



(12) 发明专利申请

(10) 申请公布号 CN 113365639 A

(43) 申请公布日 2021.09.07

(21) 申请号 201980055983.X

(22) 申请日 2019.06.27

(30) 优先权数据

62/690,441 2018.06.27 US

62/757,045 2018.11.07 US

62/814,200 2019.03.05 US

62/855,311 2019.05.31 US

(85) PCT国际申请进入国家阶段日

2021.02.25

(86) PCT国际申请的申请数据

PCT/US2019/039521 2019.06.27

(87) PCT国际申请的公布数据

W02020/006242 EN 2020.01.02

(71) 申请人 摩登纳特斯有限公司

地址 美国马萨诸塞州

(72) 发明人 钟山 本杰明·布雷顿

伊恩·麦克法迪恩

克里斯滕·霍普森

文森特·卢茨科 迈亚·加恩纳斯

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

代理人 张福誉 刘振佳

(51) Int.Cl.

A61K 31/7105 (2006.01)

A61K 31/711 (2006.01)

A61K 31/7115 (2006.01)

A61K 39/00 (2006.01)

C12Q 1/6886 (2006.01)

G01N 33/574 (2006.01)

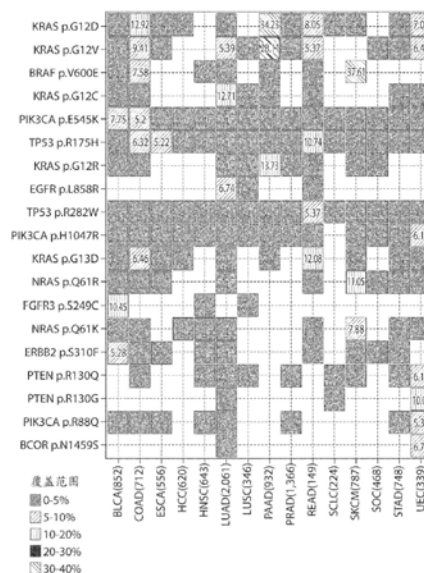
权利要求书7页 说明书79页 附图10页

(54) 发明名称

个性化癌症疫苗表位选择

(57) 摘要

本公开涉及优化的癌症疫苗,以及制备所述疫苗的方法、使用所述疫苗的方法和包含所述疫苗的组合物。所述癌症疫苗包含个性化癌症抗原,或癌症热点抗原的部分。另外,本公开涉及用于选择用于包含在优化的癌症疫苗中的核酸的计算机化系统。



1. 一种核酸癌症疫苗,所述核酸癌症疫苗包含:
一种或多种核酸,所述一种或多种核酸各自具有一个或多个编码3-130个肽表位的开放阅读框,
其中所述肽表位中的每一个均是个性化癌症抗原的部分或癌症热点抗原的部分,并且其中所述肽表位中的至少两个具有不同的长度。
2. 如权利要求1所述的核酸癌症疫苗,其中所述肽表位中的1-34个是癌症热点抗原的部分。
3. 如权利要求1所述的核酸癌症疫苗,其中所述肽表位中的5-34个是癌症热点抗原的部分。
4. 如权利要求1-3中任一项所述的核酸癌症疫苗,其中所述癌症热点抗原包含KRAS G12突变或KRAS G13突变或这两种突变。
5. 如权利要求1-4中任一项所述的核酸癌症疫苗,其中所述癌症热点新抗原的所述部分包含以下突变中的至少一者:KRAS G12突变、KRAS G13突变、NRAS Q61突变、BRAF V600突变、PIK3CA R88突变、PIK3CA E545突变、PIK3CA H1047突变、TP53 R175突变、TP53 R282突变、EGFR L858突变、FGFR3 S249突变、ERBB2 S310突变、PTEN R130突变和BCOR N1459突变。
6. 如权利要求1所述的核酸癌症疫苗,其中每个肽表位的长度被确定为使得所述核酸癌症疫苗的抗癌功效具有基于所述一种或多种核酸的长度的最大T细胞活化值。
7. 如权利要求1所述的核酸癌症疫苗,其中每个肽表位的长度被确定为使得所述核酸癌症疫苗的抗癌功效具有基于所述一种或多种核酸的长度的最大存活值。
8. 如权利要求1-7中任一项所述的核酸癌症疫苗,其中任何肽表位的最小长度为8-13个氨基酸。
9. 如权利要求1-8中任一项所述的核酸癌症疫苗,其中任何肽表位的最大长度为31-35个氨基酸。
10. 如权利要求1-9中任一项所述的核酸癌症疫苗,其中所述癌症疫苗是DNA癌症疫苗。
11. 如权利要求1-9中任一项所述的核酸癌症疫苗,其中所述癌症疫苗是RNA癌症疫苗。
12. 如权利要求11所述的核酸癌症疫苗,其中所述癌症疫苗是mRNA癌症疫苗,并且其中所述一种或多种核酸是mRNA。
13. 如权利要求12所述的核酸癌症疫苗,其中所述一种或多种mRNA各自包含5' UTR和/或3' UTR。
14. 如权利要求12或权利要求13所述的核酸癌症疫苗,其中所述一种或多种mRNA各自包含聚腺苷酸尾部。
15. 如权利要求14所述的核酸癌症疫苗,其中所述聚腺苷酸尾部包含约100个核苷酸。
16. 如权利要求12-15中任一项所述的核酸癌症疫苗,其中所述一种或多种mRNA各自包含帽结构或经修饰的帽结构。
17. 如权利要求16所述的核酸癌症疫苗,其中所述帽结构或所述经修饰的帽结构是5'帽结构、5'帽0结构、5'帽1结构或5'帽2结构。
18. 如权利要求12-17中任一项所述的核酸癌症疫苗,其中所述一种或多种mRNA包含至少一种化学修饰。
19. 如权利要求18所述的核酸癌症疫苗,其中所述化学修饰选自以下组成的组:假尿

苷、N1-甲基假尿苷、N1-乙基假尿苷、2-硫代尿苷、4'-硫代尿苷、5-甲基胞嘧啶、2-硫代-1-甲基-1-脱氮-假尿苷、2-硫代-1-甲基-假尿苷、2-硫代-5-氮杂-尿苷、2-硫代-二氢假尿苷、2-硫代-二氢尿苷、2-硫代-假尿苷、4-甲氧基-2-硫代-假尿苷、4-甲氧基-假尿苷、4-硫代-1-甲基-假尿苷、4-硫代-假尿苷、5-氮杂-尿苷、二氢假尿苷、5-甲基尿苷、5-甲基尿苷、5-甲氧基尿苷和2'-O-甲基尿苷。

20. 如权利要求18或权利要求19所述的核酸癌症疫苗,其中所述一种或多种mRNA是经完全修饰的。

21. 如权利要求1-20中任一项所述的核酸癌症疫苗,其中所述一种或多种核酸编码34个肽表位、5-10个肽表位、10-20个肽表位、20-30个肽表位、30-40个肽表位、40-50个肽表位、50-60个肽表位、60-70个肽表位、70-80个肽表位、80-90个肽表位、90-100个肽表位、100-110个肽表位、110-120个肽表位,或120-130个肽表位。

22. 如权利要求1-21中任一项所述的核酸癌症疫苗,其中所述肽表位中的每一个均由单独的开放阅读框编码。

23. 如权利要求1-22中任一项所述的核酸癌症疫苗,其中所述肽表位呈由5-130个肽表位组成的多联癌症抗原的形式。

24. 如权利要求1-23中任一项所述的核酸癌症疫苗,其中满足以下条件中的一个或多个:

- a) 所述5-130个肽表位被切割敏感性位点穿插;和/或
- b) 每个肽表位在不接头的情况下彼此直接联接;和/或
- c) 每个肽表位通过单一氨基酸接头彼此联接;和/或
- d) 每个肽表位通过短肽接头彼此联接;和/或
- e) 每个肽表位包含8-35个氨基酸,并且包含一个或多个SNP突变;和/或
- f) 每个肽表位包含8-35个氨基酸,并且包含引起独特表达的肽序列的突变;和/或
- g) 所述肽表位中没有一个对来自受试者的II类MHC分子具有最高亲和力;和/或
- h) 编码所述肽表位的所述核酸被排列成使得所述肽表位有序化以使假表位减到最少;和/或
- i) I类MHC分子肽表位与II类MHC分子肽表位的比率为至少1:1、2:1、3:1、4:1或5:1;和/或
- j) 不存在II类MHC分子肽表位;和/或
- k) 所述肽表位的至少30%对来自受试者的I类MHC分子和/或II类MHC分子具有最高亲和力;和/或
- l) 对于HLA-A、HLA-B和/或DRB1,所述肽表位的至少50%具有高于0.5%的概率百分比排名;和/或
- m) 其中所述开放阅读框编码34个肽表位,并且其中29个表位是I类MHC表位,且5个表位是II类MHC表位或者I类和II类MHC表位。

25. 如权利要求1-24中任一项所述的核酸癌症疫苗,其中所述肽表位中的至少一个是预测的T细胞反应性表位。

26. 如权利要求1-25中任一项所述的核酸癌症疫苗,其中所述肽表位中的至少一个是预测的B细胞反应性表位。

27. 如权利要求1-26中任一项所述的核酸癌症疫苗,其中所述肽表位包含预测的T细胞反应性表位和预测的B细胞反应性表位的组合。

28. 如权利要求1-27中任一项所述的核酸癌症疫苗,其中所述肽表位是预测的T细胞反应性表位和/或预测的B细胞反应性表位。

29. 如权利要求1-26中任一项所述的核酸癌症疫苗,其中所述肽表位中的至少一个是预测的新表位。

30. 如权利要求1-27中任一项所述的核酸癌症疫苗,其中至少一种核酸具有至少编码一种或多种传统癌症抗原或者一种或多种癌症/睾丸抗原的片段的开放阅读框。

31. 如权利要求1-30中任一项所述的核酸癌症疫苗,其中每种核酸被配制在脂质纳米粒子中。

32. 如权利要求31所述的核酸癌症疫苗,其中每种核酸被配制在不同的脂质纳米粒子中。

33. 如权利要求31所述的核酸癌症疫苗,其中每种核酸被配制在相同的脂质纳米粒子中。

34. 如权利要求1-33中任一项所述的核酸癌症疫苗,其中所述一种或多种核酸的总长度编码50-100个氨基酸、100-200个氨基酸、200-300个氨基酸、300-400个氨基酸、400-500个氨基酸、500-600个氨基酸、600-700个氨基酸、700-800个氨基酸、800-900个氨基酸、900-1000个氨基酸、1000-1100个氨基酸,或1100-1200个氨基酸的总蛋白质长度。

35. 如权利要求1-34中任一项所述的核酸癌症疫苗,其中所述抗癌功效至少部分地基于选自由以下组成的组的一个或多个因素来计算:基因表达、RNA Seq、转录物丰度、DNA等位基因频率、氨基酸保守性、生理化学相似性、致癌基因、与特定HLA等位基因的预测结合亲和力和克隆性、结合效率以及在indel中的存在。

36. 如权利要求35所述的核酸癌症疫苗,其中所述一个或多个因素被输入到统计模型中。

37. 一种核酸癌症疫苗,所述核酸癌症疫苗包含:

一种或多种核酸,所述一种或多种核酸各自具有一个或多个编码5-130个肽表位的开放阅读框,

其中所述肽表位中的每一个均是个性化癌症抗原的部分或癌症热点抗原的部分,并且其中每个肽表位具有相等的长度。

38. 一种制备癌症疫苗的方法,所述方法包括:

a) 在1-34个癌症热点之间进行鉴定;

b) 针对患者在5-130种个性化癌症抗原之间进行鉴定;

c) 针对所述5-130种个性化癌症抗原中的每一种,确定至少两个肽表位的抗肿瘤功效;以及

d) 制备癌症疫苗,其中所述癌症疫苗的总抗癌功效在给定的所述癌症疫苗的总长度下最大化,并且其中所述疫苗包含1-34种癌症热点新抗原的部分。

39. 一种用于治疗患有癌症的患者的方法,所述方法包括:

a) 分析源自患者的样品以便鉴定出一种或多种个性化癌症抗原;

b) 针对所述经鉴定的个性化癌症抗原中的每一种,确定至少两个肽表位的抗肿瘤功

效；

c) 制备癌症疫苗，其中所述癌症疫苗的总抗癌功效在给定的所述癌症疫苗的总长度下最大化，其中所述癌症疫苗进一步包含1-34种癌症热点抗原的部分；以及

d) 向所述患者施用所述癌症疫苗。

40. 如权利要求38或权利要求39所述的方法，其中所述1-34种癌症热点新抗原的部分包含以下突变中的至少一者：KRAS G12突变、KRAS G13突变、NRAS Q61突变、BRAF V600突变、PIK3CA R88突变、PIK3CA E545突变、PIK3CA H1047突变、TP53 R175突变、TP53 R282突变、EGFR L858突变、FGFR3 S249突变、ERBB2 S310突变、PTEN R130突变和BCOR N1459突变。

41. 如权利要求38或权利要求39所述的方法，其中所述1-34种癌症热点新抗原的部分包含KRAS G12突变或KRAS G13突变或这两种突变。

42. 如权利要求38或权利要求39所述的方法，其中所述癌症疫苗是包含一种或多种核酸的核酸癌症疫苗，所述一种或多种核酸各自具有一个或多个开放阅读框。

43. 如权利要求38-42中任一项所述的方法，其中所述癌症疫苗是DNA癌症疫苗。

44. 如权利要求38-43中任一项所述的方法，其中所述癌症疫苗是RNA癌症疫苗。

45. 如权利要求44所述的方法，其中所述癌症疫苗是mRNA癌症疫苗。

46. 如权利要求38或权利要求39所述的方法，其中所述癌症疫苗是肽癌症疫苗。

47. 如权利要求39-46中任一项所述的方法，其中所述癌症疫苗以足以向所述受试者递送介于0.02-1.0mg之间的所述癌症疫苗的剂量水平施用。

48. 如权利要求47所述的方法，其中将所述癌症疫苗施用给所述受试者两次、三次、四次或更多次。

49. 如权利要求39-48中任一项所述的方法，其中所述癌症疫苗通过真皮内、肌内、血管内、肿瘤内和/或皮下施用来施用。

50. 如权利要求49所述的方法，其中所述癌症疫苗通过肌内施用来施用。

51. 如权利要求39-50中任一项所述的方法，其中所述癌症选自由以下组成的组：非小细胞肺癌 (NSCLC)、小细胞肺癌、黑素瘤、膀胱尿路上皮癌、HPV阴性头颈鳞状细胞癌 (HNSCC)、有高度微卫星不稳定性 (MSI H) / 错配修复 (MMR) 缺陷的实体恶性肿瘤、肾癌、胃癌和高肿瘤突变负荷肿瘤。

52. 如权利要求51所述的方法，其中所述NSCLC缺乏EGFR敏化突变和/或ALK易位。

53. 如权利要求51所述的方法，其中所述有高度微卫星不稳定性 (MSI H) / 错配修复 (MMR) 缺陷的实体恶性肿瘤选自由以下组成的组：结肠直肠癌、胃腺癌、食道腺癌和子宫内膜癌。

54. 权利要求45-53中任一项所述的方法，其中所述一种或多种mRNA各自包含5' UTR和/或3' UTR。

55. 权利要求45-54中任一项所述的方法，其中所述一种或多种mRNA各自包含聚腺苷酸尾部。

56. 如权利要求55所述的方法，其中所述聚腺苷酸尾部包含约100个核苷酸。

57. 如权利要求45-56中任一项所述的方法，其中所述一种或多种mRNA各自包含帽结构或经修饰的帽结构。

58. 如权利要求57所述的核酸癌症疫苗，其中所述帽结构或所述经修饰的帽结构是5'

帽结构、5'帽0结构、5'帽1结构或5'帽2结构。

59. 如权利要求45-58中任一项所述的方法,其中所述一种或多种mRNA包含至少一种化学修饰。

60. 如权利要求59所述的方法,其中所述化学修饰选自以下组成的组:假尿苷、N1-甲基假尿苷、N1-乙基假尿苷、2-硫代尿苷、4'-硫代尿苷、5-甲基胞嘧啶、2-硫代-1-甲基-1-脱氮-假尿苷、2-硫代-1-甲基-假尿苷、2-硫代-5-氮杂-尿苷、2-硫代-二氢假尿苷、2-硫代-二氢尿苷、2-硫代-假尿苷、4-甲氧基-2-硫代-假尿苷、4-甲氧基-假尿苷、4-硫代-1-甲基-假尿苷、4-硫代-假尿苷、5-氮杂-尿苷、二氢假尿苷、5-甲基尿苷、5-甲基尿苷、5-甲氧基尿苷和2'-O-甲基尿苷。

61. 如权利要求59或权利要求60所述的方法,其中所述一种或多种mRNA是经完全修饰的。

62. 如权利要求42-45中任一项所述的方法,其中所述一种或多种核酸编码5-10个肽表位、10-20个肽表位、20-30个肽表位、30-40个肽表位、40-50个肽表位、50-60个肽表位、60-70个肽表位、70-80个肽表位、80-90个肽表位、90-100个肽表位、100-110个肽表位、110-120个肽表位,或120-130个肽表位。

63. 如权利要求38-62中任一项所述的方法,其中所述肽表位中的每一个均由单独的开放阅读框编码。

64. 如权利要求38-63中任一项所述的方法,其中所述肽表位呈由5-130个肽表位组成的多联癌症抗原的形式。

65. 如权利要求38-64中任一项所述的方法,其中满足以下条件中的一个或多个:

- a) 所述5-130个肽表位被切割敏感性位点穿插;和/或
- b) 每个肽表位在不接头的情况下彼此直接联接;和/或
- c) 每个肽表位通过单一氨基酸接头彼此联接;和/或
- d) 每个肽表位通过短接头彼此联接;和/或
- e) 每个肽表位包含8-35个氨基酸,并且包含一个或多个SNP突变;和/或
- f) 每个肽表位包含8-35个氨基酸,并且包含引起独特表达的肽序列的突变;和/或
- g) 所述肽表位中没有一个对来自受试者的II类MHC分子具有最高亲和力;和/或
- h) 编码所述肽表位的所述核酸被排列成使得所述肽表位有序化以使假表位减到最少;和/或
- i) I类MHC分子肽表位与II类MHC分子肽表位的比率为至少1:1、2:1、3:1、4:1或5:1;和/或
- j) 不存在II类MHC分子肽表位;和/或
- k) 所述肽表位的至少30%对来自受试者的I类MHC分子和/或II类MHC分子具有最高亲和力;和/或
- l) 对于HLA-A、HLA-B和/或DRB1,所述肽表位的至少50%具有高于0.5%的概率百分比排名;和/或
- m) 其中所述开放阅读框编码34个肽表位,并且其中29个表位是I类MHC表位,且5个表位是II类MHC表位或者I类和II类MHC表位。

66. 如权利要求38-65中任一项所述的方法,其中所述肽表位中的至少一个是预测的T

细胞反应性表位。

67. 如权利要求38-66中任一项所述的方法,其中所述肽表位中的至少一个是预测的B细胞反应性表位。

68. 如权利要求38-67中任一项所述的方法,其中所述肽表位包含预测的T细胞反应性表位和预测的B细胞反应性表位的组合。

69. 如权利要求38-67中任一项所述的方法,其中所述肽表位是预测的T细胞反应性表位和/或预测的B细胞反应性表位。

70. 如权利要求38-69中任一项所述的方法,其中所述肽表位中的至少一个是预测的新表位。

71. 如权利要求42-45或62-69中任一项所述的方法,其中至少一种核酸具有至少编码一种或多种传统癌症抗原或者一种或多种癌症/睾丸抗原的片段的开放阅读框。

72. 如权利要求42-45或62-71中任一项所述的方法,其中每种核酸被配制在脂质纳米粒子中。

73. 如权利要求72所述的方法,其中每种核酸被配制在不同的脂质纳米粒子中。

74. 如权利要求72所述的方法,其中每种核酸被配制在相同的脂质纳米粒子中。

75. 如权利要求42-45或62-74中任一项所述的方法,其中所述一种或多种核酸的总长度编码50-100个氨基酸、100-200个氨基酸、200-300个氨基酸、300-400个氨基酸、400-500个氨基酸、500-600个氨基酸、600-700个氨基酸、700-800个氨基酸、800-900个氨基酸、900-1000个氨基酸、1000-1100个氨基酸,或1100-1200个氨基酸的总蛋白质长度。

76. 如权利要求38-75中任一项所述的方法,其中所述抗癌功效至少部分地基于选自以下组成的组的一个或多个因素来计算:基因表达、RNA Seq、转录物丰度、DNA等位基因频率、氨基酸保守性、生理化学相似性、致癌基因、与特定HLA等位基因的预测结合亲和力、克隆性、结合效率以及在indel中的存在。

77. 如权利要求76所述的方法,其中所述一个或多个因素被输入到统计模型中。

78. 一种用于选择用于包含在具有最大长度的核酸癌症疫苗中的核酸的计算机化系统,所述系统包括:

通信接口,所述通信接口被配置为接收编码多个肽表位的多个核酸序列,其中所述肽表位中的每一个均是个性化癌症抗原的部分;和

至少一个计算机处理器,所述至少一个计算机处理器被编程为:

针对所述多个肽表位中的每一个,计算各自包含一个或多个肽表位中的至少一个肽表位的肽中多种核酸中的每一种的得分,其中所述核酸序列中的至少两个具有不同的长度;以及

基于所计算的得分,对多种肽中的多个核酸序列进行排名;以及

基于所述疫苗的排名和最大长度,选择用于包含在所述疫苗中的核酸序列。

79. 如权利要求78所述的计算机化系统,其中任何肽表位的最小长度为8个氨基酸。

80. 如权利要求78或权利要求79所述的计算机化系统,其中任何肽表位的最大长度为31个氨基酸。

81. 如权利要求78-80中任一项所述的计算机化系统,其中所述多种核酸编码5-10个肽表位、10-20个肽表位、20-30个肽表位、30-40个肽表位、34个表位、40-50个肽表位、50-60个

肽表位、60-70个肽表位、70-80个肽表位、80-90个肽表位、90-100个肽表位、100-110个肽表位、110-120个肽表位,或120-130个肽表位。

82. 如权利要求78-81中任一项所述的计算机化系统,其中满足以下条件中的一个或多个:

- a) 每个肽表位包含8-31个氨基酸,并且包含一个或多个SNP突变;和/或
- b) 每个肽表位包含8-31个氨基酸,并且包含引起独特表达的肽序列的突变;和/或
- c) 所述肽表位中没有一个对来自受试者的II类MHC分子具有最高亲和力;和/或
- d) I类MHC分子肽表位与II类MHC分子肽表位的比率为至少1:1、2:1、3:1、4:1或5:1;和/或
- e) 不存在II类MHC分子肽表位;
- f) 所述肽表位的至少30%对来自受试者的I类MHC分子和/或II类MHC分子具有最高亲和力;和/或
- g) 对于HLA-A、HLA-B和/或DRB1,所述肽表位的至少50%具有高于0.5%的概率百分比排名。

