸适体可以是或包含RNA。不希望受任何理论束缚,设想适体中的核苷酸链形成分子内相互作用,所述分子内相互作用将分子折叠成复杂的三维形状,并且此三维形状允许适体与其靶分子的表面紧密结合。在一些实施方案中,肽适体可以被描述为具有由蛋白质支架展示的一个或多个可变序列肽环。肽适体可以从组合文库中分离出来,并且通常随后通过定向突变或多轮可变区诱变和选择进行改进。鉴于存在于所有可能的核苷酸和/或肽序列的范畴中的分子形状的非凡多样性,可以获得用于广泛分子靶标的适体,包括蛋白质和小分子。除了高特异性之外,适体通常对其靶标具有非常高的亲和力(例如,对于蛋白质或多肽的亲和力在皮摩尔至低纳摩尔范围内)。因为适体通常是合成分子,所以适体适于进行各种修饰,这可以优化它们在特定应用中的功能。

[0090] 相关:如果一个事件或实体的存在、水平和/或形式与另一个事件或实体的存在、水平和/或形式相关,则两个事件或实体彼此“相关”,如本文所用的术语。例如,如果特定生

生物学现象(例如,特定生物标志物的表达)的存在与肺癌(例如,特定类型的肺癌和/或分期的肺癌)的发病和/或易患性(例如,在有关人群中)相关联,则认为特定生物学现象与肺癌相关。

[0091] 生物实体:在适当的情况下,如本领域技术人员将从上下文中清楚的那样,术语“生物实体”可以用于指存在于例如在一些实施方案中来源于或获得自受试者的生物样品中的实体或组分,所述生物样品在一些实施方案中可以是或包含细胞或生物体,诸如动物或人,或在一些实施方案中可以是或包含生物组织或流体。在一些实施方案中,生物实体是或包含细胞或微生物,或其部分、提取物或组分(包括例如,由细胞或微生物分泌的细胞内组分和/或分子)。例如,在一些实施方案中,生物实体是或包含细胞。在一些实施方案中,生物实体是或包含细胞外囊泡。在一些实施方案中,生物实体是或包含生物分析物(例如,代谢物、碳水化合物、蛋白质或多肽、酶、脂质、细胞器、细胞因子、受体、配体及其任何组合)。在一些实施方案中,存在于样品中的生物实体处于天然状态(例如,蛋白质或多肽保持在天然存在的构象结构中)。在一些实施方案中,例如通过从样品中分离或来源于天然存在的生物实体来处理生物实体。例如,可以用一种或多种化学剂处理生物实体,使得利用本文提供的技术进行检测更为理想。仅作为实例,生物实体可以是与固定剂(例如但不限于甲醇和/或甲醛)接触以使存在于细胞或细胞外囊泡中的蛋白质和/或肽形成交联的细胞或细胞外囊泡。在一些实施方案中,生物实体呈分离的或纯的形式(例如,从体液样品(例如像血液、血清、血浆样品等)中分离)。在一些实施方案中,生物实体可以存在于复杂基质(例如,体液样品(例如像血液、血清或血浆样品等))中。

[0092] 生物标志物:术语“生物标志物”通常是指其存在、水平、程度、类型和/或形式与特定的生物事件或关注的状态相关联的实体、事件或特征,因此其被认为是所述事件或状态的“标志物”。仅举几个实例,在一些实施方案中,生物标志物可以是或包含针对特定疾病状态或针对特定疾病、病症或病状可能发展、发病或复发的可能性的标志物。在一些实施方案中,生物标志物可以是或包含针对特定疾病或治疗结果或其可能性的标志物。在一些实施方案中,生物标志物可以是或包含针对特定组织(例如但不限于脑、乳腺、结肠、卵巢和/或与女性生殖系统相关的其他组织、胰腺、前列腺和/或与男性生殖系统相关的其他组织、肝脏、肺和皮肤)的标志物。在一些实施方案中,针对特定组织的此类标志物可以对健康组织具有特异性,对患病组织具有特异性,或在一些实施方案中可以存在于正常健康组织和患病组织(例如,肿瘤)中;阅读本公开的本领域技术人员将理解每种此类生物标志物的适当上下文。在一些实施方案中,生物标志物可以是或包含癌症特异性标志物(例如,对特定癌症具有特异性的标志物)。在一些实施方案中,生物标志物可以是或包含非特异性癌症标志物(例如,存在于至少两种或更多种癌症中的标志物)。在一些实施方案中,非特异性癌症标志物可以是或包含针对癌症的通用标志物(例如,通常存在于癌症中的标志物,无论组织类型如何),或在一些实施方案中,针对特定组织(例如但不限于脑、乳腺、结肠、卵巢和/或与女性生殖系统相关的其他组织、胰腺、前列腺和/或与男性生殖系统相关的其他组织、肝脏、肺和皮肤)的癌症的标志物。因此,在一些实施方案中,生物标志物是预测性的;在一些实施方案中,生物标志物是预后的;在一些实施方案中,生物标志物是有关生物事件或关注的状态的诊断物。生物标志物可以是或包含任何化学类别的实体,并且可以是或包含实体的组合。例如,在一些实施方案中,生物标志物可以是或包含核酸、多肽、脂质、碳水化合物、小分

子、无机剂(例如,金属或离子)或其组合。在一些实施方案中,生物标志物是或包含特定分子、复合物或结构的一部分;例如,在一些实施方案中,生物标志物可以是或包含表位。在一些实施方案中,生物标志物是与肺癌相关的细胞外囊泡的表面标志物(例如,表面蛋白标志物)。在一些实施方案中,生物标志物是囊泡内的(例如,存在于细胞外囊泡内的蛋白质或RNA标志物)。在一些实施方案中,生物标志物可以是或包含遗传或表观遗传特征。在一些实施方案中,生物标志物可以是或包含基因表达特征。在一些实施方案中,适合根据本公开使用的“生物标志物”可以指存在于靶标志物中的分子实体(例如,表位)的存在、水平和/或形式。例如,在一些实施方案中,作为分子实体(例如,表位)的两个或更多个“生物标志物”可以存在于相同的靶标志物(例如,标志物蛋白,诸如存在于细胞外囊泡中的表面蛋白)上。

[0093] 血液来源样品:如本文所用,术语“血液来源样品”是指来源于有需要的受试者的血液样品(即,全血样品)的样品。血液来源样品的实例包括但不限于血浆(包括例如新鲜冷冻血浆)、血清、血液部分、血浆部分、血清部分、血液部分(包含红细胞(RBC)、血小板、白细胞等),以及包括其部分的细胞裂解物(例如,细胞(诸如红细胞、白细胞等)可以被收获并裂解以获得细胞裂解物)。在一些实施方案中,与本文所述的方法、系统和/或试剂盒一起使用的血液来源样品是血浆样品。

[0094] 癌症:术语“癌症”在本文中一般用于指其中关注的组织的细胞表现出相对异常、不受控制和/或自主生长,使得它们表现出特征为细胞增殖显著失去控制的异常生长表型的疾病或病状。在一些实施方案中,癌症可以包含癌前(例如,良性)、恶性、转移前、转移性和/或非转移性细胞。本公开提供了用于检测肺癌的技术。

[0095] 分类截止值:如本文所用,术语“分类截止值”是指例如通过在群体的两个或更多个子集(例如,正常健康受试者和患有炎症病状的受试者与肺癌受试者)之间定义一条或多条分界线,用于预测受试者患疾病或病状(例如,肺癌)的风险的水平、值或评分,或一组值,或指标。在一些实施方案中,可以参考本文所述的靶生物标志物特征的至少一个参考阈值水平(例如,参考截止值),任选地结合其他适当的变量,例如年龄、生活史相关风险因素、遗传因素、受试者的身体和/或医疗状况来确定分类截止值。在分类基于单一靶生物标志物特征(例如,如本文所述)的一些实施方案中,分类截止值可以与单一靶生物标志物特征预先确定的参考阈值(例如,截止值)相同。在分类基于两个或更多个靶生物标志物特征的一些实施方案中,分类截止值可以参考两个或更多个参考阈值(例如,截止值),每个参考阈值是针对相应的靶生物标志物特征单独预先确定的,并且任选地并入一个或多个适当的变量,例如年龄、生活史相关风险因素、遗传因素、受试者的身体和/或医疗状况。在一些实施方案中,可以通过参考本文所述的靶生物标志物特征的至少一个参考阈值水平(例如,参考截止值),任选地结合其他适当的变量,例如年龄、生活史相关风险因素、遗传因素、受试者的身体和/或医疗状况的计算机算法介导的分析来确定分类截止值。

[0096] 紧密邻近:如本文所用,术语“紧密邻近”是指两个检测探针(例如,成对的两个检测探针)之间的距离足够近,使得预期检测探针之间的相互作用(例如,通过各自的寡核苷酸结构域)可能发生。例如,在一些实施方案中,当两个检测探针在特定条件下(例如,当检测探针结合至细胞外囊泡中的相应靶标时)彼此足够紧密邻近时,它们在一段时间内彼此相互作用(例如,通过各自的寡核苷酸结构域)的概率为至少50%或更多,包括例如至少60%、至少70%、至少80%、至少90%或更多。在一些实施方案中,当两个检测探针彼此足够

紧密邻近时,它们之间的距离可以在大约0.1-1000nm,或0.5-500nm,或1-250nm之间的范围内。在一些实施方案中,当两个检测探针彼此足够紧密邻近时,它们之间的距离可以在大约0.1-10nm之间或大约0.5-5nm之间的范围内。在一些实施方案中,当两个检测探针彼此足够紧密邻近时,它们之间的距离可以小于100nm或更短,包括例如小于90nm、小于80nm、小于70nm、小于60nm、小于50nm、小于40nm、小于30nm、小于20nm、小于10nm、小于5nm、小于1nm或更短。在一些实施方案中,当两个检测探针彼此足够紧密邻近时,它们之间的距离可以在大约40-1000nm或40nm-500nm之间的范围内。

[0097] 相当的:如本文所用,术语“相当的”是指两种或更多种剂、实体、情况、条件组等可能彼此不相同,但是足够类似以允许在它们之间进行比较,使得本领域技术人员将理解,可以基于观察到的差异或类似性合理地得出结论。在一些实施方案中,可比较的条件组、环境、个体或群体的特征在于多个基本上相同的特征和一个或少量的不同特征。本领域普通技术人员将理解,在上下文中,在任何给定情况下,两种或更多种此类剂、实体、情况、条件组等被认为是相当的需何种程度的同一性。例如,本领域普通技术人员将理解,当以足够数量和类型的基本相同的特征为特征时,环境、个体或群体组彼此可比,以保证合理结论,即在不同的环境、个体或群体组的情况下获得的结果或观察到的现象的差异是由那些变化的特征的变化引起或指示的。

[0098] 互补:如本文所用,术语“互补”用于指通过碱基配对规则相关的寡核苷酸杂交。例如,序列“C-A-G-T”与序列“G-T-C-A”互补。互补性可以是部分的或全部的。因此,只要部分互补性允许寡核苷酸杂交,任何程度的部分互补性都旨在包括在术语“互补性”的范围内。部分互补性是其中根据碱基配对规则,一个或多个核酸碱基不匹配。核酸之间的全部或完全互补性是其中每个核酸碱基在碱基配对规则下与另一个碱基匹配。

[0099] 检测:术语“检测”在本文中广泛使用以包括确定存在或不存在表达肺癌的靶生物标志物特征的细胞外囊泡或指示此类细胞外囊泡的任何形式的测量的适当方式。因此,“检测”可以包括确定、测量、评估或测定以任何方式对应于靶生物标志物特征的部分的关注的实体(例如,表面蛋白生物标志物、囊泡内蛋白生物标志物或囊泡内RNA生物标志物)的存在或不存在、水平、量和/或位置。在一些实施方案中,“检测”可以包括确定、测量、评估或定量指示关注的实体(例如,指示表面蛋白质生物标志物和/或囊泡内蛋白质生物标志物的连接模板,或指示囊泡内mRNA的PCR扩增产物)的测量的形式。包括定量和定性确定、测量或评估,包括半定量。此类确定、测量或评估可以是相对的,例如当相对于对照参考或绝对检测关注的实体(例如,表面蛋白生物标志物、囊泡内蛋白生物标志物或囊泡内RNA生物标志物)或指示其测量的形式时。因此,当术语“定量”用在定量关注的实体(例如,表面蛋白生物标志物、囊泡内蛋白质生物标志物或囊泡内RNA生物标志物)或指示其测量的形式的上下文中时,可以指绝对或相对定量。绝对定量可以通过将关注的实体(例如,表面蛋白生物标志物、囊泡内蛋白生物标志物或囊泡内RNA生物标志物)或指示其测量的形式的检测水平与已知的对照标准相关联(例如,通过生成标准曲线)来完成。替代地,相对定量可以通过比较两个或更多个不同的关注的实体(例如,不同的表面蛋白生物标志物,或囊泡内蛋白生物标志物,或囊泡内RNA生物标志物)之间的检测水平或量来完成,以提供两个或更多个不同的关注的实体中的每个的相对定量,即相对于彼此。

[0100] 检测标签:如本文所用,术语“检测标签”是指可检测的任何元素、分子、官能团、化

合物、片段或部分。在一些实施方案中,单独地提供或使用检测标签。在一些实施方案中,关联(例如,结合至)另一种剂提供和/或使用检测标签。检测标签的实例包括但不限于:各种配体、放射性核素(例如, ^3H 、 ^{14}C 、 ^{18}F 、 ^{19}F 、 ^{32}P 、 ^{35}S 、 ^{135}I 、 ^{125}I 、 ^{123}I 、 ^{64}Cu 、 ^{187}Re 、 ^{111}In 、 ^{90}Y 、 $^{99\text{m}}\text{Tc}$ 、 ^{177}Lu 、 ^{89}Zr 等)、荧光染料、化学发光剂(例如像吖啶酯、稳定的二氧杂环丁烷等)、生物发光剂、光谱可分辨的无机荧光半导体纳米晶体(即,量子点)、金属纳米颗粒(例如,金、银、铜、铂等)纳米簇、顺磁性金属离子、酶、比色标签(例如像染料、胶体金等)、生物素、地高辛、半抗原以及其中抗血清或单克隆抗体可用的蛋白质。

[0101] 检测探针:术语“检测探针”通常是指针对特定靶标的检测和/或量化的探针。在一些实施方案中,检测探针是定量探针,其提供表示特定靶标水平的指标。根据本公开,检测探针是指包含直接或间接偶联至寡核苷酸结构域的靶结合实体的组合物,其中靶结合实体特异性结合至相应的靶标(例如,分子靶标),并且其中寡核苷酸结构域的至少一部分被设计成允许与针对不同靶标的另一检测探针的寡核苷酸结构域的一部分杂交。在许多实施方案中,适合根据本公开使用的寡核苷酸结构域包含双链部分和至少一个单链悬突。在一些实施方案中,寡核苷酸结构域可以包含双链部分和位于双链部分的每一端的单链悬突。

[0102] 双链:如本文所用,在寡核苷酸结构域的上下文中,术语“双链”被本领域技术人员理解为以氢键螺旋排列形式存在的一对寡核苷酸,其通常与例如核酸(诸如DNA)相关。除了双链寡核苷酸的100%互补形式之外,如本文所用的术语“双链”还意指包括错配(例如,部分互补性)和/或作为凸起、环或发夹的结构特征的那些形式。

[0103] 双链复合物:如本文所用,术语“双链复合物”通常是指包含至少两个或更多个(包括例如,2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或更多个)检测探针(例如,如本文提供和/或使用的)的复合物,每个检测探针针对靶标(其可以是相同的靶标或不同的靶标),通过检测探针的互补单链悬突的杂交以线性排列彼此连接或偶联。在一些实施方案中,此类双链复合物可以包含细胞外囊泡,其中检测探针各自的靶结合部分同时结合至细胞外囊泡。

[0104] 表位:如本文所用,术语“表位”包括由免疫球蛋白(例如,抗体或受体)结合组分或适体特异性识别的任何部分。在一些实施方案中,表位由抗原上的多个化学原子或基团组成。在一些实施方案中,当抗原采用有关三维构象时,此类化学原子或基团暴露于表面。在一些实施方案中,当抗原采用此类构象时,此类化学原子或基团在空间上在物理上彼此接近。在一些实施方案中,当抗原采用替代构象(例如,线性化)时,至少一些此类化学原子或基团在物理上彼此分离。

[0105] 细胞外囊泡:如本文所用,术语“细胞外囊泡”通常是指细胞外的囊泡,例如由细胞分泌的。分泌囊泡的实例包括但不限于外泌体、微囊泡、微粒、核外粒体、原癌小体和凋亡小体。不希望受理论束缚,外泌体是内吞起源的纳米尺寸囊泡(例如,在40nm和120nm之间),其可通过多囊泡内体(MVE)的界膜向内出芽形成,而微囊泡通常从细胞表面出芽,并且它们的大小可以在50nm和1000nm之间变化。在一些实施方案中,细胞外囊泡是或包含外泌体和/或微囊泡。在一些实施方案中,包含细胞外囊泡的样品基本上不含凋亡小体。在一些实施方案中,包含细胞外囊泡的样品可以包含从一个或多个组织(例如,癌组织和/或非癌或健康组织)脱落或衍生的细胞外囊泡。在一些实施方案中,样品中的细胞外囊泡可以从肺癌肿瘤脱落或衍生;在一些实施方案中,细胞外囊泡从非肺癌肿瘤脱落或衍生。在一些实施方案中,

细胞外囊泡从健康组织脱落或衍生。在一些实施方案中,细胞外囊泡从良性肺肿瘤脱落或衍生。在一些实施方案中,细胞外囊泡从具有与肺癌相关的症状(例如,非特异性症状)的受试者的组织脱落或衍生。

[0106] 细胞外囊泡相关膜结合多肽:如本文所用,此术语是指存在于细胞外囊泡的膜中的多肽。在一些实施方案中,此类多肽可以是肿瘤特异性的。在一些实施方案中,此类多肽可以是组织特异性的(例如,肺组织特异性的)。在一些实施方案中,此类多肽可以是非特异性的,例如,其存在于一种或多种非靶肿瘤和/或一种或多种非靶组织中。

[0107] 杂交:如本文所用,术语“杂交(hybridizing/hybridize/hybridization)”、“退火(annealing/anneal)”在提及使用核酸链通过碱基配对与互补链结合以形成杂交复合物的任何过程配对互补核酸时可互换使用。杂交和杂交强度(例如,核酸之间的结合强度)受各种因素影响,包括例如核酸之间的互补程度、所涉及条件的严格性、所形成的杂交复合物的解链温度(T)以及核酸内的G:C比率。

[0108] 囊泡内蛋白生物标志物:如本文所用,术语“囊泡内蛋白生物标志物”是指指示存在于生物实体(例如,细胞或细胞外囊泡)内的多肽的状态(例如,存在、水平和/或活性)的标志物。在许多实施方案中,囊泡内蛋白生物标志物与细胞外囊泡相关或存在于细胞外囊泡内。

[0109] 囊泡内RNA生物标志物:如本文所用,术语“囊泡内RNA生物标志物”是指指示存在于生物实体(例如,细胞或细胞外囊泡)内的RNA(例如,mRNA)的状态(例如,存在和/或水平)的标志物。在许多实施方案中,囊泡内RNA生物标志物与细胞外囊泡相关或存在于细胞外囊泡内。

[0110] 连接酶:如本文所用,术语“连接酶”或“核酸连接酶”是指用于连接核酸的酶。在一些实施方案中,连接酶是用于将多核苷酸的3'端连接至多核苷酸的5'端的酶。在一些实施方案中,连接酶是用于进行粘端连接的酶。在一些实施方案中,连接酶是用于进行平端连接的酶。在一些实施方案中,连接酶是或包含DNA连接酶。

[0111] 生活史相关风险因素:如本文所用,术语“生活史风险因素”是指个体在他们的生活中的行为、经历、病史和/或暴露,相比于在他们的生活中没有此类行为、经历、病史和/或暴露的个体,所述行为、经历、病史和/或暴露可能直接或间接增加此类病状(例如,肺癌)的个体风险。在一些实施方案中,生活史相关风险因素的非限制性实例包括吸烟、酒精、药物、致癌剂、饮食、肥胖、糖尿病、慢性阻塞性肺部疾病(COPD)、身体活动、光暴露、放射暴露、沥青烟气暴露、暴露于诸如病毒和细菌的感染性因子和/或职业危害(参见例如Jyoti Malhotra等人,“Risk Factors for Lung Cancer Worldwide”European Respiratory journal (2016) 48:889-902;其出于本文所述目的以引用的方式并入本文)。本领域技术人员认识到,上述导致癌症(例如,肺癌)易感性的生活史相关风险因素列表并不是详尽的,而是不断发展的。

[0112] 连接:如本文所用,术语“连接(ligate/ligating/ligation)”是指本领域已知的用于连接两个寡核苷酸或多核苷酸的方法或组合物。连接可以是或包括粘端连接或平端连接。在一些实施方案中,提供的技术中涉及的连接是或包括粘端连接。在一些实施方案中,连接是指将多核苷酸的3'端连接至多核苷酸的5'端。在一些实施方案中,通过使用核酸连接酶促进连接。

[0113] 非癌症受试者:如本文所用,术语“非癌症受试者”一般是指未患有非良性肺癌的受试者。例如,在一些实施方案中,非癌症受试者是健康受试者。在一些实施方案中,非癌症受试者是年龄低于55岁的健康受试者。在一些实施方案中,非癌症受试者是年龄为55岁或以上的健康受试者。在一些实施方案中,非癌症受试者是患有非肺相关健康疾病、病症或病状的受试者。在一些实施方案中,非癌症受试者是患有良性肺肿瘤(例如,在胸腔或肺空洞中观察到的良性肿块)的受试者。

[0114] 核酸/寡核苷酸:如本文所用,术语“核酸”是指具有至少10个或更多个核苷酸的聚合物。在一些实施方案中,核酸是或包含DNA。在一些实施方案中,核酸是或包含RNA。在一些实施方案中,核酸是或包含肽核酸(PNA)。在一些实施方案中,核酸是或包含单链核酸。在一些实施方案中,核酸是或包含双链核酸。在一些实施方案中,核酸包含单链和双链部分。在一些实施方案中,核酸包含含有一个或多个磷酸二酯键联的主链。在一些实施方案中,核酸包含含有磷酸二酯键联和非磷酸二酯键联的主链。例如,在一些实施方案中,核酸可以包含含有一个或多个硫代磷酸酯或5'-N-亚磷酰胺键联和/或一个或多个肽键的主链,例如,如在“肽核酸”中。在一些实施方案中,核酸包含一种或多种或所有天然残基(例如,腺嘌呤、胞嘧啶、脱氧腺苷、脱氧胞苷、脱氧鸟苷、脱氧胸苷、鸟嘌呤、胸腺嘧啶、尿嘧啶)。在一些实施方案中,核酸包含一种或多种或所有非天然残基。在一些实施方案中,非天然残基包含核苷类似物(例如,2-氨基腺苷、2-硫代胸苷、肌苷、吡咯并嘧啶、3-甲基腺苷、5-甲基胞苷、C-5丙炔基-胞苷、C-5丙炔基-尿苷、2-氨基腺苷、C5-溴尿苷、C5-氟尿苷、C5-碘尿苷、C5-丙炔基-尿苷、C5-丙炔基-胞苷、C5-甲基胞苷、2-氨基腺苷、7-脱氮腺苷、7-脱氮鸟苷-氧代腺苷、8-氧代腺苷、8-氧代鸟苷、6-0-甲基鸟嘌呤、2-硫代胞苷、甲基化碱基、插入的碱基及其组合)。在一些实施方案中,与天然残基中的糖相比,非天然残基包含一种或多种修饰的糖(例如,2'-氟核糖、核糖、2'-脱氧核糖、阿拉伯糖和己糖)。在一些实施方案中,核酸具有编码功能基因产物诸如RNA或多肽的核苷酸序列。在一些实施方案中,核酸具有包含一个或多个内含子的核苷酸序列。在一些实施方案中,核酸可以通过从天然来源分离、酶促合成(例如,通过基于互补模板的聚合,例如在体内或体外)、在重组细胞或系统中再生或化学合成来制备。在一些实施方案中,核酸为至少3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、110、120、130、140、150、160、170、180、190、200、225、250、275、300、325、350、375、400、425、450、475、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、5500、6000、6500、7000、7500、8000、8500、9000、9500、10,000、10,500、11,000、11,500、12,000、12,500、13,000、13,500、14,000、14,500、15,000、15,500、16,000、16,500、17,000、17,500、18,000、18,500、19,000、19,500或20,000个或更多个残基或核苷酸长。

[0115] 核苷酸:如本文所用,术语“核苷酸”是指其领域公认的含义。当使用多个核苷酸作为大小(例如寡核苷酸的大小)的指示时,一定数量的核苷酸是指单链上的核苷酸(例如寡核苷酸)的数量。

[0116] 患者:如本文所用,术语“患者”是指患有疾病或病症或病状或处于疾病或病症或病状的风险的任何生物体。典型的患者包括动物(例如,哺乳动物,诸如小鼠、大鼠、兔、非人灵长类动物和/或人)。在一些实施方案中,患者是人。在一些实施方案中,患者患有或易患一种或多种疾病或病症或病状。在一些实施方案中,患者表现出疾病或病症或病状的一种

或多种症状。在一些实施方案中,患者已被诊断为患有的一种或多种疾病或病症或病状。在一些实施方案中,适用提供的技术的疾病或病症或病状是或包括癌症,或一种或多种肿瘤的存在。在一些实施方案中,患者正在接受或已经接受某些疗法以诊断和/或治疗疾病、病症或病状。

[0117] 多肽:如本文所用,术语“多肽”通常具有至少三个或更多个氨基酸的聚合物的其领域公认的含义。本领域普通技术人员将理解,术语“多肽”旨在足够通用,以便不仅涵盖具有本文所述的具有完整序列的多肽,而且还涵盖表示此类完整多肽的功能性、生物活性或特征性片段、部分或结构域(例如,保留至少一种活性的片段、部分或结构域)的多肽。在一些实施方案中,多肽可以含有L-氨基酸、D-氨基酸或两者和/或可以含有本领域已知的多种氨基酸修饰或类似物中的任一种。有用的修饰包括例如末端乙酰化、酰胺化、甲基化等。在一些实施方案中,多肽可以包含天然氨基酸、非天然氨基酸、合成氨基酸及其组合(例如,可以是或包含拟肽)。

[0118] 预防(prevent/prevention):如本文所用,当与疾病、病症和/或病状的发生相关联时,“预防(prevent/prevention)”是指降低发展疾病、病症和/或病状的风险和/或延迟疾病、病症或病状的一个或多个特征或症状的发作。当疾病、病症或病状的发作已延迟预定时间段时,可以认为预防是完全的。

[0119] 引物:如本文所用,术语“引物”是指当置于诱导与核酸链互补的引物延伸产物的合成的条件下(例如,在核苷酸和诱导剂(诸如DNA聚合酶)的存在下以及在合适的温度和pH下)时能够充当合成起始点的寡核苷酸。引物优选为单链以最大化扩增效率。引物必须足够长以在诱导剂的存在下引发延伸产物的合成。引物的确切长度可以取决于许多因素,例如温度。

[0120] 参考:如本文所用,“参考”描述了相对于其进行比较的标准或对照。例如,在一些实施方案中,将关注的剂、动物、个体、群体、样品、序列或值与参考或对照的剂、动物、个体、群体、样品、序列或值进行比较。在一些实施方案中,测试和/或确定参考或对照与所关注的测试或确定基本上同时进行。在一些实施方案中,参考或对照是历史参考或对照,任选地体现在有形介质中。在一些实施方案中,在靶标的参考水平的上下文中的参考或对照是指正常健康受试者或正常健康受试者群体中的靶标水平。在一些实施方案中,在靶标的参考水平的上下文中的参考或对照是指治疗前受试者体内的靶标水平。通常,如本领域技术人员所理解的,参考或对照是在与评估中的那些相当的状况或情况下确定或表征的。当存在足够的类似性来证明对特定的可能参考或对照的依赖和/或比较时,本领域技术人员将理解。

[0121] 风险:从上下文中可以理解,疾病、病症和/或病状的“风险”是指特定个体发展疾病、病症和/或病状的可能性。在一些实施方案中,风险表示为百分比。在一些实施方案中,风险为0%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、20%、30%、40%、50%、60%、70%、80%、90%至100%。在一些实施方案中,风险表示为相对于与参考样品或参考样品组相关的风险的风险。在一些实施方案中,参考样品或参考样品组具有疾病、病症、病状和/或事件的已知风险。在一些实施方案中,参考样品或参考样品组来自与特定个体相当的个体。在一些实施方案中,相对风险为0、1、2、3、4、5、6、7、8、9、10或更大。

[0122] 样品:如本文所用,术语“样品”通常是指从关注的来源获得或衍生的材料的等分试样。在一些实施方案中,样品是从关注的生物来源(例如,组织或生物体或细胞培养物)获

得或衍生的。在一些实施方案中,关注的来源可以是或包括细胞或生物体,诸如动物或人。在一些实施方案中,关注的来源是或包括生物组织或流体。在一些实施方案中,生物组织或流体可以是或包含羊水、房水、腹水、胆汁、骨髓、血液、母乳、脑脊液、耳垢、乳糜、食糜(chyme)、精液(ejaculate)、内淋巴、渗出物、粪便、胃酸、胃液、淋巴、粘液、心包液、外淋巴、腹膜液、胸膜液、脓液、泪水、唾液、皮脂、精液(semen)、血清、包皮垢、痰、滑液、汗液、眼泪、尿液、阴道分泌物、玻璃体液、呕吐物和/或其组合或一个或多个组分。在一些实施方案中,生物流体可以是或包含细胞内液、细胞外液、囊泡内液(血浆)、间质液、淋巴液和/或跨细胞液。在一些实施方案中,生物组织或样品可以通过例如抽吸、活检(例如,细针或组织活检)、拭子(例如,口腔、鼻、皮肤或阴道拭子)、刮擦、手术、洗涤或灌洗(例如,支气管肺泡、导管、鼻、眼、口腔、子宫、阴道或其他洗涤或灌洗)获得。在一些实施方案中,生物样品是或包含液体活检。在一些实施方案中,生物样品是或包含从个体获得的细胞。在一些实施方案中,样品是通过任何适当的方式直接从关注的来源获得的“原始样品”。在一些实施方案中,从上下文中将清楚,术语“样品”是指通过处理原始样品(例如,通过去除其中的一种或多种组分和/或通过向其中添加一种或多种剂)获得的制剂。例如,样品是通过使用半透膜或基于亲和力的方法(诸如基于抗体的方法)处理以将关注的生物实体与其他非靶实体分离的制剂。例如,此类“经处理的样品”在一些实施方案中可以包含细胞外囊泡,而在一些实施方案中,可以包含从样品中提取的核酸和/或蛋白质等。在一些实施方案中,可以通过使原始样品经受一种或多种技术(诸如核酸的扩增或逆转录、某些组分的分离和/或纯化等)来获得经处理的样品。

[0123] 选择性或特异性:当在本文中用于提及具有活性的剂时,术语“选择性”或“特异性”将被本领域技术人员理解为意指所述剂区分潜在的靶实体、状态或细胞。例如,在一些实施方案中,如果在存在一个或多个竞争性替代靶标的情况下,剂优先与其靶标结合,则称所述剂“特异性”结合其靶标。在许多实施方案中,特异性相互作用取决于靶实体的特定结构特征(例如,表位、裂口、结合位点)的存在。应当理解,特异性不必是绝对的。在一些实施方案中,可以相对于一个或多个其他潜在靶实体(例如,竞争者)的靶结合部分的特异性评估特异性。在一些实施方案中,相对于参考特异性结合部分的特异性评估特异性。在一些实施方案中,相对于参考非特异性结合部分的特异性评估特异性。在一些实施方案中,靶结合部分在与其靶实体结合的条件下不可检测地结合至竞争性替代靶标。在一些实施方案中,与一个或多个竞争性替代靶标相比,靶结合部分以更高的结合速率、更低的解离速率、增加的亲和力、减少的解离和/或增加的稳定性与其靶实体结合。

[0124] 小分子:如本文所用,术语“小分子”意指低分子量有机和/或无机化合物。一般来讲,“小分子”是大小小于约5千道尔顿(kD)的分子。在一些实施方案中,小分子小于约4kD、3kD、约2kD或约1kD。在一些实施方案中,小分子小于约800道尔顿(D)、约600D、约500D、约400D、约300D、约200D或约100D。在一些实施方案中,小分子小于约2000g/mol、小于约1500g/mol、小于约1000g/mol、小于约800g/mol或小于约500g/mol。在一些实施方案中,小分子不是聚合物。在一些实施方案中,小分子不包含聚合部分。在一些实施方案中,小分子不是蛋白质或多肽(例如,不是寡肽或肽)。在一些实施方案中,小分子不是多核苷酸(例如,不是寡核苷酸)。在一些实施方案中,小分子不是多糖。在一些实施方案中,小分子不包含多糖(例如,不是糖蛋白、蛋白聚糖、糖脂等)。在一些实施方案中,小分子不是脂质。在一些实

施方案中,小分子具有生物活性。在一些实施方案中,合适的小分子可以通过诸如以下方法鉴定:筛选化合物大库(Beck-Sickinger&Weber(2001) *Combinational Strategies in Biology and Chemistry* (John Wiley&Sons,Chichester,Sussex));通过核磁共振的构效关系(Shuker等人(1996) "Discovering high-affinity ligands for proteins:SAR by NMR." *Science* 274:1531-1534);编码的自组装化学库(Melkko等人(2004) "Encoded self-assembling chemical libraries." *Nature Biotechnol.* 22:568-574);DNA模板化学(Gartner等人(2004) "DNA-templated organic synthesis and selection of a library of macrocycles." *Science* 305:1601-1605);动态组合化学(Ramstrom&Lehn(2002) "Drug discovery by dynamic combinatorial libraries." *Nature Rev. Drug Discov.* 1:26-36);拴系(Arkin&Wells(2004) "Small-molecule inhibitors of protein-protein interactions:progressing towards the dream." *Nature Rev. Drug Discov.* 3:301-317);和速度筛选(Muckenschnabel等人(2004) "SpeedScreen:label-free liquid chromatography-mass spectrometry-based high-throughput screening for the discovery of orphan protein ligands." *Anal. Biochem.* 324:241-249)。在一些实施方案中,小分子可以具有纳摩尔范围内的靶标的解离常数。

[0125] 特异性结合:如本文所用,术语“特异性结合”是指在发生结合的环境中区分可能的结合配偶体的能力。在存在其他潜在靶标时与一个特定靶标相互作用的靶结合部分称为“特异性结合”至与其相互作用的靶标。在一些实施方案中,通过检测或确定靶结合部分与其配偶体之间的缔合程度来评估特异性结合;在一些实施方案中,通过检测或确定靶结合部分-配偶体复合物的解离程度来评估特异性结合;在一些实施方案中,通过检测或确定靶结合部分竞争其配偶体与另一实体之间的替代相互作用的能力来评估特异性结合。在一些实施方案中,通过在一定浓度范围内进行此类检测或确定来评估特异性结合。

[0126] 癌症分期:如本文所用,术语“癌症分期”是指对癌症(例如,肺癌)进展水平的定性或定量评估。在一些实施方案中,用于确定癌症分期的标准可以包括但不限于以下中的一项或多项:癌症在体内的位置、肿瘤大小、癌症是否已扩散至淋巴结、癌症是否已扩散到体内的一个或多个不同部位等。在一些实施方案中,可以使用AJCC分期系统对癌症进行分期。AJCC分期系统是由美国癌症联合委员会(American Joint Committee on Cancer)开发的用于描述癌症患者中疾病进展的程度的分类系统,所述系统部分地利用了TNM评分系统:肿瘤大小、受影响的淋巴结、转移。在一些实施方案中,可以使用部分涉及TNM评分系统的分类系统对癌症进行分期,根据所述系统,T是指主要肿瘤(通常称为原发性肿瘤)的大小和程度;N是指附近有癌症的淋巴结数量;并且M是指癌症是否已转移。在一些实施方案中,癌症可以被称为0期(存在异常细胞但未扩散到附近组织,也称为原位癌,或CIS;CIS不是癌症,但它可能变成癌症)、I-III期(存在癌症;数字越大,肿瘤越大,并且扩散到附近组织的程度越高)或IV期(癌症已扩散到身体的远处部位)。在一些实施方案中,癌症可以被指定为选自以下组成的组的分期:原位(存在异常细胞但未扩散到附近组织);局部(癌症仅限于它开始的地方,没有扩散的迹象);区域性(癌症已扩散到附近的淋巴结、组织或器官);远处(癌症已经扩散到身体的远处部位);以及未知(没有足够的信息来指出分期)。

[0127] 受试者:如本文所用,术语“受试者”是指从其获得样品的生物体,以例如用于实验、诊断、预防和/或治疗目的。典型的受试者包括动物(例如,哺乳动物,诸如小鼠、大鼠、

兔、非人灵长类动物、家养宠物等)和人。在一些实施方案中,受试者是人受试者,例如,人男性或女性受试者。在一些实施方案中,受试者患有肺癌。在一些实施方案中,受试者易患肺癌。在一些实施方案中,受试者表现出肺癌的一种或多种症状或特征。在一些实施方案中,受试者表现出肺癌的一种或多种非特异性症状。在一些实施方案中,受试者不表现出肺癌的任何症状或特征。在一些实施方案中,受试者是具有肺癌的易感性所特有的一个或多个特征或者患有肺癌的风险的受试者。在一些实施方案中,受试者是患者。在一些实施方案中,受试者是施用和/或已施用诊断和/或治疗的个体。在一些实施方案中,受试者是确定具有一个或多个胸部或肺部肿块的受试者(例如,男性或女性受试者)。在一些实施方案中,受试者是无症状受试者。此类有症状的受试者可以是处于平均群体风险、具有生活史相关风险或具有遗传风险的受试者(例如,男性或女性受试者)。例如,此类无症状受试者可以是具有癌症家族病史的受试者、先前接受过癌症治疗的受试者、在癌症治疗后处于癌症复发风险的受试者、在癌症治疗后有所缓解的受试者和/或先前或定期筛查至少一种癌症生物标志物的存在的受试者。替代地,在一些实施方案中,无症状受试者可以是先前未经癌症筛查的受试者、未诊断为癌症的受试者和/或先前未接受癌症治疗的受试者。在一些实施方案中,适用提供的技术的受试者是基于诸如以下的一个或多个特征选择的个体:年龄、种族、遗传史、病史、个人史(例如,吸烟、酒精、药物、致癌剂、饮食、肥胖、身体活动、光暴露、放射暴露、暴露于诸如病毒的感染性因子和/或职业危害)。

[0128] 患有:“患有”疾病、病症和/或病状的个体已被诊断为患有和/或表现出疾病、病症和/或病状的一种或多种症状。

[0129] 表面多肽或表面蛋白:如在本文中可互换使用的,术语“表面多肽”、“表面蛋白”和“膜结合多肽”是指具有存在于生物实体(例如,细胞、细胞外囊泡等)的膜表面中和/或其上的一个或多个结构域或区域的多肽或蛋白质。在一些实施方案中,表面蛋白可以包含跨越生物实体(例如,细胞、细胞外囊泡等)的质膜和/或与其相关的一个或多个结构域或区域。在一些实施方案中,表面蛋白可以包含跨越生物实体(例如,细胞、细胞外囊泡等)的质膜和/或与其相关并且还突入细胞内和/或囊泡内空间的一个或多个结构域或区域。在一些实施方案中,表面蛋白可以包含例如通过一个或多个非肽键联与生物实体(例如,细胞、细胞外囊泡等)的质膜相关的一个或多个结构域或区域。在一些实施方案中,表面蛋白可以包含锚定到生物实体(例如,细胞、细胞外囊泡等)的质膜任一侧的一个或多个结构域或区域。在一些实施方案中,表面蛋白与细胞外囊泡相关或存在于细胞外囊泡内。在一些实施方案中,表面多肽或膜结合多肽可以与肺癌相关细胞外囊泡(例如,从患有或易患肺癌的受试者的血液或血液来源样品获得或衍生的细胞外囊泡)相关或存在于其内。如本领域技术人员将理解的,检测细胞外囊泡上/内的表面多肽或表面蛋白的至少一部分的存在可以促进肺癌相关细胞外囊泡与来自受试者的生物样品(例如,血液或血液来源样品)的分开和/或分离。在一些实施方案中,检测表面多肽或表面蛋白的存在可以是或包括检测此类表面多肽或表面蛋白的囊泡内部分(例如,囊泡内表位)。在一些实施方案中,检测表面多肽或表面蛋白的存在可以是或包括检测此类表面多肽或表面蛋白的跨膜部分。在一些实施方案中,检测表面多肽或表面蛋白的存在可以是或包括检测此类表面多肽或表面蛋白的囊泡外部分。

[0130] 表面蛋白生物标志物:如本文所用,术语“表面蛋白生物标志物”是指指示生物实体(例如,细胞或细胞外囊泡)的表面蛋白(例如,如本文所述)的状态(例如,存在、水平和/

或活性)的标志物。在一些实施方案中,表面蛋白是指具有位于生物实体(例如,细胞或细胞外囊泡)的膜表面内或表面上的一个或多个结构域或区域的多肽或蛋白质。在一些实施方案中,表面蛋白生物标志物可以是或包含存在于膜内侧(囊泡内)或外侧(囊泡外)的表位。在一些实施方案中,表面蛋白生物标志物与细胞外囊泡相关或存在于细胞外囊泡中。

[0131] 易患:“易患”疾病、病症和/或病状的个体是比公众成员有更高风险发展所述疾病、病症和/或病状的人。在一些实施方案中,易患疾病、病症和/或病状的个体可能未被诊断为患有所述疾病、病症和/或病状。在一些实施方案中,易患疾病、病症和/或病状的个体可以表现出所述疾病、病症和/或病状的症状。在一些实施方案中,易患疾病、病症和/或病状的个体可以不表现出所述疾病、病症和/或病状的症状。在一些实施方案中,易患疾病、病症和/或病状的个体会发展所述疾病、病症和/或病状。在一些实施方案中,易患疾病、病症和/或病状的个体不会发展所述疾病、病症和/或病状。

[0132] 靶结合部分:一般来讲,术语“靶结合部分”和“结合部分”在本文中可互换使用以指结合至关注的靶标(例如,关注的分子靶标,诸如生物标志物或表位)的任何实体或部分。在许多实施方案中,关注的靶结合部分是与其靶标(例如,靶生物标志物)特异性结合的部分,因为它将其靶标与特定相互作用环境中的其他潜在结合配偶体区分开来。一般来讲,靶结合部分可以是或包含任何化学类别的实体或部分(例如,聚合物、非聚合物、小分子、多肽、碳水化合物、脂质、核酸等)。在一些实施方案中,靶结合部分是单个化学实体。在一些实施方案中,靶结合部分是在相关条件下通过非共价相互作用彼此缔合的两个或更多个离散化学实体的复合物。例如,本领域技术人员将理解,在一些实施方案中,靶结合部分可以包含“通用”结合部分(例如,生物素/抗生物素蛋白/链霉抗生物素和/或类别特异性抗体中的一种)和与通用结合部分的配偶体连接的“特异性”结合部分(例如,具有特定分子靶标的抗体或适体)。在一些实施方案中,此类方法可以允许通过不同特异性结合部分与通用结合部分配偶体的连接来模块化组装多个靶结合部分。

[0133] 靶生物标志物特征:如本文所用,术语“靶生物标志物特征”是指(例如,至少2个或更多个,包括例如至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少14个、至少15个、至少16个、至少17个、至少18个、至少19个、至少20个、至少25个、至少30个或更多个)生物标志物的组合,所述组合与特定生物事件或关注的状态相关联,使得本领域技术人员将理解,其可以适当地被认为是所述事件或状态的“特征”。仅举几个实例,在一些实施方案中,靶生物标志物特征可以与特定疾病或疾病状态相关联,和/或与特定疾病、病症或病状可能发展、发生或复发的可能性相关联。在一些实施方案中,靶生物标志物特征可以与特定疾病或治疗结果或其可能性相关联。在一些实施方案中,靶生物标志物特征可以与特定癌症和/或其分期相关联。在一些实施方案中,靶生物标志物特征可以与肺癌和/或其分期和/或其亚型相关联。在一些实施方案中,靶生物标志物特征包含(例如,至少2个或更多个,包括例如至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少14个、至少15个、至少16个、至少17个、至少18个、至少19个、至少20个、至少25个、至少30个或更多个)生物标志物的组合,它们一起对肺癌或其亚型和/或其疾病分期具有特异性,尽管此类组合中的一个或多个生物标志物可以针对对肺癌不具有特异性的靶标(例如,表面蛋白生物标志物、囊泡内蛋白生物标志物和/或囊泡内RNA)。例如,在一些实施方案中,靶生物标志物特

征可以包含对肺癌或其分期和/或亚型具有特异性的至少一种生物标志物(即,肺癌特异性靶标),并且还可以包含不一定或完全对肺癌具有特异性的生物标志物(例如,也可以在一些或所有生物实体(例如像细胞、细胞外囊泡等)上发现,所述生物实体不是癌性的、不是有关癌症的和/或不属于关注的特定分期和/或亚型)。也就是说,如阅读本说明书的本领域技术人员将理解的,只要在靶生物标志物特征中使用的生物标志物的组合是或包含多个生物标志物,所述生物标志物一起对关注的有关靶生物实体(例如,关注的肺癌细胞或肺癌细胞分泌的细胞外囊泡)具有特异性(即,充分区分用于检测的有关靶生物实体(例如,关注的肺癌细胞或肺癌细胞分泌的细胞外囊泡)与用于检测的其他不受关注的生物实体),根据本公开的某些实施方案,生物标志物的此类组合是有用的靶生物标志物特征。

[0134] 治疗剂:如在本文中可互换使用的,短语“治疗剂”或“疗法”是指当施用于受试者或患者时具有治疗效果和/或引发所需生物学和/或药理学效果的剂或干预。在一些实施方案中,治疗剂或疗法是可以用于减轻、改善、缓解、抑制、预防、延迟疾病、病症和/或病状的一种或多种症状或特征的发作、降低其严重程度和/或降低其发病率的任何物质。在一些实施方案中,治疗剂或疗法是可以进行以减轻、缓解、抑制、呈现、延迟疾病、病症和/或病状的一种或多种症状或特征的发作、降低其严重程度和/或降低其发病率的医学干预(例如,手术、放射、光照疗法)。

[0135] 阈值水平(例如,截止值):如本文所用,术语“阈值水平”是指用作获得关于测量结果(例如,在测定中获得的测量结果)的信息和/或对其进行分类的参考的水平。例如,在一些实施方案中,阈值水平(例如,截止值)意指在定义群体的两个子集(例如,正常和/或非肺癌与肺癌)之间的分界线的测定中测量的值。因此,等于或高于阈值水平的值定义了群体的一个子集,并且低于阈值水平的值定义了群体的另一个子集。阈值水平可以基于一个或多个对照样品或整个对照样品群体来确定。可以在进行关注的测量之前、同时或之后确定阈值水平。在一些实施方案中,阈值水平可以是值的范围。

[0136] 治疗:如本文所用,术语“治疗(treat、treatment或treating)”是指用于部分或完全减轻、改善、缓解、抑制、预防、延迟疾病、病症和/或病状的一种或多种症状或特征的发作、降低其严重程度和/或降低其发病率的任何方法。可以向未表现出疾病、病症和/或病状迹象的受试者施用治疗。在一些实施方案中,可以向仅表现出疾病、病症和/或病状的早期迹象的受试者施用治疗,例如为了降低发展与疾病、病症和/或病状相关的病理的风险的目的。在一些实施方案中,可以向处于疾病、病症和/或病状的晚期的受试者施用治疗。

[0137] 标准技术可以用于重组DNA、寡核苷酸合成以及组织培养和转化(例如,电穿孔、脂质转染)。酶促反应和纯化技术可以根据制造商的说明书或如本领域中通常完成的或如本文所述来进行。前述技术和程序一般可以根据本领域众所周知的常规方法进行,并且如在本说明书通篇引用和讨论的各种一般性和更具体的参考文献中所述。参见例如Sambrook等人,Molecular Cloning:A Laboratory Manual(第2版,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.(1989)),其出于本文所述目的以引用的方式并入本文。

具体实施方式

[0138] 2016年美国估计有148,869人死于肺癌(美国癌症数据工作组(U.S.Cancer statistics working group),2019;其出于本文所述目的以引用的方式并入本文)。这些死

亡中的大部分归因于晚期诊断;2001-2007年美国肺癌的5年总生存率为15.6%。诊断时有局部疾病的患者具有52%的5年生存率;然而,大多数患者在已经形成远处转移时才得到初步诊断,并且这些患者具有3.6%的5年生存率(Cruz等人,2011;其出于本文所述目的以引用的方式并入本文)。

[0139] 遗憾的是,没有廉价的、可广泛使用的针对平均风险人群的推荐肺癌筛查测试。虽然目前许多具有遗传或生活史相关风险的个体通过低剂量CT扫描进行筛查,但这些测试是次优的筛查,因为它们相对昂贵并且可及性有限。例如,前列腺癌、肺癌、结肠直肠癌和卵巢癌筛查随机试验发现,相对高性价比和可广泛使用的胸部X射线成像和痰液测试在改变肺癌死亡率方面是无效的,而低剂量CT扫描确实成功地降低了死亡率,但它通常会增加不必要的手术次数,并且可以被认为是昂贵的,而且可及性有限。

[0140] 本公开尤其用某些现有技术(包括例如检测和诊断肺癌的某些常规方法)鉴定问题的根源。例如,本公开认识到许多常规的诊断测定,例如X射线成像、痰液测试、低剂量CT扫描和/或基于游离核酸、血清蛋白(例如,CEA、CYFRA 21-1、NSE、ProGRP和/或SCCA)的分子测试,和/或细胞外囊泡的整体分析,可能耗时、昂贵和/或缺乏足以提供可靠和全面的诊断评估的灵敏度和/或特异性。在一些实施方案中,本公开提供了解决此类问题的技术(包括系统、组合物和方法),尤其通过基于生物信息学分析鉴定预测对肺癌表现出高灵敏度和特异性的生物标志物组合。在一些实施方案中,本公开提供通过检测个体细胞外囊泡中肺癌的靶生物标志物特征的共定位(例如,通过生物信息学分析鉴定)来解决此类问题的技术(包括系统、组合物和方法),所述细胞外囊泡包含至少一个细胞外囊泡相关膜结合多肽和选自以下组成的组的至少一个靶生物标志物:存在于与肺癌相关的细胞外囊泡中的表面蛋白生物标志物、内部蛋白生物标志物和RNA生物标志物。在一些实施方案中,本公开尤其提供了通过使用基于个体细胞外囊泡中靶生物标志物特征的相互作用和/或共定位的靶实体检测方法(由申请人开发并描述于均在2020年2月28日提交且名为“Systems, Compositions, and Methods for Target Entity Detection”的美国申请号16/805,637和国际申请PCT/US2020/020529中)检测肺癌的此类生物标志物特征来解决此类问题的技术(包括系统、组合物和方法)。前述各公开内容全文以引用的方式并入本文。

[0141] 本公开尤其提供了用于实现有效肺癌筛查,(例如,用于肺癌的早期检测)的见解和技术。在一些实施方案中,本公开提供了用于可能正在经历与肺癌相关的一种或多种症状的受试者中的肺癌早期检测的技术。在一些实施方案中,本公开提供了用于处于肺癌遗传风险的受试者中的肺癌早期检测的技术。在一些实施方案中,本公开提供了用于可能处于遗传风险和/或正在经历与肺癌相关的一种或多种症状的受试者中的肺癌早期检测的技术。在一些实施方案中,本公开提供了用于可能具有生活史风险因素(例如,但不限于吸烟)的受试者中的肺癌早期检测的技术。在一些实施方案中,本公开提供了用于筛查个体(例如,具有某些风险(例如,遗传风险、生活史相关风险或平均风险)的个体)早期非小细胞肺癌(例如像肺腺癌(LUAD)和肺鳞状细胞癌(LUSC))的技术。非小细胞肺癌是肺癌最常见的亚型,非小细胞肺癌病例中的54%在晚期被检测到(SEER癌症统计回顾1975-2017)。在一些实施方案中,提供的技术对于检测早期肺癌是有效的。在一些实施方案中,提供的技术在应用于包含有可能与肺癌相关的一种或多种症状的个体或由所述个体组成的群体时是有效的。在一些实施方案中,提供的技术即使在应用于包括无症状或有症状个体或由所述个体

组成的群体时也是有效的(例如,由于足够高的灵敏度和/或低的假阳性和/或假阴性结果率)。在一些实施方案中,提供的技术在应用于包含没有发展肺癌的遗传风险和/或生活史相关风险的个体(例如,无症状或有症状个体)或由所述个体组成的群体时是有效的。在一些实施方案中,提供的技术在应用于包含具有发展肺癌的遗传风险和/或生活史相关风险的个体(例如,无症状或有症状个体)或由所述个体组成的群体时是有效的。在一些实施方案中,提供的技术在应用于包括易患肺癌的个体(例如,具有已知遗传、环境或经验风险等的个体)或由所述个体组成的群体时是有效的。在一些实施方案中,提供的技术可以是或包括一种或多种组合物(例如,分子复合物、系统、集合、组合、试剂盒等)和/或方法(例如,制造、使用、评估等),如对于阅读本文提供的公开的本领域技术人员来说将是清楚的。

[0142] 在一些实施方案中,提供的技术实现对肺癌的一个或多个特征(例如,发病、进展、对治疗的反应性、复发等)的检测(例如,早期检测,例如,在一个或多个无症状个体和/或一个或多个群体中),其中灵敏度和/或特异性(例如,假阳性和/或假阴性结果的比率)适用于允许将提供的技术有用地应用于单次和/或定期(例如,周期性)评估。在一些实施方案中,提供的技术可用于与个体的常规医疗检查结合,所述常规医疗检查诸如但不限于:体检、全科医生就诊、胆固醇/血脂测试、糖尿病(2型)筛查、结肠镜检查、血压筛查、甲状腺功能测试、前列腺癌筛查、乳腺图像、HPV/Pap涂片和/或预防接种。在一些实施方案中,提供的技术可用于与一种或多种治疗方案结合;在一些实施方案中,提供的技术可以改进此类一种或多种治疗方案的一个或多个特征(例如,根据接受的参数的成功率)。

[0143] 在一些实施方案中,本公开尤其提供了以下见解:对无症状个体进行筛查(例如,在出现一种或多种症状之前或以其他方式在没有出现一种或多种症状的情况下进行定期筛查)可能是有益的,并且甚至对于肺癌的有效管理(例如,成功治疗)是重要的。在一些实施方案中,本公开提供了肺癌筛查系统,在一些实施方案中,其可以被实施以检测无症状个体(例如,没有肺癌遗传风险和/或生活史相关风险)中的肺癌,包括早期癌症。在一些实施方案中,提供的技术被实施以实现无症状个体(例如,有或没有肺癌遗传风险和/或生活史相关风险)的定期筛查。在一些实施方案中,提供的技术被实施以实现对有症状个体(例如,有或没有肺癌遗传风险和/或生活史相关风险)的定期筛查。本公开提供了例如组合物(例如,试剂、试剂盒、组分等)以及提供和/或使用它们的方法,包括涉及对一个或多个个体(例如,无症状个体)进行定期测试的策略。本公开定义了此类系统的有用性,并提供了用于实施它们的组合物和方法。

[0144] I. 肺癌检测

[0145] 目前没有任何一种肺癌筛查测试被CDC推荐用于筛查平均风险的无症状个体,而在美国,肺癌的年龄标化发病率为62/10万男性和女性/年。每年死于肺癌的美国人几乎和死于前列腺癌、乳腺癌和结肠癌的总和一样多。仅在2010年,美国中估计有约240,000肺癌新发病例和约161,000肺癌死亡病例(Cruz等人,2011;其出于本文所述目的以引用的方式并入本文)。尽管自约1985年以来,美国肺癌死亡总数已在下降,但全球肺癌率却在稳定上升,并且从1985年至2010年已经上升了约51%(Cruz等人,2011;其出于本文所述目的以引用的方式并入本文)。在全球范围内,肺癌是癌症新发病例的最大贡献者,并且在2010年全球大约有1,350,000例肺癌新发病例(约占所有新发癌症病例的12.4%)和约1,180,000例死亡病例(约占癌症相关死亡总数的17.6%)。这些新发肺癌病例中大约有一半发生在发展

中国。而欧洲、中国和发展中国家的5年肺癌生存率估计仅为8.9% (Cruz等人, 2011; 其出于本文所述目的以引用的方式并入本文)。

[0146] 2013-2017年的监测、流行病学和最终结果 (SEER) 数据广泛报道了过去45年美国肺癌的患病率和流行病学。SEER报告了肺癌和支气管癌的中位诊断年龄为约71岁。肺癌起源于呼吸道上皮细胞, 并且可以分为两大类。小细胞肺癌 (SCLC) 恶性程度高, 并且来源于表现出神经内分泌特征的细胞, SCLC 占有肺癌病例的约12% (图3)。非小细胞肺癌 (NSCLC) 占病例的剩余85%, 并且分为三种病理亚型: 腺癌、鳞状细胞癌和大细胞癌。腺癌本身占美国所有新发肺癌病例的约50% (图3), 鳞状细胞癌占约23%, 并且大细胞癌和其他肺癌和支气管癌占剩余的约15% (图3)。

[0147] 所有侵袭性非小细胞肺癌和侵袭性非小细胞支气管癌患者的5年生存率约为25%。初步诊断时患有局部肺癌的患者具有约63%的5年生存率, 初步诊断时患有区域肺癌转移的患者具有约35%的5年生存率, 而患有远处肺癌转移的患者具有仅约7%的5年生存率 (SEER 1975-2017综述, 表15.12)。这些数据表明早期肺癌的诊断很重要, 因为它可以提高肺癌患者的生存率。

[0148] 肺癌的某些风险因素包括年龄和吸烟史。1964年美国公共卫生部部长的开创性报告指出: (1) 吸烟与男性年龄别死亡率70%的增加和女性死亡率较少的增加相关; (2) 吸烟与男性肺癌的发生有因果关系, 其影响程度远大于导致肺癌发生的所有其他因素, 并且其风险随着吸烟时间的延长和每日吸烟数量的增加而直接增加; (3), 吸烟在一般群体的肺癌原因中被认为比职业暴露更为重要; (4) 吸烟被报道为美国慢性支气管炎的重要病因; (5) 男性吸烟者具有高于男性非吸烟者的冠状动脉疾病死亡率。

[0149] 国际癌症研究机构 (IARC) 已在烟草烟气中鉴定出至少50种已知致癌物。此类致癌物的实例包括但不限于烟草加工和吸烟过程中尼古丁亚硝化形成的烟草特有的N-亚硝胺 (TSNA)。已知化学物4-(甲基亚硝胺基)-1(3-吡啶基)-1-丁酮 (NNK) 在实验动物中诱导肺腺癌。已知NNK与DNA结合并产生DNA加合物, 导致DNA损伤。不能修复这种损伤可能导致永久性突变。NNK与导致K-ras癌基因激活的DNA突变相关, 在24%的人肺腺癌中检测到NNK。

[0150] 据估计, 每九个吸烟者中就有一个最终发展为肺癌。估计长期吸烟者的肺癌相对风险比终生非吸烟者高10倍至30倍。在美国, 超过80%的肺癌发生在有烟草暴露的人群中, 而在全球范围内, 15%的男性肺癌和多至53%的女性肺癌不归因于吸烟, 从未吸烟者占全世界所有肺癌病例的约25%。

[0151] 如CDC定义的高风险个体和/或群体为55-77岁、有>30包年的吸烟史、现在吸烟者, 或在过去15年内戒烟。对于这些个体, 低剂量CT扫描是目前推荐的肺癌筛查工具。然而, 低剂量CT在高风险的 (例如, 如CDC指南定义的患者) 群体中可以被认为是相对昂贵的、有限可及的, 以及具有不合理的高水平假阳性 (例如, 所有阳性测试中假阳性的比例可能高达97.5%; Raghu等人, 2020和Kinsinger等人, 2017, 其均出于本文所述目的以引用的方式并入本文)。本研究尤其提供了具有足够高的特异性和/或灵敏度的成本效率高的筛查测定。

[0152] 在某些实施方案中, 本公开尤其提供了以下见解: 需要开发用于筛查具有肺癌遗传风险和/或生活史相关风险的受试者和/或可能经历一种或多种肺癌相关症状的受试者的肺癌液体活检测定。在某些实施方案中, 本公开提供了以下见解: 有必要开发用于筛查有症状或无症状患者的肺癌液体活检测定, 例如, 在其他筛查方法 (例如, 用于肺癌检测的成

像方法(例如,MRI、CT扫描等))之前。

[0153] 在一些实施方案中,本公开提供了用于在有遗传风险的个体或具有生活史相关风险的个体中有效筛查肺癌的技术。在一些实施方案中,本公开提供了用于在平均风险个体中有效筛查肺癌的技术。在一些实施方案中,本公开提供了用于在具有一种或多种肺癌相关症状的个体中有效筛查肺癌的技术。在一些实施方案中,本公开提供了用于在无症状个体中有效筛查肺癌的技术。尽管肺癌是所有癌症中男性和女性的最大杀手,但目前尚无推荐的用于无症状和/或平均风险个体(例如,55岁以下的个体,或55岁以上无吸烟史或已戒烟15年以上的个体)的肺癌筛查工具。这部分是由于现有肺癌筛查技术的成本、有限的可用性、潜在的副作用和/或不佳的性能(例如,高假阳性率或无效性)。鉴于平均风险个体的肺癌发病率,测试特异性不足(<99.5%)可能导致假阳性结果数量超过真阳性多于一个数量级。这给医疗保健系统和接受筛查的个体带来了显著负担,因为假阳性结果会导致额外的测试、不必要的手术和情绪/身体困扰(Wu等人,2016)。

[0154] 在一些实施方案中,本公开提供了以下见解:特别有用的肺癌筛查测试将具有以下特征:(1)最小化假阳性数量的超高特异性(>99.5%);(2)对I期和II期肺癌的高灵敏度(>40%)(即当预后最有利时)。

[0155] 例如,在一些实施方案中,特别有用的肺癌筛查测试的特征可以是>98%的特异性和>50%的灵敏度,例如对于I期和II期肺癌。在一些实施方案中,特别有用的肺癌筛查测试的特征可以是>98%的特异性和>60%的灵敏度,例如对于I期和II期肺癌。在一些实施方案中,特别有用的肺癌筛查测试的特征可以是>98%的特异性和>70%的灵敏度,例如对于I期和II期肺癌。在一些实施方案中,特别有用的肺癌筛查测试的特征可以是>99.5%的特异性和>65%的灵敏度,例如对于I期和II期肺癌。在一些实施方案中,特别有用的肺癌筛查测试的特征可以是>99.5%的特异性和>60%的灵敏度,例如对于I期和II期肺癌。在一些实施方案中,特别有用的肺癌筛查测试的特征可以是99%或更高的特异性和>10%或更高(包括例如,>15%、>20%、>25%)的灵敏度。在一些实施方案中,特别有用的肺癌筛查测试的特征可以是99%或更高的特异性和50%或更高的灵敏度。

[0156] 在一些实施方案中,本公开提供了以下见解:与通过一组生物标志物组合来实现的肺癌筛查测试相比,涉及多于一组生物标志物组合(例如,如本文所述的至少两种正交生物标志物组合)的肺癌筛查测试可以增加此类测定的灵敏度。例如,在一些实施方案中,涉及至少两种正交生物标志物组合的肺癌筛查测试可以实现至少98%的特异性和至少50%的灵敏度。在一些实施方案中,涉及至少两种正交生物标志物组合的肺癌筛查测试可以实现至少98%的特异性和至少60%的灵敏度。在一些实施方案中,涉及至少两种正交生物标志物组合的肺癌筛查测试可以实现99%的特异性和50%或更高的灵敏度。

[0157] 在一些实施方案中,本公开提供了以下见解:特别有用的肺癌筛查测试的特征可以是以经济合理的成本获得可接受的阳性预测值(PPV)。PPV是患者在测试呈阳性后患有疾病的可能性,并且PPV受灵敏度、特异性和/或疾病患病率的影响。肺癌筛查所需最低PPV的临床医师共识为10%。PPV为10%时,对每一个真阳性将有九个假阳性(Lung Cancer Screening:Recommendation Statement.,Am Fam Physician.2005Mar15;71(6):1165-1168)。这些假阳性给医疗保健系统和接受筛查的受试者均带来了显著负担,因为这些假阳性导致额外的测试、不必要的手术以及情绪和身体困扰。在一些实施方案中,本文所述的测

定(具有对于有肺癌遗传风险的受试者至少为98%的特异性截止值,或者具有对于经历肺癌相关的一种或多种症状的受试者至少为99.5%的特异性截止值)对于实现大于10%或更高(包括,例如大于15%、大于20%或大于25%或更高)的PPV的早期肺癌检测特别有用。

[0158] 在一些实施方案中,本文所述的测定可以用于实现大于2%或更高(包括,例如大于3%、大于4%、大于5%、大于6%、大于7%、大于8%、大于9%、大于10%、大于15%、大于20%或大于25%或更高)的PPV的早期肺癌检测。在一些此类实施方案中,本文所述的测定可以实现至少95%或更高(例如,对于有肺癌风险的受试者至少为98%的特异性截止值,或具有对于经历一种或多种肺癌相关的症状的受试者至少99.5%的特异性截止值)的特异性截止值。

[0159] 已经针对肺癌液体活检测定研究了几种不同的生物标志物类别,包括循环肿瘤DNA(ctDNA)、循环肿瘤细胞(CTC)、批量蛋白质和细胞外囊泡(EV)。相对于ctDNA和CTC,EV由于其在血流中的丰度和稳定性而特别有希望,这表明EV对早期癌症的灵敏度有所提高。此外,EV含有源自同一细胞的货物(即,蛋白质、RNA、代谢物),从而提供优于批量蛋白质测量的特异性。虽然已经研究了EV的诊断效用,但大部分工作都与批量EV测量或低通量单EV分析有关。

[0160] II. 用于检测肺癌的提供的生物标志物和/或靶生物标志物特征

[0161] 本公开尤其提供了肺癌的各种靶生物标志物或其组合(例如,靶生物标志物特征)。此类被预测对肺癌表现出高灵敏度和特异性的靶生物标志物特征通过多管齐下的生物信息学分析和生物学方法被发现,例如,在一些实施方案中,所述信息学分析和生物学方法涉及不同数据集的计算分析,例如在一些实施方案中,所述数据集包括测序数据、表达数据、质谱、组织学、翻译后修饰数据和/或通过机器学习和/或计算建模的体外和/或体内实验数据中的一个或多个。

[0162] 在一些实施方案中,肺癌的靶生物标志物特征包括至少一个或多个(例如,1、2、3、4、5、6、7、8个或更多个)胞外囊泡相关膜结合多肽(例如,存在于与肺癌相关的细胞外囊泡中的表面多肽)和选自由以下组成的组的至少一个或多个(例如,1、2、3、4、5、6、7、8个或更多个)靶生物标志物:一个或多个表面蛋白生物标志物、一个或多个囊泡内蛋白生物标志物和一个或多个囊泡内RNA生物标志物,使得一个或多个此类细胞外囊泡相关膜结合多肽和一个或多个此类靶生物标志物的组合呈现肺癌的靶生物标志物特征,所述组合提供(a)最小化假阳性数量的高特异性(例如,大于98%或更高,诸如大于99%,或大于99.5%),和(b)对预后最有利的I期和II期肺癌的高灵敏度(例如,大于40%、大于50%、大于60%、大于70%、大于80%)。

[0163] 在一些实施方案中,本公开认识到,在某些实施方案中,对具有不同肺风险水平的受试者的灵敏度和特异性率可以根据主治医师的风险承受能力和/或关注的医联体提出的指导而变化。在某些实施方案中,有肺癌风险的受试者可以在99.5%的特异性率和70%的灵敏度或在98%的特异性率和80%的灵敏度下得到最佳对待。在某些实施方案中,有生活史风险的受试者可以在99.5%的特异性率和70%的灵敏度或在98%的特异性率和80%的灵敏度下得到最佳对待。在一些实施方案中,本文所述的用于检测有风险受试者的肺癌的测定可以具有小于80%灵敏度的设定灵敏度率,包括例如小于70%、小于60%、小于50%或更低的灵敏度率。在某些实施方案中,无症状受试者可以在99.5%的特异性率和70%的灵

敏度或在98%的特异性率和80%的灵敏度下得到最佳对待。在一些实施方案中,本文所述的用于检测无症状受试者的肺癌的测定可以具有小于80%灵敏度的设定灵敏度率,包括例如小于70%、小于60%、小于50%或更低的灵敏度率。在一些实施方案中,本文所述的用于检测有症状受试者的肺癌的技术和/或测定可以具有比检测无症状受试者的肺癌更小的灵敏度和/或特异性要求。在一些实施方案中,本文所述的用于检测有症状受试者的肺癌的测定可以具有小于99.5%特异性的设定特异性率,包括例如小于99%、小于95%、小于90%或小于85%的特异性率。在一些实施方案中,本文所述的用于检测有症状受试者的肺癌的测定可以具有小于80%灵敏度的设定灵敏度率,包括例如小于70%或小于60%的灵敏度率。

[0164] 本公开尤其观察到用于筛查高风险吸烟者的金标准是胸部CT,其在全国肺筛查试验研究中具有在此类高风险群体中3.8%的报告阳性预测值(National Lung Screening Trial Research Team(2013)“Results of initial low-dose computed tomographic screening for lung cancer.New England Journal of Medicine,”368(21):1980-1991)。在一些实施方案中,本公开尤其认识到提供3.8%或更高阳性预测值(PPV)的肺癌的生物标志物特征对筛查有风险个体的肺癌特别有用。在一些实施方案中,肺癌的靶生物标志物特征包含至少一个细胞外囊泡相关膜结合多肽(例如,存在于与肺癌相关的细胞外囊泡中的表面多肽)和选自由以下组成的组的至少一个靶生物标志物:一个或多个表面蛋白生物标志物、一个或多个囊泡内蛋白生物标志物和一个或多个囊泡内RNA生物标志物,使得一个或多个此类细胞外囊泡相关膜结合多肽和一个或多个此类靶生物标志物的组合呈现肺癌的靶生物标志物特征,所述组合在高风险群体中提供至少3.8%或更高,包括例如至少4%、至少5%、至少6%、至少7%、至少8%、至少9%、至少10%或更高、至少15%或更高、至少20%或更高、至少25%或更高和/或至少30%或更高的阳性预测值(PPV)。

[0165] 在一些实施方案中,肺癌的靶生物标志物特征包含至少一个细胞外囊泡相关膜结合多肽(例如,存在于与肺癌相关的细胞外囊泡中的表面多肽)和选自由以下组成的组的至少一个靶生物标志物:一个或多个表面蛋白生物标志物、一个或多个囊泡内蛋白生物标志物和一个或多个囊泡内RNA生物标志物,使得一个或多个此类细胞外囊泡相关膜结合多肽和一个或多个此类靶生物标志物的组合对肺癌具有特异性。在一些实施方案中,肺癌的靶生物标志物特征是或包含:CD166抗原(ALCAM)多肽、UDP-GlcNAc:βGalβ-1,3-N-乙酰氨基葡萄糖转移酶3(B3GNT3)基因编码的N-乙酰乳糖胺β-1,3-N-乙酰氨基葡萄糖转移酶3多肽、含有CUB结构域的蛋白1(CDCP1)基因编码的含有CUB结构域的蛋白1多肽、钙粘蛋白-1(CDH1)多肽、钙粘蛋白3(CDH3)基因编码的钙粘蛋白3多肽、补体衰变加速因子(CD55)多肽、程序性细胞死亡1配体1(CD274;也称为PD-L1)多肽、癌胚抗原细胞粘附分子5(CEACAM5)基因编码的癌胚抗原细胞粘附分子5多肽、癌胚抗原细胞粘附分子6(CEACAM6)基因编码的癌胚抗原细胞粘附分子6多肽、claudin 3(CLDN3)基因编码的claudin 3多肽、(CLDN4)基因编码的claudin 4多肽、桥粒芯蛋白2(DSG2)基因编码的桥粒芯蛋白2多肽、表皮生长因子受体(EGFR)多肽、上皮细胞粘附分子(EPCAM)基因编码的上皮细胞粘附分子多肽、叶酸受体α(FOLR1)基因编码的叶酸受体α多肽、缝隙连接蛋白β1(GJB1)基因编码的缝隙连接β-1蛋白多肽、缝隙连接蛋白β2(GJB2)基因编码的缝隙连接β-2蛋白多肽、肝细胞生长因子受体(MET)多肽、胰岛素样生长因子1受体(IG1FR)多肽、层粘连蛋白亚基β3(LAMB3)基因编码的层粘连蛋白亚基β-3多肽、间皮素(MSLN)多肽、粘蛋白-1(MUC1)多肽、磷脂酰肌醇聚糖锚定