83. 如权利要求78-82中任一项所述的计算机化系统,其中所述肽表位中的至少一个是预测的T细胞反应性表位。

84. 如权利要求78-83中任一项所述的计算机化系统,其中所述肽表位中的至少一个是预测的B细胞反应性表位。

85. 如权利要求78-84中任一项所述的计算机化系统,其中所述肽表位包含预测的T细胞反应性表位和预测的B细胞反应性表位的组合。

86. 如权利要求78-85中任一项所述的计算机化系统,其中所述肽表位是预测的T细胞反应性表位和/或预测的B细胞反应性表位。

87. 如权利要求78-86中任一项所述的计算机化系统,其中所述肽表位中的至少一个是预测的新表位。

88. 如权利要求78-87中任一项所述的计算机化系统,其中至少一种核酸具有至少编码一种或多种传统癌症抗原或者一种或多种癌症/睾丸抗原的片段的开放阅读框。

89. 如权利要求78-88中任一项所述的计算机化系统,其中所述疫苗的总长度编码50-100个氨基酸、100-200个氨基酸、200-300个氨基酸、300-400个氨基酸、400-500个氨基酸、500-600个氨基酸、600-700个氨基酸、700-800个氨基酸、800-900个氨基酸、900-1000个氨基酸、1000-1100个氨基酸,或1100-1200个氨基酸的总蛋白质长度。

90. 如权利要求78-89中任一项所述的计算机化系统,其中所述得分至少部分地基于选自以下组成的组的一个或多个因素来计算:基因表达、RNA Seq、转录物丰度、DNA等位基因频率、氨基酸保守性、生理化学相似性、致癌基因、与特定HLA等位基因的预测结合亲和力、克隆性、结合效率以及在indel中的存在。

91. 如权利要求90所述的计算机化系统,其中所述一个或多个因素被输入到统计模型中。

个性化癌症疫苗表位选择

[0001] 相关申请

[0002] 本申请根据35 U.S.C. §119(e) 要求2018年6月27日提交的美国临时申请号62/690,441、2018年11月7日提交的美国临时申请号62/757,045、2019年3月5日提交的美国临时申请号62/814,200以及2019年5月31日提交的美国临时申请号62/855,311的优先权,其中的每一篇均以引用方式整体并入本文。

背景技术

[0003] 最近的肿瘤进化理论集中于三个步骤,包括应激诱导的基因组不稳定性、种群多样性或异质性以及基因组介导的宏观进化。该理论解释了大多数已知的分子机制可促成癌症的原因,但大多数临床病例没有单一主要机制。然而,共同的机制表明,癌症疫苗可提供治疗癌症的通用解决方案。

[0004] 癌症疫苗包括预防性或防治性疫苗,其意图预防健康人群罹患癌症;以及治疗性疫苗,其意图通过加强身体对癌症的天然防御来治疗现有癌症。癌症预防性疫苗可例如针对导致或促使罹患癌症的感染因子(infectious agent)以防止感染性疾病导致癌症。**Gardasil®**和**Cervarix®**是可预防HPV感染和由此导致的癌症的可商购获得的防治性疫苗的两个实例。其他预防性癌症疫苗可针对经预测会使个体在未来罹患癌症的可能性增加的宿主蛋白或片段。

[0005] 许多商业疫苗或开发中的疫苗是基于整个微生物、蛋白抗原、肽或多糖以及它们的组合。某些开发中的疫苗还基于核酸疫苗(例如,脱氧核糖核酸(DNA)疫苗或核糖核酸(RNA)疫苗)。此类核酸疫苗一般未被优选成对于其尺寸或长度具有最大功效。

发明内容

[0006] 发明内容

[0007] 本文提供了一种核酸(例如,核糖核酸(RNA))癌症疫苗,所述癌症疫苗在给定长度下具有最大化抗癌功效,并且包含一种或多种可指导身体细胞机制产生几乎任何感兴趣的癌蛋白或其片段的核酸。在一些实施方案中,本公开还提供了制备在给定长度下具有最大化抗癌功效的核酸癌症疫苗的方法。在一些实施方案中,本公开还提供了用在给定长度下具有最大化抗癌功效的癌症疫苗治疗患有癌症的患者的方法。另外,在某些实施方案中,本公开提供了一种用于创建在给定长度下具有最大化抗癌功效的核酸癌症疫苗的计算机化系统。

[0008] 在一方面,本公开提供了一种核酸癌症疫苗,所述核酸癌症疫苗包含:一种或多种核酸,所述一种或多种核酸各自具有一个或多个编码5-130个肽表位的开放阅读框,其中肽表位中的每一个均是个性化癌症抗原的部分,并且其中至少两个肽表位具有不同的长度。在另一方面,本公开提供了一种核酸癌症疫苗,所述核酸癌症疫苗包含:一种或多种核酸,所述一种或多种核酸各自具有一个或多个编码5-130个、20-40个、30-35个或34个肽表位的开放阅读框,其中肽表位中的每一个均是个性化癌症抗原的部分,并且其中肽表位中的每

一个均具有不同的长度。在另一方面,本公开提供了一种核酸癌症疫苗,所述核酸癌症疫苗包含:一种或多种核酸,所述一种或多种核酸各自具有一个或多个编码5-130个、20-40个、30-35个或34个肽表位的开放阅读框,其中肽表位中的每一个均是个性化癌症抗原的部分,并且其中肽表位中的每一个均具有相等的长度。在一些实施方案中,所述癌症疫苗组合物包含一种或多种mRNA,所述一种或多种mRNA各自具有一个或多个编码34个肽表位的开放阅读框,并且其中29个表位是I类MHC表位,且5个表位是II类MHC表位或者I类和II类MHC表位。

[0009] 在一些实施方案中,每个肽表位的长度被确定为使得所述核酸癌症疫苗的抗癌功效在给定的所述一种或多种核酸的总长度下最大化。在一些实施方案中,任何肽表位的最小长度为8个氨基酸。在一些实施方案中,任何肽表位的最大长度为31个氨基酸。在一些实施方案中,任何或所有肽表位的最小长度为13个氨基酸。在一些实施方案中,任何或所有肽表位的最大长度为35个氨基酸。在一些实施方案中,任何或所有肽表位的长度为25个氨基酸。

[0010] 在一些实施方案中,所述癌症疫苗是DNA癌症疫苗。在一些实施方案中,所述癌症疫苗是RNA癌症疫苗。在一些实施方案中,所述癌症疫苗是mRNA癌症疫苗,并且所述一种或多种核酸是mRNA。在一些实施方案中,所述一种或多种mRNA各自包含5' UTR和/或3' UTR。在一些实施方案中,所述一种或多种mRNA各自包含聚腺苷酸尾部。在一些实施方案中,所述聚腺苷酸尾部包含约100个核苷酸。在一些实施方案中,所述一种或多种mRNA各自包含帽结构或经修饰的帽结构。在一些实施方案中,所述帽结构或所述经修饰的帽结构是5' 帽结构、5' 帽0结构、5' 帽1结构或5' 帽2结构。

[0011] 在一些实施方案中,所述一种或多种mRNA包含至少一种化学修饰。在某些实施方案中,所述化学修饰选自由以下组成的组:假尿苷、N1-甲基假尿苷、N1-乙基假尿苷、2-硫代尿苷、4'-硫代尿苷、5-甲基胞嘧啶、2-硫代-1-甲基-1-脱氮-假尿苷、2-硫代-1-甲基-假尿苷、2-硫代-5-氮杂-尿苷、2-硫代-二氢假尿苷、2-硫代-二氢尿苷、2-硫代-假尿苷、4-甲氧基-2-硫代-假尿苷、4-甲氧基-假尿苷、4-硫代-1-甲基-假尿苷、4-硫代-假尿苷、5-氮杂-尿苷、二氢假尿苷、5-甲基尿苷、5-甲氧基尿苷和2'-O-甲基尿苷。在一些实施方案中,所述一种或多种mRNA是经完全修饰的。

[0012] 在一些实施方案中,所述一种或多种核酸编码3-10个肽表位、5-10个肽表位、10-20个肽表位、20-30个肽表位、30-40个肽表位、40-50个肽表位、50-60个肽表位、60-70个肽表位、70-80个肽表位、80-90个肽表位、90-100个肽表位、100-110个肽表位、110-120个肽表位,或120-130个肽表位。在一些实施方案中,肽表位中的每一个均由单独的开放阅读框编码。在一些实施方案中,所述肽表位呈由3-130个肽表位组成的多联(concatemeric)癌症抗原的形式。在一些实施方案中,所述癌症疫苗组合物包含一种mRNA,所述mRNA具有一个编码15个肽表位的开放阅读框。

[0013] 在一些实施方案中,满足以下条件中的一个或多个:a)所述3-130个肽表位被切割敏感性位点(例如接头,诸如包含切割敏感性位点或作为相邻表位的一部分的切割敏感性位点的肽接头)穿插;和/或b)每个肽表位在无接头的情况下彼此直接联接(linked);和/或c)每个肽表位通过单一氨基酸接头彼此联接;和/或d)每个肽表位通过短接头彼此联接;和/或e)每个肽表位包含8-31个氨基酸,并且包含一个或多个SNP突变;和/或f)每个肽表位包含8-31个氨基酸,并且包含引起独特表达的肽序列的突变;和/或g)所述肽表位的至少

30%对来自受试者的I类MHC分子具有最高亲和力;和/或h)所述肽表位的至少30%对来自受试者的II类MHC分子具有最高亲和力;和/或i)所述肽表位中没有一个对来自受试者的II类MHC分子具有最高亲和力;和/或j)所述肽表位的至少50%对HLA-A、HLA-B和/或DRB1具有 $IC_{50} < 500\text{nM}$ 的预测结合亲和力;和/或k)编码所述肽表位的核酸被排列成使得所述肽表位有序化以使假表位减到最少;和/或l) I类MHC分子肽表位与II类MHC分子肽表位的比率为至少1:1、2:1、3:1、4:1或5:1;和/或m)不存在II类MHC分子肽表位。在其他实施方案中,所述肽表位的至少30%对来自受试者的I类MHC分子和/或II类MHC分子具有最高亲和力。在其他实施方案中,对于HLA-A、HLA-B和/或DRB1,所述肽表位的至少50%具有高于0.5%的概率百分比排名。概率百分比排名提供了用于确定强结合物的阈值,并且是频率分布中等于或低于它的得分的百分比的计算结果。

[0014] 在一些实施方案中,所述肽表位中的至少一个是预测的T细胞反应性表位。在某些实施方案中,所述肽表位中的至少一个是预测的B细胞反应性表位。在一些实施方案中,所述肽表位包含预测的T细胞反应性表位和预测的B细胞反应性表位的组合。在一些实施方案中,所述肽表位是预测的T细胞反应性表位和/或预测的B细胞反应性表位。在一些实施方案中,所述肽表位中的至少一个是预测的新表位。在某些实施方案中,至少一种核酸具有至少编码一种或多种传统癌症抗原或者一种或多种癌症/睾丸抗原的片段的开放阅读框。

[0015] 在一些实施方案中,每种核酸被配制在脂质纳米粒子中。在一些实施方案中,每种核酸被配制在不同的脂质纳米粒子中。在一些实施方案中,每种核酸被配制在相同的脂质纳米粒子中。

[0016] 在一些实施方案中,所述一种或多种核酸的总长度编码50-100个氨基酸、100-200个氨基酸、200-300个氨基酸、300-400个氨基酸、400-500个氨基酸、500-600个氨基酸、600-700个氨基酸、700-800个氨基酸、800-900个氨基酸、900-1000个氨基酸、1000-1100个氨基酸,或1100-1200个氨基酸的总蛋白质长度。

[0017] 在一些实施方案中,抗癌功效至少部分地基于一个或多个选自由以下组成的组的因素来计算:基因表达、RNA Seq、转录物丰度、DNA等位基因频率、氨基酸保守性、生理化学相似性、致癌基因、与特定HLA等位基因的预测结合亲和力、克隆性、结合效率以及在indel中的存在。在一些实施方案中,所述一个或多个因素被输入到统计模型(例如,回归模型(诸如线性回归模型、逻辑回归模型、广义线性模型等)、广义线性模型(诸如逻辑回归模型、机率单位回归模型等)、随机森林回归模型、神经网络、支持向量机、高斯混合模型、分层贝叶斯模型和/或任何其他适合的统计模型)中。

[0018] 在另一方面,本公开提供了一种制备癌症疫苗的方法,所述方法包括:a)针对患者在3-130种个性化癌症抗原之间进行鉴定;b)针对该3-130种个性化癌症抗原中的每一种,确定至少两个肽表位的抗肿瘤功效;以及c)制备癌症疫苗,其中所述癌症疫苗的总抗癌功效在给定的所述癌症疫苗的总长度下最大化。

[0019] 在另一方面,本公开提供了一种用于治疗患有癌症的患者的方法,所述方法包括:a)分析源自患者的样品以便鉴定出一种或多种个性化癌症抗原;b)针对经鉴定的个性化癌症抗原中的每一种,确定至少两个肽表位的抗肿瘤功效;c)制备癌症疫苗,其中所述癌症疫苗的总抗癌功效在给定的所述癌症疫苗的总长度下最大化;以及d)向患者施用所述癌症疫苗。任选地,本文所述的任何方法可包括制造癌症疫苗。

[0020] 在一些实施方案中,所述癌症疫苗是包含一种或多种核酸的核酸癌症疫苗,所述一种或多种核酸各自具有一个或多个开放阅读框。在一些实施方案中,所述癌症疫苗是DNA癌症疫苗。在一些实施方案中,所述癌症疫苗是RNA癌症疫苗。在一些实施方案中,所述癌症疫苗是mRNA癌症疫苗。在一些实施方案中,所述癌症疫苗是肽癌症疫苗。

[0021] 在一些实施方案中,所述癌症疫苗以足以向所述受试者递送介于0.02-1.0mg之间的所述癌症疫苗的剂量水平施用。在一些实施方案中,将所述癌症疫苗施用给所述受试者两次、三次、四次或更多次。在一些实施方案中,所述癌症疫苗通过真皮内、肌内、血管内、肿瘤内和/或皮下施用来施用。在一些实施方案中,所述癌症疫苗通过肌内施用来施用。

[0022] 在某些实施方案中,本文所述的方法和组合物可与任何类型的癌症一起使用或者用于任何类型的癌症。在一些实施方案中,所述癌症选自以下组成的组:非小细胞肺癌(NSCLC)、小细胞肺癌、黑素瘤、膀胱尿路上皮癌、HPV阴性头颈鳞状细胞癌(HNSCC)、有高度微卫星不稳定性(MSI H)/错配修复(MMR)缺陷的实体恶性肿瘤、肾癌、胃癌和高肿瘤突变负荷肿瘤。在一些实施方案中,NSCLC缺乏EGFR敏化突变和/或ALK易位。在一些实施方案中,有高度微卫星不稳定性(MSI H)/错配修复(MMR)缺陷的实体恶性肿瘤选自以下组成的组:结肠直肠癌、胃腺癌、食道腺癌和子宫内膜癌。在一些实施方案中,所述癌症是黑素瘤、膀胱癌、HPV阴性HNSCC、NSCLC、SCLC、高MSI肿瘤或高TMB(肿瘤突变负荷)癌症中的任一种。

[0023] 在某些实施方案中,所述一种或多种mRNA各自包含5' UTR和/或3' UTR。在一些实施方案中,所述一种或多种mRNA各自包含聚腺苷酸尾部。在一些实施方案中,所述聚腺苷酸尾部包含约100个核苷酸。在一些实施方案中,所述一种或多种mRNA各自包含帽结构或经修饰的帽结构。在一些实施方案中,所述帽结构或所述经修饰的帽结构是5' 帽结构、5' 帽0结构、5' 帽1结构或5' 帽2结构。在某些实施方案中,所述一种或多种mRNA包含至少一种化学修饰。在一些实施方案中,所述化学修饰选自以下组成的组:假尿苷、N1-甲基假尿苷、N1-乙基假尿苷、2-硫代尿苷、4'-硫代尿苷、5-甲基胞嘧啶、2-硫代-1-甲基-1-脱氮-假尿苷、2-硫代-1-甲基-假尿苷、2-硫代-5-氮杂-尿苷、2-硫代-二氢假尿苷、2-硫代-二氢尿苷、2-硫代-假尿苷、4-甲氧基-2-硫代-假尿苷、4-甲氧基-假尿苷、4-硫代-1-甲基-假尿苷、4-硫代-假尿苷、5-氮杂-尿苷、二氢假尿苷、5-甲基尿苷、5-甲基尿苷、5-甲氧基尿苷和2'-O-甲基尿苷。在一些实施方案中,所述一种或多种mRNA是经完全修饰的。

[0024] 在某些实施方案中,所述一种或多种核酸编码3-10个肽表位、5-10个肽表位、10-20个肽表位、20-30个肽表位、30-40个肽表位、40-50个肽表位、50-60个肽表位、60-70个肽表位、70-80个肽表位、80-90个肽表位、90-100个肽表位、100-110个肽表位、110-120个肽表位,或120-130个肽表位。在一些实施方案中,肽表位中的每一个均由单独的开放阅读框编码。在一些实施方案中,所述肽表位呈由5-130个肽表位组成的多联癌症抗原的形式。

[0025] 在一些实施方案中,满足以下条件中的一个或多个:a)所述3-130个肽表位被切割敏感性位点穿插;和/或b)每个肽表位在不接头的情况下彼此直接联接;和/或c)每个肽表位通过单一氨基酸接头彼此联接;和/或d)每个肽表位通过短接头彼此联接;和/或e)每个肽表位包含8-31个氨基酸,并且包含一个或多个SNP突变;和/或f)每个肽表位包含8-31个氨基酸,并且包含引起独特表达的肽序列的突变;和/或g)所述肽表位的至少30%对来自受试者的I类MHC分子具有最高亲和力;和/或h)所述肽表位的至少30%对来自受试者的II类MHC分子具有最高亲和力;和/或i)所述肽表位中没有一个对来自受试者的II类MHC分子具

有最高亲和力;和/或j)所述肽表位的至少50%对HLA-A、HLA-B和/或DRB1具有 $IC_{50} < 500nM$ 的预测结合亲和力;和/或k)编码所述肽表位的核酸被排列成使得所述肽表位有序化以使假表位减到最少;和/或l) I类MHC分子肽表位与II类MHC分子肽表位的比率为至少1:1、2:1、3:1、4:1或5:1;和/或m)不存在II类MHC分子肽表位。

[0026] 在一些实施方案中,所述肽表位中的至少一个是预测的T细胞反应性表位。在某些实施方案中,所述肽表位中的至少一个是预测的B细胞反应性表位。在一些实施方案中,所述肽表位包含预测的T细胞反应性表位和预测的B细胞反应性表位的组合。在某些实施方案中,所述肽表位是预测的T细胞反应性表位和/或预测的B细胞反应性表位。在一些实施方案中,所述肽表位中的至少一个是预测的新表位。在一些实施方案中,至少一种核酸具有至少编码一种或多种传统癌症抗原或者一种或多种癌症/睾丸抗原的片段的开放阅读框。

[0027] 在一些实施方案中,每种核酸被配制在脂质纳米粒子中。在一些实施方案中,每种核酸被配制在不同的脂质纳米粒子中。在某些实施方案中,每种核酸被配制在相同的脂质纳米粒子中。

[0028] 在一些实施方案中,所述一种或多种核酸的总长度编码50-100个氨基酸、100-200个氨基酸、200-300个氨基酸、300-400个氨基酸、400-500个氨基酸、500-600个氨基酸、600-700个氨基酸、700-800个氨基酸、800-900个氨基酸、900-1000个氨基酸、1000-1100个氨基酸,或1100-1200个氨基酸的总蛋白质长度。在一些实施方案中,抗癌功效至少部分地基于一个或多个选自由以下组成的组的因素来计算:基因表达、RNA Seq、转录物丰度、DNA等位基因频率、氨基酸保守性、生理化学相似性、致癌基因、与特定HLA等位基因的预测结合亲和力、克隆性、结合效率以及在indel中的存在。在某些实施方案中,所述一个或多个因素被输入到统计模型(例如,回归模型(诸如线性回归模型、逻辑回归模型、广义线性模型等)、广义线性模型(诸如逻辑回归模型、机率单位回归模型等)、随机森林回归模型、神经网络、支持向量机、高斯混合模型、分层贝叶斯模型和/或任何其他适合的统计模型)中。

[0029] 在另一方面,本公开提供了一种用于选择用以包含在具有最大长度的核酸癌症疫苗中的核酸的计算机化系统,所述系统包括:通信接口,所述通信接口被配置为接收编码多个肽表位的多个核酸序列,其中所述肽表位中的每一个均是个性化癌症抗原的部分;和至少一个计算机处理器,所述至少一个计算机处理器被编程为:针对所述多个肽表位中的每一个,计算各自包含一个或多个肽表位中的至少一个肽表位的肽中多种核酸中的每一种的得分,其中所述核酸序列中的至少两个具有不同的长度;以及基于所计算的得分,对多种肽中的多个核酸序列进行排名;以及基于疫苗的排名和最大长度,选择用于包含在所述疫苗中的核酸序列。

[0030] 在一些实施方案中,任何肽表位的最小长度为8个氨基酸。在一些实施方案中,任何肽表位的最大长度为31个氨基酸。在某些实施方案中,所述多个核酸编码3-10个肽表位、5-10个肽表位、10-20个肽表位、20-30个肽表位、30-40个肽表位、40-50个肽表位、50-60个肽表位、60-70个肽表位、70-80个肽表位、80-90个肽表位、90-100个肽表位、100-110个肽表位、110-120个肽表位,或120-130个肽表位。

[0031] 在一些实施方案中,满足以下条件中的一个或多个:a)每个肽表位包含8-31个氨基酸,并且包含一个或多个SNP突变;和/或b)每个肽表位包含8-31个氨基酸,并且包含引起独特表达的肽序列的突变;和/或c)所述肽表位的至少30%对来自受试者的I类MHC分子具

有最高亲和力;和/或d)所述肽表位的至少30%对来自受试者的II类MHC分子具有最高亲和力;和/或e)所述肽表位中没有一个对来自受试者的II类MHC分子具有最高亲和力;和/或f)所述肽表位的至少50%对HLA-A、HLA-B和/或DRB1具有 $IC_{50} < 500nM$ 的预测结合亲和力;和/或g) I类MHC分子肽表位与II类MHC分子肽表位的比率为至少1:1、2:1、3:1、4:1或5:1;和/或h)不存在II类MHC分子肽表位。

[0032] 在一些实施方案中,所述肽表位中的至少一个是预测的T细胞反应性表位。在一些实施方案中,所述肽表位中的至少一个是预测的B细胞反应性表位。在一些实施方案中,所述肽表位包含预测的T细胞反应性表位和预测的B细胞反应性表位的组合。在某些实施方案中,所述肽表位是预测的T细胞反应性表位和/或预测的B细胞反应性表位。在一些实施方案中,所述肽表位中的至少一个是预测的新表位。在一些实施方案中,至少一种核酸具有至少编码一种或多种传统癌症抗原或者一种或多种癌症/睾丸抗原的片段的开放阅读框。

[0033] 在一些实施方案中,所述疫苗的总长度编码50-100个氨基酸、100-200个氨基酸、200-300个氨基酸、300-400个氨基酸、400-500个氨基酸、500-600个氨基酸、600-700个氨基酸、700-800个氨基酸、800-900个氨基酸、900-1000个氨基酸、1000-1100个氨基酸,或1100-1200个氨基酸的总蛋白质长度。在一些实施方案中,所述得分至少部分地基于选自由以下组成的组的一个或多个因素来计算:基因表达、RNA Seq、转录物丰度、DNA等位基因频率、氨基酸保守性、生理化学相似性、致癌基因、与特定HLA等位基因的预测结合亲和力、克隆性、结合效率以及在indel中的存在。在某些实施方案中,所述一个或多个因素被输入到统计模型(例如,回归模型(诸如线性回归模型、逻辑回归模型、广义线性模型等)、广义线性模型(诸如逻辑回归模型、机率单位回归模型等)、随机森林回归模型、神经网络、支持向量机、高斯混合模型、分层贝叶斯模型和/或任何其他适合的统计模型)中。