生物合成类T (PIGT) 基因编码的GPI转酰胺酶组分PIG-T多肽、足糖萼蛋白样 (podocalyxin-like) 2 (PODXL2) 基因编码的足糖萼蛋白样蛋白2多肽、ROS原癌基因1编码的原癌基因酪氨酸蛋白激酶ROS多肽、受体酪氨酸激酶 (ROS1) 基因、粘蛋白聚糖1 (SDC1) 基因编码的粘蛋白聚糖1多肽、溶质运载转运体家族34 (磷酸钠) 成员2 (SLC34A2) 基因编码的钠依赖型磷转运蛋白2B多肽、鞘磷脂磷酸二酯酶样3B (SMPDL3B) 基因编码的酸性鞘磷脂酶样磷酸二酯酶3b多肽、ST14跨膜丝氨酸蛋白酶蛋白裂解酶 (ST14) 基因编码的瘤变抑制因子14蛋白多肽、肿瘤相关钙信号转导因子2 (TACSTD2) 多肽、跨膜丝氨酸蛋白酶2 (TMPRSS2) 基因编码的跨膜丝氨酸蛋白酶2多肽、肿瘤坏死因子受体超家族成员10B (TNFRSF10B) 多肽、四次跨膜蛋白8 (TSPAN8) 基因编码的四次跨膜蛋白-8多肽、sTn糖基化多肽、Tn糖基化多肽、T糖基化多肽或其组合。

[0166] 在一些实施方案中, 肺癌的靶生物标志物特征是或包含选自由以下组成的组的表面蛋白生物标志物: 磷脂转运ATP酶ABCA3 (ABCA3) 多肽、多药耐药相关蛋白1 (ABCC1) 多肽、ATP结合盒亚家族C成员3 (ABCC3) 多肽、高尔基体常驻蛋白GCP60 (ACBD3) 多肽、长链脂肪酸CoA连接酶5 (ACSL5) 多肽、晚期糖基化终末产物特异性受体 (AGER) 多肽、CD166抗原 (ALCAM) 多肽、AP-1复合体亚基 μ -2 (AP1M2) 多肽、 γ -分泌酶亚基-1A (APH1A) 多肽、MICOS复合体亚基MIC26 (APOO) 多肽、磷脂转运ATP酶1H (ATP11A) 多肽、磷脂转运ATP酶1F (ATP11B) 多肽、钠/钾转运ATP酶亚基 β -1 (ATP1B1) 多肽、肾素受体 (ATP6AP2) 多肽、 β -1,4-半乳糖转移酶4 (B4GALT4) 多肽、B细胞受体相关蛋白31 (BCAP31) 多肽、含B框和SPRY结构域蛋白 (BSPRY) 多肽、CD109抗原 (CD109) 多肽、补体衰退加速因子 (CD55) 多肽、CD9抗原 (CD9) 多肽、细胞分裂控制蛋白42同源物 (CDC42) 多肽、钙粘蛋白-1 (CDH1) 多肽、钙粘蛋白-3 (CDH3) 多肽、苏氨酸酰基氨基甲酰腺苷tRNA甲硫基转移酶 (CDKAL1) 多肽、癌胚抗原相关细胞粘附分子5 (CEACAM5) 多肽、癌胚抗原相关细胞粘附分子6 (CEACAM6) 多肽、钙粘蛋白EGF LAG七经G型受体1 (CELSR1) 多肽、蛋白CIP2A (CIP2A) 多肽、CDGSH含铁硫结构域蛋白2 (CISD2) 多肽、细胞骨架相关蛋白4 (CKAP4) 多肽、钙激活氯离子通道调节剂2 (CLCA2) 多肽、Claudin-1 (CLDN1) 多肽、胞内氯离子通道蛋白6 (CLIC6) 多肽、唇腭裂跨膜蛋白1样蛋白 (CLPTM1L) 多肽、钙结合蛋白 (Calsyntenin) -1 (CLSTN1) 多肽、接触蛋白-1 (CNTN1) 多肽、羧肽酶D (CPD) 多肽、细胞色素P450 2S1 (CYP2S1) 多肽、细胞色素P450 4F11 (CYP4F11) 多肽、细胞色素P450 4F3 (CYP4F3) 多肽、可能的C-甘露糖基转移酶DPY19L1 (DPY19L1) 多肽、桥粒芯胶蛋白-2 (DSC2) 多肽、桥粒芯胶蛋白-3 (DSC3) 多肽、桥粒芯蛋白-2 (DSG2) 多肽、桥粒芯蛋白-3 (DSG3) 多肽、表皮生长因子受体 (EGFR) 多肽、上皮细胞粘附分子 (EPCAM) 多肽、肝配蛋白B型受体3 (EPHB3) 多肽、原钙粘蛋白Fat 2 (FAT2) 多肽、含F框/SPRY结构域蛋白1 (FBX045) 多肽、Fermitin家族同源物1 (FERMT1) 多肽、叶酸受体 α (FOLR1) 多肽、卷曲基因-6 (FZD6) 多肽、多肽N-乙酰氨基半乳糖转移酶1 (GALNT1) 多肽、多肽N-乙酰氨基半乳糖转移酶3 (GALNT3) 多肽、多肽N-乙酰氨基半乳糖转移酶5 (GALNT5) 多肽、多肽N-乙酰氨基半乳糖转移酶6 (GALNT6) 多肽、维生素K依赖性 γ -羧化酶 (GGCX) 多肽、高尔基体跨膜糖蛋白1 (GOLM1) 多肽、高尔基体磷酸化蛋白3样 (GOLPH3L) 多肽、粒状头样蛋白2同源物 (GRHL2) 多肽、超长链 (3R) -3-羟基酰基-CoA脱水酶3 (HACD3) 多肽、立即早期反应3相互作用蛋白1 (IER3IP1) 多肽、免疫球蛋白超家族成员3 (IGSF3) 多肽、白细胞介素-1受体辅助蛋白 (IL1RAP) 多肽、整合素 α -2 (ITGA2) 多肽、整合素 β -6 (ITGB6) 多肽、杀伤细胞凝集素样受体亚家族G成员2 (KLRG2) 多肽、输入蛋白亚基 α -1

(KPNA2)多肽、角质形成细胞相关蛋白3(KRTCAP3)多肽、Ladinin-1(LAD1)多肽、层粘连蛋白亚基 β -3(LAMB3)多肽、层粘连蛋白亚基 γ -2(LAMC2)多肽、溶酶体相关膜糖蛋白3(LAMP3)多肽、Ragulator复合体蛋白LAMTOR2(LAMTOR2)多肽、溶血磷脂酰基转移酶1(LCLAT1)多肽、溶血磷脂酰胆碱酰基转移酶1(LPCAT1)多肽、脂解刺激脂蛋白受体(LSR)多肽、镁转运蛋白1(MAGT1)多肽、MARCKS相关蛋白(MARCKSL1)多肽、肝细胞生长因子受体(MET)多肽、 α -1,3-甘露糖基-糖蛋白-2- β -N-乙酰氨基葡萄糖转移酶(MGAT1)多肽、间皮素(MSLN)多肽、粘蛋白-1(MUC1)多肽、粘蛋白-4(MUC4)多肽、呆蛋白(Nicastrin)(NCSTN)多肽、结合素-1(NECTIN1)多肽、结合素-4(NECTIN4)多肽、GTP酶NRas(NRAS)多肽、5'-核苷酸酶(NT5E)多肽、核孔膜糖蛋白210(NUP210)多肽、早老素相关菱形样蛋白、线粒体(PARL)多肽、过氧化物膜蛋白PEX13(PEX13)多肽、GPI乙醇胺磷酸转移酶1(PIGN)多肽、GPI转酰胺酶组分PIG-T(PIGT)多肽、胞浆型磷脂酶A2(PLA2G4A)多肽、1-磷脂酰肌醇-4,5-二磷酸磷酸二酯酶 ζ -1(PLCH1)多肽、网蛋白(PLEC)多肽、26S蛋白酶体非ATP酶调节亚基2(PSMD2)多肽、磷脂酰丝氨酸合成酶1(PTDSS1)多肽、前列腺素F2受体负调控因子(PTGFRN)多肽、受体型酪氨酸蛋白磷酸酶F(PTPRF)多肽、巯基氧化酶1(QSOX1)多肽、Ras相关蛋白Rab-25(RAB25)多肽、Ras相关蛋白Rab-38(RAB38)多肽、Ras-相关蛋白Rab-6B;(RAB6B)多肽、Ras相关蛋白Rap-2b(RAP2B)多肽、蛋白RCC2(RCC2)多肽、GTP结合蛋白Rit1(RIT1)多肽、分泌型载体相关膜蛋白3(SCAMP3)多肽、粘蛋白聚糖-1(SDC1)多肽、蛋白sel-1同源物3(SEL1L3)多肽、蛋白Shroom2(SHROOM2)多肽、溶质载体家族2、易化葡萄糖转运蛋白成员1(SLC2A1)多肽、钠依赖性磷酸盐转运蛋白2B(SLC34A2)多肽、腺苷3'-磷酸-5'-磷酸硫酸盐转运蛋白1(SLC35B2)多肽、锌转运蛋白ZIP11(SLC39A11)多肽、酸性鞘磷脂酶样磷酸二酯酶3b(SMPDL3B)多肽、甾醇O-酰基转移酶1(SOAT1)多肽、Spastin(SPAST)多肽、转位子相关蛋白亚基 α (SSR1)多肽、转位子相关蛋白亚基 δ (SSR4)多肽、Surfeit基因座蛋白4;(SURF4)多肽、突触回蛋白-2(SYNGR2)多肽、肿瘤相关钙信号转导因子2(TACSTD2)多肽、钙调磷酸酶B同源蛋白3(TESC)多肽、转铁蛋白受体蛋白1(TFRC)多肽、跨膜通道样蛋白5(TMC5)多肽、钙负荷激活钙通道(TMC01)多肽、含跨膜emp24结构域蛋白2(TMED2)多肽、含跨膜emp24结构域蛋白3(TMED3)多肽、跨膜蛋白132A(TMEM132A)多肽、跨膜蛋白33(TMEM33)多肽、跨膜丝氨酸蛋白酶4(TMPRSS4)多肽、蛋白O-甘露糖基-转移酶TMTC3(TMTC3)多肽、线粒体输入受体亚基TOM22同源物(TOMM22)多肽、Torsin-1A相互作用蛋白2、同种型IFRG15(TOR1AIP2)多肽、转位链相关膜蛋白1(TRAM1)多肽、瞬时受体电位阳离子通道亚家族V成员4(TRPV4)多肽、四肽重复蛋白33(TTC33)多肽、UDP-葡萄糖醛酸基转移酶1-6(UGT1A6)多肽、尿路斑块蛋白-1b(UPK1B)多肽、囊泡膜相关蛋白基因8(VAMP8)多肽、液泡ATP酶组装完整膜蛋白VMA21(VMA21)多肽、丝氨酸/苏氨酸-蛋白激酶VRK2(VRK2)多肽、含von Willebrand因子A结构域的蛋白1(VWA1)多肽、异嗜性和多嗜性逆转录病毒受体1(XPR1)多肽、木糖苷木糖基转移酶1(XXYL1)多肽、含解离素和金属蛋白酶结构域蛋白28(ADAM28)多肽、酪氨酸-蛋白激酶受体UF0(AXL)多肽、Basigin(BSG)多肽、程序性细胞死亡1配体1(CD274;也称为PD-L1)多肽、白细胞表面抗原CD47(CD47)多肽、聚集素(CLU)多肽、Dickkopf相关蛋白1(DKK1)多肽、受体酪氨酸蛋白激酶erbB-3(ERBB3)多肽、血管内皮生长因子受体3(FLT4)多肽、(N-甘氨酸神经氨酸(NeuGc,NGNA)-神经节苷脂GM3)(GM3)多肽、肝细胞生长因子(HGF)多肽、胰岛素样生长因子1受体(IGF1R)多肽、白细胞介素-6(IL6)多肽、血管内皮生长因子受体2(KDR)多肽、淋巴细胞激活基因3蛋白(LAG3)多肽、

Lewis Y/B抗原多肽、淋巴细胞抗原6E (LY6E) 多肽、神经源性基因座缺口同源蛋白2 (NOTCH2) 多肽、神经源性基因座缺口同源蛋白3 (NOTCH3) 多肽、磷脂酰丝氨酸递呈多肽、具有Ig和ITIM结构域的T细胞免疫受体 (TIGIT) 多肽、肿瘤坏死因子受体超家族成员10A (TNFRSF10A) 多肽、肿瘤坏死因子受体超家族成员10B (TNFRSF10B) 多肽、肿瘤坏死因子受体超家族成员18 (TNFSF18) 多肽、滋养层糖蛋白 (TPBG) 多肽、血管内皮生长因子A (VEGFA) 多肽、 α GalNAc-Ser/Thr (Tn) 抗原糖基化多肽、Lewis Y/CD174糖基化多肽、唾液酸化Lewis X (sLex) (也称为唾液酸化SSEA-1 (SLX)) 糖基化多肽、N-羟乙酰基GM3神经节苷脂 (NeuGcGM3) 糖基化多肽及其组合。

[0167] 在一些实施方案中,包含在肺癌的靶生物标志物特征中的一个或多个细胞外囊泡相关膜结合多肽是或包含:SLC34A2多肽、CEACAM5多肽、CEACAM6多肽、EpCAM多肽和/或其组合。SLC34A2多肽是一种多通道膜转运蛋白,其以往研究作为非小细胞肺癌的治疗靶标(Lin等人,2015;其出于本文所述目的以引用的方式并入本文)。CEACAM5多肽,是细胞粘附分子(CAM)的癌胚抗原(CEA)家族的成员,是与胃肠道癌症有关并且被认为与细胞分化、凋亡和极性有关的细胞表面糖蛋白。CEACAM6多肽是与CEACAM5同蛋白家族的成员,并且与克罗恩病(克罗恩病)和胰腺癌有关,其被认为涉及先天免疫系统和细胞表面相互作用。EpCAM多肽与胃肠道肿瘤有关,并且被认为作为同型钙非依赖性细胞粘附分子发挥作用。在一些实施方案中,SLC34A2多肽、CEACAM5多肽、CEACAM6多肽和/或EpCAM多肽被检测为完整EV相关跨膜蛋白。在本公开的一些实施方案中,SLC34A2多肽、CEACAM5多肽、CEACAM6多肽和/或EPCAM多肽被检测为EV相关跨膜多肽。

[0168] 在一些实施方案中,包含在肺癌的靶生物标志物特征中的靶生物标志物是或包含选自自由以下组成的组的表面蛋白生物标志物:CD166抗原(ALCAM)多肽、ATP结合盒亚家族C成员3(ABCC3)基因编码的微管多特异性有机阴离子转运蛋白2多肽、芳香基硫酸酯酶L(ARSL)基因编码的芳香基硫酸酯酶L多肽、UDP-GlcNAc: β Gal β -1,3-N-乙酰氨基葡萄糖转移酶3(B3GNT3)基因编码的N-乙酰乳糖胺 β -1,3-N-乙酰氨基葡萄糖转移酶3多肽、含CUB结构域蛋白1(CDCP1)基因编码的含CUB结构域蛋白1多肽、钙粘蛋白1(CDH1)基因编码的钙粘蛋白1多肽、钙粘蛋白3(CDH3)基因编码的钙粘蛋白3多肽、补体衰退加速因子(CD55)多肽、程序性细胞死亡1配体1(CD274;也称为PD-L1)多肽、癌胚抗原细胞粘附分子5多肽(CEACAM5)基因编码的癌胚抗原细胞粘附分子5多肽、癌胚抗原细胞粘附分子6多肽(CEACAM6)基因编码的癌胚抗原细胞粘附分子6多肽、钙粘蛋白EGF LAG七经G型受体1(CELSR1)基因编码的钙粘蛋白EGF LAG七经G型受体1多肽、claudin 18(CLDN18)基因编码的claudin 18多肽、claudin 3(CLDN3)基因编码的claudin 3多肽、(CLDN4)基因编码的claudin 4多肽、claudin 7(CLDN7)基因编码的claudin 7多肽、胞内氯离子通道蛋白6(CLIC6)基因编码的胞内氯离子通道蛋白6多肽、恶性脑肿瘤缺失1(DMBT1)基因编码的恶性脑肿瘤缺失1蛋白多肽、桥粒芯蛋白2(DSG2)基因编码的桥粒芯蛋白2多肽、表皮生长因子受体(EGFR)多肽、上皮细胞粘附分子(EPCAM)基因编码的上皮细胞粘附分子多肽、环氧化物酶3(EPHX3)基因编码的环氧化物酶3多肽、eva-1同源A编码的蛋白eva-1同源A多肽、程序性细胞死亡调节因子(EVA1A)基因、叶酸受体 α (FOLR1)基因编码的叶酸受体 α 多肽、缝隙连接蛋白 β 1(GJB1)基因编码的缝隙连接 β -1蛋白多肽、缝隙连接蛋白 β 2(GJB2)基因编码的缝隙连接 β -2蛋白多肽、肝细胞生长因子受体(MET)多肽、胰岛素样生长因子1受体(IG1FR)多肽、糖基磷脂酰肌醇蛋

白聚糖4 (GPC4) 基因编码的糖基磷脂酰肌醇蛋白聚糖4多肽、硫酸乙酰肝素6-0-磺基转移酶2 (HS6ST2) 基因编码的硫酸乙酰肝素6-0-磺基转移酶2多肽、ER管腔蛋白滞留受体3 (KDEL3) 基因编码的ER管腔蛋白滞留受体3多肽、角质形成细胞相关蛋白3 (KRTCAP3) 基因编码的角质形成细胞相关蛋白3多肽、层粘连蛋白亚基β3 (LAMB3) 基因编码的层粘连蛋白亚基β-3多肽、LFNG 0-岩藻糖基肽3-β-N-乙酰氨基葡萄糖转移酶 (LFNG) 基因编码的β-1,3-N-乙酰氨基葡萄糖转移酶lunatic边缘多肽、脂解刺激脂质蛋白受体 (LSR) 基因编码的脂解刺激脂质蛋白受体多肽、甘露糖苷酶内切α样 (MANEAL) 基因编码的糖蛋白内切-α-1,2-甘露糖苷酶样蛋白多肽、间皮素 (MSLN) 基因编码的间皮素多肽、粘蛋白1编码的粘蛋白1多肽、细胞表面相关 (MUC1) 基因、粘蛋白21编码的粘蛋白21多肽、细胞表面相关 (MUC21) 基因、磷脂酰肌醇聚糖锚定生物合成类T (PIGT) 基因编码的GPI转酰胺酶组分PIG-T多肽、足糖萼蛋白样2 (PODXL2) 基因编码的足糖萼蛋白样蛋白2多肽、富含脯氨酸和G1a结构域4 (PRRG4) 基因编码的跨膜γ-羧基谷氨酸蛋白4多肽、ROS原癌基因1编码的原癌基因酪氨酸-蛋白激酶ROS多肽、受体酪氨酸蛋白激酶 (ROS1) 基因、粘蛋白聚糖1 (SDC1) 基因编码的粘蛋白聚糖1多肽、丝氨酸整合因子2 (SERINC2) 基因编码的丝氨酸整合因子2多肽、惊厥相关6同源物样2 (SEZ6L2) 基因编码的惊厥6-样蛋白2多肽、溶质载体家族34 (磷酸钠) 成员2 (SLC34A2) 基因编码的钠依赖型磷转运蛋白2B多肽、溶质载体家族44成员4 (SLC44A4) 基因编码的胆碱转运蛋白样蛋白4多肽、溶质载体家族6成员14 (SLC6A14) 基因编码的钠离子和氯离子依赖的中性和碱性氨基酸转运蛋白B(0+) 多肽、溶质载体家族7成员7 (SLC7A7) 基因编码的Y+L氨基酸转运蛋白1多肽、小膜内在蛋白22 (SMIM22) 基因编码的小膜内在蛋白22多肽、鞘磷脂磷酸二酯酶酸样3B (SMPDL3B) 基因编码的酸性鞘磷脂酶样磷酸二酯酶3b多肽、ST14跨膜丝氨酸蛋白酶蛋白裂解酶 (ST14) 基因编码的瘤变抑制因子14蛋白多肽、肿瘤相关钙信号传导因子2 (TACSTD2) 多肽、跨膜通道样4 (TMC4) 基因编码的跨膜通道样蛋白4多肽、跨膜通道样5 (TMC5) 基因编码的跨膜通道样蛋白5多肽、跨膜蛋白45B (TMEM45B) 基因编码的跨膜蛋白45B多肽、跨膜丝氨酸蛋白酶2 (TMPRSS2) 基因编码的跨膜丝氨酸蛋白酶2多肽、跨膜丝氨酸蛋白酶4 (TMPRSS4) 基因编码的跨膜丝氨酸蛋白酶4多肽、肿瘤坏死因子受体超家族成员10B (TNFRSF10B) 多肽、四次跨膜蛋白1 (TSPAN1) 基因编码的四次跨膜蛋白-1多肽、四次跨膜蛋白8 (TSPAN8) 基因编码的四次跨膜蛋白-8多肽、sTn抗原糖基化多肽、Tn抗原糖基化多肽、T抗原糖基化多肽及其组合。

[0169] 在一些实施方案中,包含在肺癌的靶生物标志物特征中的靶生物标志物是或包含选自由以下组成的组的表面蛋白生物标志物:SLC34A2多肽、CEACAM5多肽、CEACAM6多肽、EpCAM多肽及其组合。

[0170] 在一些实施方案中,包含在肺癌的靶生物标志物特征中的靶生物标志物是或包含选自由以下组成的组的表面蛋白生物标志物:ALCAM多肽、CD55多肽、CDH1多肽、CDH3多肽、CD274 (PD-L1) 多肽、CEACAM5多肽、CEACAM6多肽、DSG2多肽、EGFR多肽、EPCAM多肽、FOLR1多肽、IG1FR多肽、MET多肽、MSLN多肽、MUC1多肽、SLC34A2多肽、sTn抗原糖基化多肽、Tn抗原糖基化多肽、T抗原糖基化多肽、TACSTD2多肽、TNFRSF10B多肽及其组合。

[0171] 在一些实施方案中,靶生物标志物特征包含一个或多个细胞外囊泡相关膜结合多肽生物标志物,其选自由以下组成的列表:ABCA3多肽、ABCC1多肽、ABCC3多肽、ACBD3多肽、ACSL5多肽、AGER多肽、ALCAM多肽、AP1M2多肽、APH1多肽、APO0多肽、ATP11多肽、ATP11B多

肽、ATP1B1多肽、ATP6AP2多肽、B4GALT4多肽、BCAP31多肽、BSPRY多肽、CD109多肽、CD55多肽、CD9多肽、CDC42多肽、CDH1多肽、CDH3多肽、CDKAL1多肽、CEACAM5多肽、CEACAM6多肽、CELSR1多肽、CIP2多肽、CISD2多肽、CKAP4多肽、CLCA2多肽、CLDN1多肽、CLIC6多肽、CLPTM1L多肽、CLSTN1多肽、CNTN1多肽、CPD多肽、CYP2S1多肽、CYP4F11多肽、CYP4F3多肽、DPY19L1多肽、DSC2多肽、DSC3多肽、DSG2多肽、DSG3多肽、EGFR多肽、EPCAM多肽、EPHB3多肽、FAT2多肽、FBXO45多肽、FERMT1多肽、FOLR1多肽、FZD6多肽、GALNT1多肽、GALNT3多肽、GALNT5多肽、GALNT6多肽、GGCX多肽、GOLM1多肽、GOLPH3L多肽、GRHL2多肽、HACD3多肽、IER3IP1多肽、IGSF3多肽、IL1RAP多肽、ITGA2多肽、ITGB6多肽、KLRG2多肽、KPNA2多肽、KRTCAP3多肽、LAD1多肽、LAMB3多肽、LAMC2多肽、LAMP3多肽、LAMTOR2多肽、LCLAT1多肽、LPCAT1多肽、LSR多肽、MAGT1多肽、MARCKSL1多肽、MET多肽、MGAT1多肽、MSLN多肽、MUC1多肽、MUC4多肽、NCSTN多肽、NECTIN1多肽、NECTIN4多肽、NRAS多肽、NT5E多肽、NUP210多肽、PARL多肽、PEX13多肽、PIGN多肽、PIGT多肽、PLA2G4多肽、PLCH1多肽、PLEC多肽、PSMD2多肽、PTDSS1多肽、PTGFRN多肽、PTPRF多肽、QSOX1多肽、RAB25多肽、RAB38多肽、RAB6B多肽、RAP2B多肽、RCC2多肽、RIT1多肽、SCAMP3多肽、SDC1多肽、SEL1L3多肽、SHROOM2多肽、SLC2A1多肽、SLC34A2多肽、SLC35B2多肽、SLC39A11多肽、SMPDL3B多肽、SOAT1多肽、SPAST多肽、SSR1多肽、SSR4多肽、SURF4多肽、SYNGR2多肽、TACSTD2多肽、TESC多肽、TFRC多肽、TMC5多肽、TMC01多肽、TMED2多肽、TMED3多肽、TMEM132A多肽、TMEM33多肽、TMPRSS4多肽、TMTC3多肽、TOMM22多肽、TOR1AIP2多肽、TRAM1多肽、TRPV4多肽、TTC33多肽、UGT1A6多肽、UPK1B多肽、VAMP8多肽、VMA21多肽、VRK2多肽、VWA1多肽、XPR1多肽、XXYLT1多肽、ADAM28多肽、AXL多肽、BSG多肽、CD274多肽、CD47多肽、CLU多肽、DKK1多肽、ERBB3多肽、FLT4多肽、GM3多肽、HGF多肽、IGF1R多肽、IL6多肽、KDR多肽、LAG3多肽、Lewis Y/B抗原多肽、LY6E多肽、NOTCH2多肽、NOTCH3多肽、磷脂酰丝氨酸呈递多肽、TIGIT多肽、TNFRSF10多肽、TNFRSF10B多肽、TNFSF18多肽、TPBG多肽、VEGF多肽、Tn抗原糖基化多肽、Lewis Y/CD174糖基化多肽、唾液酸化Lewis X(sLex)(也称为唾液酸化SSEA-1(SLX))糖基化多肽、NeuGcGM3糖基化多肽及其组合。

[0172] 在一些实施方案中,靶生物标志物特征包含一个或多个细胞外囊泡相关膜结合多肽生物标志物,其选自由以下组成的列表:HS6ST2多肽、CYP2S1多肽、HAS3多肽、LAMC2多肽、ADAM23多肽、ABCA13多肽、TMPRSS4多肽、UGT1A6多肽、ILDR1多肽、CYP4F11多肽、PIGT多肽、LAMB3多肽、PRSS21多肽、DSG3多肽、SDK2多肽及其组合。

[0173] 在一些实施方案中,靶生物标志物特征包含一个或多个细胞外囊泡相关膜结合多肽生物标志物,其选自由以下组成的列表:HS6ST2多肽、CYP2S1多肽、HAS3多肽、LAMC2多肽、ADAM23多肽、ABCA13多肽、TMPRSS4多肽、UGT1A6多肽、ILDR1多肽、CYP4F11多肽、PIGT多肽、LAMB3多肽、PRSS21多肽、DSG3多肽、SDK2多肽、FERMT1多肽、EPCAM多肽、SDC1多肽、PANX2多肽、ULBP2多肽、ECE2多肽、KRTCAP3多肽、CLCA2多肽、KPNA2多肽、TMEM132A多肽、ABCC1多肽、UPK1B多肽、DSG2多肽、NECTIN1多肽、SHISA2多肽及其组合。

[0174] 在一些实施方案中,包含在肺癌的靶生物标志物特征中的靶生物标志物是或包含选自由以下组成的组的囊泡内蛋白生物标志物:胺氧化酶含铜1(AOC1)基因编码的阿米洛利(amiloride)敏感的胺氧化酶[含铜]多肽、染色体12开放阅读框45(C12orf45)基因编码的未表征蛋白C12orf45多肽、细胞视黄酸结合蛋白2(CRABP2)基因编码的细胞视黄酸结合蛋白2多肽、胱蛋白SN(CST1)基因编码的胱蛋白SN多肽、ETS变体转录因子4(ETV4)基因编码

的ETS易位变体4多肽、序列类似性83家族成员A (FAM83A) 基因编码的蛋白FAM83A多肽、叉头框蛋白A2 (FOXA2) 基因编码的肝细胞核因子3- β 多肽、高迁移率组蛋白3 (HMGB3) 基因编码的高迁移率组蛋白B3多肽、半乳凝素3结合蛋白 (LGALS3BP) 基因编码的半乳凝素-3结合蛋白多肽、巨噬细胞游走抑制因子 (MIF) 基因编码的巨噬细胞游走抑制因子多肽、napsin A天冬氨酸肽酶 (NAPSA) 基因编码的napsin-A多肽、蛋白磷酸酶1调节抑制亚基14D (PPP1R14D) 基因编码的蛋白磷酸酶1调节亚基14D多肽、S100钙结合蛋白A14 (S100A14) 基因编码的蛋白S100-A14多肽、SH3结构域结合激酶1 (SBK1) 基因编码的丝氨酸/苏氨酸蛋白激酶SBK1多肽、分泌珠蛋白家族3A成员2 (SCGB3A2) 基因编码的分泌珠蛋白家族3A成员2多肽、表面活性物质结合2 (SFTA2) 基因编码的表面活性物质结合蛋白2多肽、表面活性物质蛋白A1 (SFTPA1) 基因编码的肺表面活性物质相关蛋白A1多肽、表面活性蛋白A2 (SFTPA2) 基因编码的肺表面活性物质相关蛋白A2多肽、表面活性物质蛋白B (SFTPB) 基因编码的肺表面活性物质相关蛋白B多肽、丝氨酸蛋白酶抑制剂Kazal 1型 (SPINK1) 基因编码的丝氨酸蛋白酶抑制剂Kazal 1型多肽、转化生长因子 α (TGFA) 基因编码的促转化生长因子 α 多肽、含锌指CCCH型11A (ZC3H11A) 基因编码的含锌指CCCH结构域蛋白11A多肽及其组合。

[0175] 在一些实施方案中, 肺癌的靶生物标志物特征中的靶生物标志物是或包含选自由以下组成的组的囊泡内蛋白生物标志物: ABRACL多肽、ACP5多肽、ADH7多肽、AGR2多肽、AIF1多肽、AKR1C1多肽、AKR1C2多肽、AKR1C3多肽、ALDH1A1多肽、ALDH3A1多肽、ALDH3B2多肽、ALG1L多肽、AP1M2多肽、APOBEC3B多肽、APOBEC3C多肽、ARNTL2多肽、ASF1B多肽、AURKB多肽、BAIAP2L1多肽、BIRC5多肽、C15orf48多肽、C19orf33多肽、C1S多肽、C8orf4多肽、CA9多肽、CALML3多肽、CAPNS2多肽、CBLC多肽、CCL19多肽、CCNB2多肽、CDC20多肽、CDC45多肽、CDCA4多肽、CDCA5多肽、CDK1多肽、CDKN2多肽、CDKN2B多肽、CENPW多肽、CEP55多肽、CES1多肽、CHMP4C多肽、CNN2多肽、CPA3多肽、CRABP2多肽、CST多肽、CTSC多肽、CTSE多肽、CYP2S1多肽、DPYSL3多肽、EFS多肽、EGLN3多肽、EHF多肽、ELF3多肽、ELF4多肽、ENAH多肽、ESRP1多肽、EVPL多肽、FAM129B多肽、FAM60多肽、FAM83D多肽、FAM83H多肽、FBP1多肽、FERMT1多肽、FOX E1多肽、FOX M1多肽、GBP6多肽、GNA15多肽、GPX2多肽、GRHL2多肽、GSTA1多肽、HCK多肽、HOXB7多肽、ID1多肽、IGF2BP2多肽、IMPA2多肽、IRF6多肽、IVL多肽、JUP多肽、KIAA1522多肽、KIF2C多肽、KIFC1多肽、KLF4多肽、KLF5多肽、KRT13多肽、KRT14多肽、KRT15多肽、KRT16多肽、KRT17多肽、KRT18多肽、KRT19多肽、KRT5多肽、KRT6多肽、KRT6B多肽、KRT6C多肽、KRT7多肽、KRT8多肽、LGALS7B多肽、LSP1多肽、MAGEA4多肽、MAGEA6多肽、MCM2多肽、MDFI多肽、MYBL2多肽、MYH14多肽、MZB1多肽、NCF2多肽、NNMT多肽、NRARP多肽、NUP210多肽、NUSAP1多肽、OSGIN1多肽、PALLD多肽、PITX1多肽、PKP1多肽、PKP3多肽、PLEK多肽、PLEK2多肽、POSTN多肽、PPP1R14C多肽、PRAME多肽、PTPN6多肽、RBP1多肽、RIN2多肽、RIPK4多肽、RPS4Y1多肽、RRM2多肽、S100A11多肽、S100A14多肽、S100A16多肽、S100A2多肽、S100P多肽、SERPINB13多肽、SERPINB3多肽、SERPINB5多肽、SH3BP4多肽、SNAI2多肽、SOX2多肽、SPI1多肽、SPINT1多肽、SPRR1多肽、SPRR1B多肽、SPRR2多肽、SPRR2D多肽、SPRR2E多肽、SPRR3多肽、SULF1多肽、SYK多肽、SYTL1多肽、TBC1D2多肽、TEAD2多肽、TEAD3多肽、TFAP2C多肽、THBS2多肽、TK1多肽、TOP2多肽、TP63多肽、TPD52多肽、TPX2多肽、TRIM29多肽、TRIP13多肽、UBE2C多肽、YAP1多肽、ZC3H11多肽、ZNF217多肽、ZNF750多肽及其组合。

[0176] 在一些实施方案中, 包含在肺癌的靶生物标志物特征中的靶生物标志物是或包含

选自由以下组成的组的囊泡内RNA(例如,mRNA)生物标志物:ABCC3 RNA、AOC1 RNA、ARSL RNA、B3GNT3 RNA、C12orf45 RNA、CDCP1 RNA、CDH1 RNA、CDH3 RNA、CEACAM5 RNA、CEACAM6 RNA、CELSR1 RNA、CLDN18 RNA、CLDN3 RNA、CLDN4 RNA、CLDN7 RNA、CLIC6 RNA、CRABP2 RNA、CST1 RNA、DMBT1 RNA、DSG2 RNA、EPCAM RNA、EPHX3 RNA、ETV4 RNA、EVA1A RNA、FAM83A RNA、FOLR1 RNA、FOXA2 RNA、GJB1 RNA、GJB2 RNA、GPC4 RNA、HMGB3 RNA、HS6ST2 RNA、KDELR3 RNA、KRTCAP3 RNA、LAMB3 RNA、LFNG RNA、LGALS3BP RNA、LSR RNA、MANEAL RNA、MIF RNA、MSLN RNA、MUC1 RNA、MUC21 RNA、NAPSA RNA、PIGT RNA、PODXL2 RNA、PPP1R14D RNA、PRRG4 RNA、ROS1 RNA、S100A14 RNA、SBK1 RNA、SCGB3A2 RNA、SDC1 RNA、SERINC2 RNA、SEZ6L2 RNA、SFTA2 RNA、SFTPA1 RNA、SFTPA2 RNA、SFTPB RNA、SLC34A2 RNA、SLC44A4 RNA、SLC6A14 RNA、SLC7A7 RNA、SMIM22 RNA、SMPDL3B RNA、SPINK1 RNA、ST14 RNA、TGFA RNA、TMC4 RNA、TMC5 RNA、TMEM45B RNA、TMPRSS2 RNA、TMPRSS4 RNA、TSPAN1 RNA、TSPAN8 RNA、ZC3H11A RNA及其组合。

[0177] 在一些实施方案中,靶生物标志物特征包含由以下组成的列表中的一种或多种囊泡内RNA(例如,mRNA)生物标志物:ABCA3 RNA、ABCC1 RNA、ABRACL RNA、ACP5 RNA、ADAM23 RNA、ADH7 RNA、AGR2 RNA、AIF1 RNA、AKR1C1 RNA、AKR1C2 RNA、AKR1C3 RNA、ALDH1A1 RNA、ALDH3A1 RNA、ALDH3B2 RNA、ALG1L RNA、ANTXR1 RNA、AP1M2 RNA、APOBEC3B RNA、APOBEC3C RNA、AQP3 RNA、AREG RNA、ARNTL2 RNA、ASF1B RNA、ATP8B1 RNA、AURKB RNA、B3GNT5 RNA、BAIAP2L1 RNA、BCAM RNA、BIK RNA、BIRC5 RNA、C15orf48 RNA、C19orf33 RNA、C1S RNA、C8orf4 RNA、CA12 RNA、CA9 RNA、CALML3 RNA、CAPNS2 RNA、CBLC RNA、CCL19 RNA、CCL5 RNA、CCNB2 RNA、CD109 RNA、CD24 RNA、CD53 RNA、CD74 RNA、CD9 RNA、CDC20 RNA、CDC42EP1 RNA、CDC45 RNA、CDCA4 RNA、CDCA5 RNA、CDCP1 RNA、CDH1 RNA、CDH3 RNA、CDK1 RNA、CDKN2A RNA、CDKN2B RNA、CEACAM5 RNA、CEACAM6 RNA、CELSR1 RNA、CENPW RNA、CEP55 RNA、CES1 RNA、CHMP4C RNA、CLCA2 RNA、CLDN1 RNA、CLDN4 RNA、CLDN7 RNA、CNN2 RNA、COL17A1 RNA、CPA3 RNA、CRABP2 RNA、CSTA RNA、CTSC RNA、CTSE RNA、CX3CL1 RNA、CXADR RNA、CXCR4 RNA、CYBB RNA、CYP2S1 RNA、CYP4F11 RNA、DAPL1 RNA、DPYSL3 RNA、DSC2 RNA、DSC3 RNA、DSG2 RNA、DSG3 RNA、DSP RNA、EFNA1 RNA、EFS RNA、EGFR RNA、EGLN3 RNA、EHD2 RNA、EHF RNA、ELF3 RNA、ELF4 RNA、EMP1 RNA、EMP2 RNA、ENAH RNA、EPCAM RNA、EPHA2 RNA、EPHB3 RNA、EPHX1 RNA、ESRP1 RNA、EVPL RNA、F11R RNA、F2R RNA、F2RL1 RNA、F3 RNA、FAM129B RNA、FAM60A RNA、FAM83D RNA、FAM83H RNA、FAT1 RNA、FAT2 RNA、FBLIM1 RNA、FBP1 RNA、FCER1G RNA、FERMT1 RNA、FGFR2 RNA、FGFR3 RNA、FOX E1 RNA、FOX M1 RNA、FX YD3 RNA、GALNT3 RNA、GBP6 RNA、GJA1 RNA、GJB2 RNA、GJB3 RNA、GJB5 RNA、GJB6 RNA、GNA15 RNA、GPC1 RNA、GPC3 RNA、GPNMB RNA、GPR87 RNA、GPCR5A RNA、GPX2 RNA、GRHL2 RNA、GSTA1 RNA、HAS3 RNA、HCK RNA、HOXB7 RNA、ID1 RNA、IGF2BP2 RNA、IGSF9 RNA、IL2RG RNA、IMPA2 RNA、IRF6 RNA、ITGA2 RNA、ITGA6 RNA、ITGB4 RNA、ITGB6 RNA、IVL RNA、JAG2 RNA、JUP RNA、KCNS3 RNA、KIAA1522 RNA、KIF2C RNA、KIFC1 RNA、KITLG RNA、KLF4 RNA、KLF5 RNA、KRT13 RNA、KRT14 RNA、KRT15 RNA、KRT16 RNA、KRT17 RNA、KRT18 RNA、KRT19 RNA、KRT5 RNA、KRT6A RNA、KRT6B RNA、KRT6C RNA、KRT7 RNA、KRT8 RNA、KRTCAP3 RNA、LAMP3 RNA、LAPTM5 RNA、LGALS7B RNA、LRP11 RNA、LRRC4 RNA、LSP1 RNA、LSR RNA、LYPD3 RNA、MAGEA4 RNA、MAGEA6 RNA、MAL2 RNA、MAOA RNA、MARCO RNA、MCM2 RNA、MDFI RNA、MET RNA、MMP14

RNA、MPZL2 RNA、MUC1 RNA、MYBL2 RNA、MYH14 RNA、MYOF RNA、MZB1 RNA、NCF2 RNA、NKG7 RNA、NNMT RNA、NOTCH3 RNA、NRARP RNA、NTRK2 RNA、NUP210 RNA、NUSAP1 RNA、OSGIN1 RNA、OSMR RNA、PALLD RNA、PDPN RNA、PDZK1IP1 RNA、PECAM1 RNA、PERP RNA、PIGR RNA、PIGT RNA、PITX1 RNA、PKP1 RNA、PKP3 RNA、PLEKRNA、PLEK2 RNA、PLVAP RNA、PMP22 RNA、POSTN RNA、PPL RNA、PPP1R14C RNA、PRAME RNA、PROM2 RNA、PRRG4 RNA、PRSS8 RNA、PTGES RNA、PTGFRN RNA、PTPN6 RNA、PTPRF RNA、PTPRZ1 RNA、RAB25 RNA、RAB38 RNA、RAET1L RNA、RARRES1 RNA、RBP1 RNA、RGS1 RNA、RHCG RNA、RHOF RNA、RIN2 RNA、RIPK4 RNA、RPS4Y1 RNA、RRM2 RNA、S100A10 RNA、S100A11 RNA、S100A14 RNA、S100A16 RNA、S100A2 RNA、S100P RNA、SCNN1A RNA、SDC1 RNA、SERINC2 RNA、SERPINB13 RNA、SERPINB3 RNA、SERPINB5 RNA、SEZ6L2 RNA、SH3BP4 RNA、SHISA2 RNA、SLC1A5 RNA、SLC2A1 RNA、SLC34A2 RNA、SLC40A1 RNA、SLC6A8 RNA、SLC7A8 RNA、SNAI2 RNA、SOX2 RNA、SPI1 RNA、SPINT1 RNA、SPINT2 RNA、SPRR1A RNA、SPRR1B RNA、SPRR2A RNA、SPRR2D RNA、SPRR2E RNA、SPRR3 RNA、ST14 RNA、STEAP1 RNA、SULF1 RNA、SYK RNA、SYTL1 RNA、TACSTD2 RNA、TBC1D2 RNA、TEAD2 RNA、TEAD3 RNA、TFAP2C RNA、THBD RNA、THBS2 RNA、TK1 RNA、TM4SF1 RNA、TMC4 RNA、TMEM30B RNA、TMEM54 RNA、TMPRSS11D RNA、TMPRSS11E RNA、TMPRSS4 RNA、TNFRSF18 RNA、TNS4 RNA、TOP2A RNA、TP53I11 RNA、TP63 RNA、TPD52 RNA、TPX2 RNA、TREM2 RNA、TRIM29 RNA、TRIP13 RNA、TSPAN1 RNA、TSPAN13 RNA、TSPAN6 RNA、TSPAN7 RNA、TUSC3 RNA、TYROBP RNA、UBE2C RNA、UPK1B RNA、VAMP8 RNA、VANG2 RNA、WLS RNA、YAP1 RNA、ZC3H11A RNA、ZNF217 RNA、ZNF750 RNA及其组合。

[0178] 在一些实施方案中，肺癌的靶生物标志物特征包含至少一个或多个(例如，1、2、3、4、5、6、7、8个或更多个)细胞外囊泡相关膜结合多肽(例如，本文所述的那些)和至少一个或多个(例如，1、2、3、4、5、6、7、8个或更多个)表面蛋白生物标志物(例如，本文所述的那些)。在一些实施方案中，至少一个细胞外囊泡相关膜结合多肽和至少一个表面蛋白生物标志物是相同的。在一些实施方案中，肺癌的靶生物标志物特征的至少一个细胞外囊泡相关膜结合多肽和至少一个表面蛋白生物标志物是不同的。例如，在一些实施方案中，肺癌的靶生物标志物特征包含至少一个细胞外囊泡相关膜结合多肽(其是或包含SLC34A2多肽)和至少一个表面蛋白生物标志物(其是或包含CEACAM6多肽和/或EpCAM多肽)。在一些实施方案中，肺癌的靶生物标志物特征包含至少一个细胞外囊泡相关膜结合多肽(其是或包含CEACAM5多肽)和至少一个表面蛋白生物标志物(其是或包含CEACAM6多肽和/或SLC34A2多肽)。

[0179] 在一些实施方案中，肺癌的靶生物标志物特征包含至少一个或多个(例如，1、2、3、4、5、6、7、8个或更多个)细胞外囊泡相关膜结合多肽(例如，本文所述的那些)和至少一个或多个(例如，1、2、3、4、5、6、7、8个或更多个)囊泡内蛋白生物标志物(例如，本文所述的那些)。在一些此类实施方案中，一个或多个细胞外囊泡相关膜结合多肽和一个或多个囊泡内蛋白生物标志物可以由相同的基因编码，而前者在细胞外囊泡的膜中表达，并且后者在细胞外囊泡内表达。在一些此类实施方案中，一个或多个细胞外囊泡相关膜结合多肽和一个或多个囊泡内蛋白生物标志物可以由不同的基因编码。

[0180] 在一些实施方案中，肺癌的靶生物标志物特征包含至少一个或多个(例如，1、2、3、4、5、6、7、8个或更多个)细胞外囊泡相关膜结合多肽(例如，本文所述的那些)和至少一个或多个(例如，1、2、3、4、5、6、7、8个或更多个)囊泡内RNA(例如，mRNA)生物标志物(例如，本文

所述的那些)。在一些此类实施方案中,一个或多个细胞外囊泡相关膜结合多肽和一个或多个囊泡内RNA(例如,mRNA)生物标志物可以由相同的基因编码。在一些此类实施方案中,一个或多个细胞外囊泡相关膜结合多肽和一个或多个囊泡内RNA(例如,mRNA)生物标志物可以由不同的基因编码。

[0181] 在一些实施方案中,肺癌的靶生物标志物特征包括如表4中描述的生物标志物的组合。在一些实施方案中,此类组合中的生物标志物被用作多肽靶标的捕获探针(作为细胞外囊泡相关膜结合多肽),例如,如表4中描述的。在一些实施方案中,此类组合中的生物标志物被用作多肽靶标的检测探针(作为靶表面蛋白生物标志物),例如,如表4中描述的。

[0182] 在一些实施方案中,肺癌的靶生物标志物特征包含至少一个SLC34A2多肽(作为细胞外囊泡相关膜结合多肽)和一个CEACAM6多肽(作为靶表面蛋白生物标志物)。

[0183] 在一些实施方案中,肺癌的靶生物标志物特征包含至少一个SLC34A2多肽(作为细胞外囊泡相关膜结合多肽)和一个EpCAM多肽(作为靶表面蛋白生物标志物)。

[0184] 在一些实施方案中,肺癌的靶生物标志物特征包含至少一个CEACAM5多肽(作为细胞外囊泡相关膜结合多肽)和一个CEACAM6多肽(作为靶表面蛋白生物标志物)。

[0185] 在一些实施方案中,肺癌的靶生物标志物特征包含至少一个CEACAM5多肽(作为细胞外囊泡相关膜结合多肽)和一个SLC34A2多肽(作为靶表面蛋白生物标志物)。

[0186] 在一些实施方案中,肺癌的靶生物标志物特征包含至少一个SLC34A2多肽(作为细胞外囊泡相关膜结合多肽)和至少两个靶表面蛋白生物标志物,其可以是或包含CEACAM6多肽和EpCAM多肽。

[0187] 在一些实施方案中,肺癌的靶生物标志物特征包含至少一个CEACAM5多肽(作为细胞外囊泡相关膜结合多肽)和至少两个靶表面蛋白生物标志物,其可以是或包含CEACAM6多肽和SLC34A2多肽。

[0188] 在一些实施方案中,可以通过野生型形式的蛋白质和/或RNA(例如,mRNA)表达水平来检测和/或测量提供的生物标志物中的任何一个。

[0189] 在一些实施方案中,可以通过突变型形式的蛋白质和/或RNA(例如,mRNA)表达水平来检测和/或测量提供的生物标志物中的任何一个。因此,在一些实施方案中,可以包括对提供的生物标志物(例如,蛋白质和/或RNA,例如像mRNA)的突变型特异性检测。

[0190] 如本文所提到的,在一些实施方案中,生物标志物是或包含一种或多种多肽或蛋白质的特定形式(例如,替代形式、截短形式、修饰形式诸如糖基化、磷酸化、脂化形式等)。在一些实施方案中,此类形式的检测检测以该形式存在的多种(并且,在一些实施方案中,基本上所有)多肽(例如,含有特定修饰,诸如像特定糖基化,例如唾液酸化Tn(sTn)糖基化,例如含有与GalNAc α -0-Ser/Thr连接的唾液酸 α -2,6的截短的0-聚糖)。

[0191] 因此,在一些实施方案中,表面蛋白生物标志物可以是或包含糖基化部分(例如,sTn抗原部分、Tn抗原部分或T抗原部分)。汤普森-努维尔(Thompson-nouvelle)(Tn)抗原是被认为与多种肿瘤相关的0-连接聚糖。Tn是添加至Ser或Thr作为主要0-连接糖基化途径的第一步的单个 α -连接GalNAc。熟练技术人员将理解,在某些实施方案中,T抗原通常是指具有Gal β 1-3GalNAc-结构的0-连接聚糖。

[0192] 在一些实施方案中,表面蛋白生物标志物可以是或包含肿瘤相关翻译后修饰。在一些实施方案中,此类翻译后修饰可以是或包括肿瘤特异性糖基化模式,诸如具有在初始

GalNAc (例如, Tn) 处异常截短的聚糖的粘蛋白或其组合。

[0193] 在一些实施方案中, 靶生物标志物特征包含如表4中描述的组合的靶标, 其中靶标可以用在捕获探针和/或检测探针中。在一些实施方案中, 靶生物标志物特征可以包括如表4中描述的捕获探针的靶标和至少一种或多种 (包括, 例如至少两种或更多种) 检测探针 (例如, 检测探针1和/或检测探针2) 的靶标。

[0194] 在一些实施方案中, 如表4中描述的可能对肺癌检测特别有用 (例如, 具有更高的灵敏度、特异性和/或PPV) 的某些生物标志物组合可以使用进展期 (例如, 晚期, 例如, III期和/或IV期) 肺癌样品池和健康对照样品池作为参考进行初始轮筛查。在一些实施方案中, 可以使用早期肺癌样品池 (例如, I期和/或II期, 任选地通过低或适当区分)、良性肺肿瘤血浆样品池 (例如, 如本文所述)、非肺癌样品池 (例如, 如本文所述) 和/或其任何组合进一步测试选择的组合。在一些实施方案中, 可以通过计算健康样品池和肺癌样品池之间的测定信号 (例如, 在Ct基础上) 差异来确定生物标志物组合性能。

[0195] 在一些实施方案中, 可以用大于组间变异性的 ΔCt 来选择某些用于肺癌检测的生物标志物组合。例如, 在一些实施方案中, 具有大于2.0 (对应于四倍差值) 或1.0 (对应于两倍差值) 的 ΔCt 的生物标志物组合被认为提供特别有效的诊断效用 (例如, 提供大于组间变异性的信号)。参见例如实施例10, 其提供了本文所述的某些组合的示例性分析。

[0196] 在一些实施方案中, 肺癌的靶生物标志物特征可以是或包含如表4中所述的组合的靶标。在某些实施方案中, 肺癌的靶生物标志物特征包含区分晚期肺癌样品与对照样品 (例如, 与健康吸烟者样品相比, 和/或与非吸烟者样品相比; 参见例如, 表4) 的一组标志物。在某些实施方案中, 肺癌的靶生物标志物特征包含区分早期肺癌样品与对照样品 (例如, 与健康吸烟者样品相比, 和/或与非吸烟者样品相比; 参见例如, 表4) 的一组标志物。在一些实施方案中, 针对检测肺癌的靶生物标志物特征的测定可以包括如表4中所述的捕获和检测探针的组合。

[0197] 在一些实施方案中, 肺癌的靶生物标志物特征可以是或包含如实施例10的表4中所述的区分患有早期肺癌 (例如, 早期非小细胞肺癌诸如LUAD和/或LUSC) 的受试者和不具有肺癌的受试者 (例如, 健康受试者或具有不是肺癌或与肺癌不相关的病状的受试者) 的组合中的靶标。在一些实施方案中, 肺癌的靶生物标志物特征可以是或包含如实施例10的表4中所述的区分患有晚期肺癌 (例如, 晚期非小细胞肺癌诸如LUAD和/或LUSC) 的受试者和不具有肺癌的受试者 (例如, 健康受试者或具有不是肺癌或与肺癌不相关的病状的受试者) 的组合中的靶标。在一些实施方案中, 肺癌的靶生物标志物特征可以是或包含如实施例10的表4中所述的区分患有早期肺癌 (例如, 早期非小细胞肺癌诸如LUAD和/或LUSC) 的受试者和患有晚期肺癌 (例如, 晚期非小细胞肺癌诸如LUAD和/或LUSC) 的受试者的组合中的靶标。

[0198] 在一些实施方案中, 肺癌的靶生物标志物特征可以是或包含在8种测试的情况的至少8种中区分肺癌和健康样品的如实施例10的表4中所述的组合中的靶标。在一些实施方案中, 肺癌的靶生物标志物特征可以是或包含在8种测试的情况的至少7种中区分肺癌和健康样品的如实施例10的表4中所述的组合中的靶标。在一些实施方案中, 肺癌的靶生物标志物特征可以是或包含在8种测试的情况的至少6种中区分肺癌和健康样品的如实施例10的表4中所述的组合中的靶标。在一些实施方案中, 肺癌的靶生物标志物特征可以是或包含在8种测试的情况的至少5种中区分肺癌和健康样品的如实施例10的表4中所述的组合中的靶

标。在一些实施方案中,肺癌的靶生物标志物特征可以是或包含在8种测试的情况的至少4种中区分肺癌和健康样品的如实施例10的表4中所述的组合中的靶标。在一些实施方案中,肺癌的靶生物标志物特征可以是或包含在8种测试的情况的至少3种中区分肺癌和健康样品的如实施例10的表4中所述的组合中的靶标。在一些实施方案中,肺癌的靶生物标志物特征可以是或包含在8种测试的情况的至少2种中区分肺癌和健康样品的如实施例10的表4中所述的组合中的靶标。在一些实施方案中,肺癌的靶生物标志物特征可以是或包含在8种测试的情况的至少1种中区分肺癌和健康样品的如实施例10的表4中所述的组合中的靶标。

[0199] 在某些实施方案中,肺癌检测的靶生物标志物特征包含:TNFRSF10B生物标志物和PD-L1生物标志物;或TNFRSF10B生物标志物和CEACAM6生物标志物;或TNFRSF10B生物标志物和EGFR生物标志物;或TNFRSF10B生物标志物和IGF1R生物标志物;或ALCAM生物标志物和EPCAM生物标志物;或CEACAM6生物标志物和MUC1生物标志物;或EGFR生物标志物和T抗原生物标志物;或EPCAM生物标志物和T生物标志物;或FOLR1生物标志物和T抗原生物标志物;或Tn抗原生物标志物和TACSTD2生物标志物;或TNFRSF10B生物标志物和FOLR1生物标志物;或TNFRSF10B生物标志物和sTn抗原生物标志物;或ALCAM生物标志物和PD-L1生物标志物;或EPCAM生物标志物和MUC1生物标志物;或TNFRSF10B生物标志物和CD55生物标志物;或TNFRSF10B生物标志物和MUC1生物标志物;或FOLR1生物标志物和TACSTD2生物标志物;或MET生物标志物和MUC1生物标志物;或MET生物标志物和sTn抗原生物标志物;或MUC1生物标志物和TACSTD2生物标志物;或PD-L1生物标志物和MUC1生物标志物;或PD-L1生物标志物和Tn抗原生物标志物;或SLC34A2生物标志物和MET生物标志物;或SLC34A2生物标志物和T抗原生物标志物;或TNFRSF10B生物标志物和CEACAM5生物标志物;或TNFRSF10B生物标志物和MSLN生物标志物;或TNFRSF10B生物标志物和Tn生物标志物;或其组合。

[0200] 在某些实施方案中,肺癌检测的靶生物标志物特征包含TNFRSF10B生物标志物和PD-L1生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含TNFRSF10B生物标志物和CEACAM6生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含TNFRSF10B生物标志物和EGFR生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含TNFRSF10B生物标志物和IGF1R生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含ALCAM生物标志物和EPCAM生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含CEACAM6生物标志物和MUC1生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含EGFR生物标志物和T抗原生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含EPCAM生物标志物和T抗原生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含FOLR1生物标志物和T抗原生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含Tn抗原生物标志物和TACSTD2生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含TNFRSF10B生物标志物和FOLR1生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含TNFRSF10B生物标志物和sTn抗原生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含ALCAM生物标志物和PD-L1生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含EPCAM生物标志物和MUC1生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含TNFRSF10B生物标志物和CD55生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含TNFRSF10B生物标志物和MUC1生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含

FOLR1生物标志物和TACSTD2生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含MET生物标志物和MUC1生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含MET生物标志物和sTn抗原生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含MUC1生物标志物和TACSTD2生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含PD-L1生物标志物和MUC1生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含PD-L1生物标志物和Tn抗原生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含SLC34A2生物标志物和MET生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含SLC34A2生物标志物和T抗原生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含TNFRSF10B生物标志物和CEACAM5生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含TNFRSF10B生物标志物和MSLN生物标志物。在某些实施方案中,肺癌检测的靶生物标志物特征包含TNFRSF10B生物标志物和Tn抗原生物标志物。

[0201] 在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物的组合,所述组合可以选自以下:CYP2S1多肽和HS6ST2多肽;或ADAM23多肽和CYP2S1多肽;或ADAM23多肽和CYP4F11多肽;或ADAM23多肽和UGT1A6多肽;或ADAM23多肽和TMPRSS4多肽;或ADAM23多肽和ILDR1多肽;或DSG3多肽和UPK1B多肽;或ABCA13多肽和CYP2S1多肽;或ABCA13多肽和ADAM23多肽;或ADAM23多肽和LAMC2多肽;或ADAM23多肽和HAS3多肽;或HS6ST2多肽和LAMC2多肽;或CYP4F11多肽和HS6ST2多肽;或ADAM23多肽和ULBP2多肽;或ABCA13多肽和HS6ST2多肽;或CYP4F11多肽和LAMC2多肽;或ABCA13多肽和UPK1B多肽;或ABCA13多肽和CYP4F11多肽;或CYP2S1多肽和LAMC2多肽;或CYP2S1多肽和CYP4F11多肽;或HS6ST2多肽和PIGT多肽;或CYP4F11多肽和SHISA2多肽;或ABCA13多肽和DSG2多肽;或ADAM23多肽和LAMB3多肽;或CYP2S1多肽和VTCN1多肽;或CYP4F11多肽和HAS3多肽;或CYP4F11多肽和SLC7A11多肽;或ADAM23多肽和DSG3多肽;或ADAM23多肽和FERMT1多肽;或ADAM23多肽和HS6ST2多肽;或HS6ST2多肽和ULBP2多肽;或CYP4F11多肽和RAP2B多肽;或RACGAP1多肽和TFRC多肽;或CYP2S1多肽和ULBP2多肽;或KLRG2多肽和UPK1B多肽;或CLCA2多肽和CYP2S1多肽;或ADAM23多肽和PANX2多肽;或ABCA13多肽和HAS3多肽;或ADAM23多肽和SLC7A11多肽;或CYP4F11多肽和DSG3多肽;或ADAM23多肽和CYP4F3多肽;或CYP2S1多肽和KLRG2多肽;或HS6ST2多肽和SLC7A11多肽;或ADAM23多肽和RAP2B多肽;或CYP4F11多肽和PIGT多肽;或ADAM23多肽和SDC1多肽;或ADAM23多肽和ECE2多肽;或DSG2多肽和MARVELD2多肽;或CYP2S1多肽和SHISA2多肽;或ABCA13多肽和UGT1A6多肽;或ABCC1多肽和CYP4F11多肽;或ADAM23多肽和PIGT多肽;或CYP2S1多肽和PIGT多肽;或CYP2S1多肽和ILDR1多肽;或CYP2S1多肽和NECTIN1多肽;或CYP4F11多肽和RAB6B多肽;或CYP4F11多肽和KLRG2多肽;或ADAM23多肽和FXRD3多肽;或CLCA2多肽和CYP4F11多肽;或ABCA13多肽和DSG3多肽;或CYP4F11多肽和XXYLT1多肽;或ADAM23多肽和PRSS21多肽;或FGFBP1多肽和MARVELD3多肽;或CYP2S1多肽和HAS3多肽;或ADAM23多肽和CLCA2多肽;或CYP4F11多肽和NECTIN1多肽;或CLCA2多肽和HS6ST2多肽;或HS6ST2多肽和PRSS21多肽;或CYP2S1多肽和DSG2多肽;或CD9多肽和CYP4F11多肽;或CYP4F11多肽和FERMT1多肽;或CYP4F11多肽和TFRC多肽;或ADAM23多肽和EPCAM多肽;或DSG2多肽和FBXO45多肽;或CYP2S1多肽和PANX2多肽;或CYP2S1多肽和PRSS21多肽;或CYP4F11多肽和SDK1多肽;或CYP4F11多肽和ULBP2多肽;或ABCA13多肽和PIGT多肽;或

HS6ST2多肽和RAP2B多肽;或CYP2S1多肽和ECE2多肽;或ADAM23多肽和NECTIN1多肽;或CYP2S1多肽和SLC7A11多肽;或ECE2多肽和HS6ST2多肽;或CYP4F11多肽和SDK2多肽;或PACC1多肽和TFRC多肽;或CYP4F11多肽和KPNA2多肽;或ADAM23多肽和SLC12A8多肽;或ADAM23多肽和APO0多肽;或APO0多肽和MARVELD2多肽;或CD9多肽和PACC1多肽;或ECE2多肽和UPK1B多肽;或ILDR1多肽和MARVELD2多肽;或ITGA2多肽和TMEM158多肽;或ADAM23多肽和UCHL1多肽;或CYP4F11多肽和PANX2多肽;或CYP4F11多肽和ILDR1多肽;或CYP4F11多肽和NRCAM多肽;或ADAM23多肽和CDH1多肽;或CYP4F11多肽和ECE2多肽;或其组合。在一些实施方案中,前述组合中的靶生物标志物可以用作本文所述测定的捕获探针的靶标和/或检测探针的靶标。

[0202] 在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含CYP2S1多肽和HS6ST2多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽和CYP2S1多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽和TMPRSS4多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽和ILDR1多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ABCA13多肽和CYP2S1多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ABCA13多肽和ADAM23多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含HS6ST2多肽和LAMC2多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含CYP4F11多肽和HS6ST2多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ABCA13多肽和HS6ST2多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽和CYP4F11多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽和UGT1A6多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽和LAMC2多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽和HAS3多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽和LAMB3多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽和ULBP2多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽和DSG3多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含CYP4F11多肽和LAMC2多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含CYP4F11多肽和SLC7A11多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽和FERMT1多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽和CYP4F3多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含RACGAP1多肽

和TFRC多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含CYP4F11多肽和HAS3多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含CD9多肽和DSG3多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽和UPK1B多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽和RACGAP1多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含EPCAM多肽和LAMP3多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽和ITGA2多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含CDH1多肽和DSG3多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含CDH3多肽和DSG3多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含KPNA2多肽和ULBP2多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含TFRC多肽和ULBP2多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽和VTCN1多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ULBP2多肽和UPK1B多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽和NECTIN1多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽和PTPRZ1多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含KPNA2多肽和UPK1B多肽。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含RACGAP1多肽和ULBP2多肽。在一些实施方案中,前述组合中的靶生物标志物可以用作本文所述测定的捕获探针的靶标和/或检测探针的靶标。

[0203] 在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含TNFRSF10B和PD-L1生物标志物。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含TNFRSF10B和CEACAM6生物标志物。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含TNFRSF10B和EGFR生物标志物。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含TNFRSF10B和IGF1R生物标志物。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含ALCAM和EPCAM生物标志物。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含CEACAM6和MUC1生物标志物。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含EGFR和T抗原生物标志物。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含EPCAM和T生物标志物。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含FOLR1和T生物标志物。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含Tn抗原和TACSTD2生物标志物。在一些实施方案中,靶生物标志物特征包含至

[illegible]

[illegible]

[illegible]

物标志物特征是或包含MUC1和Tn抗原生物标志物。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含SLC34A2和EGFR生物标志物。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含TNFRSF10B和CDH1生物标志物。在一些实施方案中,靶生物标志物特征包含至少两种靶生物标志物,所述靶生物标志物特征是或包含TNFRSF10B和SLC34A2生物标志物。在一些实施方案中,前述组合中的靶生物标志物可以用作本文所述测定的捕获探针的靶标和/或检测探针的靶标。

[0204] 在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物的组合,所述组合可以选自以下:ADAM23多肽、CYP2S1多肽和LAMC2多肽;或CYP2S1多肽、HS6ST2多肽和LAMC2多肽;或CYP2S1多肽、HS6ST2多肽和PIGT多肽;或ADAM23多肽、ILDR1多肽和LAMC2多肽;或ABCA13多肽、CYP2S1多肽和DSG2多肽;或CYP2S1多肽、HS6ST2多肽和KPNA2多肽;或ABCA13多肽、ADAM23多肽和UGT1A6多肽;或CYP2S1多肽、HS6ST2多肽和ULBP2多肽;或ADAM23多肽、CYP2S1多肽和NECTIN1多肽;或ADAM23多肽、UGT1A6多肽和ULBP2多肽;或ADAM23多肽、CYP2S1多肽和VTCN1多肽;或ABCA13多肽、ADAM23多肽和CYP2S1多肽;或CLCA2多肽、CYP2S1多肽和HS6ST2多肽;或ABCA13多肽、CYP2S1多肽和UGT1A6多肽;或CYP2S1多肽、HS6ST2多肽和NECTIN1多肽;或ABCA13多肽、CYP2S1多肽和FERMT1多肽;或ABCA13多肽、DSG2多肽和HAS3多肽;或ABCA13多肽、ADAM23多肽和DSG3多肽;或ADAM23多肽、RAP2B多肽和TMPRSS4多肽;或ADAM23多肽、PIGT多肽和TMPRSS4多肽;或ADAM23多肽、TMPRSS4多肽和ULBP2多肽;或CYP2S1多肽、HAS3多肽和PIGT多肽;或CYP2S1多肽、HS6ST2多肽和XXYLT1多肽;或ADAM23多肽、ILDR1多肽和ULBP2多肽;或ADAM23多肽、LAMC2多肽和UGT1A6多肽;或ADAM23多肽、CLCA2多肽和CYP2S1多肽;或ABCA13多肽、ADAM23多肽和FERMT1多肽;或ADAM23多肽、CYP2S1多肽和SDC1多肽;或ABCA13多肽、ADAM23多肽和HAS3多肽;或ABCA13多肽、CYP2S1多肽和HAS3多肽;或ABCA13多肽、DSG2多肽和HS6ST2多肽;或CYP2S1多肽、HS6ST2多肽和SDK2多肽;或CYP2S1多肽、HS6ST2多肽和KLRG2多肽;或ADAM23多肽、CYP2S1多肽和FERMT1多肽;或ABCA13多肽、ADAM23多肽和CDH3多肽;或ABCA13多肽、CYP2S1多肽和ECE2多肽;或ADAM23多肽、CDH3多肽和EPCAM多肽;或CELSR1多肽、CYP2S1多肽和HS6ST2多肽;或ADAM23多肽、CDH3多肽和ILDR1多肽;或ADAM23多肽、CYP2S1多肽和ILDR1多肽;或ABCA13多肽、CYP2S1多肽和NECTIN1多肽;或ADAM23多肽、LAMC2多肽和TMPRSS4多肽;或ABCA13多肽、CYP2S1多肽和DSG3多肽;或ADAM23多肽、LAMP3多肽和UGT1A6多肽;或CYP2S1多肽、HS6ST2多肽和RAP2B多肽;或CYP2S1多肽、ECE2多肽和HS6ST2多肽;或ADAM23多肽、CYP2S1多肽和HAS3多肽;或ADAM23多肽、ILDR1多肽和LAMB3多肽;或ADAM23多肽、ILDR1多肽和RAP2B多肽;或CYP2S1多肽、FBX045多肽和HS6ST2多肽;或ABCA13多肽、CYP2S1多肽和HS6ST2多肽;或CYP2S1多肽、LAMC2多肽和PIGT多肽;或ADAM23多肽、ECE2多肽和ILDR1多肽;或ADAM23多肽、KLRG2多肽和TMPRSS4多肽;或ABCA13多肽、FERMT1多肽和HS6ST2多肽;或ABCA13多肽、DSG2多肽和PIGT多肽;或CDH3多肽、CYP2S1多肽和ILDR1多肽;或ADAM23多肽、ILDR1多肽和PIGT多肽;或ADAM23多肽、CYP2S1多肽和ECE2多肽;或CYP2S1多肽、HS6ST2多肽和PRSS21多肽;或HS6ST2多肽、LAMC2多肽和LSR多肽;或ADAM23多肽、CYP2S1多肽和HS6ST2多肽;或ADAM23多肽、ECE2多肽和TMPRSS4多肽;或HS6ST2多肽、LAMC2多肽和RAP2B多肽;或ADAM23多肽、HAS3多肽和ILDR1多肽;或ABCA13多肽、ADAM23多肽和NECTIN1多肽;或CYP2S1多肽、DSG3多肽和UPK1B多肽;或

ADAM23多肽、SHISA2多肽和TMPRSS4多肽；或ADAM23多肽、CYP2S1多肽和ITGA2多肽；或ADAM23多肽、KLRG2多肽和UGT1A6多肽；或CYP2S1多肽、ILDR1多肽和ULBP2多肽；或ADAM23多肽、SHISA2多肽和UGT1A6多肽；或CLCA2多肽、CYP2S1多肽和DSG2多肽；或ADAM23多肽、CYP2S1多肽和LAMB3多肽；或ADAM23多肽、NRCAM多肽和UGT1A6多肽；或CYP2S1多肽、FERMT1多肽和ULBP2多肽；或CYP2S1多肽、PIGT多肽和VTCN1多肽；或ADAM23多肽、PIGT多肽和UGT1A6多肽；或ADAM23多肽、ECE2多肽和UGT1A6多肽；或ABCA13多肽、ADAM23多肽和SDC1多肽；或ABCA13多肽、CYP2S1多肽和SDC1多肽；或ADAM23多肽、TMPRSS4多肽和XXYLT1多肽；或CYP2S1多肽、HS6ST2多肽和LAMB3多肽；或CYP2S1多肽、LAMB3多肽和ULBP2多肽；或CYP2S1多肽、UGT1A6多肽和ULBP2多肽；或ADAM23多肽、KRTCAP3多肽和LAMC2多肽；或CLCA2多肽、HS6ST2多肽和PIGT多肽；或CYP2S1多肽、HS6ST2多肽和LSR多肽；或CYP2S1多肽、HS6ST2多肽和RACGAP1多肽；或CYP2S1多肽、HS6ST2多肽和LAPTM4B多肽；或ADAM23多肽、LAMC2多肽和PRSS21多肽；或CLCA2多肽、CYP2S1多肽和ULBP2多肽；或ADAM23多肽、LAMB3多肽和LAMC2多肽；或ADAM23多肽、ILDR1多肽和XXYLT1多肽；或CYP2S1多肽、HS6ST2多肽和LARGE2多肽；或CYP2S1多肽、LAMC2多肽和ULBP2多肽；或ADAM23多肽、TMEM132A多肽和TMPRSS4多肽；或CYP2S1多肽、LAMC2多肽和SHISA2多肽；或ADAM23多肽、CYP2S1多肽和FAT2多肽；或CYP2S1多肽、CYP4F11多肽和KLRG2多肽；或其组合。在一些实施方案中，前述组合中的至少一种靶生物标志物可以用作捕获探针的靶标，并且至少两种靶生物标志物可以用作检测探针的靶标。

[0205] 在一些实施方案中，靶生物标志物特征包含至少三种靶生物标志物，所述靶生物标志物特征是或包含ADAM23多肽、CYP2S1多肽和LAMC2多肽。在一些实施方案中，靶生物标志物特征包含至少三种靶生物标志物，所述靶生物标志物特征是或包含CYP2S1多肽、HS6ST2多肽和LAMC2多肽。在一些实施方案中，靶生物标志物特征包含至少三种靶生物标志物，所述靶生物标志物特征是或包含CYP2S1多肽、HS6ST2多肽和PIGT多肽。在一些实施方案中，靶生物标志物特征包含至少三种靶生物标志物，所述靶生物标志物特征是或包含ADAM23多肽、ILDR1多肽和LAMC2多肽。在一些实施方案中，靶生物标志物特征包含至少三种靶生物标志物，所述靶生物标志物特征是或包含ABCA13多肽、CYP2S1多肽和DSG2多肽。在一些实施方案中，靶生物标志物特征包含至少三种靶生物标志物，所述靶生物标志物特征是或包含CYP2S1多肽、HS6ST2多肽和KPNA2多肽。在一些实施方案中，靶生物标志物特征包含至少三种靶生物标志物，所述靶生物标志物特征是或包含ABCA13多肽、ADAM23多肽和UGT1A6多肽。在一些实施方案中，靶生物标志物特征包含至少三种靶生物标志物，所述靶生物标志物特征是或包含CYP2S1多肽、HS6ST2多肽和ULBP2多肽。在一些实施方案中，靶生物标志物特征包含至少三种靶生物标志物，所述靶生物标志物特征是或包含ADAM23多肽、CYP2S1多肽和NECTIN1多肽。在一些实施方案中，靶生物标志物特征包含至少三种靶生物标志物，所述靶生物标志物特征是或包含ADAM23多肽、UGT1A6多肽和ULBP2多肽。在一些实施方案中，靶生物标志物特征包含至少三种靶生物标志物，所述靶生物标志物特征是或包含ADAM23多肽、LAMC2多肽和UGT1A6多肽。在一些实施方案中，靶生物标志物特征包含至少三种靶生物标志物，所述靶生物标志物特征是或包含ADAM23多肽、CDH3多肽和EPCAM多肽。在一些实施方案中，靶生物标志物特征包含至少三种靶生物标志物，所述靶生物标志物特征是或包含ADAM23多肽、LAMP3多肽和UGT1A6多肽。在一些实施方案中，靶生物标志物特征包