[0034] 在一些实施方案中,针对每种新抗原评估抗肿瘤T细胞应答。在一些实施方案中,所述评估是基于:来自WES和RNA-Seq数据的变体识别置信度;来自RNA-Seq数据的mRNA转录物丰度;来自WES和RNA-Seq数据的变异等位基因频率;以及来自NetMHCpan和NetMHCIIpan的预测HLA结合亲和力。

[0035] 在一些实施方案中,鉴定患者的HLA同种异型,并且并入预测与患者的HLA结合的抗原。在一些实施方案中,可将更大的权重分配给HLA-A、HLA-B和HLA-DR(核心靶标)的预测结合物,并且将更低的(但非零)权重分配给患者的其他HLA同种异型(补充靶标)。几乎所有个体都具有至少一种HLA-A、HLA-B和HLA-DR功能同种异型(即核心MHC等位基因),这些是所有已知人表位中的约90%人表位的限制因子(图5)。极少观察到HLA-C限制性同种异体反应性T细胞,并且HLA-C的细胞表面表达是针对HLA-A和HLA-B观察到的细胞表面表达的10%。其余补充靶标编码II类分子,并且个体可能不具有编码它们的基因。此外,即使使用最先进的NGS和其他基于序列的分型方法,对这些补充II类靶标的4位精度分型也常常是不明确的。在一些实施方案中,如果对核心或补充HLA靶标的基于NGS的等位基因分型不明确,那么在对新抗原进行排名时可能不考虑所述一种或多种等位基因。

[0036] 在一些实施方案中,可对每种新抗原执行自我性检查(selfness check)。在一些实施方案中,通过根据患者自身的种系蛋白质编码变体集合定制序列,使用参照来自人基因组注释的蛋白质编码转录物氨基酸序列创建特定于患者的转录物集合。在一些实施方案中,这种特定于患者的外显子组(不包括含有新抗原的基因)可用于检查每个I类HLA结合新

抗原表位 (8至11聚体) 是否具有100%精确自我匹配。在一些实施方案中,可将使用此工具在基因组和/或转录组中其他地方鉴定为100%自我匹配的任何新抗原从mRNA构建体中排除。

[0037] 可对所有未被自我性检查排除的变体进行评估以确定是否将其包含于特定于患者的mRNA构建体设计中。在一些实施方案中,基于MHC结合预测不完善并且RNA-Seq灵敏度可能受活检物肿瘤含量和测序深度限制的知识,可使用预定义权重而不是硬过滤器 (hard filter)。

附图说明

[0038] 图1是通过指示描绘热点突变的表。

[0039] 图2示出了netMHCpan v3.0与netMHCpan 4.0EL对HLA-A*02:01的预测%排名的比较。大量肽移入和移出0.5%排名,该值一般被认为是“强结合物”的截止值。

[0040] 图3A至图3B示出了不同的结合预测方法。图3A是示出主要HLA等位基因的预测结合物的均匀性的图。切换到百分比排名(%排名)导致所预测结合物在不同HLA等位基因上的分布更加平衡。同样,图3B是示出使用用于预测MHC结合的不同方法获得的不同样品的曲线下面积(AUC)的图。百分比排名方法已被证明与其他替代方法(例如,IC₅₀)相比提高了预测性能。

[0041] 图4A至图4C示出了体内免疫原性研究的结果。检测到20聚体/31侧翼和34聚体/25侧翼疫苗对I类表位具有相当的免疫应答,但对40聚体/21侧翼在3μg剂量和10μg剂量下均未检测到。对于数个再刺激,只有34聚体构建体在测试条件下显示出可检测到的应答。

[0042] 图5A至图5B示出了新抗原的核心和补充HLA靶标。图5A:使用免疫表位数据库(IEDB;www.iedb.org/)对所有已知的人T细胞表位(在人T细胞刺激测定中呈阳性)进行的分析,该分析揭示了HLA限制因子具有清晰的分级,其中HLA-A、HLA-B和HLA-DR占数据库中所有描述的人表位的约90%(n=8101)。图5B:将IEDB搜索工具限于仅病毒表位(n=4472)能增强T细胞对这些核心I类和II类基因座的明显偏好。此分析表明,针对预测与患者的HLA-A、HLA-B和HLA-DRB1同种异型结合的突变的新抗原选择可被赋予优先权。

[0043] 图6示出了对体细胞突变负荷的群体分析。来自cBioPortal的癌症组织学群组中非同义突变的分布。红色、蓝色和绿色线分别代表20、34和100种突变。

[0044] 图7A至图7D示出了下一代测序(NGS)和生物信息学系统输出的可重现性。使用对来自单名患者的4个相关肿瘤样品的独立处理。通过NGS、变体识别和生物信息学系统(图7A)运行原发性肿瘤样品和源自其的3个肿瘤细胞系。观察到在4个样品之间识别出的变体中的最小差异(图7B)。鉴定出的369种突变的原始新抗原得分之间的相关性(Spearman排名相关系数:肿瘤对比细胞系1: $\rho=0.86$; $p=1.92E-101$;肿瘤对比细胞系2: $\rho=0.84$; $p=3.01E-89$;以及肿瘤对比细胞系3: $\rho=0.84$; $p=5.77E-91$) (图7C)。被选择用于包含在代表性mRNA序列中的共同的独特新抗原的维恩图(图7D)。

具体实施方式

[0045] 本公开的实施方案提供了核酸(例如,DNA或RNA,诸如mRNA)疫苗,所述核酸疫苗包含一种或多种核酸,所述一种或多种核酸具有一个或多个编码肽表位的开放阅读框。如本

文所提供,编码长度不均匀的肽表位的核酸癌症疫苗可用于诱导平衡的免疫应答,包括细胞和/或体液免疫力。本文还提供了制备在给定长度下具有最大化抗癌功效的核酸癌症疫苗的方法,以及使用在给定长度下具有最大化抗癌功效的癌症疫苗治疗患有癌症的患者的方法。另外,本文提供了一种用于创建在给定长度下具有最大化抗癌功效的核酸癌症疫苗的计算机化系统。可通过基于表位或编码表位的核酸的长度鉴定T细胞活化值或存活值(诸如最大T细胞活化值或存活值)来确定最大化抗癌功效。T细胞活化值或存活值可使用本领域中已知的任何方法,例如使用可商购获得的测定(Thermo Fisher Scientific、Promega Corporation等)来确定。通常,T细胞活化值是基于细胞因子(诸如与T细胞活化相关的干扰素 γ)的表达水平变化,或细胞表面活化标记物(诸如41BB和/或OX40)的上调来确定的。存活值可相对于对照组中的存活或基于群体的存活数据来评定。

[0046] 尽管已经尝试生产核酸癌症疫苗,诸如RNA(例如,mRNA)癌症疫苗,但是此类疫苗的功效仍然是可变的。相当令人惊讶地是,发明人已经发现,对此类癌症疫苗的免疫应答可通过评估和选择不同大小的肽表位以包含在癌症疫苗中(与选择均匀长度/大小的肽表位相反)来优化。

[0047] 在疫苗开发中,生成引发针对靶向多肽序列的所需免疫应答(例如,T细胞应答)的癌症抗原仍然是一项具有挑战性的任务。本发明涉及克服与疫苗开发相关的阻碍的技术。在一些实施方案中,本发明的核酸疫苗优于常规疫苗(例如,编码均匀长度的肽表位的那些疫苗)至少10倍、20倍、40倍、50倍、100倍、500倍或1,000倍。

[0048] 作为非限制性实例,当将如本文所述的RNA(例如,mRNA)核酸癌症疫苗递送至细胞时,所述RNA(例如,mRNA)将由细胞内机制加工成多肽,所述细胞内机制然后将该多肽加工成能够刺激针对肿瘤或癌性细胞群的免疫应答的免疫敏感性片段。

[0049] 肽表位

[0050] 本公开的核酸癌症疫苗可编码一个或多个肽表位(所述一个或多个肽表位是个性化癌症抗原的部分)。个性化癌症抗原的部分是小于全长个性化癌症抗原的个性化癌症抗原的区段。个性化癌症抗原是肿瘤特异性抗原,也称为存在于个体的肿瘤中的新抗原,该抗原在个体的正常非癌组织中不表达或以低水平表达。该抗原可能存在或不存在于其他个体的肿瘤中。

[0051] 在一个实施方案中,核酸癌症疫苗由可含有任何数目的肽表位的开放阅读框构成。在一些实施方案中,核酸癌症疫苗由编码2个或更多个、3个或更多个、4个或更多个、5个或更多个、6个或更多个、7个或更多个、8个或更多个、9个或更多个、10个或更多个、11个或更多个、12个或更多个、13个或更多个、14个或更多个、15个或更多个、16个或更多个、17个或更多个、18个或更多个、19个或更多个、20个或更多个、21个或更多个、22个或更多个、23个或更多个、24个或更多个、25个或更多个、26个或更多个、27个或更多个、28个或更多个、29个或更多个、30个或更多个、31个或更多个、32个或更多个、33个或更多个、34个或更多个、35个或更多个、36个或更多个、37个或更多个、38个或更多个、39个或更多个、40个或更多个、45个或更多个、50个或更多个、55个或更多个、60个或更多个、65个或更多个、70个或更多个、75个或更多个、80个或更多个、85个或更多个、90个或更多个、95个或更多个、100个或更多个、105个或更多个、110个或更多个、115个或更多个、120个或更多个、125个或更多个、130个或更多个、135个或更多个、140个或更多个、145个或更多个、150个或更多个、155

个或更多个、160个或更多个、165个或更多个、170个或更多个、175个或更多个、180个或更多个、185个或更多个、190个或更多个、195个或更多个,或200个或更多个肽表位的开放阅读框构成。在其他实施方案中,核酸癌症疫苗由编码200个或更少、195个或更少、190个或更少、185个或更少、180个或更少、175个或更少、170个或更少、165个或更少、160个或更少、155个或更少、150个或更少、145个或更少、140个或更少、135个或更少、130个或更少、125个或更少、120个或更少、115个或更少、110个或更少、100个或更少、95个或更少、90个或更少、85个或更少、80个或更少、75个或更少、70个或更少、65个或更少、60个或更少、55个或更少、50个或更少、45个或更少、40个或更少、35个或更少、30个或更少、25个或更少、20个或更少、15个或更少、10个或更少,或5个或更少的肽表位的开放阅读框构成。在其他实施方案中,核酸癌症疫苗由编码多达200个、多达195个、多达190个、多达185个、多达180个、多达175个、多达170个、多达165个、多达160个、多达155个、多达150个、多达145个、多达140个、多达135个、多达130个、多达125个、多达120个、多达115个、多达110个、多达100个、多达95个、多达90个、多达85个、多达80个、多达75个、多达70个、多达65个、多达60个、多达55个、多达50个、多达45个、多达40个、多达35个、多达30个、多达25个、多达20个、多达15个、多达10个肽表位、多达5个肽表位,或多达3个肽表位的开放阅读框构成。

[0052] 在某些实施方案中,核酸癌症疫苗编码3-10个肽表位、5-10个肽表位、10-20个肽表位、20-30个肽表位、30-40个肽表位、40-50个肽表位、50-60个肽表位、60-70个肽表位、70-80个肽表位、80-90个肽表位、90-100个肽表位、100-110个肽表位、110-120个肽表位、120-130个肽表位、130-140个肽表位、140-150个肽表位、150-160个肽表位、160-170个肽表位、170-180个肽表位、180-190个肽表位,或190-200个肽表位。

[0053] 在某些实施方案中,核酸癌症疫苗编码2-200个、5-200个、8-200个、10-200个、2-190个、5-190个、8-190个、10-190个、2-180个、5-180个、8-180个、10-180个、2-170个、5-170个、8-170个、10-170个、2-160个、5-160个、8-160个、10-160个、2-150个、5-150个、8-150个、10-150个、2-145个、5-145个、8-145个、10-145个、2-140个、5-140个、8-140个、10-140个、2-139个、5-139个、8-139个、10-139个、2-138个、5-138个、8-138个、10-138个、2-137个、5-137个、8-137个、10-137个、2-136个、5-136个、8-136个、10-136个、2-135个、5-135个、8-135个、10-135个、2-134个、5-134个、8-134个、10-134个、2-133个、5-133个、8-133个、10-133个、2-132个、5-132个、8-132个、10-132个、2-131个、5-131个、8-131个、10-131个、2-130个、5-130个、8-130个、10-130个、2-129个、5-129个、8-129个、10-129个、2-128个、5-128个、8-128个、10-128个、2-127个、5-127个、8-127个、10-127个、2-126个、5-126个、8-126个、10-126个、2-125个、5-125个、8-125个、10-125个、2-124个、5-124个、8-124个、10-124个、2-123个、5-123个、8-123个、10-123个、2-122个、5-122个、8-122个、10-122个、2-121个、5-121个、8-121个、10-121个、2-120个、5-120个、8-120个、10-120个、2-119个、5-119个、8-119个、10-119个、2-118个、5-118个、8-118个、10-118个、2-117个、5-117个、8-117个、10-117个、2-116个、5-116个、8-116个、10-116个、2-115个、5-115个、8-115个、10-115个、2-114个、5-114个、8-114个、10-114个、2-113个、5-113个、8-113个、10-113个、2-112个、5-112个、8-112个、10-112个、2-111个、5-111个、8-111个、10-111个、2-110个、5-110个、8-110个、10-110个、2-100个、5-100个、8-100个,或10-100个肽表位。

[0054] 在其他实施方案中,核酸癌症疫苗编码2-95个、5-95个、8-95个、10-95个、2-90个、

5-90个、8-90个、10-85个、2-85个、5-85个、8-85个、10-85个、2-80个、5-80个、8-80个、10-80个、2-85个、5-85个、8-85个、10-85个、2-80个、5-80个、8-80个、10-80个、2-75个、5-75个、8-75个、10-75个、2-70个、5-70个、8-70个、10-70个、2-65个、5-65个、8-65个、10-65个、2-60个、5-60个、8-60个、10-60个、2-55个、5-55个、8-55个、10-55个、2-50个、5-50个、8-50个、10-50个、2-45个、5-45个、8-45个、10-45个、2-40个、5-40个、8-40个、10-40个、2-39个、5-39个、8-39个、10-39个、2-38个、5-38个、8-38个、10-38个、2-37个、5-37个、8-37个、10-37个、2-36个、5-36个、8-36个、10-36个、2-35个、5-35个、8-35个、10-35个、2-34个、5-34个、8-34个、10-34个、2-33个、5-33个、8-33个、10-33个、2-32个、5-32个、8-32个、10-32个、2-31个、5-31个、8-31个、10-31个、2-30个、5-30个、8-30个、10-30个、2-29个、5-29个、8-29个、10-29个、2-28个、5-28个、8-28个、10-28个、2-27个、5-27个、8-27个、10-27个、2-26个、5-26个、8-26个、10-26个、2-25个、5-25个、8-25个、10-25个、2-24个、5-24个、8-24个、10-24个、2-23个、5-23个、8-23个、10-23个、2-22个、5-22个、8-22个、10-22个、2-21个、5-21个、8-21个、10-21个、2-20个、5-20个、8-20个、10-20个、2-19个、5-19个、8-19个、10-19个、2-18个、5-18个、8-18个、10-18个、2-17个、5-17个、8-17个、10-17个、2-16个、5-16个、8-16个、10-16个、2-15个、5-15个、8-15个、10-15个、2-14个、5-14个、8-14个、10-14个、2-13个、5-13个、8-13个、10-13个、2-12个、5-12个、8-12个、10-12个、2-11个、5-11个、8-11个、10-11个、2-10个、5-10个,或8-10个肽表位。

[0055] 在又一些其他实施方案中,核酸癌症疫苗编码20-200个、30-200个、40-200个、50-200个、20-180个、30-180个、40-180个、50-180个、20-170个、30-170个、40-170个、50-170个、20-160个、30-160个、40-160个、20-150个、30-150个、40-150个、50-150个、20-140个、30-140个、40-140个、50-140个、20-130个、30-130个、40-130个、50-130个、20-120个、30-120个、40-120个、50-120个、20-110个、30-110个、40-110个、50-110个、20-100个、30-100个、40-100个,或50-100个肽表位。在一个实施方案中,核酸疫苗编码34个肽表位。

[0056] 在一些实施方案中,本文所述的核酸癌症疫苗和疫苗接种方法包括编码基于特定突变的表位或抗原(新表位)和/或由癌症种系基因表达的那些表位或抗原(在多名患者中发现的肿瘤所共同具有的抗原)的开放阅读框。

[0057] 如本文所用,表位(也称为抗原决定簇)是在适当的背景下被免疫系统(特别是被抗体、B细胞或T细胞)辨识的抗原的一部分。表位可包括B细胞表位(例如,预测的B细胞反应性表位)和T细胞表位(例如,预测的T细胞反应性表位)。B细胞表位(例如,预测的B细胞反应性表位)是被产生特异性抗体的B细胞辨识所必需的肽序列。B细胞表位(例如,预测的B细胞反应性表位)是指被抗体辨识的抗原的特定区域。T细胞表位(例如,预测的T细胞反应性表位)是这样的肽序列,其与APC上的蛋白质缔合,是被特异性T细胞辨识所必需的。T细胞表位(例如,预测的T细胞反应性表位)在细胞内被加工并被呈递在APC的表面上,在那里它们与MHC分子(包括II类MHC分子和I类MHC分子)结合。抗体的与该表位结合的那部分称为互补位。基于结构及与互补位的相互作用,表位可以是构象表位或线性表位。线性或连续表位由蛋白质的特定区域的一级氨基酸序列限定。与抗体相互作用的序列循序地彼此相邻地位于该蛋白质上,并且该表位通常可由单个肽模拟。构象表位是由原生蛋白质的构象结构限定的表位。这些表位可以是连续的或不连续的(即,该表位的组分可能可位于该蛋白质的迥然不同的部分上,该迥然不同的部分在折叠的原生蛋白质结构中得以彼此接近)。

[0058] 每个肽表位可以是对于表位合理的任何长度。在一些实施方案中,每个肽表位的长度不一定相等。在一些实施方案中,核酸癌症疫苗中的每个肽表位具有不同的长度。在某些实施方案中,核酸癌症疫苗中的肽表位中的至少两个(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,以及多达且包括全部)具有不同的长度。

[0059] 在一些实施方案中,所述肽表位中的至少一个的长度为至少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、至少21个、至少22个、至少23个、至少24个、至少25个、至少26个、至少27个、至少28个、至少29个、至少30个、至少31个、至少32个、至少33个、至少34个、至少35个、至少36个、至少37个、至少38个、至少39个、至少40个、至少45个、至少50个、至少55个、至少60个、至少65个、至少70个、至少75个、至少80个、至少85个、至少90个、至少95个,或至少100个氨基酸。在其他实施方案中,所述肽表位中的至少一个的长度为100个或更少、95个或更少、90个或更少、85个或更少、80个或更少、75个或更少、70个或更少、65个或更少、60个或更少、55个或更少、50个或更少、45个或更少、40个或更少、35个或更少、30个或更少、25个或更少、20个或更少、15个或更少、14个或更少、13个或更少、12个或更少、11个或更少、10个或更少、9个或更少、8个或更少、7个或更少、6个或更少、5个或更少、4个或更少、3个或更少,或2个或更少的氨基酸。在其他实施方案中,所述肽表位中的至少一个的长度为多达100个、多达95个、多达90个、多达85个、多达80个、多达75个、多达70个、多达65个、多达60个、多达55个、多达50个、多达45个、多达40个、多达35个、多达30个、多达25个、多达20个、多达15个,或多达10个氨基酸。

[0060] 在一些实施方案中,每个肽表位的长度可以是5-100个氨基酸(包括端点)。在一些实施方案中,所述肽表位中的至少一个的长度为5-100个、5-95个、5-90个、5-85个、5-80个、5-75个、5-70个、5-65个、5-60个、5-55个、5-50个、5-45个、5-40个、5-39个、5-38个、5-37个、5-36个、5-35个、5-34个、5-33个、5-32个、5-31个、5-30个、5-29个、5-28个、5-27个、5-26个、5-25个、5-24个、5-23个、5-22个、5-21个、5-20个、8-100个、8-95个、8-90个、8-85个、8-80个、8-75个、8-70个、8-65个、8-60个、8-55个、8-50个、8-45个、8-40个、8-39个、8-38个、8-37个、8-36个、8-35个、8-34个、8-33个、8-32个、8-31个、8-30个、8-29个、8-28个、8-27个、8-26个、8-25个、8-24个、8-23个、8-22个、8-21个、8-20个、10-100个、10-95个、10-90个、10-85个、10-80个、10-75个、10-70个、10-65个、10-60个、10-55个、10-50个、10-45个、10-40个、10-39个、10-38个、10-37个、10-36个、10-35个、10-34个、10-33个、10-32个、10-31个、10-30个、10-29个、10-28个、10-27个、10-26个、10-25个、10-24个、10-23个、10-22个、10-21个,或10-20个氨基酸。

[0061] 在一些实施方案中,由核酸癌症疫苗编码的肽表位中的每一个均可具有不同的长度。在某些实施方案中,所述肽表位中的至少一个具有与由核酸癌症疫苗编码的另一个肽表位的长度不同的长度。每个肽表位可以是对于表位合理的任何长度。

[0062] 在一些实施方案中,核酸编码不同百分比的肽表位长度。以下清单中描述的所有百分比都可能是近似的(即,在所陈述的量的5%之内)。术语“近似”和“约”的使用是等效的。

[0063] 在一些实施方案中,由核酸编码的肽表位长度的百分比可以如下:约100%<15个

氨基酸,约0% $\geq$ 15个氨基酸;约95% $<$ 15个氨基酸,约5% $\geq$ 15个氨基酸;约90% $<$ 15个氨基酸,约10% $\geq$ 15个氨基酸;约85% $<$ 15个氨基酸,约15% $\geq$ 15个氨基酸;约80% $<$ 15个氨基酸,约20% $\geq$ 15个氨基酸;约75% $<$ 15个氨基酸,约25% $\geq$ 15个氨基酸;约70% $<$ 15个氨基酸,约30% $\geq$ 15个氨基酸;约65% $<$ 15个氨基酸,约35% $\geq$ 15个氨基酸;约60% $<$ 15个氨基酸,约40% $\geq$ 15个氨基酸;约55% $<$ 15个氨基酸,约45% $\geq$ 15个氨基酸;约50% $<$ 15个氨基酸,约50% $\geq$ 15个氨基酸;约45% $<$ 15个氨基酸,约55% $\geq$ 15个氨基酸;约40% $<$ 15个氨基酸,约60% $\geq$ 15个氨基酸;约35% $<$ 15个氨基酸,约65% $\geq$ 15个氨基酸;约30% $<$ 15个氨基酸,约70% $\geq$ 15个氨基酸;约25% $<$ 15个氨基酸,约75% $\geq$ 15个氨基酸;约20% $<$ 15个氨基酸,约80% $\geq$ 15个氨基酸;约15% $<$ 15个氨基酸,约85% $\geq$ 15个氨基酸;约10% $<$ 15个氨基酸,约90% $\geq$ 15个氨基酸;约5% $<$ 15个氨基酸,约95% $\geq$ 15个氨基酸;或约0% $<$ 15个氨基酸,约100% $\geq$ 15个氨基酸。