含至少三种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽、NRCAM多肽和UGT1A6多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽、PIGT多肽和UGT1A6多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽、ECE2多肽和UGT1A6多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽、LAMC2多肽和PRSS21多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽、LAMB3多肽和LAMC2多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽、FXYP3多肽和ULBP2多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含ADAM23多肽、RAP2B多肽和UGT1A6多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽、EPCAM多肽和UPK1B多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽、KPNA2多肽和UPK1B多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽、LAMP3多肽和UPK1B多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽、ULBP2多肽和UPK1B多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽、RAP2B多肽和UPK1B多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽、LMNB2多肽和UPK1B多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含ABCC1多肽、DSG3多肽和UPK1B多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽、TFRC多肽和UPK1B多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含CD9多肽、DSG3多肽和UPK1B多肽。在一些实施方案中,靶生物标志物特征包含至少三种靶生物标志物,所述靶生物标志物特征是或包含DSG3多肽、LAMB3多肽和UPK1B多肽。在一些实施方案中,前述组合中的至少一种靶生物标志物可以用作捕获探针的靶标,并且至少两种靶生物标志物可以用作检测探针的靶标。

[0206] 在某些实施方案中,肺癌检测的靶生物标志物特征包含至少三种靶生物标志物的组合,所述组合可以选自以下:CEACAM6多肽、MUC1多肽和sTn抗原多肽;或TNFRSF10B多肽、MUC1多肽和CEACAM6多肽;或SLC34A2多肽、MUC1多肽和CEACAM6多肽;或CEACAM6多肽、MUC1多肽和MSLN多肽;或FOLR1多肽、T抗原多肽和EGFR多肽;或DSG2多肽、MUC1多肽和CEACAM6多肽;或其组合。在一些实施方案中,前述组合中的至少一种靶生物标志物可以用作捕获探针的靶标,并且至少两种靶生物标志物可以用作检测探针的靶标。

[0207] 在某些实施方案中,肺癌检测的靶生物标志物特征包含CEACAM6多肽、MUC1和sTn抗原多肽。在某些实施方案中,肺癌检测的靶多肽特征包含TNFRSF10B多肽、MUC1多肽和CEACAM6多肽。在某些实施方案中,肺癌检测的靶多肽特征包含SLC34A2多肽、MUC1多肽和CEACAM6多肽。在某些实施方案中,肺癌检测的靶多肽特征包含CEACAM6多肽、MUC1多肽和MSLN多肽。在某些实施方案中,肺癌检测的靶多肽特征包含FOLR1多肽、T抗原多肽和EGFR多

肽。在某些实施方案中,肺癌检测的靶多肽特征包含DSG2多肽、MUC1多肽和CEACAM6多肽。在一些实施方案中,前述组合中的至少一种靶生物标志物可以用作捕获探针的靶标,并且至少两种靶生物标志物可以用作检测探针的靶标。

[0208] III. 检测提供的肺癌的标志物和/或靶生物标志物特征的示例性方法

[0209] 一般来讲,本公开提供了分析和/或评估在来自有需要的受试者的包含细胞外囊泡的血液来源样品中的靶生物标志物特征所根据的技术;在一些实施方案中,基于此类分析和/或评估做出诊断或治疗决定。

[0210] 在一些实施方案中,检测靶生物标志物特征的方法包括用于检测作为蛋白质的靶生物标志物特征的一种或多种提供的标志物的方法。检测一种或多种提供的标志物的示例性基于蛋白质的方法包括但不限于邻近连接测定、质谱(MS)和免疫测定,诸如免疫沉淀;蛋白质印迹;ELISA;免疫组化;免疫细胞化学;流式细胞术;和免疫PCR。在一些实施方案中,免疫测定可以是化学发光免疫测定。在一些实施方案中,免疫测定可以是高通量和/或自动化免疫测定平台。

[0211] 在一些实施方案中,在样品中检测作为蛋白质的一种或多种提供的标志物的方法包括使样品与针对关注的提供的标志物的一种或多种抗体剂接触。在一些实施方案中,此类方法还包括使样品与一种或多种检测标记接触。在一些实施方案中,抗体剂用一种或多种检测标记进行标记。

[0212] 在一些实施方案中,检测关注的生物标志物和针对关注的生物标志物的抗体剂之间的结合包括确定一种或多种检测剂的吸光度值或发射值。例如,吸光度值或发射值指示由细胞外囊泡表达的关注的生物标志物的量和/或浓度(例如,较高的吸光度指示细胞外囊泡表达的关注的生物标志物的水平较高)。在一些实施方案中,检测剂的吸光度值或发射值高于阈值。在一些实施方案中,检测剂的吸光度值或发射值为阈值的至少1.3倍、至少1.4倍、至少1.5倍、至少1.6倍、至少1.7倍、至少1.8倍、至少1.9倍、至少2.0倍、至少2.5倍、至少3.0倍、至少3.5倍或更大。在一些实施方案中,阈值是在对照组或参考组(例如,非癌症受试者)群体中确定的。

[0213] 在一些实施方案中,检测一种或多种提供的标志物的方法包括用于检测作为核酸的一种或多种提供的标志物的方法。检测一种或多种提供的标志物的示例性基于核酸的方法包括但不限于进行核酸扩增方法,诸如聚合酶链反应(PCR)、逆转录聚合酶链反应(RT-PCR)、转录介导的扩增(TMA)、连接酶链反应(LCR)、链置换扩增(SDA)和基于核酸序列的扩增(NASBA)。在一些实施方案中,检测一种或多种提供的标志物的基于核酸的方法包括检测一个或多个核酸探针与编码关注的生物标志物的一个或多个核苷酸之间的杂交。在一些实施方案中,核酸探针每一个都与编码关注的生物标志物的一个或多个核苷酸中的一个的至少一部分互补。在一些实施方案中,编码关注的生物标志物的核苷酸包括DNA(例如,cDNA)。在一些实施方案中,编码关注的生物标志物的核苷酸包括RNA(例如,mRNA)。

[0214] 在一些实施方案中,检测一种或多种提供的标志物的方法涉及邻近连接免疫定量聚合酶链反应(pliq-PCR)。与其他技术相比,pliq-PCR在分析EV方面可以具有某些优点。例如,pliq-PCR可以具有比其他标准免疫测定(诸如ELISA)高三个数量级的灵敏度(Darmanis等人,2010;其出于本文所述目的以引用的方式并入本文)。在一些实施方案中,pliq-PCR反应可以被设计成具有超低LOD,这使得能够检测到痕量水平的肿瘤来源的EV,例如低至

1000EV/mL。

[0215] 在一些实施方案中,用于检测一个或多个提供的标志物的方法可以涉及用于检测EV的其他技术,包括例如纳米胞质外泌体(nPLEX)传感器(Im等人,2014;其出于本文所述目的以引用的方式并入本文)和集成磁-电化学外泌体(iMEX)传感器(Jeong等人,2016;其出于本文所述目的以引用的方式并入本文),所述技术分别报告了约 10^3 的LOD和约 10^4 的EV(Shao等人,2018;其出于本文所述目的以引用的方式并入本文)。

[0216] 在一些实施方案中,用于检测细胞外囊泡中的一种或多种提供的生物标志物的方法可以基于批量EV样品分析。

[0217] 在一些实施方案中,用于检测细胞外囊泡中一种或多种提供的生物标志物的方法可以基于分析个体EV(例如,单个EV谱测定),这在下文标题为“用于分析个体细胞外囊泡(EV)的示例性方法”的部分中进一步讨论。

[0218] 在一些实施方案中,在检测根据本公开的一种或多种提供的生物标志物之前,样品中的细胞外囊泡可以捕获或固定在固体基底上。在一些实施方案中,细胞外囊泡可以通过非特异性相互作用(包括例如吸附)被捕获在固体基底表面上。在一些实施方案中,细胞外囊泡可以被选择性地捕获在固体基底表面上。例如,在一些实施方案中,固体基底表面可以涂覆有与细胞外囊泡特异性结合的剂(例如,特异性靶向细胞外囊泡(例如,与肺癌相关)的抗体剂)。在一些实施方案中,固体基底表面可以涂覆有亲和结合对的成员,并且待捕获的关注实体(例如,细胞外囊泡)可以缀合至亲和结合对的互补成员。在一些实施方案中,示例性亲和结合对包括但不限于例如生物素和抗生物素蛋白样分子,诸如链霉亲和素。如本领域技术人员将理解的,其他适当的亲和结合对也可以用于促进将关注的实体捕获至固体基底表面上。在一些实施方案中,关注的实体可以通过施加电流而捕获至固体基底表面上,例如,如描述于Ibsen等人ACS Nano., 11:6641-6651(2017)和Lewis等人ACS Nano., 12:3311-3320(2018),两者出于本文所述目的以引用的方式并入本文,并且两者都描述了从未稀释的人血液或血浆样品中分离细胞外囊泡的交流电动微阵列芯片装置的用途。

[0219] 固体基底可以以适用于捕获细胞外囊泡并且不干扰下游处理、加工和/或检测的形式提供。例如,在一些实施方案中,固体基底可以是或包含珠(例如,磁珠)。在一些实施方案中,固体基底可以是或包含表面。例如,在一些实施方案中,此类表面可以是测定室(包括例如管、孔、微孔、板、过滤器、膜、基质等)的捕获表面。因此,在一些实施方案中,本文所述的方法包括在检测样品中提供的生物标志物之前,将细胞外囊泡捕获或固定到固体基底上。

[0220] 在一些实施方案中,在将细胞外囊泡捕获至固体基底表面上之前,可以处理样品,以例如去除不需要的实体,诸如细胞碎片或细胞。例如,在一些实施方案中,可以对此类样品进行离心,以例如去除细胞碎片、细胞和/或其他颗粒。另外或替代地,在一些实施方案中,可以对此类样品进行基于尺寸排阻的纯化或过滤。各种基于尺寸排阻的纯化或过滤是本领域已知的,并且本领域技术人员将理解,在一些情况下,可以基于特定分子量或粒径截止值对样品进行离心柱纯化。本领域技术人员还将理解,可以例如基于细胞外囊泡的大小来选择用于纯化目的的适当分子量或粒径截止值。例如,在一些实施方案中,尺寸排阻分离方法可以应用于包含细胞外囊泡的样品以分离具有某一尺寸(例如,大于30nm且不超过1000nm或大于70nm且不超过200nm)的细胞外囊泡的一部分。通常,细胞外囊泡的直径可以

在30nm至几微米的范围内。参见例如出于本文所述目的以引用的方式并入本文的Chuo等人,“Imaging extracellular vesicles:current and emerging methods”Journal of Biomedical Sciences 25:91(2018),其提供不同细胞外囊泡(EV)亚型的大小信息:迁移体(0.5-3 μ m)、微泡(0.1-1 μ m)、癌小体(1-10 μ m)、外泌颗粒(<50nm)、小外泌体(60-80nm)和大外泌体(90-120nm)。在一些实施方案中,大小范围为约30nm至1000nm的纳米粒子可以被分离(例如,在一些实施方案中,通过一种或多种尺寸排阻分离方法)用于检测测定。在一些实施方案中,一种或多种特定EV亚型可以被分离(例如,在一些实施方案中,通过一种或多种尺寸排阻分离方法)用于检测测定。

[0221] 在一些实施方案中,可以在检测肺癌的靶生物标志物特征中的一种或多种提供的生物标志物之前处理样品中的细胞外囊泡。可以进行不同的样品处理和/或制备以例如稳定待检测的细胞外囊泡中的靶标(例如,靶生物标志物),和/或促进靶标(例如,囊泡内蛋白质和/或RNA,诸如mRNA)暴露于检测测定(例如,如本文所述),和/或减少非特异性结合。此类样品处理和/或制备的实例是本领域已知的,并且包括但不限于交联分子靶标(例如,固定)、使生物实体(例如,细胞或细胞外囊泡)透化和/或封闭非特异性结合位点。

[0222] 在一个方面,本公开提供了一种用于检测肺癌的靶生物标志物特征是否存在于来自有需要的受试者的生物样品中的方法,所述生物样品在一些实施方案中可以是包含细胞外囊泡的血液来源样品。在一些实施方案中,此类方法包括(a)在生物样品诸如来自受试者的血液来源样品(例如,血浆样品)中检测表达肺癌的靶生物标志物特征的关注的生物实体(包括例如细胞外囊泡);以及(b)将指示生物样品(例如,血液来源样品)中表达靶生物标志物特征的关注的生物实体(例如,细胞外囊泡)的水平的样品信息与包括参考阈值水平的参考信息进行比较。在一些实施方案中,参考阈值水平对应于在来自参考受试者(例如,非癌症受试者)群体的可比样品中表达此类靶生物标志物特征的关注的生物实体(例如,细胞外囊泡)的水平。在一些实施方案中,示例性非癌症受试者包括健康受试者(例如,特定年龄范围(例如像低于55岁或高于55岁)的健康受试者)、患有非肺相关健康疾病、病症或病状的受试者(包括,例如,具有非肺癌诸如卵巢癌、结肠直肠癌等的受试者,或具有慢性阻塞性肺部疾病或慢性肺部感染的受试者)、具有良性肺肿瘤(例如,通过成像(诸如胸部X射线或低剂量CT扫描)在肺中观察到良性肿块)的受试者,及其组合。

[0223] 在一些实施方案中,在用于本文所述的测定之前,预先筛查样品的某些特征。在一些实施方案中,符合某些预筛查标准的样品比不符合预筛查标准的样品更适用于诊断应用。例如,在一些实施方案中,使用已知标准目视检查样品外观,例如,样品是否正常、溶血(红色)、黄疸(黄色)和/或脂血(混浊)。在一些实施方案中,然后可按照已知的标准等级(例如,1、2、3、4、5)对样品进行分级并记录结果。在一些实施方案中,目视检查样品的溶血(例如,血红素)并以1-5的等级分级,其中目视检查与已知浓度相关,例如,其中1表示大约0mg/dL,2表示大约50mg/dL,3表示大约150mg/dL,4表示大约250mg/dL,5表示大约525mg/dL。在一些实施方案中,目视检查样品的黄疸水平(例如,胆红素)并以1-5的等级分级,其中目视检查与已知浓度相关,例如,其中1表示大约0mg/dL,2表示大约1.7mg/dL,3表示大约6.6mg/dL,4表示大约16mg/dL,5表示大约30mg/dL。在一些实施方案中,目视检查样品的浊度(例如脂血)并以1-5的等级评定,其中目视检查与已知浓度相关,例如,其中1表示大约0mg/dL,2表示大约125mg/dL,3表示大约250mg/dL,4表示大约500mg/dL,5表示大约1000mg/dL。

[0224] 在一些实施方案中,在一个或多个量度上得分低于某个水平的样品(例如,等于或低于4分)可用于如本文所述的测定中。在一些实施方案中,在一个或多个量度上得分低于某个水平的样品(例如,等于或低于3分)可用于如本文所述的测定中。在一些实施方案中,在一个或多个量度上得分低于某个水平的样品(例如,等于或低于2分)可用于如本文所述的测定中。在一些实施方案中,在所有三个量度(例如,溶血、黄疸、脂血)上得分低于某个水平的样品(例如,等于或低于2分)可用于如本文所述的测定中。在一些实施方案中,预筛查标准诸如溶血、胆红素和/或脂血的低目视检查评分(例如,等于或低于2分)可能对在如本文所述的测定中产生的诊断特性(例如,Ct值)没有明显影响(例如,不相关)。

[0225] 在一些实施方案中,当样品示出相对于参考阈值水平升高水平的靶生物标志物特征表达细胞外囊泡(例如,本文所述的那些)时,确定样品具有靶生物标志物特征表达细胞外囊泡(例如,本文所述的那些)。在一些实施方案中,如与参考阈值水平相比,如果靶生物标志物特征表达细胞外囊泡的水平为至少30%或更高,包括例如至少40%、至少50%、至少60%、至少70%、至少80%、至少90%、至少95%或更高,则确定样品对靶生物标志物特征表达细胞外囊泡呈阳性。在一些实施方案中,如与参考阈值水平相比,如果靶生物标志物特征表达细胞外囊泡的水平为至少2倍或更高,包括例如至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少50倍、至少100倍、至少250倍、至少500倍、至少750倍、至少1000倍、至少2500倍、至少5000倍或更高,则确定样品对靶生物标志物特征表达细胞外囊泡呈阳性。

[0226] 在一些实施方案中,二元分类系统可用于确定样品是否对靶生物标志物特征表达细胞外囊泡呈阳性。例如,在一些实施方案中,如果靶生物标志物特征表达细胞外囊泡的水平处于或高于参考阈值水平(例如截止值),则确定样品对靶生物标志物特征表达细胞外囊泡呈阳性。在一些实施方案中,此类参考阈值水平(例如,截止值)可以通过选择偏离从对照受试者获得的平均值一定量的标准偏差来确定,使得可以实现肺癌检测测定(例如,本文所述的那些)的期望灵敏度和/或特异性。在一些实施方案中,此类参考阈值水平(例如,截止值)可以通过选择偏离从对照受试者获得的最大测定信号一定量的标准偏差来确定,使得可以实现肺癌检测测定(例如,本文所述的那些)的期望灵敏度和/或特异性。在一些实施方案中,此类参考阈值水平(例如,截止值)可以通过选择以下限制性较小的一个来确定:(i) 偏离从对照受试者获得的平均值的一定量的标准偏差,或(ii) 偏离从对照受试者获得的最大测定信号相差一定量的标准偏差,使得可以实现肺癌检测测定(例如,本文所述的那些)的期望灵敏度和/或特异性。在一些实施方案中,用于确定参考阈值水平(例如,截止值)的对照受试者可以包括但不限于健康受试者、患有炎性病状(例如,慢性阻塞性肺部疾病(COPD))的受试者、患有良性肺肿瘤的受试者及其组合。在一些实施方案中,健康受试者和患有炎性病状(例如,COPD)的受试者包括在参考阈值水平(例如,截止值)的确定中。在一些实施方案中,患有良性肺肿瘤的受试者不包括在参考阈值水平(例如,截止值)的确定中。在一些实施方案中,参考阈值(例如,截止值)可以通过选择偏离(i) 从对照受试者获得的平均值,或(ii) 从对照受试者获得的最大测定信号至少1.5个标准偏差(SD)或更高(包括,例如,至少1.6、至少1.7、至少1.8、至少1.9、至少2、至少2.1、至少2.2、至少2.3、至少2.4、至少2.5、至少2.6、至少2.7、至少2.8、至少2.9、至少3、至少3.1、至少至少3.2、至少3.3、至少3.4、至少3.5、至少3.6个或更高的SD)来确定,使得可以实现肺癌检测测定(例如,本文所述

的那些)的期望特异性(例如,至少95%或更高的特异性[包括,例如,至少96%、至少97%、至少98%、至少99%或更高的特异性],诸如在一些实施方案中至少99.8%的特异性)。在一些实施方案中,参考阈值水平(例如,截止值)可以通过选择偏离(i)从对照受试者获得的平均值,或(ii)从对照受试者获得的最大测定信号至少2.9个SD(例如,至少2.93个SD)来确定,使得可以实现肺癌检测测定(例如,本文所述的那些)的期望特异性(例如,至少99%或更高的特异性)。在一些实施方案中,参考阈值水平(例如,截止值)可以通过选择偏离(i)从对照受试者获得的平均值,或(ii)从对照受试者获得的最大测定信号中限制性较小的一个至少2.9个SD(例如,至少2.93个SD)来确定,使得可以实现肺癌检测测定(例如,本文所述的那些)的期望特异性(例如,至少99%或更高的特异性)。在一些实施方案中,此类参考阈值水平(例如,截止值)可以基于正常健康组织中与肺癌样品中的靶生物标志物的表达水平(例如,转录物水平)来确定,使得可以实现关注的特异性和/或灵敏度(例如,如本文所述)。在一些实施方案中,参考阈值水平(例如,截止值)可以根据例如肺癌分期和/或亚型和/或患者特征(例如,患者年龄、肺癌的风险因素(例如,遗传风险与平均风险、生活史相关风险因素)、有症状/无症状状态)及其组合而变化。

[0227] 在一些实施方案中,此类参考阈值水平(例如,截止值)可以基于健康受试者(例如,特定年龄范围的)和任选的患有炎性病状(例如,COPD)的受试者的对数正态分布和实现关注的特异性(例如,至少95%或更高的特异性[包括例如至少96%、至少97%、至少98%、至少99%或更高的特异性],诸如在一些实施方案中为至少99.8%的特异性)所需的标准偏差(SD)的量的选择(例如,至少1.5、至少1.6、至少1.7、至少1.8、至少1.9、至少2、至少2.1、至少2.2、至少2.3、至少2.4、至少2.5、至少2.6、至少2.7、至少2.8、至少2.9、至少3、至少3.1、至少3.2、至少3.3、至少3.4、至少3.5、至少3.6个或更高的SD)来确定,例如,基于肺癌或其亚型患病率。

[0228] 本公开尤其还提供了用于确定受试者是否患有或易患肺癌的技术。例如,在一些实施方案中,当来自有需要的受试者的血液来源样品示出处于或高于参考阈值水平(例如,截止值(例如,如根据本公开确定的))的靶生物标志物特征表达细胞外囊泡的水平时,受试者被分类为患有或易患肺癌。在一些此类实施方案中,参考阈值水平(例如,截止值)可以基于健康受试者(例如,特定年龄范围的)和任选的患有炎性病状(例如,COPD)的受试者的对数正态分布和实现关注的特异性(例如,至少95%或更高的特异性[包括例如至少96%、至少97%、至少98%、至少99%或更高的特异性],诸如在一些实施方案中为至少99.8%的特异性)所需的标准偏差(SD)的量的选择(例如,至少1.5、至少1.6、至少1.7、至少1.8、至少1.9、至少2、至少2.1、至少2.2、至少2.3、至少2.4、至少2.5、至少2.6、至少2.7、至少2.8、至少2.9、至少3、至少3.1、至少3.2、至少3.3、至少3.4、至少3.5、至少3.6个或更高的SD)来确定,例如,基于肺癌或其亚型患病率。在一些此类实施方案中,参考阈值水平(例如,截止值)可以基于正常健康组织中与肺癌样品中的靶生物标志物特征中的一种或多种个体靶生物标志物的表达水平(例如,转录物水平)来确定,使得可以实现关注的特异性和/或灵敏度(例如,如本文所述)。在一些实施方案中,参考阈值水平(例如,截止值)可以根据例如肺癌分期和/或亚型和/或患者特征(例如,患者年龄、肺癌的风险因素(例如,遗传风险与平均风险、生活史相关风险因素)、有症状/无症状状态)及其组合而变化。

[0229] 在一些实施方案中,当来自有需要的受试者的血液来源样品示出相对于参考阈值

水平升高的靶生物标志物特征表达细胞外囊泡水平时,受试者被分类为患有或易患肺癌。在一些实施方案中,当有需要的受试者(男性/女性)的血液来源样品示出与参考阈值水平相比至少30%或更高,包括例如至少40%、至少50%、至少60%、至少70%、至少80%、至少90%、至少95%或更高的靶生物标志物特征表达细胞外囊泡水平时,有需要的受试者被分类为患有或易患肺癌。在一些实施方案中,当有需要的受试者(男性/女性)的血液来源样品示出与参考阈值水平相比至少2倍或更高,包括例如至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少20倍、至少30倍、至少40倍、至少50倍、至少60倍、至少70倍、至少80倍、至少90倍、至少100倍、至少250倍、至少500倍、至少750倍、至少1000倍或更高的靶生物标志物特征表达细胞外囊泡水平时,有需要的受试者被分类为患有或易患肺癌。当来自有需要的受试者的血液来源样品示出与参考阈值水平可比的水平(例如,在10%至20%内)时,有需要的受试者被分类为患有或易患肺癌。在一些此类实施方案中,参考阈值水平对应于在来自参考受试者(例如,非癌症受试者)群体的可比样品中靶生物标志物特征表达细胞外囊泡的水平。在一些实施方案中,示例性非癌症受试者包括健康受试者(例如,特定年龄范围(例如像低于55岁或高于55岁)的健康受试者)、患有非肺相关健康疾病、病症或病状的受试者(包括,例如,具有非肺癌诸如胰腺癌、结肠直肠癌等的受试者,或具有慢性阻塞性肺部疾病或慢性肺部感染的受试者)、具有良性肺肿瘤(例如,通过成像(诸如胸部X射线或低剂量CT扫描)在肺中观察到良性肿块)的受试者,及其组合。

[0230] IV. 用于分析个体细胞外囊泡(EV)的示例性方法

[0231] 在一些实施方案中,用于分析个体细胞外囊泡的测定(例如,单个EV谱测定)可以用于检测肺癌的一种或多种靶生物标志物特征中的一个或多个提供的生物标志物。例如,在一些实施方案中,此类测定可以涉及(i)通过靶向肺癌的靶生物标志物特征的一种或多种提供的标志物的捕获测定和(ii)此类肺癌的靶生物标志物特征的至少一种或多种其他提供的标志物的检测测定,其中此类捕获测定在此类检测测定之前进行。

[0232] 在一些实施方案中,进行捕获测定以从有需要的受试者的血液或血液来源样品(例如,血浆样品)中选择性地捕获肿瘤相关细胞外囊泡(例如,肺肿瘤相关细胞外囊泡)。在一些实施方案中,进行捕获测定以选择性捕获具有一定大小范围和/或一个或多个一定特征的细胞外囊泡,例如与肺癌相关的细胞外囊泡。在一些此类实施方案中,在捕获测定之前,可以对血液或血液来源样品进行预处理以去除非细胞外囊泡,包括例如但不限于可溶性蛋白质和干扰实体,例如像细胞碎片。例如,在一些实施方案中,使用尺寸排阻色谱法从受试者的血液或血液来源样品中纯化细胞外囊泡。在一些此类实施方案中,可以使用尺寸排阻色谱法从血液或血液来源样品中直接纯化细胞外囊泡,这在一些实施方案中可以去除至少90%或更高(包括例如至少93%、95%、97%、99%或更高)的可溶性蛋白质和其他干扰剂,例如像细胞碎片。

[0233] 在一些实施方案中,捕获测定包括使血液或血液来源样品与至少一种捕获剂接触的步骤,所述捕获剂包含与肺癌的靶生物标志物特征中的至少一种或多种提供的生物标志物结合的靶捕获部分。在一些实施方案中,可以多路复用捕获测定,其包括使血液或血液来源样品与一组捕获剂接触的步骤,每种捕获剂包含与肺癌的靶生物标志物特征的不同提供的生物标志物结合的靶捕获部分。在一些实施方案中,靶捕获部分针对细胞外囊泡相关膜结合多肽(例如,如本文所述和/或使用的那些)。

[0234] 在一些实施方案中,此类靶捕获部分可以固定在固体基底上。因此,在一些实施方案中,用于捕获测定的捕获剂是或包含固体基底,所述固体基底包含至少一个或多个(例如,1、2、3、4、5个或更多个)与其缀合的靶捕获部分,每个靶捕获部分针对细胞外囊泡相关膜结合多肽(例如,如本文所述和/或使用的那些)。固体基底可以以适合于捕获细胞外囊泡并且不干扰下游处理、加工和/或检测的形式提供。例如,在一些实施方案中,固体基底可以是或包含珠(例如,磁珠)。在一些实施方案中,固体基底可以是或包含表面。例如,在一些实施方案中,此类表面可以是测定室(包括例如管、孔、微孔、板、过滤器、膜、基质等)的捕获表面。在一些实施方案中,捕获剂是或包含磁珠,所述磁珠包含与其缀合的靶捕获部分。

[0235] 在一些实施方案中,进行检测测定以检测通过捕获测定捕获的细胞外囊泡(例如,如上所述)中肺癌的靶生物标志物特征的一种或多种提供的生物标志物(例如,与捕获测定中靶向的生物标志物不同的生物标志物)。在一些实施方案中,检测测定包括免疫PCR。在一些实施方案中,免疫-PCR可以涉及靶向肺癌的靶生物标志物特征的单个提供的生物标志物(例如,本文所述的那些)的至少一个探针。在一些实施方案中,免疫PCR可以涉及针对靶生物标志物特征的相同生物标志物(例如,本文所述的那些)的不同表位的多个(例如,至少两个、至少三个、至少四个或更多个)探针。在一些实施方案中,免疫PCR可以涉及多个(例如,至少两个、至少三个、至少四个或更多个)探针,每个探针针对本文所述的不同提供的生物标志物。

[0236] 在一些实施方案中,检测测定包括逆转录聚合酶链反应(RT-PCR)。在一些实施方案中,RT-PCR可以涉及靶向本文所述的单个提供的生物标志物的至少一个引物/探针组。在一些实施方案中,RT-PCR可以涉及多个(例如,至少两个、至少三个、至少四个或更多个)引物/探针组,每组针对本文所述的不同提供的生物标志物。

[0237] 在一些实施方案中,检测测定包括邻近连接免疫定量聚合酶链反应(pliq-PCR),以例如确定细胞外囊泡(例如,表达至少一个细胞外囊泡相关膜结合多肽的捕获的细胞外囊泡)内肺癌的靶生物标志物特征的一种或多种提供的生物标志物的共定位。

[0238] 在一些实施方案中,检测测定采用由申请人开发并在美国申请号16/805,637和国际申请PCT/US2020/020529中描述的靶实体检测系统,所述申请均于2020年2月28日提交,并且标题为“Systems, Compositions, and Methods for Target Entity Detection”(‘637申请和’529申请;两者均以引用的方式整体并入本文),所述系统部分基于个体细胞外囊泡中的靶生物标志物特征的相互作用和/或共定位。例如,此类靶实体检测系统(如’637申请和’529申请中所述,并且还在下面标题为“提供的靶实体检测系统和涉及相同靶实体检测系统的方法”的部分中进一步描述)可以在一些实施方案中,在单个实体水平上检测样品中(例如,生物、环境或其他样品中)包含至少一个或多个(例如,至少两个或更多个)靶标(例如,分子靶标)的关注的实体(例如,关注的生物或化学实体,诸如细胞外囊泡或分析物)。阅读本公开的本领域技术人员将认识到,提供的靶实体检测系统可以用于很多种应用和/或目的,包括例如用于检测肺癌。例如,在一些实施方案中,提供的靶实体检测系统可以用于医学应用和/或目的。在一些实施方案中,提供的靶实体检测系统可以用于筛查(例如,定期筛查)个体(例如,在一些实施方案中可以是无症状个体,或在一些实施方案中可以是经历与肺癌相关的一种或多种症状的个体或在一些实施方案中可以是处于肺癌风险的个体,例如像具有肺癌遗传风险和/或生活史相关风险因素的个体,或吸烟个体)的疾病或病状(例

如,肺癌)。在一些实施方案中,提供的靶实体检测系统可以用于筛查(例如,定期筛查)个体(例如,在一些实施方案中可以是无症状个体,或在一些实施方案中可以是经历与肺癌相关的一种或多种症状的个体或在一些实施方案中可以是处于肺癌风险的个体,例如像具有肺癌遗传风险和/或生活史相关风险因素的个体,或吸烟个体)的不同类型的癌症(例如,多种不同的癌症,其中一种可以是肺癌)。在一些实施方案中,提供的靶实体检测系统即使在应用于包含无症状个体或由无症状个体组成的群体时也是有效的(例如,由于足够高的灵敏度和/或低的假阳性和/或假阴性结果率)。在一些实施方案中,提供的靶实体检测系统可以用作与疾病治疗(例如,肺癌的治疗)结合的伴随诊断。

[0239] 在一些实施方案中,可以进行多次(例如,至少两次或更多次)检测测定以检测细胞外囊泡(例如,通过捕获测定捕获的那些(例如,如上所述))中肺癌的一个或多个靶生物标志物特征的多种生物标志物(例如,至少两种或更多种)(例如,与捕获测定中靶向的生物标志物不同的生物标志物)。例如,在一些实施方案中,多此检测测定包括(i)提供的靶实体检测系统或在'637申请和'529申请中描述的系统;和(ii)免疫PCR。在一些实施方案中,多次检测测定包括(i)提供的靶实体检测系统或在'637申请和'529申请中描述的系统;和(ii)RT-PCR。

[0240] V.提供的靶实体检测系统和涉及相同靶实体检测系统的方法

[0241] 在一些实施方案中,可以用于肺癌的一个或多个靶生物标志物特征的一种或多种提供的生物标志物的检测测定的靶实体检测系统包含多个检测探针,每个检测探针针对特定靶标(例如,提供的靶生物标志物特征的生物标志物)。在一些实施方案中,此类系统可以包含至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少15个、至少20个、至少25个、至少30个、至少40个、至少50个或更多个检测探针,每个检测探针针对特定靶标(例如,提供的靶生物标志物特征的生物标志物)。在一些实施方案中,此类系统可以包含2-50个检测探针,每个检测探针针对特定靶标(例如,提供的靶生物标志物特征的生物标志物)。在一些实施方案中,此类系统可以包含2-30个检测探针,每个检测探针针对特定靶标(例如,提供的靶生物标志物特征的生物标志物)。在一些实施方案中,此类系统可以包含2-25个检测探针,每个检测探针针对特定靶标(例如,提供的靶生物标志物特征的生物标志物)。在一些实施方案中,此类系统可以包含5-30个检测探针,每个检测探针针对特定靶标(例如,提供的靶生物标志物特征的生物标志物)。在一些实施方案中,此类系统可以包含5-25个检测探针,每个检测探针针对特定靶标(例如,提供的靶生物标志物特征的生物标志物)。在一些实施方案中,一组此类检测探针中的至少两个可以针对靶生物标志物特征的相同生物标志物。在一些实施方案中,一组此类检测探针中的至少两个可以针对靶生物标志物特征的相同生物标志物的相同表位。在一些实施方案中,一组此类检测探针中的至少两个可以针对靶生物标志物特征的相同生物标志物的不同表位。

[0242] 在一些实施方案中,适合在本文提供的靶实体检测系统中使用的检测探针可以用于检测单个疾病或病状,例如肺癌。在一些实施方案中,适合在本文提供的靶实体检测系统中使用的检测探针可以允许检测至少两种或更多种疾病或病状,例如,其中一种是肺癌。在一些实施方案中,适合在本文提供的靶实体检测系统中使用的检测探针可以允许检测某些亚型的肺癌,包括例如肺腺癌肺癌、小细胞肺癌、鳞状细胞和移行细胞肺癌、大细胞肺癌、非小细胞肺癌、其他特定的肺癌、肉瘤型肺癌和其他本领域已知的特定类型的肺癌(SEER癌症

统计回顾1975-2017)。在一些实施方案中,适合在本文提供的靶实体检测系统中使用的检测探针可以允许检测某些分期(包括例如I期、II期、III期和/或IV期)的肺癌。因此,在一些实施方案中,适合在本文提供的靶实体检测系统中使用的检测探针包含多(例如,至少2、至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10或更多)组检测探针,其中每组针对不同疾病或不同类型的疾病或病状的检测。例如,在一些实施方案中,适合在本文提供的靶实体检测系统中使用的检测探针包含多(例如,至少2、至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10或更多)组检测探针,其中在一些实施方案中,每组针对不同类型的癌症(其中一种为肺癌)的检测,或在一些实施方案中,每组针对不同亚型和/或分期的肺癌的检测。

[0243] 检测探针

[0244] 在一些实施方案中,如本文所提供和/或使用的检测探针包含靶结合部分和偶联至靶结合部分的寡核苷酸结构域。在一些实施方案中,偶联至靶结合部分的寡核苷酸结构域包含双链部分和从寡核苷酸结构域的至少一端延伸的单链悬突。在一些实施方案中,偶联至靶结合部分的寡核苷酸结构域包含双链部分和从寡核苷酸结构域的每一端延伸的单链悬突。在一些实施方案中,检测探针可以适用于邻近连接免疫定量聚合酶链反应(pliq-PCR)并且被称为pliq-PCR检测探针。

[0245] A. 靶结合部分

[0246] 偶联至寡核苷酸结构域的靶结合部分是靶标(例如,提供的靶生物标志物特征的生物标志物;本领域技术人员将理解,其中靶生物标志物是特定形式或部分/组分,靶结合部分特异性地与此形式或部分/组分结合)特异性结合的实体或剂。在一些实施方案中,靶结合部分可以具有至少约 10^{-4} M、至少约 10^{-5} M、至少约 10^{-6} M、至少约 10^{-7} M、至少约 10^{-8} M、至少约 10^{-9} M或更低的对靶标(例如,分子靶标)的结合亲和力(例如,如通过解离常数测量的)。本领域技术人员将理解,在一些情况下,结合亲和力(例如,如通过解离常数测量的)可以被两个分子之间的非共价分子相互作用(诸如氢键、静电相互作用、疏水键和范德华力)影响。替代地或另外,配体与其靶分子之间的结合亲和力可以被其他分子的存在影响。本领域技术人员将熟悉用于根据本公开测量结合亲和力和/或解离常数的多种技术,包括例如但不限于ELISA、凝胶迁移测定、下拉测定、平衡透析、分析超速离心、表面等离子体共振(SPR)、生物层干涉测量、光栅耦合干涉技术和光谱测定。

[0247] 在一些实施方案中,靶结合部分可以是或包含任何化学类别(例如像碳水化合物、核酸、脂质、金属、多肽、小分子等和/或其组合)的剂。在一些实施方案中,靶结合部分可以是或包含抗体剂和/或适体。在一些实施方案中,靶结合部分是或包含抗体剂,例如,特异性结合靶标或其表位(例如,提供的肺癌的靶生物标志物特征的生物标志物或其表位)的抗体剂。在一些实施方案中,提供的生物标志物的靶结合部分可以是可商购获得的。在一些实施方案中,为了如本文所述的测定中使用的目的,可以设计和创造提供的生物标志物的靶结合部分。在一些实施方案中,靶结合部分是或包含适体,例如,特异性结合至靶标或其表位(例如,提供的肺癌的靶生物标志物特征的生物标志物或其表位)的适体。在一些实施方案中,靶结合部分是或包含affimer分子,所述affimer分子与靶标或其表位(例如,提供的肺癌的靶生物标志物特征的生物标志物或其表位)特异性结合。在一些实施方案中,此类affimer分子可以是或包含与靶标或其表位结合的肽或多肽(例如,如本文所述),所述

affimer分子具有与相应抗体的特异性和亲和力类似的特异性和亲和力。在一些实施方案中,靶标可以是或包含与肺癌相关的靶标。例如,在一些此类实施方案中,癌症相关靶标可以是或包含与多于一种癌症(即,至少两种或更多种癌症)相关的靶标。在一些实施方案中,癌症相关靶标可以是或包含通常与癌症相关的靶标。在一些实施方案中,癌症相关靶标可以是或包含与特定组织的癌症(例如,肺癌)相关的靶标。在一些实施方案中,癌症相关靶标可以是或包含对特定癌症(例如,特定的肺癌)具有特异性的靶标。

[0248] 在一些实施方案中,靶结合部分识别并特异性结合存在于生物实体(包括例如但不限于细胞和/或细胞外囊泡)中的靶标。例如,在一些实施方案中,靶结合部分可以识别并特异性结合肿瘤相关抗原或其表位。在一些实施方案中,肿瘤相关抗原可以是或包含与癌症相关的抗原,所述癌症例如像皮肤癌、脑癌(包括例如胶质母细胞瘤)、乳腺癌、结肠直肠癌、肝癌、肺癌、卵巢癌、胰腺癌、前列腺癌和皮肤癌。在一些实施方案中,靶结合部分可以识别与肺癌(例如,肺腺癌、小细胞肺癌、非小细胞肺癌等)相关的肿瘤抗原。在一些实施方案中,靶结合部分可以识别与肺腺癌相关的肿瘤抗原。

[0249] 在一些实施方案中,靶结合部分可以特异性结合囊泡内靶标,例如提供的囊泡内蛋白质或RNA(例如,mRNA)。在一些实施方案中,靶结合部分可以特异性结合存在于细胞外囊泡上/内的表面靶标,例如存在于肺癌相关细胞外囊泡上的膜结合多肽。

[0250] 在一些实施方案中,靶结合部分针对特定病状或疾病(例如,癌症)的生物标志物,所述生物标志物通过或已通过例如分析患者活检的群体或文库(例如,数十个、数百个、数千个、数万个、数十万个或更多个)和/或鉴定此类生物标志物(例如,预测性生物标志物)的患者数据来确定。

[0251] 在一些实施方案中,有关生物标志物可以是例如通过数据分析鉴定和/或表征的生物标志物。在一些实施方案中,例如,可以通过机器学习和/或计算建模来分析一组不同的数据(例如,在一些实施方案中,包括批量RNA测序、单细胞RNA(scRNA)测序、质谱、组织学、翻译后修饰数据、体外和/或体内实验数据中的一个或多个)进行分析,以鉴定对疾病或病状(例如,癌症)具有高度特异性的生物标志物(例如,预测标志物)。

[0252] 在一些实施方案中,靶结合部分针对组织特异性靶标,例如与特定组织相关的靶标,所述组织例如像脑、乳腺、结肠、卵巢和/或与女性生殖系统相关的其他组织、胰腺、前列腺和/或与男性生殖系统相关的其他组织、肝脏、肺和皮肤。在一些实施方案中,此类组织特异性靶标可以与正常健康组织和/或患病组织(诸如肿瘤)相关。在一些实施方案中,靶结合部分针对与受试者的正常健康状况特异性相关的靶标。

[0253] 在一些实施方案中,在多个检测探针(例如,如本文所述和/或使用的)中使用的个体靶结合实体针对不同的靶标。在一些实施方案中,此类不同的靶标可以代表不同的标志物蛋白或多肽。在一些实施方案中,此类不同的靶标可以代表相同的标志物蛋白或多肽的不同表位。在一些实施方案中,在多个检测探针(例如,如本文所述和/或使用的)中使用的两个或更多个个体靶结合实体可以针对相同的靶标。

[0254] 在一些实施方案中,在用于检测肺癌的多个检测探针中使用的个体靶结合实体可以针对肺癌的靶生物标志物特征的不同靶生物标志物(例如,如以上标题为“用于检测肺癌的提供的生物标志物和/或靶生物标志物特征”的部分中所述的那些)。例如,在一些实施方案中,多个检测探针中的至少两个可以具有其分别针对CEACAM6和EpCAM的靶结合实体。在

一些实施方案中,多个检测探针中的至少两个可以具有其分别针对CEACAM6和SLC34A2的靶结合实体。在一些实施方案中,多个检测探针中的至少两个可以具有其分别针对MUC1和CEACAM6的靶结合实体。

[0255] 在一些实施方案中,用于检测肺癌的多个检测探针中使用的个体靶结合实体可以针对肺癌的靶生物标志物特征的相同靶生物标志物(例如,如以上标题为“用于检测肺癌的提供的生物标志物和/或靶生物标志物特征”的部分中所述的那些)。在一些实施方案中,此类靶结合实体可以针对肺癌的此类靶生物标志物特征的相同靶生物标志物的相同或不同表位。例如,在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对CEACAM6的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对ALCAM的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对CD55的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对CD1的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对CDH3的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对CD274(PD-L1)的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对CEACAM5的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对DSG2的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对EGFR的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对EPCAM的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对FOLR1的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对IG1FR的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对MET的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对MSLN的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对MUC1的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对SLC34A2的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相

同表位或在不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对sTn抗原的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或在不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对Tn抗原的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或在不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对T抗原的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或在不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对TACSTD的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或在不同表位)。在一些实施方案中,多个检测探针中的至少两个可以具有其各自针对TNFRSF10B的靶结合实体(例如,以其完整的蛋白质形式或其片段,例如,其胞外域和/或在相同表位或在不同表位)。

[0256] B. 寡核苷酸结构域

[0257] 在一些实施方案中,根据本公开使用的寡核苷酸结构域(例如,其可以偶联至靶结合部分)包含双链部分和从寡核苷酸结构域的一端或两端延伸的单链悬突。在寡核苷酸结构域包含从每一端延伸的单链悬突的一些实施方案中,单链悬突从双链部分的不同链延伸。在寡核苷酸结构域包含从寡核苷酸结构域的一端延伸的单链悬突的一些实施方案中,寡核苷酸结构域的另一端可以是平端。

[0258] 在一些实施方案中,寡核苷酸结构域包含核糖核苷酸、脱氧核糖核苷酸、能够参与沃森-克里克(Watson-Crick)类型或类似碱基对相互作用的合成核苷酸残基及其任何组合。在一些实施方案中,寡核苷酸结构域是或包含DNA。在一些实施方案中,寡核苷酸结构域是或包含肽核酸(PNA)。

[0259] 在一些实施方案中,寡核苷酸可以具有至少部分地例如通过例如待检测的关注的实体(例如,生物实体诸如细胞外囊泡)的物理特征和/或在待检测的关注的实体(例如,生物实体诸如细胞外囊泡)中分子靶标的选择和定位来确定的长度。在一些实施方案中,检测探针的寡核苷酸结构域被配置成具有这样的长度,使得当第一检测探针和第二检测探针结合关注的实体(例如,生物实体诸如细胞外囊泡)时,第一单链悬突和第二单链悬突足够紧密邻近以允许单链悬突之间的相互作用(例如,杂交)。例如,当关注的实体(例如,生物实体)是细胞外囊泡(例如,外泌体)时,检测探针的寡核苷酸结构域可以各自独立地具有这样的长度,使得当相应的检测探针结合相同细胞外囊泡时,它们各自的单链悬突足够紧密邻近以彼此退火或相互作用。例如,在一些实施方案中,用于检测细胞外囊泡(例如,外泌体)的检测探针的寡核苷酸结构域可以各自独立地具有约20nm至约200nm、约40nm至约500nm、约40nm至约300nm或约50nm至约150nm的长度。在一些实施方案中,用于检测细胞外囊泡(例如,外泌体)的检测探针的寡核苷酸结构域可以各自独立地具有约20nm至约200nm的长度。在一些实施方案中,一组中检测探针的寡核苷酸结构域的长度可以各自独立地变化,以增加和/或最大化在它们同时结合至同一关注的实体时它们相互发现的概率。

[0260] 因此,在一些实施方案中,用于本文提供的技术的寡核苷酸结构域的长度可在约20个至约1000个核苷酸的范围内。在一些实施方案中,寡核苷酸结构域可以具有在约30个至约1000个核苷酸范围内的长度。在一些实施方案中,寡核苷酸结构域可以具有约30个至约500个核苷酸、约30个至约250个核苷酸、约30个至约200个核苷酸、约30个至约150个核

苷酸、约40个至约150个核苷酸、约40个至约125个核苷酸、约40个至约100个核苷酸、约40个至约60个核苷酸、约50个至约90个核苷酸、约50个至约80个核苷酸范围内的长度。在一些实施方案中,寡核苷酸结构域可以具有至少20个或更多个核苷酸,包括例如至少30个、至少40个、至少50个、至少60个、至少70个、至少80个、至少90个、至少100个、至少250个、至少500个、至少750个、至少1000个核苷酸或更多的长度。在一些实施方案中,寡核苷酸结构域可以具有不超过1000个核苷酸或更少,包括例如不超过900个、不超过800个、不超过700个、不超过600个、不超过500个、不超过400个、不超过300个、不超过200个、不超过100个、不超过90个、不超过80个、不超过70个、不超过60个、不超过50个、不超过40个、不超过30个核苷酸、不超过20个核苷酸或更少的长度。

[0261] 在一些实施方案中,寡核苷酸结构域的长度可为约20nm至约500nm。在一些实施方案中,寡核苷酸结构域的长度可为约20nm至约400nm、约30nm至约200nm、约50nm至约100nm、约30nm至约70nm或约40nm至约60nm。在一些实施方案中,寡核苷酸结构域可以具有至少约20nm或更长,包括例如至少约30nm、至少约40nm、至少约50nm、至少约60nm、至少约70nm、至少约80nm、至少约90nm、至少约100nm、至少约200nm、至少约300nm、至少约400nm或更长的长度。在一些实施方案中,寡核苷酸结构域可以具有不超过1000nm或更短,包括例如不超过900nm、不超过800nm、不超过700nm、不超过600nm、不超过500nm、不超过400nm、不超过300nm、不超过200nm、不超过100nm或更短的长度。

[0262] 在一些实施方案中,用于本文提供的技术的寡核苷酸结构域的双链部分的长度可在约30个至约1000个核苷酸的范围内。在一些实施方案中,寡核苷酸结构域的双链部分的长度可在约30个至约500个核苷酸、约30个至约250个核苷酸、约30个至约200个核苷酸、约30个至约150个核苷酸、约40个至约150个核苷酸、约40个至约125个核苷酸、约40个至约100个核苷酸、约50个至约90个核苷酸、约50个至约80个核苷酸的范围内。在一些实施方案中,寡核苷酸结构域的双链部分可以具有至少30个或更多个核苷酸,包括例如至少40个、至少50个、至少60个、至少70个、至少80个、至少90个、至少100个、至少250个、至少500个、至少750个、至少1000个核苷酸或更多的长度。在一些实施方案中,寡核苷酸结构域的双链部分可以具有不超过1000个核苷酸或更少,包括例如不超过900个、不超过800个、不超过700个、不超过600个、不超过500个、不超过400个、不超过300个、不超过200个、不超过100个、不超过90个、不超过80个、不超过70个、不超过60个、不超过50个、不超过40个核苷酸或更少的长度。

[0263] 在一些实施方案中,寡核苷酸结构域的双链部分的长度可为约20nm至约500nm。在一些实施方案中,寡核苷酸结构域的双链部分的长度可为约20nm至约400nm、约30nm至约200nm、约50nm至约100nm、约30nm至约70nm或约40nm至约60nm。在一些实施方案中,寡核苷酸结构域的双链部分可以具有至少约20nm或更长,包括例如至少约30nm、至少约40nm、至少约50nm、至少约60nm、至少约70nm、至少约80nm、至少约90nm、至少约100nm、至少约200nm、至少约300nm、至少约400nm或更长的长度。在一些实施方案中,寡核苷酸结构域的双链部分可以具有不超过1000nm或更短,包括例如不超过900nm、不超过800nm、不超过700nm、不超过600nm、不超过500nm、不超过400nm、不超过300nm、不超过200nm、不超过100nm或更短的长度。

[0264] 在一些实施方案中,寡核苷酸结构域的双链部分的特征在于,当检测探针通过各

自的互补单链悬突(例如,如本文所述和/或使用的)的杂交而彼此连接时,相应寡核苷酸结构域的组合长度(如果有的话,包括将靶结合部分连接到寡核苷酸结构域的接头)足够长,以允许各自的靶结合实体基本上跨越关注的实体(例如,细胞外囊泡)的全部特征长度(例如,直径)。例如,在细胞外囊泡是关注的实体的一些实施方案中,当检测探针彼此完全连接时,检测探针的寡核苷酸结构域的组合长度(如果有的话,包括将靶结合部分连接到寡核苷酸结构域的接头)可以是大约50nm至200nm。

[0265] 在一些实施方案中,寡核苷酸结构域的双链部分包含引物的结合位点。在一些实施方案中,引物的此类结合位点包含被设计成降低或最小化错误引发或引物二聚体的可能性的核苷酸序列。在一些实施方案中,此类特征可以降低检测下限,并因此增加本文提供的系统的灵敏度。在一些实施方案中,引物的结合位点包含被设计成具有与另一引物结合位点类似的退火温度的核苷酸序列。

[0266] 在一些实施方案中,寡核苷酸结构域的双链部分包含被设计成减少或最小化与通常与受试者(例如,人受试者)的基因组和/或基因转录物(例如,基因组DNA和/或RNA,诸如基因的mRNA)相关的核酸序列(例如,DNA和/或RNA序列)的重叠的核苷酸序列。在一些实施方案中,此类特征可以减少或最小化在检测期间可能存在(例如,作为污染物)在样品中的受试者的任何基因组DNA和/或mRNA转录物的干扰。

[0267] 在一些实施方案中,寡核苷酸结构域的双链部分可以具有被设计成减少或最小化自身二聚体、同源二聚体或异源二聚体的形成的核苷酸序列。

[0268] 在一些实施方案中,用于本文提供的技术的寡核苷酸结构域的单链悬突可以具有约2个至约20个核苷酸的长度。在一些实施方案中,寡核苷酸结构域的单链悬突可以具有约2个至约15个核苷酸、约2个至约10个核苷酸、约3个至约20个核苷酸、约3个至约15个核苷酸、约3个至约10个核苷酸的长度。在一些实施方案中,单链悬突可以具有至少1个至5个核苷酸的长度。在一些实施方案中,寡核苷酸结构域的单链悬突可以具有至少2个或更多个核苷酸,包括例如至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少14个、至少15个、至少20个核苷酸或更多的长度。在一些实施方案中,寡核苷酸结构域的单链悬突可以具有不超过20个核苷酸或更少,包括例如不超过15个、不超过14个、不超过13个、不超过12个、不超过11个、不超过10个、不超过9个、不超过8个、不超过7个、不超过6个、不超过5个、不超过4个核苷酸或更少的长度。

[0269] 在一些实施方案中,寡核苷酸结构域的单链悬突可以具有约1nm至约10nm的长度。在一些实施方案中,寡核苷酸结构域的单链悬突可以具有约1nm至约5nm的长度。在一些实施方案中,寡核苷酸结构域的单链悬突可以具有至少约0.5nm或更长,包括例如至少约1nm、至少约1.5nm、至少约2nm、至少约3nm、至少约4nm、至少约5nm、至少约6nm、至少约7nm、至少约8nm、至少约9nm、至少约10nm或更长的长度。在一些实施方案中,寡核苷酸结构域的单链悬突可以具有不超过10nm或更短,包括例如不超过9nm、不超过8nm、不超过7nm、不超过6nm、不超过5nm、不超过4nm、不超过3nm、不超过2nm、不超过1nm或更短的长度。

[0270] 寡核苷酸结构域的单链悬突被设计成包含与第二检测探针的单链悬突的至少一部分互补的核苷酸序列,使得包含第一检测探针和第二检测探针的双链复合物可以通过互补单链悬突的杂交形成。在一些实施方案中,选择互补单链悬突的核苷酸序列以在适当核酸连接酶存在下获得最佳连接效率。在一些实施方案中,单链悬突具有优先选择用于通过

关注的特定核酸连接酶(例如,DNA连接酶,诸如T4或T7连接酶)有效连接的核苷酸序列。例如,此类单链悬突可以具有核苷酸序列GAGT,例如,如描述于Song等人,“Enzyme-guided DNA sewing architecture”Scientific Reports 5:17722(2015),其出于本文所述目的以引用的方式并入本文。

[0271] 当两个检测探针通过各自的互补单链悬突的杂交偶联在一起时,它们各自的包含杂交单链悬突的寡核苷酸结构域在一些实施方案中可以具有关注的实体(例如,生物实体)的特征长度(例如,直径)约90%至110%或约95%至105%的组合长度。例如,在一些实施方案中,当生物实体为外泌体时,组合长度可以是约50nm至约200nm、或约75nm至约150nm、或约80nm至约120nm。

[0272] C. 靶结合部分与寡核苷酸结构域之间的偶联

[0273] 寡核苷酸结构域和靶结合部分可以通过共价键联和/或通过非共价结合(例如像蛋白质-蛋白质相互作用,诸如链霉素亲和素-生物素相互作用和/或离子相互作用)在检测探针中偶联在一起。在一些实施方案中,适合根据本公开使用的检测探针为包含靶结合部分和寡核苷酸结构域的缀合物分子,其中所述两种组分通常例如直接通过键或间接通过一个或多个接头彼此共价偶联。在一些实施方案中,靶结合部分通过共价键联(例如直接通过键或间接通过一个或多个接头)和/或通过非共价结合(例如像蛋白质-蛋白质相互作用,诸如链霉素亲和素-生物素相互作用和/或离子相互作用)偶联至寡核苷酸结构域的两条链中的一条。

[0274] 在使用接头的情况下,在一些实施方案中,选择接头以提供靶结合部分与寡核苷酸结构域的一条或两条链通过所选接头的共价连接。在一些实施方案中,选择接头使得所得靶结合部分与寡核苷酸结构域的一条或两条链的共价连接保持靶结合部分对其靶标的期望的结合亲和力。在一些实施方案中,选择接头以增强靶结合部分对其靶标的结合特异性。关注的接头和/或缀合方法可以根据靶结合部分(例如,其大小和/或电荷)而广泛变化。在一些实施方案中,接头是生物学惰性的。

[0275] 用于将靶结合部分偶联至寡核苷酸的多种接头和/或方法是本领域普通技术人员已知的,并且可以根据本公开使用。在一些实施方案中,接头可以包含在任一端具有能够共价连接至靶结合部分的反应性官能团的间隔基团。可以用于接头的间隔基团的实例包括但不限于脂肪族和不饱和烃链(包括例如C4、C5、C6、C7、C8、C9、C10、C11、C12、C13、C14、C15、C16、C17、C18、C19、C20或更长)、含有杂原子诸如氧(例如,醚诸如聚乙二醇)或氮(聚胺)的间隔基、肽、碳水化合物、可能含有杂原子的环状或非环状系统。促进共价连接的反应性官能团的非限制性实例包括亲核官能团(例如,胺、醇、硫醇和/或酰肼)、亲电官能团(例如,醛、酯、乙烯基酮、环氧化物、异氰酸酯和/或马来酰亚胺)、能够进行环加成反应形成二硫键或与金属结合的官能团。但在一些实施方案中,示例性反应性官能团不限于伯胺和仲胺、异羟肟酸、N-羧基琥珀酰亚胺(NHS)酯、二苯并环辛炔(DBCO)-NHS酯、叠氮基-NHS酯、叠氮乙酸 NHS酯、炔丙基-NHS酯、反式环辛烯-NHS酯、N-羧基琥珀酰亚胺碳酸酯、氧羰基咪唑、硝基苯酯、三氟乙酯、缩水甘油醚、乙烯基砜、马来酰亚胺、叠氮苯甲酰肼、N-[4-(对叠氮水杨酸氨基)丁基]-3-[2'-吡啶基二硫基]丙酰胺)、双(磺基琥珀酸亚酰胺)辛二酸酯、己二亚氨酸二甲酯、酒石酸二琥珀酰亚胺、N-马来酰亚胺基丁酰氧基琥珀酰亚胺酯、N-羧基磺基琥珀酰亚胺基-4-叠氮苯甲酸酯、N-琥珀酰亚胺基[4-叠氮苯基]-1,3'-二硫代丙酸酯、N-琥珀酰亚胺

基[4-碘乙酰基]氨基苯甲酸酯、戊二醛和4-[N-马来酰亚胺基甲基]环己烷-1-羧酸琥珀酰亚胺基酯、3-(2-吡啶基二硫基)丙酸N-羟基琥珀酰亚胺酯(SPDP)、4-(N-马来酰亚胺基)-环己烷-1-羧酸N-羟基琥珀酰亚胺酯(SMCC)及其任何组合。

[0276] 在一些实施方案中,使用N-羟基琥珀酰亚胺(NHS)酯化学法将靶结合部分(例如,靶结合抗体剂)偶联或缀合至寡核苷酸结构域的一条或两条链。NHS酯与游离伯胺反应,并导致稳定的共价连接。在一些实施方案中,伯氨基可以位于具有间隔基团的末端,例如但不限于脂肪族和不饱和烃链(例如,C6或C12间隔基团)。

[0277] 在一些实施方案中,可以使用本领域已知的位点特异性缀合方法将靶结合部分(例如,靶结合抗体剂)偶联或缀合至寡核苷酸结构域的一条或两条链,以例如增强缀合的靶结合部分(例如,缀合的靶结合抗体剂)的结合特异性。位点特异性缀合方法的实例包括但不限于通过二硫键、C端、碳水化合物残基或聚糖和/或非天然氨基酸标记的偶联或缀合。在靶结合部分是或包含抗体剂或肽适体的一些实施方案中,寡核苷酸可以通过存在于靶结合部分中的至少一个或多个游离胺基偶联或缀合至靶结合部分。在一些实施方案中,寡核苷酸可以通过存在于靶结合部分中的至少一个或多个反应性硫醇基团偶联或缀合至是或包含抗体剂或肽适体的靶结合部分。在一些实施方案中,寡核苷酸可以通过存在于靶结合部分中的至少一个或多个碳水化合物残基偶联或缀合至是或包含抗体剂或肽适体的靶结合部分。

[0278] 在一些实施方案中,多个寡核苷酸(例如,至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个或更多个)可以偶联或缀合至靶结合部分(例如,靶结合抗体剂)。

[0279] 示例性双链体靶实体检测系统

[0280] 在一些实施方案中,如本公开所提供(并且例如可用于检测例如在单个实体水平上与肺癌相关的细胞外囊泡)的靶实体检测系统包含针对提供的靶生物标志物(例如,本文所述的那些)的第一检测探针(例如,如本文所述和/或使用的)的第一群体以及针对提供的靶生物标志物(例如,本文所述的那些)的第二检测探针(例如,如本文所述和/或使用的)的第二群体。在一些实施方案中,第一检测探针和第二检测探针针对相同的提供的靶生物标志物。在一些实施方案中,第一检测探针和第二检测探针针对不同的提供的靶生物标志物。

[0281] 图2说明了用于在单个实体水平检测包含(i)至少一个表达水平足够高,使得发现相同靶标(例如,提供的肺癌的靶生物标志物特征的生物标志物)的两个分子紧密邻近的靶标(例如,提供的肺癌的靶生物标志物特征的生物标志物),或(ii)至少两个或更多个不同的靶标(例如,提供的肺癌的靶生物标志物特征的生物标志物)的关注的实体(例如,生物实体诸如细胞外囊泡)的示例性双链体靶实体检测系统。第一检测探针包含第一靶结合部分(例如,针对靶癌症标志物1)和偶联至第一靶结合部分的第一寡核苷酸结构域,第一寡核苷酸结构域包含第一双链部分和从第一寡核苷酸结构域的一端延伸的第一单链悬突。如图2所示,第一寡核苷酸结构域可由较长链(链3)和较短链(链1)杂交产生,从而在一端形成双链部分和单链悬突。在一些实施方案中,第一靶结合部分(例如,针对靶癌症标志物1)偶联(例如,共价偶联)至第一寡核苷酸结构域的链(例如,链1)的5'端或3'端。在一些实施方案中,偶联至第一靶结合部分的链的5'端或3'端可以用接头(例如,如本文所述和/或使用的,具有或不具有间隔基团)修饰。在一些实施方案中,第一寡核苷酸结构域的另一条链(例如,

链3)的5'端具有游离磷酸酯基团。

[0282] 在图2所描绘的实施方案中,第二检测探针包含第二靶结合部分(例如,针对靶癌症标志物2)和偶联至第二靶结合部分的第二寡核苷酸结构域,第二寡核苷酸结构域包含第二双链部分和从第二寡核苷酸结构域的一端延伸的第二单链悬突。如图2所示,第二寡核苷酸结构域可以由较长链(链4)和较短链(链2)杂交而得,从而在一端形成双链部分和单链悬突。在一些实施方案中,第二靶结合部分(例如,针对靶癌症标志物2)偶联(例如,共价偶联)至第二寡核苷酸结构域的链(例如,链2)的5'端。在一些实施方案中,偶联至第二靶结合部分的链的5'端可以用接头(例如,如本文所述和/或使用的,具有或不具有间隔基团)修饰。在一些实施方案中,第二寡核苷酸结构域的另一条链(例如,链4)的5'端具有游离磷酸酯基团。

[0283] 至少第一单链悬突和第二单链悬突的部分彼此互补,使得当它们足够紧密邻近时(例如,当第一检测探针和第二检测探针同时结合相同的关注的实体(例如,生物实体,诸如细胞外囊泡)时),它们可以杂交以形成双链复合物。在一些实施方案中,第一单链悬突和第二单链悬突具有相等的长度,使得当它们杂交以形成双链复合物时,它们相应的寡核苷酸结构域之间没有缺口(除了待连接的切口),并且每个相应的靶结合部分位于双链复合物的相对端。例如,在一些实施方案中,双链复合物在连接发生之前形成,其中双链复合物包含通过它们相应的单链悬突(例如,具有4个核苷酸的长度)的直接杂交而彼此偶联的第一检测探针和第二检测探针,其中每个相应的靶结合部分(例如,分别针对靶癌症标志物1和靶癌症标志物2)存在于双链复合物的相对端。在此类实施方案中,双链复合物的两条链(在各自的寡核苷酸结构域之间包含切口)都是可连接的,例如用于扩增和检测。在一些实施方案中,双链复合物(例如,在连接发生之前)可以包含关注的实体(例如,生物实体诸如细胞外囊泡),其中第一靶结合部分(例如,针对靶癌症标志物1)和第二靶结合部分(例如,针对靶癌症标志物2)同时结合关注的实体。

[0284] 在用于检测肺癌的双链体靶实体检测系统的一些实施方案中,第一检测探针的第一靶结合部分可以针对第一靶表面蛋白生物标志物(例如,在标题为“用于检测肺癌的提供的生物标志物和/或靶生物标志物特征”部分中提供的那些),而第二检测探针的第二靶结合部分可以针对第二靶表面蛋白生物标志物(例如,在标题为“用于检测肺癌的提供的生物标志物和/或靶生物标志物特征”的部分中提供的那些)。在一些实施方案中,第一检测探针的第一靶结合部分可以针对第一靶囊泡内蛋白生物标志物(例如,在标题为“用于检测肺癌的提供的生物标志物和/或靶生物标志物特征”的部分中提供的那些),而第二检测探针的第二靶结合部分可以针对第二靶囊泡内蛋白生物标志物(例如,在标题为“用于检测肺癌的提供的生物标志物和/或靶生物标志物特征”的部分中提供的那些)。在一些实施方案中,第一靶结合部分和第二靶结合部分可以针对相同靶表面蛋白生物标志物或相同靶囊泡内蛋白生物标志物的相同或不同表位。在一些实施方案中,第一靶结合部分和第二靶结合部分可以针对不同靶表面蛋白生物标志物或不同靶囊泡内蛋白生物标志物。

[0285] 在用于检测肺癌的双链体靶实体检测系统的一些实施方案中,第一检测探针包含针对缀合至第一寡核苷酸结构域的CEACAM6(例如,为完整蛋白形式其片段,诸如其胞外域)的第一靶结合部分;而第二检测探针包含针对缀合至第二寡核苷酸结构域的CEACAM6(例如,为完整蛋白形式其片段,诸如其胞外域)的第二靶结合部分。在一些此类实施方案中,第

一靶结合部分和第二靶结合部分可以针对CEACAM6(例如,为完整蛋白形式其片段,诸如其胞外域)的一个或多个相同或不同表位。在一些实施方案中,第一寡核苷酸结构域和第二寡核苷酸结构域的双链部分可以是相同的。在一些实施方案中,第一寡核苷酸结构域和第二寡核苷酸结构域的双链部分可以是不同的。

[0286] 在用于检测肺癌的双链体靶实体检测系统的一些实施方案中,第一检测探针包含针对缀合至第一寡核苷酸结构域的CEACAM6多肽的第一靶结合部分;而第二检测探针包含针对缀合至第二寡核苷酸结构域的EpCAM多肽的第二靶结合部分。在一些实施方案中,第一寡核苷酸结构域和第二寡核苷酸结构域的双链部分可以是相同的。在一些实施方案中,第一寡核苷酸结构域和第二寡核苷酸结构域的双链部分可以是不同的。

[0287] 在用于检测肺癌的双链体靶实体检测系统的一些实施方案中,第一检测探针包含针对缀合至第一寡核苷酸结构域的CEACAM6多肽的第一靶结合部分;而第二检测探针包含针对缀合至第二寡核苷酸结构域的SLC34A2多肽的第二靶结合部分。在一些实施方案中,第一寡核苷酸结构域和第二寡核苷酸结构域的双链部分可以是相同的。在一些实施方案中,第一寡核苷酸结构域和第二寡核苷酸结构域的双链部分可以是不同的。

[0288] 在一些实施方案中,用于检测肺癌的双链体靶实体检测系统包含至少两组不同的检测探针。例如,在一些实施方案中,每组可以针对包含一个或多个靶生物标志物(例如,本文所述的那些)的不同靶生物标志物特征。

[0289] 在一些实施方案中,包含至少两组不同的检测探针的双链体靶实体检测系统还可以包括捕获测定,所述捕获测定包括针对细胞外囊泡相关膜结合多肽的捕获剂。

[0290] 在一些实施方案中,包括如本文所述的捕获探针或检测探针的生物标志物探针(例如,生物标志物特征)的任何组合可以与包括如本文所述的捕获探针或检测探针的任何其他生物标志物探针(例如,生物标志物特征)组组合使用。

[0291] 示例性三链体或多链体($n \geq 3$)靶实体检测系统

[0292] 在一些实施方案中,如本公开提供(并且可用于例如在单个实体水平上检测肺癌相关的细胞外囊泡)靶实体检测系统包含 n 个不同检测探针的群体(例如,如本文所述和/或使用的),其中 $n \geq 3$ 。例如,在一些实施方案中,当 $n=3$ 时,靶实体检测系统包含针对第一靶标的第一检测探针(例如,如本文所述和/或使用的)、针对第二靶标的第二检测探针群体(例如,如本文所述和/或使用的)和针对第三靶标的第三检测探针群体(例如,如本文所述和/或使用的)。

[0293] 图13说明了用于在单个实体水平上检测包含三个不同分子靶标的关注的实体(例如,生物实体,诸如细胞外囊泡)的示例性三链体靶实体检测系统。第一检测探针包含第一靶结合部分(例如,抗癌标志物1抗体剂)和偶联至第一靶结合部分的第一寡核苷酸结构域,第一寡核苷酸结构域包含第一双链部分和从第一寡核苷酸结构域的一端延伸的第一单链悬突。如图13所示,第一寡核苷酸结构域可以由较长链(链8)和较短链(链1)杂交产生,从而在一端形成双链部分和单链悬突。在一些实施方案中,第一靶结合部分(例如,抗癌标志物1抗体剂)偶联(例如,共价偶联)至第一寡核苷酸结构域的链(例如,链1)的5'端。在一些实施方案中,偶联至第一靶结合部分的链的5'端可以用接头(例如,如本文所述和/或使用的,具有或不具有间隔基团)修饰。在一些实施方案中,第一寡核苷酸结构域的另一条链(例如,链8)的5'端具有游离磷酸酯基团。

[0294] 在图13所描绘的实施方案中,第二检测探针包含第二靶结合部分(例如,抗癌标志物3抗体剂)和偶联至第二靶结合部分的第二寡核苷酸结构域,第二寡核苷酸结构域包含第二双链部分和从第二寡核苷酸结构域的一端延伸的第二单链悬突。如图13所示,第二寡核苷酸结构域可以由较长链(链4)和较短链(链2)杂交而得,从而在一端形成双链部分和单链悬突。在一些实施方案中,第二靶结合部分(例如,抗癌标志物3抗体剂)偶联(例如,共价偶联)至第二寡核苷酸结构域的链(例如,链2)的5'端。在一些实施方案中,偶联至第二靶结合部分的链的5'端可以用接头(例如,如本文所述和/或使用的,具有或不具有间隔基团)修饰。在一些实施方案中,第二寡核苷酸结构域的另一条链(例如,链4)的5'端不具有游离磷酸酯基团。

[0295] 第三检测探针包含第三靶结合部分(例如,抗癌标志物2抗体剂)和偶联至第三靶结合部分的第三寡核苷酸结构域,第三寡核苷酸结构域包含第三双链部分和从第三寡核苷酸结构域的每一端延伸的单链悬突。例如,一个单链悬突从第三寡核苷酸结构域的链的一端延伸,而另一个单链悬突从第三寡核苷酸结构域的不同链的相对端延伸。如图13所示,第三寡核苷酸结构域可以由两条链(例如,链9和链10)的部分杂交产生,从而在每一端形成双链部分和单链悬突。例如,在第三检测探针的链9的5'端形成单链悬突(3A),其中链9的5'端具有游离磷酸酯基团。另外,在相同的第三检测探针的链10的5'端形成单链悬突(3B),并且第三靶结合部分(例如,抗靶标2抗体剂)也偶联(例如,共价偶联)至链10的5'端。在一些实施方案中,偶联至第三靶结合部分的链(例如,链10)的5'端可以用接头(例如,如本文所述和/或使用的,具有或不具有间隔基团)修饰。

[0296] 当所有三个检测探针足够紧密邻近时,例如,当所有三个检测探针同时结合相同的关注的实体(例如,生物实体)时,(i)第三检测探针的单链悬突(例如,3A)的至少一部分与第二检测探针的单链悬突的相应互补部分杂交,并且(ii)第三检测探针的另一个单链悬突(例如,3B)的至少一部分与第一检测探针的单链悬突的相应互补部分杂交。因此,通过相应单链悬突的直接杂交形成包含以线性排列彼此偶联的所有三个检测探针的双链复合物。参见例如图13。

[0297] 在涉及在提供的技术中使用至少三个或更多个($n \geq 3$)检测探针的一些实施方案中,当检测探针的单链悬突退火于每个相应的配偶体以形成双链复合物时,至少($n-2$)个靶结合部分存在于双链复合物的一个或多个内部位置。在此类实施方案中,期望在双链复合物的单链中存在内部靶结合部分,使得双链复合物的另一条链不含任何内部靶结合部分,并且因此可连接以形成连接模板。例如,用于扩增和检测。参见例如图13(使用三个检测探针)、图14(使用四个检测探针)和图15(使用 n 个检测探针)。

[0298] 在双链复合物的链包含至少一个或多个内部靶结合部分的一些实施方案中,所述链包含内部靶结合部分所偶联的检测探针的寡核苷酸链末端与另一个检测探针的寡核苷酸链末端之间的缺口。缺口的大小足够大,使得所述链在核酸连接酶的存在下变得不可连接。在一些实施方案中,缺口的大小可以是2-8个核苷酸或2-6个核苷酸。在一些实施方案中,缺口的大小是6个核苷酸。在一些实施方案中,重叠(单链悬突之间的杂交区)的长度可以是2-15个核苷酸或4-10个核苷酸。在一些实施方案中,重叠(单链悬突之间的杂交区)的长度是8个核苷酸。选择缺口和/或杂交区的大小以提供与连接模板(不包含内部靶结合部分)和非连接模板(包含至少一个内部靶结合部分)的最佳信号分离。应当注意,虽然图13至

15未示出检测探针与关注的实体(例如,生物实体)的结合,但双链复合物(例如,在连接发生之前)可以包含关注的实体(例如,生物实体,诸如细胞外囊泡),其中至少三个或更多个靶结合部分同时结合关注的实体。

[0299] 在一些实施方案中,用于本文提供的靶实体检测系统(例如,本文所述的双链体、三链体或多链体靶实体检测系统)的检测探针(例如,检测探针和/或特定生物标志物的数量)组合(例如,一组)选择基于例如被认为对于特定应用是最佳的所需特异性和/或所需灵敏度。例如,在一些实施方案中,选择检测探针的组合来检测肺癌(例如,I期、II期、III期或IV期),使得其提供至少95%或更高,包括例如至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.7%、至少99.8%或更高的特异性。在一些实施方案中,选择检测探针的组合来检测肺癌(例如,I期、II期、III期或IV期),使得其提供至少30%或更高,包括例如至少40%、至少50%、至少60%、至少70%、至少80%、至少90%、至少95%或更高的灵敏度。在一些实施方案中,选择检测探针的组合来检测肺癌(例如,I期、II期、III期或IV期),使得其提供至少8%或更高,包括例如至少9%、至少10%、至少15%、至少20%、至少25%、至少30%、至少40%、至少50%或更高的阳性预测值。在一些实施方案中,选择检测探针的组合来检测肺癌(例如,I期、II期、III期或IV期),使得其提供至少2%或更高,包括例如至少3%、至少4%、至少5%、至少6%、至少7%、至少8%、至少9%、至少10%、至少15%、至少20%、至少25%、至少30%、至少40%、至少50%或更高的阳性预测值。在一些实施方案中,选择检测探针的组合来检测肺癌(例如,I期、II期、III期或IV期),使得其提供低于 1×10^7 EV/mL样品或更低,包括例如低于 7×10^6 EV/mL样品、低于 6×10^6 EV/mL样品、低于 5×10^6 EV/mL样品、低于 4×10^6 EV/mL样品、低于 3×10^6 EV/mL样品、低于 2×10^6 EV/mL样品、低于 1×10^6 EV/mL样品或更低的检测限(LOD)。在一些实施方案中,此类肺癌检测测定可以用于检测肺癌的不同亚型,包括例如肺腺癌、小细胞肺癌、鳞状细胞和移行细胞肺癌、大细胞肺癌、非小细胞肺癌、其他特定的肺癌、肉瘤型肺癌和其他本领域已知的特定类型的肺癌(SEER癌症统计回顾1975-2017)。在一些实施方案中,此类肺癌检测测定可以用于检测上皮来源的肺癌。在一些实施方案中,此类肺癌检测测定可以用于检测非小细胞肺癌(例如,肺腺癌和/或肺鳞状细胞癌)。在一些实施方案中,此类肺癌检测测定可以用于检测肺腺癌。在一些实施方案中,此类肺癌检测测定可以用于检测肺鳞状细胞癌。

[0300] 在一些实施方案中,检测探针的组合(例如,一组)而不是单独的检测探针赋予检测疾病、病症或病状(例如,如本文所述的特定肺癌和/或肺癌分期)特异性,例如,一个或多个单独的探针可以针对本身对肺癌没有特异性的靶标。例如,在一些实施方案中,本文提供的靶实体检测系统(例如,本文所述的双链体、三链体或多链体靶实体检测系统)中检测探针的有用组合包含针对对有关疾病、病症或病状具有特异性的靶标(即,对有关疾病、病症或病状具有特异性的靶标)至少一个检测探针,并且还可以包含针对不一定或不完全对有关疾病、病症或病状具有特异性的靶标(例如,也可以在健康的、不具有特定疾病、病症或病状和/或不具有关注的特定疾病分期的一些或所有细胞上发现)的至少一个检测探针。也就是说,如阅读本说明书的本领域技术人员将理解的,只要根据本发明使用的检测探针组是或包含一起对检测相关疾病、病症或病状具有特异性的多个单独的检测探针(即,充分区分与相关疾病、病症或病状相关的用于检测的生物实体与用于检测的不受关注的其他生物实体),则所述组根据本公开的某些实施方案是有用的。

[0301] 在一些实施方案中,本文提供的靶实体检测系统(例如,本文所述的双链体、三链体或多链体靶实体检测系统)可以包含至少一个或多个(例如,至少2个或更多个)对照探针(除了靶特异性检测探针,例如,如本文所述和/或使用的,以例如在一些实施方案中识别疾病特异性生物标志物,诸如癌症特异性生物标志物和/或组织特异性生物标志物)。在一些实施方案中,对照探针被设计成使得其与感兴趣的实体(例如,生物实体)的结合抑制(完全或部分)检测信号的生成。

[0302] 在一些实施方案中,对照探针包含对照结合部分和偶联至对照结合部分的寡核苷酸结构域(例如,如本文所述和/或使用的),所述寡核苷酸结构域包含双链部分和从寡核苷酸结构域的一端延伸的单链悬突。对照结合部分是是与对照参考结合的实体或部分。在一些实施方案中,对照参考可以是或包含优先与正常健康细胞相关的生物标志物。在一些实施方案中,对照参考可以是或包含优先与非靶组织相关的生物标志物。在一些实施方案中,包括对照探针可以选择性地去除或最小化由假阳性(例如,包含对照参考的关注实体,任选地与一个或多个待检测的靶标组合)生成的可检测信号。均于2020年2月28日提交的标题为“Systems, Compositions, and Methods for Target Entity Detection”的美国申请号16/805,637和国际申请PCT/US2020/020529(每个申请的全部内容以引用的方式并入本文)中所述的其他对照探针可以用在提供的靶实体检测系统中。

[0303] 在一些实施方案中,本公开尤其提供了以下见解:如本文所述或使用的检测探针可以非特异性地结合固体基底表面,并且甚至在多次洗涤以去除任何过量或未结合的检测探针之后,其中一些仍可以保留在测定样品中;并且此类非特异性结合检测探针可以从固体基底表面脱落并在连接反应中自由浮动,从而允许它们彼此相互作用以生成产生不期望的背景信号的非特异性连接模板。因此,在一些实施方案中,本文提供的靶实体检测系统(例如,本文所述的双链体、三链体或多链体靶实体检测)可以包含至少一种或多种(例如,至少2种或更多种)抑制剂寡核苷酸,其被设计来捕获未与关注的实体结合但在连接反应中作为游离剂保留的残留检测探针,从而防止此类自由浮动的检测探针与其他自由浮动的互补检测探针相互作用以产生不期望的背景信号。在一些实施方案中,抑制剂寡核苷酸可以是或包含单链或双链寡核苷酸,所述单链或双链寡核苷酸包含针对检测探针(例如,如本文所述或使用的)的单链悬突的结合结构域,其中抑制剂寡核苷酸不包含引物结合位点。抑制剂寡核苷酸中不存在此类引物结合位点,防止引物结合由可检测探针与抑制剂寡核苷酸的连接产生的非特异性连接模板,从而例如通过聚合酶链反应减少或抑制非特异性连接模板的扩增和/或检测。

[0304] 在一些实施方案中,抑制剂寡核苷酸包含针对检测探针(例如,如本文所述或使用的)的单链悬突的结合结构域,其中结合结构域是或包含与检测探针的单链悬突基本互补的核苷酸序列,使得具有互补单链悬突的游离的、未结合的检测探针可以结合抑制剂寡核苷酸的结合结构域。在一些实施方案中,抑制剂寡核苷酸可以在一端具有发夹。在一些实施方案中,抑制剂寡核苷酸可以是单链寡核苷酸,其在一端包含针对检测探针的单链悬突的结合结构域,其中单链寡核苷酸的一部分可以自杂交以在另一端形成发夹。

[0305] 在一些实施方案中,本文提供的靶实体检测系统(例如,本文所述的双链体、三链体或多链体靶实体检测系统)不包含将检测探针的寡核苷酸结构域与另一检测探针的寡核苷酸结构域缔合的连接器寡核苷酸。在一些实施方案中,连接器寡核苷酸被设计成桥接任

何两个检测探针的寡核苷酸结构域,所述两个检测探针在与关注的实体结合时不以其他方式彼此相互作用。在一些实施方案中,连接器寡核苷酸被设计成与检测探针的寡核苷酸结构域的至少一部分和另一检测探针的寡核苷酸结构域的至少一部分杂交。连接器寡核苷酸可以是单链的、双链的或其组合。连接器寡核苷酸不含任何靶结合部分(例如,如本文所述和/或使用的)或对照结合部分。在至少一些实施方案中,不需要连接器寡核苷酸来间接连接检测探针的寡核苷酸结构域;在一些实施方案中,不使用此类连接器寡核苷酸,部分是因为如本文提供和/或使用的检测探针被设计成使得它们各自的寡核苷酸结构域具有足够的长度以在它们足够紧密邻近时(例如,当检测探针同时结合关注的实体(例如,生物实体,诸如细胞外囊泡)时)彼此接触并相互作用。

[0306] 使用提供的靶实体检测系统的方法

[0307] 提供的靶实体检测系统可以用于检测样品中(例如,生物、环境或其他样品中)的关注的实体(例如,生物实体,诸如细胞外囊泡)以用于各种与检测肺癌相关的应用和/或目的。因此,本文提供的一些方面涉及使用适合根据本公开使用的多个(例如,至少2个、至少3个或更多个)检测探针的方法。在一些实施方案中,方法包括使样品(例如,来自人受试者的血液或血液来源样品)中的关注的实体(例如,生物实体,诸如细胞外囊泡)与包含如本文所述和/或使用的至少2个或更多个(包括例如至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少14个、至少15个、至少16个、至少17个、至少18个、至少19个、至少20个或更多个)检测探针的一组检测探针接触。在一些实施方案中,方法包括使包含关注的实体(例如,生物实体,诸如细胞外囊泡)的样品经过靶实体检测系统(例如,如本文提供的)。多个检测探针(例如,至少两个或更多个)可以同时或在不同时间(例如,顺序地)被添加到包含关注的实体(例如,生物实体,诸如细胞外囊泡)的样品中。在一些实施方案中,方法包括在与多个检测探针接触之前,使包含关注的实体的样品与至少一种针对细胞外囊泡相关膜结合多肽的捕获剂接触。

[0308] 在某些实施方案中,提供的用于本文所述方法中的靶实体检测系统包含多个(例如,至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少14个、至少15个、至少16个、至少17个、至少18个、至少19个、至少20个或更多个)不同组(例如,组合)检测探针(例如,如本文所述)。在一些实施方案中,方法包括使样品(例如,来自人受试者的血液或血液来源样品)中的关注的实体(例如,生物实体,诸如细胞外囊泡)与多个检测探针组接触,其中每组包含如本文所述和/或使用的至少2个或更多个(包括例如至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少14个、至少15个、至少16个、至少17个、至少18个、至少19个、至少20个或更多个)检测探针。在一些实施方案中,方法包括使包含关注的实体(例如,生物实体,诸如细胞外囊泡)的样品经过靶实体检测系统(例如,如本文提供的)。多个检测探针和/或检测探针组合(例如,至少两个或更多个)可以同时或在不同时间(例如,顺序地)被添加到包含关注的实体(例如,生物实体,诸如细胞外囊泡)的样品中。在一些实施方案中,方法包括在与多个检测探针接触之前,使包含关注的实体的样品与至少一种针对细胞外囊泡相关膜结合多肽的捕获剂接触。

[0309] 在一些实施方案中,来自分析样品中的一种或多种生物标志物组合的结果(例如,Ct值和/或连接的核酸模板(例如,连接的DNA模板)的相对数量)之间的关系可以与临床信

息(包括例如但不限于患者年龄、既往病史、X-射线结果、吸烟史等)和/或其他信息组合,以更好地分类患有肺癌或处于患肺癌风险的患者。可以使用各种分类算法来解释多个变量之间的关系,以提高测定的灵敏度和/或特异性。在一些实施方案中,此类算法包括但不限于逻辑回归模型、支持向量机、梯度增强机、随机森林算法、朴素贝叶斯算法、K-最近邻算法及其组合。在一些实施方案中,本文描述的测定的性能(例如,准确性)可以例如通过选择生物标志物组合(例如,如本文所述)、选择包括算法的其他因素或变量(例如,临床信息和/或生活史信息,例如,吸烟史)和/或选择算法本身的类型来提高。

[0310] 在某些实施方案中,本文描述的技术利用预测算法,所述预测算法使用如本文所述的数据组进行训练和验证。在某些实施方案中,本文描述的技术用于使用从训练样品创建的算法产生风险评分,所述评分被设计为将来自至少两个,例如至少2个、至少3个、至少4个、至少5个或多于5个包含生物标志物特征的独立测定(例如,如本文所述)的结果考虑在内。在某些实施方案中,算法生成的风险评分可以至少部分地使用来自至少一个单独测定(例如,个体生物标志物特征)的诊断数据(例如,原始和/或标准化Ct值)来生成。在某些实施方案中,参考阈值可以包括在风险评分内。在某些实施方案中,表示多种不同程度的肺癌风险的多个阈值水平可以包括在风险评分中。在一些实施方案中,独立的靶生物标志物特征测定可以作为一系列测定中的单独测定进行,并且可以在预测算法中对单独的测定进行同等或不同的加权。在一些实施方案中,例如,在算法中组合的单独测定的权重(例如,一组生物标志物测定)可以由许多因素确定,包括但不限于单独测定的灵敏度、单独测定的特异性、单独测定的重现性、单独测定的可变性、单独测定的阳性预测值和/或特定测定的最低检测限。在一些实施方案中,可以根据特征(例如,灵敏度、特异性、最低检测限等)对一组生物标志物测定进行排序,然后可以基于它们的相对排序对生物标志物测定进行加权。

[0311] 在一些实施方案中,由算法(如本文所述)生成的风险评分可以以合适的方式呈现,例如,以标称尺度,例如,以反映多种可能性的0-100的尺度,例如,所述可能性包括但不限于受试者患有肺癌的可能性、受试者将患肺癌的可能性和/或肺癌的可能分期。在一些实施方案中,较高的风险评分可以表示疾病病变的可能性增加,例如,从较低到较高的值可以反映健康对照、良性对照、I期、II期、III期和IV期肺癌。在一些实施方案中,当与其他癌症类型相比时,风险评分可以用于降低如本文所述的技术的交叉反应性的可能性。

[0312] 在一些实施方案中,风险评分可以由来源于如本文所述的测定的数据与其他适用的诊断数据诸如年龄、生活史、X-射线结果、MRI结果、低剂量CT扫描或其任何组合的联合生成。在一些实施方案中,风险评分提供高于和超过常规标准的护理诊断测定预测值(例如,高于独立使用或与另一个诊断测定组合使用的低剂量CT扫描测定提供的预测值)的预测值。在一些实施方案中,可以生成对肺癌(例如,肺腺癌和/或肺鳞状细胞癌)具有高特异性并且对其他癌症具有低灵敏度的风险评分。

[0313] 在一些实施方案中,风险评分可以具有用于检测肺癌的相关临床截止值。例如,在一些实施方案中,用于检测的风险评分临床截止值可能需要对早期和晚期肺癌的检测产生至少40%,例如至少50%、至少60%或更高的灵敏度,以及在20至89岁的一般健康受试者群体中具有最小95%的特异性,例如至少96%、至少97%、至少98%、至少99%或更高的特异性的测定。在一些实施方案中,灵敏度和特异性目标是目标为77%灵敏度和99.5%特异性的两侧95%置信区间的近似下限。

[0314] 在一些实施方案中,进行训练研究以提供编程风险评分算法所需的必要数据。在一些实施方案中,此类训练研究包括来自一系列供应商(至少包括商业供应商、目的驱动的研究和/或医生)的一组样品。在一些实施方案中,训练研究包括来自肺腺癌和/或肺鳞状细胞癌患者(例如,I期、II期、III期和/或IV期)的阳性样品、来自某些肺癌细胞系的阳性对照样品、来自良性肺肿瘤患者的阴性样品、来自非肺癌患者(如脑癌、乳腺癌、结肠直肠癌、子宫内膜癌、卵巢癌、黑色素瘤、非霍奇金淋巴瘤、胰腺癌、皮肤癌等)的阴性样品、来自炎性病患者(例如,克罗恩氏病、子宫内膜异位症、II型糖尿病、狼疮、胰腺炎、类风湿性关节炎、溃疡性结肠炎、慢性阻塞性肺疾病等)的阴性样品、来自健康患者的阴性样品或其任何组合。在一些实施方案中,训练研究包括来自任何适当年龄范围(例如<31岁、31-40岁、41-50岁、51-60岁、61-70岁、71-80岁,或>80岁)的患者的样品。在一些实施方案中,训练研究包括来自任何种族/民族/派别的患者(例如,白种人、非洲人、亚洲人等)的样品。

[0315] 在一些实施方案中,进行验证研究以提供确认风险评分算法的效用所需的必要数据。在一些实施方案中,此类验证研究包括来自一系列供应商(至少包括商业供应商、目的驱动的研究和/或医生)的一组样品。在一些实施方案中,验证研究包括来自肺腺癌和/或肺鳞状细胞癌患者(例如,I期、II期、III期和/或IV期)的阳性样品、来自某些肺癌细胞系的阳性对照样品、来自良性肺肿瘤患者的阴性样品、来自非肺癌患者(如脑癌、乳腺癌、结肠直肠癌、子宫内膜癌、卵巢癌、黑色素瘤、非霍奇金淋巴瘤、胰腺癌、皮肤癌等)的阴性样品、来自炎性病患者(例如,克罗恩氏病、子宫内膜异位症、II型糖尿病、狼疮、胰腺炎、类风湿性关节炎、溃疡性结肠炎、慢性阻塞性肺疾病等)的阴性样品、来自健康患者的阴性样品或其任何组合。在一些实施方案中,验证研究包括来自任何适当年龄范围(例如<31岁、31-40岁、41-50岁、51-60岁、61-70岁、71-80岁,或>80岁)的患者的样品。在一些实施方案中,验证研究包括来自任何种族/民族/派别的患者(例如,白种人、非洲人、亚洲人等)的样品。

[0316] 在某些实施方案中,至少一个包含至少一个细胞外囊泡相关膜结合多肽和至少一个(包括例如至少两个或更多个)靶生物标志物(其可以选自任何本文所述的表面蛋白生物标志物、本文所述的囊泡内蛋白生物标志物和/或本文所述的囊泡内RNA生物标志物)的靶生物标志物特征可以体现在肺癌检测测定中。在一些此类实施方案中,至少一种捕获剂针对细胞外囊泡相关膜结合多肽,并且至少一组检测探针针对一个或多个本文所述的此类靶生物标志物。例如,图5至7公开了靶生物标志物特征的某些实例,其中的每一个都可以体现在肺癌检测测定(例如,本文所述的那些)中。

[0317] 在某些实施方案中,至少两个(包括例如,至少三个或更多个)包含至少一个细胞外囊泡相关膜结合多肽和至少一个(包括例如,至少两个或更多个)靶生物标志物(其可以选自任何本文所述的表面蛋白生物标志物、本文所述的囊泡内蛋白生物标志物和/或本文所述的囊泡内RNA生物标志物)的不同的靶生物标志物特征可以体现在肺癌检测测定中。在一些实施方案中,每个不同的靶生物标志物特征可以具有不同的用于单独确定样品是否对肺癌呈阳性的预定截止值。在一些实施方案中,如果测定读出高于多个(例如,至少2个或更多个)靶生物标志物特征的截止值中的至少一个,则确定样品对肺癌呈阳性。在一些实施方案中,可以利用截止值的组合(例如,至少2个、至少3个或更多个)来产生具有推论提高的灵敏度和/或特异性的诊断值。

[0318] 因此,在一些实施方案中,可以将样品分成等分试样,使得可以将不同的捕获剂

和/或不同组的检测探针(例如,每个不同疾病或病状的检测)添加到不同的等分试样中。在此类实施方案中,提供的技术可以与一次一个等分试样或一次多个等分试样一起实施(例如,用于平行测定以增加吞吐率)。

[0319] 在一些实施方案中,添加到样品中的检测探针的量提供在混合物中足够低浓度的检测探针,以确保检测探针在没有与关注的实体(例如,生物实体)结合的情况下不会随机地彼此紧密邻近,至少没有达到任何很大或实质性的程度。因此,在许多实施方案中,当检测探针通过检测探针的各自的靶结合部分与关注的实体(例如,生物实体)的结合位点之间的结合相互作用同时结合相同的关注的实体(例如,生物实体)时,检测探针彼此足够紧密邻近以形成双链复合物(例如,如本文所述)。在一些实施方案中,与样品组合后的混合物中检测探针的浓度可以在约1fM至1 μ M的范围内,诸如约1pM至约1nM,包括约1pM至约100nM。

[0320] 在一些实施方案中,样品中关注的实体(例如,生物实体)的浓度足够低,使得在不存在结合到相同的关注的实体(例如,生物实体)的相应检测探针的情况下,结合到一个关注的实体(例如,生物实体)的检测探针不会随机地与结合到另一个关注的实体(例如,生物实体)的另一个检测探针紧密邻近,至少没有达到任何很大或实质性的程度。仅举例来说,样品中关注的实体(例如,生物实体)的浓度足够低,使得结合到关注的非靶实体(例如,非癌性生物实体,诸如包含第一靶标的细胞外囊泡)的第一靶检测探针不会随机地与结合到另一个关注的非靶实体(例如,非癌性生物实体,诸如细胞外囊泡)的另一个不同的靶检测探针紧密邻近,至少没有达到任何很大或实质性的程度来生成假阳性可检测信号。

[0321] 在使样品中的关注的实体(例如,生物实体)与一组检测探针接触后,可以将此类混合物孵育足以使检测探针结合到关注的实体中的相应靶标(例如,分子靶标)(如果存在的话)的一段时间,以形成双链复合物(例如,如本文所述)。在一些实施方案中,在约10 $^{\circ}$ C至约50 $^{\circ}$ C范围内(包括约20 $^{\circ}$ C至约37 $^{\circ}$ C)的温度下将此类混合物孵育约5分钟至约5小时范围内(包括约30分钟至约2小时)的时间段。

[0322] 然后可以使双链复合物(由使关注的实体诸如生物实体与检测探针接触而产生)随后与核酸连接酶接触以进行检测探针的寡核苷酸链的游离3'端羟基和5'端磷酸端的核酸连接,从而生成包含至少两个或更多检测探针的寡核苷酸链的连接模板。在一些实施方案中,在使包含双链复合物的测定样品与核酸连接酶接触之前,可以将至少一种或多种抑制剂寡核苷酸(例如,从如本文所述)添加到测定样品,使得抑制剂寡核苷酸可以捕获可以在连接反应期间以其他方式彼此相互作用的任何残留的自由浮动检测探针。

[0323] 如本领域中已知的,当两个紧邻的核酸退火或杂交到与它们互补的第三核酸序列时,连接酶催化两个紧邻的核酸的并列的3'-羟基和5'-磷酸末端之间形成磷酸二酯键。可以使用任何已知的核酸连接酶(例如,DNA连接酶),包括但不限于温度敏感和/或热稳定连接酶。温度敏感连接酶的非限制性实例包括噬菌体T4DNA连接酶、噬菌体T7连接酶和大肠杆菌(*E. coli*)连接酶。热稳定连接酶的非限制性实例包括Taq连接酶、Tth连接酶和Pfu连接酶。热稳定连接酶可以从嗜热或超嗜热生物获得,包括但不限于原核生物、真核生物或古细菌生物。在一些实施方案中,核酸连接酶是DNA连接酶。在一些实施方案中,核酸连接酶可以是RNA连接酶。

[0324] 在一些实施方案中,在连接步骤中,合适的核酸连接酶(例如,DNA连接酶)和必要和/或期望的任何试剂与反应混合物组合,并保持在足以发生杂交的连接寡核苷酸

的条件下。连接反应条件是本领域技术人员所熟知的。在连接期间,在一些实施方案中,可以将反应混合物保持在约20℃至约45℃范围内(诸如约25℃至约37℃)的温度下约5分钟至约16小时范围内(诸如约1小时至约4小时)的时间段。在其他实施方案中,可以将反应混合物保持在约35℃至约45℃范围内(诸如约37℃至约42℃,例如,是或约38℃、39℃、40℃或41℃)的温度下约5分钟至约16小时范围内(诸如约1小时至约10小时,包括约2小时至约8小时)的时间段。

[0325] 检测此类连接模板可以提供关于样品中关注的实体(例如,生物实体)对于检测探针所针对的靶标是阳性还是阴性的信息。例如,此类连接模板的可检测水平指示包含关注的靶标(例如,分子靶标)的关注的受试实体(例如,生物实体)。在一些实施方案中,可检测水平是高于参考水平例如至少10%或更多(包括例如至少20%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%或更多)的水平。在一些实施方案中,参考水平可以是在阴性对照样品(诸如其中不存在包含此类靶标的关注的实体的样品)中观察到的水平。相反,此类连接模板的非可检测水平(例如,低于可检测水平的阈值的水平)指示关注的受试实体(例如,生物实体)中不存在关注的靶标(例如,分子靶标)中的至少一个。本领域技术人员将理解,可基于例如被认为对于每个应用和/或目的是最佳的期望灵敏度水平和/或期望特异性水平来确定将可检测水平与不可检测水平分开的阈值。例如,在一些实施方案中,可以使用本文提供的系统来实现99.7%的特异性,例如通过设置比参考水平(例如,在阴性对照样品(例如像来自一个或多个正常健康个体的样品)中观察到的水平)高三个标准偏差的阈值。另外或替代地,本领域技术人员将理解,可检测水平的阈值(例如,如由检测信号强度所反映)可以比参考水平高1至100倍。

[0326] 在一些实施方案中,本文提供的方法包括在连接之后,检测例如作为样品中关注的实体的存在和/或量的量度的连接模板。在各种实施方案中,检测连接模板可以是定性的或定量的。因此,在检测是定性的一些实施方案中,方法提供了对被测定的样品中是否存在包含至少两个或更多个靶标(例如,分子靶标)的关注的实体(例如,生物实体)的读数或评价,例如评估。在其他实施方案中,方法提供了对被测定的样品中是否存在包含至少两个或更多个靶标(例如,分子靶标)的关注的实体(例如,生物实体)的定量检测,例如评价或评估被测定的样品中包含至少两个或更多个靶标(例如,分子靶标)的关注的实体(例如,生物实体)的实际量。在一些实施方案中,此类定量检测可以是绝对的或相对的。

[0327] 可以通过本领域已知的适当方法检测通过使用本文提供的技术而形成的连接模板。本领域技术人员将理解,可以基于例如期望的灵敏度水平和/或实践方法的应用来选择适当的检测方法。在一些实施方案中,可以直接检测连接模板而无需任何扩增,而在其他实施方案中,可以扩增连接模板使得连接模板的拷贝数增加,以例如增强特定测定的灵敏度。在不进行扩增的检测是可行的情况下,可以以多种不同的方式检测连接模板。例如,检测探针的寡核苷酸结构域(例如,如本文所述和/或使用的)可以被直接标记,例如荧光标记或放射性同位素标记,使得连接模板被直接标记。例如,在一些实施方案中,检测探针的寡核苷酸结构域(例如,如本文提供和/或使用的)可以包含可检测标记。可检测标记可以通过光谱、光化学、生物化学、免疫化学、电、光学或化学方式可检测的组合物。此类标记包括用于标记的链霉亲和素缀合物染色的生物素、磁珠(例如, Dynabeads®)、荧光染料(例如, 荧光素、德克萨斯红、罗丹明、绿色荧光蛋白等)、放射性标记(例如, ^3H 、 ^{125}I 、 ^{34}S 、 ^{14}C 或 ^{32}P)、酶

(例如,辣根过氧化物酶、碱性磷酸酶和ELISA中常用的其他酶)和量热标记诸如胶体金或有色玻璃或塑料(例如,聚苯乙烯、聚丙烯、乳胶等)珠。在一些实施方案中,直接标记的连接模板可以与反应混合物的其余部分(包括未连接的直接标记的连接寡核苷酸)大小分离,以检测连接模板。

[0328] 在一些实施方案中,连接模板的检测可以包括扩增步骤,其中连接核酸的拷贝数增加,以例如增强测定的灵敏度。根据需要,扩增可以是线性的或指数的,其中扩增可以包括但不限于聚合酶链反应(PCR);定量PCR、等温扩增、NASBA、数字液滴PCR等。

[0329] 用于实现PCR扩增的各种技术是本领域已知的;本领域技术人员将非常熟悉PCR技术的多种实施方案,并且将能够容易地选择适合扩增使用本文提供的技术生成的连接模板的实施方案。例如,在一些实施方案中,包含连接模板的反应混合物与在引物延伸反应中使用的一个或多个引物(例如PCR引物(诸如在几何(或指数)扩增中使用的正向和反向引物或在线性扩增中使用的单个引物))组合。与一个或多个连接模板接触的寡核苷酸引物应具有足够的长度,以在适当的退火条件下提供与互补模板DNA的杂交。引物的长度通常为至少10bp,包括例如长度为至少15bp、长度为至少20bp、长度为至少25bp、长度为至少30bp或更长。在一些实施方案中,引物的长度通常可以在长度为约15bp至50bp、长度为约18bp至30bp或约20bp至35bp的范围内。连接模板可以与单个引物或一组两个引物(正向和反向引物)接触,这取决于是否需要模板DNA的引物延伸、线性或指数扩增。

[0330] 除了上述组分之外,包含连接模板的反应混合物通常还包括聚合酶和脱氧核糖核苷三磷酸(dNTP)。所需的聚合酶活性可以由一种或多种不同的聚合酶提供。在制备例如用于连接模板的扩增的反应混合物中,各种组成成分可以以任何方便的顺序组合。例如,适当的缓冲液可以与一个或多个引物、一种或多种聚合酶和待检测的连接模板组合,或各种组成成分中的所有可以同时组合以产生反应混合物。

[0331] VI. 用途

[0332] 在一些实施方案中,可以例如使用如本文所述的检测方法和/或测定在包含生物实体(包括例如细胞、循环肿瘤细胞、游离DNA、细胞外囊泡等)的样品中检测肺癌的一个或多个靶生物标志物特征的一个或多个提供的生物标志物。在一些实施方案中,可以例如使用如本文所述的检测方法和/或测定在包含细胞外囊泡的样品中检测肺癌的一个或多个靶生物标志物特征的一个或多个提供的生物标志物。

[0333] 在一些实施方案中,样品可以是或包含生物样品。在一些实施方案中,生物样品来源于需要此类测定的受试者(例如,人受试者)的血液或血液来源样品。在一些实施方案中,生物样品可以是或包含来自需要此类测定的受试者(例如,人受试者)的原始样品(例如,组织或肿瘤样品)。在一些实施方案中,可以对生物样品进行处理,以将一个或多个关注的实体(例如,生物实体)与关注的非靶实体分开和/或使一个或多个关注的实体(例如,生物实体)富集。在一些实施方案中,存在于样品中的关注的实体可以是或包含生物实体,例如细胞或细胞外囊泡(例如,外泌体)。在一些实施方案中,可以对此类生物实体(例如,细胞外囊泡)进行处理或将其与化学试剂接触,以例如稳定和/或交联生物实体中待测定的靶标(例如,提供的靶生物标志物)和/或减少与检测探针的非特异性结合。在一些实施方案中,生物实体是或包含细胞,可以任选地例如用化学试剂对其进行处理用于稳定和/或交联靶标(例如,分子靶标)和/或用于减少非特异性结合。在一些实施方案中,生物实体是或包含

细胞外囊泡(例如,外泌体),可以任选地例如用化学试剂对其进行处理用于稳定和/或交联靶标(例如,分子靶标)和/或用于减少非特异性结合。

[0334] 在一些实施方案中,本文提供的技术可以用于管理对例如一个或多个个体受试者和/或整个受试者群体患者护理。仅举例来说,在一些实施方案中,提供的技术可以用于筛查,例如,其可以定期进行,诸如每年、每半年、每两年或以本领域技术人员认为适当的一些其他频率进行。在一些实施方案中,此类筛查可以是临时动机的或偶然动机的。例如,在一些实施方案中,提供的技术可以用于临时动机地筛查大于某个年龄(例如,超过40、45、50、55、60、65、70、75、80岁或以上)的一个或多个个体受试者或整个受试者(例如,无症状受试者)群体。如本领域技术人员将理解的,在一些实施方案中,可以基于例如但不限于疾病、病症或病状(例如,癌症诸如肺癌)的患病率来确定筛查年龄和/或频率。在一些实施方案中,提供的技术可以用于临时动机地筛查可能已经历促使筛查特定疾病、病症或病状(例如,癌症诸如肺癌)的事故或事件的个体受试者。例如,在一些实施方案中,与确定疾病、病症或病状(例如,癌症诸如肺癌)或其易患性的一个或多个指标相关的偶然动机可以是或包括例如,基于他们的家族病史的事故(例如,近亲诸如血缘亲属先前被诊断为此类疾病、病症或病状诸如肺癌);鉴定疾病、病症或病状(例如,肺癌)的一个或多个生活史相关风险因素;和/或来自基因测试(例如,基因组测序)和/或成像诊断测试(例如,X射线胸片或低剂量CT扫描)的先前偶然发现;特征为特定疾病、病症或病状(例如,慢性阻塞性肺部疾病,和/或症状诸如血痰、久咳、呼吸短促、反复和/或慢性呼吸道感染、胸痛、不明原因体重减轻和/或疲劳)一种或多种体征或症状的发展;患者患有良性肿瘤(例如,通过成像诸如X射线胸片或低剂量CT扫描在肺部观察到良性肿块)及其组合,和/或如本领域技术人员将理解的其他事故或事件。

[0335] 在一些实施方案中,提供的用于管理患者护理的技术可通知治疗和/或支付(例如,治疗的报销)决定和/或行动。例如,在一些实施方案中,提供的技术可以提供对个体受试者是否具有疾病、病症或病状(例如,癌症诸如肺癌)的风险、发病或复发的一个或多个指标的确定,从而根据此类发现通知医师和/或患者何时提供/接受治疗性或预防性建议和/或何时开始此类治疗。在一些实施方案中,此类个体受试者可以是无症状受试者,可以以定期频率(例如,每年、每半年、每两年或本领域技术人员认为适当的其他频率)对其进行暂时动机或偶然动机的筛查。在一些实施方案中,此类个体受试者可能正在经历一种或多种可能与肺癌相关的症状,可以以定期频率(例如,每年、每半年、每两年或本领域技术人员认为适当的其他频率)对其进行暂时动机或偶然动机的筛查。在一些实施方案中,此类个体受试者可以是患有良性肺肿瘤和/或慢性炎性病状的受试者,可以以定期频率(例如,每年、每半年、每两年或本领域技术人员认为适当的其他频率)对其进行暂时动机或偶然动机的筛查。在一些实施方案中,此类个体受试者可以是处于肺癌遗传风险的受试者,可以以定期频率(例如,每年、每半年、每两年或本领域技术人员认为适当的其他频率)对其进行暂时动机或偶然动机的筛查。在一些实施方案中,此类个体受试者可以是具有生活史相关风险的受试者,可以以定期频率(例如,每年、每半年、每两年或本领域技术人员认为适当的其他频率)对其进行暂时动机或偶然动机的筛查。在一些实施方案中,此类个体受试者可以是吸烟受试者(例如,重度吸烟者),可以以定期频率(例如,每年、每半年、每两年或本领域技术人员认为适当的其他频率)对其进行暂时动机或偶然动机的筛查。

[0336] 另外或替代地, 在一些实施方案中, 提供的技术可以例如基于特定反应性生物标志物(例如, 癌症反应性生物标志物)的发现来通知医师和/或患者治疗选择。在一些实施方案中, 提供的技术可以例如基于与疾病相关的一个或多个分子靶标水平的变化的发现来提供个体受试者是否对当前治疗有反应的确定, 从而根据此类发现通知医师和/或患者此类治疗的功效和/或保持或改变治疗的决定。在一些实施方案中, 提供的技术可以例如基于预测推荐的治疗对个体受试者的治疗效果的分子靶标(例如, 肺癌的靶生物标志物特征的一个或多个提供的生物标志物)的发现来提供个体受试者是否可能对推荐的治疗有反应的确定, 从而根据此类发现通知医师和/或患者此类治疗的潜在功效和/或实施或改变治疗的决定。

[0337] 在一些实施方案中, 提供的技术可通知关于健康保险提供者是否报销(或不报销)例如以下的决定: (1) 筛查本身(例如, 报销仅可用于周期性/定期筛查或仅可用于临时和/或偶然动机的筛查); 和/或(2) 根据通过提供的技术的发现开始、保持和/或改变治疗。例如, 在一些实施方案中, 本公开提供了与以下有关的方法: (a) 接收采用提供的技术的筛查结果, 并且还接收对筛查和/或特定治疗方案的报销请求; (b) 如果根据适当的时间表(例如, 基于筛查年龄, 诸如超过某个年龄, 例如超过40岁、45岁、50岁、55岁、60岁、65岁、70岁、75岁、80岁或以上, 和/或筛查频率, 例如像每3个月、每6个月、每年、每2年、每3年或以一些其他频率)或对相关事件的反应对受试者进行筛查, 则批准报销筛查; 和/或如果治疗方案根据接收的筛查结果代表适当治疗, 则批准治疗的报销方案; 以及任选地(c) 实现报销或提供报销被拒绝的通知。在一些实施方案中, 如果接收的筛查结果检测到代表相关治疗方案的批准的生物标志物(例如, 如可以在处方信息标签中和/或通过批准的伴随诊断指出的生物标志物)的生物标志物, 则根据接收的筛查结果, 治疗方案是适当的。

[0338] 替代地或另外, 本公开设想允许或促进如本文所述的筛查结果(例如, 如根据本公开生成的)和/或报销决定的报告和/或处理的报告系统(例如, 通过一个或多个适当的电子装置和/或一个或多个通信系统实施)。各种报告系统是本领域已知的; 本领域技术人员将非常熟悉各种此类实施方案, 并且将能够容易地选择适合实施的实施方案。

[0339] 示例性用途

[0340] A. 检测肺癌发病或复发

[0341] 本公开尤其认识到, 基于批量样品而不是以单个生物实体(例如, 个体细胞外囊泡)的分辨率检测生物实体(例如, 细胞外囊泡)中的单个癌症相关生物标志物或多个癌症相关生物标志物通常在确定从其获得生物实体的受试者是否可能患有或易患癌症(例如, 肺癌)方面不提供足够的特异性和/或灵敏度。本公开尤其提供了解决此类问题的技术, 包括组合物和/或方法, 包括例如通过具体要求用于检测的实体(例如, 细胞外囊泡)的特征是存在至少两个或更多个靶标(例如, 至少两个或更多个提供的肺癌的靶生物标志物特征的生物标志物)的组合。在特定实施方案中, 本公开教导要求此类实体(例如, 细胞外囊泡)的特征是存在(例如, 通过表达)对癌症具有特异性的分子靶标(即, 相关癌症(例如, 肺癌)的“靶生物标志物特征”)的组合的技术, 而不包含靶向组合(例如, 靶生物标志物特征)的生物实体(例如, 细胞外囊泡)不会产生可检测信号。因此, 在一些实施方案中, 本文提供的技术可以用于检测受试者中癌症的风险、发病和/或复发。在一些此类实施方案中, 本文提供的技术可用于检测受试者中肺癌的风险、发病和/或复发。例如, 在一些实施方案中, 选择两

个或更多个提供的生物标志物的组合用于检测特定癌症(例如,肺癌)或各种癌症(其中一种包括肺癌)。在一些实施方案中,提供的用于检测肺癌的生物标志物的特定组合可以通过分析鉴定此类预测组合的肺癌患者活检的群体或文库(例如,数十、数百、数千、数万、数十万或更多)和/或患者数据来确定。在一些实施方案中,生物标志物的相关组合可以是例如通过数据分析鉴定和/或表征的那些。例如,在一些实施方案中,数据分析可以包括生物信息学分析,例如如实施例7-9所描述的。在一些实施方案中,例如,可以通过机器学习和/或计算建模来分析一组不同的肺癌相关数据(例如,在一些实施方案中,包括批量RNA测序、单细胞RNA(scRNA)测序、质谱、组织学、翻译后修饰数据、体外和/或体内实验数据中的一个或多个),以鉴定对肺癌具有高度特异性的预测标志物的组合。在一些实施方案中,可以基于对与癌症(例如,肺癌)不同分期相关的不同数据(例如,在一些实施方案中,包括批量RNA测序、scRNA测序、质谱、组织学、翻译后修饰数据、体外和/或体内实验数据)的比较和分析来用计算机确定区分癌症(例如,肺癌)分期的预测标志物的组合。例如,在一些实施方案中,本文提供的技术可以用于区分肺癌受试者与非肺癌受试者,包括例如健康个体、诊断患有良性肿瘤或附件肿块的受试者以及患有非肺相关疾病、病症和/或病状的受试者(例如,患有非肺癌的受试者,或患有炎性病状(例如,慢性阻塞性肺疾病或慢性肺部感染)的受试者)。在一些实施方案中,本文提供的技术可以用于肺癌的早期检测,例如I期或II期的肺癌的检测。在一些实施方案中,本文提供的技术可以用于检测如本领域已知的一种或多种肺癌亚型,包括例如,肺腺癌、小细胞肺癌、鳞状细胞和移行细胞肺癌、大细胞肺癌、非小细胞癌、其他特定的肺癌、肺肉瘤和其他特定类型的肺癌(SEER癌症统计回顾1975-2017)。在一些实施方案中,本文提供的技术可以用于筛查有遗传风险、生活史相关风险或早期肺癌(例如,肺腺癌和肺鳞状细胞癌)的平均风险的个体。

[0342] 在一些实施方案中,本文提供的技术可以用于在单次测定中筛查受试者的特定癌症的风险、发病或复发。例如,在一些实施方案中,本文提供的技术可以用于筛查受试者的肺癌的风险、发病或复发。在一些实施方案中,本文提供的技术可以用于在单次测定中筛查受试者的特定癌症或多种(例如,至少2种、至少3种或更多种)癌症的风险或发病。例如,在一些实施方案中,本文提供的技术可以用于在单次测定中筛查受试者的多种癌症,其中一种包括肺癌,并且其他待筛查的癌症可以选自脑癌(包括例如胶质母细胞瘤)、乳腺癌、结肠直肠癌、胰腺癌、前列腺癌、肝癌、卵巢癌和皮肤癌组成的组。

[0343] 在一些实施方案中,可以定期(例如,每年、每两年、每三年等)使用提供的技术来筛查人受试者(例如,人受试者)的肺癌(例如,早期肺癌)或癌症复发。在一些实施方案中,适用此类筛查的人受试者可以是成年人或老年人。在一些实施方案中,适用此类筛查的人受试者可以大于特定年龄,例如,45岁或以上、55岁或以上、65岁或以上、70岁或以上、75岁或以上或80岁或以上。在一些实施方案中,适用此类筛查的人受试者的年龄可以是约50岁或以上。在一些实施方案中,适用此类筛查的人受试者的年龄可以是50岁或以下。在一些实施方案中,适用此类筛查的人受试者的年龄可以是35岁以上。

[0344] 在一些实施方案中,适用提供的用于检测肺癌的发病或复发的技术的受试者可以是有吸烟史的受试者(例如,重度吸烟者),所述受试者在一些实施方案中可能正在经历一种或多种与肺癌相关的症状。在一些实施方案中,适用提供的用于检测肺癌的发病或复发的技术的受试者可以是被确定患有良性肺肿瘤和/或一种或多种慢性炎性病状(例如,

COPD)的人受试者。在一些实施方案中,适用提供的用于检测肺癌的发病或复发的技术的受试者可以是具有肺癌家族病史(例如,有一个或多个有肺癌史的一级亲属的受试者)的受试者、之前接受过癌症(例如,肺癌)治疗的受试者、在癌症治疗后处于肺癌复发风险的受试者、肺癌治疗后处于缓解期的受试者,和/或先前或定期筛查肺癌(例如,通过筛查至少一个肺癌生物标志物的存在(例如,如本文所述))的受试者。

[0345] 在一些实施方案中,本公开尤其提供了以下见解:本文描述和/或利用的技术对于筛查某些处于较高的肺癌易感性的受试者群体可能特别有用。在一些实施方案中,本公开尤其认识到本文描述和/或使用的肺检测技术的所得PPV在肺癌易发群体或易感群体中可能更高。在一些实施方案中,本公开尤其提供了以下见解:筛查吸烟个体(例如,在出现症状之前或以其他方式在没有出现一种或多种发展的症状的情况下进行定期筛查)可以是有益的,并且甚至对于肺癌的有效管理(例如,成功治疗)是重要的。在一些实施方案中,本公开提供了肺癌筛查系统,在一些实施方案中,所述肺癌筛查系统可以被实施以检测吸烟患者(例如,有或没有肺癌遗传风险和/或生活史风险,和/或有或没有与肺癌相关的症状)中的肺癌,包括早期癌症。在一些实施方案中,可以实施提供的技术以实现吸烟个体(例如,有或没有肺癌遗传风险和/或生活史风险,和/或有或没有与肺癌相关的症状)的定期筛查。在一些实施方案中,提供的技术实现对肺癌的一个或多个特征(例如,发病、进展、对治疗的反应性、复发等)的检测(例如,早期检测,例如,在一个或多个有症状或无症状个体和/或一个或多个群体中),其中灵敏度和/或特异性(例如,假阳性和/或假阴性结果的比率)适用于允许将提供的技术有用地应用于单次和/或定期(例如,周期性)评估。在一些实施方案中,提供的技术与受试者的定期体格检查(例如,每年、每隔一年或以主治医师批准的间隔)结合是有用的。在一些实施方案中,提供的技术可用于与一种或多种治疗方案结合;在一些实施方案中,提供的技术可以改进此类一种或多种治疗方案的一个或多个特征(例如,根据接受的参数的成功率)。

[0346] 在一些实施方案中,适用提供的用于检测肺癌的发病或复发的技术的受试者可以是无症状人受试者和/或整个无症状受试者群体。此类无症状受试者和/或整个无症状受试者群体可以是有癌症(诸如乳腺癌、卵巢癌、白血病和/或肺癌)家族病史的受试者(例如,具有一个或多个已知与遗传风险因素相关的癌症史的一级亲属的个体);先前接受过癌症(例如,肺癌)治疗的受试者;癌症治疗后有肺癌复发风险的受试者;肺癌治疗后处于缓解期的受试者,和/或先前或定期筛查肺癌(例如,通过筛查至少一种肺癌生物标志物或通过胸部影像学(例如,X射线成像、痰液测试、低剂量CT扫描和/或基于游离核酸、血清蛋白例如CEA、CYFRA 21-1、NSE、ProGRP和/或SCCA的分子测试))的受试者。替代地,在一些实施方案中,无症状受试者可以是先前未经肺癌筛查的受试者、未诊断为肺癌的受试者和/或先前未接受肺癌治疗的受试者。在一些实施方案中,无症状受试者可以是患有良性肺肿瘤的受试者。在一些实施方案中,无症状受试者可以是易患肺癌(例如,处于平均群体风险、处于升高的生活史相关风险或具有肺癌遗传风险)的受试者。

[0347] 在一些实施方案中,可以基于一个或多个特征来选择适用提供的用于检测肺癌的技术的受试者或受试者群体,所述一个或多个特征诸如年龄、种族、遗传史、病史、个人史(例如,吸烟、酒精、药物、致癌剂、饮食、肥胖、身体活动、光暴露、放射暴露、暴露于诸如病毒的感染性因子和/或职业危害)。例如,在一些实施方案中,能够适用提供的肺癌检测技术的

受试者或受试者群体可以目前是或曾经是吸烟者(例如,香烟、雪茄、烟斗和/或水烟)的受试者或受试者群体。

[0348] 在一些实施方案中,适用提供的肺癌检测技术的受试者或受试者群体可以是确定在已知暴露于可吸入致癌物质的条件下工作的受试者或受试者群体,所述条件包括但不限于:铜矿冶炼、铅矿冶炼、锌矿冶炼、杀虫剂制造、砷矿开采、石棉开采、石棉纺织生产、制动衬片加工、水泥生产、施工作业、保温作业、造船作业、陶瓷制造、电子和航空航天设备制造、化学制造、铬酸盐生产、铬电镀、皮革鞣制、颜料生产、镍矿开采、镍精炼、镍电镀、不锈钢和耐热钢生产、多环芳烃生产、铝生产、烃类化合物生产、焦炭生产、铬铁合金生产、含镍矿冶炼、盖屋顶、氦矿开采、陶瓷和玻璃生产、花岗岩加工、金属矿冶炼、硅石开采和采石。

[0349] 在一些实施方案中,适用提供的肺癌检测技术的受试者或受试者群体可以是被确定为在肺癌相关基因(例如,与DNA修复途径相关的基因,诸如ATM或BRCA和/或在潜在原癌基因表皮生长因子受体(EGFR)中的胚系突变)中具有一个或多个胚系突变的受试者或受试者群体及其组合。

[0350] 在一些实施方案中,适用提供的用于肺癌检测的技术的受试者或受试者群体可以是诊断有影像学证实的胸部肿块或肺部肿块的受试者或受试者群体。

[0351] 在一些实施方案中,适用提供的用于肺癌检测的技术的受试者或受试者群体可以是在经历降低风险的肺活检之前处于遗传风险和生活史相关风险的受试者或受试者群体。

[0352] 在一些实施方案中,适用提供的用于肺癌检测的技术的受试者或受试者群体可以是确定具有COPD或肺纤维化的受试者或受试者群体。在一些实施方案中,适用提供的用于肺癌检测的技术的受试者或受试者群体可以是有慢性支气管炎、结核病和/或肺炎史的受试者或受试者群体。在一些实施方案中,适用提供的用于肺癌检测的技术的受试者或受试者群体可以是确定具有HIV和/或AIDS的受试者或受试者群体。在一些实施方案中,适用提供的用于肺癌检测的技术的受试者或受试者群体可以是目前或曾经饮酒量高的受试者或受试者群体。在一些实施方案中,适用提供的用于肺癌检测的技术的受试者或受试者群体可以是确定在EGFR、细胞色素p450酶和/或DNA修复基因中具有遗传突变的受试者或受试者群体。在一些实施方案中,适用提供的用于肺癌检测的技术的受试者或受试者群体可以是暴露于放射疗法和/或化学治疗的受试者或受试者群体。

[0353] 在一些实施方案中,适用提供的用于肺癌检测的技术的受试者或受试者群体可以是有的一种或多种肺癌非特异性症状的受试者或受试者群体。在一些实施方案中,肺癌的示例性非特异性症状可以包括类似于慢性阻塞性肺疾病的症状,和/或症状诸如血痰、持续性咳嗽、呼吸短促、反复和/或慢性呼吸道感染、胸痛、不明原因体重减轻、咯血、气道阻塞和/或疲劳。

[0354] 在一些实施方案中,适用提供的用于肺癌检测的技术的受试者或受试者群体可以是不同血统诸如亚洲人、非裔美国人、白种人、夏威夷原住民或其他太平洋岛民、西班牙裔或拉丁裔、美国印第安人或阿拉斯加原住民、非西班牙裔黑人或非西班牙裔白人的受试者或受试者群体。在一些实施方案中,适用提供的用于肺癌检测的技术的受试者或受试者群体可以是不同血统诸如亚太岛民、西班牙裔、美国印第安人/阿拉斯加原住民、非西班牙裔黑人或非西班牙裔白人的受试者或受试者群体。在一些实施方案中,适用提供的用于检测肺癌的技术的受试者或受试者群体可以是任何种族和/或任何族裔的受试者或受试者群

体。

[0355] 在一些实施方案中,适用提供的用于检测肺癌的技术的受试者或受试者群体可以是确定具有正常的X射线成像、痰液测试、低剂量CT扫描和/或基于细胞游离核酸、血清蛋白(例如,CEA、CYFRA 21-1、NSE、ProGRP和/或SCCA)的分子测试。在一些实施方案中,此类受试者可以从此类诊断测试中获得肺癌的阴性指征。在一些实施方案中,此类受试者可以从此类诊断测试中获得肺癌的阳性指征。在一些实施方案中,本文提供的技术可以与其他诊断测定组合使用,包括例如,但不限于:(i) 体格检查、全科医师就诊、胆固醇/血脂测试、糖尿病(2型)筛查、纤维结肠镜检查、血压筛查、甲状腺功能测试、前列腺癌筛查、乳腺摄影、HPV/Pap涂片和/或疫苗接种;(ii) 胸部X射线成像、痰液测试、胸部低剂量CT扫描和/或基于游离核酸、血清蛋白(例如,CEA、CYFRA 21-1、NSE、ProGRP和/或SCCA)的分子测试;(iii) 筛选与癌症相关的循环肿瘤DNA和/或蛋白质生物标志物中基因突变的血浆的基因测定;(iv) 涉及鉴定细胞表型和标志物表达的免疫荧光染色,随后通过下一代测序进行扩增和分析的测定;和/或(v) EGFR、KRAS、ALK、ROAS1、HER2、BRAF和RET种系、体细胞和循环细胞突变测定,或涉及游离肿瘤DNA、液体活检、血清蛋白和游离DNA和/或循环肿瘤细胞的测定。

[0356] B. 选择癌症疗法(例如,肺癌疗法)

[0357] 在一些实施方案中,提供的技术可以用于为癌症患者(例如,患有或易患肺癌的患者)选择适当的疗法。例如,本文提供的一些实施方案涉及用于癌症治疗(例如,肺癌和/或辅助治疗)的患者分类的伴随诊断测定,其包括使用本文提供的技术在患者样品(例如,来自肺癌患者的血液或血液来源样品)中评估提供的生物标志物的选定组合。基于此类测定结果,那些被确定更可能对癌症治疗(例如,肺癌治疗和/或辅助治疗,包括例如:白蛋白结合型紫杉醇(Abraxane)、马来酸阿法替尼、艾乐替尼(Alectinib)、阿特珠单抗(Atezolizumab)、贝伐单抗(Bevacizumab)、布格替尼(Brigatinib)、盐酸卡马替尼(Capmatinib Hydrochloride)、卡铂、色瑞替尼(Ceritinib)、克唑替尼(Crizotinib)、甲磺酸达拉非尼(Dabrafenib Mesylate)、达可替尼(Dacomitinib)、多西紫杉醇(Docetaxel)、盐酸阿霉素、德瓦鲁单抗(Durvalumab)、恩曲替尼(Entrectinib)、盐酸埃罗替尼(Erlotinib Hydrochloride)、伊维莫司(Everolimus)、吉非替尼(Gefitinib)、盐酸吉西他滨(Gemcitabine Hydrochloride)、易普利单抗(Ipilimumab)、劳拉替尼(Lorlatinib)、甲氮芥盐酸盐(Mechlorethamine Hydrochloride)、氨甲蝶呤、奈西木单抗(Necitumumab)、纳武利尤单抗、甲磺酸奥希替尼(Osimertinib Mesylate)、紫杉醇、紫杉醇白蛋白稳定的纳米粒制剂、派姆单抗(Pembrolizumab)、培美曲塞二钠(Pemetrexed Disodium)、雷莫卢单抗(Ramucirumab)、塞尔帕替尼(Selpercatinib)、曲美替尼(Trametinib)和/或酒石酸长春瑞滨(Vinorelbine Tartrate))有反应的患者可以被施用此类治疗,或者被确定对特定的此类治疗没有反应的患者可以被施用不同的治疗。

[0358] C. 评估治疗功效(例如,癌症治疗功效)

[0359] 在一些实施方案中,本文提供的技术可以用于监测和/或评估施用于癌症患者(例如,肺癌患者)的抗癌疗法的功效。例如,可以在肺癌患者首次接受抗癌治疗(例如,白蛋白结合型紫杉醇、马来酸阿法替尼、艾乐替尼、阿特珠单抗、贝伐单抗、布格替尼、盐酸卡马替尼、卡铂、色瑞替尼、克唑替尼、甲磺酸达拉非尼、达可替尼、多西紫杉醇、盐酸阿霉素、德瓦鲁单抗、恩曲替尼、盐酸埃罗替尼、伊维莫司、吉非替尼、盐酸吉西他滨、易普利单抗、劳拉替

尼、甲氮芥盐酸盐、氨甲蝶呤、奈西木单抗、纳武利尤单抗、甲磺酸奥希替尼、紫杉醇、紫杉醇白蛋白稳定的纳米粒制剂、派姆单抗、培美曲塞二钠、雷莫卢单抗、塞尔帕替尼、曲美替尼和/或酒石酸长春瑞滨)之前或之后收集血液或血液来源样品来检测或测量肿瘤负荷,例如,通过检测包含对肺癌检测有特异性的生物标志物选定组合的细胞外囊泡的存在或数量。一段时间的治疗后,可以从同一肺癌患者收集第二血液或血液来源样品以检测肿瘤负荷的变化,例如通过检测包含对肺癌检测具有特异性的生物标志物的选定组合的细胞外囊泡不存在或量的减少。通过在治疗过程中监测肿瘤负荷的水平和/或变化,可以采取适当的行动过程,例如增加或减少治疗剂的剂量,和/或施用不同的治疗剂。

[0360] VII. 试剂盒

[0361] 还提供了可用于实践如上所述的技术的试剂盒。在一些实施方案中,试剂盒包含多个检测探针(例如,如本文所述和/或使用的)。在一些实施方案中,提供的试剂盒包含两个或更多个(例如,3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或更多个)检测探针。在一些实施方案中,单独的检测探针可针对不同的靶标。在一些实施方案中,两个或更多个单独的检测探针可针对相同的靶标。在一些实施方案中,提供的试剂盒包含针对不同靶标的两个或更多个不同的检测探针,并且任选地可包含也针对另一检测探针所针对的靶标的至少一个另外的检测探针。在一些实施方案中,提供的试剂盒包含多个检测探针子集,其中每个子集包含针对相同的靶标的两个或更多个检测探针。在一些实施方案中,多个检测探针可以作为容器中的混合物提供。在一些实施方案中,多个检测探针子集可以作为独立容器中的单独混合物提供。在一些实施方案中,每个检测探针在独立容器中单独地提供。

[0362] 在一些实施方案中,用于检测肺癌的试剂盒包含:(a)捕获剂,其包含针对细胞外囊泡相关膜结合多肽的靶捕获部分;以及(b)一组检测探针,所述组包含至少两个检测探针,每个检测探针针对肺癌的靶生物标志物特征的靶生物标志物,其中检测探针各自包含:(i)靶结合部分,其针对肺癌的靶生物标志物特征的靶生物标志物;以及(ii)偶联至靶结合部分的寡核苷酸结构域,所述寡核苷酸结构域包含双链部分和从寡核苷酸结构域的一端延伸的单链悬突部分,其中所述至少两个检测探针的单链悬突部分的特征在于它们在所述至少两个检测探针结合至相同的细胞外囊泡时可以彼此杂交。在这些实施方案中,此类肺癌的靶生物标志物特征包含:

[0363] 至少一个细胞外囊泡相关膜结合多肽和选自由以下组成的组的至少一个靶生物标志物:表面蛋白生物标志物、囊泡内蛋白生物标志物和囊泡内RNA生物标志物,其中:

[0364] ●所述表面蛋白生物标志物选自:ABCC3、ALCAM、ARSL、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、DMBT1、DSG2、EGFR、EpCAM、EPHX3、EVA1A、FOLR1、GJB1、GJB2、GPC4、HS6ST2、IG1FR、KDEL3、KRTCAP3、LAMB3、LFNG、LSR、MANEAL、MET、MSLN、MUC1、MUC21、PIGT、PODXL2、PRRG4、ROS1、SDC1、SERINC2、SEZ6L2、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、TNFRSF10B及其组合;

[0365] ●所述囊泡内蛋白生物标志物选自:AOC1、C12orf45、CRABP2、CST1、ETV4、FAM83A、FOXA2、HMGB3、LGALS3BP、MIF、NAPSA、PPP1R14D、S100A14、SBK1、SCGB3A2、SFTA2、SFTPA1、

SFTPA2、SFTPB、SPINK1、TGFA、ZC3H11A及其组合；

[0366] ●所述囊泡内RNA生物标志物选自ABCC3、AOC1、ARSL、B3GNT3、C12orf45、CDCP1、CDH1、CDH3、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CRABP2、CST1、DMBT1、DSG2、EPCAM、EPHX3、ETV4、EVA1A、FAM83A、FOLR1、FOXA2、GJB1、GJB2、GPC4、HMGB3、HS6ST2、KDEL3、KRTCAP3、LAMB3、LFNG、LGALS3BP、LSR、MANEAL、MIF、MSLN、MUC1、MUC21、NAPSA、PIGT、PODXL2、PPP1R14D、PRRG4、ROS1、S100A14、SBK1、SCGB3A2、SDC1、SERINC2、SEZ6L2、SFTA2、SFTPA1、SFTPA2、SFTPB、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、SPINK1、ST14、TGFA、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、ZC3H11A及其组合。

[0367] 在一些实施方案中，用于检测肺癌的试剂盒包含：(a) 捕获剂，其包含针对细胞外囊泡相关膜结合多肽的靶捕获部分；以及(b) 一组检测探针，所述组包含至少两个检测探针，每个检测探针针对肺癌的靶生物标志物特征的靶生物标志物，其中检测探针各自包含：(i) 靶结合部分，其针对肺癌的靶生物标志物特征的靶生物标志物；以及(ii) 偶联至靶结合部分的寡核苷酸结构域，所述寡核苷酸结构域包含双链部分和从寡核苷酸结构域的一端延伸的单链悬突部分，其中所述至少两个检测探针的单链悬突部分的特征在于它们在所述至少两个检测探针结合至相同的细胞外囊泡时可以彼此杂交。在这些实施方案中，此类肺癌的靶生物标志物特征包含至少一个细胞外囊泡相关膜结合多肽(例如，如本文所述)和至少一个选自由以下组成的组的靶生物标志物：表面蛋白生物标志物(例如，如本文所述)、囊泡内蛋白生物标志物(例如，如本文所述)和囊泡内RNA生物标志物(例如，如本文所述)。在一些实施方案中，在提供的试剂盒中使用的一个或多个表面蛋白生物标志物选自：ABCA3、ABCC1、ABCC3、ACBD3、ACSL5、AGER、ALCAM、AP1M2、APH1A、APOO、ATP11A、ATP11B、ATP1B1、ATP6AP2、B4GALT4、BCAP31、BSPRY、CD109、CD55、CD9、CDC42、CDH1、CDH3、CDKAL1、CEACAM5、CEACAM6、CELSR1、CIP2A、CISD2、CKAP4、CLCA2、CLDN1、CLIC6、CLPTM1L、CLSTN1、CNTN1、CPD、CYP2S1、CYP4F11、CYP4F3、DPY19L1、DSC2、DSC3、DSG2、DSG3、EGFR、EPCAM、EPHB3、FAT2、FBXO45、FERMT1、FOLR1、FZD6、GALNT1、GALNT3、GALNT5、GALNT6、GGCX、GOLM1、GOLPH3L、GRHL2、HACD3、IER3IP1、IGSF3、IL1RAP、ITGA2、ITGB6、KLRG2、KPNA2、KRTCAP3、LAD1、LAMB3、LAMC2、LAMP3、LAMTOR2、LCLAT1、LPCAT1、LSR、MAGT1、MARCKSL1、MET、MGAT1、MSLN、MUC1、MUC4、NCSTN、NECTIN1、NECTIN4、NRAS、NT5E、NUP210、PARL、PEX13、PIGN、PIGT、PLA2G4A、PLCH1、PLEC、PSMD2、PTDSS1、PTGFRN、PTPRF、QSOX1、RAB25、RAB38、RAB6B、RAP2B、RCC2、RIT1、SCAMP3、SDC1、SEL1L3、SHROOM2、SLC2A1、SLC34A2、SLC35B2、SLC39A11、SMPDL3B、SOAT1、SPAST、SSR1、SSR4、SURF4、SYNGR2、TACSTD2、TESC、TFRC、TMC5、TMC01、TMED2、TMED3、TMEM132A、TMEM33、TMPRSS4、TMT3、TOMM22、TOR1AIP2、TRAM1、TRPV4、TTC33、UGT1A6、UPK1B、VAMP8、VMA21、VRK2、VWA1、XPR1、XXYLT1、ADAM28、AXL、BSG、CD274、CD47、CLU、DKK1、ERBB3、FLT4、GM3、HGF、IGF1R、IL6、KDR、LAG3、Lewis Y/B抗原、LY6E、NOTCH2、NOTCH3、磷脂酰丝氨酸、TIGIT、TNFRSF10A、TNFRSF10B、TNFSF18、TPBG、VEGFA、Tn抗原、Lewis Y/CD174、唾液酸化Lewis X(sLex)(也称为唾液酸化SSEA-1(SLX))、NeuGcGM3及其组合。在一些实施方案中，在提供的试剂盒中使用的一个或多个囊泡内蛋白生物标志物选自：ABRACL、ACP5、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、AOC1、AP1M2、APOBEC3B、APOBEC3C、ARNTL2、ASF1B、AURKB、BAIAP2L1、BIRC5、C12orf45、C15orf48、C19orf33、C1S、

C8orf4、CA9、CALML3、CAPNS2、CBLC、CCL19、CCNB2、CDC20、CDC45、CDCA4、CDCA5、CDK1、CDKN2A、CDKN2B、CENPW、CEP55、CES1、CHMP4C、CNN2、CPA3、CRABP2、CST1、CSTA、CTSC、CTSE、CYP2S1、DPYSL3、EFS、EGLN3、EHF、ELF3、ELF4、ENAH、ESRP1、ETV4、EVPL、FAM129B、FAM60A、FAM83A、FAM83D、FAM83H、FBP1、FERMT1、FOXA2、FOX E1、FOX M1、GBP6、GNA15、GPX2、GRHL2、GSTA1、HCK、HMGB3、HOXB7、ID1、IGF2BP2、IMPA2、IRF6、IVL、JUP、KIAA1522、KIF2C、KIFC1、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、LGALS3BP、LGALS7B、LSP1、MAGEA4、MAGEA6、MCM2、MDFI、MIF、MYBL2、MYH14、MZB1、NAPSA、NCF2、NNMT、NRARP、NUP210、NUSAP1、OSGIN1、PALLD、PITX1、PKP1、PKP3、PLEK、PLEK2、POSTN、PPP1R14C、PPP1R14D、PRAME、PTPN6、RBP1、RIN2、RIPK4、RPS4Y1、RRM2、S100A11、S100A14、S100A16、S100A2、S100P、SBK1、SCGB3A2、SERPINB13、SERPINB3、SERPINB5、SFTA2、SFTPA1、SFTPA2、SFTPB、SH3BP4、SNAI2、SOX2、SPI1、SPINK1、SPINT1、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、SULF1、SYK、SYTL1、TBC1D2、TEAD2、TEAD3、TFAP2C、TGFA、THBS2、TK1、TOP2A、TP63、TPD52、TPX2、TRIM29、TRIP13、UBE2C、YAP1、ZC3H11A、ZNF217、ZNF750及其组合。在一些实施方案中,在提供的试剂盒中使用的一个或多个囊泡内RNA生物标志物选自:ABCA3、ABCC1、ABCC3、ABRACL、ACP5、ADAM23、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、ANTXR1、AOC1、AP1M2、APOBEC3B、APOBEC3C、AQP3、AREG、ARNTL2、ARSL、ASF1B、ATP8B1、AURKB、B3GNT3、B3GNT5、BAIAP2L1、BCAM、BIK、BIRC5、C12orf45、C15orf48、C19orf33、C1S、C8orf4、CA12、CA9、CALML3、CAPNS2、CBLC、CCL19、CCL5、CCNB2、CD109、CD24、CD53、CD74、CD9、CDC20、CDC42EP1、CDC45、CDCA4、CDCA5、CDCP1、CDH1、CDH3、CDK1、CDKN2A、CDKN2B、CEACAM5、CEACAM6、CELSR1、CENPW、CEP55、CES1、CHMP4C、CLCA2、CLDN1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CNN2、COL17A1、CPA3、CRABP2、CST1、CSTA、CTSC、CTSE、CX3CL1、CXADR、CXCR4、CYBB、CYP2S1、CYP4F11、DAPL1、DMBT1、DPYSL3、DSC2、DSC3、DSG2、DSG3、DSP、EFNA1、EFS、EGFR、EGLN3、EHD2、EHF、ELF3、ELF4、EMP1、EMP2、ENAH、EPCAM、EPHA2、EPHB3、EPHX1、EPHX3、ESRP1、ETV4、EVA1A、EVPL、F11R、F2R、F2RL1、F3、FAM129B、FAM60A、FAM83A、FAM83D、FAM83H、FAT1、FAT2、FBLIM1、FBP1、FCER1G、FERMT1、FGFR2、FGFR3、FOLR1、FOXA2、FOX E1、FOX M1、FX YD3、GALNT3、GBP6、GJA1、GJB1、GJB2、GJB3、GJB5、GJB6、GNA15、GPC1、GPC3、GPC4、GPNMB、GPR87、GPRC5A、GPX2、GRHL2、GSTA1、HAS3、HCK、HMGB3、HOXB7、HS6ST2、ID1、IGF2BP2、IGSF9、IL2RG、IMPA2、IRF6、ITGA2、ITGA6、ITGB4、ITGB6、IVL、JAG2、JUP、KCNS3、KDEL R3、KIAA1522、KIF2C、KIFC1、KITLG、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、KRTCAP3、LAMB3、LAMP3、LAPTM5、LFNG、LGALS3BP、LGALS7B、LRP11、LRRC4、LSP1、LSR、LYPD3、MAGEA4、MAGEA6、MAL2、MANEAL、MAOA、MARCO、MCM2、MDFI、MET、MIF、MMP14、MPZL2、MSLN、MUC1、MUC21、MYBL2、MYH14、MYOF、MZB1、NAPSA、NCF2、NKG7、NNMT、NOTCH3、NRARP、NTRK2、NUP210、NUSAP1、OSGIN1、OSMR、PALLD、PDPN、PDZK1IP1、PECAM1、PERP、PIGR、PIGT、PITX1、PKP1、PKP3、PLEK、PLEK2、PLVAP、PMP22、PODXL2、POSTN、PPL、PPP1R14C、PPP1R14D、PRAME、PROM2、PRRG4、PRSS8、PTGES、PTGFRN、PTPN6、PTPRF、PTPRZ1、RAB25、RAB38、RAET1L、RARRES1、RBP1、RGS1、RHCG、RHOF、RIN2、RIPK4、ROS1、RPS4Y1、RRM2、S100A10、S100A11、S100A14、S100A16、S100A2、S100P、SBK1、SCGB3A2、SCNN1A、SDC1、SERINC2、SERPINB13、SERPINB3、SERPINB5、SEZ6L2、SFTA2、SFTPA1、SFTPA2、SFTPB、SH3BP4、SHISA2、SLC1A5、SLC2A1、SLC34A2、SLC40A1、