[0064] 在一些实施方案中,由核酸编码的肽表位长度的百分比可以如下:约100% $<$ 17个氨基酸,约0% $\geq$ 17个氨基酸;约95% $<$ 17个氨基酸,约5% $\geq$ 17个氨基酸;约90% $<$ 17个氨基酸,约10% $\geq$ 17个氨基酸;约85% $<$ 17个氨基酸,约17% $\geq$ 17个氨基酸;约80% $<$ 17个氨基酸,约20% $\geq$ 17个氨基酸;约75% $<$ 17个氨基酸,约25% $\geq$ 17个氨基酸;约70% $<$ 17个氨基酸,约30% $\geq$ 17个氨基酸;约65% $<$ 17个氨基酸,约35% $\geq$ 17个氨基酸;约60% $<$ 17个氨基酸,约40% $\geq$ 17个氨基酸;约55% $<$ 17个氨基酸,约45% $\geq$ 17个氨基酸;约50% $<$ 17个氨基酸,约50% $\geq$ 17个氨基酸;约45% $<$ 17个氨基酸,约55% $\geq$ 17个氨基酸;约40% $<$ 17个氨基酸,约60% $\geq$ 17个氨基酸;约35% $<$ 17个氨基酸,约65% $\geq$ 17个氨基酸;约30% $<$ 17个氨基酸,约70% $\geq$ 17个氨基酸;约25% $<$ 17个氨基酸,约75% $\geq$ 17个氨基酸;约20% $<$ 17个氨基酸,约80% $\geq$ 17个氨基酸;约17% $<$ 17个氨基酸,约85% $\geq$ 17个氨基酸;约10% $<$ 17个氨基酸,约90% $\geq$ 17个氨基酸;约5% $<$ 17个氨基酸,约95% $\geq$ 17个氨基酸;或约0% $<$ 17个氨基酸,约100% $\geq$ 17个氨基酸。

[0065] 在一些实施方案中,由核酸编码的肽表位长度的百分比可以如下:约100% $<$ 19个氨基酸,约0% $\geq$ 19个氨基酸;约95% $<$ 19个氨基酸,约5% $\geq$ 19个氨基酸;约90% $<$ 19个氨基酸,约10% $\geq$ 19个氨基酸;约85% $<$ 19个氨基酸,约19% $\geq$ 19个氨基酸;约80% $<$ 19个氨基酸,约20% $\geq$ 19个氨基酸;约75% $<$ 19个氨基酸,约25% $\geq$ 19个氨基酸;约70% $<$ 19个氨基酸,约30% $\geq$ 19个氨基酸;约65% $<$ 19个氨基酸,约35% $\geq$ 19个氨基酸;约60% $<$ 19个氨基酸,约40% $\geq$ 19个氨基酸;约55% $<$ 19个氨基酸,约45% $\geq$ 19个氨基酸;约50% $<$ 19个氨基酸,约50% $\geq$ 19个氨基酸;约45% $<$ 19个氨基酸,约55% $\geq$ 19个氨基酸;约40% $<$ 19个氨基酸,约60% $\geq$ 19个氨基酸;约35% $<$ 19个氨基酸,约65% $\geq$ 19个氨基酸;约30% $<$ 19个氨基酸,约70% $\geq$ 19个氨基酸;约25% $<$ 19个氨基酸,约75% $\geq$ 19个氨基酸;约20% $<$ 19个氨基酸,约80% $\geq$ 19个氨基酸;约19% $<$ 19个氨基酸,约85% $\geq$ 19个氨基酸;约10% $<$ 19个氨基酸,约90% $\geq$ 19个氨基酸;约5% $<$ 19个氨基酸,约95% $\geq$ 19个氨基酸;或约0% $<$ 19个氨基酸,约100% $\geq$ 19个氨基酸。

[0066] 在一些实施方案中,肽表位长度可归类为以下组之一(总计100%):8-12个氨基酸、13-17个氨基酸、18-21个氨基酸、22-26个氨基酸,或27-31个氨基酸。由核酸的开放阅读框编码的肽表位的约0%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%,或100%的长度可以是8-12个氨基酸。由核

酸的开放阅读框编码的肽表位的约0%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%，或100%的长度可以是13-17个氨基酸。由核酸的开放阅读框编码的肽表位的约0%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%，或100%的长度可以是18-21个氨基酸。由核酸的开放阅读框编码的肽表位的约0%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%，或100%的长度可以是22-26个氨基酸。由核酸的开放阅读框编码的肽表位的约0%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%，或100%的长度可以是27-31个氨基酸。下面是由核酸的开放阅读框编码的肽表位长度百分比的几个非限制性实例。

[0067] 在一些实施方案中，由核酸编码的肽表位长度的百分比可以如下：50%8-12个氨基酸，50%13-17个氨基酸，0%18-21个氨基酸，0%22-26个氨基酸，和0%27-31个氨基酸；0%8-12个氨基酸，50%13-17个氨基酸，50%18-21个氨基酸，0%22-26个氨基酸，和0%27-31个氨基酸；0%8-12个氨基酸，0%13-17个氨基酸，50%18-21个氨基酸，50%22-26个氨基酸，和0%27-31个氨基酸；0%8-12个氨基酸，0%13-17个氨基酸，0%18-21个氨基酸，50%22-26个氨基酸，和50%27-31个氨基酸；50%8-12个氨基酸，0%13-17个氨基酸，50%18-21个氨基酸，0%22-26个氨基酸，和0%27-31个氨基酸；50%8-12个氨基酸，0%13-17个氨基酸，0%18-21个氨基酸，50%22-26个氨基酸，和0%27-31个氨基酸；50%8-12个氨基酸，0%13-17个氨基酸，0%18-21个氨基酸，0%22-26个氨基酸，和50%27-31个氨基酸；0%8-12个氨基酸，50%13-17个氨基酸，50%18-21个氨基酸，0%22-26个氨基酸，和0%27-31个氨基酸；0%8-12个氨基酸，50%13-17个氨基酸，0%18-21个氨基酸，50%22-26个氨基酸，和0%27-31个氨基酸；0%8-12个氨基酸，50%13-17个氨基酸，0%18-21个氨基酸，0%22-26个氨基酸，和50%27-31个氨基酸；或0%8-12个氨基酸，0%13-17个氨基酸，50%18-21个氨基酸，0%22-26个氨基酸，和50%27-31个氨基酸。

[0068] 在一些实施方案中，由核酸编码的肽表位长度的百分比可以如下：10%8-12个氨基酸，40%13-17个氨基酸，40%18-21个氨基酸，10%22-26个氨基酸，和0%27-31个氨基酸；10%8-12个氨基酸，10%13-17个氨基酸，40%18-21个氨基酸，40%22-26个氨基酸，和0%27-31个氨基酸；40%8-12个氨基酸，40%13-17个氨基酸，10%18-21个氨基酸，10%22-26个氨基酸，和0%27-31个氨基酸；10%8-12个氨基酸，40%13-17个氨基酸，10%18-21个氨基酸，40%22-26个氨基酸，和0%27-31个氨基酸；40%8-12个氨基酸，10%13-17个氨基酸，40%18-21个氨基酸，10%22-26个氨基酸，和0%27-31个氨基酸；0%8-12个氨基酸，10%13-17个氨基酸，40%18-21个氨基酸，40%22-26个氨基酸，和10%27-31个氨基酸；0%8-12个氨基酸，10%13-17个氨基酸，10%18-21个氨基酸，40%22-26个氨基酸，和40%27-31个氨基酸；0%8-12个氨基酸，40%13-17个氨基酸，40%18-21个氨基酸，10%22-26个氨基酸，和10%27-31个氨基酸；0%8-12个氨基酸，10%13-17个氨基酸，40%18-21个氨基酸，10%22-26个氨基酸，和40%27-31个氨基酸；0%8-12个氨基酸，40%13-17个氨基酸，10%18-21个氨基酸，40%22-26个氨基酸，和10%27-31个氨基酸。

[0069] 在一些实施方案中，由核酸编码的肽表位长度的百分比可以如下：25%8-12个氨基酸，25%13-17个氨基酸，25%18-21个氨基酸，25%22-26个氨基酸，和0%27-31个氨基

酸;25%8-12个氨基酸,25%13-17个氨基酸,25%18-21个氨基酸,0%22-26个氨基酸,和25%27-31个氨基酸;25%8-12个氨基酸,25%13-17个氨基酸,0%18-21个氨基酸,25%22-26个氨基酸,和25%27-31个氨基酸;25%8-12个氨基酸,0%13-17个氨基酸,25%18-21个氨基酸,25%22-26个氨基酸,和25%27-31个氨基酸;0%8-12个氨基酸,25%13-17个氨基酸,25%18-21个氨基酸,25%22-26个氨基酸,和25%27-31个氨基酸。

[0070] 在一些实施方案中,由核酸编码的肽表位长度的百分比可以如下:15%8-12个氨基酸,15%13-17个氨基酸,15%18-21个氨基酸,15%22-26个氨基酸,和40%27-31个氨基酸;15%8-12个氨基酸,15%13-17个氨基酸,15%18-21个氨基酸,15%22-26个氨基酸,和40%27-31个氨基酸;15%8-12个氨基酸,15%13-17个氨基酸,15%18-21个氨基酸,15%22-26个氨基酸,和40%27-31个氨基酸;15%8-12个氨基酸,15%13-17个氨基酸,15%18-21个氨基酸,15%22-26个氨基酸,和40%27-31个氨基酸;15%8-12个氨基酸,15%13-17个氨基酸,15%18-21个氨基酸,15%22-26个氨基酸,和40%27-31个氨基酸;40%8-12个氨基酸,15%13-17个氨基酸,15%18-21个氨基酸,15%22-26个氨基酸,和15%27-31个氨基酸;40%8-12个氨基酸,15%13-17个氨基酸,15%18-21个氨基酸,15%22-26个氨基酸,和15%27-31个氨基酸;40%8-12个氨基酸,15%13-17个氨基酸,15%18-21个氨基酸,15%22-26个氨基酸,和15%27-31个氨基酸;40%8-12个氨基酸,15%13-17个氨基酸,15%18-21个氨基酸,15%22-26个氨基酸,和15%27-31个氨基酸;40%8-12个氨基酸,15%13-17个氨基酸,15%18-21个氨基酸,15%22-26个氨基酸,和15%27-31个氨基酸。

[0071] 在一些实施方案中,由核酸编码的肽表位长度的百分比可以如下:10%8-12个氨基酸,10%13-17个氨基酸,10%18-21个氨基酸,10%22-26个氨基酸,和60%27-31个氨基酸;10%8-12个氨基酸,10%13-17个氨基酸,10%18-21个氨基酸,10%22-26个氨基酸,和60%27-31个氨基酸;10%8-12个氨基酸,10%13-17个氨基酸,10%18-21个氨基酸,10%22-26个氨基酸,和60%27-31个氨基酸;10%8-12个氨基酸,10%13-17个氨基酸,10%18-21个氨基酸,10%22-26个氨基酸,和60%27-31个氨基酸;10%8-12个氨基酸,10%13-17个氨基酸,10%18-21个氨基酸,10%22-26个氨基酸,和60%27-31个氨基酸;60%8-12个氨基酸,10%13-17个氨基酸,10%18-21个氨基酸,10%22-26个氨基酸,和10%27-31个氨基酸;60%8-12个氨基酸,10%13-17个氨基酸,10%18-21个氨基酸,10%22-26个氨基酸,和10%27-31个氨基酸;60%8-12个氨基酸,10%13-17个氨基酸,10%18-21个氨基酸,10%22-26个氨基酸,和10%27-31个氨基酸;60%8-12个氨基酸,10%13-17个氨基酸,10%18-21个氨基酸,10%22-26个氨基酸,和10%27-31个氨基酸;60%8-12个氨基酸,10%13-17个氨基酸,10%18-21个氨基酸,10%22-26个氨基酸,和10%27-31个氨基酸。

[0072] 在一些实施方案中,由核酸编码的肽表位长度的百分比可以如下:15%8-12个氨基酸,20%13-17个氨基酸,20%18-21个氨基酸,15%22-26个氨基酸,和30%27-31个氨基酸;15%8-12个氨基酸,15%13-17个氨基酸,20%18-21个氨基酸,20%22-26个氨基酸,和30%27-31个氨基酸;20%8-12个氨基酸,20%13-17个氨基酸,15%18-21个氨基酸,15%22-26个氨基酸,和30%27-31个氨基酸;15%8-12个氨基酸,20%13-17个氨基酸,15%18-21个氨基酸,20%22-26个氨基酸,和30%27-31个氨基酸;20%8-12个氨基酸,15%13-17个氨基酸,20%18-21个氨基酸,15%22-26个氨基酸,和30%27-31个氨基酸;30%8-12个氨基酸,15%13-17个氨基酸,20%18-21个氨基酸,20%22-26个氨基酸,和15%27-31个氨基酸;

30%8-12个氨基酸,15%13-17个氨基酸,15%18-21个氨基酸,20%22-26个氨基酸,和20%27-31个氨基酸;30%8-12个氨基酸,20%13-17个氨基酸,20%18-21个氨基酸,15%22-26个氨基酸,和15%27-31个氨基酸;30%8-12个氨基酸,15%13-17个氨基酸,20%18-21个氨基酸,15%22-26个氨基酸,和20%27-31个氨基酸;30%8-12个氨基酸,20%13-17个氨基酸,15%18-21个氨基酸,20%22-26个氨基酸,和15%27-31个氨基酸。

[0073] 在一些实施方案中,由核酸编码的肽表位长度的百分比可以如下:35%8-12个氨基酸,35%13-17个氨基酸,10%18-21个氨基酸,10%22-26个氨基酸,和10%27-31个氨基酸;10%8-12个氨基酸,35%13-17个氨基酸,35%18-21个氨基酸,10%22-26个氨基酸,和10%27-31个氨基酸;10%8-12个氨基酸,10%13-17个氨基酸,35%18-21个氨基酸,35%22-26个氨基酸,和10%27-31个氨基酸;10%8-12个氨基酸,10%13-17个氨基酸,10%18-21个氨基酸,35%22-26个氨基酸,和35%27-31个氨基酸;35%8-12个氨基酸,10%13-17个氨基酸,35%18-21个氨基酸,10%22-26个氨基酸,和10%27-31个氨基酸;35%8-12个氨基酸,10%13-17个氨基酸,10%18-21个氨基酸,35%22-26个氨基酸,和10%27-31个氨基酸;35%8-12个氨基酸,10%13-17个氨基酸,10%18-21个氨基酸,10%22-26个氨基酸,和35%27-31个氨基酸;10%8-12个氨基酸,35%13-17个氨基酸,10%18-21个氨基酸,35%22-26个氨基酸,和10%27-31个氨基酸;10%8-12个氨基酸,35%13-17个氨基酸,10%18-21个氨基酸,10%22-26个氨基酸,和35%27-31个氨基酸。

[0074] 在一些实施方案中,由核酸编码的肽表位长度的百分比可以如下:30%8-12个氨基酸,30%13-17个氨基酸,30%18-21个氨基酸,5%22-26个氨基酸,和5%27-31个氨基酸;5%8-12个氨基酸,30%13-17个氨基酸,30%18-21个氨基酸,30%22-26个氨基酸,和5%27-31个氨基酸;5%8-12个氨基酸,5%13-17个氨基酸,30%18-21个氨基酸,30%22-26个氨基酸,和30%27-31个氨基酸;30%8-12个氨基酸,5%13-17个氨基酸,5%18-21个氨基酸,30%22-26个氨基酸,和30%27-31个氨基酸;30%8-12个氨基酸,30%13-17个氨基酸,5%18-21个氨基酸,5%22-26个氨基酸,和30%27-31个氨基酸;5%8-12个氨基酸,30%13-17个氨基酸,5%18-21个氨基酸,30%22-26个氨基酸,和30%27-31个氨基酸;5%8-12个氨基酸,30%13-17个氨基酸,30%18-21个氨基酸,5%22-26个氨基酸,和30%27-31个氨基酸;30%8-12个氨基酸,30%13-17个氨基酸,5%18-21个氨基酸,30%22-26个氨基酸,和5%27-31个氨基酸;30%8-12个氨基酸,5%13-17个氨基酸,30%18-21个氨基酸,5%22-26个氨基酸,和30%27-31个氨基酸。

[0075] 在一些实施方案中,由核酸编码的肽表位长度的百分比可以如下:20%8-12个氨基酸,20%13-17个氨基酸,20%18-21个氨基酸,20%22-26个氨基酸,和20%27-31个氨基酸。

[0076] 在一些实施方案中,肽表位的最佳长度可通过以下程序获得:合成V5标签多联体-测试蛋白酶位点,将其引入到DC细胞中(例如,使用RNA Squeeze程序),裂解细胞,然后运行抗V5蛋白质印迹法以评定蛋白酶位点处的切割。

[0077] RNA Squeeze技术是一种细胞内递送方法,通过该方法可将多种材料递送至范围广泛的活细胞。细胞经受微流体构造,这引起快速的机械变形。变形导致暂时的膜破裂和新形成的瞬时孔。材料然后经由这些瞬时孔被动地扩散到细胞胞质溶胶中。该技术可用于包括原代成纤维细胞、胚胎干细胞和大量免疫细胞在内的多种细胞类型,并且已被证实在大

(A)₁₋₅₀、(C)₁₋₅₀(E)₁₋₅₀(D)₁₋₅₀(B)₁₋₅₀(A)₁₋₅₀、(C)₁₋₅₀(E)₁₋₅₀(D)₁₋₅₀(A)₁₋₅₀(B)₁₋₅₀、(D)₁₋₅₀(B)₁₋₅₀、
 (C)₁₋₅₀(A)₁₋₅₀(E)₁₋₅₀、(D)₁₋₅₀(B)₁₋₅₀(C)₁₋₅₀(E)₁₋₅₀(A)₁₋₅₀、(D)₁₋₅₀(B)₁₋₅₀(A)₁₋₅₀(C)₁₋₅₀(E)₁₋₅₀、
 (D)₁₋₅₀(B)₁₋₅₀(A)₁₋₅₀(E)₁₋₅₀(C)₁₋₅₀、(D)₁₋₅₀(B)₁₋₅₀(E)₁₋₅₀(C)₁₋₅₀(A)₁₋₅₀、(D)₁₋₅₀(B)₁₋₅₀(E)₁₋₅₀、
 (A)₁₋₅₀(C)₁₋₅₀、(D)₁₋₅₀(C)₁₋₅₀(A)₁₋₅₀(E)₁₋₅₀(B)₁₋₅₀、(D)₁₋₅₀(C)₁₋₅₀(A)₁₋₅₀(B)₁₋₅₀(E)₁₋₅₀、(D)₁₋₅₀、
 (C)₁₋₅₀(E)₁₋₅₀(A)₁₋₅₀(B)₁₋₅₀、(D)₁₋₅₀(C)₁₋₅₀(E)₁₋₅₀(B)₁₋₅₀(A)₁₋₅₀、(D)₁₋₅₀(C)₁₋₅₀(B)₁₋₅₀(E)₁₋₅₀、
 (A)₁₋₅₀、(D)₁₋₅₀(C)₁₋₅₀(B)₁₋₅₀(A)₁₋₅₀(E)₁₋₅₀、(D)₁₋₅₀(A)₁₋₅₀(C)₁₋₅₀(B)₁₋₅₀(E)₁₋₅₀、(D)₁₋₅₀(A)₁₋₅₀、
 (C)₁₋₅₀(E)₁₋₅₀(B)₁₋₅₀、(D)₁₋₅₀(A)₁₋₅₀(B)₁₋₅₀(C)₁₋₅₀(E)₁₋₅₀、(D)₁₋₅₀(A)₁₋₅₀(B)₁₋₅₀(E)₁₋₅₀(C)₁₋₅₀、
 (D)₁₋₅₀(A)₁₋₅₀(E)₁₋₅₀(B)₁₋₅₀(C)₁₋₅₀、(D)₁₋₅₀(A)₁₋₅₀(E)₁₋₅₀(C)₁₋₅₀(B)₁₋₅₀、(D)₁₋₅₀(E)₁₋₅₀(C)₁₋₅₀、
 (B)₁₋₅₀(A)₁₋₅₀、(D)₁₋₅₀(E)₁₋₅₀(C)₁₋₅₀(A)₁₋₅₀(B)₁₋₅₀、(D)₁₋₅₀(E)₁₋₅₀(B)₁₋₅₀(C)₁₋₅₀(A)₁₋₅₀、(D)₁₋₅₀、
 (E)₁₋₅₀(B)₁₋₅₀(A)₁₋₅₀(C)₁₋₅₀、(D)₁₋₅₀(E)₁₋₅₀(A)₁₋₅₀(B)₁₋₅₀(C)₁₋₅₀、(D)₁₋₅₀(E)₁₋₅₀(A)₁₋₅₀(C)₁₋₅₀、
 (B)₁₋₅₀、(E)₁₋₅₀(B)₁₋₅₀(C)₁₋₅₀(D)₁₋₅₀(A)₁₋₅₀、(E)₁₋₅₀(B)₁₋₅₀(C)₁₋₅₀(A)₁₋₅₀(D)₁₋₅₀、(E)₁₋₅₀(B)₁₋₅₀、
 (D)₁₋₅₀(C)₁₋₅₀(A)₁₋₅₀、(E)₁₋₅₀(B)₁₋₅₀(D)₁₋₅₀(A)₁₋₅₀(C)₁₋₅₀、(E)₁₋₅₀(B)₁₋₅₀(A)₁₋₅₀(C)₁₋₅₀(D)₁₋₅₀、
 (E)₁₋₅₀(B)₁₋₅₀(A)₁₋₅₀(D)₁₋₅₀(C)₁₋₅₀、(E)₁₋₅₀(C)₁₋₅₀(D)₁₋₅₀(A)₁₋₅₀(B)₁₋₅₀、(E)₁₋₅₀(C)₁₋₅₀(D)₁₋₅₀、
 (B)₁₋₅₀(A)₁₋₅₀、(E)₁₋₅₀(C)₁₋₅₀(A)₁₋₅₀(D)₁₋₅₀(B)₁₋₅₀、(E)₁₋₅₀(C)₁₋₅₀(A)₁₋₅₀(B)₁₋₅₀(D)₁₋₅₀、(E)₁₋₅₀、
 (C)₁₋₅₀(B)₁₋₅₀(A)₁₋₅₀(D)₁₋₅₀、(E)₁₋₅₀(C)₁₋₅₀(B)₁₋₅₀(D)₁₋₅₀(A)₁₋₅₀、(E)₁₋₅₀(D)₁₋₅₀(C)₁₋₅₀(B)₁₋₅₀、
 (A)₁₋₅₀、(E)₁₋₅₀(D)₁₋₅₀(C)₁₋₅₀(A)₁₋₅₀(B)₁₋₅₀、(E)₁₋₅₀(D)₁₋₅₀(B)₁₋₅₀(C)₁₋₅₀(A)₁₋₅₀、(E)₁₋₅₀(D)₁₋₅₀、
 (B)₁₋₅₀(A)₁₋₅₀(C)₁₋₅₀、(E)₁₋₅₀(D)₁₋₅₀(A)₁₋₅₀(B)₁₋₅₀(C)₁₋₅₀、(E)₁₋₅₀(D)₁₋₅₀(A)₁₋₅₀(C)₁₋₅₀(B)₁₋₅₀、
 (E)₁₋₅₀(A)₁₋₅₀(C)₁₋₅₀(B)₁₋₅₀(D)₁₋₅₀、(E)₁₋₅₀(A)₁₋₅₀(C)₁₋₅₀(D)₁₋₅₀(B)₁₋₅₀、(E)₁₋₅₀(A)₁₋₅₀(B)₁₋₅₀、
 (C)₁₋₅₀(D)₁₋₅₀、(E)₁₋₅₀(A)₁₋₅₀(B)₁₋₅₀(D)₁₋₅₀(C)₁₋₅₀、(E)₁₋₅₀(A)₁₋₅₀(D)₁₋₅₀(B)₁₋₅₀(C)₁₋₅₀、或
 (E)₁₋₅₀(A)₁₋₅₀(D)₁₋₅₀(C)₁₋₅₀(B)₁₋₅₀