SLC44A4、SLC6A14、SLC6A8、SLC7A7、SLC7A8、SMIM22、SMPDL3B、SNAI2、SOX2、SPI1、SPINK1、SPINT1、SPINT2、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、ST14、STEAP1、SULF1、SYK、SYTL1、TACSTD2、TBC1D2、TEAD2、TEAD3、TFAP2C、TGFA、THBD、THBS2、TK1、TM4SF1、TMC4、TMC5、TMEM30B、TMEM45B、TMEM54、TMPRSS11D、TMPRSS11E、TMPRSS2、TMPRSS4、TNFRSF18、TNS4、TOP2A、TP53I11、TP63、TPD52、TPX2、TREM2、TRIM29、TRIP13、TSPAN1、TSPAN13、TSPAN6、TSPAN7、TSPAN8、TUSC3、TYROBP、UBE2C、UPK1B、VAMP8、VANGL2、WLS、YAP1、ZC3H11A、ZNF217、ZNF750及其组合。

[0368] 在一些实施方案中,当至少一个靶生物标志物选自提供的表面蛋白生物标志物中的一个或多个时,选定的一个或多个表面蛋白生物标志物和至少一种细胞外囊泡相关膜结合多肽是不同的。在一些实施方案中,当至少一个靶生物标志物选自提供的表面蛋白生物标志物中的一个或多个时,选定的一个或多个表面蛋白生物标志物和至少一种细胞外囊泡相关膜结合多肽是相同的(具有相同或不同的表位)。

[0369] 在一些实施方案中,在试剂盒中提供的捕获剂包含针对细胞外囊泡相关膜结合多肽生物标志物的靶捕获部分,其是或包含以下中的一个或多个:ALCAM、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CLDN3、CLDN4、DSG2、EGFR、EPCAM、FOLR1、GJB1、GJB2、IG1FR、LAMB3、MET、MSLN、MUC1、PIGT、PODXL2、ROS1、SDC1、SLC34A2、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMPRSS4、TSPAN8、TNFRSF10B及其组合。在一些实施方案中,试剂盒中提供的捕获剂包含针对细胞外囊泡相关膜结合多肽的靶标捕获部分,其是或包含选自SLC34A2、CEACAM5、CEACAM6、EpCAM和/或其组合的一种或多种多肽。

[0370] 在一些实施方案中,试剂盒中提供的至少两个检测探针的靶结合部分各自针对靶生物标志物特征的相同靶生物标志物。在一些实施方案中,此类相同的靶生物标志物是或包含CEACAM6。在一些此类实施方案中,这样的至少两个检测探针的寡核苷酸结构域是不同的。

[0371] 在一些实施方案中,试剂盒中提供的至少两个检测探针的靶结合部分各自针对靶生物标志物特征的不同靶生物标志物。例如,在一些实施方案中,试剂盒包含至少两个分别针对CEACAM6和EpCAM的检测探针。在一些实施方案中,试剂盒包含至少两个分别针对CEACAM6和SLC34A2的检测探针。

[0372] 在一些实施方案中,检测探针的靶结合部分是或包含抗体(例如,单克隆抗体)。

[0373] 在一些实施方案中,试剂盒针对检测如本文所述的肺癌的靶生物标志物特征(包括例如但不限于表4至5中描述的那些)。在一些实施方案中,试剂盒针对检测包含如本文所述的至少一个表面蛋白生物标志物(包括例如但不限于表3中提供的作为细胞外囊泡相关膜结合多肽的那些)和至少一个如本文所述的表面蛋白生物标志物(包括例如但不限于表3中提供的作为靶表面蛋白生物标志物的那些)的肺癌的靶生物标志物特征。

[0374] 在一些实施方案中,试剂盒针对包含至少一个SLC34A2多肽(作为细胞外囊泡相关膜结合多肽)和一个CEACAM6多肽(作为靶表面蛋白生物标志物)的肺癌的靶生物标志物特征。

[0375] 在一些实施方案中,试剂盒针对包含至少一个SLC34A2多肽(作为细胞外囊泡相关膜结合多肽)和一个EpCAM多肽(作为靶表面蛋白生物标志物)的肺癌的靶生物标志物特征。

[0376] 在一些实施方案中,试剂盒针对包含至少一个CEACAM5多肽(作为细胞外囊泡相关

膜结合多肽) 和一个CEACAM6多肽(作为靶表面蛋白生物标志物) 的肺癌的靶生物标志物特征。

[0377] 在一些实施方案中, 试剂盒针对包含至少一个CEACAM5多肽(作为细胞外囊泡相关膜结合多肽) 和一个SLC34A2多肽(作为靶表面蛋白生物标志物) 的肺癌的靶生物标志物特征。

[0378] 在一些实施方案中, 试剂盒针对包含至少一个CEACAM5多肽(作为细胞外囊泡相关膜结合多肽) 和一个CEACAM6多肽以及一个SLC34A2多肽(作为两个靶表面蛋白生物标志物) 的肺癌的靶生物标志物特征。

[0379] 在一些实施方案中, 试剂盒针对包含至少一个SLC34A2多肽(作为细胞外囊泡相关膜结合多肽) 和一个CEACAM6多肽以及一个EpCAM多肽(作为两个靶表面蛋白生物标志物) 的肺癌的靶生物标志物特征。

[0380] 在一些实施方案中, 试剂盒包含至少一种化学试剂, 诸如固定剂、透化剂和/或封闭剂。

[0381] 在一些实施方案中, 试剂盒包含一种或多种核酸连接试剂(例如, 核酸连接酶, 诸如DNA连接酶和/或缓冲溶液)。

[0382] 在一些实施方案中, 试剂盒包含至少一种或多种扩增试剂, 诸如PCR扩增试剂。在一些实施方案中, 试剂盒包含一种或多种核酸聚合酶(例如, DNA聚合酶)、一对或多对引物、核苷酸和/或缓冲溶液。

[0383] 在一些实施方案中, 试剂盒包含用于捕获关注的实体(例如, 生物实体) 的固体基底。例如, 此类固体基底可以是或包含珠(例如, 磁珠)。在一些实施方案中, 此类固体基底可以是或包含表面。在一些实施方案中, 表面可以是或包含测定室的捕获表面(例如, 实体捕获表面), 例如像过滤器、基质、膜、板、管、孔(例如但不限于微孔) 等。在一些实施方案中, 固体基底的表面(例如, 捕获表面) 可以涂覆有助于关注的实体(例如, 生物实体) 的捕获剂(例如, 多肽或抗体剂)。

[0384] 在一些实施方案中, 可以选择试剂盒中提供的一组检测探针用于诊断肺癌。

[0385] 在一些实施方案中, 试剂盒包含多个检测探针组, 其中每组检测探针针对特定癌症的检测并且包含至少2个或更多个检测探针。例如, 此类试剂盒可以用于在单次测定中筛查受试者的各种癌症, 其中一种是肺癌, 而其他癌症可选自皮肤癌、卵巢癌、乳腺癌、结肠直肠癌、胰腺癌、前列腺癌、脑癌和肝癌。

[0386] 在一些实施方案中, 本文提供的试剂盒可以包括用于实践本文所述的方法的说明书。这些说明书可以以多种形式存在于试剂盒中, 所述形式中的一种或多种可以存在于试剂盒中。这些说明书可以存在的一种形式是在合适的介质或基底上打印信息, 例如, 在试剂盒的包装、包装说明书等中打印信息的一张或多张纸。又一方式可以是计算机可读介质, 例如在其上记录有说明信息的磁盘、CD、USB驱动器等。可以存在的又一方式是网站地址, 其可通过互联网用以存取说明信息。试剂盒中可以存在任何方便的方式。

[0387] 在试剂盒用作伴随诊断的一些实施方案中, 此类试剂盒可以包括用于鉴定可能对抗治疗剂作出反应的患者的说明(例如, 鉴定指示患者对抗治疗剂反应性的生物标志物)。在一些实施方案中, 此类试剂盒可以包含与伴随诊断测试串联使用的治疗剂。

[0388] 本发明的其他特征将在以下示例性实施方案的描述过程中变得明显, 给出所述示

例性实施方案是为了说明本发明,并且不旨在限制本发明。

[0389] 例示

[0390] 实施例1:检测与肺腺癌相关的个体细胞外囊泡中的示例性靶生物标志物特征

[0391] 本实施例描述了针对靶标(例如,一个或多个靶生物标志物)的检测探针的合成,每个探针包含靶结合部分和偶联至靶结合部分的寡核苷酸结构域(包含双链部分和单链悬突)。本实施例还展示了使用此类检测探针来检测包含两个或更多个不同靶标的生物实体(例如,细胞外囊泡)的存在或不存在。

[0392] 在一些实施方案中,检测探针可以包含双链寡核苷酸,其一端具有对靶癌症生物标志物具有特异性的抗体剂且另一端具有单链悬突。当两个或更多个检测探针结合到同一生物实体(例如,细胞外囊泡)时,检测探针的单链悬突紧密邻近,使得它们可以相互杂交以形成双链复合物,其可以随后被连接并扩增以进行检测。

[0393] 此研究使用了一组中的至少两个检测探针。在一些实施方案中,此类至少两个检测探针针对相同的靶生物标志物。在一些实施方案中,此类针对相同的靶标的至少两个检测探针针对相同靶标的不同表位或相同靶标的相同表位。在一些实施方案中,此类至少两个检测探针针对不同的靶标。阅读本公开的技术人员将理解两个检测探针可以针对不同的靶生物标志物,或者可以使用三个或更多个检测探针,每个针对不同的靶蛋白。此外,描述于本实施例中的组合物和方法可以扩展到不同生物样品(例如,包含细胞外囊泡)中的应用。

[0394] 本实施例示出来自某些实验的实验数据,证明本文提供的技术能够使用如本文所述的示例性生物标志物组合(例如,SLC34A2、CEACAM5、CEACAM6和EpCAM,例如,在一些实施方案中,使用PBS中的SLC34A2捕获与CEACAM6+CEACAM6细胞外囊泡)使用本文所述的双链体系系统测定检测患者样品中的肺癌(例如,肺腺癌和/或肺鳞状细胞癌)。参见例如图1和2。

[0395] 用此类能够检测细胞系来源的细胞外囊泡(CLD-EV)的双链体系系统,进行了含有来自患有各分期肺癌(I-IV期)和/或亚型的患者的初级群组(参见图4)以及来自各对照组(例如,健康受试者)的患者样品的研究。

[0396] 示例性测定概述

[0397] 在一些实施方案中,本文所述的靶实体检测系统是双链体系系统。在一些实施方案中,此类双链体系系统(例如,如图2所示)使用各自识别不同表位的两种抗体。成对的双链模板DNA也用于qPCR,每个双链模板DNA都具有与其配偶体上的5'悬突互补的特定四碱基5'悬突。每个抗体都与两个双链DNA模板中的一个缀合。当抗体结合它们的靶表位时,相应模板的粘性末端可以杂交。然后在PCR扩增之前,这些粘性末端通过T7连接酶连接在一起。为了在两个DNA模板之间发生杂交,两个抗体需要彼此足够紧密地结合(DNA接头和抗体的长度在50nm至60nm内)。结合但保持未连接的任何模板都不会产生PCR产物,如图2所示。

[0398] 健康对照与I、II、III和IV期肺腺癌血浆:

[0399] 处理来自健康对照和肺腺癌(LUAD)患者的血浆样品以获得纯化的细胞外囊泡,使用如下所述的示例性测定对其进行查询。

[0400] 使用与抗SLC34A2或抗CEACAM5抗体共价缀合的磁珠捕获纯化的EV。使用一组两个检测探针针对由珠捕获的EV进行分析,每个检测探针包含针对靶生物标志物(例如,CEACAM6、EpCAM或SLC34A2)的抗体和不同的寡核苷酸结构域(例如,如本文所述的那些)。

[0401] 仔细选择生物标志物组合以最小化与健康组织来源的细胞外囊泡的交叉反应性，所述生物标志物组合是 (i) SLC34A2和CEACAM6、(ii) SLC34A2、CEACAM6和EPCAM以及 (iii) CEACAM5、CEACAM6和SLC34A2。此类生物标志物组合与健康组织的交叉反应性部分地是通过使用肺腺癌中差异表达的mRNA的热图生物信息学预测的。因此，可以预测不同标志物组合在肺癌细胞外囊泡表面上比在来自健康组织的细胞外囊泡表面上要丰富得多。在一些实施方案中，此类生物标志物组合可以选自：(i) SLC34A2捕获探针和CEACAM6+CEACAM6检测探针、(ii) SLC34A2捕获探针和CEACAM6+EPCAM检测探针，以及 (iii) CEACAM5捕获探针和CEACAM6+SLC34A2检测探针。表1表示如在肺癌细胞系与阴性对照细胞系(例如，非肺癌细胞系)中表达的指示的生物标志物的转录表达评分。

[0402] 表1. 如在某些肺癌细胞系与阴性对照细胞系(例如，非肺癌细胞系)中表达的以下生物标志物组合的转录表达评分

基因	肺癌细胞系1	阴性非肺癌
CEACAM6	++++	-
SLC34A2	+++	-
EPCAM	+++	+
CEACAM5	++	-

[0404] 示例性方法：

[0405] 寡核苷酸

[0406] 在一些实施方案中，寡核苷酸可以具有以下序列结构和修饰。应当注意，下面的链数对应于与图2中所示的链相关的数值。

[0407] 链1 v1：

[0408] /5AzideN/CAGTCTGACACAGCAGTCGTTAATCGTCGCTGCTACCCTTGACATCCGTGACTGGCTAGACAGAGGTGT，其中/5AzideN/是指通过NHS酯接头连接到5'寡核苷酸末端的叠氮基团，或

[0409] /5AmMC12/CAGTCTGACACAGCAGTCGTTAATCGTCGCTGCTACCCTTGACATCCGTGACTGGCTAGACAGAGGTGT，其中/5AmMC12/是指通过12碳间隔基连接到5'寡核苷酸末端的氨基(例如，伯氨基)，或

[0410] /5ThiolMC6/CAGTCTGACACAGCAGTCGTTAATCGTCGCTGCTACCCTTGACATCCGTGACTGGCTAGACAGAGGTGT，其中/5ThiolMC6/是指通过6碳间隔基连接到5'寡核苷酸末端的硫醇。

[0411] 链2 v1：

[0412] /5AzideN/GACCTGACCTACAGTGACCATAGCCTTGCCTGATTAGCCACTGTCCAGTTTGGCTCCTGTCTACTAG，其中/5AzideN/是指通过NHS酯接头连接到5'寡核苷酸末端的叠氮基团，或

[0413] /5AmMC12/GACCTGACCTACAGTGACCATAGCCTTGCCTGATTAGCCACTGTCCAGTTTGGCTCCTGTCTACTAG，其中/5AmMC12/是指通过12碳间隔基连接到5'寡核苷酸末端的氨基(例如，伯氨基)，或

[0414] /5ThiolMC6/GACCTGACCTACAGTGACCATAGCCTTGCCTGATTAGCCACTGTCCAGTTTGGCTCCTGTCTACTAG，其中/5ThiolMC6/是指通过6碳间隔基连接到5'寡核苷酸末端的硫醇

[0415] 链3 v1：

[0416] /5Phos/GAGTACACCTCTGTCTAGCCAGTCACGGATGTCAAGGGTAGCAGCGACGATTAACGACTGCTGTGTCAGACTG，其中/5Phos/是指连接到5'寡核苷酸末端的磷酸酯基团

[0417] 链4 v1:

[0418] /5Phos/ACTCCTAGTGAGACCAGGAGCCAACTGGACAGTGGCTAATCAGGCAAGGCTATGGTCACTGTAGGTCAGGTC,其中/5Phos/是指连接到5'寡核苷酸末端的磷酸酯基团

[0419] 链5 v1:

[0420] CAGTCTGACACAGCAGTCGT

[0421] 链6 v1:

[0422] GACCTGACCTACAGTGACCA

[0423] 链7(探针)v1:

[0424] /56-FAM/TGGCTAGAC/ZEN/AGAGGTGTACTCCTAGTGAGA/3IABkFQ/,其中/56-FAM/是指位于5'寡核苷酸末端的荧光素(例如,6-FAM);并且/3IABkFQ/是指位于3'寡核苷酸末端的荧光素猝灭剂。

[0425] 在一些实施方案中,寡核苷酸可以具有以下序列结构和修饰。应当注意,下面的链数对应于与图2中所示的链相关的数值。

[0426] 链1 v2:

[0427] /5AzideN/CAGTCTGACTCACCCTCGTTAATCGTCGCTGCTACCCTTGACATCCGTGACTGGCTAGACAGAGGTGT,其中/5AzideN/是指通过NHS酯接头连接到5'寡核苷酸末端的叠氮基团,或

[0428] /5AmMC12/CAGTCTGACTCACCCTCGTTAATCGTCGCTGCTACCCTTGACATCCGTGACTGGCTAGACAGAGGTGT,其中/5AmMC12/是指通过12碳间隔基连接到5'寡核苷酸末端的氨基(例如,伯氨基),或

[0429] /5ThiolMC6/CAGTCTGACTCACCCTCGTTAATCGTCGCTGCTACCCTTGACATCCGTGACTGGCTAGACAGAGGTGT,其中/5ThiolMC6/是指通过6碳间隔基连接到5'寡核苷酸末端的硫醇

[0430] 链2 v2:

[0431] /5AzideN/CACCAGACCTACGAAGTCCATAGCCTTGCCTGATTAGCCACTGTCCAGTTTGGCTCCTGTCTCTACTAG,其中/5AzideN/是指通过NHS酯接头连接到5'寡核苷酸末端的叠氮基团,或

[0432] /5AmMC12/CACCAGACCTACGAAGTCCATAGCCTTGCCTGATTAGCCACTGTCCAGTTTGGCTCCTGTCTCTACTAG,其中/5AmMC12/是指通过12碳间隔基连接到5'寡核苷酸末端的氨基(例如,伯氨基),或

[0433] /5ThiolMC6/CACCAGACCTACGAAGTCCATAGCCTTGCCTGATTAGCCACTGTCCAGTTTGGCTCTGGTCTCTACTAG,其中/5ThiolMC6/是指通过6碳间隔基连接到5'寡核苷酸末端的硫醇

[0434] 链3 v2:

[0435] /5Phos/GAGTACACCTCTGTCTAGCCAGTCACGGATGTCAAGGGTAGCAGCGACGATTAACGAGTGGTGAGTCAGACTG,其中/5Phos/是指连接到5'寡核苷酸末端的磷酸酯基团

[0436] 链4 v2:

[0437] /5Phos/GAGTACACCTCTGTCTAGCCAGTCACGACTGCTGTGTCAGACTG,其中/5Phos/是指连接到5'寡核苷酸末端的磷酸酯基团

[0438] 链5 v2:

[0439] CAGTCTGACTCACCCTCGT

[0440] 链6 v2:

[0441] CACCAGACCTACGAAGTCCA

[0442] 链7(探针)v2:

[0443] /56-FAM/TGGCTAGAC/ZEN/AGAGGTGTACTCCTAGTGAGA/3IABkFQ/,其中/56-FAM/是指位于5'寡核苷酸末端的荧光素(例如,6-FAM);并且/3IABkFQ/是指位于3'寡核苷酸末端的荧光素猝灭剂。

[0444] 在一些实施方案中,寡核苷酸可以具有以下序列结构和修饰。应当注意,下面的链数对应于与图2中所示的链相关的数值。

[0445] 链1 v1-med:

[0446] /5AzideN/CAGTCTGACACAGCAGTCGTGACTGGCTAGACAGAGGTGT,其中/5AzideN/是指通过NHS酯接头连接到5'寡核苷酸末端的叠氮基团,或

[0447] /5AmMC12/CAGTCTGACACAGCAGTCGTGACTGGCTAGACAGAGGTGT,其中/5AmMC12/是指通过12碳间隔基连接到5'寡核苷酸末端的氨基(例如,伯氨基),或

[0448] /5ThiolMC6/CAGTCTGACACAGCAGTCGTGACTGGCTAGACAGAGGTGT,其中/5ThiolMC6/是指通过6碳间隔基连接到5'寡核苷酸末端的硫醇。

[0449] 链2 v1-med:

[0450] /5AzideN/GACCTGACCTACAGTGACCATTGGCTCCTGGTCTCACTAG,其中/5AzideN/是指通过NHS酯接头连接到5'寡核苷酸末端的叠氮基团,或

[0451] /5AmMC12/GACCTGACCTACAGTGACCATTGGCTCCTGGTCTCACTAG,其中/5AmMC12/是指通过12碳间隔基连接到5'寡核苷酸末端的氨基(例如,伯氨基),或

[0452] /5ThiolMC6/GACCTGACCTACAGTGACCATTGGCTCCTGGTCTCACTAG,其中/5ThiolMC6/是指通过6碳间隔基连接到5'寡核苷酸末端的硫醇

[0453] 链3 v1-med:

[0454] /5Phos/GAGTACACCTCTGTCTAGCCAGTCACGACTGCTGTGTCAGACTG,其中/5Phos/是指连接到5'寡核苷酸末端的磷酸酯基团

[0455] 链4 v1-med:

[0456] /5Phos/ACTCCTAGTGAGACCAGGAGCCAATGGTCACTGTAGGTCAGGTC,其中/5Phos/是指连接到5'寡核苷酸末端的磷酸酯基团

[0457] 链5 v1:

[0458] CAGTCTGACACAGCAGTCGT

[0459] 链6 v1:

[0460] GACCTGACCTACAGTGACCA

[0461] 链7(探针)v1:

[0462] /56-FAM/TGGCTAGAC/ZEN/AGAGGTGTACTCCTAGTGAGA/3IABkFQ/,其中/56-FAM/是指位于5'寡核苷酸末端的荧光素(例如,6-FAM);并且/3IABkFQ/是指位于3'寡核苷酸末端的荧光素猝灭剂。

[0463] 抗体-寡核苷酸(例如,抗体-DNA)缀合:

[0464] 例如,使用无铜点击化学将25-100μg范围内的抗体等分试样与寡核苷酸链缀合,例如,60μg抗抗体等分试样与杂交的1+3和2+4链缀合。第一步是制备DBCO功能化抗体以参与与叠氮化物修饰的寡核苷酸结构域(例如,DNA结构域)的缀合反应。这开始于抗体与DBCO-PEG5-NHS异源双功能交联剂的反应。在室温下使NHS酯和可用赖氨酸基团之间的反应

进行2小时,之后使用离心超滤去除未反应的交联剂。为了完成缀合,允许叠氮化物修饰的寡核苷酸结构域(例如,DNA结构域)和DBCO功能化抗体在室温下反应过夜。使用Qubit蛋白测定测量缀合抗体的浓度。

[0465] 细胞培养

[0466] 使阴性对照细胞(例如,非肺癌细胞,诸如黑色素瘤细胞或健康细胞)在具有10%无外泌体FBS和50单位的青霉素/链霉素/mL的伊格尔(Eagle's)最低必须培养基(EMEM)中生长。使肺腺癌细胞在具有10%无外泌体FBS和50单位的青霉素/链霉素/mL的罗斯威尔公园纪念研究所(Roswell Park Memorial Institute) (RPMI 1640)培养基中生长。可以用于开发用于检测肺癌的测定的示例性肺癌细胞系(例如,如本文所述的那些)包括但不限于HCC4006、PC9、L068、LUDLU-1、COR-L105、SKLU1、SKMES1、NCI-H727、LC-2/AD、NCIH358、ChaGo-K-1、MOR/CPR、MOR/0.4R、MOR/0.2R、NCIH-322和描述和讨论于以下的细胞系:Gazdar等人,“Lung Cancer Cell Lines as Tools for Biomedical Discovery and Research” Journal of the National Cancer Institute:2010 September 8;102(17):1310-1321,其出于本文所述目的以引用的方式并入本文。所有细胞系均保持在5%CO₂和37℃下,并且传代数低于20。

[0467] 从细胞培养基中纯化细胞外囊泡

[0468] 在一些实施方案中,肺癌细胞和阴性对照细胞在它们各自的培养基中生长,直到它们达到约80%汇合。收集细胞培养基并在室温(RT)下以300 x rcf旋转5分钟以去除细胞和碎片。然后收集上清液并在-80℃下冷冻。

[0469] 在使用前,使用室温水浴解冻储存在-80℃下的冷冻上清液;间歇地颠倒摇动样品以破冰并加速解冻。将类似样品合并,使得每个细胞系管中含有约50mL培养基。

[0470] 在一些实施方案中,如用于细胞培养基的MACS Miltenyi外泌体提取试剂盒说明书所述澄清培养基。

[0471] 在一些实施方案中,使用尺寸排阻纯化柱(例如,使用单个15mL Amicon过滤器以过滤3份添加的细胞系培养基;例如,每一份在室温下以2500至3000×rcf离心10min)浓缩澄清的细胞培养基(例如,至约500μL)。对每个样品收集具有约65nm至约1000nm大小范围的纳米颗粒。在一些实施方案中,可能需要较小的颗粒范围。

[0472] 颗粒计数:

[0473] 例如,用使用TS400芯片的SpectroDyne颗粒计数仪器获得颗粒计数,以测量65nm和1000nm之间的范围的纳米颗粒。在一些实施方案中,可能需要较小的颗粒范围。

[0474] 产生患者血浆池:

[0475] 在一些实施方案中,使用混合患者血浆池。简而言之,将1mL的患者血浆等分试样在室温下解冻至少30分钟。将管短暂涡旋并向下旋转以将血浆固结到每个管的底部。将来自给定患者群组的血浆样品合并适当大小的容器中,并通过上下混合彻底混合。每个血浆池在低蛋白吸附(Protein Lo-bind)的1.5mL Eppendorf管中分成1mL等分试样,并在-80℃下重新冷冻。

[0476] 全血浆澄清(任选):

[0477] 在一些实施方案中,在EV纯化之前,不参与样品处理的人员对样品进行盲检。患者身份信息仅在实验完成后才显示出来,以便进行数据分析。从-80℃下的储存中取出1mL等

分的全血浆,并对其进行三次澄清旋转以去除细胞、血小板和碎片。

[0478] 从澄清血浆中纯化EV的尺寸排阻色谱法:

[0479] 每个澄清的血浆样品(个体样品或混合样品)都通过单次使用的尺寸排阻纯化柱来分离EV。对每个样品收集具有约65nm至约1000nm大小范围的纳米颗粒。在一些实施方案中,可能需要较小的颗粒范围。

[0480] 捕获抗体缀合至磁性捕获珠:

[0481] 抗体缀合至磁珠(例如,环氧官能化的Dynabeads™)。简而言之,在无菌环境中对珠进行称重并将其重悬于缓冲液中。将抗体(每mg珠大约8μg Ab)与功能化珠混合,并在37℃下以直立圆筒混合发生缀合反应过夜。使用缀合试剂盒提供的洗涤缓冲液将珠洗涤数次,并4℃下储存在提供的储存缓冲液中,或在-20℃下储存在基于甘油的储存缓冲液中。

[0482] 使用抗体缀合的磁珠直接捕获纯化的血浆EV:

[0483] 在某些实施方案中,从澄清的血浆样品中直接捕获纯化的血浆EV。例如,对于SLC34A2或CEACAM5捕获,例如在室温下,将纯化的血浆EV的稀释样品与缀合有相应抗体的磁珠一起孵育适当的时间段。

[0484] 抗体-寡核苷酸缀合物与结合在磁性捕获珠上的EV的结合:

[0485] 在适当的缓冲液中以其最佳浓度稀释抗体-寡核苷酸缀合物(例如,抗SLC34A2、CEACAM6或EPCAM抗体-寡核苷酸缀合物;“抗体探针”)。允许抗体探针与包含结合在磁性捕获珠上的EV的样品相互作用。

[0486] 结合后洗涤:

[0487] 在一些实施方案中,将样品在适当的缓冲液中洗涤(例如,多次)。

[0488] 连接:

[0489] 洗涤以去除未结合的抗体-寡核苷酸缀合物后,将具有结合的细胞外囊泡和结合的抗体-寡核苷酸缀合物的珠与连接混合物接触。将混合物在室温下孵育20分钟。

[0490] PCR:

[0491] 连接后,将具有结合的细胞外囊泡和结合的抗体-寡核苷酸缀合物的珠与PCR混合物接触。例如,在Quant Studio 3上用以下示例性PCR方案在96孔板中进行PCR:在95℃下保持1分钟,进行在95℃下持续5秒和在62℃下持续15秒的50个循环。选择的温度变化率是标准的(每秒2℃)。每次实验重复进行单次qPCR反应,并且将ROX用作被动参考以将qPCR信号归一化。然后从Quant Studio 3机器下载数据,并在Python 3.7中进行分析 and 绘制。

[0492] 数据分析:

[0493] 在一些实施方案中,二元分类系统可以用于数据分析。在一些实施方案中,可以基于参考信号将来自检测测定的信号归一化。例如,在一些实施方案中,通过选择参考样品来计算单个抗体双链体的归一化信号。在一些实施方案中,用于计算任意样品i的归一化信号的方程如下,其中信号_{最大}是来自最高浓度细胞系EV标准的信号。

[0494] $\Delta Ct_i = Ct_{ref} - Ct_i$

[0495] 信号_i = $2^{\Delta Ct_i}$

[0496] 归一化信号_i = $\frac{\text{信号}_i}{\text{信号}_{\text{最大}}}$

[0497] 代表性结果:

[0498] LUAD细胞系实验

[0499] 将纯化的细胞系EV在适当的缓冲液中稀释至最佳浓度,并使用适当的捕获探针(例如,抗SLC34A2功能化珠或抗CEACAM5功能化珠(1mL平行测定))进行捕获。使用抗体探针(例如,如本文所述)分析捕获的EV。观察到本文所述的生物标志物组合(例如,与示例性测定例如像如本实施例中所述和图1至2中所示组合)能够区分LUAD衍生的EV与阴性对照细胞系,信号强度与两个标志物的表达密切相关(参见表1)。

[0500] LUAD试验性患者血浆研究

[0501] 图4提供了LUAD患者血浆样品试验性研究中包括的患者的统计。小心地在不同的样品群组中尽可能地匹配年龄和性别。

[0502] 如上所述澄清一毫升或更少(例如,500 μ l或更少、400 μ l或更少、300 μ l或更少、200 μ l或更少或100 μ l或更少)的患者样品血浆的平行测定,并使用尺寸排阻色谱法纯化EV。使用抗SLC34A2磁珠或抗CEACAM5磁珠捕获EV。使用CEACAM6+CEACAM6检测探针或CEACAM6+EPCAM检测探针(参见图5和图6)对抗SLC34A2磁珠捕获的EV进行分析。使用CEACAM6+SLC34A2检测探针(参见图7)对抗CEACAM5磁珠捕获的EV进行分析。

[0503] 讨论:

[0504] 本实施例表明,如本文所述的生物标志物组合(例如,SLC34A2+CEACAM6,或SLC34A2+CEACAM6+EPCAM,或CEACAM5+CEACAM6+SLC34A2的组合)(例如,与如本实施例中所述并在图1至2中说明的双链体测定结合)能够以>99.9%的特异性检测肺腺癌。在一些此类实施方案中,生物标志物组合包含SLC34A2捕获和CEACAM6+CEACAM6检测探针。在一些此类实施方案中,生物标志物组合包含SLC34A2捕获和CEACAM6+EpCAM检测探针。在一些此类实施方案中,生物标志物组合包含CEACAM5捕获和CEACAM6+SLC34A2检测探针。在一些实施方案中,在测定中使用两个或更多个生物标志物组合可以增加该测定的灵敏度。

[0505] 在一些实施方案中,可以使用可以使每个抗体添加多达16条寡核苷酸结构域(例如,DNA)链的树枝化基元而不是每个抗体的一条或两条DNA链,以例如增强信噪比。

[0506] 实施例2:用于肺癌检测的靶生物标志物特征的实施方案

[0507] 在一些实施方案中,包含SLC34A2、CEACAM5、CEACAM6和/或EPCAM(例如,在一些实施例中,SLC34A2捕获和CEACAM6+CEACAM6检测探针,或SLC34A2捕获和CEACAM6+EPCAM检测探针,或CEACAM5捕获和CEACAM6+SLC34A2检测探针)的生物标志物组合被用于检测各种受试者群体(包括例如,健康对照、I期LUAD、II期LUAD、III期LUAD、和IV期LUAD)中的肺癌(例如,在如实施例1中所述的测定之后)。

[0508] 在一些实施方案中,本文所述的两个或更多个生物标志物组合可以一起用于检测肺癌,以例如增加测定的灵敏度。例如,在一些实施方案中,用于肺癌检测的测定可以涉及至少两个生物标志物组合,其中此类至少两个生物标志物组合可以各自包括本文所述的不同生物标志物组合(例如,但不限于表4至5中包括的那些)。

[0509] 在一些实施方案中,本文所述的三种或更多种生物标志物组合可以一起用于检测肺癌,以例如增加测定的灵敏度。例如,在一些实施方案中,用于肺癌检测的测定可以涉及至少三个生物标志物组合,其中此类至少三个生物标志物组合可以各自包括本文所述的不同生物标志物组合(例如,但不限于表4至5中包括的那些)。

[0510] 实施例3:评估作为肺癌生物标志物的细胞外囊泡(EV)表面蛋白

[0511] 在一些实施方案中,肺癌检测包括在免疫亲和捕获细胞外囊泡后检测至少一个或多个EV表面蛋白。

[0512] 在一些实施方案中,存在于细胞外囊泡上的一种或多种表面蛋白或细胞外膜蛋白(“捕获蛋白”)可以用于免疫亲和捕获肺癌相关细胞外囊泡。此类捕获蛋白的实例可以包括但不限于ALCAM、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CLDN3、CLDN4、DSG2、EGFR、EPCAM、FOLR1、IG1FR、GJB1、GJB2、LAMB3、MET、MSLN、MUC1、PIGT、PODXL2、ROS1、SDC1、SLC34A2、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMPRSS4、TSPAN8、TNFRSF10B及其组合。此类捕获蛋白质的其他实例包括但不限于ABCA3、ABCC1、ABCC3、ACBD3、ACSL5、AGER、ALCAM、AP1M2、APH1A、APOO、ATP11A、ATP11B、ATP1B1、ATP6AP2、B4GALT4、BCAP31、BSPRY、CD109、CD55、CD9、CDC42、CDH1、CDH3、CDKAL1、CEACAM5、CEACAM6、CELSR1、CIP2A、CISD2、CKAP4、CLCA2、CLDN1、CLIC6、CLPTM1L、CLSTN1、CNTN1、CPD、CYP2S1、CYP4F11、CYP4F3、DPY19L1、DSC2、DSC3、DSG2、DSG3、EGFR、EPCAM、EPHB3、FAT2、FBXO45、FERMT1、FOLR1、FZD6、GALNT1、GALNT3、GALNT5、GALNT6、GGCX、GOLM1、GOLPH3L、GRHL2、HACD3、IER3IP1、IGSF3、IL1RAP、ITGA2、ITGB6、KLRG2、KPNA2、KRTCAP3、LAD1、LAMB3、LAMC2、LAMP3、LAMTOR2、LCLAT1、LPCAT1、LSR、MAGT1、MARCKSL1、MET、MGAT1、MSLN、MUC1、MUC4、NCSTN、NECTIN1、NECTIN4、NRAS、NT5E、NUP210、PARL、PEX13、PIGN、PIGT、PLA2G4A、PLCH1、PLEC、PSMD2、PTDSS1、PTGFRN、PTPRF、QSOX1、RAB25、RAB38、RAB6B、RAP2B、RCC2、RIT1、SCAMP3、SDC1、SEL1L3、SHROOM2、SLC2A1、SLC34A2、SLC35B2、SLC39A11、SMPDL3B、SOAT1、SPAST、SSR1、SSR4、SURF4、SYNGR2、TACSTD2、TESC、TFRC、TMC5、TMC01、TMED2、TMED3、TMEM132A、TMEM33、TMPRSS4、TMTC3、TOMM22、TOR1AIP2、TRAM1、TRPV4、TTC33、UGT1A6、UPK1B、VAMP8、VMA21、VRK2、VWA1、XPR1、XXYL1、ADAM28、AXL、BSG、CD274、CD47、CLU、DKK1、ERBB3、FLT4、GM3、HGF、IGF1R、IL6、KDR、LAG3、Lewis Y/B抗原、LY6E、NOTCH2、NOTCH3、磷脂酰丝氨酸、TIGIT、TNFRSF10A、TNFRSF10B、TNFSF18、TPBG、VEGFA、Tn糖基化多肽、Lewis Y/CD174糖基化多肽、唾液酸化Lewis X(sLex) (也称为唾液酸化SSEA-1 (SLX))糖基化多肽、NeuGcGM3糖基化多肽及其组合。

[0513] 在一些实施方案中,EV免疫测定法(例如,本文诸如实施例1中描述的那些)和生物标志物验证过程(例如,本文诸如实施例1中描述的那些)可以用于评估作为肺癌的生物标志物的另外的表面蛋白。在一些实施方案中,将针对捕获蛋白(例如,肺癌相关EV中存在的表面蛋白)的抗体与磁珠缀合,并任选地首先在细胞系EV上然后在患者样品上评估其结合特异性靶蛋白的能力。评估抗体涂覆的珠捕获肺癌相关EV的能力,并使用靶实体检测系统(例如,如本文所述的双链体系统,其涉及一组两个检测探针,每个检测探针针对靶标志物(例如,如本文所述))读取由抗体涂覆的珠捕获的EV。

[0514] 在一些实施方案中,可以使用以下表面蛋白生物标志物中的至少一个或多个(例如,1个、2个、3个或更多个)来读取捕获的EV:ABCC3、ALCAM、ARSL、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、DMBT1、DSG2、EGFR、EPCAM、EPHX3、EVA1A、FOLR1、GJB1、GJB2、GPC4、HS6ST2、IG1FR、KDEL3、KRTCAP3、LAMB3、LFNG、LSR、MANEAL、MET、MSLN、MUC1、MUC21、PIGT、PODXL2、PRRG4、ROS1、SDC1、SERINC2、SEZ6L2、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、ST14、sTn抗

原、Tn抗原、T抗原、TACSTD2、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TNFRSF10B、TSPAN1、TSPAN8及其组合。在一些实施方案中,可以使用一组检测探针(例如,如本文使用和/或所述的)来读取捕获的EV,所述检测探针中的至少两个针对以下表面蛋白生物标志物中的一个或多个(例如,1个、2个、3个或更多个):ABCC3、ALCAM、ARSL、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、DMBT1、DSG2、EGFR、EPCAM、EPHX3、EVA1A、FOLR1、GJB1、GJB2、GPC4、HS6ST2、IG1FR、KDEL3、KRTCAP3、LAMB3、LFNG、LSR、MANEAL、MET、MSLN、MUC1、MUC21、PIGT、PODXL2、PRRG4、ROS1、SDC1、SERINC2、SEZ6L2、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TNFRSF10B、TSPAN1、TSPAN8及其组合。在一些实施方案中,一组检测探针包含两个检测探针,每个探针针对相同的表面蛋白。在一些实施方案中,一组检测探针包含两个检测探针,每个探针针对不同的表面蛋白。

[0515] 在一些实施方案中,可以通过使用以下表面蛋白生物标志物中的至少一种或多种(例如,1、2、3或更多种)读取捕获的EV:ABCA3、ABCC1、ABCC3、ACBD3、ACSL5、AGER、ALCAM、AP1M2、APH1A、APOO、ATP11A、ATP11B、ATP1B1、ATP6AP2、B4GALT4、BCAP31、BSPRY、CD109、CD55、CD9、CDC42、CDH1、CDH3、CDKAL1、CEACAM5、CEACAM6、CELSR1、CIP2A、CISD2、CKAP4、CLCA2、CLDN1、CLIC6、CLPTM1L、CLSTN1、CNTN1、CPD、CYP2S1、CYP4F11、CYP4F3、DPY19L1、DSC2、DSC3、DSG2、DSG3、EGFR、EPCAM、EPHB3、FAT2、FBX045、FERMT1、FOLR1、FZD6、GALNT1、GALNT3、GALNT5、GALNT6、GGCX、GOLM1、GOLPH3L、GRHL2、HACD3、IER3IP1、IGSF3、IL1RAP、ITGA2、ITGB6、KLRG2、KPNA2、KRTCAP3、LAD1、LAMB3、LAMC2、LAMP3、LAMTOR2、LCLAT1、LPCAT1、LSR、MAGT1、MARCKSL1、MET、MGAT1、MSLN、MUC1、MUC4、NCSTN、NECTIN1、NECTIN4、NRAS、NT5E、NUP210、PARL、PEX13、PIGN、PIGT、PLA2G4A、PLCH1、PLEC、PSMD2、PTDSS1、PTGFRN、PTPRF、QSOX1、RAB25、RAB38、RAB6B、RAP2B、RCC2、RIT1、SCAMP3、SDC1、SEL1L3、SHROOM2、SLC2A1、SLC34A2、SLC35B2、SLC39A11、SMPDL3B、SOAT1、SPAST、SSR1、SSR4、SURF4、SYNGR2、TACSTD2、TESC、TFRC、TMC5、TMC01、TMED2、TMED3、TMEM132A、TMEM33、TMPRSS4、TMTC3、TOMM22、TOR1AIP2、TRAM1、TRPV4、TTC33、UGT1A6、UPK1B、VAMP8、VMA21、VRK2、VWA1、XPR1、XXYLT1、ADAM28、AXL、BSG、CD274、CD47、CLU、DKK1、ERBB3、FLT4、GM3、HGF、IGF1R、IL6、KDR、LAG3、Lewis Y/B抗原、LY6E、NOTCH2、NOTCH3、磷脂酰丝氨酸、TIGIT、TNFRSF10A、TNFRSF10B、TNFSF18、TPBG、VEGFA、Tn糖基化多肽、Lewis Y/CD174糖基化多肽、唾液酸化Lewis X(sLex)(也称为唾液酸化SSEA-1 (SLX))糖基化多肽、NeuGcGM3糖基化多肽及其组合。在一些实施方案中,可以使用一组检测探针(例如,如本文使用和/或所述的)来读取捕获的EV,所述检测探针中的至少两个针对以下表面蛋白生物标志物中的一个或多个(例如,1个、2个、3个或更多个):ABCA3、ABCC1、ABCC3、ACBD3、ACSL5、AGER、ALCAM、AP1M2、APH1A、APOO、ATP11A、ATP11B、ATP1B1、ATP6AP2、B4GALT4、BCAP31、BSPRY、CD109、CD55、CD9、CDC42、CDH1、CDH3、CDKAL1、CEACAM5、CEACAM6、CELSR1、CIP2A、CISD2、CKAP4、CLCA2、CLDN1、CLIC6、CLPTM1L、CLSTN1、CNTN1、CPD、CYP2S1、CYP4F11、CYP4F3、DPY19L1、DSC2、DSC3、DSG2、DSG3、EGFR、EPCAM、EPHB3、FAT2、FBX045、FERMT1、FOLR1、FZD6、GALNT1、GALNT3、GALNT5、GALNT6、GGCX、GOLM1、GOLPH3L、GRHL2、HACD3、IER3IP1、IGSF3、IL1RAP、ITGA2、ITGB6、KLRG2、KPNA2、KRTCAP3、LAD1、LAMB3、LAMC2、LAMP3、LAMTOR2、LCLAT1、LPCAT1、LSR、MAGT1、MARCKSL1、MET、

MGAT1、MSLN、MUC1、MUC4、NCSTN、NECTIN1、NECTIN4、NRAS、NT5E、NUP210、PARL、PEX13、PIGN、PIGT、PLA2G4A、PLCH1、PLEC、PSMD2、PTDSS1、PTGFRN、PTPRF、QSOX1、RAB25、RAB38、RAB6B、RAP2B、RCC2、RIT1、SCAMP3、SDC1、SEL1L3、SHROOM2、SLC2A1、SLC34A2、SLC35B2、SLC39A11、SMPDL3B、SOAT1、SPAST、SSR1、SSR4、SURF4、SYNGR2、TACSTD2、TESC、TFRC、TMC5、TMC01、TMED2、TMED3、TMEM132A、TMEM33、TMPRSS4、TMTC3、TOMM22、TOR1AIP2、TRAM1、TRPV4、TTC33、UGT1A6、UPK1B、VAMP8、VMA21、VRK2、VWA1、XPR1、XXYL1、ADAM28、AXL、BSG、CD274、CD47、CLU、DKK1、ERBB3、FLT4、GM3、HGF、IGF1R、IL6、KDR、LAG3、Lewis Y/B抗原、LY6E、NOTCH2、NOTCH3、磷脂酰丝氨酸、TIGIT、TNFRSF10A、TNFRSF10B、TNFSF18、TPBG、VEGFA、Tn糖基化多肽、Lewis Y/CD174糖基化多肽、唾液酸化Lewis X(sLex) (也称为唾液酸化SSEA-1 (SLX))糖基化多肽、NeuGcGM3糖基化多肽及其组合。在一些实施方案中,一组检测探针包含两个检测探针,每个探针针对相同的表面蛋白。在一些实施方案中,一组检测探针包含两个检测探针,每个探针针对不同的表面蛋白。

[0516] 实施例4:作为肺癌生物标志物的细胞外囊泡中的mRNA(囊泡内mRNA)的评估

[0517] 在一些实施方案中,肺癌检测包括在免疫亲和捕获细胞外囊泡后检测至少一个或多个囊泡内mRNA。

[0518] 在一些实施方案中,存在于细胞外囊泡上的一种或多种表面蛋白或细胞外膜蛋白(“捕获蛋白”)可以用于免疫亲和捕获肺癌相关细胞外囊泡。此类捕获蛋白生物标志物的实例可以包括但不限于ALCAM、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CLDN3、CLDN4、DSG2、EGFR、EPCAM、FOLR1、GJB1、GJB2、IG1FR、LAMB3、MET、MSLN、MUC1、PIGT、PODXL2、ROS1、SDC1、SLC34A2、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMPRSS4、TSPAN8、TNFRSF10B及其组合。

[0519] 在一些实施方案中,EV核酸检测测定(例如,使用引物-探针组的逆转录PCR)和生物标志物验证过程(例如,本文诸如实施例1中所述的那些)可以用于评估肺癌的mRNA生物标志物候选物。在一些实施方案中,将针对捕获蛋白(例如,肺癌相关EV中存在的表面蛋白)的抗体与磁珠缀合,并任选地首先在细胞系EV上然后在患者样品上评估其结合特异性靶蛋白的能力。评估抗体涂覆的珠捕获肺癌相关EV的能力,并分析抗体涂覆的珠捕获的EV的mRNA含量,例如,使用一步定量逆转录PCR(RT-qPCR)主混合物。

[0520] 在一些实施方案中,可以通过检测以下mRNA中的至少一个或多个(例如,1个、2个、3个或更多个)读取捕获的EV:ABCC3、AOC1、ARSL、B3GNT3、C12orf45、CDCP1、CDH1、CDH3、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CRABP2、CST1、DMBT1、DSG2、EPCAM、EPHX3、ETV4、EVA1A、FAM83A、FOLR1、FOXA2、GJB1、GJB2、GPC4、HMGB3、HS6ST2、KDEL3、KRTCAP3、LAMB3、LFNG、LGALS3BP、LSR、MANEAL、MIF、MSLN、MUC1、MUC21、NAPSA、PIGT、PODXL2、PPP1R14D、PRRG4、ROS1、S100A14、SBK1、SCGB3A2、SDC1、SERINC2、SEZ6L2、SFTA2、SFTP1A1、SFTP2A2、SFTP2B、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、SPINK1、ST14、TGFA、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、ZC3H11A及其组合。在一些实施方案中,可以通过检测使用RT-qPCR的以下mRNA中的至少一个或多个(例如,1个、2个、3个或更多个)读取捕获的EV:ABCA3、ABCC1、ABRACL、ACP5、ADAM23、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、ANTXR1、AP1M2、APOBEC3B、APOBEC3C、AQP3、AREG、ARNTL2、ASF1B、ATP8B1、AURKB、B3GNT5、BAIAP2L1、BCAM、BIK、BIRC5、C15orf48、

C19orf33、C1S、C8orf4、CA12、CA9、CALML3、CAPNS2、CBLC、CCL19、CCL5、CCNB2、CD109、CD24、CD53、CD74、CD9、CDC20、CDC42EP1、CDC45、CDCA4、CDCA5、CDCP1、CDH1、CDH3、CDK1、CDKN2A、CDKN2B、CEACAM5、CEACAM6、CELSR1、CENPW、CEP55、CES1、CHMP4C、CLCA2、CLDN1、CLDN4、CLDN7、CNN2、COL17A1、CPA3、CRABP2、CSTA、CTSC、CTSE、CX3CL1、CXADR、CXCR4、CYBB、CYP2S1、CYP4F11、DAPL1、DPYSL3、DSC2、DSC3、DSG2、DSG3、DSP、EFNA1、EFS、EGFR、EGLN3、EHD2、EHF、ELF3、ELF4、EMP1、EMP2、ENAH、EPCAM、EPAH2、EPHB3、EPHX1、ESRP1、EVPL、F11R、F2R、F2RL1、F3、FAM129B、FAM60A、FAM83D、FAM83H、FAT1、FAT2、FBLIM1、FBP1、FCER1G、FERMT1、FGFR2、FGFR3、FOXO1、FOXO3、FXR1、GALNT3、GBP6、GJA1、GJB2、GJB3、GJB5、GJB6、GNA15、GPC1、GPC3、GPNMB、GPR87、GPRC5A、GPX2、GRHL2、GSTA1、HAS3、HCK、HOXB7、ID1、IGF2BP2、IGSF9、IL2RG、IMPA2、IRF6、ITGA2、ITGA6、ITGB4、ITGB6、IVL、JAG2、JUP、KCNS3、KIAA1522、KIF2C、KIF11、KITLG、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、KRTCAP3、LAMP3、LAPTM5、LGALS7B、LRP11、LRRC4、LSP1、LSR、LYPD3、MAGEA4、MAGEA6、MAL2、MAOA、MARCO、MCM2、MDF1、MET、MMP14、MPZL2、MUC1、MYBL2、MYH14、MYO6、MZB1、NCF2、NKG7、NNMT、NOTCH3、NRARP、NTRK2、NUP210、NUSAP1、OSGIN1、OSMR、PALLD、PDPN、PDZK1IP1、PECAM1、PERP、PIGR、PIGT、PITX1、PKP1、PKP3、PLEK、PLEK2、PLVAP、PMP22、POSTN、PPL、PPP1R14C、PRAME、PROM2、PRRG4、PRSS8、PTGES、PTGFRN、PTPN6、PTPRF、PTPRZ1、RAB25、RAB38、RAET1L、RARRES1、RBP1、RGS1、RHCG、RHOF、RIN2、RIPK4、RPS4Y1、RRM2、S100A10、S100A11、S100A14、S100A16、S100A2、S100P、SCNN1A、SDC1、SERINC2、SERPINB13、SERPINB3、SERPINB5、SEZ6L2、SH3BP4、SHISA2、SLC1A5、SLC2A1、SLC34A2、SLC40A1、SLC6A8、SLC7A8、SNAI2、SOX2、SPI1、SPINT1、SPINT2、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、ST14、STEAP1、SULF1、SYK、SYTL1、TACSTD2、TBC1D2、TEAD2、TEAD3、TFAP2C、THBD、THBS2、TK1、TM4SF1、TMC4、TMEM30B、TMEM54、TMPRSS11D、TMPRSS11E、TMPRSS4、TNFRSF18、TNS4、TOP2A、TP53I11、TP63、TPD52、TPX2、TREM2、TRIM29、TRIP13、TSPAN1、TSPAN13、TSPAN6、TSPAN7、TUSC3、TYROBP、UBE2C、UPK1B、VAMP8、VANGL2、WLS、YAP1、ZC3H11A、ZNF217、ZNF750及其组合。

[0521] 在一些实施方案中,可以通过检测以下mRNA中的至少一个或多个(例如,1个、2个、3个或更多个):ABCC3、AOC1、ARSL、B3GNT3、C12orf45、CDCP1、CDH1、CDH3、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CRABP2、CST1、DMBT1、DSG2、EPCAM、EPHX3、ETV4、EVA1A、FAM83A、FOLR1、FOXA2、GJB1、GJB2、GPC4、HMGB3、HS6ST2、KDEL3、KRTCAP3、LAMB3、LFNG、LGALS3BP、LSR、MANEAL、MIF、MSLN、MUC1、MUC21、NAPSA、PIGT、PODXL2、PPP1R14D、PRRG4、ROS1、S100A14、SBK1、SCGB3A2、SDC1、SERINC2、SEZ6L2、SFTA2、SFTP1A1、SFTP1A2、SFTP1B、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、SPINK1、ST14、TGFA、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、ZC3H11A及其组合;和实施例3中所述的EV表面蛋白生物标志物中的至少一个或多个(例如,1个、2个、3个或更多个)读取捕获的EV。在一些实施方案中,可以通过检测使用RT-qPCR的以下mRNA中的至少一个或多个(例如,1个、2个、3个或更多个):ABCA3、ABCC1、ABRACL、ACP5、ADAM23、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、ANTXR1、AP1M2、APOBEC3B、APOBEC3C、AQP3、AREG、ARNTL2、ASF1B、ATP8B1、AURKB、B3GNT5、BAIAP2L1、BCAM、BIK、BIRC5、C15orf48、C19orf33、C1S、C8orf4、CA12、CA9、CALML3、CAPNS2、CBLC、CCL19、CCL5、CCNB2、CD109、CD24、

CD53、CD74、CD9、CDC20、CDC42EP1、CDC45、CDCA4、CDCA5、CDCP1、CDH1、CDH3、CDK1、CDKN2A、CDKN2B、CEACAM5、CEACAM6、CELSR1、CENPW、CEP55、CES1、CHMP4C、CLCA2、CLDN1、CLDN4、CLDN7、CNN2、COL17A1、CPA3、CRABP2、CSTA、CTSC、CTSE、CX3CL1、CXADR、CXCR4、CYBB、CYP2S1、CYP4F11、DAPL1、DPYSL3、DSC2、DSC3、DSG2、DSG3、DSP、EFNA1、EFS、EGFR、EGLN3、EHD2、EHF、ELF3、ELF4、EMP1、EMP2、ENAH、EPCAM、EPAH2、EPHB3、EPHX1、ESRP1、EVPL、F11R、F2R、F2RL1、F3、FAM129B、FAM60A、FAM83D、FAM83H、FAT1、FAT2、FBLIM1、FBP1、FCER1G、FERMT1、FGFR2、FGFR3、FOX E1、FOX M1、FX YD3、GALNT3、GBP6、GJA1、GJB2、GJB3、GJB5、GJB6、GNA15、GPC1、GPC3、GPNMB、GPR87、GPRC5A、GPX2、GRHL2、GSTA1、HAS3、HCK、HOXB7、ID1、IGF2BP2、IGSF9、IL2RG、IMPA2、IRF6、ITGA2、ITGA6、ITGB4、ITGB6、IVL、JAG2、JUP、KCNS3、KIAA1522、KIF2C、KIFC1、KITLG、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、KRTCAP3、LAMP3、LAPTM5、LGALS7B、LRP11、LRR C4、LSP1、LSR、LYPD3、MAGEA4、MAGEA6、MAL2、MAOA、MARCO、MCM2、MDFI、MET、MMP14、MPZL2、MUC1、MYBL2、MYH14、MYOF、MZB1、NCF2、NKG7、NNMT、NOTCH3、NRARP、NTRK2、NUP210、NUSAP1、OSGIN1、OSMR、PALLD、PDPN、PDZK1IP1、PECAM1、PERP、PIGR、PIGT、PITX1、PKP1、PKP3、PLEK、PLEK2、PLVAP、PMP22、POSTN、PPL、PPP1R14C、PRAME、PROM2、PRRG4、PRSS8、PTGES、PTGFRN、PTPN6、PTPRF、PTPRZ1、RAB25、RAB38、RAET1L、RARRES1、RBP1、RGS1、RHCG、RHOV、RIN2、RIPK4、RPS4Y1、RRM2、S100A10、S100A11、S100A14、S100A16、S100A2、S100P、SCNN1A、SDC1、SERINC2、SERPINB13、SERPINB3、SERPINB5、SEZ6L2、SH3BP4、SHISA2、SLC1A5、SLC2A1、SLC34A2、SLC40A1、SLC6A8、SLC7A8、SNAI2、SOX2、SPI1、SPINT1、SPINT2、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、ST14、STEAP1、SULF1、SYK、SYTL1、TACSTD2、TBC1D2、TEAD2、TEAD3、TFAP2C、THBD、THBS2、TK1、TM4SF1、TMC4、TMEM30B、TMEM54、TMPRSS11D、TMPRSS11E、TMPRSS4、TNFRSF18、TNS4、TOP2A、TP53I11、TP63、TPD52、TPX2、TREM2、TRIM29、TRIP13、TSPAN1、TSPAN13、TSPAN6、TSPAN7、TUSC3、TYROBP、UBE2C、UPK1B、VAMP8、VANG L2、WLS、YAP1、ZC3H11A、ZNF217、ZNF750及其组合；和实施例3中所述的EV表面蛋白生物标志物中的至少一个或多个(例如，1个、2个、3个或更多个)读取捕获的EV。

[0522] 在一些实施方案中，可以(i)通过检测以下使用RT-qPCR的mRNA中的一个或多个(例如，1个、2个、3个或更多个)：ABCC3、AOC1、ARSL、B3GNT3、C12orf45、CDCP1、CDH1、CDH3、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CRABP2、CST1、DMBT1、DSG2、EPCAM、EPHX3、ETV4、EVA1A、FAM83A、FOLR1、FOXA2、GJB1、GJB2、GPC4、HMGB3、HS6ST2、KDEL R3、KRTCAP3、LAMB3、LFNG、LGALS3BP、LSR、MANEAL、MIF、MSLN、MUC1、MUC21、NAPSA、PIGT、PODXL2、PPP1R14D、PRRG4、ROS1、S100A14、SBK1、SCGB3A2、SDC1、SERINC2、SEZ6L2、SFTA2、SFTPA1、SFTPA2、SFTP B、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、SPINK1、ST14、TGFA、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、ZC3H11A及其组合；和(ii)通过使用一组检测探针(例如，如本文使用和/或所述的)(所述检测探针中至少一个针对实施例3中所述的EV表面蛋白生物标志物中的一个或多个(例如，1个、2个、3个或更多个))读取捕获的EV。在一些实施方案中，可以(i)通过检测以下使用RT-qPCR的mRNA中的一个或多个(例如，1个、2个、3个或更多个)：ABCA3、ABCC1、ABRACL、ACP5、ADAM23、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、ANTXR1、AP1M2、APOBEC3B、APOBEC3C、AQP3、AREG、ARNTL2、ASF1B、ATP8B1、AURKB、B3GNT5、BAIAP2L1、BCAM、BIK、BIRC5、C15orf48、

C19orf33、C1S、C8orf4、CA12、CA9、CALML3、CAPNS2、CBLC、CCL19、CCL5、CCNB2、CD109、CD24、CD53、CD74、CD9、CDC20、CDC42EP1、CDC45、CDCA4、CDCA5、CDCP1、CDH1、CDH3、CDK1、CDKN2A、CDKN2B、CEACAM5、CEACAM6、CELSR1、CENPW、CEP55、CES1、CHMP4C、CLCA2、CLDN1、CLDN4、CLDN7、CNN2、COL17A1、CPA3、CRABP2、CSTA、CTSC、CTSE、CX3CL1、CXADR、CXCR4、CYBB、CYP2S1、CYP4F11、DAPL1、DPYSL3、DSC2、DSC3、DSG2、DSG3、DSP、EFNA1、EFS、EGFR、EGLN3、EHD2、EHF、ELF3、ELF4、EMP1、EMP2、ENAH、EPCAM、EPA2、EPHB3、EPHX1、ESRP1、EVPL、F11R、F2R、F2RL1、F3、FAM129B、FAM60A、FAM83D、FAM83H、FAT1、FAT2、FBLIM1、FBP1、FCER1G、FERMT1、FGFR2、FGFR3、FOX E1、FOX M1、FX YD3、GALNT3、GBP6、GJA1、GJB2、GJB3、GJB5、GJB6、GNA15、GPC1、GPC3、GPNMB、GPR87、GPRC5A、GPX2、GRHL2、GSTA1、HAS3、HCK、HOXB7、ID1、IGF2BP2、IGSF9、IL2RG、IMPA2、IRF6、ITGA2、ITGA6、ITGB4、ITGB6、IVL、JAG2、JUP、KCNS3、KIAA1522、KIF2C、KIFC1、KITLG、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、KRTCAP3、LAMP3、LAPTM5、LGALS7B、LRP11、LRRC4、LSP1、LSR、LYPD3、MAGEA4、MAGEA6、MAL2、MAOA、MARCO、MCM2、MDFI、MET、MMP14、MPZL2、MUC1、MYBL2、MYH14、MYOF、MZB1、NCF2、NKG7、NNMT、NOTCH3、NRARP、NTRK2、NUP210、NUSAP1、OSGIN1、OSMR、PALLD、PDPN、PDZK1IP1、PECAM1、PERP、PIGR、PIGT、PITX1、PKP1、PKP3、PLEK、PLEK2、PLVAP、PMP22、POSTN、PPL、PPP1R14C、PRAME、PROM2、PRRG4、PRSS8、PTGES、PTGFRN、PTPN6、PTPRF、PTPRZ1、RAB25、RAB38、RAET1L、RARRES1、RBP1、RGS1、RHCG、RHOV、RIN2、RIPK4、RPS4Y1、RRM2、S100A10、S100A11、S100A14、S100A16、S100A2、S100P、SCNN1A、SDC1、SERINC2、SERPINB13、SERPINB3、SERPINB5、SEZ6L2、SH3BP4、SHISA2、SLC1A5、SLC2A1、SLC34A2、SLC40A1、SLC6A8、SLC7A8、SNAI2、SOX2、SPI1、SPINT1、SPINT2、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、ST14、STEAP1、SULF1、SYK、SYTL1、TACSTD2、TBC1D2、TEAD2、TEAD3、TFAP2C、THBD、THBS2、TK1、TM4SF1、TMC4、TMEM30B、TMEM54、TMPRSS11D、TMPRSS11E、TMPRSS4、TNFRSF18、TNS4、TOP2A、TP53I11、TP63、TPD52、TPX2、TREM2、TRIM29、TRIP13、TSPAN1、TSPAN13、TSPAN6、TSPAN7、TUSC3、TYROBP、UBE2C、UPK1B、VAMP8、VANGL2、WLS、YAP1、ZC3H11A、ZNF217、ZNF750及其组合；和(ii)通过使用一组检测探针(例如,如本文使用和/或所述的)(所述检测探针中至少一个针对实施例3中所述的EV表面蛋白生物标志物中的一个或多个(例如,1个、2个、3个或更多))读取捕获的EV。

[0523] 在一些此类实施方案中,一组检测探针包含至少一个针对EV表面蛋白的检测探针。在一些此类实施方案中,一组检测探针包含至少两个针对相同EV表面蛋白(具有相同或不同表位)的检测探针。在一些实施方案中,一组检测探针包含至少两个针对不同EV表面蛋白的检测探针。在一些实施方案中,可以使包含EV表面蛋白和囊泡内mRNA的样品与缀合至单链寡核苷酸(例如,DNA)的抗EV表面蛋白抗体(例如,如实施例3中所述的针对EV表面蛋白的抗体)接触,所述单链寡核苷酸用作一对囊泡内mRNA生物标志物(例如,实施例4中所述)中的两个引物之一,使得抗EV表面蛋白抗体与EV表面蛋白结合,同时缀合的单链寡核苷酸与存在于同一样品中的囊泡内mRNA生物标志物杂交。然后添加所述囊泡内mRNA生物标志物对的第二个引物和RT-qPCR探针以进行RT-qPCR,用于检测单个样品中囊泡内mRNA和EV表面蛋白的存在。

[0524] 在一些实施方案中,可以通过检测以下mRNA中的至少一个或多个(例如,1个、2个、3个或更多个):ABCC3、AOC1、ARSL、B3GNT3、C12orf45、CDCP1、CDH1、CDH3、CEACAM5、CEACAM6、

CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CRABP2、CST1、DMBT1、DSG2、EPCAM、EPHX3、ETV4、EVA1A、FAM83A、FOLR1、FOXA2、GJB1、GJB2、GPC4、HMGB3、HS6ST2、KDELR3、KRTCAP3、LAMB3、LFNG、LGALS3BP、LSR、MANEAL、MIF、MSLN、MUC1、MUC21、NAPSA、PIGT、PODXL2、PPP1R14D、PRRG4、ROS1、S100A14、SBK1、SCGB3A2、SDC1、SERINC2、SEZ6L2、SFTA2、SFTP A1、SFTP A2、SFTP B、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、SPINK1、ST14、TGFA、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、ZC3H11A及其组合；和实施例5中所述的EV囊泡内蛋白中的至少一个或多个（例如，1个、2个、3个或更多个）读取捕获的EV。在一些实施方案中，可以（i）通过检测以下使用RT-qPCR的mRNA中的一个或多个（例如，1个、2个、3个或更多个）：ABCC3、AOC1、ARSL、B3GNT3、C12orf45、CDCP1、CDH1、CDH3、CEACAM5、CEACAM6、CELSR1、CLDN18、CLDN3、CLDN4、CLDN7、CLIC6、CRABP2、CST1、DMBT1、DSG2、EPCAM、EPHX3、ETV4、EVA1A、FAM83A、FOLR1、FOXA2、GJB1、GJB2、GPC4、HMGB3、HS6ST2、KDELR3、KRTCAP3、LAMB3、LFNG、LGALS3BP、LSR、MANEAL、MIF、MSLN、MUC1、MUC21、NAPSA、PIGT、PODXL2、PPP1R14D、PRRG4、ROS1、S100A14、SBK1、SCGB3A2、SDC1、SERINC2、SEZ6L2、SFTA2、SFTP A1、SFTP A2、SFTP B、SLC34A2、SLC44A4、SLC6A14、SLC7A7、SMIM22、SMPDL3B、SPINK1、ST14、TGFA、TMC4、TMC5、TMEM45B、TMPRSS2、TMPRSS4、TSPAN1、TSPAN8、ZC3H11A及其组合；和（ii）通过使用一组检测探针（例如，如本文使用和/或所述的）（所述检测探针中至少一个针对实施例5中所述的EV囊泡内蛋白中的一个或多个（例如，1个、2个、3个或更多个））读取捕获的EV。

[0525] 在一些实施方案中，可以通过检测以下mRNA中的至少一个或多个（例如，1个、2个、3个或更多个）：ABCA3、ABCC1、ABRACL、ACP5、ADAM23、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、ANTXR1、AP1M2、APOBEC3B、APOBEC3C、AQP3、AREG、ARNTL2、ASF1B、ATP8B1、AURKB、B3GNT5、BAIAP2L1、BCAM、BIK、BIRC5、C15orf48、C19orf33、C1S、C8orf4、CA12、CA9、CALML3、CAPNS2、CBLC、CCL19、CCL5、CCNB2、CD109、CD24、CD53、CD74、CD9、CDC20、CDC42EP1、CDC45、CDCA4、CDCA5、CDCP1、CDH1、CDH3、CDK1、CDKN2A、CDKN2B、CEACAM5、CEACAM6、CELSR1、CENPW、CEP55、CES1、CHMP4C、CLCA2、CLDN1、CLDN4、CLDN7、CNN2、COL17A1、CPA3、CRABP2、CSTA、CTSC、CTSE、CX3CL1、CXADR、CXCR4、CYBB、CYP2S1、CYP4F11、DAPL1、DPYSL3、DSC2、DSC3、DSG2、DSG3、DSP、EFNA1、EFS、EGFR、EGLN3、EHD2、EHF、ELF3、ELF4、EMP1、EMP2、ENAH、EPCAM、EPAH2、EPHB3、EPHX1、ESRP1、EVPL、F11R、F2R、F2RL1、F3、FAM129B、FAM60A、FAM83D、FAM83H、FAT1、FAT2、FBLIM1、FBP1、FCER1 G、FERMT1、FGFR2、FGFR3、FOX E1、FOX M1、FXYD3、GALNT3、GBP6、GJA1、GJB2、GJB3、GJB5、GJB6、GNA15、GPC1、GPC3、GPNMB、GPR87、GPRC5A、GPX2、GRHL2、GSTA1、HAS3、HCK、HOXB7、ID1、IGF2BP2、IGSF9、IL2RG、IMPA2、IRF6、ITGA2、ITGA6、ITGB4、ITGB6、IVL、JAG2、JUP、KCNS3、KIAA1522、KIF2C、KIFC1、KITLG、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、KRTCAP3、LAMP3、LAPTM5、LGALS7B、LRP11、LRRC4、LSP1、LSR、LYPD3、MAGEA4、MAGEA6、MAL2、MAOA、MARCO、MCM2、MDFI、MET、MMP14、MPZL2、MUC1、MYBL2、MYH14、MYOF、MZB 1、NCF2、NKG7、NNMT、NOTCH3、NRARP、NTRK2、NUP210、NUSAP1、OSGIN1、OSMR、PALLD、PDPN、PDZK1 IP1、PECAM1、PERP、PIGR、PIGT、PITX1、PKP1、PKP3、PLEK、PLEK2、PLVAP、PMP22、POSTN、PPL、PPP1R14C、PRAME、PROM2、PRRG4、PRSS8、PTGES、PTGFRN、PTPN6、PTPRF、PTPRZ1、RAB25、RAB38、RAET1L、RARRES1、RBP1、RGS1、RHCG、RHOV、RIN2、RIPK4、RPS4Y1、RRM2、S100A10、S100A11、S100A14、S100A16、S100A2、S100P、SCNN1A、SDC1、SERINC2、SERPINB13、

SERPINB3、SERPINB5、SEZ6L2、SH3BP4、SHISA2、SLC1A5、SLC2A1、SLC34A2、SLC40A1、SLC6A8、SLC7A8、SNAI2、SOX2、SPI1、SPINT1、SPINT2、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、ST14、STEAP1、SULF1、SYK、SYTL1、TACSTD2、TBC1D2、TEAD2、TEAD3、TFAP2C、THBD、THBS2、TK1、TM4SF1、TMC4、TMEM30B、TMEM54、TMPRSS11D、TMPRSS 11E、TMPRSS4、TNFRSF18、TNS4、TOP2A、TP53I11、TP63、TPD52、TPX2、TREM2、TRIM29、TRIP13、TSPAN1、TSPAN13、TSPAN6、TSPAN7、TUSC3、TYROBP、UBE2C、UPK1B、VAMP8、VANG2、WLS、YAP1、ZC3H11A、ZNF217、ZNF750及其组合；和实施例5中所述的EV囊泡内蛋白中的至少一个或多个(例如，1个、2个、3个或更多个)读取捕获的EV。在一些实施方案中，可以(i)通过检测以下使用RT-qPCR的mRNA中的一个或多个(例如，1个、2个、3个或更多个)：ABCA3、ABCC1、ABRACL、ACP5、ADAM23、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、ANTXR1、AP1M2、APOBEC3B、APOBEC3C、AQP3、AREG、ARNTL2、ASF1B、ATP8B1、AURKB、B3GNT5、BAIAP2L1、BCAM、BIK、BIRC5、C15orf48、C19orf33、C1S、C8orf4、CA12、CA9、CALML3、CAPNS2、CBLC、CCL19、CCL5、CCNB2、CD109、CD24、CD53、CD74、CD9、CDC20、CDC42EP1、CDC45、CDCA4、CDCA5、CDCP1、CDH1、CDH3、CDK1、CDKN2A、CDKN2B、CEACAM5、CEACAM6、CELSR1、CENPW、CEP55、CES1、CHMP4C、CLCA2、CLDN1、CLDN4、CLDN7、CNN2、COL17A1、CPA3、CRABP2、CSTA、CTSC、CTSE、CX3CL1、CXADR、CXCR4、CYBB、CYP2S1、CYP4F11、DAPL1、DPYSL3、DSC2、DSC3、DSG2、DSG3、DSP、EFNA1、EFS、EGFR、EGLN3、EHD2、EHF、ELF3、ELF4、EMP1、EMP2、ENAH、EPCAM、EPHA2、EPHB3、EPHX1、ESRP1、EVPL、F11R、F2R、F2RL1、F3、FAM129B、FAM60A、FAM83D、FAM83H、FAT1、FAT2、FBLIM1、FBP1、FCER1G、FERMT1、FGFR2、FGFR3、FOXO1、FOXO3、FOXO4、GALNT3、GBP6、GJA1、GJB2、GJB3、GJB5、GJB6、GNA15、GPC1、GPC3、GPNMB、GPR87、GPRC5A、GPX2、GRHL2、GSTA1、HAS3、HCK、HOXB7、ID1、IGF2BP2、IGSF9、IL2RG、IMPA2、IRF6、ITGA2、ITGA6、ITGB4、ITGB6、IVL、JAG2、JUP、KCNS3、KIAA1522、KIF2C、KIF14、KITLG、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、KRTCAP3、LAMP3、LAPTM5、LGALS7B、LRP11、LRRC4、LSP1、LSR、LYPD3、MAGEA4、MAGEA6、MAL2、MAOA、MARCO、MCM2、MDF1、MET、MMP14、MPZL2、MUC1、MYBL2、MYH14、MYO6、MZB1、NCF2、NKG7、NNMT、NOTCH3、NRARP、NTRK2、NUP210、NUSAP1、OSGIN1、OSMR、PALLD、PDPN、PDZK1IP1、PECAM1、PERP、PIGR、PIGT、PITX1、PKP1、PKP3、PLEK、PLEK2、PLVAP、PMP22、POSTN、PPL、PPP1R14C、PRAME、PROM2、PRRG4、PRSS8、PTGES、PTGFRN、PTPN6、PTPRF、PTPRZ1、RAB25、RAB38、RAET1L、RARRES1、RBP1、RGS1、RHC3、RHOF、RIN2、RIPK4、RPS4Y1、RRM2、S100A10、S100A11、S100A14、S100A16、S100A2、S100P、SCNN1A、SDC1、SERINC2、SERPINB13、SERPINB3、SERPINB5、SEZ6L2、SH3BP4、SHISA2、SLC1A5、SLC2A1、SLC34A2、SLC40A1、SLC6A8、SLC7A8、SNAI2、SOX2、SPI1、SPINT1、SPINT2、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、ST14、STEAP1、SULF1、SYK、SYTL1、TACSTD2、TBC1D2、TEAD2、TEAD3、TFAP2C、THBD、THBS2、TK1、TM4SF1、TMC4、TMEM30B、TMEM54、TMPRSS11D、TMPRSS11E、TMPRSS4、TNFRSF18、TNS4、TOP2A、TP53I11、TP63、TPD52、TPX2、TREM2、TRIM29、TRIP13、TSPAN1、TSPAN13、TSPAN6、TSPAN7、TUSC3、TYROBP、UBE2C、UPK1B、VAMP8、VANG2、WLS、YAP1、ZC3H11A、ZNF217、ZNF750及其组合；和(ii)通过使用一组检测探针(例如，如本文使用和/或所述的)(所述检测探针中至少一个针对实施例5中所述的囊泡内蛋白中的一个或多个(例如，1个、2个、3个或更多个))读取捕获的EV。

[0526] 在一些实施方案中，一组检测探针包含至少一个针对囊泡内蛋白的检测探针(例

如,如本文所述)。在一些实施方案中,一组检测探针包含至少两个检测探针,每个检测探针针对相同的囊泡内蛋白(例如,具有相同或不同表位)。在一些实施方案中,一组检测探针包含至少两个检测探针,每个探针针对不同的囊泡内蛋白(例如,如本文所述)。在一些实施方案中,可以使包含EV囊泡内蛋白和囊泡内mRNA的样品与缀合至单链寡核苷酸(例如,DNA)的抗EV囊泡内蛋白抗体(例如,如实施例5中所述的针对EV囊泡内蛋白的抗体)接触,所述单链寡核苷酸用作一对囊泡内mRNA生物标志物(例如,实施例4中所述)中的两个引物之一,使得抗EV囊泡内蛋白抗体与EV囊泡内蛋白结合,同时缀合的单链寡核苷酸与存在于同一样品中的囊泡内mRNA生物标志物杂交。然后添加所述囊泡内mRNA生物标志物对的第二个引物和RT-qPCR探针以进行RT-qPCR,用于检测单个样品中囊泡内mRNA和囊泡内蛋白的存在。

[0527] 实施例5:评估作为卵巢癌生物标志物的囊泡内蛋白

[0528] 在一些实施方案中,肺癌检测包括在免疫亲和捕获细胞外囊泡后至少检测一个或多个囊泡内蛋白。

[0529] 在一些实施方案中,存在于细胞外囊泡上的一种或多种表面蛋白或细胞外膜蛋白(“捕获蛋白”)可以用于免疫亲和捕获肺癌相关细胞外囊泡。此类捕获蛋白生物标志物的实例可以包括但不限于ALCAM、B3GNT3、CDCP1、CDH1、CDH3、CD55、CD274 (PD-L1)、CEACAM5、CEACAM6、CLDN3、CLDN4、DSG2、EGFR、EPCAM、FOLR1、GJB1、GJB2、IG1FR、LAMB3、MET、MSLN、MUC1、PIGT、PODXL2、ROS1、SDC1、SLC34A2、SMPDL3B、ST14、sTn抗原、Tn抗原、T抗原、TACSTD2、TMPRSS4、TSPAN8、TNFRSF10B及其组合。

[0530] 在一些实施方案中,EV免疫测方法(例如,本文诸如实施例1中所述的那些)和生物标志物验证过程(例如,本文诸如实施例1中描述)可以用于评估作为肺癌的生物标志物的囊泡内蛋白。在一些实施方案中,将针对捕获蛋白(例如,肺癌相关EV中存在的表面蛋白)的抗体与磁珠缀合,并首先在细胞系EV上然后在患者样品上评估其结合特定靶蛋白的能力。评估抗体涂覆的珠捕获肺癌相关EV的能力,并且被抗体涂覆的珠捕获的EV在使用靶实体检测系统(例如,如本文所述的双链体系统,所述系统涉及一组两个检测探针,每个检测探针针对与捕获蛋白不同的靶标志物)分析其囊泡内蛋白前被固定和/或透化。

[0531] 在一些实施方案中,可以使用以下囊泡内蛋白中的至少一个或多个(例如,1个、2个、3个或更多个)读取固定和/或透化后的捕获的EV:AOC1、C12orf45、CRABP2、CST1、ETV4、FAM83A、FOXA2、HMGB3、LGALS3BP、MIF、NAPSA、PPP1R14D、S100A14、SBK1、SCGB3A2、SFTA2、SFTPA1、SFTPA2、SFTPB、SPINK1、TGFA、ZC3H11A及其组合。在一些实施方案中,可以使用一组检测探针(例如,如本文所用和/或描述的)读取固定和/或透化后的捕获的EV,所述检测探针中的至少两个针对以下囊泡内蛋白中的一个或多个(例如,1个、2个、3个或更多个):AOC1、C12orf45、CRABP2、CST1、ETV4、FAM83A、FOXA2、HMGB3、LGALS3BP、MIF、NAPSA、PPP1R14D、S100A14、SBK1、SCGB3A2、SFTA2、SFTPA1、SFTPA2、SFTPB、SPINK1、TGFA、ZC3H11A及其组合。在一些实施方案中,一组检测探针包含两个检测探针,每个探针针对相同的囊泡内蛋白。在一些实施方案中,一组检测探针包含两个检测探针,每个探针针对不同的囊泡内蛋白。

[0532] 在一些实施方案中,可以使用以下囊泡内蛋白中的至少一个或多个(例如,1个、2个、3个或更多个)读取固定和/或透化后的捕获的EV:ABRACL、ACP5、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、AP1M2、APOBEC3B、APOBEC3C、

ARNTL2、ASF1B、AURKB、BAIAP2L1、BIRC5、C15orf48、C19orf33、C1S、C8orf4、CA9、CALML3、CAPNS2、CBLC、CCL19、CCNB2、CDC20、CDC45、CDCA4、CDCA5、CDK1、CDKN2A、CDKN2B、CENPW、CEP55、CES1、CHMP4C、CNN2、CPA3、CRABP2、CSTA、CTSC、CTSE、CYP2S1、DPYSL3、EFS、EGLN3、EHF、ELF3、ELF4、ENAH、ESRP1、EVPL、FAM129B、FAM60A、FAM83D、FAM83H、FBP1、FERMT1、FOX E1、FOX M1、GBP6、GNA15、GPX2、GRHL2、GSTA1、HCK、HOXB7、ID1、IGF2BP2、IMPA2、IRF6、IVL、JUP、KIAA1522、KIF2C、KIFC1、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、LGALS7B、LSP1、MAGEA4、MAGEA6、MCM2、MDFI、MYBL2、MYH14、MZB1、NCF2、NNMT、NRARP、NUP210、NUSAP1、OSGIN1、PALLD、PITX1、PKP1、PKP3、PLEK、PLEK2、POSTN、PPP1R14C、PRAME、PTPN6、RBP1、RIN2、RIPK4、RPS4Y1、RRM2、S100A11、S100A14、S100A16、S100A2、S100P、SERPINB13、SERPINB3、SERPINB5、SH3BP4、SNAI2、SOX2、SPI1、SPINT1、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、SULF1、SYK、SYTL1、TBC1D2、TEAD2、TEAD3、TFAP2C、THBS2、TK1、TOP2A、TP63、TPD52、TPX2、TRIM29、TRIP13、UBE2C、YAP1、ZC3H11A、ZNF217、ZNF750及其组合。在一些实施方案中,可以使用一组检测探针(例如,如本文所用和/或描述的)读取固定和/或透化后的捕获的EV,所述检测探针中的至少两个针对以下囊泡内蛋白中的一个或多个(例如,1个、2个、3个或更多个):ABRACL、ACP5、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、AP1M2、APOBEC3B、APOBEC3C、ARNTL2、ASF1B、AURKB、BAIAP2L1、BIRC5、C15orf48、C19orf33、C1S、C8orf4、CA9、CALML3、CAPNS2、CBLC、CCL19、CCNB2、CDC20、CDC45、CDCA4、CDCA5、CDK1、CDKN2A、CDKN2B、CENPW、CEP55、CES1、CHMP4C、CNN2、CPA3、CRABP2、CSTA、CTSC、CTSE、CYP2S1、DPYSL3、EFS、EGLN3、EHF、ELF3、ELF4、ENAH、ESRP1、EVPL、FAM129B、FAM60A、FAM83D、FAM83H、FBP1、FERMT1、FOX E1、FOX M1、GBP6、GNA15、GPX2、GRHL2、GSTA1、HCK、HOXB7、ID1、IGF2BP2、IMPA2、IRF6、IVL、JUP、KIAA1522、KIF2C、KIFC1、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、LGALS7B、LSP1、MAGEA4、MAGEA6、MCM2、MDFI、MYBL2、MYH14、MZB1、NCF2、NNMT、NRARP、NUP210、NUSAP1、OSGIN1、PALLD、PITX1、PKP1、PKP3、PLEK、PLEK2、POSTN、PPP1R14C、PRAME、PTPN6、RBP1、RIN2、RIPK4、RPS4Y1、RRM2、S100A11、S100A14、S100A16、S100A2、S100P、SERPINB13、SERPINB3、SERPINB5、SH3BP4、SNAI2、SOX2、SPI1、SPINT1、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、SULF1、SYK、SYTL1、TBC1D2、TEAD2、TEAD3、TFAP2C、THBS2、TK1、TOP2A、TP63、TPD52、TPX2、TRIM29、TRIP13、UBE2C、YAP1、ZC3H11A、ZNF217、ZNF750及其组合。在一些实施方案中,一组检测探针包含两个检测探针,每个探针针对相同的囊泡内蛋白。在一些实施方案中,一组检测探针包含两个检测探针,每个探针针对不同的囊泡内蛋白。

[0533] 在一些实施方案中,可以使用(i)以下囊泡内蛋白中的至少一个或多个(例如,1个、2个、3个或更多个):AOC1、C12orf45、CRABP2、CST1、ETV4、FAM83A、FOXA2、HMGB3、LGALS3BP、MIF、NAPSA、PPP1R14D、S100A14、SBK1、SCGB3A2、SFTA2、SFTPA1、SFTPA2、SFTPB、SPINK1、TGFA、ZC3H11A及其组合;和(ii)实施例3中所述的EV表面蛋白中的至少一个或多个(例如,1个、2个、3个或更多个)读取固定和/或透化后的捕获的EV。在一些实施方案中,可以使用一组检测探针(例如,如本文所用和/或描述的)读取固定和/或透化后的捕获的EV,所述检测探针包括(i)针对以下囊泡内蛋白中的一个或多个(例如,1个、2个、3个或更多个)的第一检测探针:AOC1、C12orf45、CRABP2、CST1、ETV4、FAM83A、FOXA2、HMGB3、LGALS3BP、MIF、

NAPSA、PPP1R14D、S100A14、SBK1、SCGB3A2、SFTA2、SFTPA1、SFTPA2、SFTPB、SPINK1、TGFA、ZC3H11A及其组合；和(ii)针对实施例3中所述的EV表面蛋白中的一个或多个(例如,1个、2个、3个或更多个)的第二检测探针。

[0534] 在一些实施方案中,可以使用(i)以下囊泡内蛋白中的至少一个或多个(例如,1个、2个、3个或更多个):ABRACL、ACP5、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、AP1M2、APOBEC3B、APOBEC3C、ARNTL2、ASF1B、AURKB、BAIAP2L1、BIRC5、C15orf48、C19orf33、C1S、C8orf4、CA9、CALML3、CAPNS2、CBLC、CCL19、CCNB2、CDC20、CDC45、CDCA4、CDCA5、CDK1、CDKN2A、CDKN2B、CENPW、CEP55、CES1、CHMP4C、CNN2、CPA3、CRABP2、CSTA、CTSC、CTSE、CYP2S1、DPYSL3、EFS、EGLN3、EHF、ELF3、ELF4、ENAH、ESRP1、EVPL、FAM129B、FAM60A、FAM83D、FAM83H、FBP1、FERMT1、FOX E1、FOX M1、GBP6、GNA15、GPX2、GRHL2、GSTA1、HCK、HOXB7、ID1、IGF2BP2、IMPA2、IRF6、IVL、JUP、KIAA1522、KIF2C、KIFC1、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、LGALS7B、LSP1、MAGEA4、MAGEA6、MCM2、MDFI、MYBL2、MYH14、MZB1、NCF2、NNMT、NRARP、NUP210、NUSAP1、OSGIN1、PALLD、PITX1、PKP1、PKP3、PLEK、PLEK2、POSTN、PPP1R14C、PRAME、PTPN6、RBP1、RIN2、RIPK4、RPS4Y1、RRM2、S100A11、S100A14、S100A16、S100A2、S100P、SERPINB13、SERPINB3、SERPINB5、SH3BP4、SNAI2、SOX2、SPI1、SPINT1、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、SULF1、SYK、SYTL1、TBC1D2、TEAD2、TEAD3、TFAP2C、THBS2、TK1、TOP2A、TP63、TPD52、TPX2、TRIM29、TRIP13、UBE2C、YAP1、ZC3H11A、ZNF217、ZNF750及其组合；和(ii)实施例3中所述的EV表面蛋白中的至少一个或多个(例如,1个、2个、3个或更多个)读取固定和/或透化后的捕获的EV。在一些实施方案中,可以使用一组检测探针(例如,如本文所用和/或描述的)读取固定和/或透化后的捕获的EV,所述检测探针包括(i)针对以下囊泡内蛋白中的一个或多个(例如,1个、2个、3个或更多个)的第一检测探针:ABRACL、ACP5、ADH7、AGR2、AIF1、AKR1C1、AKR1C2、AKR1C3、ALDH1A1、ALDH3A1、ALDH3B2、ALG1L、AP1M2、APOBEC3B、APOBEC3C、ARNTL2、ASF1B、AURKB、BAIAP2L1、BIRC5、C15orf48、C19orf33、C1S、C8orf4、CA9、CALML3、CAPNS2、CBLC、CCL19、CCNB2、CDC20、CDC45、CDCA4、CDCA5、CDK1、CDKN2A、CDKN2B、CENPW、CEP55、CES1、CHMP4C、CNN2、CPA3、CRABP2、CSTA、CTSC、CTSE、CYP2S1、DPYSL3、EFS、EGLN3、EHF、ELF3、ELF4、ENAH、ESRP1、EVPL、FAM129B、FAM60A、FAM83D、FAM83H、FBP1、FERMT1、FOX E1、FOX M1、GBP6、GNA15、GPX2、GRHL2、GSTA1、HCK、HOXB7、ID1、IGF2BP2、IMPA2、IRF6、IVL、JUP、KIAA1522、KIF2C、KIFC1、KLF4、KLF5、KRT13、KRT14、KRT15、KRT16、KRT17、KRT18、KRT19、KRT5、KRT6A、KRT6B、KRT6C、KRT7、KRT8、LGALS7B、LSP1、MAGEA4、MAGEA6、MCM2、MDFI、MYBL2、MYH14、MZB1、NCF2、NNMT、NRARP、NUP210、NUSAP1、OSGIN1、PALLD、PITX1、PKP1、PKP3、PLEK、PLEK2、POSTN、PPP1R14C、PRAME、PTPN6、RBP1、RIN2、RIPK4、RPS4Y1、RRM2、S100A11、S100A14、S100A16、S100A2、S100P、SERPINB13、SERPINB3、SERPINB5、SH3BP4、SNAI2、SOX2、SPI1、SPINT1、SPRR1A、SPRR1B、SPRR2A、SPRR2D、SPRR2E、SPRR3、SULF1、SYK、SYTL1、TBC1D2、TEAD2、TEAD3、TFAP2C、THBS2、TK1、TOP2A、TP63、TPD52、TPX2、TRIM29、TRIP13、UBE2C、YAP1、ZC3H11A、ZNF217、ZNF750及其组合；和(ii)针对实施例3中所述的EV表面蛋白中的一个或多个(例如,1个、2个、3个或更多个)的第二检测探针。

[0535] 在一些实施方案中,可通过在单个样品中一起检测EV囊泡内蛋白和EV囊泡内mRNA来读取固定和/或透化后的捕获的EV,如上文实施例4中所述。

[0536] 实施例6:开发肺癌液体活检测定

[0537] 本实施例描述了例如用于筛查遗传风险、环境风险和平均风险个体的肺癌液体活检测定的开发。尽管肺癌是所有癌症中男性和女性的头号杀手(Barta等人,2019;其出于本文所述目的以引用的方式并入本文),尽管低剂量胸部X射线断层摄影术(LDCT)是美国预防服务工作组(USPSTF)认可的,但目前还没有低成本、推荐的以及可及性高的肺癌筛查工具(肺癌筛查(Lung Cancer Screening),PDQ®,2020)。这部分是由于提出的肺癌筛查技术诸如X射线和/或痰液分析的性能不佳(Oken等人,2011;其出于本文所述目的以引用的方式并入本文)。考虑到肺癌的发病率,不足的测试特异性(<99.5%)导致假阳性结果,并可能导致鉴定一些低攻击性的肺癌。这给医疗保健系统和被筛查的个体带来了巨大的负担,因为假阳性结果导致额外的测试、不必要的手术和情绪/身体困扰。因此,期望开发可以表现出两个提供临床效用的特征的肺癌筛查测试:(1)超高的特异性(>99.5%),以最小化假阳性的数量;和(2)预后最佳时对I期和II期肺癌的高灵敏度(>40%)。开发此类测试有可能每年挽救数万人的生命。

[0538] 已经针对肺癌液体活检测定研究了几种不同的生物标志物类别,包括循环肿瘤DNA(ctDNA)、循环肿瘤细胞(CTC)、批量蛋白质和细胞外囊泡(EV)。相对于ctDNA和CTC,EV由于其在血流中的丰度和稳定性而特别有希望,这表明EV对早期癌症的灵敏度有所提高。此外,EV含有来源于同一细胞的货物(例如,蛋白质、RNA、代谢物),从而提供优于批量蛋白质测量的特异性。虽然已经研究了EV的诊断效用,但大部分工作都与批量EV测量或低通量单EV分析有关。

[0539] 本实施例描述了通过对人血浆中的个体细胞外囊泡(EV)进行分析来进行早期肺癌检测的示例性方法的一个方面。EV(包括外泌体和微囊泡)含有共定位的蛋白质、RNA、代谢物和代表其起源细胞的其他化合物(Kosaka等人,2019;其出于本文所述目的以引用的方式并入本文)。检测单个EV内的战略性选择的共定位标志物可以使得能够鉴定具有超高特异性的细胞类型,包括区分癌细胞与正常组织的能力。与依赖细胞死亡以使生物标志物进入血液(即cfDNA)的其他癌症诊断方法相反,EV由功能细胞以高速率释放。单个细胞已被证明每天在体外释放多达10,000个EV(Balaj等人,2011;其出于本文所述目的以引用的方式并入本文)。此外,普遍认为癌细胞释放EV的速率高于健康细胞(Bebelman等人2018;其出于本文所述目的以引用的方式并入本文)。

[0540] 一方面,本公开提供了涉及使用申请人的专有生物信息学生物标志物发现过程鉴定在肺癌与健康组织中上调的基因的见解和技术。从上调的生物标志物列表中,设计了预计对肺癌表现出高灵敏度和特异性的生物标志物组合。使用示例性单个EV测定(参见例如,图1或2中所示和/或本文所述),检测到此类生物标志物在单独囊泡上的共定位,表明生物标志物的分组来源于相同的细胞。鉴于许多上调的癌症生物标志物可以由一种或多种健康组织表达,这为批量生物标志物测量(包括批量EV测定)提供了优越的特异性。通过仔细设计生物标志物组合,可以减少或消除来自竞争组织的信号,包括与肺癌密切相关的信号。在一些实施方案中,本公开提供了具有特别有助于肺癌筛查测试的高或超高特异性的技术,所述技术的阳性预测值寻求至少与胸部CT相当的阳性预测值,所述胸部CT是用于筛查高风险吸烟者的金标准。参见例如,美国国家肺癌筛查试验研究小组(2013)“Results of initial low-dose computed tomographic screening for lung cancer”New England

Journal of Medicine, 368(21):1980-1991。在一些实施方案中,本公开提供了有助于肺癌筛查的技术,在高风险群体中的阳性预测值至少为3.8%或更高(包括例如,至少4%、至少5%、至少6%、至少7%、至少8%、至少9%、至少10%、至少15%、至少20%、至少25%或更高)。

[0541] 生物标志物发现

[0542] 在一些实施方案中,生物标志物发现过程利用对大型数据库的生物信息学分析以及对肺癌和细胞外囊泡的生物学的理解。

[0543] 个体细胞外囊泡分析

[0544] 检测血液中肿瘤来源的EV需要具有足够选择性和灵敏度的测定,以在来自各种各样的健康组织的100亿个EV的背景下检测每毫升血浆中相对较少的肿瘤来源的EV。本公开尤其提供了解决此挑战的技术。例如,在一些实施方案中,用于个体细胞外囊泡分析的测定在图1中说明,其在如下概述的三个关键步骤中进行:

[0545] 1. 使用尺寸排阻色谱 (SEC) 从患者血浆中纯化EV,所述尺寸排阻色谱去除超过99%的可溶性蛋白质和其他干扰化合物。

[0546] 2. 使用对膜结合蛋白具有特异性的抗体功能化磁珠捕获肿瘤特异性EV。

[0547] 3. 进行邻近连接免疫定量聚合酶链反应 (pliq-PCR) 的改良版本,以确定包含在捕获的EV上或其中的另外的蛋白生物标志物的共定位。

[0548] 在pliq-PCR测定的改良版本的许多实施方案中,将两种或多种不同的抗体-寡核苷酸缀合物添加到由抗体功能化磁珠捕获的EV,并且抗体随后与其蛋白质靶标结合。寡核苷酸由具有互补的单链悬突的dsDNA组成,并因此在紧密邻近时(即,当相应的蛋白质靶标位于同一EV上时)能够杂交。在洗去未结合的抗体-寡核苷酸物质后,使用标准DNA连接酶反应连接相邻结合的抗体-寡核苷酸物质。连接模板链的后续qPCR使得能够对共定位的蛋白质物种进行检测和相对定量。在一些实施方案中,可以将二至二十个不同的抗体-寡核苷酸探针并入此类测定中,例如,如均于2020年2月28日提交,标题为“Systems, Compositions, and Methods for Target Entity Detection”的美国申请号16/805,637和国际申请PCT/US2020/020529中所述;其出于任何目的都以引用的方式并入本文。

[0549] 与其他技术相比,pliq-PCR在分析EV方面具有许多优点。例如,pliq-PCR具有比其他标准免疫测定(诸如ELISA)高三个数量级的灵敏度(Darmanis等人,2010;其出于本文所述目的以引用的方式并入本文)。良好优化的pliq-PCR反应的超低LOD(低至一千个EV/mL)使得能够检测痕量水平的肿瘤来源的EV。这与其他新兴EV分析技术相比是有利的,所述技术包括纳米胞质外泌体 (nPLEX) 传感器 (Im等人,2014;其出于本文所述目的以引用的方式并入本文) 和集成磁电化学外泌体 (iMEX) 传感器 (Jeong等人,2016;其出于本文所述目的以引用的方式并入本文),它们的报告LOD分别为约 10^3 和约 10^4 EV (Shao等人,2018;其出于本文所述目的以引用的方式并入本文)。此外,在一些实施方案中,pliq-PCR方法的改良版本不需要复杂的设备,并且可以独特地检测多个生物标志物在单个EV上的共定位。

[0550] 在一些实施方案中,为了进一步提高单个EV谱测定的灵敏度和特异性,包括mRNA和囊泡内蛋白质(除了EV表面蛋白质之外)的其他类别的EV生物标志物录被鉴定并包括在测定中。

[0551] 初步工作

[0552] 通过初步研究,开发了工作流程,其中验证生物标志物候选物存在于EV中,并且能够被可商购获得的抗体或mRNA引物探针组检测到。对于给定的关注的生物标志物,表达(阳性对照)和不表达关注的生物标志物(阴性对照)的一个或多个细胞系可以通过浓缩其细胞培养基并进行纯化以分离具有关注的大小范围的纳米颗粒(例如,使用SEC)来培养以收获其EV。通常,细胞外囊泡的直径可以在30nm至几微米的范围内。参见例如出于本文所述目的以引用的方式并入本文的Chuo等人,“Imaging extracellular vesicles:current and emerging methods”*Journal of Biomedical Sciences* 25:91(2018),并且提供不同细胞外囊泡(EV)亚型的大小信息:迁移小体(0.5 μ m至3 μ m)、微囊泡(0.1 μ m至1 μ m)、原癌小体(1 μ m至10 μ m)、外泌颗粒(exomere)(<50nm)、小外泌体(60nm至80nm)和大外泌体(90nm至120nm)。在一些实施方案中,可以分离大小范围为约30nm至1000nm的纳米颗粒用于检测测定。在一些实施方案中,可以分离一个或多个特定EV亚型用于检测测定。

[0553] 通过专有的生物标志物发现过程,鉴定了在LUAD与健康组织中上调的四种膜结合蛋白生物标志物(SLC34A2、CEACAM5、CEACAM6和EPCAM),并将其用于细胞系EV和肺癌患者样品的概念验证实验。

[0554] 为了检测来自含有共定位的SLC34A2、CEACAM5、CEACAM6和EPCAM标志物的EV的测定信号,在一些实施方案中,开发了涉及SLC34A2免疫亲和捕获和两种不同的对CEACAM6具有特异性的抗体-寡核苷酸探针的测定配置。在一些实施方案中,开发了涉及SLC34A2免疫亲和捕获和两种不同的对CEACAM6和EPCAM具有特异性的抗体-寡核苷酸探针的测定配置。在一些实施方案中,开发了涉及CEACAM5免疫亲和捕获和两种不同的对CEACAM6和SLC34A2具有特异性的抗体-寡核苷酸探针的测定配置。

[0555] 使用抗SLC34A2功能化磁珠或抗EPCAM功能化磁珠捕获纯化的细胞系EV。观察到相对于低表达细胞系,高表达细胞系表现出明显的信号增加。这些观察表明,在一些实施方案中,单个EV谱测定(例如,本文所述的那些)能够以非常高的灵敏度检测单个EV上的共定位的膜结合蛋白标志物。

[0556] 在验证肺癌细胞系EV中的SLC34A2、CEACAM5、CEACAM6和EPCAM的有效性后,使用优化的和操作人员盲法的测定方案,对肺癌患者血浆样品进行了试验群组研究和扩大群组研究。所有血浆样品均购自同一来源,根据相同的采血方案对其进行处理,并且在开始任何治疗之前(即治疗初始(treatment naive))收集患者样品。本实施例的研究中包括的患者群组描述于图4和11。临床试验研究提供于图5至10,而扩大群组研究结果呈现于图12。

[0557] 在临床预试验研究中,通过假设所有健康对照(n=20)的对数正态分布来确定特异性。为了确定截断值,确定了健康对照分布的平均值,并添加了适当量的标准偏差,使得基本上100%的值预期低于截断值。在一些涉及SLC34A2免疫亲和捕获和两个针对CEACAM6具有特异性的不同抗体寡核苷酸探针的检测配置的一些实施方案中,实现了16.7%的I期LUAD灵敏度、60%的II期LUAD灵敏度、50%的III期LUAD灵敏度和100%的IV期LUAD灵敏度。在一些涉及SLC34A2免疫亲和捕获和两个针对CEACAM6和EPCAM具有特异性的不同抗体寡核苷酸探针的检测配置的一些实施方案中,实现了20%的II期LUAD灵敏度、50%的III期LUAD灵敏度和75%的IV期LUAD灵敏度。在一些涉及CEACAM5免疫亲和捕获和两个针对CEACAM6和SLC34A2具有特异性的不同抗体寡核苷酸探针的检测配置的一些实施方案中,实现了16.7%的I期LUAD灵敏度、20%的II期LUAD灵敏度、50%的III期LUAD灵敏度和75%的IV期

LUAD灵敏度。

[0558] 比较这三种示例性LUAD诊断测定相互之间的预测能力,并使用皮尔逊积矩相关系数确定相关性。特别是在III期和IV期LUAD样品(图8至10)中观察到强相关性。

[0559] 图8是如本文所述的示例性肺腺癌诊断测定之间相关性的图形表示。来自基于SLC34A2抗体的用CEACAM6+CEACAM6检测探针的捕获的信号沿x轴进行描绘,而来自基于SLC34A2抗体的用CEACAM6+EpCAM检测探针的捕获的信号沿y轴进行描绘。使用皮尔逊积矩相关系数确定相关性。观察到强相关性,特别是对III期和IV期LUAD样品。

[0560] 图9是如本文所述的示例性肺腺癌诊断测定之间相关性的图形表示。来自基于SLC34A2抗体的用CEACAM6+CEACAM6检测探针的捕获的信号沿x轴进行描绘,而来自基于CEACAM5抗体的用CEACAM6+SLC34A2检测探针的捕获的信号沿y轴进行描绘。使用皮尔逊积矩相关系数确定相关性。观察到强相关性,特别是对III期和IV期LUAD样品。

[0561] 图10是如本文所述的示例性肺腺癌诊断测定之间相关性的图形表示。来自基于SLC34A2抗体的用CEACAM6+EPCAM检测探针的捕获的信号沿x轴进行描绘,而来自基于CAECAM5抗体的用CEACAM6+SLC34A2检测探针的捕获的信号沿y轴进行描绘。使用皮尔逊积矩相关系数确定相关性。观察到强相关性,特别是对III期和IV期LUAD样品。

[0562] 在临床扩大群组研究期间,通过假设所有健康对照($n=90$)的对数正态分布来确定特异性。为了确定截断值,确定了健康对照分布的平均值,并添加了适当量的标准偏差,使得99.9%的值预期低于截断值。在一些涉及SLC34A2免疫亲和捕获和两个针对CEACAM6具有特异性的不同抗体寡核苷酸探针的检测配置的一些实施方案中,实现了50%的II期LUAD灵敏度、55.5%的III期LUAD灵敏度和71.4%的IV期LUAD灵敏度。

[0563] 98%、98.5%、99%、99.5%、99.9%或100%的特异性抗原用于评估用于在风险群体中筛查不同LUAD的PPV,例如,大约每10万吸烟者的发病率,或大约每10万非吸烟者的发病率。

[0564] 为了进一步提高用于检测肺癌的示例性单一EV谱测定(例如,本文所述的那些)的性能,在一些实施方案中,可以鉴定包括膜结合蛋白和囊泡内mRNA/蛋白的其他生物标志物候选物。

[0565] 在一些实施方案中,此前申请人证明了使用批量纯化的细胞系EV的EV-mRNA检测的可行性。通过对膜结合蛋白标志物的免疫亲和捕获,此类方法使得能够检测两个共定位的生物标志物。此外,EV-mRNA检测需要更简单的方案,因为RT-qPCR可以在免疫亲和捕获后直接进行。在一些实施方案中,使用EV的mRNA检测可以通过使用捕获探针(例如,如本文所述)捕获EV和检测特定的肺癌mRNA生物标志物来实现。选择性检测表达捕获探针标志物和肺癌mRNA生物标志物两者的EV。

[0566] 实施例7:生物信息学鉴定的生物标志物和生物标志物组合

[0567] 本实施例展示了用于鉴定某些可以用于肺癌诊断的生物标志物和生物标志物组合的示例性生物信息驱动方法。

[0568] 生物信息过滤

[0569] 在基因型-组织表达(GTEx)数据库(例如,正常组织基因表达的原始数据资源)和癌症基因组图谱(TCGA)数据库(例如,癌症组织基因表达的原始数据资源)中捕获了超过55,000个转录子。为了鉴定对检测肺癌有用的生物标志物,对数据应用了两个过滤步骤。

[0570] 在一些实施方案中,使用UniProt过滤。在分析(例如,没有为膜相关的证据的蛋白质被任选过滤掉)中考虑了具有有效的UniProt条目的生物标志物,其包括发现生物标志物蛋白与膜相关的某种类型的证据。此类过滤步骤可以任选地区分不同关注的膜或提供的证据的置信水平。

[0571] 在一些实施方案中,使用Vesiclepedia过滤。使用Vesiclepedia(细胞外囊泡出版物的储存库)以过滤结果。Vesiclepedia列出了每个基因(例如,Entrez)发表的EV相关参考的数量。这些参考文献被用作EV中或EV上存在给定生物标志物的代表。如果没有针对给定生物标志物的EV相关出版物,那么它不太可能是实际的EV生物标志物,因此从生物标志物列表中过滤出来以供进一步考虑。

[0572] 最小化表达和差异化过滤

[0573] 在一些实施方案中,考虑生物标志物的最低表达水平。低生物标志物表达可以产生随机噪声并使稳定的信号检测变得困难和不可靠。为了克服这一挑战,应用了以下表达过滤器中的一个或多个(包括所有的)。在特定实施方案中,应用了四个表达过滤器。

[0574] 关注的癌症中的最小化表达

[0575] 在一些实施方案中,使用了最少数量的样品来显示在血浆中可检测的表达水平,同时为发现潜在具有差异基因表达谱的亚型留下了空间。为实现此过滤,在一些实施方案中,计算了关注的TCGA癌症(例如,肺癌)中基因表达的第80百分位数,并且在一些实施方案中,考虑了具有>15的每百万转录本(TPM)值的生物标志物。

[0576] 最小表达相关细胞系

[0577] 在一些实施方案中,阳性对照细胞系被用于测试针对生物信息学预测的生物标志物的抗体。在一些实施方案中,使用含有>1000个细胞系谱的癌细胞系百科全书(CCLE)基因表达集来将生物标志物列表减少到存在那些表达关注的生物标志物的细胞系。在一些实施方案中,计算了每个生物标志物在相关细胞系中表达的第90百分位数,并且在一些实施方案中,考虑了TPM>15的生物标志物。

[0578] 在关注的癌症中在蛋白质水平上的最低表达

[0579] 本领域技术人员会理解,并非所有表达的基因最终都被翻译成蛋白质。因此,在一些实施方案中,来自临床蛋白质组学肿瘤分析联盟(CPTAC)的质谱数据被用来过滤蛋白质表达基因。在一些实施方案中,光谱计数大于10的生物标志物被认为是表达的。

[0580] 针对正常组织的最小分化

[0581] 在一些实施方案中,本文所述的测定通过要求至少两个生物标志物(在一些实施方案中,至少三个生物标志物)的共表达在相同细胞外囊泡上实现了优越的特异性。正常组织的简单差异基因表达产生太多的假阴性值。相反,在一些实施方案中,生物标志物特征包含生物标志物的组合,实施生物标志物可以包括在多种组织类型中高表达的生物标志物,但只有当它们与其他生物标志物配对时,才能提供额外的辨别力(例如,高组织特异性和/或高癌症特异性)。然而,此类分析可以捕获普遍存在的管家基因,诸如GAPDH,因此不一定是有用的辨别生物标志物。为了去除此类标志物,在一些实施方案中,计算了比较癌组织(例如,肺癌)和GTEx中每种组织类型对于给定生物标志物的Z评分。在一些实施方案中,选择了具有最大z评分5的生物标志物(例如,至少有一个正常组织被一个生物标志物候选物明确排除)。

[0582] 模拟和随机抽样

[0583] 通过模拟和比较此类生物标志物特征候选物在正常受试者(例如,被确定未患有肺癌的受试者)和癌症受试者中的表达,可以确定生物标志物特征候选物或包含至少两个或更多个(包括例如,至少三个或更多个)生物标志物的生物标志物组合候选物的判别力。基于过滤后的生物标志物组生成至少2个和至少3个生物标志物的组合。通过乘以给定组合中所有标志物的TPM值,为给定组合计算EV评分,所述评分估计由分析的组织生成的EV数量。

[0584] 为模拟正常受试者群体,创建了来自5000个“健康个体”的5000个血浆样品的群组。通过从GTEx数据库中的54个组织中的每一个中随机选取一个组织样品,将这些组织合并成个体,并将表达的基因的TPM值与基于健康个体的每个器官以克计的估计重量相乘来创建个体样品。然后对个体跨组织的EV评分进行加和。然后对模拟个体进行跨组织的EV评分加和。除了健康群组外,通过重复“健康”池生成技术创建了5000个来自“癌症个体”的样品,但增加了一个步骤:将随机选择的来自TCGA的肺癌(例如,LUAD和/或LUSC)样品的EV评分相加,将样品对应1g、10g或100g的肿瘤乘以1、10或100。使用“健康”和“癌症”个体的两个样品池,计算每个生物标志物组合在99%特异性下的灵敏度。然后使用该指标对生物标志物组合候选物进行排序。

[0585] 对于生物标志物特征选择,在一些实施方案中,随机抽样3种生物标志物的100万个组合,并且在一些实施方案中,使用100g肿瘤进行模拟,并购且癌症池和健康池中的每一个有1000个个体。在一些实施方案中,生物标志物组合随后基于其灵敏度值进行排序。在一些实施方案中,单个生物标志物随后基于其在组合列表中排名的前0.5个百分位进行排序。

[0586] 实施例8:生物信息学鉴定的生物标志物和生物标志物组合与本领域已知途径的相关性

[0587] 本实施例描述了用于确定某些生物信息学预测的生物标志物和已公开的基因通路之间的重叠的基因组富集分析。本领域技术人员将认识到,在某些情况下,单个基因的列表可能很难正确翻译。幸运的是,有提供基因的功能列表的资源,例如像编码蛋白质的基因列表,所述基因列表是同一生化通路或现象的组成部分。比较生物信息学鉴定的生物标志物列表和已知基因组和生化通路可以对生物标志物列表施加结构。

[0588] 表2示出了与来自博德研究所(Broad Institute)的分子特征数据库类别2-Canonical通路(v.7.4.)中的所有基因组相比时,某些生物信息学鉴定的生物标志物的富集分析。此数据库包括KEGG、Biocarta和Reactome数据等其他资源。每个p值都是卡方检验(Chi-square test)的结果,比较特定的基因组与MSigDB C2-CP数据库中所有基因的背景下的某些生物信息学鉴定的生物标志物列表。将重叠最高的生物标志物排在第一位,在一些实施方案中,考虑了名义p值为0.05的重叠。

[0589] 表2示出了生物信息学鉴定的生物标志物列表中富集的某些分子通路,在多重测试校正后,多个分子通路表现出小于0.05的错误发现率(FDR)。此类分子通路为某些生物信息学鉴定的生物标志物提供了生物学主题。

[0590] 表2-某些生物信息学鉴定的生物标志物的富集分析

[0591]

基因本体通路、来源和说明	原始P值	FDR	通路基因#	包括的示例性生物标志物
REACTOME_APOPTOTIC_CLEAVAGE_OF_CELL_ADHESION_PROTEINS	0.0E+00	0.0E+00	11	CDH1、DSG2、DSG3
REACTOME_MISCELLANEOUS_SUBSTRATES	0.0E+00	0.0E+00	12	CYP2S1、CYP4F11、CYP4F3
REACTOME_CELL_JUNCTION_ORGANIZATION	0.0E+00	0.0E+00	92	CDH1、CDH3、LAMB3、LAMC2、NECTIN1、SDK1、SDK2
REACTOME_SYNTHESIS_OF_LEUKOTRIENES_LT_AND_EOXINS_EX	0.0E+00	0.0E+00	21	ABCC1、CYP4F11、CYP4F3
REACTOME_CELL_CELL_JUNCTION_ORGANIZATION	0.0E+00	0.0E+00	65	CDH1、CDH3、NECTIN1、SDK1、SDK2
REACTOME_CELL_CELL_COMMUNICATION	0.0E+00	0.0E+00	130	CDH1、CDH3、LAMB3、LAMC2、NECTIN1、SDK1、SDK2
PID_A6B1_A6B4_INTEGRIN_PATHWAY	0.0E+00	4.0E-07	46	CD9、CDH1、LAMB3、LAMC2
KEGG_ECM_RECEPTOR_INTERACTION	0.0E+00	1.1E-05	84	ITGA2、ITGB6、LAMB3、LAMC2、SDC1
PID_INTEGRIN4_PATHWAY	0.0E+00	1.9E-05	11	LAMB3、LAMC2
REACTOME_TYPE_I_HEMIDESMOsome_ASSEMBLY	0.0E+00	1.9E-05	11	LAMB3、LAMC2
REACTOME_LAMININ_INTERACTIONS	0.0E+00	3.5E-05	30	ITGA2、LAMB3、LAMC2
REACTOME_MET_ACTIVATES_PTK2_SIGNALING	0.0E+00	3.5E-05	30	ITGA2、LAMB3、LAMC2
REACTOME_EICOSANOIDS	0.0E+00	9.0E-05	12	CYP4F11、CYP4F3
REACTOME_NON	0.0E+00	9.2E-05	59	ITGA2、

[0592]

基因本体通路、来源和说明	原始P值	FDR	通路基因#	包括的示例性生物标志物
_INTEGRIN_MEMBRANE_ECM_INTERACTIONS				LAMB3、LAMC2、SDC1
REACTOME_ADHERENS_JUNCTIONS_INTERACTIONS	1.0E-07	2.0E-04	33	CDH1、CDH3、NECTIN1
REACTOME_APOPTOTIC_CLEAVAGE_OF_CELLULAR_PROTEINS	7.0E-07	1.9E-03	38	CDH1、DSG2、DSG3
REACTOME_FATTY_ACIDS	1.0E-06	2.9E-03	15	CYP4F11、CYP4F3
REACTOME_ANCHORING_FIBRIL_FORMATION	1.0E-06	2.9E-03	15	LAMB3、LAMC2
REACTOME_MET_PROMOTES_CELL_MOTILITY	2.0E-06	5.9E-03	41	ITGA2、LAMB3、LAMC2
PID_INTEGRIN5_PATHWAY	5.1E-06	1.5E-02	17	ITGB6、SDC1
KEGG_CELL_ADHESION_MOLECULES_CAMS	1.2E-05	3.5E-02	133	CDH1、CDH3、NECTIN1、NRCAM、SDC1
REACTOME_APOPTOTIC_EXECUTION_PHASE	4.1E-05	1.2E-01	52	CDH1、DSG2、DSG3
REACTOME_ARACHIDONIC_ACID_METABOLISM	1.5E-04	4.4E-01	59	ABCC1、CYP4F11、CYP4F3
REACTOME_OTHER_INTERLEUKIN_SIGNALING	1.9E-04	5.5E-01	24	PTPRZ1、SDC1
REACTOME_O_LINKED_GLYCOSYLATION	2.4E-04	7.1E-01	111	B3GNT5、GALNT14、GALNT3、LARGE2
REACTOME_O_LINKED_GLYCOSYLATION_OF_MUCINS	2.5E-04	7.2E-01	62	B3GNT5、GALNT14、GALNT3
WP_OXIDATION_BY_CYTOCHROME_P450	2.9E-04	8.3E-01	63	CYP2S1、CYP4F11、CYP4F3
PID_INTEGRIN_CS_PATHWAY	3.8E-04	1.0E+00	26	ITGA2、ITGB6
PID_INTEGRIN1_P	4.4E-04	1.0E+00	66	ITGA2、

[0593]

基因本体通路、来源和说明	原始P值	FDR	通路基因#	包括的示例性生物标志物
ATHWAY				LAMB3、LAMC2
REACTOME_CYT OCHROME_P450_ ARRANGED_BY_S UBSTRATE_TYPE	4.4E-04	1.0E+00	66	CYP2S1、 CYP4F11、 CYP4F3
REACTOME_SYN DECAN_INTERAC TIONS	5.1E-04	1.0E+00	27	ITGA2、SDC1
WP_METAPATHW AY_BIOTRANSFO RMATION_PHASE I_AND_II	6.1E-04	1.0E+00	185	CYP2S1、 CYP4F11、 CYP4F3、 HS6ST2、 UGT1A6
WP_HIPPOMERLI N_SIGNALING_D YSREGULATION	6.6E-04	1.0E+00	123	CDH1、CDH3、 ITGA2、ITGB6
KEGG_O_GLYCA N_BIOSYNTHESIS	1.1E-03	1.0E+00	30	GALNT14、 GALNT3
PID_NECTIN_PAT HWAY	1.1E-03	1.0E+00	30	CDH1、NECTIN1
KEGG_ARRHYTH MOGENIC_RIGHT VENTRICULAR_ CARDIOMYOPAT HY_ARVC	1.2E-03	1.0E+00	74	DSG2、ITGA2、 ITGB6
WP_ARRHYTHMO GENIC_RIGHT_VE NTRICULAR_CAR DIOMYOPATHY	1.4E-03	1.0E+00	76	DSG2、ITGA2、 ITGB6
REACTOME_HS_G AG_BIOSYNTHESI S	1.4E-03	1.0E+00	31	HS6ST2、SDC1
REACTOME_SIGN ALING_BY_MET	1.9E-03	1.0E+00	79	ITGA2、 LAMB3、LAMC2
REACTOME_EXT RACELLULAR_M ATRIX_ORGANIZ ATION	3.0E-03	1.0E+00	301	CDH1、ITGA2、 ITGB6、 LAMB3、 LAMC2、SDC1
KEGG_SMALL_CE LL_LUNG_CANCE R	3.0E-03	1.0E+00	84	ITGA2、 LAMB3、LAMC2
WP_ALPHA_6_BE TA_4_SIGNALING PATHWAY	3.2E-03	1.0E+00	35	LAMB3、LAMC2
REACTOME_INTE GRIN_CELL_SURF ACE_INTERACTIO	3.2E-03	1.0E+00	85	CDH1、ITGA2、 ITGB6

[0594]

基因本体通路、来源和说明	原始P值	FDR	通路基因#	包括的示例性生物标志物
NS				
KEGG_HEMATOP OIETIC_CELL_LIN EAGE	3.8E-03	1.0E+00	87	CD9、ITGA2、 TFRC
REACTOME_RAC 2_GTPASE_CYCL E	4.1E-03	1.0E+00	88	DSG2、 RACGAP1、TFRC
REACTOME_ELEC TRIC_TRANSMISS ION_ACROSS_GA P_JUNCTIONS	4.7E-03	1.0E+00	5	PANX2
REACTOME_HYA LURONAN_BIOSY NTHESIS_AND_E XPORT	4.7E-03	1.0E+00	5	HAS3
REACTOME_RNA POLYMERASE_II TRANSCRIPTION	5.9E-03	1.0E+00	1374	
REACTOME_POST TRANSLATIONA L_MODIFICATION SYNTHESIS_OF GPI_ANCHORED PROTEINS	6.3E-03	1.0E+00	94	PIGT、PRSS21、 ULBP2
REACTOME_RAC 3_GTPASE_CYCL E	6.3E-03	1.0E+00	94	DSG2、 RACGAP1、TFRC
NABA_BASEMEN T_MEMBRANES	6.9E-03	1.0E+00	40	LAMB3、LAMC2
WP_FERROPTOSI S	6.9E-03	1.0E+00	40	SLC7A11、TFRC
WP_SMALL_CELL LUNG_CANCER	8.2E-03	1.0E+00	98	ITGA2、 LAMB3、LAMC2
BIOCARTA_MRP_ PATHWAY	1.1E-02	1.0E+00	6	ABCC1
REACTOME_ACR OSOME_REACTIO N_AND_SPERM_O OCYTE_MEMBRA NE_BINDING	1.1E-02	1.0E+00	6	CD9
REACTOME_UPT AKE_AND_FUNCT ION_OF_DIPHTHE RIA_TOXIN	1.1E-02	1.0E+00	6	CD9
REACTOME_SENS ING_OF_DNA_DO UBLE_STRAND_B REAKS	1.1E-02	1.0E+00	6	KPNA2
REACTOME_VLD	1.1E-02	1.0E+00	6	LSR

[0595]

基因本体通路、来源和说明	原始P值	FDR	通路基因#	包括的示例性生物标志物
L_CLEARANCE				
KEGG_ABC_TRANSPORTERS	1.1E-02	1.0E+00	44	ABCA13、ABCC1
REACTOME_PHASE_I_FUNCTIONALIZATION_OF_COMPOUNDS	1.3E-02	1.0E+00	106	CYP2S1、CYP4F11、CYP4F3
WP_MECHANOREGULATION_AND_PATHOLOGY_OF_YAPTAZ_VIA_HIPPO_AND_NONHIPPO_MECHANISMS	1.7E-02	1.0E+00	48	CDH1、ITGB6
PID_ARF6_TRAFFICKING_PATHWAY	1.9E-02	1.0E+00	49	CDH1、ITGA2
BIOCARTA_EEA1_PATHWAY	2.0E-02	1.0E+00	7	TFRC
REACTOME_ATTACHMENT_OF_GPI_ANCHOR_TO_UPAR	2.0E-02	1.0E+00	7	PIGT
REACTOME_NECTIN_NECL_TRANS_HETERODIMERIZATION	2.0E-02	1.0E+00	7	NECTIN1
REACTOME_NRCAM_INTERACTIONS	2.0E-02	1.0E+00	7	NRCAM
REACTOME_NEUROFASCIN_INTERACTIONS	2.0E-02	1.0E+00	7	NRCAM
KEGG_FOCAL_ADHESION	2.3E-02	1.0E+00	199	ITGA2、ITGB6、LAMB3、LAMC2
REACTOME_TRANSPORT_OF_SMALL_MOLECULES	2.4E-02	1.0E+00	728	ABCC1、CLCA2、FXYP3、GNG4、LSR、PARL、SLC15A2、SLC7A11、TFRC
WP_FOCAL_ADHESION	2.5E-02	1.0E+00	202	ITGA2、ITGB6、LAMB3、LAMC2
PID_RAC1_PATHWAY	2.9E-02	1.0E+00	54	CDH1、RACGAP1
REACTOME_GLYCOSAMINOGLYCAN_METABOLISM	2.9E-02	1.0E+00	124	HAS3、HS6ST2、SDC1

[0596]

基因本体通路、来源和说明	原始P值	FDR	通路基因#	包括的示例性生物标志物
M				
BIOCARTA_CTBP1_PATHWAY	3.1E-02	1.0E+00	8	CDH1
REACTOME_CREB1_PHOSPHORYLATION_THROUGH_THE_ACTIVATION_OF_CAMKII_CAMKK_CAMKIV_CASCASDE	3.1E-02	1.0E+00	8	KPNA2
REACTOME_PHOSPHATE_BOND_HYDROLYSIS_BY_NTPDASE_PROTEINS	3.1E-02	1.0E+00	8	ENTPD3
REACTOME_HEPARAN_SULFATE_HEPARIN_HS_GAG_METABOLISM	3.1E-02	1.0E+00	55	HS6ST2、SDC1
WP_FOCAL_ADHESIONPI3KAKTMTORSIGNALING_PATHWAY	3.3E-02	1.0E+00	309	GNG4、ITGA2、ITGB6、LAMB3、LAMC2
REACTOME_BIOLOGICAL_OXIDATIONS	4.1E-02	1.0E+00	222	CYP2S1、CYP4F11、CYP4F3、UGT1A6
REACTOME_CHL1_INTERACTIONS	4.5E-02	1.0E+00	9	ITGA2
REACTOME_INLA_MEDIATED_ENTRY_OF_LISTERIA_MONOCYTOGENES INTO_HOST_CELLS	4.5E-02	1.0E+00	9	CDH1
REACTOME_ASSEMBLY_OF_COLLAGEN_FIBRILS_AND_OTHER_MULTIMERIC_STRUCTURES	4.7E-02	1.0E+00	61	LAMB3、LAMC2

[0597] 实施例9:生物信息学鉴定的生物标志物和生物标志物组合与临床协变量和已知突变驱动因素的相关性。

[0598] 本实施例说明了已知的肺癌临床协变量和某些生物信息学预测的生物标志物之间的潜在关联;以及已知的肺癌突变驱动因素与某些生物信息学预测的生物标志物之间的潜在关联。

[0599] 在一些实施方案中,热图可以用于此类分析。例如,生成了一个热图,生成了使用癌症基因组图谱(TCGA) LUSC样品示出某些鉴定的生物标志物基因的行尺度基因表达的热

图,其中每一行代表一个生物标志物候选物基因,并且每一列代表LUSC样品中的基因表达。使用皮尔逊相关性度量对行进行聚类,使用欧氏距离 (Euclidean distance) 对列进行聚类。两个维度均采用Ward’s算法进行聚类,并且采用最优叶序算法。

[0600] 在一些实施方案中,除了某些生物信息鉴定的生物标志物的基因表达外,还考虑了一个或多个临床协变量。此类分析可以用于为潜在的亚组提供指示,包括分期、淋巴结转移程度、微卫星不稳定性等。

[0601] 在一些实施方案中,临床协变量包括淋巴结转移程度(例如,n0、n1、n2、n3或n4)、癌症分期和/或吸烟量(例如,包年)。在一些实施方案中,癌症分期包括I期、II期、III期或IV期癌症。在一些实施方案中,吸烟量以包年(例如,如本文所述,例如,其中1包年等于每天1包香烟吸烟一年)来量化。

[0602] 此临床协变量分析在TCGA样品中没有鉴定出任何强富集,证明某些生物信息学鉴定的生物标志物组合可以特别有用地鉴定肺癌样品(例如,LUSC样品),而不考虑特定的临床协变量(数据未显示)。

[0603] 在一些实施方案中,除了某些生物信息学鉴定的生物标志物的基因表达外,还考虑了一个或多个突变驱动因素(包括例如,改变分析的突变和拷贝数)。例如,某些已知的主要肺癌突变驱动因子包括但不限于TP53、RB1、PTEN、PIK3CA、NOTCH1、NFE2L2、KMT2D、KEAP1或其组合中的突变。对于这些驱动因素中的每一个,癌症相关突变可以包括拷贝数改变(CNA;包括例如,但不限于扩增和/或删除)和/或突变(包括例如,但不限于非移码突变、错义突变、剪接和/或截短突变等)。进行聚类分析以鉴定生物信息学预测的生物标志物、生物标志物组合和肺癌的某些主要突变驱动因素之间的关联。

[0604] 这种突变驱动因素分析在TCGA样品中没有鉴定出任何强富集,证明某些生物信息学预测的生物标志物和/或生物标志物组合可以特别有用地鉴定肺癌样品(例如,LUSC样品),而不考虑特定的突变驱动因素(数据未显示)。

[0605] 实施例10:开发示例性肺癌液体活检测定

[0606] 本实施例描述了示例性肺癌液体活检测定的开发,例如,用于筛选具有平均风险、遗传风险和/或生活史相关风险的有或无症状的受试者。如本公开的实施例7至9中所述进行导致生物标志物发现的步骤。样品制备和测定如实施例1和实施例2进行。

[0607] 使用生物信息学途鉴定径肺癌生物标志物;使用市售靶向细胞系EV的单克隆抗体对某些生物标志物(呈现于表3)进行测试,以鉴定在示例性实验(数据未显示)中表现出高亲和力和高选择性的抗体克隆。对于表3中列出的每个生物标志物,坚定了至少一个合适的抗体克隆。

[0608] 表3-本文所述示例性双链体测定中测试的肺癌生物标志物

[0609]

#	生物标志物	#	生物标志物
1	ALCAM	12	IG1FR
2	CD55	13	MET
3	CDH1	14	MSLN
4	CDH3	15	MUC1
5	CD274 (PD-L1)	16	SLC34A2
6	CEACAM5	17	sTn抗原

7	CEACAM6	18	Tn抗原
8	DSG2	19	T抗原
9	EGFR	20	TACSTD2
10	EPCAM	21	TNFRSF10B
11	FOLR1		

[0610] 下表4示出了某些评估的双链体生物标志物组合。每个双链体生物标志物组合包含上表3中列出的两个不同的生物标志物。

[0611] 表4-用于示例性双链体测定的某些生物标志物组合

[0612]

测试	捕获探针靶标	检测探针靶标	测试	捕获探针靶标	检测探针靶标
1	ALCAM	CD274	104	FOLR1	MSLN
2	ALCAM	CD55	105	FOLR1	MUC1
3	ALCAM	CDH1	106	FOLR1	SLC34A2
4	ALCAM	CDH3	107	FOLR1	sTn抗原
5	ALCAM	CEACAM5	108	FOLR1	T抗原
6	ALCAM	CEACAM6	109	FOLR1	TACSTD2
7	ALCAM	EGFR	110	FOLR1	Tn抗原
8	ALCAM	EPCAM	111	IGF1R	CDH1
9	ALCAM	FOLR1	112	IGF1R	T抗原
10	ALCAM	IGF1R	113	MET	CDH1
11	ALCAM	MET	114	MET	CDH3
12	ALCAM	MSLN	115	MET	IGF1R
13	ALCAM	MUC1	116	MET	MSLN
14	ALCAM	SLC34A2	117	MET	MUC1
15	ALCAM	sTn抗原	118	MET	sTn抗原
16	ALCAM	T抗原	119	MET	T抗原
17	ALCAM	TACSTD2	120	MET	TACSTD2
18	ALCAM	Tn抗原	121	MET	Tn抗原
19	CDH3	CDH1	122	MSLN	CDH1
20	CDH3	IGF1R	123	MSLN	IGF1R
21	CDH3	MSLN	124	MSLN	sTn抗原
22	CDH3	sTn抗原	125	MSLN	T抗原
23	CDH3	T抗原	126	MUC1	CDH1
24	CDH3	TACSTD2	127	MUC1	CDH3
25	CEACAM5	CDH1	128	MUC1	IGF1R
26	CEACAM5	CDH3	129	MUC1	MSLN
27	CEACAM5	IGF1R	130	MUC1	sTn抗原
28	CEACAM5	MET	131	MUC1	T抗原
29	CEACAM5	MSLN	132	MUC1	TACSTD2
30	CEACAM5	MUC1	133	MUC1	Tn抗原
31	CEACAM5	sTn抗原	134	SLC34A2	CD274
32	CEACAM5	T抗原	135	SLC34A2	CDH1
33	CEACAM5	TACSTD2	136	SLC34A2	CDH3
34	CEACAM5	Tn抗原	137	SLC34A2	CEACAM5
35	CEACAM6	CD274	138	SLC34A2	EGFR
36	CEACAM6	CDH1	139	SLC34A2	IGF1R
37	CEACAM6	CDH3	140	SLC34A2	MET
38	CEACAM6	CEACAM5	141	SLC34A2	MSLN
39	CEACAM6	EGFR	142	SLC34A2	MUC1
40	CEACAM6	EPCAM	143	SLC34A2	sTn抗原
41	CEACAM6	FOLR1	144	SLC34A2	T抗原

[0613]

测试	捕获探针靶标	检测探针靶标	测试	捕获探针靶标	检测探针靶标
42	CEACAM6	IGF1R	145	SLC34A2	TACSTD2
43	CEACAM6	MET	146	SLC34A2	Tn抗原
44	CEACAM6	MSLN	147	sTn抗原	CDH1
45	CEACAM6	MUC1	148	sTn抗原	IGF1R
46	CEACAM6	SLC34A2	149	T抗原	CDH1
47	CEACAM6	sTn抗原	150	TACSTD2	CDH1
48	CEACAM6	T抗原	151	TACSTD2	IGF1R
49	CEACAM6	TACSTD2	152	TACSTD2	MSLN
50	CEACAM6	Tn抗原	153	TACSTD2	sTn抗原
51	DSG2	ALCAM	154	TACSTD2	T抗原
52	DSG2	CD274	155	Tn抗原	CDH1
53	DSG2	CD55	156	Tn抗原	CDH3
54	DSG2	CDH1	157	Tn抗原	IGF1R
55	DSG2	CDH3	158	Tn抗原	MSLN
56	DSG2	CEACAM5	159	Tn抗原	TACSTD2
57	DSG2	CEACAM6	160	TNFRSF10B	ALCAM
58	DSG2	EGFR	161	TNFRSF10B	CD274
59	DSG2	EPCAM	162	TNFRSF10B	CD55
60	DSG2	FOLR1	163	TNFRSF10B	CDH1
61	DSG2	IGF1R	164	TNFRSF10B	CDH3
62	DSG2	MET	165	TNFRSF10B	CEACAM5
63	DSG2	MSLN	166	TNFRSF10B	CEACAM6
64	DSG2	MUC1	167	TNFRSF10B	EGFR
65	DSG2	SLC34A2	168	TNFRSF10B	EPCAM
66	DSG2	sTn抗原	169	TNFRSF10B	FOLR1
67	DSG2	T抗原	170	TNFRSF10B	IGF1R
68	DSG2	TACSTD2	171	TNFRSF10B	MET
69	DSG2	Tn抗原	172	TNFRSF10B	MSLN
70	EGFR	CD274	173	TNFRSF10B	MUC1
71	EGFR	CDH1	174	TNFRSF10B	SLC34A2
72	EGFR	CDH3	175	TNFRSF10B	sTn抗原
73	EGFR	CEACAM5	176	TNFRSF10B	T抗原
74	EGFR	IGF1R	177	TNFRSF10B	TACSTD2
75	EGFR	MET	178	TNFRSF10B	Tn抗原
76	EGFR	MSLN	179	CD274	CDH1
77	EGFR	MUC1	180	CD274	CDH3
78	EGFR	sTn抗原	181	CD274	CEACAM5
79	EGFR	T抗原	182	CD274	IGF1R
80	EGFR	TACSTD2	183	CD274	MET
81	EGFR	Tn抗原	184	CD274	MSLN
82	EPCAM	CD274	185	CD274	MUC1
83	EPCAM	CDH1	186	CD274	sTn抗原

[0614]	测试	捕获探针靶标	检测探针靶标	测试	捕获探针靶标	检测探针靶标
	84	EPCAM	CDH3	187	CD274	T抗原
	85	EPCAM	CEACAM5	188	CD274	TACSTD2
	86	EPCAM	EGFR	189	CD274	Tn抗原
	87	EPCAM	FOLR1	190	CD55	CD274
	88	EPCAM	IGF1R	191	CD55	CDH1
	89	EPCAM	MET	192	CD55	CDH3
	90	EPCAM	MSLN	193	CD55	CEACAM5
	91	EPCAM	MUC1	194	CD55	CEACAM6
	92	EPCAM	SLC34A2	195	CD55	EGFR
	93	EPCAM	sTn抗原	196	CD55	EPCAM
	94	EPCAM	T抗原	197	CD55	FOLR1
	95	EPCAM	TACSTD2	198	CD55	IGF1R
	96	EPCAM	Tn抗原	199	CD55	MET
	97	FOLR1	CD274	200	CD55	MSLN
	98	FOLR1	CDH1	201	CD55	MUC1
	99	FOLR1	CDH3	202	CD55	SLC34A2
	100	FOLR1	CEACAM5	203	CD55	sTn抗原
	101	FOLR1	EGFR	204	CD55	T抗原
	102	FOLR1	IGF1R	205	CD55	TACSTD2
	103	FOLR1	MET	206	CD55	Tn抗原

[0615] 下图16A至20B和表5总结了使用上表4所列的某些双链体生物标志物组合检测肺癌(例如,不同类型和/或不同分期的肺癌)的结果。如实施例1和实施例2进行测定。表5所示结果比较健康非吸烟者混合样品、健康吸烟者混合样品、早期肺腺癌混合样品、晚期肺腺癌混合样品、早期肺鳞状细胞癌混合样品和晚期肺鳞状细胞癌混合样品。对每个生物标志物组合进行了合适的阳性和阴性对照,以验证测定是否适当的生效(例如,包含来自肺癌细胞系的细胞外囊泡的阳性对照样品和不包含细胞外囊泡的阴性对照样品)。结果证明,某些生物标志物组合成功地区分肺癌样品与所指出的参考样品(例如,当用描述的生物标志物组合以如实施例1中所述进行的扩增的至少2倍差异测量时)。

[0616] 健康非吸烟者池包括来自没有肺癌病史的受试者(例如,女性和男性受试者)的血清样品。在一些情况下,此类受试者的医疗状况可以不是肺癌或与肺部病症无关,包括例如,慢性胃炎、冠状动脉疾病、食管溃疡、高血压、骨关节炎、慢性胆囊炎、静脉曲张、肥胖、慢性前列腺炎和/或高血压。所有样品均来自白种人,其年龄为55至78岁,并且体重指数(BMI)在22至31的范围内。

[0617] 健康吸烟者池包括来自没有肺癌病史但有相当于20至30包年的吸烟史的受试者(例如,女性和男性受试者)的血清样品。在一些情况下,此类受试者的医疗状况可以不是肺癌或与肺部病症无关,包括例如,慢性支气管炎、高血压和/或胃十二指肠炎。所有样品均来自白种人,其年龄为50至67岁,并且BMI在23至29的范围内。

[0618] 早期LUAD池包括来自已被确定患有早期LUAD的受试者(例如,女性和男性受试者)的血清样品。所有样品均来自白种人,其年龄为36至77岁,有非吸烟者至45包年范围内的不同吸烟史,并且BMI在19至41的范围内。在一些情况下,此类受试者至少有以下一种或多种医疗病状(除早期LUAD外):高血压、肥胖、哮喘、缺血性心脏病、肾癌、动脉粥样硬化、2型糖

尿病、慢性支气管炎、子宫癌、乳腺癌、慢性肾盂肾炎、宫颈癌、胃溃疡、肾癌、慢性胃炎和/或骨关节炎。

[0619] 晚期LUAD池包括来自已被确定患有请晚期LUAD的受试者(例如,女性和男性受试者)的血清样品。所有样品均来自白种人,其年龄为51至72岁,有非吸烟者至54包年范围内的不同吸烟史,并且BMI在19至40的范围内。在一些情况下,此类受试者至少有以下一种或多种医疗病状(除晚期LUAD外):脑血管病、缺血性心脏病、动脉粥样硬化、高血压、冠状动脉疾病、肥胖、结节性甲状腺肿、慢性胆囊炎、2型糖尿病、中风、慢性支气管炎、慢性胃炎、心绞痛、甲状腺毒症和/或慢性阻塞性支气管炎。

[0620] 早期LUSC池包括来自已被确定患有早期LUSC的受试者(例如,女性和男性受试者)的血清样品。所有样品均来自白种人,其年龄为49至82岁,有非吸烟者至50包年范围内的不同吸烟史,并且BMI在19至33的范围内。在一些情况下,此类受试者至少有以下一种或多种医疗病状(除早期LUSC外):中风、慢性支气管炎、2型糖尿病、动脉粥样硬化、慢性萎缩性胃炎、冠状动脉疾病、高血压、慢性胃炎、慢性静脉功能不全、慢性阻塞性肺疾病、脂肪性肝炎、胃癌、阵发性房颤、肾癌、病态窦房结综合征、心律不齐、感觉性听力损失、振动病、结肠炎、糖耐量异常、消化性溃疡、十二指肠溃疡和/或肥胖。

[0621] 晚期LUSC池包括来自已被确定患有请晚期LUSC的受试者(例如,女性和男性受试者)的血清样品。所有样品均来自白种人,其年龄为49至66岁,有20至52包年范围内的吸烟史,并且BMI在23至31的范围内。在一些情况下,此类受试者至少有以下一种或多种医疗病状(除了晚期LUSC):慢性支气管炎、肺硬化、肺功能衰竭、慢性胃炎、慢性酒精中毒、胃溃疡、冠状动脉疾病、动脉粥样硬化、肺气肿、十二指肠溃疡、慢性阻塞性肺疾病、静脉曲张疾病、疼痛综合征、高血压和/或动脉粥样硬化心硬化。

[0622] 本实施例表明,本文所述测定能够使用本文提供的某些生物标志物来区分肺癌来源的EV和来源于健康组织的EV。因此,在一些实施方案中,本文提供的某些生物标志物特征(例如,如表4所列)对于检测肺癌,特别是非小细胞肺癌(例如,肺腺癌和/或肺鳞状细胞癌)特别有用。在某些实施方案中,表4中列出的生物标志物组合结合如本文所述的EV测定能够检测出肺癌,例如,在一些实施方案中,信号至少大于来自健康个体(例如,非吸烟者和/或吸烟者)血浆样品池的两倍。

[0623] 如下所示,在一些实施方案中,某些生物标志物组合区分晚期肺癌池和健康池(不吸烟者和/或吸烟者)。此类发现尤其证明提供的根据本公开使用的生物标志物组合区分肺恶性肿瘤与非癌症病状,从而证明了所公开技术的特别有利的特征。在一些情况下,此类特征特别有益,因为肺肿瘤可以生成与和吸烟有关的其他疾病(诸如慢性支气管炎和/或慢性阻塞性肺疾病)类似的症状,并且使用常规的筛查方法往往难以区分。

[0624] 如下所示,在一些实施方案中,某些生物标志物组合区分早期肺癌池和健康池(不吸烟者和吸烟者)。此类发现尤其证明提供的根据本公开使用的生物标志物组合区分肺恶性肿瘤与无肺癌的个体,从而证明了所公开技术的特别有利的特征。在一些情况下,此类特征特别有益,因为在早期正确区分肺癌样品和健康个体(吸烟者和/或不吸烟者),可以增加成功治疗疾病的机会,同时保持患者健康状态。

[0625] 尤其在本实施例中展示的结果证明提供的技术检测出肺癌(例如,非小细胞肺癌,例如像肺腺癌和/或肺鳞状细胞癌),并特别证明提供的技术(例如,使用如本文所述的生物

标志物组合来检测EV中和/或EV上的生物标志物)区分肺癌来源的EV和来源于非肺癌组织的EV。

[0626] 在一些实施方案中,本文提供的用于检测肺癌的目标生物标志物特征可以包含至少三个靶生物标志物,其中至少一个是如本文所述的表面蛋白生物标志物。在一些实施方案中,至少两个或至少三个此类靶生物标志物是本文描述的不同的表面蛋白生物标志物。在一些实施方案中,可以用于检测肺癌(例如,非小细胞肺癌)的靶生物标志物特征包含TNFRSF10B、MUC1和CEACAM6。在一些实施方案中,可以用于检测肺癌(例如,非小细胞肺癌)的靶生物标志物特征包含DSG2、MUC1和CEACAM6。在一些实施方案中,可以用于检测肺癌(例如,非小细胞肺癌)的靶生物标志物特征包含CEACAM6、MUC1和sTn抗原。在一些实施方案中,可以用于检测肺癌(例如,非小细胞肺癌)的靶生物标志物特征包含SLC34A2、MUC1和CEACAM6。在一些实施方案中,可以用于检测肺癌(例如,非小细胞肺癌)的靶生物标志物特征包含CEACAM6、MUC1和MSLN。在一些实施方案中,可以用于检测肺癌(例如,非小细胞肺癌)的靶生物标志物特征包含FOLR1、T抗原和EGFR。

[0627] 在一些实施方案中,本文所述的靶生物标志物特征可以通过本文描述的测定(例如,本文所述的基于邻近性的连接测定)检测。在一些实施方案中,基于邻近性的连接测定可以包括针对TNFRSF10B的捕获探针、针对MUC1的检测探针和针对CEACAM6的检测探针。在某些实施方案中,本文所述的此类基于邻近性的连接测定能够区分早期肺癌(例如,早期非小细胞肺癌,例如像LUAD)与健康非吸烟者和健康吸烟者(参加例如,图21A至24B;和表6)。在一些实施方案中,基于邻近性的连接测定可以包括针对DSG2的捕获探针、针对MUC1的检测探针和针对CEACAM6的检测探针。在某些实施方案中,本文所述的此类基于邻近性的连接测定能够区分早期肺癌(例如,早期非小细胞肺癌,例如像LUAD)与健康非吸烟者和健康吸烟者(参加例如,图21A至24B;和表6)。在一些实施方案中,基于邻近性的连接测定可以包括针对CEACAM6的捕获探针、针对MUC1的检测探针和针对sTn抗原的检测探针。在某些实施方案中,本文所述的此类基于邻近性的连接测定能够区分早期肺癌(例如,早期非小细胞肺癌,例如像LUAD)与健康非吸烟者和健康吸烟者(参加例如,图21A至24B;和表6)。在一些实施方案中,基于邻近性的连接测定可以包括针对SLC34A2的捕获探针、针对MUC1的检测探针和针对CEACAM6的检测探针。在某些实施方案中,本文所述的此类基于邻近性的连接测定能够区分早期肺癌(例如,早期非小细胞肺癌,例如像LUAD)与健康非吸烟者和健康吸烟者(参加例如,图21A至24B;和表6)。在一些实施方案中,基于邻近性的连接测定可以包括针对CEACAM6的捕获探针、针对MUC1的检测探针和针对MSLN的检测探针。在某些实施方案中,本文所述的此类基于邻近性的连接测定能够区分晚期肺癌(例如,晚期非小细胞肺癌,例如像LUAD)与健康非吸烟者和健康吸烟者(参加例如,图21A至24B;和表6)。在一些实施方案中,基于邻近性的连接测定可以包括针对FOLR1的捕获探针、针对T抗原的检测探针和针对EGFR的检测探针。在某些实施方案中,本文所述的此类基于邻近性的连接测定能够区分晚期肺癌(例如,晚期非小细胞肺癌,例如像LUAD)与健康非吸烟者(参加例如,图21A至24B;和表6)。