[0081] 其中8-12个氨基酸的肽表位由“A”表示,13-17个氨基酸的肽表位由“B”表示,18-21个氨基酸的肽表位由“C”表示,22-26个氨基酸的肽表位由“D”表示,并且27-31个氨基酸的肽表位由“E”表示。

[0082] 可将前述任何肽表位组合进行组合。例如,本文所述的核酸癌症疫苗中的任一种均可编码所列肽表位组中的多于一个。

[0083] 在一些实施方案中,肽表位包含至少一个I类MHC表位和至少一个II类MHC表位。在一些实施方案中,肽表位的至少10%是I类MHC表位。在一些实施方案中,肽表位的至少20%是I类MHC表位。在一些实施方案中,肽表位的至少30%是I类MHC表位。在一些实施方案中,肽表位的至少40%是I类MHC表位。在一些实施方案中,肽表位的至少0%、60%、70%、80%、90%或100%是I类MHC表位。在一些实施方案中,肽表位中无(0%)是II类MHC表位。在一些实施方案中,肽表位的至少10%是II类MHC表位。在一些实施方案中,肽表位的至少20%是II类MHC表位。在一些实施方案中,肽表位的至少30%是II类MHC表位。在一些实施方案中,肽表位的至少40%是II类MHC表位。在一些实施方案中,肽表位的至少50%、60%、70%、80%、90%或100%是II类MHC表位。在一些实施方案中,I类MHC表位与II类MHC表位的比率是选自以下的比率:约10%:约90%;约20%:约80%;约30%:约70%;约40%:约60%;约50%:约50%;约60%:约40%;约70%:约30%;约80%:约20%;约90%:约10% I类MHC表位:II类MHC表位。在一个实施方案中,I类MHC表位:II类MHC表位的比率为1:1。在一个实施方案中,I类MHC表位:II类MHC表位的比率为2:1。在一个实施方案中,I类MHC表位:II类MHC表位的比率为3:1。在一个实施方案中,I类MHC表位:II类MHC表位的比率为4:1。在一个实施

方案中,I类MHC表位:II类MHC表位的比率为5:1。在一些实施方案中,II类MHC表位与I类MHC表位的比率是选自以下的比率:约10%:约90%;约20%:约80%;约30%:约70%;约40%:约60%;约50%:约50%;约60%:约40%;约70%:约30%;约80%:约20%;约90%:约10%的II类MHC表位:I类MHC表位。在一个实施方案中,II类MHC表位:I类MHC表位的比率为1:1。在一个实施方案中,II类MHC表位:I类MHC表位的比率为1:2。在一个实施方案中,II类MHC表位:I类MHC表位的比率为1:3。在一个实施方案中,II类MHC表位:I类MHC表位的比率为1:4。在一个实施方案中,II类MHC表位:I类MHC表位的比率为1:5。在一些实施方案中,癌症疫苗的肽表位中的至少一个是B细胞表位。在一些实施方案中,癌症疫苗的一个或多个预测的T细胞反应性表位包含介于8-11个之间的氨基酸。在一些实施方案中,癌症疫苗的一个或多个预测的B细胞反应性表位包含介于13-17个之间的氨基酸。

[0084] 在一些方面,本公开的癌症疫苗包括编码多个肽表位抗原的mRNA疫苗,所述多个肽表位抗原被排列成在肽表位之间有单氨基酸间隔子、在肽表位之间有短接头,或者在肽表位之间不具有间隔子的情况下直接彼此联接。所述多个表位抗原可包括I类MHC表位和II类MHC表位的混合物。作为非限制性实例,所述多个肽表位抗原可以是具有以下结构的多肽:

[0085] $(X-G-X)_{1-10} (G-Y-G-Y)_{1-10} (G-X-G-X)_{0-10} (G-Y-G-Y)_{0-10}$ 、 $(X-G)_{1-10} (G-Y)_{1-10} (G-X)_{0-10} (G-Y)_{0-10}$ 、 $(X-G-X-G-X)_{1-10} (G-Y-G-Y)_{1-10} (X-G-X)_{0-10} (G-Y-G-Y)_{0-10}$ 、 $(X-G-X)_{1-10} (G-Y-G-Y-G-Y)_{1-10} (X-G-X)_{0-10} (G-Y-G-Y)_{0-10}$ 、 $(X-G-X)_{0-10} (G-Y-G-Y)_{0-10}$ 、 $(X-G-X-G-X-G-X)_{1-10} (G-Y-G-Y)_{1-10} (X-G-X)_{0-10} (G-Y-G-Y)_{0-10}$ 、 $(X-G-X)_{1-10} (G-Y-G-Y-G-Y-G-Y)_{1-10} (X-G-X)_{0-10} (G-Y-G-Y)_{0-10}$ 、 $(X)_{1-10} (Y)_{1-10} (X)_{0-10} (Y)_{0-10}$ 、 $(Y)_{0-10} (Y)_{1-10} (X)_{1-10} (Y)_{0-10} (X)_{0-10}$ 、 $(XX)_{1-10} (Y)_{1-10} (X)_{0-10} (Y)_{0-10}$ 、 $(YY)_{1-10} (XX)_{1-10} (Y)_{0-10} (X)_{0-10}$ 、 $(X)_{1-10} (YY)_{1-10} (X)_{0-10} (Y)_{0-10}$ 、 $(XXX)_{1-10} (YYY)_{1-10} (XX)_{0-10} (YY)_{0-10}$ 、 $(YYY)_{1-10} (XXX)_{1-10} (YY)_{0-10} (XX)_{0-10}$ 、 $(XY)_{1-10} (Y)_{1-10} (X)_{1-10} (Y)_{1-10}$ 、 $(YX)_{1-10} (Y)_{1-10} (X)_{1-10} (Y)_{1-10}$ 、 $(YX)_{1-10} (X)_{1-10} (Y)_{1-10} (Y)_{1-10}$ 、 $(Y-G-Y)_{1-10} (G-X-G-X)_{1-10} (G-Y-G-Y)_{0-10} (G-X-G-X)_{0-10}$ 、 $(Y-G)_{1-10} (G-X)_{1-10} (G-Y)_{0-10} (G-X)_{0-10}$ 、 $(Y-G-Y-G-Y)_{1-10} (G-X-G-X)_{1-10} (Y-G-Y)_{0-10} (G-X-G-X)_{0-10}$ 、 $(Y-G-Y)_{1-10} (G-X-G-X-G-X)_{1-10} (Y-G-Y)_{0-10} (G-X-G-X)_{0-10}$ 、 $(Y-G-Y-G-Y-G-Y)_{1-10} (G-X-G-X)_{1-10} (Y-G-Y)_{0-10} (G-X-G-X)_{0-10}$ 、 $(Y-G-Y)_{1-10} (G-X-G-X-G-X-G-X)_{1-10} (Y-G-Y)_{0-10} (G-X-G-X)_{0-10}$ 、 $(XY)_{1-10} (YX)_{1-10} (XY)_{0-10} (YX)_{0-10}$ 、 $(YX)_{1-10} (XY)_{1-10} (Y)_{0-10} (X)_{0-10}$ 、 $(YY)_{1-10} (X)_{1-10} (Y)_{0-10} (X)_{0-10}$ 、 $(XY)_{1-10} (X)_{0-10} (X)_{0-10} (Y)_{1-10} (YX)_{1-10} (X)_{0-10} (Y)_{0-10}$ 、 $(XYX)_{1-10} (YXX)_{1-10} (YX)_{0-10} (YY)_{0-10}$ 、或 $(YYY)_{1-10} (XXY)_{1-10} (YX)_{0-10} (XY)_{0-10}$,

[0086] 其中X是长度为5-100个氨基酸(例如,包括8-31个氨基酸在内的本文所述的任何长度)的I类MHC表位,Y是长度为5-100个氨基酸(例如,包括8-31个氨基酸在内的本文所述的任何长度)的II类MHC表位,并且G是甘氨酸。

[0087] 在一些方面,本公开的核酸癌症疫苗包含编码一个或多个肽表位的核酸,所述一个或多个肽表位包含引起独特表达的肽序列的突变。在一些实施方案中,引起独特表达的肽序列的突变可以是但不限于插入、删除、移码突变和/或剪接变体。在一些实施方案中,核酸癌症疫苗编码多个肽表位抗原,所述多个肽表位抗原包含一个或多个单核苷酸多态性(SNP)突变,所述一个或多个单核苷酸多态性(SNP)突变在SNP突变的每一侧上都具有侧接氨基酸。在一些实施方案中,SNP突变的每一侧上的侧接氨基酸的数目可以是5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、22、24、26、28或30。在一些实施方案中,SNP突变位于

中心,并且SNP突变的每一侧上的侧接氨基酸的数目近似相同。在其他实施方案中,SNP突变在每一侧上不具有相等数量的侧接氨基酸。在一个实施方案中,癌症疫苗的表位包含侧接有两个I类序列的SNP,每个序列包含七个氨基酸。在另一个实施方案中,癌症疫苗的表位包含侧接有两个II类序列的SNP,每个序列包含10个氨基酸。在一些实施方案中,表位可包含位于中心的SNP和侧翼,所述侧翼都是I类序列,都是II类序列,或者一个是I类序列而一个是II类序列。

[0088] 在另一个实施方案中,肽表位呈由肽表位组成的多联癌症抗原的形式。可使用任何数目的肽表位。在某些实施方案中,肽表位呈由5-200个肽表位组成的多联癌症抗原的形式。在某些实施方案中,肽表位呈由3-130个肽表位组成的多联癌症抗原的形式。在一些实施方案中,所述多联癌症抗原包含以下中的一者或多者:a) 肽表位(例如,3-200或3-130个肽表位)被切割敏感性位点穿插;和/或b) 每个肽表位在无接头的情况下彼此直接联接;和/或c) 每个肽表位通过单一氨基酸接头彼此联接;和/或d) 每个肽表位通过短接头彼此联接;和/或e) 每个肽表位包含8-31个氨基酸,并且包含一个或多个SNP突变(例如,位于中心的SNP突变);和/或f) 每个肽表位包含8-31个氨基酸,并且包含引起独特表达的肽序列的突变;和/或g) 肽表位的至少30%对来自受试者的I类MHC分子具有最高亲和力;和/或h) 肽表位的至少30%对来自受试者的II类MHC分子具有最高亲和力;和/或i) 肽表位中没有一个对来自受试者的II类MHC分子具有最高亲和力;和/或j) 肽表位的至少50%对HLA-A、HLA-B和/或DRB1具有 $IC_{50} < 500nM$ 的预测结合亲和力;和/或k) 编码肽表位的核酸被排列成使得所述肽表位有序化以使假表位减到最少;1) I类MHC分子肽表位与II类MHC分子肽表位的比率为至少1:1、2:1、3:1、4:1或5:1;和/或m) 不存在II类MHC分子肽表位。在一些实施方案中,对I类MHC分子具有“最高亲和力”的肽表位特异性地结合(即,以最大亲和力结合)至该I类MHC分子。在一些实施方案中,对I类MHC分子具有“最高亲和力”的肽表位对该I类MHC分子的结合亲和力比对II类MHC分子的结合亲和力大。在一些实施方案中,对II类MHC分子具有“最高亲和力”的肽表位特异性地结合(即,以最大亲和力结合)至该II类MHC分子。在一些实施方案中,对II类MHC分子具有“最高亲和力”的肽表位对该II类MHC分子的结合亲和力比对I类MHC分子的结合亲和力大。

[0089] 应当理解,2种或更多种肽(例如2种或更多种新抗原)的多联体可在肽边界处产生非预期的新表位(假表位)。为了预防或消除此类假表位,可扫描I类等位基因在多联体中肽边界上的命中。在一些实施方案中,将多联体内的肽顺序打乱以减少或消除假表位形成。在一些实施方案中,在肽之间使用接头,例如单一氨基酸接头(诸如甘氨酸),以减少或消除假表位形成。在一些实施方案中,可将锚定氨基酸用其他氨基酸替代,这将减少或消除假表位形成。在一些实施方案中,在多联体内的肽边界处修剪肽以减少或消除假表位形成。

[0090] 在一些实施方案中,对多个肽表位抗原进行排列并使该多个肽表位抗原有序化以使假表位减到最少。在其他实施方案中,所述多个肽表位抗原是不含假表位的多肽。当癌症抗原表位是多联结构中头对尾阵型(head to tail formation)排列时,在每个癌症抗原表位之间形成接合点。这包含来自肽的N末端上的表位的若干个(即,1-10个)氨基酸和在相邻的直接联接的表位的C末端上的若干个(即,1-10个)氨基酸。重要的是,接合点不是可能产生免疫应答的免疫原性肽。在一些实施方案中,接合点形成了与受试者的HLA蛋白结合的肽序列,所述个性化癌症疫苗被设计成对所述受试者的HLA蛋白具有大于约50nM的 IC_{50} 。在

其他实施方案中,接合点肽序列以大于约10nM、150nM、200nM、250nM、300nM、350nM、400nM、450nM或500nM的 IC_{50} 与受试者的HLA蛋白结合。

[0091] 个性化癌症疫苗

[0092] 在一些方面,本公开提供了一种包含一种或多种核酸的核酸癌症疫苗,其中每种核酸编码至少一种合适的癌症抗原,诸如对癌症受试者具有特异性的个性化抗原。

[0093] 例如,核酸癌症疫苗可包含编码一种或多种对每名受试者具有特异性的癌症抗原(称为新表位)的核酸。在肿瘤细胞中表达或由肿瘤细胞表达的抗原被称为“肿瘤相关抗原”。特定的肿瘤相关抗原也可在,或者也可不在非癌性细胞中表达。许多肿瘤突变在本领域中是众所周知的。在非癌性细胞中不表达或很少表达,或者其在非癌性细胞中的表达相比在癌性细胞中的表达充分降低并且在接种疫苗后诱导免疫应答的肿瘤相关抗原被称为新表位。新表位对身体完全是外来的,因而不会产生针对健康组织的免疫应答,也不会被免疫系统的保护性组分掩盖。在一些实施方案中,基于新表位的个性化疫苗是合乎需要的,因此此类疫苗制剂将使针对患者的特定肿瘤的特异性最大化。突变源性新表位可由以下产生:点突变,导致蛋白质中的氨基酸不同的非同义突变;通读突变,其中终止密码子被修饰或删除从而导致在C末端处具有新型肿瘤特异性序列的较长蛋白质的翻译;剪接位点突变,其导致在成熟mRNA中包含内含子并因而包含独特的肿瘤特异性蛋白质序列;染色体重排,其在2种蛋白质的接合点处产生具有肿瘤特异性序列的嵌合蛋白(即,基因融合);移码突变或删除,其导致具有新型肿瘤特异性蛋白质序列的新的开放阅读框;和/或易位。

[0094] 用于产生个性化癌症疫苗的方法一般涉及鉴定突变(例如,使用深度核酸或蛋白质测序技术),鉴定新表位(例如,使用经验证的肽-MHC结合预测算法或其他分析技术的应用来产生可结合患者HLA等位基因并且基于肿瘤中存在的突变的候选T细胞表位集合),任选地证明抗原特异性T细胞针对所选择的新表位,或者证明候选新表位与肿瘤表面上的HLA蛋白结合,以及开发疫苗。用于鉴定突变的技术的实例包括但不限于动态等位基因特异性杂交(DASH)、微孔板阵列对角凝胶电泳(MADGE)、焦磷酸测序、寡核苷酸特异性结扎(oligonucleotide-specific ligation)、TaqMan系统以及各种DNA“芯片”技术(例如,Affymetrix SNP芯片),以及基于通过侵入性切割产生小信号分子、然后执行质谱法或固定锁式探针和滚环扩增的方法。

[0095] 若干深度核酸和蛋白质测序技术在本领域中是已知的。可使用任何类型的序列分析方法。例如,可对整个肿瘤基因组、肿瘤外显子组(编码蛋白质的DNA)和/或肿瘤转录组执行核酸测序。实时单分子边合成边测序技术依赖于对荧光核苷酸的检测,因为它们被并入到与待测序模板互补的新生DNA链中。还存在其他快速高通量测序方法。可对肿瘤蛋白质组执行蛋白质测序。另外,蛋白质质谱法可用于鉴定或验证与肿瘤细胞上的MHC蛋白结合的突变肽的存在。可将肽从肿瘤细胞或从肿瘤免疫沉淀的HLA分子中酸洗脱,然后使用质谱法对其进行鉴定。可将测序结果与已知的对照集合或者与对患者的正常组织执行的测序分析进行比较。在一些实施方案中,这些新表位以比野生型肽更大的亲和力与I类HLA蛋白结合,并且/或者能够激活抗肿瘤CD8 T细胞。在肿瘤上很少发现任何特定基因中的同一的突变。

[0096] I类MHC蛋白存在于身体的几乎所有细胞(包括大多数肿瘤细胞)的表面上。I类MHC蛋白负载有通常起源于内源性蛋白或起源于细胞内存在的病原体,并随后被呈递给细胞毒性T淋巴细胞(CTL)的抗原。T细胞受体能够辨识和结合与I类MHC分子复合的肽。每种细胞毒

性T淋巴细胞均表达能够结合特异性MHC/肽复合物的独特的T细胞受体。

[0097] 使用计算机算法,有可能预测潜在的新表位,诸如推定的T细胞反应性表位,即,以肽呈递复合物的形式被I类或II类MHC分子结合,并以后以这种形式被T淋巴细胞的T细胞受体辨识的肽序列。可用于鉴定将与MHC结合的肽的程序的实例包括例如:Lonza Epibase、SYFPEITHI (Rammensee等人,Immunogenetics,50 (1999),213-219) 和HLA_BIND (Parker等人,J.Immunol.,152 (1994),163-175)。

[0098] 一旦选择了推定的新表位,就可使用体外和/或体内测定对它们进行进一步测试。可将常规的体外实验室测定,诸如Elispot测定与来自每名患者的分离物一起使用,以精修基于算法的预测选择的新表位列表。

[0099] 在一些实施方案中,本文所述的核酸癌症疫苗和疫苗接种方法可包括基于特定突变的肽表位或抗原(新表位)以及由癌症种系基因表达的那些肽表位或抗原(在多名患者中发现的肿瘤所共同具有的抗原,在本文中称为“传统癌症抗原”或“共有癌症抗原”)。在一些实施方案中,传统抗原是已知普遍在癌症或肿瘤中被发现,或者在特定类型的癌症或肿瘤中被发现的抗原。在一些实施方案中,传统癌症抗原是非突变的肿瘤抗原。在一些实施方案中,传统癌症抗原是突变的肿瘤抗原。

[0100] 在一些实施方案中,本文所述的核酸癌症疫苗和方法可包括基于癌症/睾丸(CT)抗原的肽表位。癌症/睾丸抗原表达仅限于健康成年人中的雄性生殖细胞,但是在多种类型的人癌症的肿瘤细胞中观察到了异位表达。由于雄性生殖细胞缺乏I类HLA分子并且不能将抗原呈递给T细胞,因此癌症/睾丸抗原当在癌细胞中表达时一般被认为是新抗原,并且具有引发具有严格癌症特异性的免疫应答的能力。与本文所述的组合物和方法一起使用的癌症/睾丸抗原可以是本领域中已知的任何此类癌症/睾丸抗原,包括但不限于MAGEA1、MAGEA2、MAGEA3、MAGEA4、MAGEA5、MAGEA6、MAGEA8、MAGEA9、MAGEA10、MAGEA11、MAGEA12、BAGE、BAGE2、BAGE3、BAGE4、BAGE5、MAGEB1、MAGEB2、MAGEB5、MAGEB6、MAGEB3、MAGEB4、GAGE1、GAGE2A、GAGE3、GAGE4、GAGE5、GAGE6、GAGE7、GAGE8、SSX1、SSX2、SSX2b、SSX3、SSX4、CTAG1B、LAGE-1b、CTAG2、MAGEC1、MAGEC3、SYCP1、BRDT、MAGEC2、SPANXA1、SPANXB1、SPANXC、SPANXD、SPANXN1、SPANXN2、SPANXN3、SPANXN4、SPANXN5、XAGE1D、XAGE1C、XAGE1B、XAGE1、XAGE2、XAGE3、XAGE-3b、XAGE-4/RP11-167P23.2、XAGE5、DDX43、SAGE1、ADAM2、PAGE5、CT16.2、PAGE1、PAGE2、PAGE2B、PAGE3、PAGE4、LIPI、VENTXP1、IL13RA2、TSP50、CTAGE1、CTAGE-2、CTAGE5、SPA17、ACRBP、CSAG1、CSAG2、DSCR8、MMA1b、DDX53、CTCFL、LUZP4、CASC5、TFDP3、JARID1B、LDHC、MORC1、DKKL1、SP011、CRISP2、FMR1NB、FTHL17、NXF2、TAF7L、TDRD1、TDRD6、TDRD4、TEX15、FATE1、TPTE、CT45A1、CT45A2、CT45A3、CT45A4、CT45A5、CT45A6、HORMAD1、HORMAD2、CT47A1、CT47A2、CT47A3、CT47A4、CT47A5、CT47A6、CT47A7、CT47A8、CT47A9、CT47A10、CT47A11、CT47B1、SLC06A1、TAG、LEMD1、HSPB9、CCDC110、ZNF165、SPACA3、CXorf48、THEG、ACTL8、NLRP4、COX6B2、LOC348120、CCDC33、LOC196993、PASD1、LOC647107、TULP2、CT66/AA884595、PRSS54、RBM46、CT69/BC040308、CT70/B1818097、SPINLW1、TSSK6、ADAM29、CCDC36、LOC440934、SYCE1、CPXCR1、TSPY3、TSGA10、HIWI、MIWI、PIWI、PIWIL2、ARMC3、AKAP3、Cxorf61、PBK、C21orf99、OIP5、CEP290、CABYR、SPAG9、MPHOSPH1、ROPN1、PLAC1、CALR3、PRM1、PRM2、CAGE1、TTK、LY6K、IMP-3、AKAP4、DPPA2、KIAA0100、DCAF12、SEMG1、POTED、POTEE、POTEA、POTEG、POTEB、POTEC、POTEH、GOLGAGL2 FA、CDCA1、PEPP2、OTOA、CCDC62、GPATCH2、

CEP55、FAM46D、TEX14、CTNNA2、FAM133A、LOC130576、ANKRD45、ELOVL4、IGSF11、TMEFF1、TMEFF2、ARX、SPEF2、GPAT2、TMEM108、NOL4、PTPN20A、SPAG4、MAEL、RQCD1、PRAME、TEX101、SPATA19、ODF1、ODF2、ODF3、ODF4、ATAD2、ZNF645、MCAK、SPAG1、SPAG6、SPAG8、SPAG17、FBX039、RGS22、细胞周期素A1、C15orf60、CCDC83、TEKT5、NR6A1、TMPRSS12、TPPP2、PRSS55、DMRT1、EDAG、NDR、DNAJB8、CSAG3B、CTAG1A、GAGE12B、GAGE12C、GAGE12D、GAGE12E、GAGE12F、GAGE12G、GAGE12H、GAGE12I、GAGE12J、GAGE13、LOC728137、MAGEA2B、MAGEA9B/LOC728269、NXF2B、SPANXA2、SPANXB2、SPANXE、SSX4B、SSX5、SSX6、SSX7、SSX9、TSPY1D、TSPY1E、TSPY1F、TSPY1G、TSPY1H、TSPY1I、TSPY2、XAGE1E、XAGE2B/CTD-2267G17.3,和/或它们的变体。

[0101] 在一些实施方案中,核酸癌症疫苗可进一步包含一种或多种编码一种或多种非突变肿瘤抗原的核酸。在一些实施方案中,核酸癌症疫苗可进一步包含一种或多种编码一种或多种突变肿瘤抗原的核酸。