[0628] 表5-使用某些示例性2-生物标志物组合检测肺癌

[0629]

生物标志物组合	区分早期(ES)或晚期(LS) LUAD样品和健康吸烟者(S)或健康非吸烟者(NS)的样品?				区分早期(ES)或晚期(LS) LUSC样品和健康吸烟者(S)或健康非吸烟者(NS)的样品?			
	ES与S	LS与S	ES与NS	LS于NS	ES与S	LS与S	ES与NS	LS于NS
TNFRSF10B + PD-L1	*	*	*	*	*	*	*	*
TNFRSF10B + CEACAM6	*	*	*	*		*	*	*
TNFRSF10B + EGFR	*	*		*	*	*	*	*
TNFRSF10B + IGF1R	*	*	*	*	*		*	*
ALCAM + EPCAM			*	*	*	*	*	*
CEACAM6 + MUC1	*	*	*	*		*		*
EGFR + T 抗原		*		*	*	*	*	*
EPCAM + T 抗原		*		*	*	*	*	*
FOLR1 + T 抗原		*		*	*	*	*	*
Tn 抗原 + TACSTD2	*	*	*	*	*	*		
TNFRSF10B + FOLR1	*		*	*		*	*	*
TNFRSF10B + sTn 抗原	*		*	*		*	*	*
ALCAM + PD-L1		*		*	*	*	*	
EPCAM + MUC1		*		*	*	*	*	
TNFRSF10B + CD55		*	*	*	*		*	
TNFRSF10B + MUC1	*	*	*	*	*			
FOLR1 + TACSTD2					*	*	*	*
MET +	*	*		*	*			

[0630]

生物标志物组合	区分早期(ES)或晚期(LS) LUAD样品和健康吸烟者(S)或健康非吸烟者(NS)的样品?				区分早期(ES)或晚期(LS) LUSC样品和健康吸烟者(S)或健康非吸烟者(NS)的样品?			
	ES与S	LS与S	ES与NS	LS于NS	ES与S	LS与S	ES与NS	LS于NS
MUC1								
MET + sTn抗原		*		*		*		*
MUC1 + TACSTD2		*		*		*		*
PD-L1 + MUC1	*	*		*	*			
PD-L1 + 抗原	*	*		*	*			
SLC34A2 + MET		*		*			*	*
SLC34A2 + T抗原				*	*		*	*
TNFRSF10B + CEACAM5			*	*			*	*
TNFRSF10B + MSLN		*		*			*	*
TNFRSF10B + Tn抗原			*	*			*	*
ALCAM + T抗原		*		*				*
ALCAM + TACSTD2		*				*		*
CD55 + PD-L1	*	*			*			
CDH3 + CDH1					*	*	*	
CEACAM5 + MUC1		*		*		*		
CEACAM6 + EGFR		*			*	*		
CEACAM6 + EPCAM		*				*		*
CEACAM6 + FOLR1		*		*		*		
CEACAM6 + MSLN		*		*				*
CEACAM6 + sTn 抗			*			*		*

[0631]

生物标志物组合	区分早期(ES)或晚期(LS) LUAD样品和健康吸烟者(S)或健康非吸烟者(NS)的样品?				区分早期(ES)或晚期(LS) LUSC样品和健康吸烟者(S)或健康非吸烟者(NS)的样品?			
	ES与S	LS与S	ES与NS	LS于NS	ES与S	LS与S	ES与NS	LS于NS
原								
CEACAM 6 + TACSTD2				*		*		*
DSG2 + CEACAM 6	*		*					*
EGFR + MUC1		*		*	*			
FOLR1 + MUC1		*		*	*			
MUC1 + sTn 抗原		*		*		*		
SLC34A2 + MSLN		*				*		*
T抗原 + CDH1				*			*	*
Tn抗原 + IGF1R	*	*			*			
TNFRSF10B + ALCAM			*	*			*	
CD55 + EPCAM	*		*					
CEACAM 6 + PD-L1		*			*			
CEACAM 6 + SLC34A2		*		*				
DSG2 + CEACAM 5			*					*
DSG2 + EPCAM			*					*
DSG2 + MET			*	*				
DSG2 + T 抗原		*		*				
EGFR + sTn抗原		*				*		
EGFR + TACSTD2				*				*
EPCAM + TACSTD2						*		*

[0632]

生物标志物组合	区分早期(ES)或晚期(LS) LUAD样品和健康吸烟者(S)或健康非吸烟者(NS)的样品?				区分早期(ES)或晚期(LS) LUSC样品和健康吸烟者(S)或健康非吸烟者(NS)的样品?			
	ES与S	LS与S	ES与NS	LS于NS	ES与S	LS与S	ES与NS	LS于NS
IGF1R + T 抗原					*		*	
MET + Tn 抗原		*		*				
MSLN + T 抗原		*		*				
PD-L1 + sTn 抗原						*		*
PD-L1 + T 抗原	*		*					
SLC34A2 + MUC1		*		*				
SLC34A2 + PD-L1		*			*			
SLC34A2 + sTn 抗原						*		*
SLC34A2 + Tn 抗原		*		*				
TACSTD2 + sTn 抗原						*		*
TNFRSF10B + CDH3			*					*
TNFRSF10B + EPCAM						*		*
TNFRSF10B + MET	*		*					
ALCAM + CDH3						*		
ALCAM + MET								*
ALCAM + MUC1		*						
ALCAM + SLC34A2							*	
ALCAM + sTn 抗原								*
CD55 + CDH1				*				
CD55 + MET			*					

[0633]

生物标志物组合	区分早期(ES)或晚期(LS) LUAD样品和健康吸烟者(S)或健康非吸烟者(NS)的样品?				区分早期(ES)或晚期(LS) LUSC样品和健康吸烟者(S)或健康非吸烟者(NS)的样品?			
	ES与S	LS与S	ES与NS	LS于NS	ES与S	LS与S	ES与NS	LS于NS
CDH3 + T抗原						*		
CEACAM 5 + MSLN				*				
CEACAM 5 + TACSTD2						*		
CEACAM 6 + Tn抗原	*							
DSG2 + CDH1			*					
DSG2 + CDH3		*						
DSG2 + MUC1			*					
DSG2 + SLC34A2			*					
DSG2 + sTn抗原			*					
EGFR + MSLN				*				
FOLR1 + sTn抗原		*						
FOLR1 + Tn抗原				*				
MET + T抗原							*	
MSLN + sTn抗原								*
MUC1 + T抗原								*
MUC1 + Tn抗原		*						
SLC34A2 + EGFR		*						
TNFRSF10B + CDH1		*						
TNFRSF10B + SLC34A2								*

[0634] 表5描述了用于检测肺癌(例如,非小细胞肺癌)的某些示例性的两个靶生物标志物组合的评估。将早期(ES)或晚期(LS)肺腺癌(LUAD)或肺鳞状细胞癌(LUSC)与健康吸烟者(S)样品池或健康非吸烟者(NS)样品池进行比较。行描述了区分肺癌样品与健康吸烟者或非吸烟者样品的生物标志物组合。检测/区分记为*,并且与标注的参考样品相比时,对应于癌症样品之间至少2倍的扩增差异(例如,至少3倍扩增差异、至少4倍扩增差异等)。值得注意的是,此类某些示例性2-生物标志物组合的2个生物标志物可以用作用于提供基于邻近

性的连接测定的捕获探针和/或检测探针的靶标。

[0635] 表6-使用某些示例性3-生物标志物组合检测肺癌

生物标志物组合	区分早期(ES)或晚期(LS) LUAD样品和健康吸烟者(S)或健康非吸烟者(NS)的样品?				区分早期(ES)或晚期(LS) LUSC样品和健康吸烟者(S)或健康非吸烟者(NS)的样品?			
	ES与S	LS与S	ES与NS	LS于NS	ES与S	LS与S	ES与NS	LS于NS
CEACAM6, MUC1 + sTn 抗原	*	*	*	*	*	*	*	*
TNFRSF10B, MUC1 + CEACAM6	*	*	*	*		*		*
SLC34A2, MUC1 + CEACAM6	*	*		*	*	*		
CEACAM6, MUC1 + MSLN		*		*		*		*
FOLR1, T抗原 + EGFR				*				
DSG2, MUC1 + CEACAM6	*	*						

[0637] 表6描述了用于检测肺癌(例如,非小细胞肺癌)的某些示例性的三个靶生物标志物组合的评估。将早期(ES)或晚期(LS)肺腺癌(LUAD)或肺鳞状细胞癌(LUSC)与健康吸烟者(S)样品池或健康非吸烟者(NS)样品池进行比较。行描述了区分肺癌样品与健康吸烟者或非吸烟者样品的生物标志物组合。检测/区分记为*,并且与标注的参考样品相比时,对应于癌症样品之间至少2倍的扩增差异(例如,至少3倍扩增差异、至少4倍扩增差异等)。对于每个生物标志物组合,列出的第一个生物标志物是捕获探针的靶标,而列出的之后的生物标志物是检测探针的靶标。值得注意的是,生物标志物可以适合作为提供的基于邻近性的连接测定的一个或多个检测探针和/或捕获探针的靶标。

[0638] 实施例11:进一步表征示例性肺癌液体活检测定

[0639] 此实施例针对使用不同细胞群体和患者群体的示例性肺癌液体活检测定中的各种生物标志物组合的进一步表征。通过筛查本文公开的在混合对照和患者血浆样品中的各种生物标志物组合(例如,表面蛋白靶标生物标志物,如实施例1至6所述),可以确定用于检测肺癌的有用生物标志物组合。混合的样品可以提供对各患者群组之间平均测定信号的估计,从而促进生物标志物组合分类。

[0640] 在一些实施方案中,用于检测肺癌的生物标志物组合可以包含一个、两个、三个、四个、五个、六个、七个或更多个生物标志物(例如,本文所述的那些),其中此类组合包含至少一个用于捕获细胞外囊泡(例如,通过免疫亲和捕获)的生物标志物和至少一个用于通过pliq-PCR测定进行检测的生物标志物(包括例如,一个、两个、三个、四个、五个、六个、七个或更多个生物标志物)。在一些实施方案中,用于捕获细胞外囊泡(例如,免疫亲和捕获)的靶生物标志物可以与用于基于pliq-PCR的分析的至少一个靶生物标志物相同。在一些实施方案中,用于捕获细胞外囊泡和基于pliq-PCR的分析的靶生物标志物可以各不相同。

[0641] 可以混合的患者群组中验证各种生物标志物组合。患者群组可以包括适当的患者和/或对照群体。在某些实施方案中,患者群组可以包括健康非吸烟受试者(例如,年龄在45-85岁之间的健康受试者)有吸烟史的健康受试者(例如,年龄在45-85岁之间的健康受试者)、患有晚期肺癌(例如,III期、IV期非小细胞肺癌,例如像肺腺癌或肺鳞状细胞癌)的受试者、患有早期肺癌(例如,I期、II期非小细胞肺癌,例如像肺腺癌或肺鳞状细胞癌)的受试者、患有良性肿瘤的受试者、患有炎症疾病、病症或病状的受试者,和/或患有其他癌症的受试者。

[0642] 在一些实施方案中,两步筛查程序可以表征各种生物标志物组合的性能。例如,最初可以在健康背景池(来自各年龄组)和晚期肺癌(例如,LUAD和/或LUSC)池中筛查各种生物标志物组合(例如,用于捕获探针(例如,用于免疫亲和捕获)和检测探针的生物标志物的各种组合)。在一些实施方案中,在健康池和肺癌池之间表现出临界分离(例如, ΔCt 小于4,对应于小于16倍的差异)的生物标志物组合可以作为用在分离中的生物标志物组合从进一步研究中排除,然而,当与如本文所述的其他生物标志物组合组合时,此类生物标志物组合可以是有用的。在一些实施方案中,在健康池和肺癌池之间表现出临界分离(例如, ΔCt 小于2,对应于小于4倍的差异)的生物标志物组合可以作为用在分离中的生物标志物组合从进一步研究中排除,然而,当与如本文所述的其他生物标志物组合组合时,此类生物标志物组合可以是有用的。在一些实施方案中,在健康池和肺癌池之间表现出临界分离(例如, ΔCt 小于1,对应于小于2倍的差异)的生物标志物组合可以作为用在分离中的生物标志物组合从进一步研究中排除,然而,当与如本文所述的其他生物标志物组合组合时,此类生物标志物组合可以是有用的。

[0643] 在一些实施方案中,两步筛查程序可以表征各种生物标志物组合的性能。例如,最初可以健康背景池(来自各年龄组)和晚期肺癌池中筛查各种生物标志物组合(例如,用于捕获探针(例如,用于免疫亲和捕获)和检测探针的生物标志物的各种组合)。在一些实施方案中,表现出良好诊断性能(例如, ΔCt 大于1,对应于大于2倍的差异)的生物标志物组合可以用来自早期肺癌、良性肿瘤、其他癌症和/或炎症病状的混合样品进行第二轮筛查。在一些实施方案中,表现出良好诊断性能(例如, ΔCt 大于2,对应于大于4倍的差异)的生物标志物组合可以用混合的早期HGSOC、良性妇科肿瘤和/或炎症病状进行第二轮筛查。在一些实施方案中,表现出良好诊断性能(例如, ΔCt 大于4,对应于大于16倍的差异)的生物标志物组合可以用混合的早期HGSOC、良性妇科肿瘤和/或炎症病状进行第二轮筛查。在一些实施方案中,可以进一步评估可以区分早期和晚期肺癌与对照池的最优生物标志物组合(例如,表现最好的大约20种生物标志物组合)。

[0644] 在示例性测定(例如,本文描述的那些)中并入一种或多种另外的生物标志物组合可以提高其灵敏度。如前所述,通过使用至少两种生物标志物组合,可以实现提高测定(例如,在99%的特异性下至少大约10%的灵敏度;或在98%的特异度下至少大约80%的灵敏度;或者在99.5%的特异性下至少大约70%的灵敏度)的性能。混合的患者样品可以近似整个单独患者群体的测定信号,这一特征可以提供真实的矩阵来以有效的方式评估大量的生物标志物组合。鉴定的最优生物标志物组合(例如,多达20种生物标志物组合)可以在基于个体患者血浆的试验性研究中进行进一步测试。在一些实施方案中,基于患者个体血浆的试验性研究可以包括是健康吸烟者或非吸烟者的对照患者、具有诱发类似肺癌的症状的条

件的对照患者、具有其他类型癌症的对照患者和/或患有良性肿瘤的对照患者。在一些实施方案中,基于个体患者血浆的试验性研究可以包括是健康吸烟者或非吸烟者的测试患者、有症状的测试患者、无症状的测试患者、患有I期或II期肺癌的测试患者和/或患有III期或IV期HGSOC的测试患者。非肺癌健康病状可以与肺癌群组进行年龄匹配。可以期望的是,对照样品(例如,健康对照、良性肿瘤和/或其他非靶病状等)可以提供对数正态分布的估计以设置与98%特异性(大约95%)有关的用于遗传风险和/或生活方式相关风险筛查测定的信号截止值,和与99.5%特异性(大约)有关的用于症状分类测定的信号截止值。可以使用组合之间的双变量关联来评估生物标志物组合性能特征以评估独立性,并且可以使用用于预测肺癌的三变量逻辑回归模型进一步测试顶级生物标志物组合。

[0645] 可以实现对可以区分肺癌病例与对照并鉴定组合之间的正交性的新的生物标志物组合的鉴定。此正交性可以帮助区分肺癌病例与对照群组,增加如本文所述的示例性测定的灵敏度。为了评估患者群组中的种族多样性,可以从来源于不同群体的供应商处获得样品。在某些实施方案中,示例性生物标志物组合可对特定种族和/或族裔群组具有特异性。在某些实施方案中,在不同种族和/或族裔群组中测试示例性生物标志物组合可以鉴定对特定种族和/或族裔群组具有适当诊断价值的生物标志物组合。

[0646] 在主要患者样品中验证其他肺癌生物标志物组合可以提高示例性测定的诊断性能。在一些实施方案中,通过纳入额外的正交生物标志物组合,可以在99%特异性或更高的情况下提高超过50%的灵敏度或更高的测定性能。在一些实施方案中,通过纳入额外的正交生物标志物组合,可以在98%特异性的情况下提高超过70%的灵敏度或更高的测定性能和在99.5%特异性的情况下提高超过60%的灵敏度或更高的测定性能。

[0647] 实施例12:验证示例性肺癌液体活检测定

[0648] 此实施例涉及在其他群体/群组中验证示例性肺癌液体活检测定。可以进行独立验证研究,以使用其他患者样品群组评估如本文所述的示例性肺癌诊断测定。样品可以从任何适当的来源获得。技术专家在进行任何结果分析之前不知道样品名称,和/或测定结果由外部独立技术专家进行分析。为确保从代表美国的群体中取样,根据样品的可用性,至少大约20%的样品可来源于非白人种族(美国人口普查局(United States Census Bureau), 2018)。

[0649] 在某些实施方案中,分析的患者群组可以包括:在接受任何风险降低手术之前具有肺癌遗传和/或生活方式相关风险的患者,和适当的相对对照群组。在某些情况下,生物标志物组合(例如,本文所述的那些)在具有遗传和/或生活方式相关风险的受试者中可以提供至少在大约98%的特异性下大约80%的灵敏度。在某些情况下,生物标志物组合(例如,本文所述的那些)在具有遗传和/或生活方式相关风险的受试者中可以提供至少在大约99.5%的特异性下大约70%的灵敏度。此外,在某些实施方案中,示例性测定可以区分患有良性肿瘤的受试者与患有肺癌的受试者,从而导致几乎没有假阳性。

[0650] 在某些实施方案中,如本文所述的示例性测定的其他验证研究可以利用来自独立来源的纵向收集的样品进行。在一些实施方案中,样品可以获得自或来源于肺癌病例的纵向收集的抽血(例如,在不同时间点收集的抽血)。此外,可以分析从年龄匹配的对照中抽取的血液。使用用于99%特异性截止值的示例性逻辑回归模型,可以评估在保持每年的特异性为99%时,示例性测定在诊断前多少年(例如,无论肺癌诊断之前多少年抽血可用)能够

检测到肺癌。测定灵敏度可以在诊断前计算为年份的函数,同时确保对照样品中的特异性保持在99%。

[0651] 引用文献

[0652] Balaj,L.,Lessard,R.,Dai,L.,Cho,Y.J.,Pomeroy,S.L.,Breakefield,X.O.and Skog,J.,2011.Tumour microvesicles contain retrotransposon elements and amplified oncogene sequences.Nature communications,2,p.180.

[0653] Bebelman,M.P.,Smit,M.J.,Pegtel,D.M.,and Baglio,S.R.,2018.Biogenesis and function of extracellular vesicles in cancer.Pharmacology&Therapeutics, 188,pp.1-11.

[0654] Geena X Wu,Dan J.Raz,Laura Brown,and Virginia Sun,2016.Psychological burden associated with lung cancer screening;a systematic review.Clin Lung Cancer,2016 Sep;17(5):315-324.

[0655] Charles S.Dela Cruz,Lynn T.Tanoue,and Richard A.Matthay.2011.Lung Cancer:Epidemiology,Etiology,and Prevention,Clin Chest Med 32(2011)605-644.doi:10.1016/j.ccm.2011.09.001

[0656] Darmanis,S.,Nong,R.Y.,Hammond,M.,Gu,J.,Alderborn,A.,Vänelid,J., Siegbahn,A,Gustafsdottir,S.,Ericsson,O.,Landegren,U.and Kamali-Moghaddam,M., 2010.Sensitive plasma protein analysis by microparticle-based proximity ligation assays.Molecular&cellular proteomics,9(2),pp.327-335.

[0657] HowladerN,NooneAM,KrapchoM,MillerD,Brest A,Yu M,Ruhl J,Tatalovich Z, Mariotto A,Lewis DR,Chen HS,Feuer EJ,Cronin KA(eds).SEER Cancer Statistics Review,1975-2016,National Cancer Institute.Bethesda,MD,https://seer.cancer.gov/csr/1975_2016/,based on November 2018 SEER data submission, posted to the SEER web site,April 2019.

[0658] Im,H.,Shao,H,Park,Y.I.,Peterson,V.M.,Castro,C.M.,Weissleder,R.and Lee,H.,2014.Label-free detection and molecular profiling of exosomes with a nano-plasmonic sensor.Nature biotechnology,32(5),p.490.

[0659] Jeong,S.,Park,J.,Pathania,D.,Castro,C.M.,Weissleder,R.andLee,H., 2016.Integrated magneto-electrochemical sensor for exosome analysis.A CS nano,10(2),pp.1802-1809.

[0660] Kinsinger LS,Anderson C,Kim J,et al.Implementation of LungCancer Screening in the Veterans Health Administration.JAMA Intern Med.2017;177(3): 399-406.

[0661] doi:10.1001/jamainternmed.2016.9022

[0662] Martin M Oken,Willam G Hocking,Paul A Kvale,Gerald L Andriole,Saundra S Buys,Timothy R Church,E David Crawford,Mona N Fouad,Claudine Isaacs,Douglas J Reding,Joel L Weissfeld,Lance A Yokochi,BarbaraO'Brien,Lawrence R Ragard,

[0663] Joshua M Rathmell,Thomas L Riley,Patrick Wright,Neil Caparaso,Ping Hu,Grant Izmirlian,Paul F Pinsky,Philip C Prorok,Barnett S Kramer,Anthony B

Miller, John K Gohagan, Christine D Berg, PLCO Project Team., 2011. Screening by Chest Radiograph and Lung Cancer Mortality: The Prostate, Lung, Colorectal, and Ovarian (PLCO) Randomized Trial. JAMA, 2011 Nov 2; 306(17): 1865-1873.

[0664] Raghu VK, Zhao W, Pu J, et al. 2019. Feasibility of lung cancer prediction from low-dose CT scan and smoking factors using causal models. Thorax 2019; 74: 643-649.

[0665] Shao, H., Im, H., Castro, C.M., Breakefield, X., Weissleder, R. and Lee, H., 2018. New technologies for analysis of extracellular vesicles. Chemical reviews, 118(4), pp. 1917-1950.

[0666] 等效方案

[0667] 本领域技术人员仅仅使用常规试验将认识到或者能够确定本文所描述的发明的具体实施方案的很多等效方案。应当了解的是, 除非另外指明或除非对于本领域的普通技术人员来说显然将会出现矛盾或不一致, 否则本发明涵盖了所有变化、组合以及置换, 其中来自所列权利要求中的一项或多项的一个或多个限制、要素、条款、描述性术语等被引入从属于同一基础权利要求(或相关的任何其他权利要求)的另一权利要求中。此外, 还应当了解的是, 本发明的任何实施方案或方面可以明确地被排除在权利要求之外, 无论在说明书中是否叙述了特定的排除。本发明的范围并非旨在限于以上描述, 而是如以下权利要求所阐述。

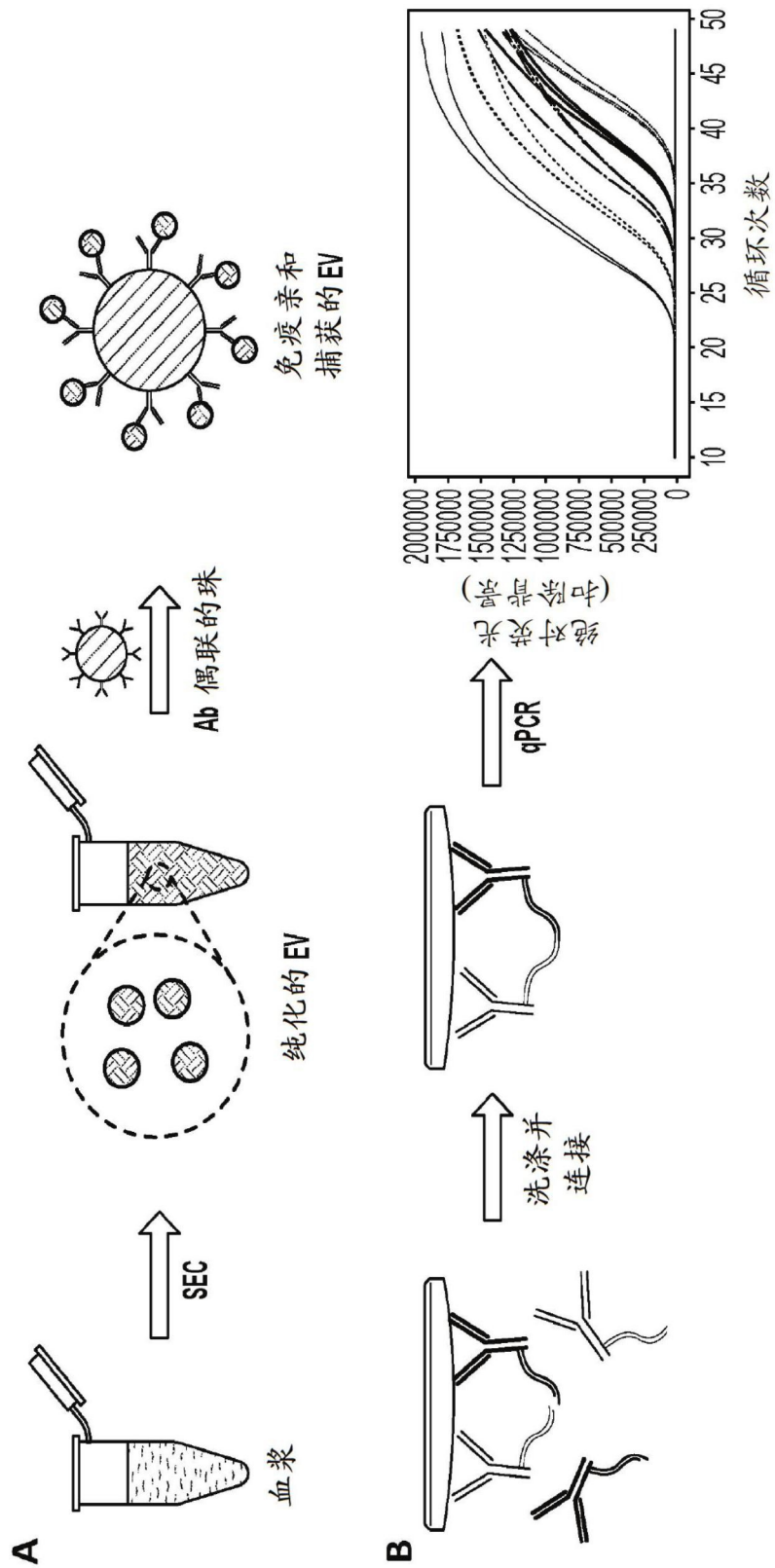


图1

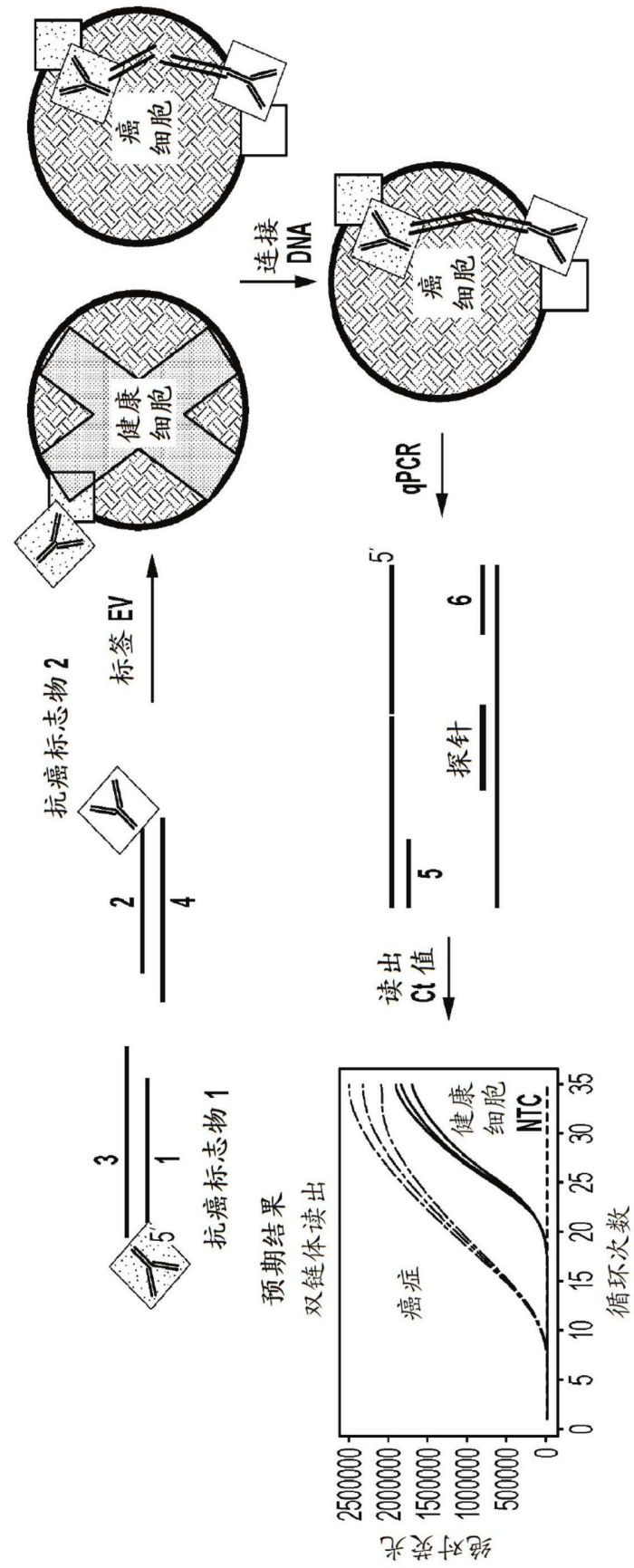


图2

肺癌和支气管癌的组织学患病率
(SEER 2013-2017 男女两性)

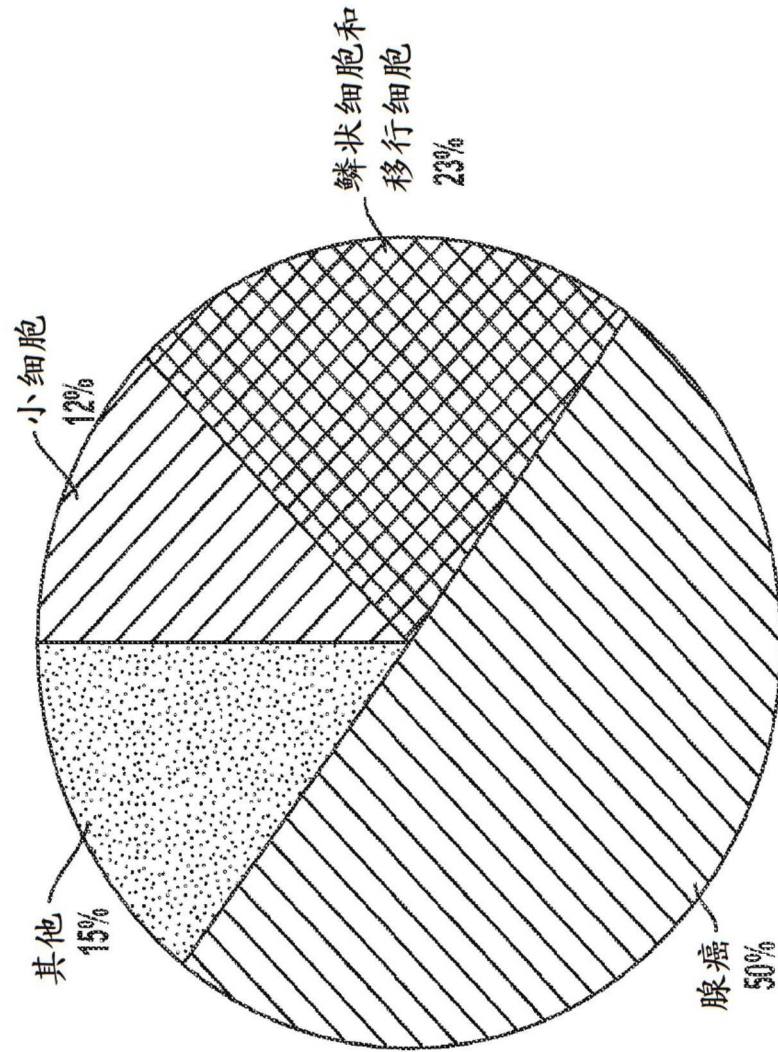


图3

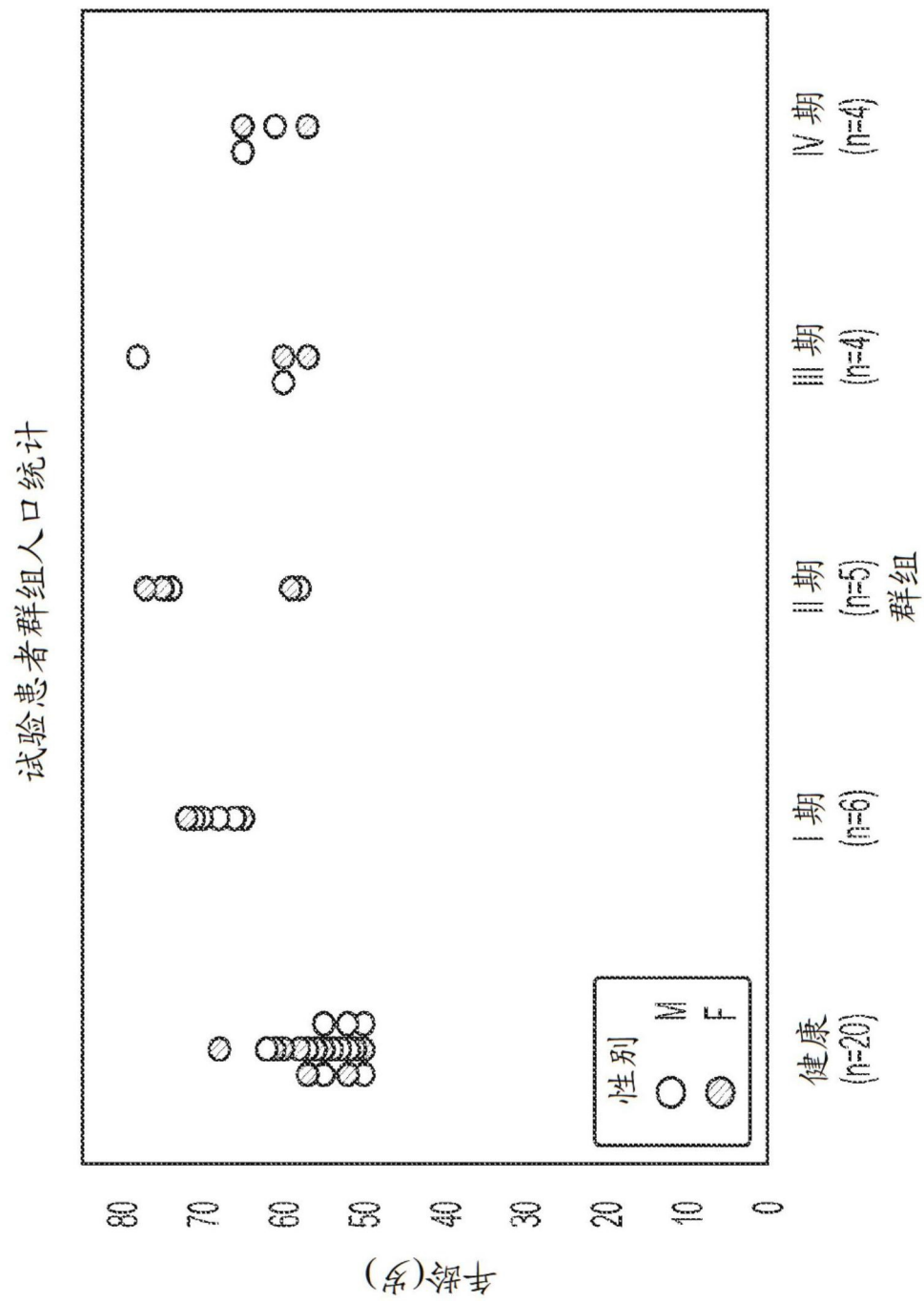


图4

使用 SLC34A2 捕获和 CEACAM6 + CEACAM6
定量探针的示例性肺腺癌诊断测定

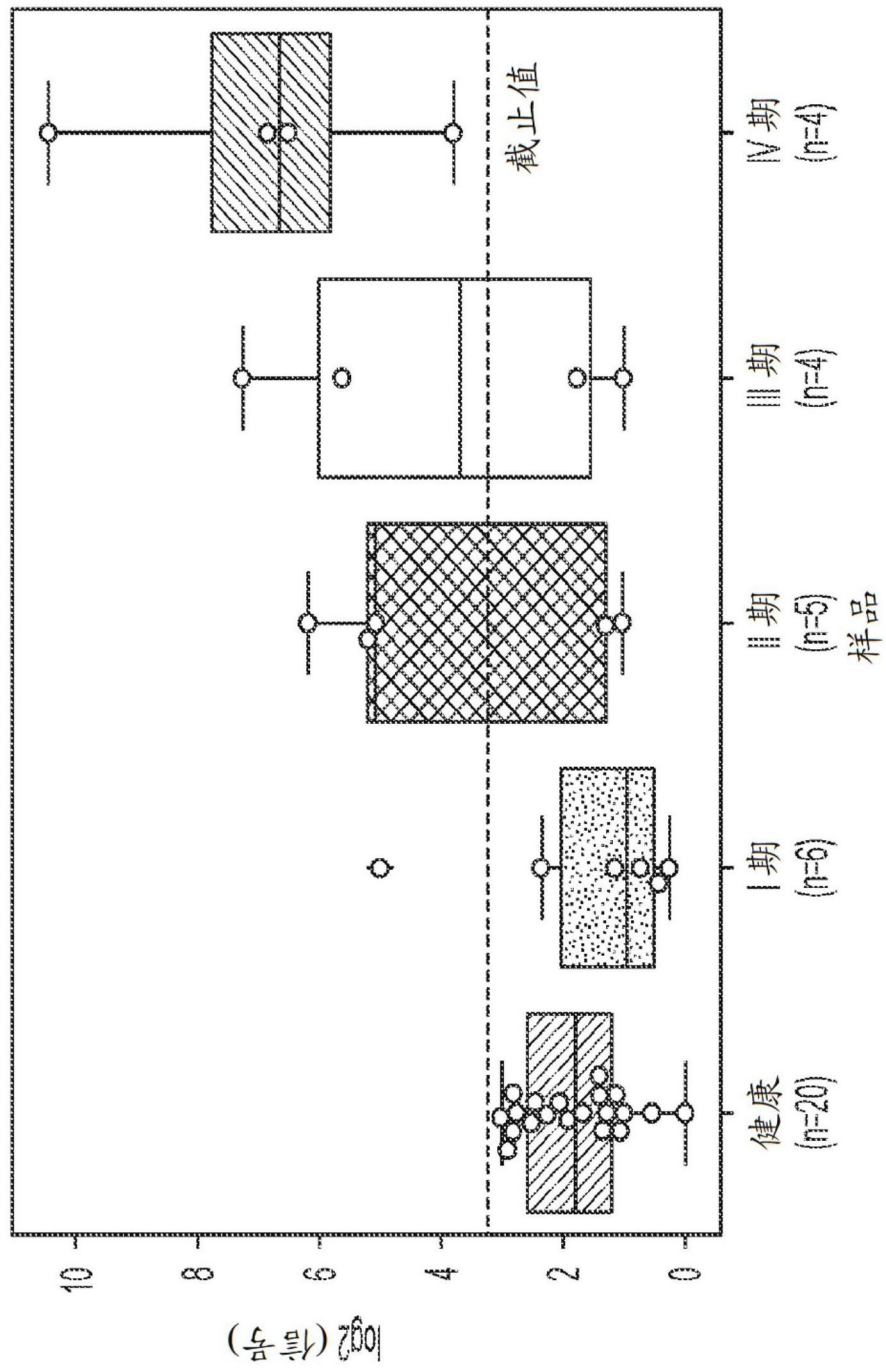


图5

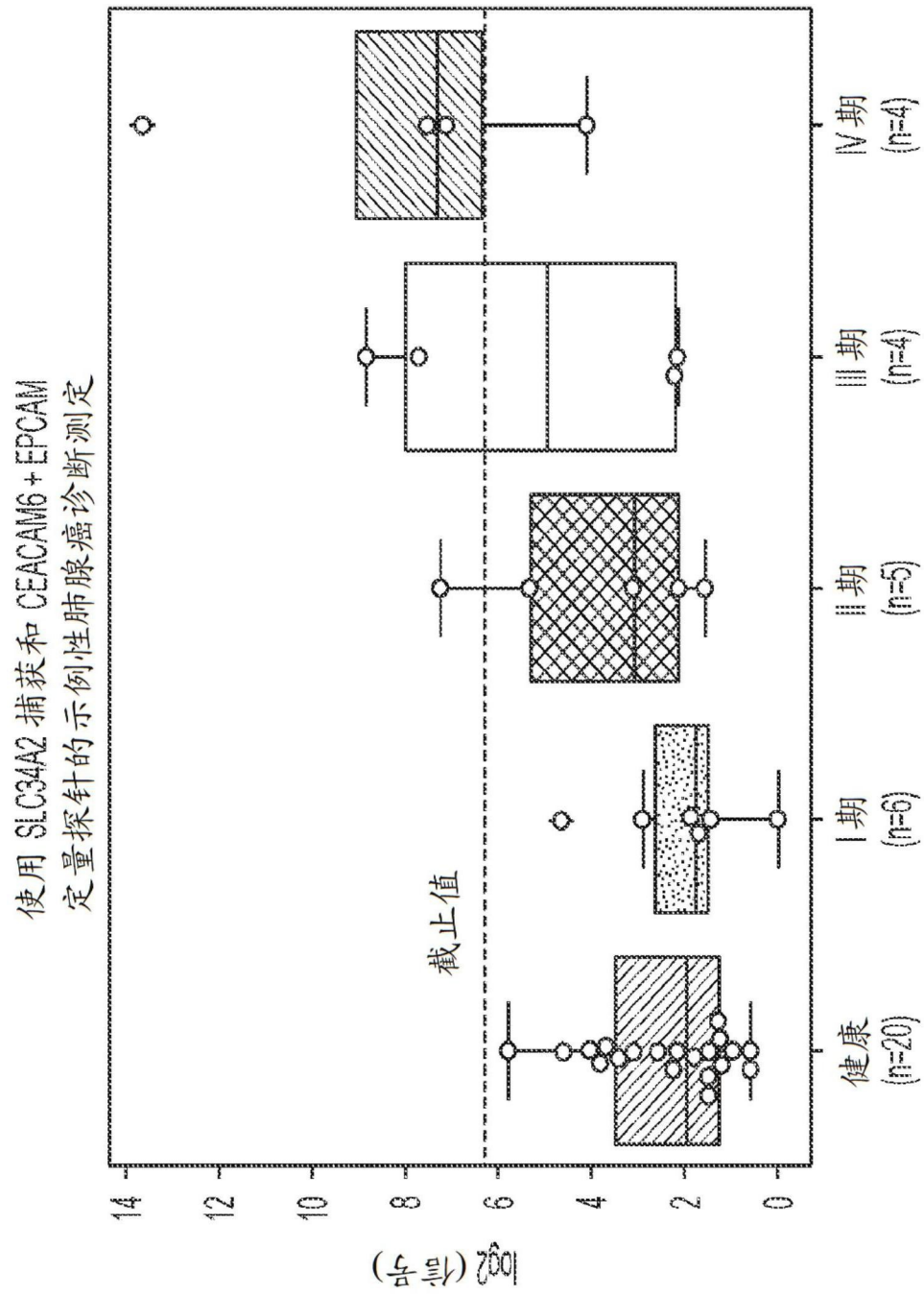


图6

使用 CEACAM5 捕获和 CEACAM6 + SLC34A2
定量探针的示例性肺癌诊断测定

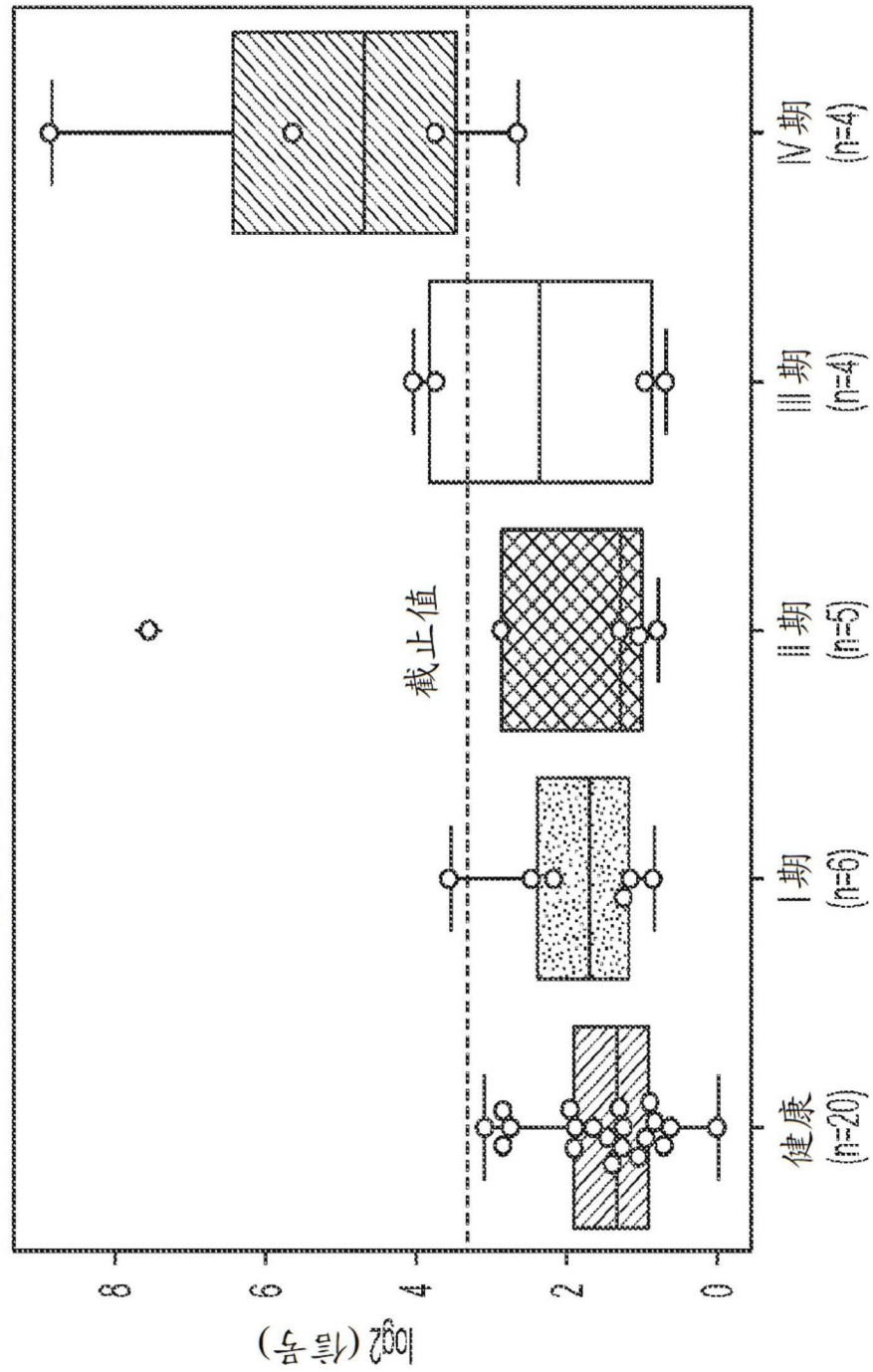


图7

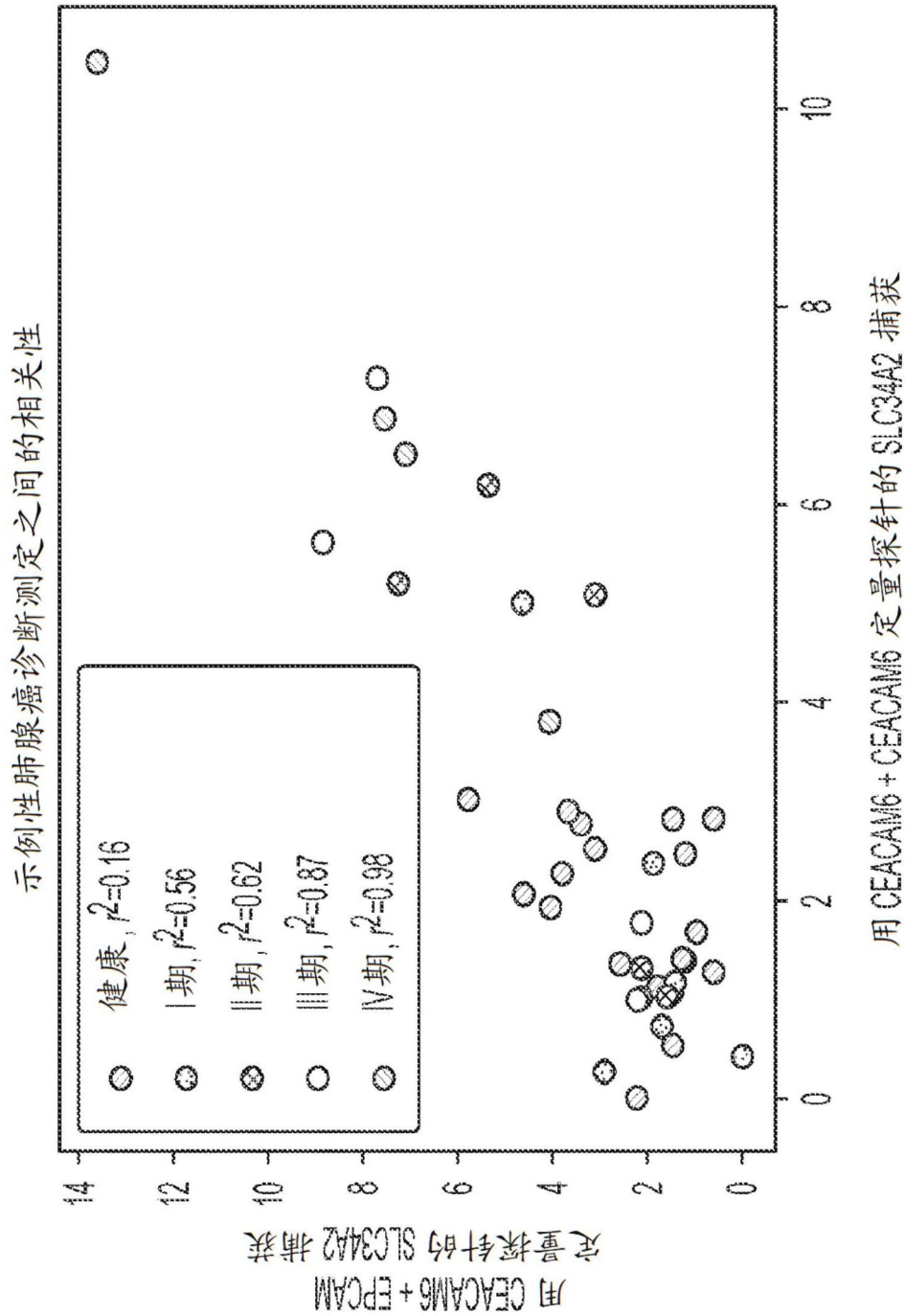


图8

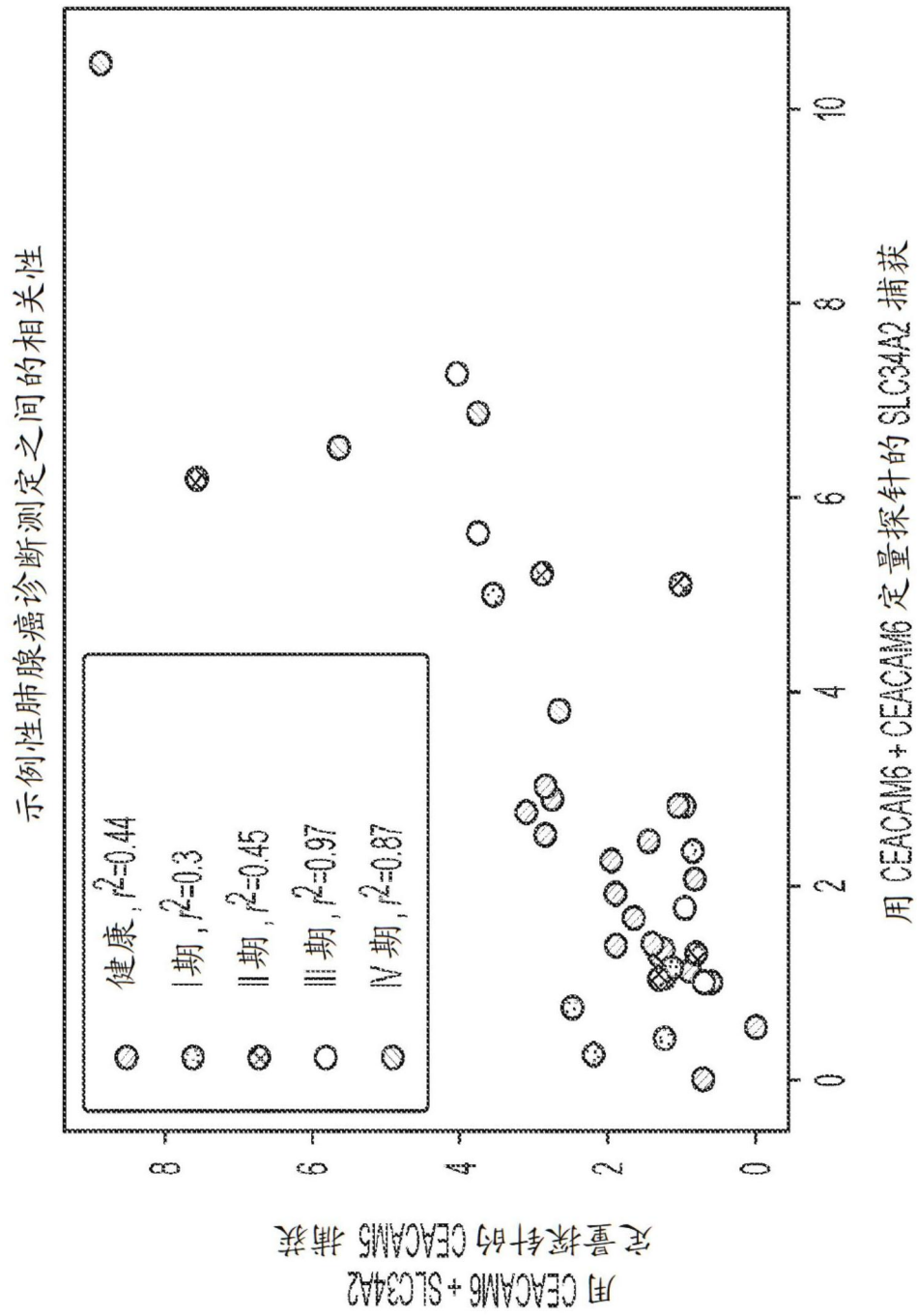


图9

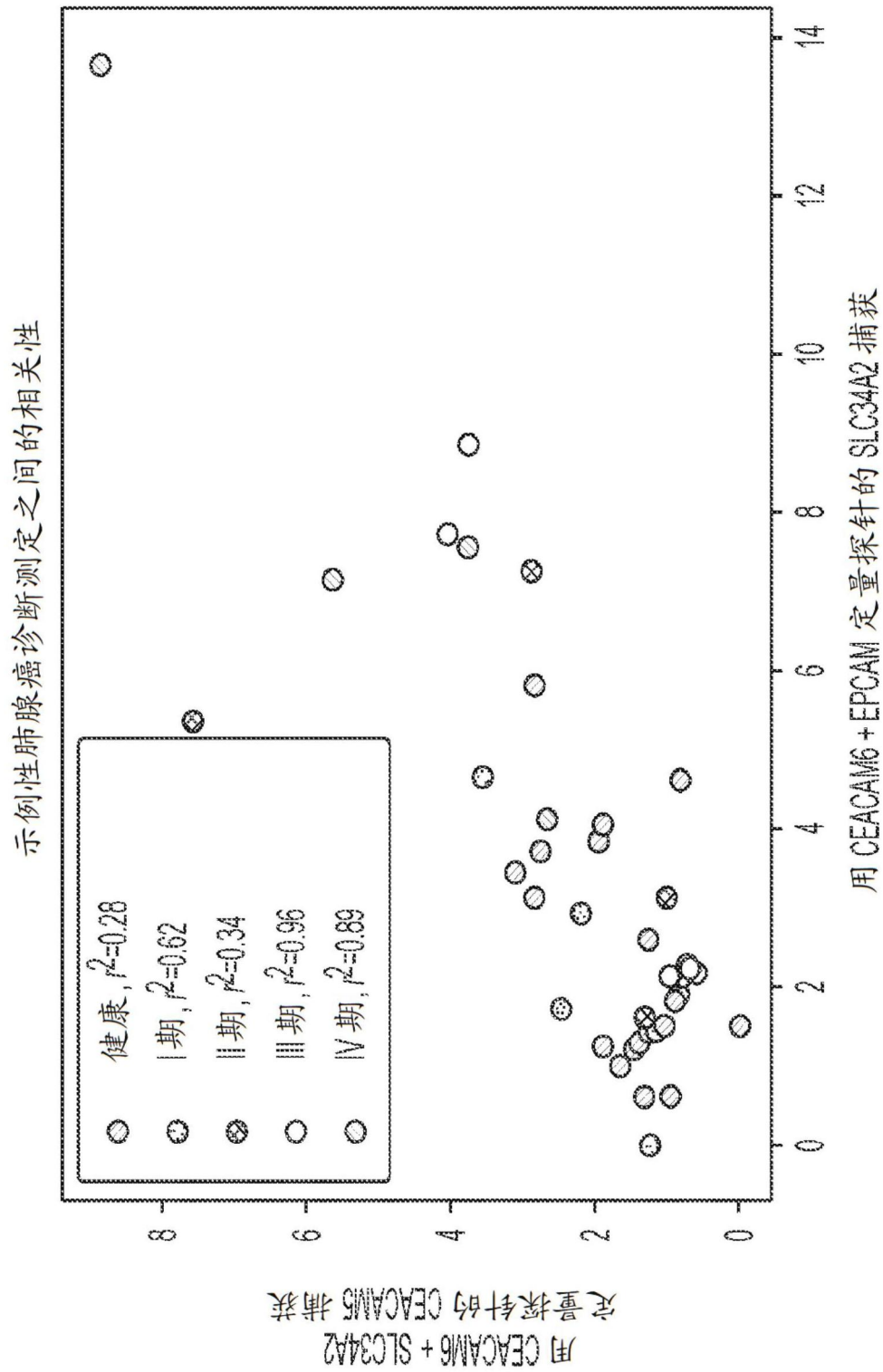


图10

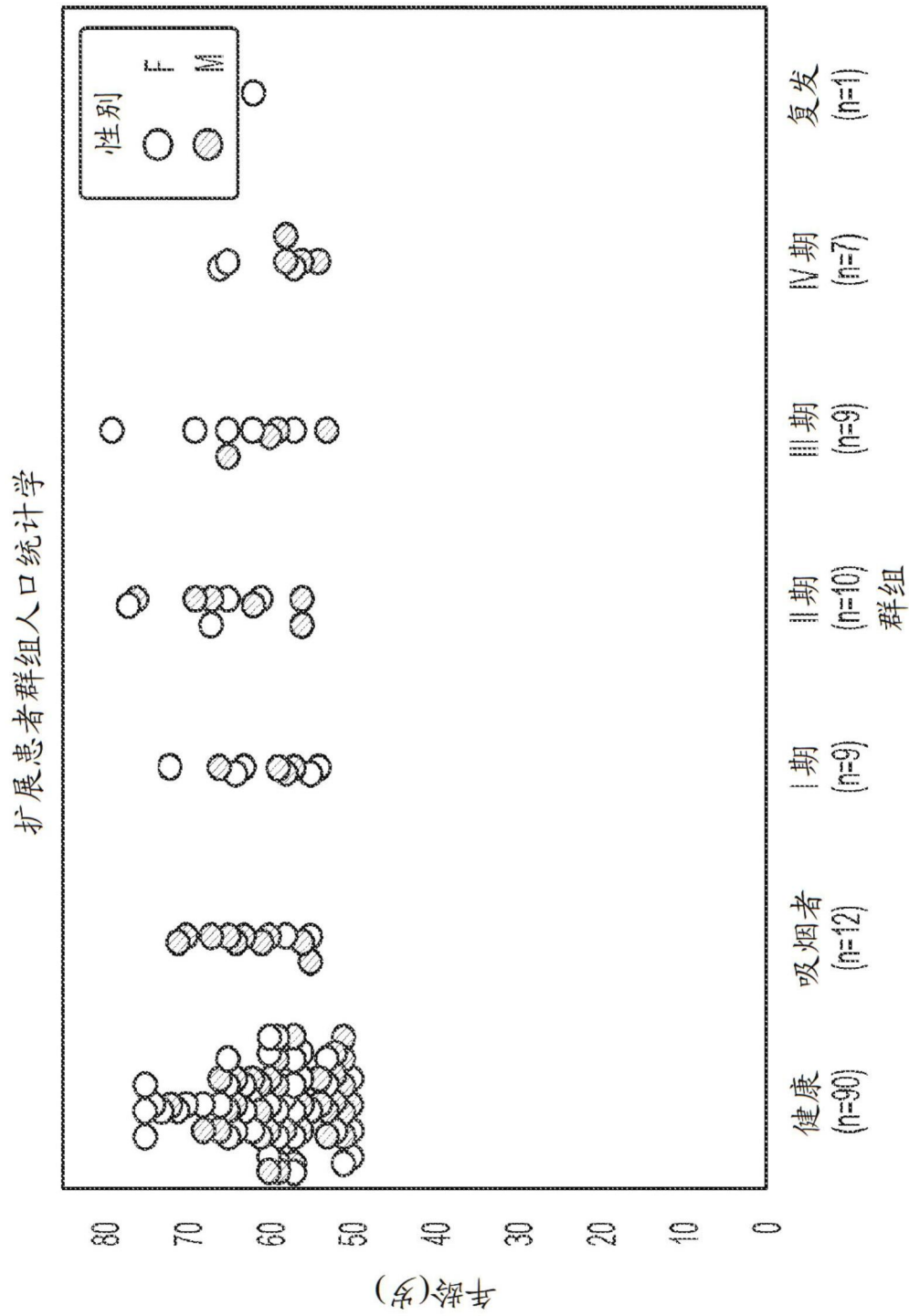


图11

使用SLC34A2捕获和CEACAM6+CEACAM6
定量探针的示例性肺癌诊断测定

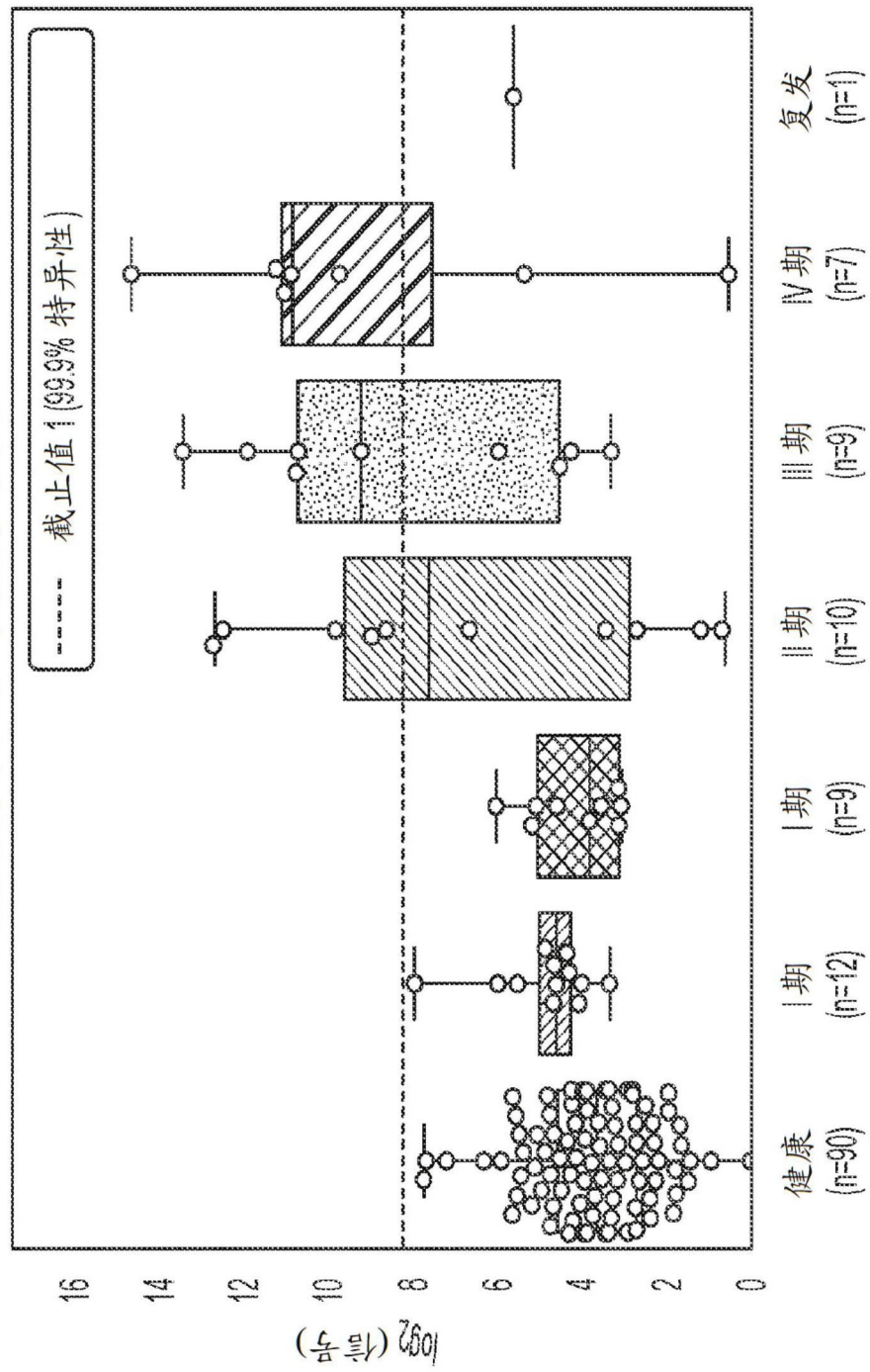


图12

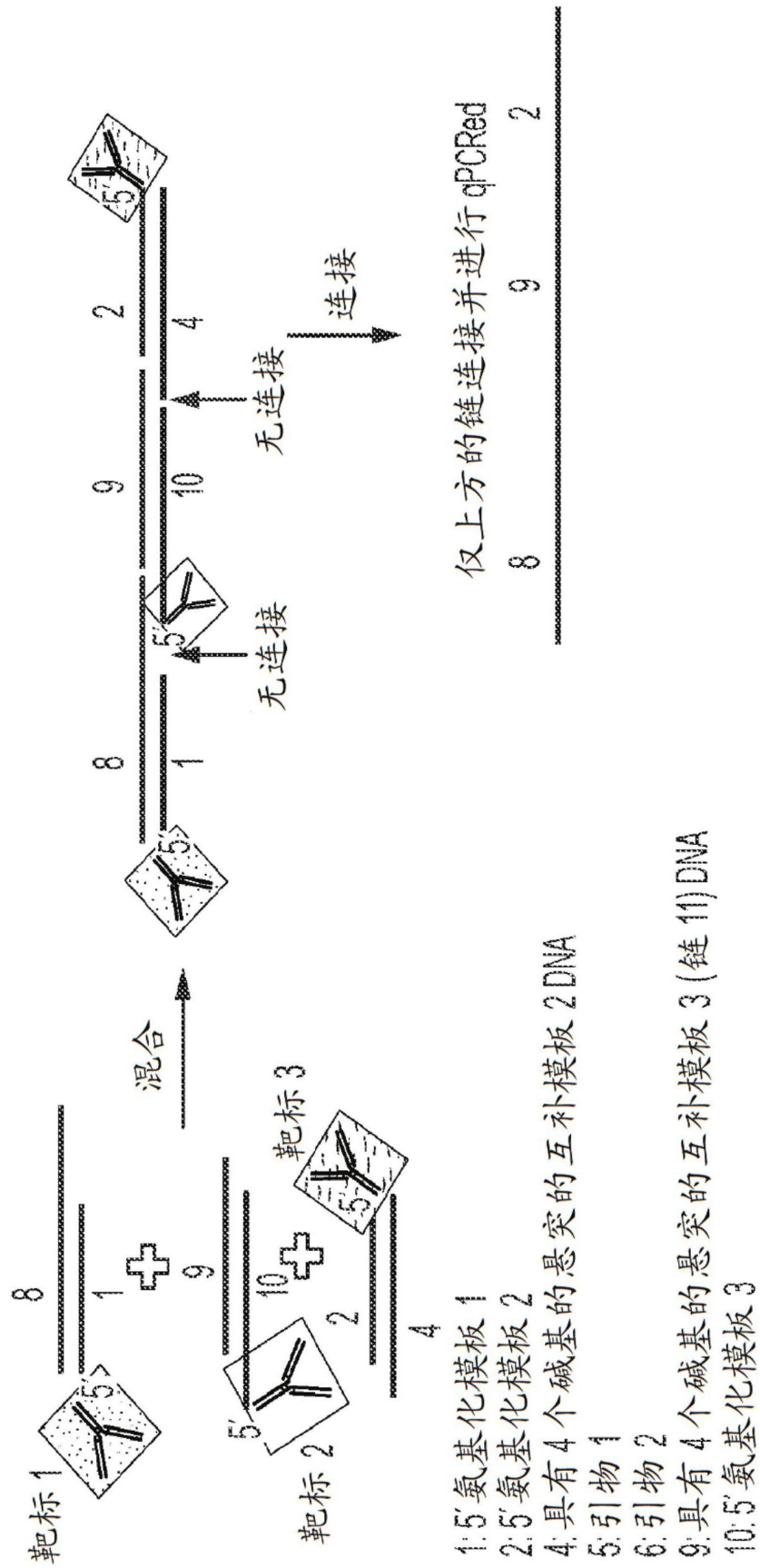


图13

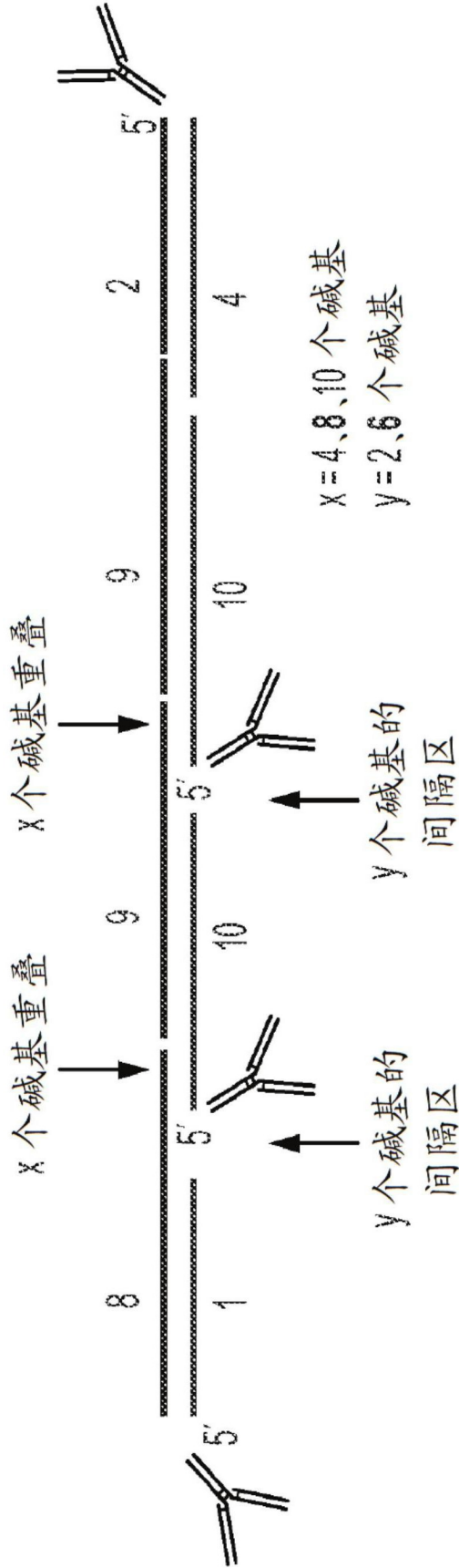


图14

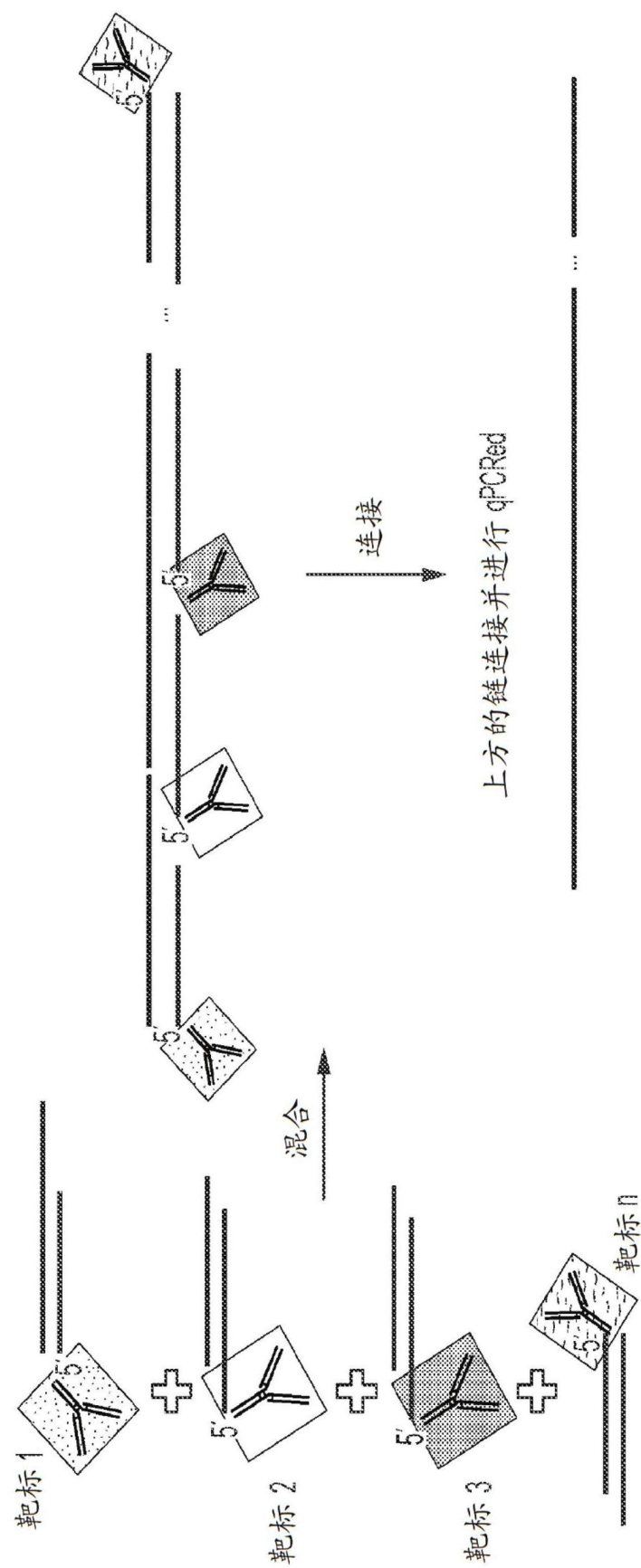


图15

具有不同大小的肿瘤的灵敏度

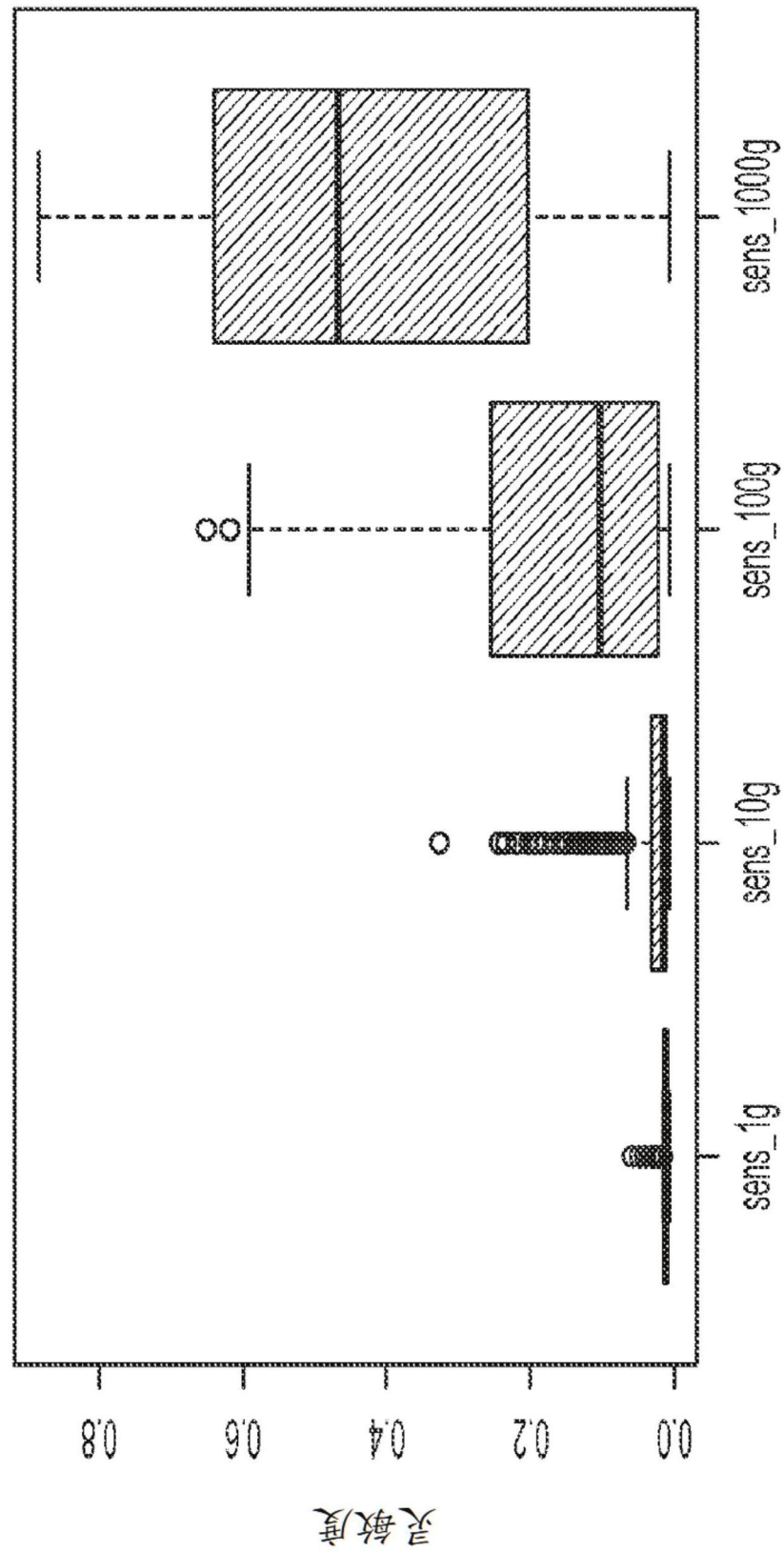


图16A

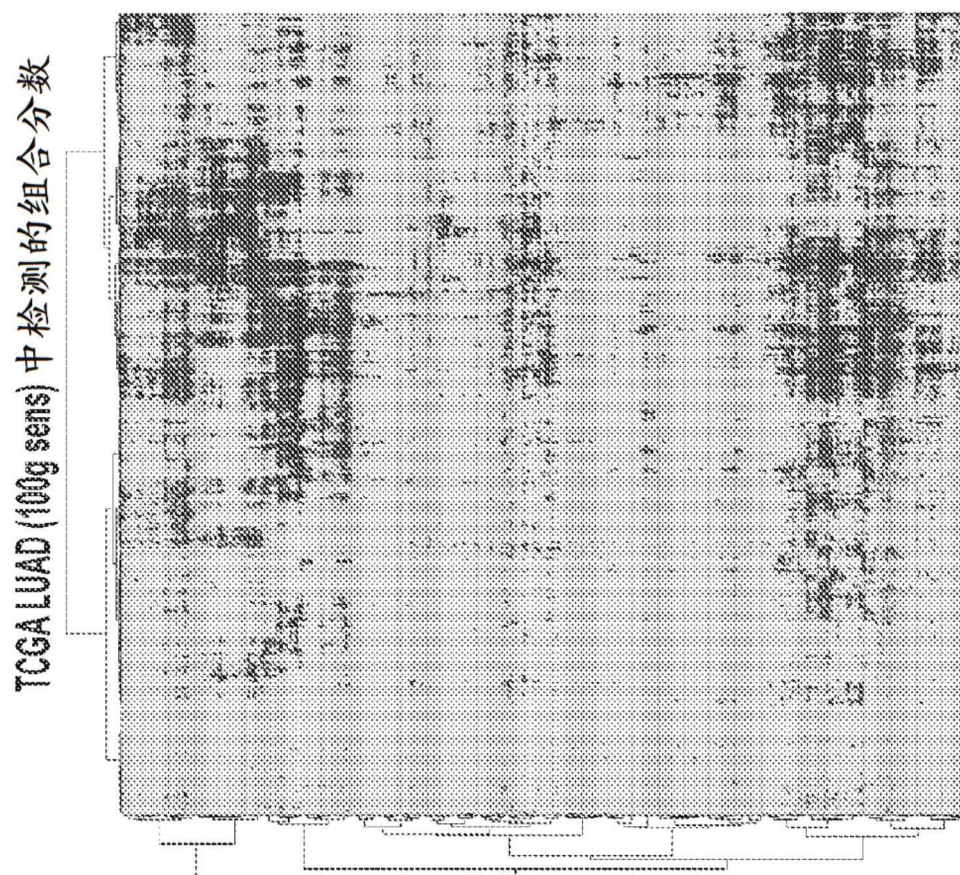


图16B

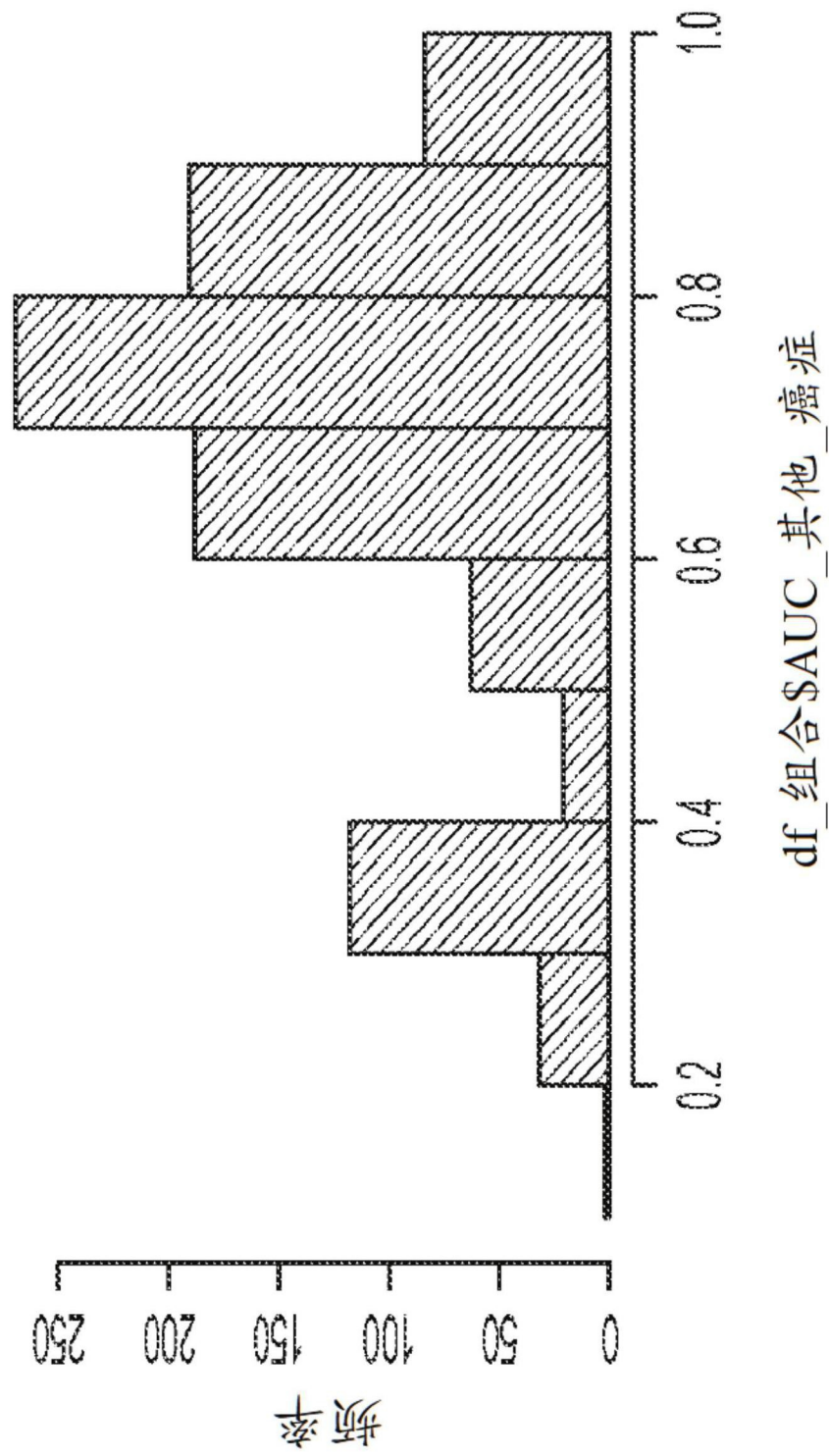


图16C

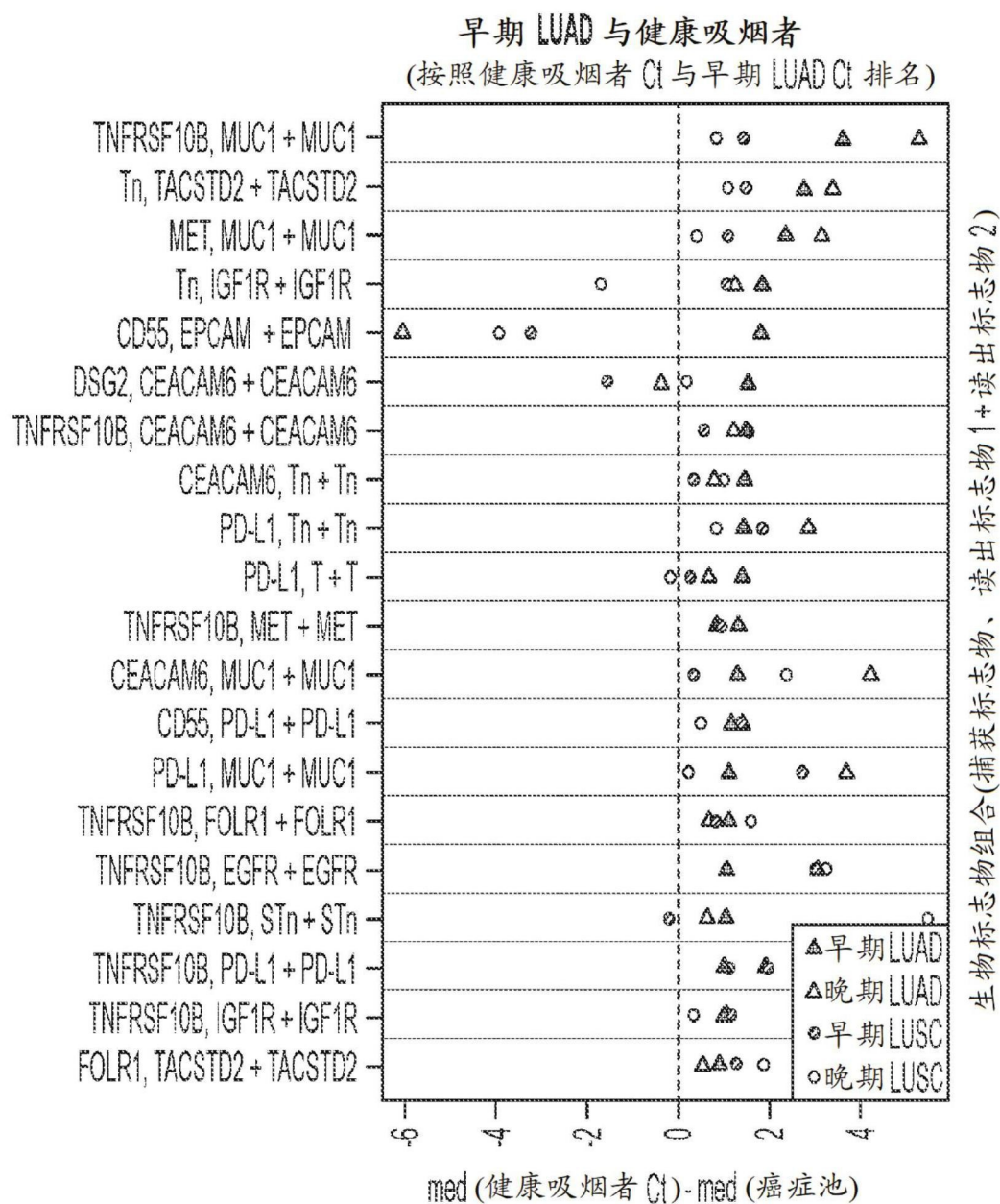


图17A

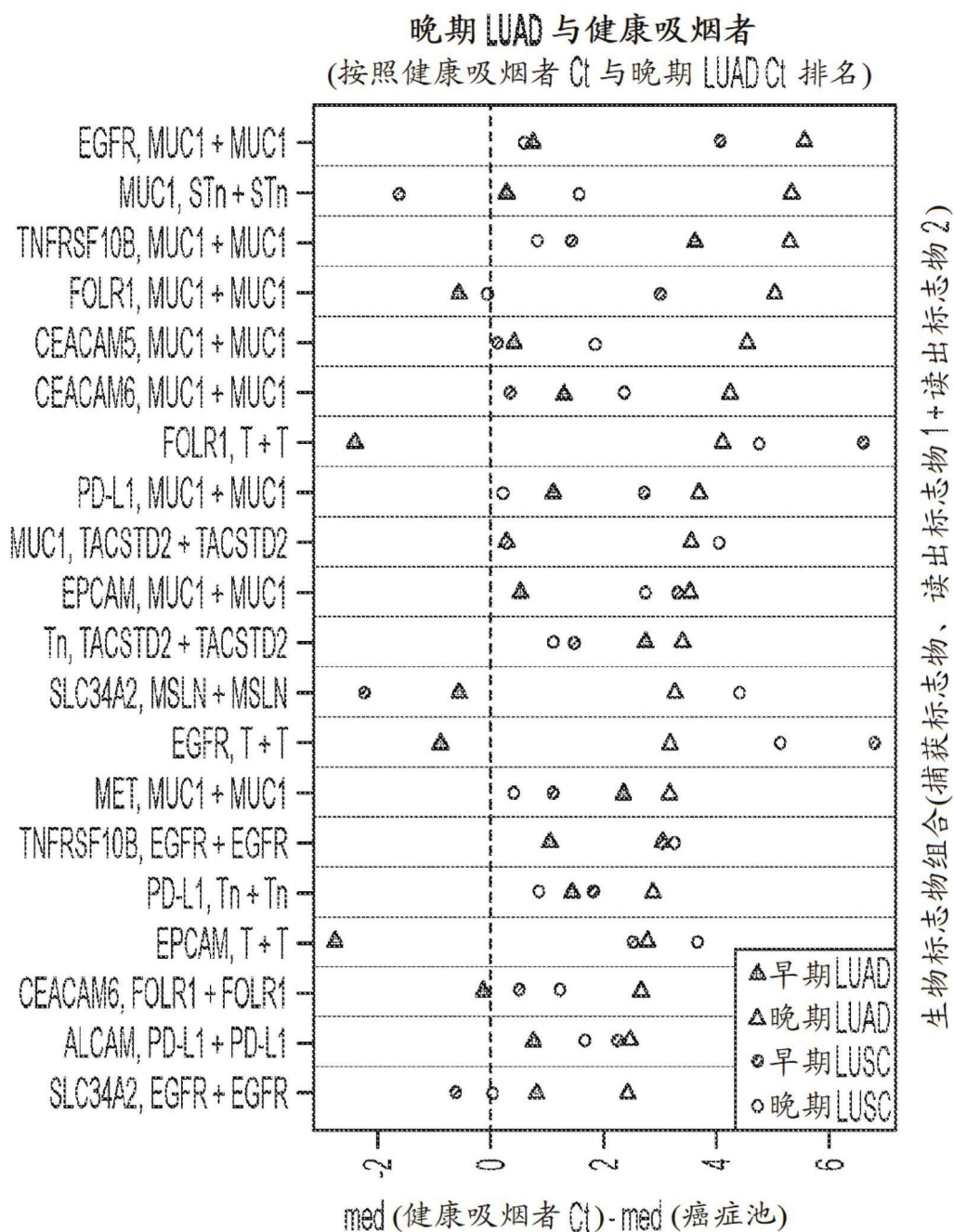


图17B

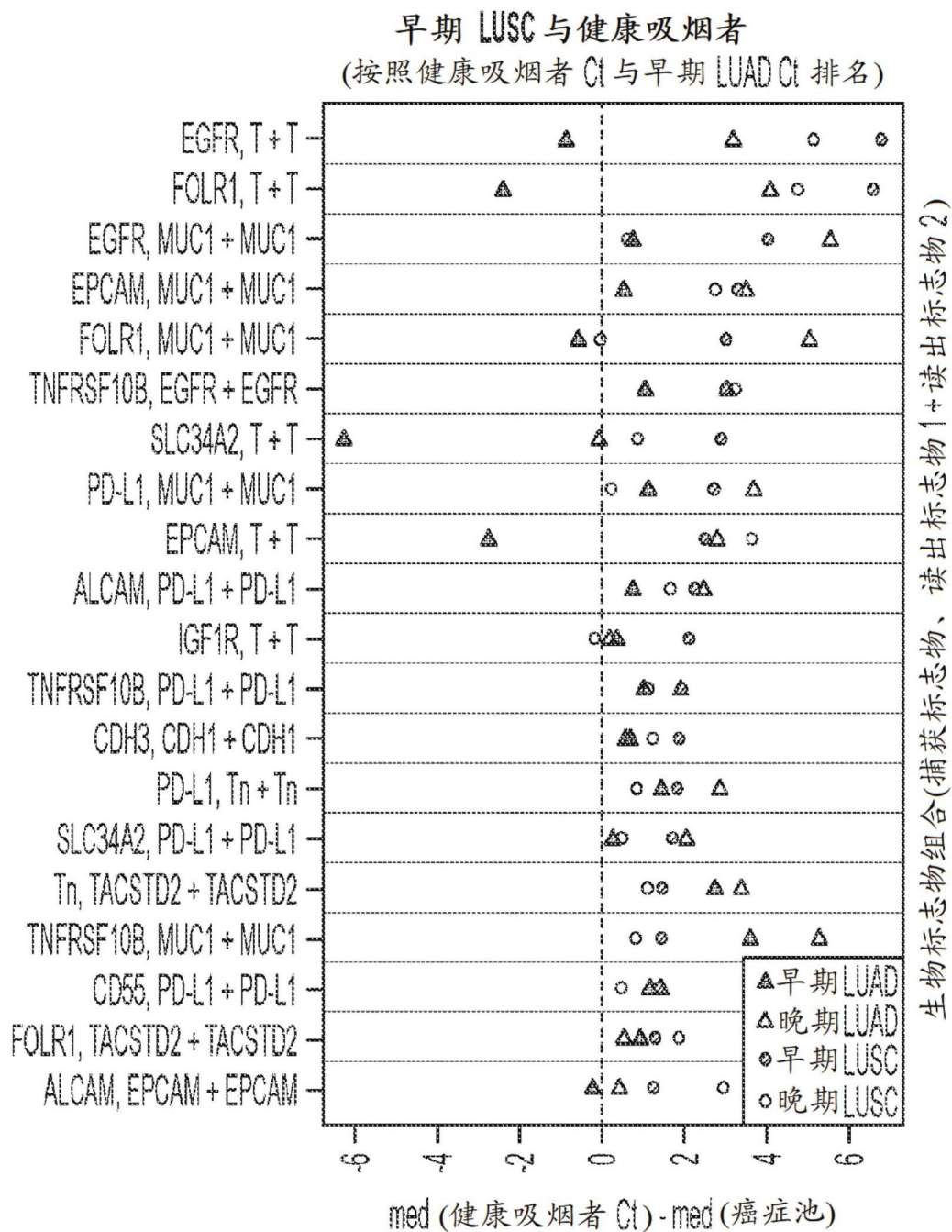


图18A

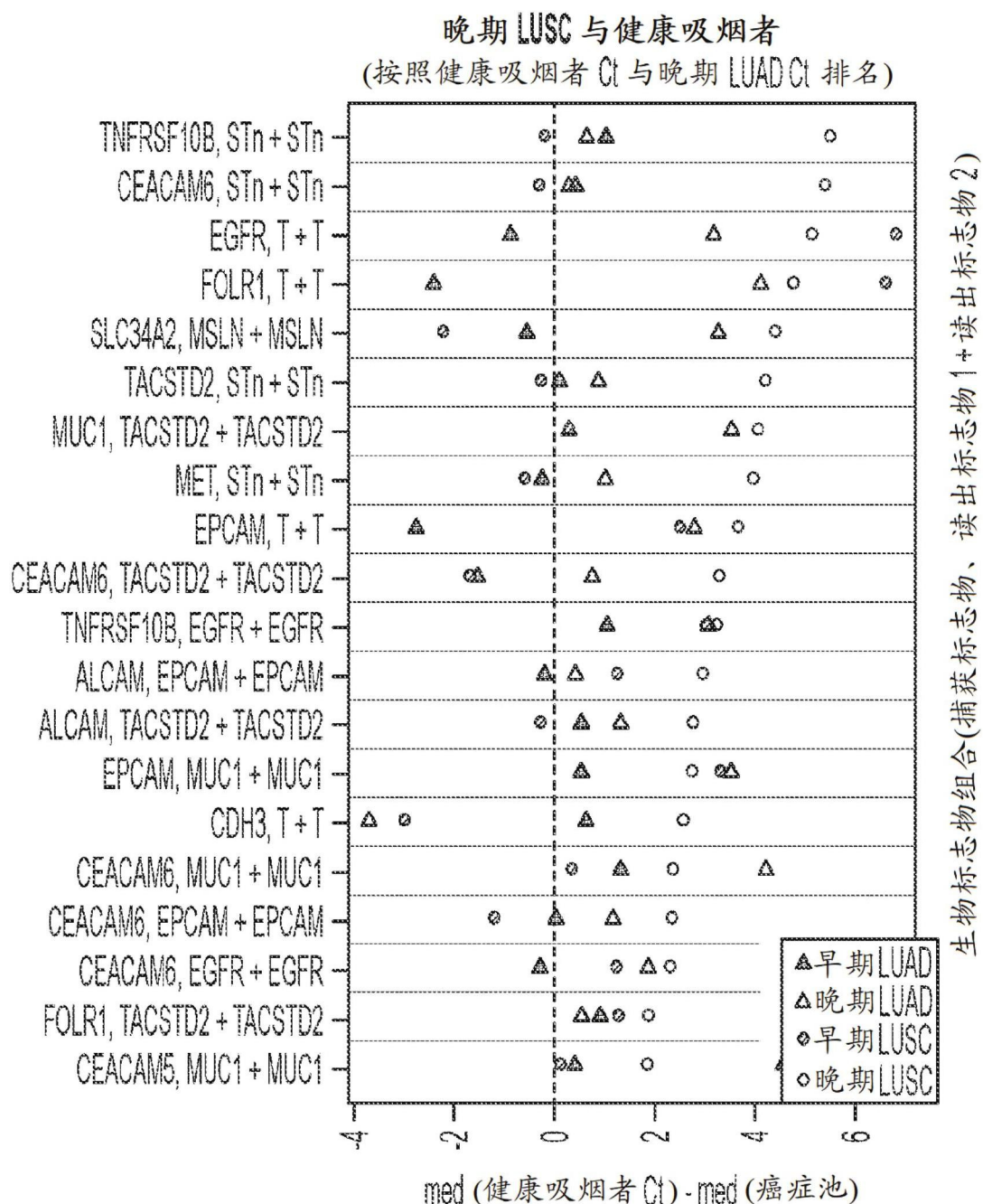


图18B

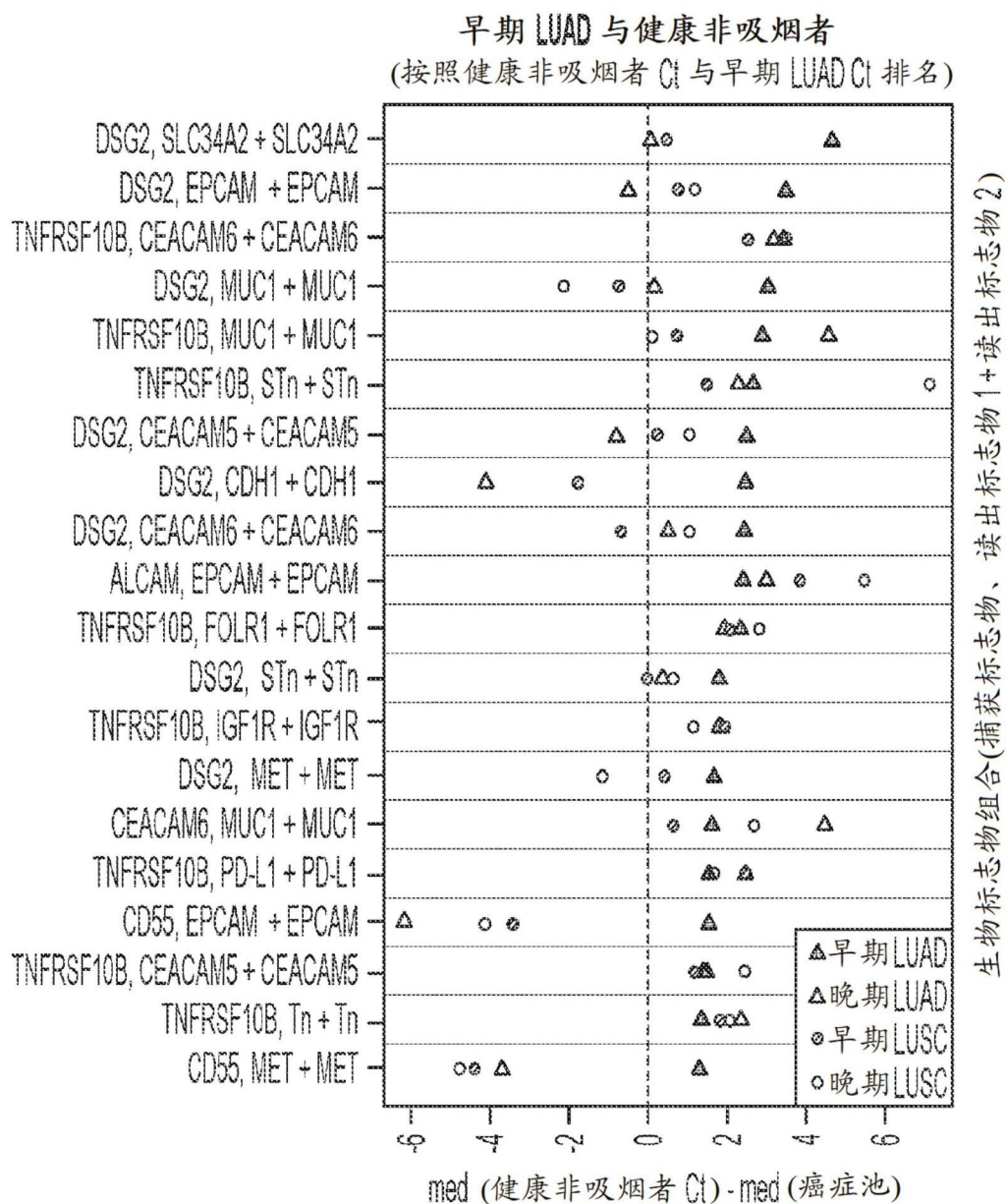


图19A

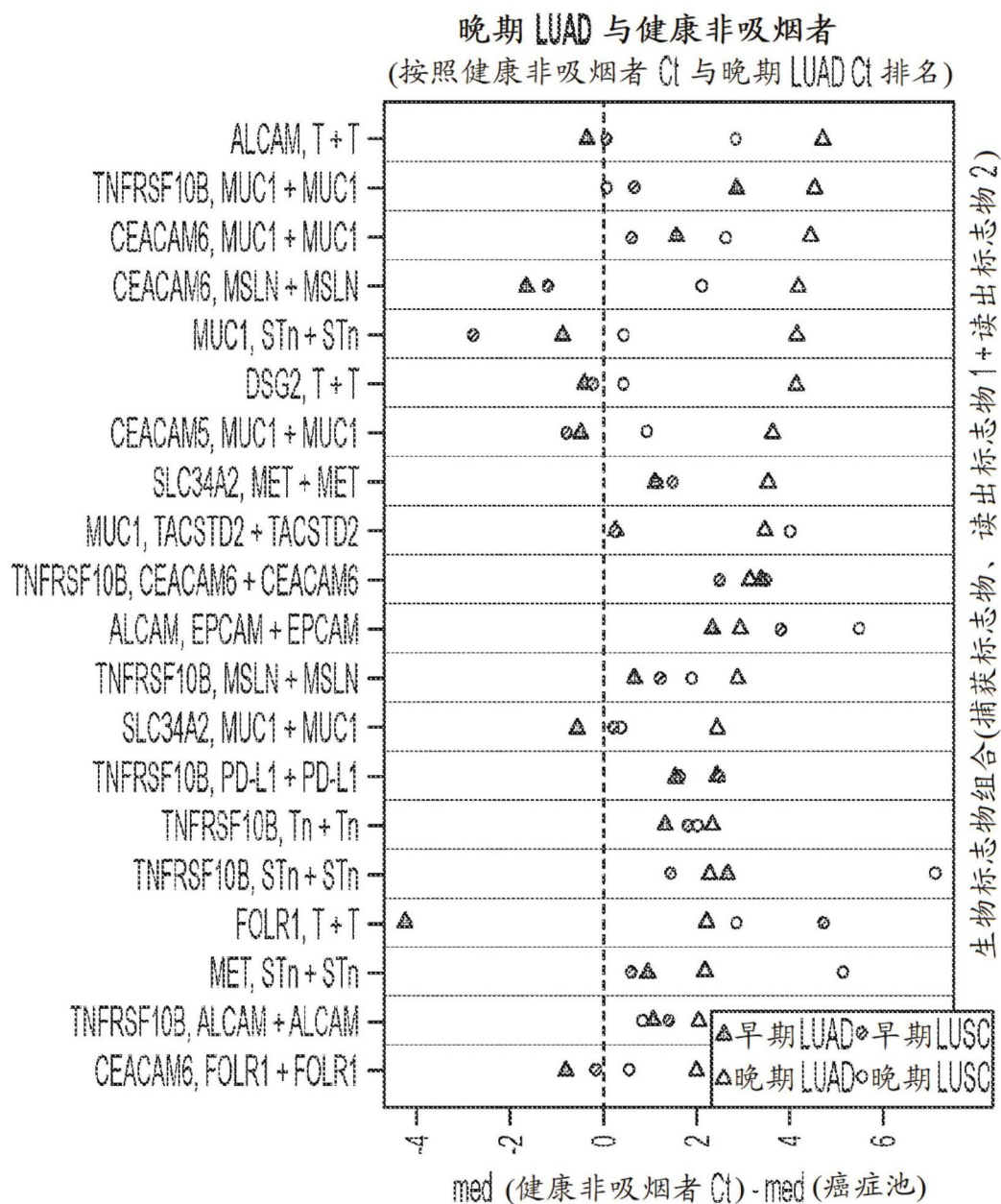


图19B

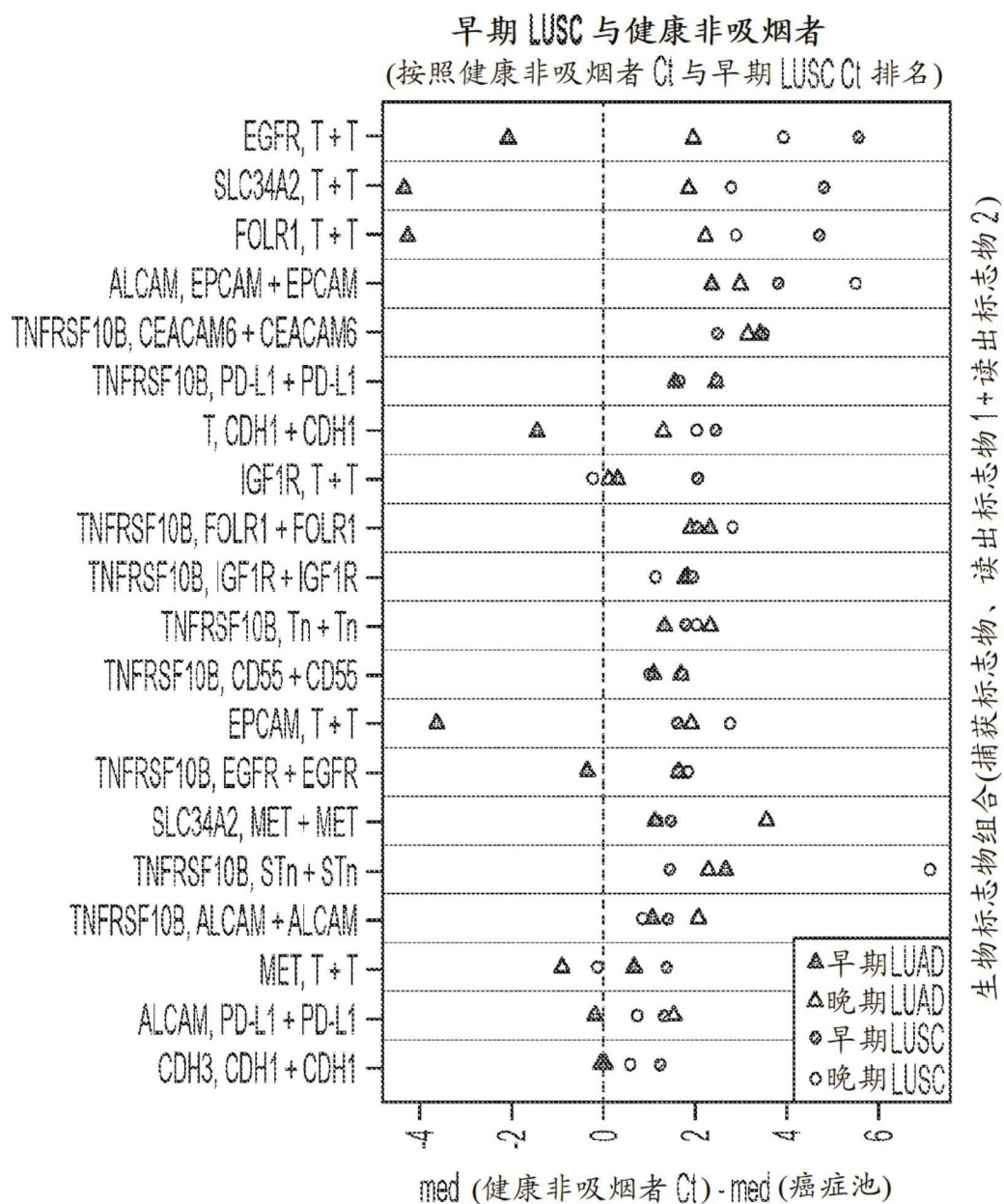


图20A

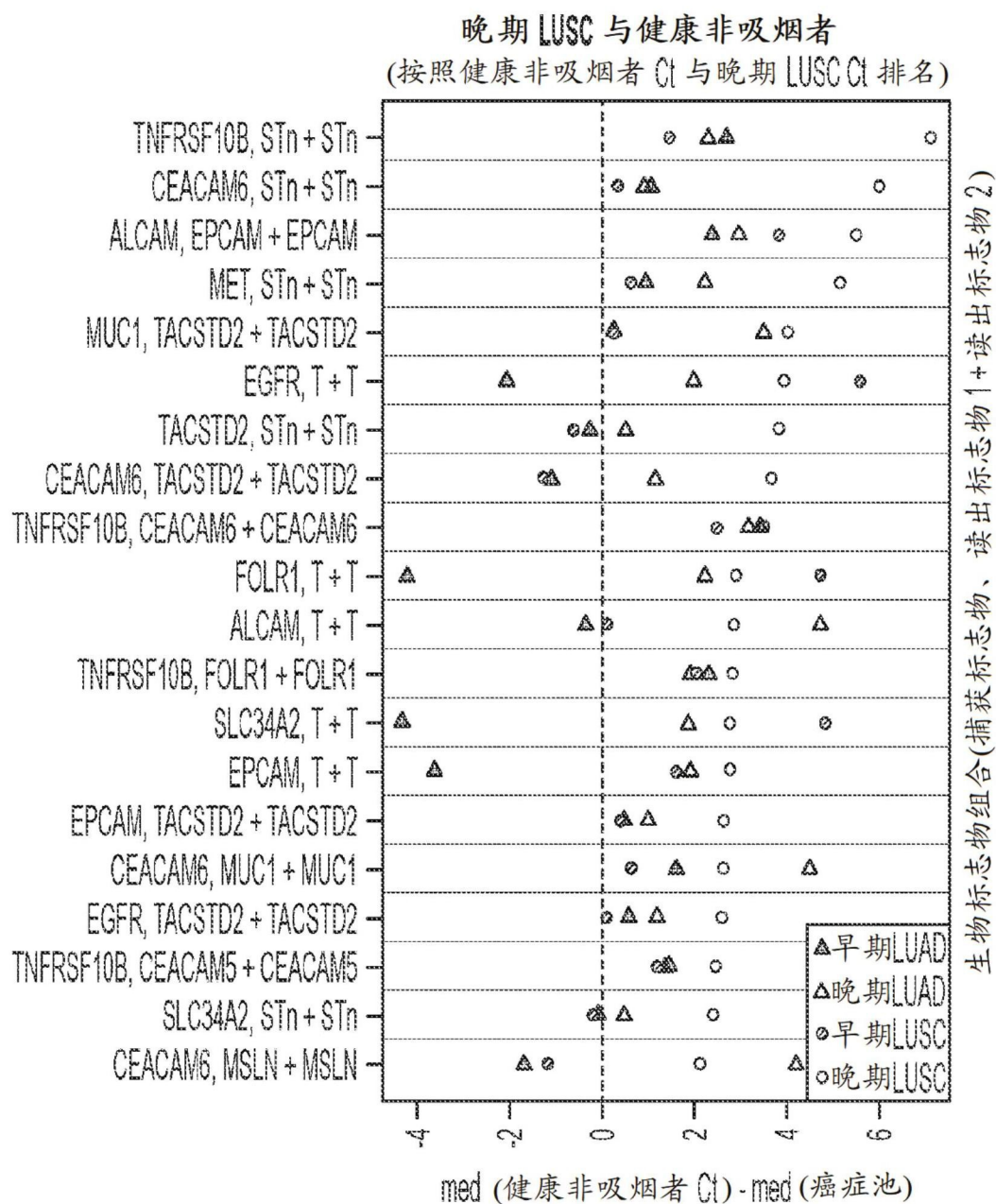


图20B

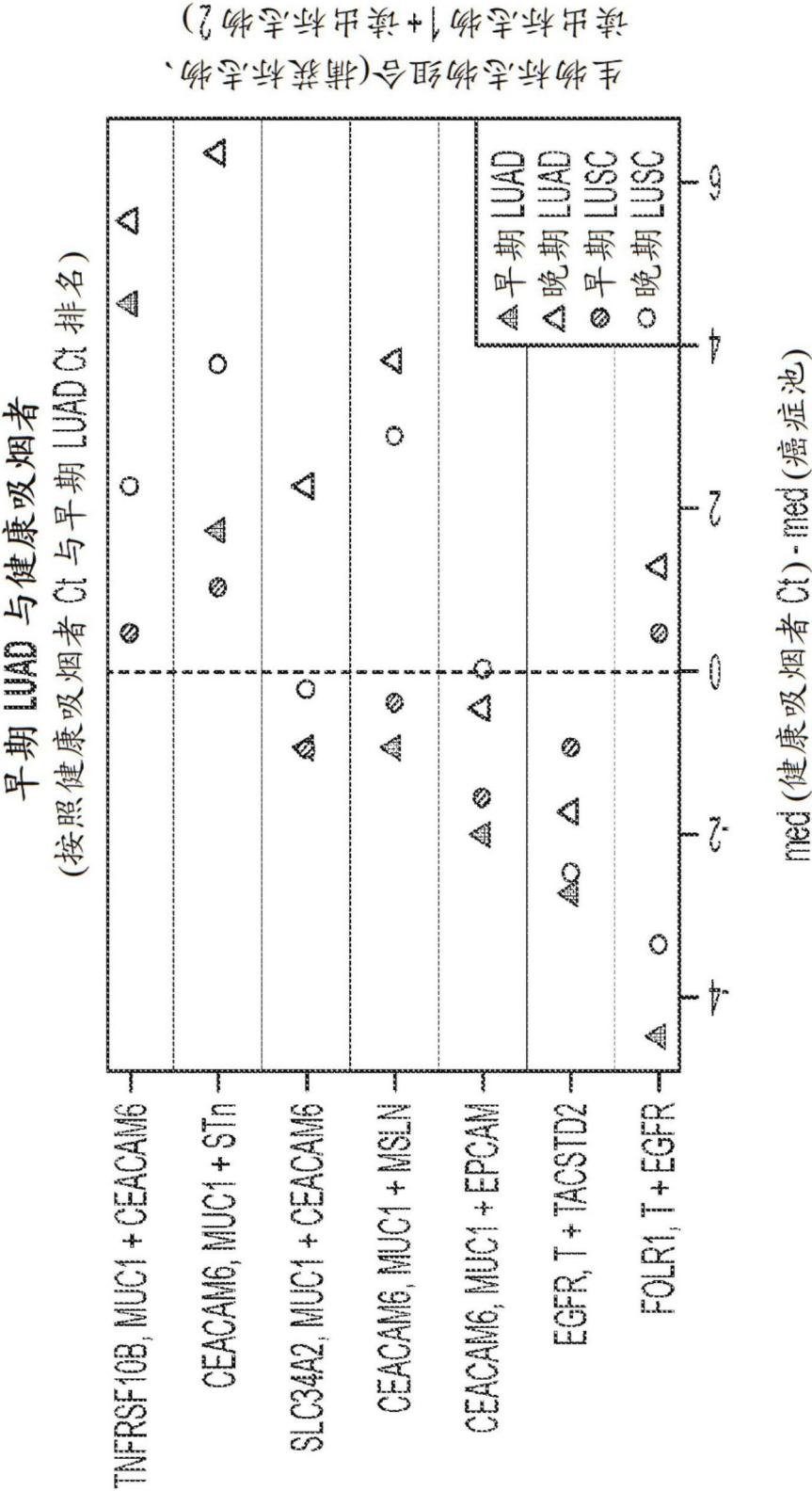


图21A

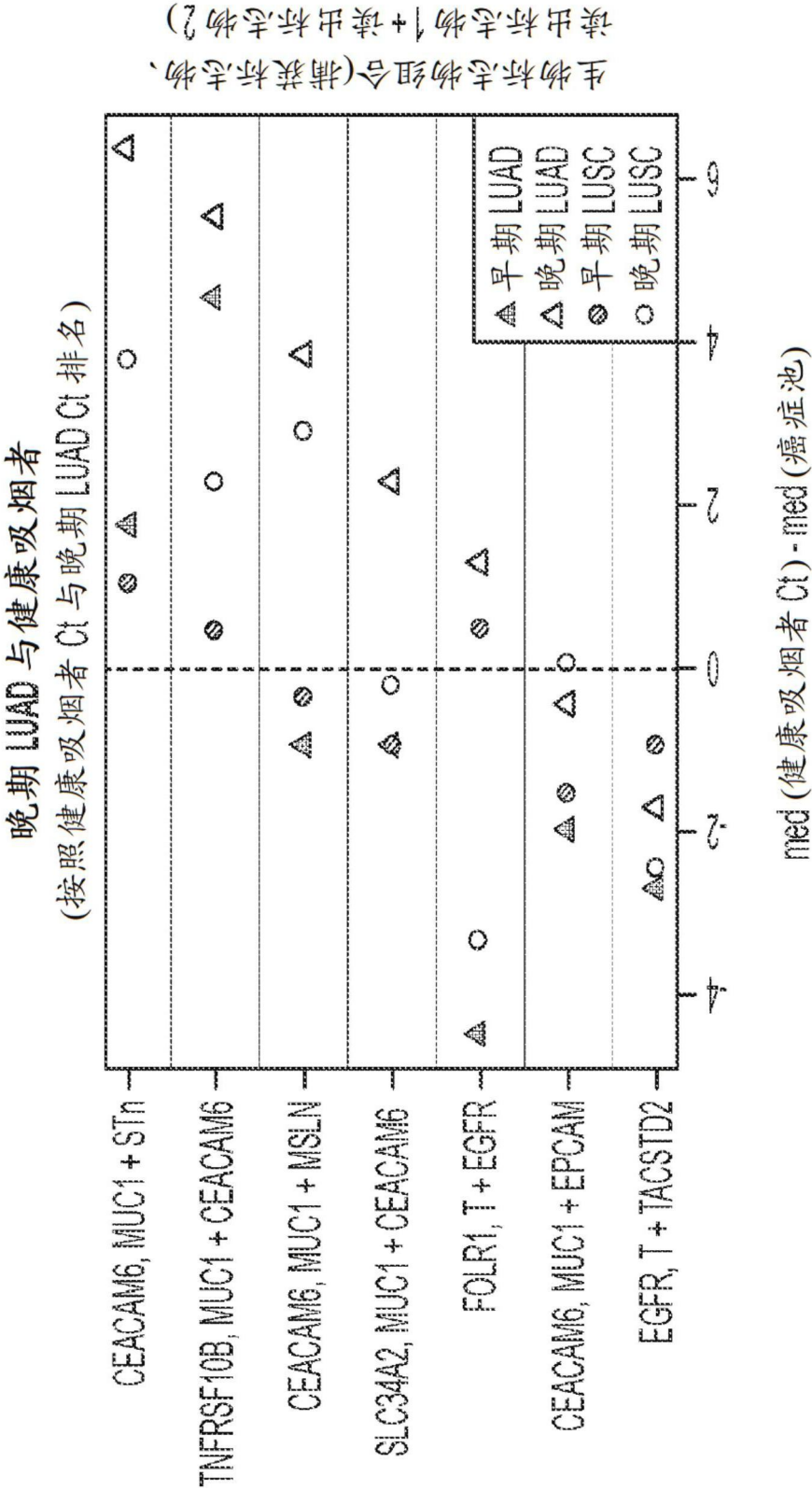


图21B

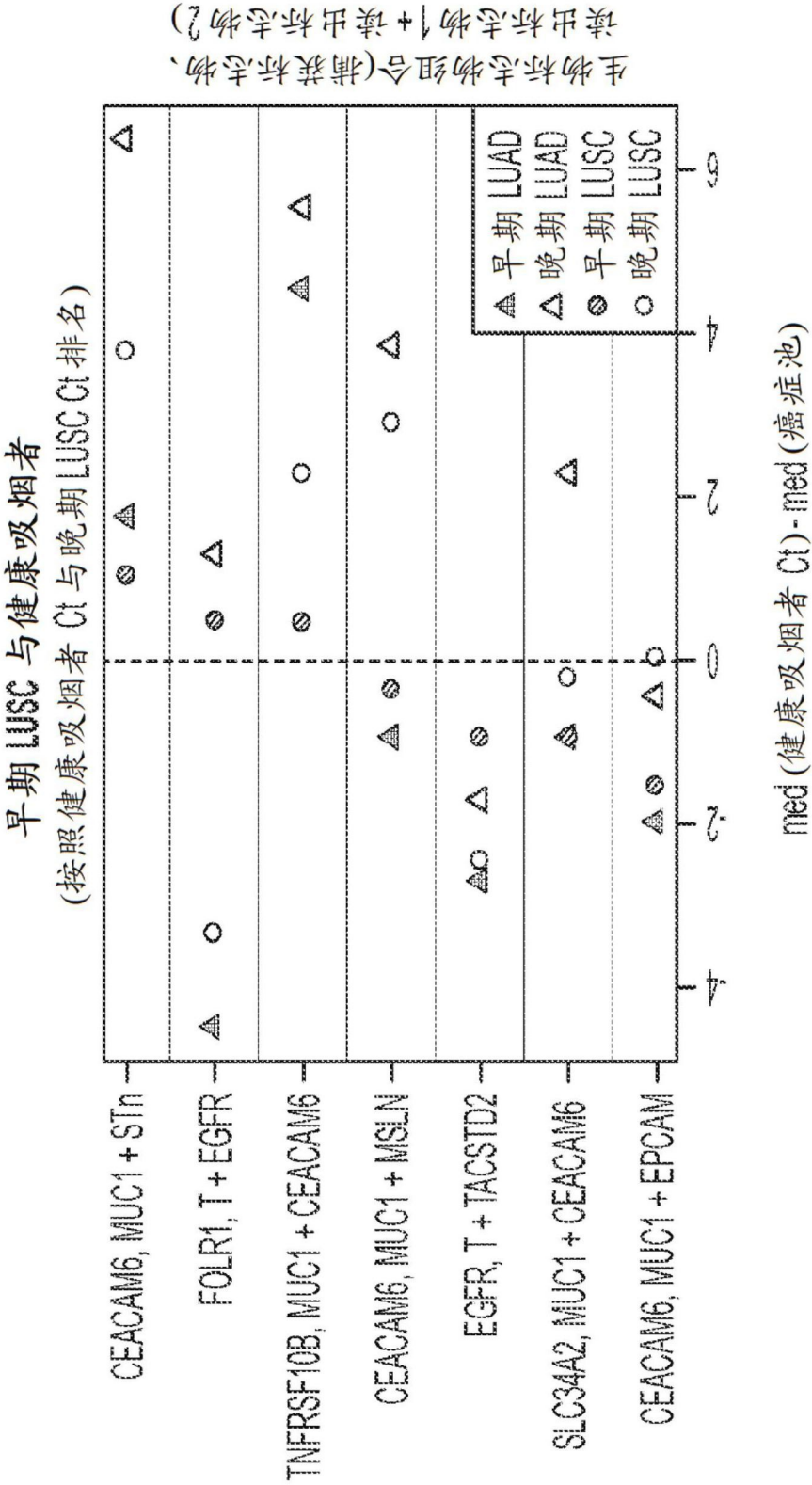


图22A

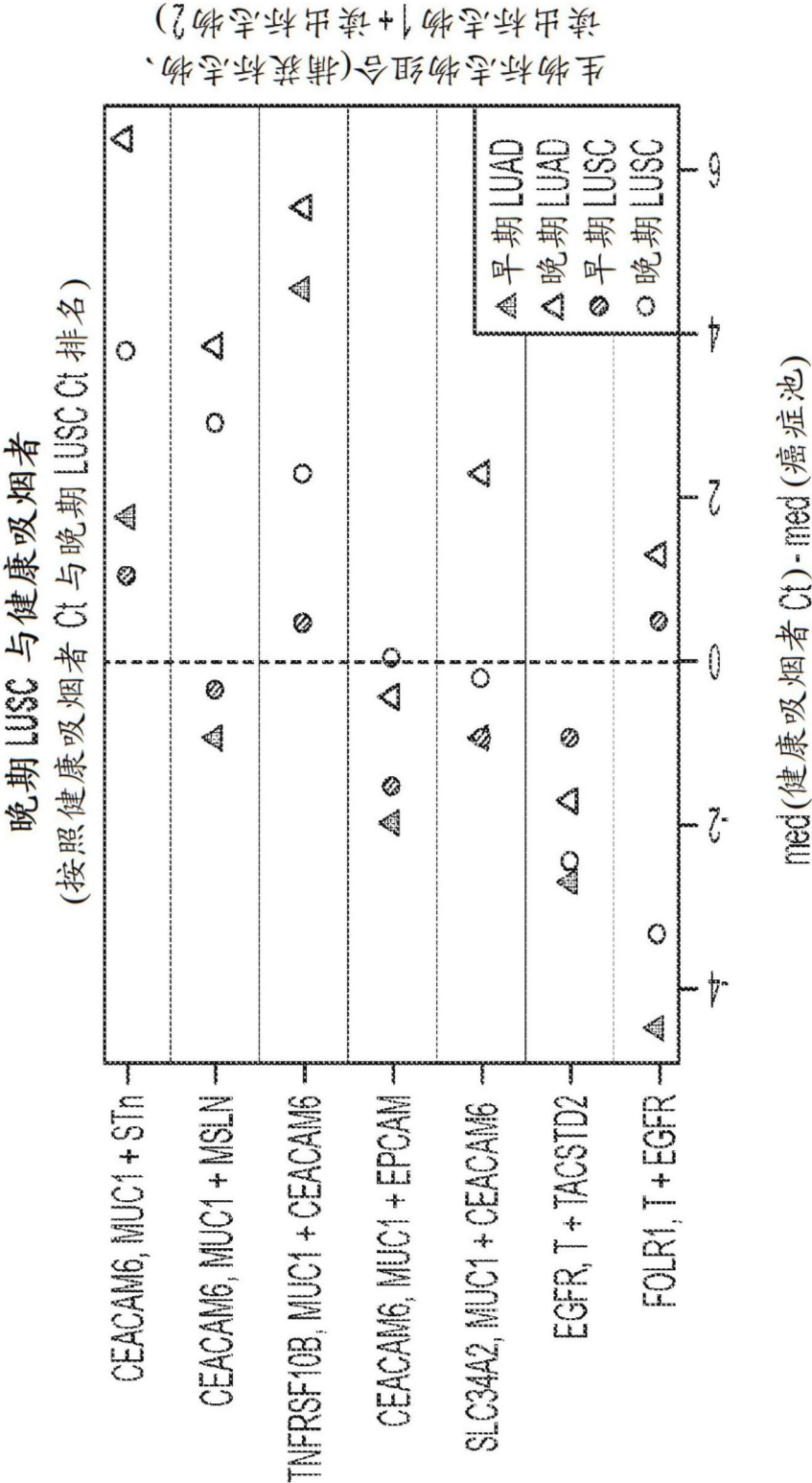


图22B

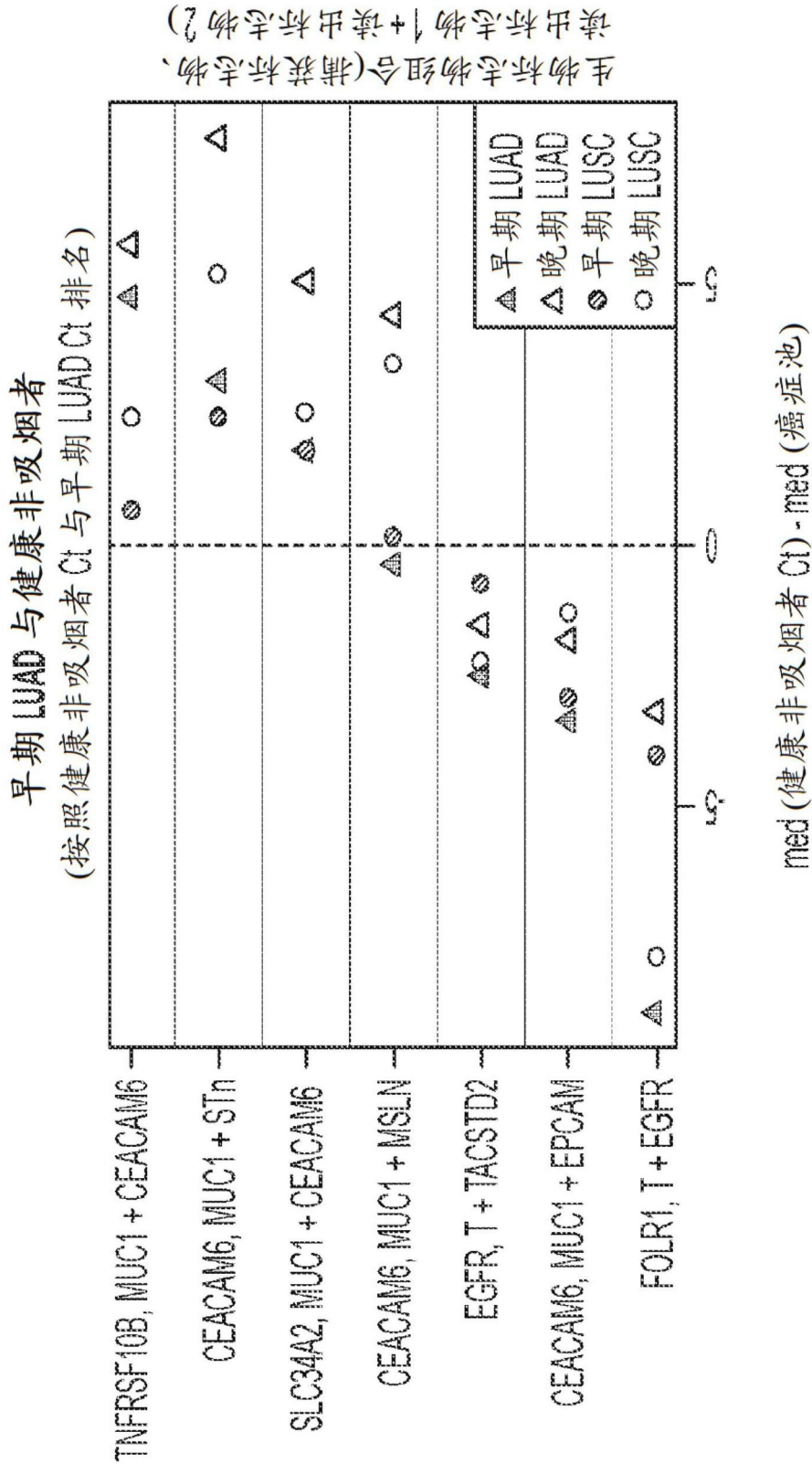


图23A

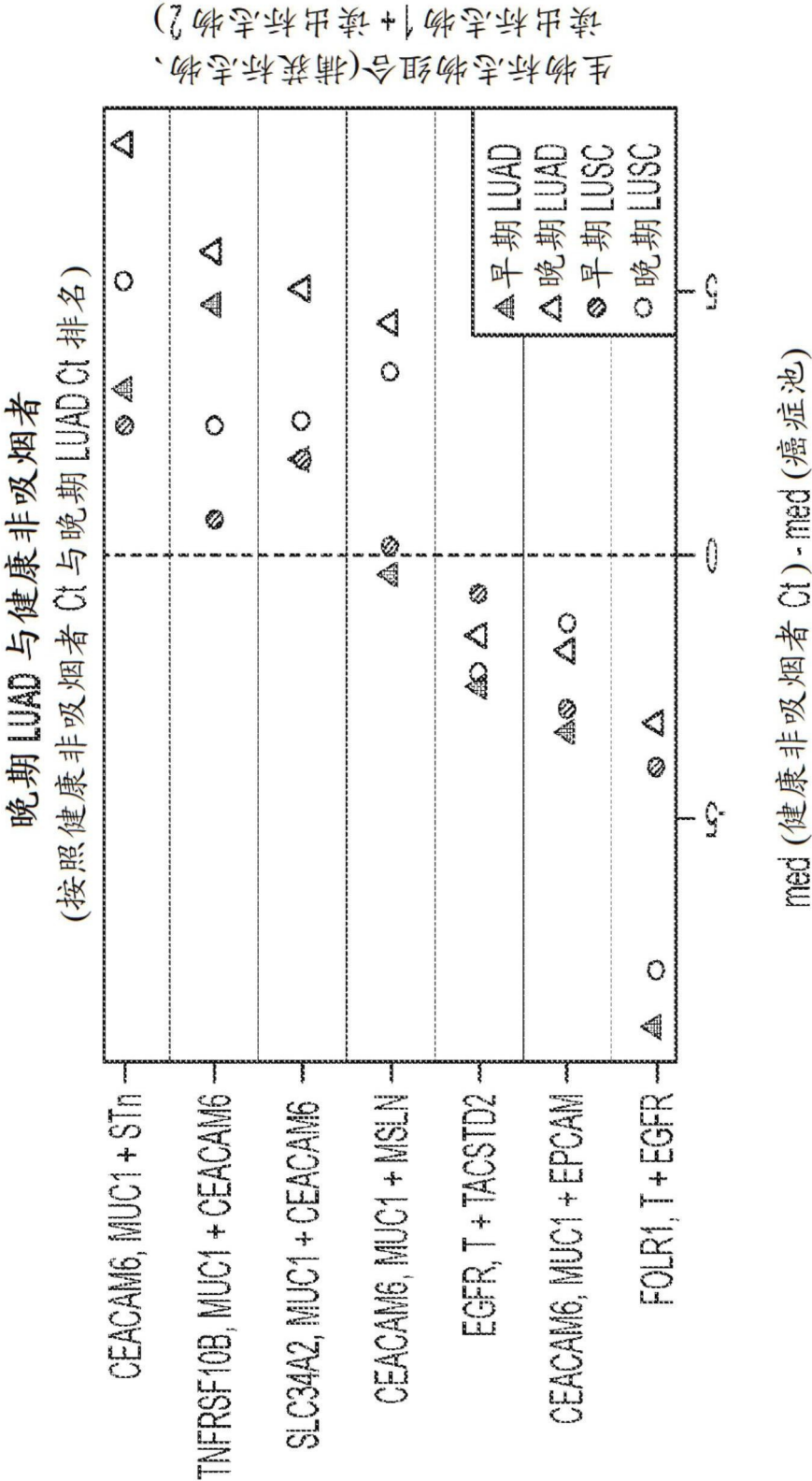


图23B

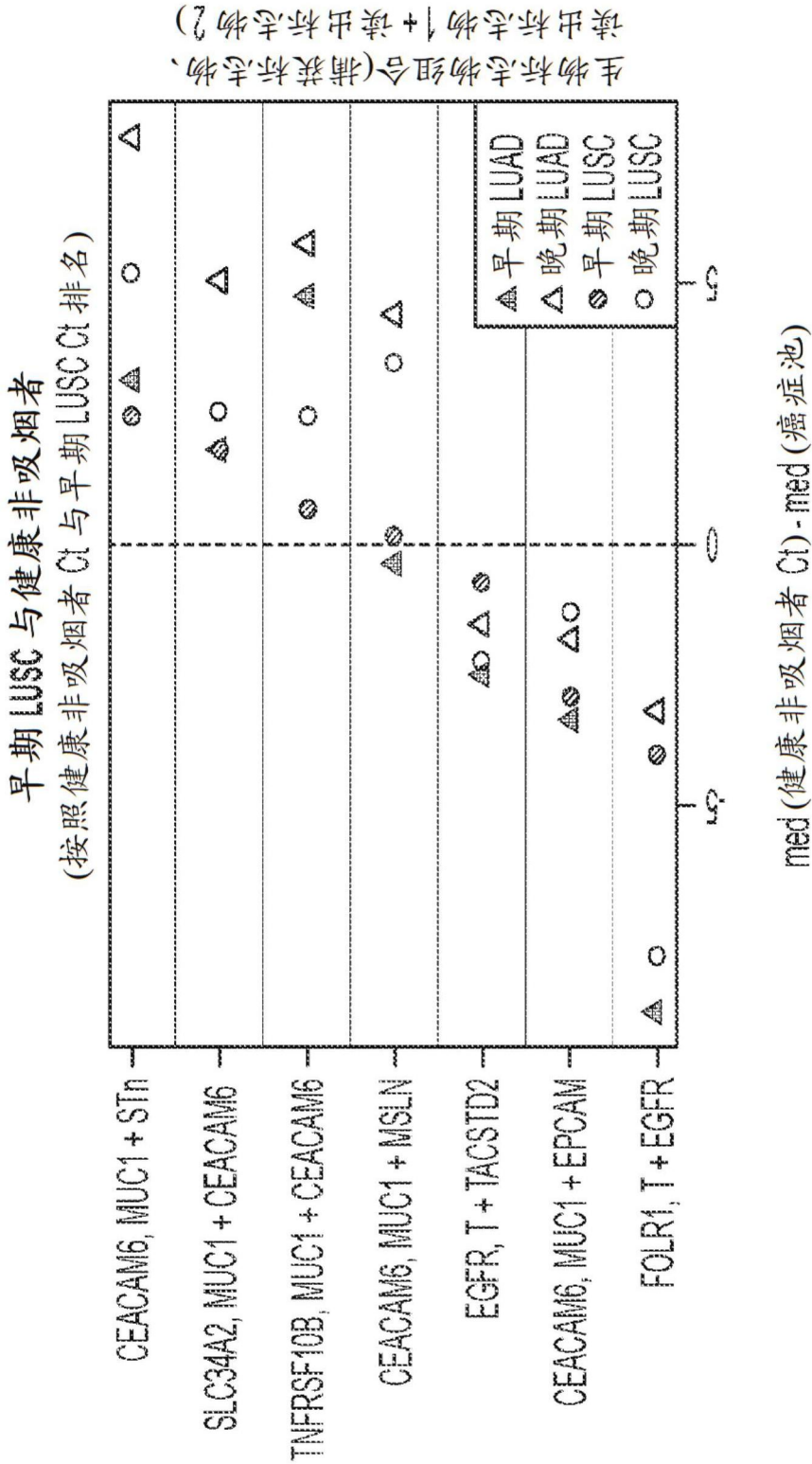


图24A

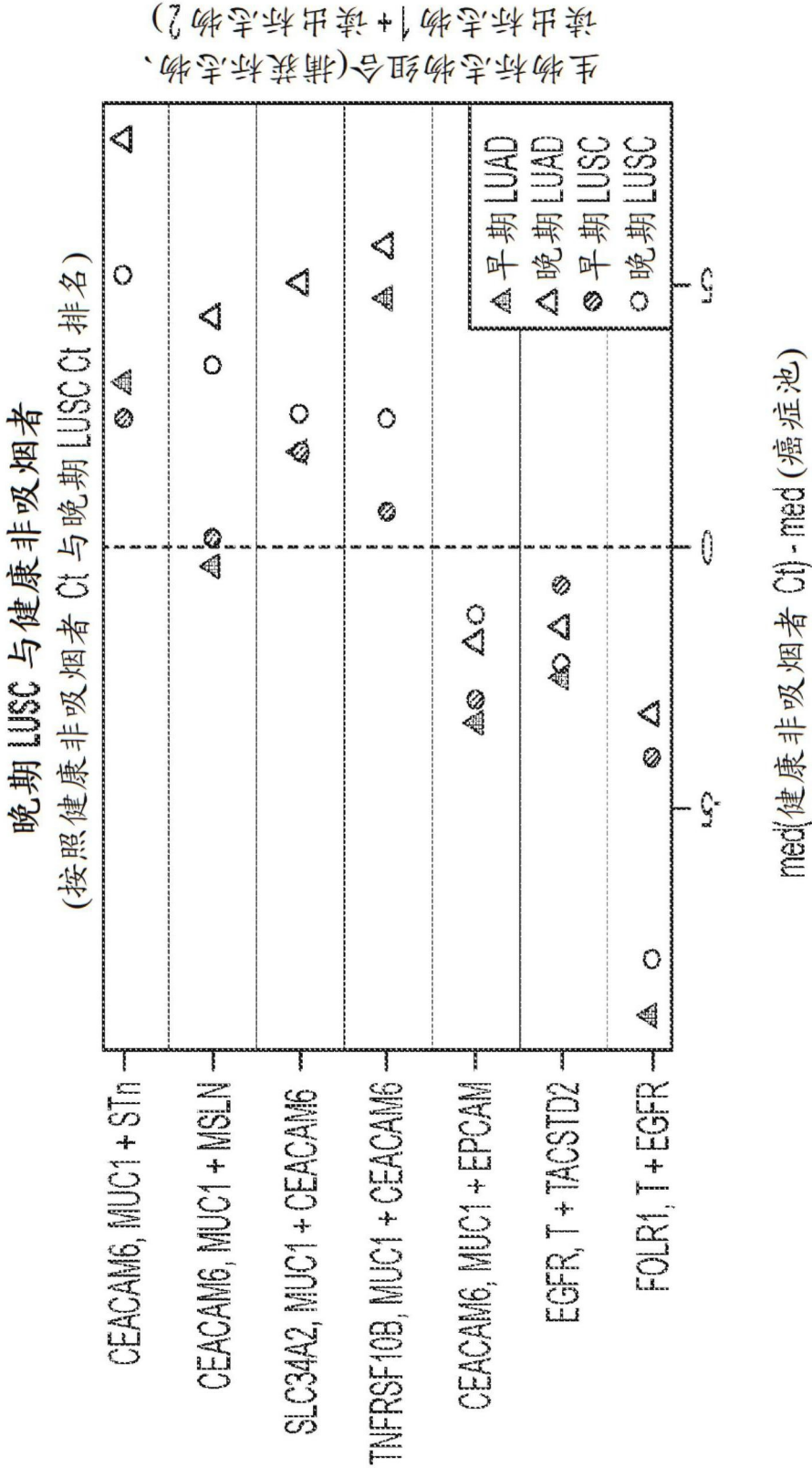


图24B