[0102] 许多肿瘤抗原在本领域中是已知的。用于与本文所述的组合物和方法一起使用的癌症或肿瘤抗原(例如,传统癌症抗原)可以是本领域中已知的任何此类癌症或肿瘤抗原。在一些实施方案中,癌症或肿瘤抗原(例如,传统癌症抗原)是以下抗原中的一种:CD2、CD19、CD20、CD22、CD27、CD33、CD37、CD38、CD40、CD44、CD47、CD52、CD56、CD70、CD79、CD137、4-1BB、5T4、AGS-5、AGS-16、血管生成素2、B2M、B7.1、B7.2、B7DC、B7H1、B7H2、B7H3、BT-062、BTLA、CAIX、癌胚抗原、CTLA4、Cripto、ED-B、ErbB1、ErbB2、ErbB3、ErbB4、EGFL7、EpCAM、EphA2、EphA3、EphB2、FAP、纤连蛋白、叶酸盐受体、神经节苷脂GM3、GD2、糖皮质激素诱导的肿瘤坏死因子受体(GITR)、gp100、gpA33、GPNMB、ICOS、IGF1R、整合素 α_v 、整合素 $\alpha_v\beta$ 、LAG-3、Lewis Y、间皮素、c-MET、MN碳酸酐酶IX、MUC1、MUC16、结合素-4(Nectin-4)、NKGD2、NOTCH、OX40、OX40L、PD-1、PDL1、PSCA、PSMA、RANKL、ROR1、ROR2、SLC44A4、多配体聚糖-1、TACI、TAG-72、腱生蛋白、TIM3、TRAILR1、TRAILR2、VEGFR-1、VEGFR-2、VEGFR-3,以及/或者它们的变体。

[0103] 可使用免费数据库或商业数据库(例如,Lonza Epibase,antitope)来鉴定表位。此类工具可用于预测靶抗原蛋白内最具免疫原性的表位。然后可在人HLA小组中合成和筛选所选择的肽,并且使用最具免疫原性的序列来构建编码一个或多个肽表位的核酸。用于绘制细胞毒性T细胞的表位的一种策略基于产生跨蛋白质的每种标称11聚体的四个C末端肽的等摩尔混合物。该策略将产生含有所有可能的活性CTL表位的文库抗原。

[0104] 新表位可被设计成最佳地结合至MHC,以便促进强烈的免疫应答。在一些实施方案中,每个肽表位均包含抗原区域和MHC稳定区域。MHC稳定区域是使MHC中的肽稳定的序列。

[0105] 表位内的所有MHC稳定区域可以是相同的或者它们可以是不同的。MHC稳定区域可位于肽的N末端部分或肽的C末端部分。替代地,MHC稳定区域可位于肽的中心区域中。

[0106] MHC稳定区域的长度可以是5-10个、5-15个、8-10个、1-5个、3-7个或3-8个氨基酸。在其他实施方案中,抗原区域的长度为5-100个氨基酸。肽在被呈递于细胞表面上之前,通过内质网内的竞争性亲和力结合与I类MHC分子相互作用。单个肽的亲和力与其氨基酸序列以及该氨基酸序列内的限定位置处特异性结合基序的存在直接相关。被呈递在MHC中的肽由肽结合沟的底部固定,位于 $\alpha 1/\alpha 2$ 异二聚体(由两个不同的亚单位构成的分子)的中心区域中。肽结合沟底部的残基序列决定了它结合的特定肽残基。

[0107] 可通过计算机辅助比较结合位点(MHC袋)对每个靶表位的结合位点中的每个氨基酸处的特定氨基酸的亲和力来鉴定最佳结合区域,以鉴定出对所有受检查抗原的理想结合

物。可使用结合位点的数据点的氨基酸预测矩阵来鉴定表位的MHC稳定区域。氨基酸预测矩阵是具有定义数据点的第一轴线和第二轴线的表格。预测矩阵可如Singh, H. 和Raghava, G.P.S. (2001), "ProPred: prediction of HLA-DR binding sites." *Bioinformatics*, 17 (12), 1236-37 中所示的那样来生成。在一些实施方案中, 预测矩阵基于进化保守性; 在另一个实施方案中, 预测矩阵使用生理化学相似性来检查体氨基酸与种系氨基酸的相似度 (例如, Kim等人, *J Immunol.* 2017: 3360-3368)。体氨基酸与种系氨基酸的相似性近似于突变如何影响结合 (例如, T细胞受体辨识)。在一些实施方案中, 较少的相似性指示改善的结合 (例如, T细胞受体辨识)。

[0108] 在一些实施方案中, 基于受试者的特定MHC来设计MHC稳定区域。以这种方式, 可针对每名患者优化MHC稳定区域。

[0109] 被选择用来包含在疫苗 (例如, 核酸癌症疫苗) 中的新表位将通常为高亲和力结合肽。在一些方面, 新表位以比野生型肽更大的亲和力结合HLA蛋白。在一些实施方案中, 新表位具有至少小于5000nM、至少小于500nM、至少小于250nM、至少小于200nM、至少小于150nM、至少小于100nM、至少小于50nM或更小的 IC_{50} 。通常, 具有 $<50nM$ 的预测 IC_{50} 的肽一般被认为是中至高亲和力的结合肽, 并且将被选择用于使用HLA结合的生物化学测定来凭经验测试它们的亲和力。最后, 将确定人免疫系统是否能够针对这些突变的肿瘤抗原产生有效的免疫应答, 从而有效地杀死肿瘤, 但不杀死正常细胞。

[0110] 在一些实施方案中, 新表位的长度为13个残基或更少, 并且该新表位可由介于约8个与约11个之间的残基 (特别是9个或10个残基) 组成。在其他实施方案中, 新表位可被设计得更长。例如, 新表位可能有2-5个氨基酸朝向每个相应基因产物的N末端和C末端延伸。使用较长的肽可允许患者细胞进行内源性加工, 并且可导致更有效的抗原呈递和对T细胞应答的诱导。

[0111] 可在必要时对具有所需活性的新表位进行修饰, 以提供某些所需的属性, 例如, 改善的药理学特征, 同时增加或至少保留未经修饰肽的基本上所有生物活性, 以结合所需的MHC分子并激活适当的T细胞或B细胞。例如, 新表位可经受各种变化, 诸如保守或非保守取代, 在这种情况下, 此类变化可能会在所述新表位的使用中提供某些优点, 诸如改善的MHC结合。所谓保守取代, 意指将一种氨基酸残基替换为另一种在生物学上和/或化学上相似的氨基酸残基, 例如, 将一种疏水性残基替换为另一种, 或将一种极性残基替换为另一种。所述取代包括诸如以下的组合: Gly、Ala; Val、Ile、Leu、Met; Asp、Glu; Asn、Gln; Ser、Thr; Lys、Arg; 和Phe、Tyr。也可使用D-氨基酸来探测单氨基酸取代的影响。此类修饰可使用如例如以下文献中所描述的众所周知的肽合成程序来进行: Merrifield, *Science* 232: 341-347 (1986), Barany和Merrifield, *The Peptides*, Gross和Meienhofer编著 (N.Y, Academic Press), 第1-284页 (1979); 以及Stewart和Young, *Solid Phase Peptide Synthesis*, (Rockford, Ill., Pierce), 第2版 (1984)。

[0112] 也可通过延长或减小化合物的氨基酸序列 (例如, 通过添加或删除氨基酸) 来修饰新表位。所述肽、多肽或类似物也可通过更改某些残基的顺序或组成来修饰, 容易理解的是, 某些对生物活性必不可少的氨基酸残基 (例如, 在关键接触位点处的那些或保守残基) 一般可能不会被更改, 而不会对生物活性产生不利影响。

[0113] 通常, 采用具有单氨基酸取代的一系列肽来确定静电荷、疏水性等对结合的影响。

例如,沿肽的长度进行一系列带正电荷(例如,LVs或Arg)或带负电荷(例如,Glu)的氨基酸取代,从而揭示对各种MHC分子和T细胞受体或B细胞受体的不同敏感模式。另外,可采用使用小的相对中性的部分诸如Ala、Gly、Pro或类似残基的多种取代。所述取代可以是同源寡聚体或异源寡聚体。被取代的或添加的残基的数目和类型取决于必需接触点之间所必需的间距和所寻求的某些功能属性(例如,疏水性对亲水性)。通过此类取代还可实现对MHC分子或T细胞受体的结合亲和力相比亲本肽的亲力的增加。在任何情况下,此类取代都应当采用所选择的氨基酸残基或其他分子片段,以避免例如可能破坏结合的空间干扰和电荷干扰。

[0114] 新表位还可包含新表位中的两个或更多个残基的等排物(isostere)。如本文所定义的等排物是可被第二序列取代的两个或更多个残基的序列,因为第一序列的空间构象与对第二序列具有特异性的结合位点相匹配。该术语具体包括本领域技术人员熟知的肽骨架修饰。此类修饰包括酰胺氮、 α 碳、酰胺羰基的修饰,酰胺键的完全替换,延伸,删除或骨架交联。一般来讲,参见Spatola, Chemistry and Biochemistry of Amino Acids, Peptides and Proteins, 第VII卷(Weinstein编著,1983)。

[0115] 免疫原性的考虑是选择最佳新表位以包含在疫苗中的重要组成部分。作为一组非限制性实例,可通过分析新表位的MHC结合能力、HLA混杂性(promiscuity)、突变位置、预测的T细胞反应性、实际的T细胞反应性、导致特定构象和由此导致的溶剂暴露的结构,以及具体氨基酸的表示来评定免疫原性。诸如NetMHC预测算法的已知算法可用于预测肽结合至共同的HLA-A和-B等位基因的能力。在某些实施方案中,NetMHC预测算法使用 IC_{50} 来确定结合能力。在其他实施方案中,NetMHC预测算法使用百分比排名和洗脱的配体数据来确定结合能力(Jurtz等人, J Immunol. 2017年11月1日; 199 (9): 3360-3368)。如图2至图3B所示,百分比排名方法导致预测结合物在不同HLA等位基因上的分布更加平衡。对MHC结合肽的结构评定也可通过经由电脑模拟(in silico)的三维分析和/或蛋白质对接程序来进行。与MHC分子结合时的预测表位结构(诸如从Rosetta算法获得的预测表位结构)的使用可用于评估当表位与MHC分子结合时该表位的氨基酸残基的溶剂暴露程度。可在体外用表位和T细胞来通过实验评定T细胞反应性。替代地,可使用T细胞应答/序列数据集来评定T细胞反应性。

[0116] 包含在疫苗中的新表位的一个重要方面是缺乏自身反应性。可筛选推定的新表位以证实该表位局限于肿瘤组织,例如,由于恶性细胞内的遗传性改变而出现。理想的是,该表位不应存在于患者的正常组织中,因而从数据集中滤出自相似表位。个性化编码基因组可作为参照用于比较新抗原候选物以确定是否缺乏自身反应性。在一些实施方案中,从个性化转录组和/或外显子组产生个性化编码基因组。

[0117] 肽组成性质也可在表位设计时加以考虑。例如,可以为每种推定表位提供关于在表位中发现的保守氨基酸与非保守氨基酸的值的得分。

[0118] 在一些实施方案中,通过本文所述的工具执行的的分析可包括比较在不同时间从患者(即,在治疗干预之前和之后)、从不同组织样品、从具有类似肿瘤的不同患者等获得的不同特性集合。在一些实施方案中,可将来自一个特性集合的峰值的平均值与来自另一个特性集合的峰值的平均值进行比较。例如,可在两个不同的分布集合之间比较HLA结合的平均值。该两个分布集合可以由例如以数天、数月或数年分隔的持续时间来确定。

[0119] 根据本文所述的技术的新表位表征系统可采用任何适合形式,因为实施方案在这

个方面不受限制。一个或多个计算机系统可用于实现上述任何功能性。所述计算机系统可包括一个或多个处理器以及一个或多个计算机可读存储介质(即,有形非暂时计算机可读介质),例如易失性存储器以及一个或多个可由任何适合的数据存储介质形成的非易失性存储介质。所述处理器可以任何适合方式控制向易失性存储器和非易失性存储装置写入数据以及从易失性存储器和非易失性存储装置读取数据,因为实施方案在这个方面不受限制。为执行任何本文所述的功能性,处理器可执行一个或多个存储在一个或多个计算机可读存储介质(例如,易失性存储器和/或非易失性存储器)中的指令,所述存储介质可充当存储用于由处理器执行的指令的有形非暂时计算机可读介质。

[0120] 制备方法

[0121] 在其他方面,本公开提供了一种用于制备癌症疫苗的方法,所述方法包括:a)针对患者在个性化癌症抗原之间进行鉴定;b)针对3-130种个性化癌症抗原中的每一种,确定至少两个肽表位的抗肿瘤功效;以及c)制备癌症疫苗,其中所述癌症疫苗的总抗癌功效(例如,所述癌症疫苗的预测总抗癌功效)在给定的所述癌症疫苗的总长度下最大化。

[0122] 根据本公开的用于产生个性化癌症疫苗的方法可涉及使用诸如如本文所述对组织样品进行的深度核酸或蛋白质测序方法的技术鉴定突变。在一些实施方案中,对受试者(例如,患者)转录组中的突变执行初始鉴定。将来自受试者(例如,患者)转录组的数据与来自受试者(例如,患者)外显子组的序列信息进行比较,以便鉴定表达的特定于患者的肿瘤特异性突变。所述比较产生推定新表位的数据集,其称为突变组(mutanome)。突变组可包括每名患者的近似100-10,000个候选突变。使用一组查询或算法对突变组进行数据探测分析,以鉴定用于产生新抗原疫苗的最佳突变集合。在一些实施方案中,设计并制造了mRNA新抗原疫苗。然后用疫苗治疗患者。在某些实施方案中,此种含新抗原的疫苗可以是包含多个新表位的多顺反子性疫苗,或者一种或多种单一RNA疫苗,或者它们的组合。

[0123] 在一些实施方案中,从突变鉴定过程的启动到开始治疗患者的整个方法在少于2个月内实现。在其他实施方案中,整个过程在7周或更短的时间、6周或更短的时间、5周或更短的时间、4周或更短的时间、3周或更短的时间、2周或更短的时间,或少于1周内实现。在一些实施方案中,整个方法在少于30天内执行。

[0124] 在个性化癌症疫苗中,可在患者样品中鉴定受试者特异性癌症抗原。术语“生物样品”是指包含诸如DNA、RNA和蛋白质的生物物质的样品。在一些实施方案中,生物样品可适合地包含来自受试者的体液。所述体液可以是来自受试者的身体中任何地方(优选是外周位置)分离的液体,包括但不限于例如血液、血浆、血清、尿液、痰液、脊髓液、脑脊髓液、胸膜液、乳头抽吸物、淋巴液、呼吸道液体、肠道液体和泌尿生殖道液体、泪液、唾液、母乳、来自淋巴系统的液体、精液、脑脊髓液、器官内系统液体、腹水液、肿瘤囊肿液体、羊膜液以及它们的组合。在一些实施方案中,样品可以是组织样品或肿瘤样品。例如,可检查具有一个或多个肿瘤细胞的样品中是否存在受试者特异性癌症抗原。

[0125] 特异性癌症抗原的鉴定过程可能涉及转录组分析和外显子组分析两者,或者仅涉及转录组分析或外显子组分析。在一些实施方案中,首先执行转录组分析,其次执行外显子组分析。对生物样品或组织样品执行分析。在一些实施方案中,生物样品或组织样品是血液或血清样品。在其他实施方案中,所述样品是组织库样品,或B细胞的EBV转化形式(EBV transformation)。

[0126] 替代地,可在受试者的外来体中鉴定受试者特异性癌症抗原。当在受试者的外来体中鉴定出用于疫苗的抗原时,此类抗原被称为受试者的外来体抗原的代表。

[0127] 外来体是由细胞排出的小型微囊泡,其通常具有近似30-100nm的直径。外来体典型地由如下方式形成:晚期内体膜向内凹入和夹断,从而导致形成载满小型脂质双层囊泡的多囊体(MVB),所述囊泡中的每一个均含有亲本细胞的细胞质的样品。MVB与细胞膜的融合导致这些外来体从细胞释放出来,并且被递送到血液、尿液、脑脊髓液或其他体液中。可从这些生物液体中的任一种中回收外来体以进行进一步分析。

[0128] 外来体内的核酸具有作为肿瘤抗原的生物标记物的作用。分析外来体以便鉴定受试者特异性癌症抗原的优势是该方法规避了对活检体的需要。这在患者需要接受包括癌症抗原鉴定和疫苗接种在内的数轮疗法时可特别有利。

[0129] 本领域中已描述了从生物样品分离外来体的许多方法。例如,可使用以下方法:差速离心、低速离心、阴离子交换和/或凝胶渗透色谱法、蔗糖密度梯度或细胞器电泳、磁激活细胞分选(MACS)、纳米膜超滤浓缩、Percoll梯度分离以及使用微流体装置。示例性方法例如描述于美国专利公布号2014/0212871中。

[0130] 一旦合成了mRNA疫苗,就将其施用于患者。在一些实施方案中,根据时间表施用疫苗长达两个月、长达三个月、长达四个月、长达五个月、长达六个月、长达七个月、长达八个月、长达九个月、长达十个月、长达十一个月、长达1年、长达1年半、长达两年、长达三年,或长达四年。时间表可相同或不同。在一些实施方案中,时间表是在前3周内每周施用,之后每月施用。

[0131] 在治疗的任何时候,都可对患者进行检查以确定疫苗中的突变是否仍然适当。基于该分析,可调整或重新配置疫苗,以使其包含一种或多种不同的突变或者以移除一种或多种突变。

[0132] 已经认识到并且理解,通过分析癌症相关突变的某些特性,可评估和/或选择最佳新表位以包含在癌症疫苗中。新表位或新表位集合的特性可包括例如:对患者RNA-Seq或其他核酸分析中的基因或转录物水平表达的评定;可用数据库中的组织特异性表达;已知的致癌基因/肿瘤阻抑因子;变体识别置信度得分;RNA-Seq等位基因特异性表达;保守与非保守氨基酸取代;点突变的位置(增大的TCR接合的中心得分);点突变的位置(差异HLA结合的锚定得分);自我性:与患者WES数据的核心表位同源性 $<100\%$,8聚体-11聚体的HLA-A和HLA-B IC_{50} ;15聚体-20聚体的HLA-DRB1 IC_{50} ;混杂性得分(即,预测结合的患者HLA的数目);8聚体-11聚体的HLA-C IC_{50} ;15聚体-20聚体的HLA-DRB3-5 IC_{50} ;15聚体-20聚体的HLA-DQB1/A1 IC_{50} ;15聚体-20聚体的HLA-DPB1/A1 IC_{50} ;I类与II类比例;涵盖的患者HLA-A、HLA-B和DRB1同种异型的多样性;点突变与复杂表位(例如,移码)的比例,和/或假表位HLA结合得分。

[0133] 在一些实施方案中,用于鉴定最佳新表位的癌症相关突变的特性是与突变类型、患者样品中的突变丰度、免疫原性、自身反应性缺乏以及肽组成的性质有关的特性。应当确定突变类型,并将其视为确定推定的表位是否应当被包含在疫苗中的一个因素。突变的类型可变化。在一些情况下,可能期望在单个疫苗中包含多种不同类型的突变。在其他情况下,可能更期望包含单一类型的突变。可对每个特定突变的值进行加权和计算。在一些实施方案中,特定突变是单核苷酸多态性(SNP)。在一些实施方案中,特定突变是复杂变体,例如

由内含子保留、复杂剪接事件,或改变序列的阅读框的插入/删除突变产生的肽序列。

[0134] 还可对患者样品中的突变的丰度进行评分,并在决定推定的表位是否应当被包含在疫苗中时将该丰度考虑进去。高度充足的突变可促进更强力的免疫应答。

[0135] 在一些实施方案中,本文所述的个性化mRNA癌症疫苗可用于治疗癌症。作为一个非限制性实例,本公开提供了用于治疗患有癌症的患者的方法,所述方法包括:a)分析源自患者的样品以便鉴定出一种或多种个性化癌症抗原;b)针对所述经鉴定的个性化癌症抗原中的每一种,确定至少两个肽表位的抗肿瘤功效;c)制备癌症疫苗,其中所述癌症疫苗的总抗癌功效(例如,所述癌症疫苗的预测总抗癌功效)在给定的所述癌症疫苗的总长度下最大化;以及d)向患者施用所述癌症疫苗。

[0136] 癌症疫苗(例如,核酸癌症疫苗)可作为主动免疫方案的一部分向健康个体或者在癌症早期或者晚期和/或转移性癌症期间防治性或治疗性地施用。在一个实施方案中,提供给细胞、组织或受试者的有效量的癌症疫苗(例如,核酸癌症疫苗)可足以用于免疫激活,特别是抗原特异性免疫激活。

[0137] 在一些实施方案中,癌症疫苗(例如,核酸癌症疫苗)可与抗癌治疗剂一起施用。可将癌症疫苗(例如,核酸癌症疫苗)和抗癌治疗剂结合,以进一步增强免疫治疗应答。癌症疫苗(例如,核酸癌症疫苗)和其他治疗剂可同时或相继施用。当同时施用其他治疗剂时,它们可以以同一制剂或分开的制剂施用,但是在同一时间施用。当其他治疗剂和癌症疫苗(例如,核酸癌症疫苗)的施用在时间上分开时,所述其他治疗剂彼此相继施用并且与癌症疫苗(例如,核酸癌症疫苗)相继施用。介于这些化合物的施用之间的时间间隔可以是大约几分钟,或者可以是更长时间,例如,几小时、几天、几周、几个月。其他治疗剂包括但不限于抗癌治疗剂、佐剂、细胞因子、抗体、抗原等。

[0138] 在一些实施方案中,可监测癌症的进展以鉴定所表达抗原的变化。因此,在一些实施方案中,所述方法还涉及在施用癌症mRNA疫苗之后至少一个月,从受试者的样品鉴定至少2种癌症抗原以产生第二癌症抗原集合,以及向所述受试者施用具有编码所述第二癌症抗原集合的开放阅读框的mRNA疫苗。在一些实施方案中,在具有编码第一癌症抗原集合的开放阅读框的mRNA疫苗之后2个月、3个月、4个月、5个月、6个月、8个月、10个月或1年,向受试者施用具有编码第二抗原集合的开放阅读框的mRNA疫苗。在其他实施方案中,在具有编码第一癌症抗原集合的开放阅读框的mRNA疫苗之后1年半、2年、2年半、3年、3年半、4年、4年半或5年,向受试者施用具有编码第二抗原集合的开放阅读框的mRNA疫苗。

[0139] 作为新抗原的热点突变

[0140] 在对癌症的群体分析中,某些突变在患者中的发生百分率比偶然预期的要高。这些“频发”或“热点”突变已常常被证明在肿瘤中具有“驱动器”作用,从而在癌细胞功能方面产生对肿瘤引发、维持或转移重要,并且因此被选择用于肿瘤的进化中的一些变化。除了在肿瘤生物学和疗法中具有重要性之外,频发突变还为精准医疗提供了机会,在精准医疗中,患者群体被分为更可能对特定疗法起响应的各组,所述疗法包括但不限于靶向突变蛋白自身。

[0141] 因此,在一些实施方案中,除个性化癌症表位之外,癌症疫苗还进一步包含一种或多种癌症热点新表位。在一些实施方案中,疫苗中包含在目标适应症中超过阈值发生率发生的癌症热点突变。在一些实施方案中,阈值发生率大于2%、3%、4%、5%、6%、7%、8%、

9%或10%。目标适应症包括但不限于膀胱尿路上皮癌 (BLCA)、结肠腺癌 (COAD)、食道癌 (ESCA)、肝细胞癌 (HCC)、头颈鳞状细胞癌 (HNSC)、肺腺癌 (LUAD)、胰腺腺癌 (PAAD)、前列腺腺癌 (PRAD)、直肠腺癌 (READ)、小细胞肺癌 (SCLC)、皮肤皮肤黑素瘤 (SKCM)、浆液性卵巢癌 (SOC)、胃腺癌 (STAD) 和子宫内膜癌 (UEC)。下表中提供了示例性突变,并且适应症的热点突变的示例性图如图1所提供。

[0142]	基因	突变位置
	KRAS	G12、G13
	NRAS	Q61
	BRAF	V600
	PIK3CA	R88、E545、H1047
	TP53	R175、R282
	EGFR	L858
	FGFR3	S249
	ERBB2	S310
	PTEN	R130
	BCOR	N1459

[0144] 关于频发突变的许多努力和研究已集中于非同义(或“错义”)单一核苷酸变体 (SNV),但群体分析已揭示多种更复杂(非SNV)变体分类,诸如同义(或“沉默”)、剪接位点、多核苷酸变体、插入和删除,也可在高频率下发生。

[0145] 在人癌症的情况下,相比于任何其他基因,p53基因(官方符号TP53)更频繁地突变。大量群组研究已证明对于大多数p53突变,基因组位置是一名患者或仅少许患者所特有的,并且该突变不能用作针对特定患者群体设计的治疗性疫苗的频发新抗原。然而,令人惊讶的是,p53基因座的小型子集确实表现出“热点”样式,其中基因中的若干位置以相对高频率发生突变。引人注目的是,这些频发突变区域中的大部分发生在外显子-内含子边界附近,从而破坏由mRNA剪接机构辨识的规范的核苷酸序列基序。剪接基序的突变可更改最终mRNA序列,即使没有预测到局部氨基酸序列的变化(即,对于同义突变或内含子突变来说)。因此,这些突变常常被常见的注释工具注释为“非编码”,并且在进一步分析时被忽略,即使它们可以不可预测方式更改mRNA剪接并且对所翻译的蛋白质造成严重功能性影响。如果选择性剪接同种型产生框内序列变化(即,不产生PTC),那么它可逃脱由NMD所致的耗竭,并且容易被HLA系统表达、加工,以及呈递在细胞表面上。进一步地,突变源性选择性剪接通常是“隐蔽的”,即不在正常组织中表达,因此可被T细胞辨识为非自身新抗原。

[0146] 突变通常是从患者的DNA测序数据获得,以衍生出现有肽疫苗的新表位。然而,mRNA表达是对可能的表位的全局空间的更直接的量度。例如,一些肿瘤特异性新表位可能由剪接变化、导致移码的插入/删除(InDel)、替代性启动子,或仅使用外显子组测序数据不容易鉴定的表观遗传修饰引起。在一些方面,相对于nsSNV,由InDel产生的新抗原对于预测的高亲和力和结合物是富集的。此类新抗原可以是免疫原性的。例如,在三个黑色素瘤群组中,移码InDel被发现与检查点抑制剂应答显著相关。可与那些由SNV产生的新表位相同的方式对所有新表位进行评分,但是每个InDel最多包含一种新抗原候选物,以便避免偏向于

InDel。鉴定用于新抗原疫苗的这些类型的复杂突变有待开发的價值，因为它们将增大能够结合患者独有的HLA同种异型的表位的数量。此外，与由单一氨基酸变化产生的变体相比，复杂变体将具有更高的免疫原性并可能引起更有效的针对肿瘤的免疫应答，这是由于它们与自身蛋白存在差异。

[0147] 在一些方面，本发明涉及一种用于鉴定特定于患者的复杂突变并将这些突变配制成有效的个性化癌症疫苗（例如，核酸癌症疫苗）的方法。所述方法涉及使用短读段RNA-Seq。使用短RNA-Seq读段固有的一个主要挑战是，由于可变剪接和其他机制，可从同一基因组位点获得多个mRNA转录物同种型。由于测序读段比全长mRNA转录物短得多，因此很难将读段集合映射回已知基因注释模型内的正确相应同种型。结果是，可能难以通过标准途径发现与已知基因注释（如在癌症中常见的已知基因注释）有趋异的复杂变体。然而，可鉴定短肽，而不是全长转录物的确切外显子组合物。鉴定将代表这些复杂突变的短肽的方法涉及针对复杂变体的新表位预测的短k聚体计数途径。

[0148] 核酸/多核苷酸

[0149] 如本文所提供的癌症疫苗（例如，核酸癌症疫苗）包含至少一种（一种或多种）具有编码至少一个肽表位的开放阅读框的核酸。术语“核酸”以它的最广泛意义包括包含核苷酸的聚合物的任何化合物和/或物质。这些聚合物也被称为多核苷酸。

[0150] 核酸可以是或者可包括例如核糖核酸（RNA）、脱氧核糖核酸（DNA）、苏糖核酸（TNA）、二醇核酸（GNA）、肽核酸（PNA）、锁核酸（LNA，包括具有 β -D-核糖构型的LNA、具有 α -L-核糖构型的 α -LNA（LNA的非对映异构体）、具有2'-氨基官能化的2'-氨基-LNA和具有2'-氨基官能化的2'-氨基- α -LNA）、亚乙基核酸（ENA）、环己烷基核酸（CeNA）或它们的嵌合体或组合。

[0151] 作为非限制性实例，当将如本文所述的DNA核酸癌症疫苗递送至细胞时，DNA被转录成RNA，RNA将由细胞内机制加工成多肽，所述细胞内机制然后将该多肽加工成能够刺激针对肿瘤或癌性细胞群的免疫应答的免疫敏感性片段。作为非限制性实例，当将如本文所述的RNA（例如，mRNA）核酸癌症疫苗递送至细胞时，所述RNA（例如，mRNA）将由细胞内机制加工成多肽，所述细胞内机制然后将该多肽加工成能够刺激针对肿瘤或癌性细胞群的免疫应答的免疫敏感性片段。

[0152] 在一些实施方案中，本公开的核酸充当信使RNA（mRNA）。“信使RNA”（mRNA）是指编码（至少一种）多肽（天然存在的、非天然存在的或经修饰的氨基酸聚合物），并且可在体外、在体内、原位或离体被翻译从而产生所编码多肽的任何核酸。

[0153] mRNA分子的基本组分通常包括至少一个编码区域、5'非翻译区域（UTR）、3'UTR、5'帽和聚腺苷酸尾部。本公开的核酸可充当mRNA，但可在它们的功能性和/或结构性设计特征方面区别于野生型mRNA，所述功能性和/或结构性设计特征用于克服使用基于核酸的治疗剂进行有效多肽表达的现有问题。

[0154] 在一些实施方案中，本公开的多核苷酸是经密码子优化的。密码子优化方法在本领域中是已知的，并且可如本文所提供的那样加以使用。在一些实施方案中，密码子优化可用于使靶标中的密码子频率和宿主生物体匹配以确保适当折叠；偏向GC含量以增加mRNA稳定性或减少二级结构；使可损害基因构建或表达的串联重复密码子或碱基延伸减到最少；定制转录和翻译控制区域；插入或移除蛋白质移行序列；在所编码蛋白质中移除/添加翻译

后修饰位点(例如,糖基化位点);添加、移除或打乱蛋白质结构域;插入或删除限制性位点;修饰核糖体结合位点和mRNA降解位点;调整翻译速率以使蛋白质的各种结构域适当折叠;或者减少或消除多核苷酸内的问题二级结构。密码子优化工具、算法和服务在本领域中是已知的-非限制性实例包括来自GeneArt (Life Technologies) 的服务、DNA2.0 (Menlo Park CA) 和/或专有方法。在一些实施方案中,使用优化算法使开放阅读框(ORF)序列优化。

[0155] 在一些实施方案中,密码子优化的序列与天然存在的或野生型序列(例如,编码目标多肽或蛋白质(例如,抗原性蛋白质或多肽)的天然存在的或野生型mRNA序列)共有小于95%的序列同一性。在一些实施方案中,密码子优化的序列与天然存在的或野生型序列(例如,编码目标多肽或蛋白质(例如,抗原性蛋白质或多肽)的天然存在的或野生型mRNA序列)共有小于90%的序列同一性。在一些实施方案中,密码子优化的序列与天然存在的或野生型序列(例如,编码目标多肽或蛋白质(例如,抗原性蛋白质或多肽)的天然存在的或野生型mRNA序列)共有小于85%的序列同一性。在一些实施方案中,密码子优化的序列与天然存在的或野生型序列(例如,编码目标多肽或蛋白质(例如,抗原性蛋白质或多肽)的天然存在的或野生型mRNA序列)共有小于80%的序列同一性。在一些实施方案中,密码子优化的序列与天然存在的或野生型序列(例如,编码目标多肽或蛋白质(例如,抗原性蛋白质或多肽)的天然存在的或野生型mRNA序列)共有小于75%的序列同一性。

[0156] 在一些实施方案中,密码子优化的序列与天然存在的或野生型序列(例如,编码目标多肽或蛋白质(例如,抗原性蛋白质或多肽)的天然存在的或野生型mRNA序列)共有介于65%与85%之间(例如,介于约67%与约85%之间,或介于约67%与约80%之间)的序列同一性。在一些实施方案中,密码子优化的序列与天然存在的或野生型序列(例如,编码目标多肽或蛋白质(例如,抗原性蛋白质或多肽)的天然存在的或野生型mRNA序列)共有介于65%与75%之间或约80%的序列同一性。

[0157] 在一些实施方案中,密码子优化的RNA可例如是其中G/C的水平得以增强的RNA。核酸分子的G/C含量可影响RNA的稳定性。相比于含有大量腺嘌呤(A)和胸腺嘧啶(T)或尿嘧啶(U)核苷酸的核酸,具有增加量的鸟嘌呤(G)和/或胞嘧啶(C)残基的RNA可能在功能上更稳定。WO02/098443公开了一种含有通过翻译区域中的序列修饰稳定化的mRNA的药物组合物。归因于遗传密码的简并性,所述修饰通过将现有密码子替换成促进更大RNA稳定性而不改变所得氨基酸的那些来起作用。该方法限于RNA的编码区域。

[0158] 抗原/抗原性多肽

[0159] 在一些实施方案中,每个肽表位的长度可以是5-100个氨基酸(包括端点)。在一些实施方案中,所述肽表位中的至少一个的长度为5-100个、5-95个、5-90个、5-85个、5-80个、5-75个、5-70个、5-65个、5-60个、5-55个、5-50个、5-45个、5-40个、5-39个、5-38个、5-37个、5-36个、5-35个、5-34个、5-33个、5-32个、5-31个、5-30个、5-29个、5-28个、5-27个、5-26个、5-25个、5-24个、5-23个、5-22个、5-21个、5-20个、8-100个、8-95个、8-90个、8-85个、8-80个、8-75个、8-70个、8-65个、8-60个、8-55个、8-50个、8-45个、8-40个、8-39个、8-38个、8-37个、8-36个、8-35个、8-34个、8-33个、8-32个、8-31个、8-30个、8-29个、8-28个、8-27个、8-26个、8-25个、8-24个、8-23个、8-22个、8-21个、8-20个、10-100个、10-95个、10-90个、10-85个、10-80个、10-75个、10-70个、10-65个、10-60个、10-55个、10-50个、10-45个、10-40个、10-39个、10-38个、10-37个、10-36个、10-35个、10-34个、10-33个、10-32个、10-31个、10-30

个、10-29个、10-28个、10-27个、10-26个、10-25个、10-24个、10-23个、10-22个、10-21个,或10-20个氨基酸。

[0160] 在一些实施方案中,由核酸癌症疫苗编码的肽表位中的每一个均可具有不同的长度。在某些实施方案中,所述肽表位中的至少一个具有与由核酸癌症疫苗编码的另一个肽表位的长度不同的长度。每个肽表位可以是对于表位合理的任何长度。

[0161] 用于与本公开一起使用的多肽包括基因产物,天然存在的多肽,合成多肽,同源物,直系同源物,旁系同源物,片段以及前述各物的其他等效物、变体和类似物。多肽可以是单个分子或者可以是多分子复合物,诸如二聚体、三聚体或四聚体。多肽还可包含单链或多链多肽,诸如抗体或胰岛素,并且可加以缔合或联接。最通常地,二硫键联见于多链多肽中。术语多肽也可适用于氨基酸聚合物,其中至少一个氨基酸残基是相应天然存在的氨基酸的人工化学类似物。

[0162] 术语“多肽变体”是指在其氨基酸序列方面不同于原生或参照序列的分子。相较于原生序列或参照序列,氨基酸序列变体可在氨基酸序列内的某些位置处具有取代、删除和/或插入。正常情况下,变体与原生序列或参照序列具有至少50%同一性。在一些实施方案中,变体与原生序列或参照序列共有至少80%或至少90%的同一性。

[0163] 在一些实施方案中,提供了“变体模拟物”。如本文所用,术语“变体模拟物”是含有至少一个将模拟活化序列的氨基酸的模拟物。例如,谷氨酸盐可充当磷酸苏氨酸和/或磷酸丝氨酸的模拟物。替代地,变体模拟物可导致去活化,或导致含有所述模拟物的产物失活,例如,苯丙氨酸可充当酪氨酸的失活替代物;或者丙氨酸可充当丝氨酸的失活替代物。

[0164] “直系同源物”是指不同物种中通过物种形成而从共同祖先基因进化而来的基因。正常情况下,直系同源物在进化的过程中保留相同功能。鉴定直系同源物对于可靠预测新测序基因组中的基因功能至关重要。

[0165] “类似物”意欲包括仍然维持亲本或起始多肽的一种或多种特性,差异在于一个或多个氨基酸更改(包括例如氨基酸残基的取代、添加或删除)的多肽变体。

[0166] 本公开提供了基于多核苷酸或多肽的若干类型的组合物,包括变体和衍生物。这些包括例如取代性、插入性、删除性和共价性的变体和衍生物。术语“衍生物”与术语“变体”同义地使用,但一般是指已相对于参照分子或起始分子以任何方式被修饰和/或改变的分子。

[0167] 因此,在本公开的范围内包括编码相对于参照序列含有取代、插入和/或添加、删除和共价修饰的肽或多肽(特别是本文所公开的多肽序列)的多核苷酸。例如,可将序列标签或氨基酸诸如一个或多个赖氨酸添加至肽序列中(例如,在N末端或C末端处)。序列标签可用于肽检测、纯化或定位。赖氨酸可用于增加肽溶解度或允许生物素化。替代地,位于肽或蛋白质的氨基酸序列的羧基和氨基末端区域处的氨基酸残基可任选地被删除,从而提供截短序列。某些氨基酸(例如,C末端残基或N末端残基)可替代地被删除,这取决于该序列的用途,如例如所述序列作为可溶性或联接于固体支撑物的较大序列的一部分进行表达。

[0168] 在提及多肽时的“取代性变体”是将原生或起始序列中的至少一个氨基酸残基去除并且在同一位置处的位置中插入了不同的氨基酸而得到的变体。取代可以是单一的,其中分子中仅一个氨基酸已被取代,或者取代可以是多重的,其中同一分子中的两个或更多个氨基酸已被取代。

[0169] 如本文所用,术语“保守氨基酸取代”是指正常情况下存在于序列中的氨基酸被具有类似大小、电荷或极性的不同氨基酸取代。保守取代的实例包括非极性(疏水性)残基诸如异亮氨酸、缬氨酸和亮氨酸对另一种非极性残基的取代。同样,保守取代的实例包括一种极性(亲水性)残基对另一种极性(亲水性)残基的取代,诸如在精氨酸与赖氨酸之间、在谷氨酰胺与天冬酰胺之间以及在甘氨酸与丝氨酸之间的取代。另外,碱性残基诸如赖氨酸、精氨酸或组氨酸对另一种碱性残基的取代,或者一种酸性残基诸如天冬氨酸或谷氨酸对另一种酸性残基的取代是保守取代的另外的实例。非保守取代的实例包括非极性(疏水性)氨基酸残基(诸如异亮氨酸、缬氨酸、亮氨酸、丙氨酸、甲硫氨酸)对极性(亲水性)残基(诸如半胱氨酸、谷氨酰胺、谷氨酸或赖氨酸)的取代,以及/或者极性残基对非极性残基的取代。

[0170] 在提及多肽或多核苷酸时的“特征”分别被定义为分子的不同的基于氨基酸序列或基于核苷酸的组分。由多核苷酸编码的多肽的特征包括表面表现形式(surface manifestation)、局部构象形状、折叠、环、半环、结构域、半结构域、位点、末端或它们的任何组合。

[0171] 如本文在提及多肽时所用,术语“结构域”是指多肽的具有一种或多种可鉴定的结构或功能特征或特性(例如,结合能力、充当蛋白质-蛋白质相互作用的位点)的基序。

[0172] 如本文在提及多肽时所用,术语“位点”在它涉及基于氨基酸的实施方案时与“氨基酸残基”和“氨基酸侧链”同义地使用。如本文在提及多核苷酸时所用,术语“位点”在它涉及基于核苷酸的实施方案时与“核苷酸”同义地使用。位点代表肽或多肽或多核苷酸内的可在基于该多肽或多核苷酸的分子内被修饰、操作、更改、衍生化或改变的位置。

[0173] 如本文所用,在提及多肽或多核苷酸时的术语“末端(termini/terminus)”分别是指多肽或多核苷酸的尽头。此种尽头不仅仅限于多肽或多核苷酸的第一位点或最终位点,而是可包括在末端区域中的另外的氨基酸或核苷酸。基于多肽的分子可被表征为具有N末端(以具有游离氨基(NH_2)的氨基酸终止)与C末端(以具有游离羧基(COOH)的氨基酸终止)两者。蛋白质在一些情况下由通过二硫键或通过非共价力(多聚体、寡聚物)聚集在一起的多个多肽链组成。这些蛋白质具有多个N末端和C末端。替代地,所述多肽的末端可被修饰成使得它们根据情况以基于非多肽的部分(诸如有机缀合物)开始或结束。

[0174] 如由本领域技术人员所认识到,蛋白质片段、功能性蛋白质结构域和同源性蛋白质也被视为在目标多肽的范围内。例如,本文提供了长度为5、10、20、30、40、50、60、70、80、90、100或大于100个氨基酸的参照蛋白质的任何蛋白质片段(意指比参照多肽序列短至少一个氨基酸残基,但在其他方面同一的多肽序列)。在另一实例中,可根据本公开来利用包含具有与本文所述的任何序列40%、50%、60%、70%、80%、90%、95%或100%同一的10个、20个、30个、40个、50个或100个氨基酸的链段的任何蛋白质。在一些实施方案中,多肽包含如本文所提供或提及的任何序列中所示的2、3、4、5、6、7、8、9、10或更多种突变。在另一个实例中,可根据本公开来利用包含具有与本文所述的任何序列大于80%、90%、95%或100%同一的20个、30个、40个、50个或100个氨基酸的链段的任何蛋白质,其中所述蛋白质具有与本文所述的任何序列小于80%、75%、70%、65%或60%同一的5个、10个、15个、20个、25个或30个氨基酸的链段。

[0175] 本公开的多肽或多核苷酸分子可与参照分子(例如,参照多肽或参照多核苷酸),例如与本领域所描述的分子(例如,经工程改造或设计的分子或者野生型分子)共有某一程

度的序列相似性或同一性。如本领域中已知的术语“同一性”是指两个或更多个多肽或多核苷酸(例如,DNA分子和/或RNA分子)的序列之间的关系,如通过将序列进行比较所确定。在本领域中,同一性也意指它们之间的序列相关性程度,如通过两个或更多个氨基酸残基或核酸残基的串段之间的匹配的数目所确定。同一性衡量两个或更多个序列之间相对于较小序列的同一匹配百分比,其中采用通过特定数学模型或计算机程序(例如“算法”)处理的空位比对(如果有的话)。相关肽的同一性可通过已知方法容易地计算。“同一性百分比”或“同一性%”在它适用于多肽或多核苷酸序列时被定义为在比对序列以及必要时引入空位以实现最大同一性百分比之后,候选氨基酸或核酸序列中与第二序列的氨基酸序列或核酸序列中的残基同一的残基(氨基酸残基或核酸残基)的百分比。用于比对的方法和计算机程序在本领域中是众所周知的。应当理解同一性取决于对同一性百分比的计算,但可能会由于在计算中引入的空位和罚分而在数值方面不同。对两个多核酸序列的同一性百分比的计算例如可通过出于最优比较目的比对该两个序列来进行(例如可在第一核酸序列和第二核酸序列中的一者或两者中引入空位以达成最优比对,并且可出于比较目的而忽视非同序列)。在某些实施方案中,出于比较目的而比对的序列的长度为参照序列的长度的至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%、至少95%或100%。然后比较相应核苷酸位置处的核苷酸。当第一序列中的位置被与第二序列中的相应位置相同的核苷酸占据时,那么在该位置处的分子是同一的。该两个序列之间的同一性百分比是在考虑需要被引入以实现两个序列的最优比对的空位的数目和每个空位的长度的情况下,该序列共有的同一位置的数目的函数。序列的比较和两个序列之间同一性百分比的确定可使用数学算法来完成。

[0176] 一般来讲,特定多核苷酸或多肽的变体与那个特定参照多核苷酸或多肽具有至少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%但小于100%序列同一性,如通过本文所述以及为本领域技术人员所知的序列比对程序和参数所确定。例如,两个核酸序列之间的同一性百分比可使用诸如以下中所述的那些方法的方法确定:Computational Molecular Biology,Lesk,A.M.编,Oxford University Press,New York,1988;Biocomputing:Informatics and Genome Projects,Smith,D.W.编,Academic Press,New York,1993;Sequence Analysis in Molecular Biology,von Heinje,G.,Academic Press,1987;Computer Analysis of Sequence Data,第I部分,Griffin,A.M.和Griffin,H.G.编,Humana Press,New Jersey,1994;以及Sequence Analysis Primer,Gribskov,M.和Devereux,J.编,M Stockton Press,New York,1991;所述文献中的每一篇均以引用方式并入本文。例如,可使用Meyers和Miller(CABIOS,1989,4:11-17)的算法来确定两个核酸序列之间的同一性百分比,所述算法已被并入ALIGN程序(2.0版)中,所述程序使用PAM120权重残基表、空位长度罚分12和空位罚分4。两个核酸序列之间的同一性百分比可替代地使用GCG软件包中的GAP程序利用NWSgapdha.CMP矩阵来确定。通常用于确定序列之间的同一性百分比的方法包括但不限于Carillo,H.和Lipman,D.,SIAM J Applied Math.,48:1073(1988)中公开的那些;所述文献以引用方式并入本文。用于确定同一性的技术被编纂在可公开获得的计算机程序中。用于确定两个序列之间的同源性的示例性计算机软件包括但不限于GCG程序包(Devereux,J.等人,Nucleic Acids Research,12(1),387(1984))、BLASTP、BLASTN和FASTA(Stephen

F.Altschul等人(1997)，“Gapped BLAST and PSI-BLAST:a new generation of protein database search programs”,Nucleic Acids Res.25:3389-3402)。另一普及局部比对技术基于史密斯-沃特曼算法(Smith-Waterman algorithm)(Smith,T.F.和Waterman,M.S.(1981)“Identification of common molecular subsequences.”J.Mol.Biol.147:195-197)。一种基于动态规划的一般性整体比对技术是尼德曼-翁施算法(Needleman-Wunsch algorithm)(Needleman,S.B.和Wunsch,C.D.(1970)“A general method applicable to the search for similarities in the amino acid sequences of two proteins.”J.Mol.Biol.48:443-453)。更新近地,已开发出快速最优整体序列比对算法(Fast Optimal Global Sequence Alignment Algorithm,FOGSAA),据称该算法比包括尼德曼-翁施算法在内的其他最优整体比对方法更快产生对核苷酸和蛋白质序列的整体比对。

[0177] 如本文所用,术语“同源性”是指聚合分子之间,例如核酸分子(例如,DNA分子和/或RNA分子)之间和/或多肽分子之间的总体相关性。共有通过比对匹配残基来确定的阈值水平的相似性或同一性的聚合分子(例如,核酸分子(例如,DNA分子和/或RNA分子)和/或多肽分子)被称为同源。同源性是描述分子之间的关系的定性术语,并且可基于定量相似性或同一性。相似性或同一性是定义两个所比较序列之间的序列匹配程度的定量术语。在一些实施方案中,如果聚合分子的序列至少25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或99%同一或相似,那么它们被视为彼此“同源”。术语“同源”必须是指至少两个序列(多核苷酸或多肽序列)之间的比较。如果两个多核苷酸序列编码的多肽就至少一个具有至少20个氨基酸的链段来说至少50%、60%、70%、80%、90%、95%或甚至99%,那么该两个多核苷酸序列被视为同源。在一些实施方案中,同源多核苷酸序列的特征在于具有编码具有至少4-5个独特指定的氨基酸的链段的能力。对于长度小于60个核苷酸的多核苷酸序列,同源性通过编码具有至少4-5个独特指定的氨基酸的链段的能力来确定。如果两个蛋白质就至少一个具有至少20个氨基酸的链段来说至少50%、60%、70%、80%或90%同一,那么该两个蛋白质序列被视为同源。

[0178] 同源性暗示所比较序列从共同起源趋异演化。术语“同源物”是指由于从共同祖先序列继承而与第二氨基酸序列或核酸序列相关的第一氨基酸序列或核酸序列(例如,基因(DNA或RNA)或蛋白质序列)。术语“同系物”可适用于通过物种形成事件而分开的基因和/或蛋白质之间的关系,或者适用于通过遗传重复事件而分开的基因和/或蛋白质之间的关系。“直系同源物”是不同物种中通过物种形成而从共同祖先基因(或蛋白质)进化的基因(或蛋白质)。通常,直系同源物在进化的过程中保留相同功能。“旁系同源物”是由于基因组内的重复而相关的基因(或蛋白质)。直系同源物在进化的过程中保留相同功能,而旁系同源物进化出新功能,即使这些功能与原始功能相关。

[0179] 化学修饰

[0180] 编码表位抗原多肽的经修饰的核苷酸序列

[0181] 在一些实施方案中,本发明的核酸癌症疫苗包含一种或多种经化学修饰的核碱基。本发明包括包含本文所述的多核苷酸的经修饰的多核苷酸(例如,包含编码一个或多个癌症肽表位的核苷酸序列的核酸)。经修饰的核酸可被化学修饰和/或结构修饰。当本发明的核酸被化学和/或结构修饰时,多核苷酸可被称为“经修饰的核酸”。

[0182] 本公开提供了编码一个或多个癌症肽表位的核酸的经修饰的核苷和核苷酸(例

如, RNA多核苷酸, 诸如mRNA多核苷酸)。“核苷”是指含有糖分子(例如, 戊糖或核糖)或其衍生物与有机碱基(例如, 嘌呤或嘧啶)或其衍生物(在本文中也称为“核碱基”)的组合的化合物。“核苷酸”是指包含磷酸酯基团的核苷。经修饰的核苷酸可通过用于包含一个或多个经修饰的或非天然的核苷的任何可用的方法(诸如例如化学、酶促或重组方法)来合成。核酸可包含联接的核苷的一个或多个区域。此类区域可具有可变主链键联。所述键联可以是标准磷酸二酯键联, 在这种情况下多核苷酸将包含核苷酸的区域。

[0183] 本文所公开的经修饰的核酸可包含各种不同的修饰。在一些实施方案中, 经修饰的多核苷酸含有一个、两个或更多个(任选地不同的)核苷或核苷酸修饰。在一些实施方案中, 引入到细胞中的经修饰的多核苷酸可表现出一种或多种期望的特性, 诸如, 例如与未经修饰的多核苷酸相比, 改善蛋白质表达、降低免疫原性或者降低细胞中的降解。

[0184] 在一些实施方案中, 本文所公开的核酸(例如, 编码一个或多个肽表位的核酸)是经结构修饰的。如本文所用, “结构”修饰是这样的修饰, 其中两个或更多个联接核苷被插入、删除、复制、反转或随机化在多核苷酸中而对所述核苷酸本身没有显著的化学修饰。因为要实现结构修饰就必须使化学键断裂并重新形成, 所以结构修饰具有化学性质并因此是化学修饰。然而, 结构修饰将产生不同的核苷酸序列。例如, 多核苷酸“ATCG”可被化学修饰成“AT-5meC-G”。同一多核苷酸可从“ATCG”被结构修饰成“ATCCCG”。在这里, 二核苷酸“CC”已被插入, 从而对核酸进行了结构修饰。

[0185] 在一些实施方案中, 本公开的多核苷酸是经化学修饰的。如本文针对核酸所使用的, 术语“化学修饰”或在适当时“化学修饰的”是指关于腺苷(A)、鸟苷(G)、尿苷(U)或胞苷(C)核糖核苷或脱氧核糖核苷的在它们的位置、模式、百分比或群体中的一者或多者上进行的修饰。一般来讲, 在本文中, 这些术语不意图指天然存在的5'末端mRNA帽部分中的核糖核苷酸修饰。

[0186] 在一些实施方案中, 本公开的核酸可具有全部或任何相同核苷类型的均匀化学修饰, 或通过在全部分或任何相同核苷类型中的相同起始修饰的仅仅向下滴定而产生的修饰群体, 或全部或任何相同核苷类型的测量百分比的化学修饰但伴有随机并入, 诸如其中所有尿苷被尿苷类似物(例如, 假尿苷或5-甲氧基尿苷)替换。在另一个实施方案中, 所述多核苷酸可具有对整个多核苷酸中的两种、三种或四种相同核苷类型的均匀化学修饰(诸如所有尿苷和所有胞嘧啶等以相同方式被修饰)。

[0187] 经修饰的核苷酸碱基配对不仅涵盖标准腺苷-胸腺嘧啶、腺苷-尿嘧啶或鸟苷-胞嘧啶碱基对, 而且还涵盖在核苷酸和/或包含非标准或经修饰的碱基的经修饰的核苷酸之间形成的碱基对, 其中氢键供体和氢键接受体的排列容许在非标准碱基与标准碱基之间或者在两个互补非标准碱基结构之间进行氢键合。此种非标准碱基配对的一个实例是经修饰的核苷酸肌苷与腺嘌呤、胞嘧啶或尿嘧啶之间的碱基配对。碱基/糖或接头的任何组合都可并入到本公开的多核苷酸中。

[0188] 本领域技术人员将理解, 除非另有说明, 否则本申请中列出的核酸序列将列举出代表性DNA序列中的“T”, 但是当该序列表示RNA时, “T”将被“U”取代。

[0189] 在一些实施方案中, 本公开的癌症疫苗包含至少一种具有编码至少一个(例如, 3-200或3-130个)肽表位的开放阅读框的核酸(例如, RNA), 其中所述核酸包含核苷酸和/或核苷, 所述核苷酸和/或核苷可以是标准的(未经修饰的)或如本领域已知的那样经修饰的。在

一些实施方案中,本公开的核苷酸和核苷包括经修饰的核苷酸或核苷。此类经修饰的核苷酸和核苷可以是天然存在的经修饰的核苷酸和核苷或者非天然存在的经修饰的核苷酸和核苷。此类修饰可包括如本领域所认可的在核苷酸和/或核苷的糖、主链或核碱基部分处的那些修饰。

[0190] 在一些实施方案中,本公开的天然存在的经修饰的核苷酸或核苷是如本领域普遍已知或认可的那种天然存在的经修饰的核苷酸或核苷。此类天然存在的经修饰的核苷酸和核苷的非限制性实例尤其可在广泛认可的MODOMICS数据库中找到。

[0191] 在一些实施方案中,本公开的非天然存在的经修饰的核苷酸或核苷是如本领域普遍已知或认可的那种非天然存在的经修饰的核苷酸或核苷。此类非天然存在的经修饰的核苷酸和核苷的非限制性实例尤其可在公布的美国申请号PCT/US2012/058519;PCT/US2013/075177;PCT/US2014/058897;PCT/US2014/058891;PCT/US2014/070413;PCT/US2015/36773;PCT/US2015/36759;PCT/US2015/36771;或PCT/IB2017/051367中找到;所述文献全部出于此目的以引用方式并入本文。

[0192] 因此,本公开的核酸(例如,DNA核酸和RNA核酸,诸如mRNA核酸)可包含标准核苷酸和核苷、天然存在的核苷酸和核苷、非天然存在的核苷酸和核苷,或它们的任何组合。

[0193] 在一些实施方案中,本公开的核酸(例如,DNA核酸和RNA核酸,诸如mRNA核酸)包含各种(超过一种)不同类型的标准和/或经修饰的核苷酸和核苷。在一些实施方案中,核酸的特定区域含有一种、两种或更多种(任选地不同)类型的标准和/或经修饰的核苷酸和核苷。

[0194] 在一些实施方案中,相对于包含标准核苷酸和核苷的未修饰核酸,引入到细胞或生物体中的经修饰的RNA核酸(例如,经修饰的mRNA核酸)分别在细胞或生物体中表现出降解降低。

[0195] 在一些实施方案中,相对于包含标准核苷酸和核苷的未修饰核酸,引入到细胞或生物体中的经修饰的RNA核酸(例如,经修饰的mRNA核酸)可分别在细胞或生物体中表现出免疫原性降低(例如,先天性应答降低)。

[0196] 在一些实施方案中,核酸(例如,RNA核酸,诸如mRNA核酸)包含在核酸的合成期间或合成后被引入以实现所需功能或特性的非天然的经修饰的核苷酸。该修饰可存在于核苷酸间键联、嘌呤或嘧啶碱基或糖上。该修饰可用化学合成或用聚合酶引入在链的末端或链中任何其他地方。核酸的任何区域都可被化学修饰。

[0197] 本公开提供了核酸(例如,DNA核酸或RNA核酸,诸如mRNA核酸)的经修饰的核苷和核苷酸。“核苷”是指含有糖分子(例如,戊糖或核糖)或其衍生物与有机碱基(例如,嘌呤或嘧啶)或其衍生物(在本文中也称为“核碱基”)的组的化合物。“核苷酸”是指包含磷酸酯基团的核苷。经修饰的核苷酸可通过用于包含一个或多个经修饰的或非天然的核苷的任何可用的方法(诸如例如化学、酶促或重组方法)来合成。核酸可包含联接的核苷的一个或多个区域。此类区域可具有可变主链键联。所述键联可以是标准磷酸二酯键联,在这种情况下核酸将包含核苷酸的区域。

[0198] 经修饰的核苷酸碱基配对不仅涵盖标准腺苷-胸腺嘧啶、腺苷-尿嘧啶或鸟苷-胞嘧啶碱基对,而且还涵盖在核苷酸和/或包含非标准或经修饰的碱基的经修饰的核苷酸之间形成的碱基对,其中氢键供体和氢键接受体的排列容许在非标准碱基与标准碱基之间或者在两个互补非标准碱基结构之间进行氢键合,诸如例如在具有至少一种化学修饰的那些

多核苷酸中。此种非标准碱基配对的一个实例是经修饰的核苷酸肌苷与腺嘌呤、胞嘧啶或尿嘧啶之间的碱基配对。碱基/糖或接头的任何组合都可并入到本公开的核酸中。

[0199] 在一些实施方案中,核酸(例如, RNA核酸, 诸如mRNA核酸)中的经修饰的核碱基包含1-甲基-假尿苷(m1ψ)、1-乙基-假尿苷(e1ψ)、5-甲氧基-尿苷(mo5U)、5-甲基-胞苷(m5C)和/或假尿苷(ψ)。在一些实施方案中,核酸(例如, RNA核酸, 诸如mRNA核酸)中的经修饰的核碱基包含5-甲氧基甲基尿苷、5-甲基硫代尿苷、1-甲氧基甲基假尿苷、5-甲基胞苷和/或5-甲氧基胞苷。在一些实施方案中,多核糖核苷酸包含至少两个(例如, 2个、3个、4个或更多个)任何前述经修饰的核碱基(包括但不限于化学修饰)的组合。

[0200] 在一些实施方案中,本公开的RNA核酸在所述核酸的一个或多个或所有尿苷位置处包含1-甲基-假尿苷(m1ψ)取代。

[0201] 在一些实施方案中,本公开的RNA核酸在所述核酸的一个或多个或所有尿苷位置处包含1-甲基-假尿苷(m1ψ)取代,并且在所述核酸的一个或多个或所有胞苷位置处包含5-甲基胞苷取代。

[0202] 在一些实施方案中,本公开的RNA核酸在所述核酸的一个或多个或所有尿苷位置处包含假尿苷(ψ)取代。

[0203] 在一些实施方案中,本公开的RNA核酸在所述核酸的一个或多个或所有尿苷位置处包含假尿苷(ψ)取代,并且在所述核酸的一个或多个或所有胞苷位置处包含5-甲基胞苷取代。

[0204] 在一些实施方案中,本公开的RNA核酸在所述核酸的一个或多个或所有尿苷位置处包含尿苷。

[0205] 在一些实施方案中,核酸(例如, RNA核酸, 诸如mRNA核酸)就特定修饰来说被均匀修饰(例如, 被完全修饰, 在整个序列上被修饰)。例如, 核酸可用1-甲基-假尿苷均匀修饰, 这意味着mRNA序列中的所有尿苷残基都用1-甲基-假尿苷替换。类似地, 可通过用经修饰的残基(诸如上面阐述的那些)替换来针对序列中存在的任何类型的核苷残基对核酸进行均匀修饰。

[0206] 本公开的核酸可沿该分子的整个长度被部分修饰或完全修饰。例如, 在本公开的核酸中, 或在其预定序列区域中(例如, 在包含或排除聚腺苷酸尾部的mRNA中), 一种或多种或所有或给定类型的核苷酸(例如嘌呤或嘧啶, 或A、G、U、C中的任何一者或多者或全部)可被均匀修饰。在一些实施方案中, 本公开的核酸(或其序列区域中)的所有核苷酸X都是经修饰的核苷酸, 其中X可以是核苷酸A、G、U、C中的任一者, 或组合A+G、A+U、A+C、G+U、G+C、U+C、A+G+U、A+G+C、G+U+C或A+G+C中的任一者。

[0207] 所述核酸可含有约1%至约100%或任何居间百分比(例如, 1%至20%、1%至25%、1%至50%、1%至60%、1%至70%、1%至80%、1%至90%、1%至95%、10%至20%、10%至25%、10%至50%、10%至60%、10%至70%、10%至80%、10%至90%、10%至95%、10%至100%、20%至25%、20%至50%、20%至60%、20%至70%、20%至80%、20%至90%、20%至95%、20%至100%、50%至60%、50%至70%、50%至80%、50%至90%、50%至95%、50%至100%、70%至80%、70%至90%、70%至95%、70%至100%、80%至90%、80%至95%、80%至100%、90%至95%、90%至100%、以及95%至100%)的经修饰的核苷酸(相对于总体核苷酸含量或者相对于一种或多种类型的核苷酸, 即, A、G、U或C中的任何一

者或多者)。应当理解,任何剩余百分比均是由未修饰A、G、U或C的存在引起的。

[0208] 所述核酸可含有最低1%且最高100%的经修饰的核苷酸,或任何居间百分比,诸如至少5%的经修饰的核苷酸、至少10%的经修饰的核苷酸、至少25%的经修饰的核苷酸、至少50%的经修饰的核苷酸、至少80%的经修饰的核苷酸,或至少90%的经修饰的核苷酸。例如,所述核酸可含有经修饰的嘧啶,诸如经修饰的尿嘧啶或胞嘧啶。在一些实施方案中,所述核酸中至少5%、至少10%、至少25%、至少50%、至少80%、至少90%或100%的尿嘧啶被经修饰的尿嘧啶(例如,5取代的尿嘧啶)替换。经修饰的尿嘧啶可被具有单一独特结构的化合物替换,或者可被具有不同结构(例如,2个、3个、4个或更多个独特结构)的多种化合物替换。在一些实施方案中,所述核酸中至少5%、至少10%、至少25%、至少50%、至少80%、至少90%或100%的胞嘧啶被经修饰的胞嘧啶(例如,5-取代的胞嘧啶)替换。经修饰的胞嘧啶可被具有单一独特结构的化合物替换,或者可被具有不同结构(例如,2个、3个、4个或更多个独特结构)的多种化合物替换。

[0209] 在一些实施方案中,所述核酸可包含在核苷之间的任何有用的接头。可用于本公开组合物中的此类接头(包括主链修饰)包括但不限于以下:3'-亚烷基磷酸酯、3'-氨基氨基磷酸酯、含烯烃的主链、氨基烷基氨基磷酸酯、氨基烷基磷酸三酯、硼烷磷酸酯、-CH₂-O-N(CH₃)-CH₂-、-CH₂-N(CH₃)-N(CH₃)-CH₂-、-CH₂-NH-CH₂-、手性磷酸酯、手性硫代磷酸酯、甲酰基主链和硫代甲酰基主链、亚甲基(甲基亚氨基)、亚甲基甲酰基主链和硫代甲酰基主链、亚甲基亚氨基主链和亚甲基胍基主链、吗啉代键联、-N(CH₃)-CH₂-CH₂-、具有杂原子核苷间键联的寡核苷、次磷酸酯、氨基磷酸酯、二硫代磷酸酯、硫代磷酸酯核苷间键联、硫代磷酸酯、磷酸三酯、PNA、硅氧烷主链、氨基磺酸酯主链、硫化亚砷主链和硫化砷主链、磺酸酯主链和磺酰胺主链、硫羰基烷基磷酸酯、硫羰基烷基磷酸三酯以及硫羰基氨基磷酸酯。

[0210] 可并入到核酸(例如, RNA或mRNA,如本文所述)中的经修饰的核苷和核苷酸(例如, 构件分子)可在核糖核酸的糖上被修饰。例如, 2' 羟基(OH)可被许多个不同的取代基修饰或替换。2' 位置处的示例性取代包括但不限于H、卤基、任选取代的C₁₋₆烷基;任选取代的C₁₋₆烷氧基;任选取代的C₆₋₁₀芳氧基;任选取代的C₃₋₈环烷基;任选取代的C₃₋₈环烷氧基;任选取代的C₆₋₁₀芳氧基;任选取代的C₆₋₁₀芳基-C₁₋₆烷氧基,任选取代的C₁₋₁₂(杂环基)氧基;糖(例如,核糖、戊糖或本文所述的任何糖);聚乙二醇(PEG), -O(CH₂CH₂O)_nCH₂CH₂OR, 其中R是H或任选取代的烷基,并且n是0至20(例如, 0至4、0至8、0至10、0至16、1至4、1至8、1至10、1至16、1至20、2至4、2至8、2至10、2至16、2至20、4至8、4至10、4至16以及4至20)的整数;“锁”核酸(LNA), 其中2'-羟基通过C₁₋₆亚烷基或C₁₋₆杂亚烷基桥连接至同一核糖的4'-碳, 其中示例性桥包括亚甲基、亚丙基、醚或氨基桥;氨基烷基;氨基烷氧基;氨基;以及氨基酸。

[0211] 一般来讲, RNA包括糖基核糖, 所述糖基核糖是具有氧的5元环。示例性的非限制性的经修饰的核苷酸包括核糖中的氧的替换(例如, 用S、Se或亚烷基诸如亚甲基或亚乙基替换);添加双键(例如, 以用环戊烯基或环己烯基替换核糖);核糖的环缩(例如, 以形成环丁烷或氧杂环丁烷的4元环);核糖的扩环(例如, 以如对于失水己糖醇、阿卓糖醇、甘露醇、环己烷基、环己烯基以及吗啉代来说形成具有另外的碳或杂原子的6元或7元环, 其也具有氨基磷酸酯主链);多环形式(例如, 三环;以及“非锁”形式, 诸如二醇核酸(GNA)(例如, R-GNA或S-GNA, 其中核糖被联接至磷酸二酯键的二醇单元替换)、苏糖核酸(TNA, 其中核糖被α-L-苏型呋喃糖基-(3'→2')替换), 和肽核酸(PNA, 其中2-氨基-乙基-甘氨酸键联替换核糖和

磷酸二酯主链)。糖基还可含有具有与核糖中相应碳的立体化学构型相反的立体化学构型的一个或多个碳。因此,多核苷酸分子可包含含有例如阿拉伯糖作为糖的核苷酸。此类糖修饰描述于例如国际专利公布号W02013052523和W02014093924中,其中每一篇的内容以引用方式整体并入本文。

[0212] 本公开的核酸(例如,编码一个或多个肽表位的核酸或其功能片段或变体)可包括对糖、核碱基和/或核苷间键联的修饰的组合。这些组合可包括本文所述的任何一个或多个修饰。

[0213] 本文所公开的核酸癌症疫苗是组合物,包括药物组合物。本公开还涵盖用于选择、设计、制备、制造、配制和/或使用如本文所提供的核酸癌症疫苗的方法。还提供了用于选择、设计和/或利用本文所述的核酸癌症疫苗的系统(例如,计算机化系统)、工艺、装置和试剂盒。

[0214] RNA(例如mRNA)的体外转录

[0215] 本公开的癌症疫苗可包含至少一种核酸(例如,RNA多核苷酸,诸如mRNA(信使RNA)或mmRNA(经修饰的mRNA))。mRNA例如是在体外从被称为“体外转录模板”的模板DNA转录的。在一些实施方案中,体外转录模板编码5'非翻译(UTR)区域,含有开放阅读框,并且编码3'UTR和聚腺苷酸尾部。体外转录模板的特定核酸序列组成和长度将取决于由模板编码的mRNA。

[0216] 在一些实施方案中,核酸包含15至3,000个核苷酸。例如,多核苷酸可包含15至50个、15至100个、15至200个、15至300个、15至400个、15至500个、15至600个、15至700个、15至800个、15至900个、15至1000个、15至1200个、15至1400个、15至1500个、15至1800个、15至2000个、15至2500个、15至3000个、50至100个、50至200个、50至300个、50至400个、50至500个、50至600个、50至700个、50至800个、50至900个、50至1000个、50至1200个、50至1400个、50至1500个、50至1800个、50至2000个、50至2500个、50至3000个、100至200个、100至300个、100至400个、100至500个、100至600个、100至700个、100至800个、100至900个、100至1000个、100至1200个、100至1400个、100至1500个、100至1800个、100至2000个、100至2500个、100至3000个、200至300个、200至400个、200至500个、200至600个、200至700个、200至800个、200至900个、200至1000个、200至1500个、200至3000个、500至1000个、500至1500个、500至2000个、500至2500个、500至3000个、1000至1500个、1000至2000个、1000至2500个、1000至3000个、1500至3000个、2500至3000个,或2000至3000个核苷酸。

[0217] 在其他方面,本公开涉及通过IVT方法来制备核酸癌症疫苗(例如,mRNA癌症疫苗)的方法。体外转录(IVT)方法容许具有几乎任何序列的RNA分子的模板导向合成。可使用IVT方法合成的RNA分子的大小在从短寡核苷酸至具有数千碱基的长核酸聚合物的范围内。IVT方法容许合成大量RNA转录物(例如,从微克至毫克数量)。参见Beckert等人,Synthesis of RNA by in vitro transcription,Methods Mol Biol.703:29-41(2011);Rio等人RNA: A Laboratory Manual.Cold Spring Harbor:Cold Spring Harbor Laboratory Press,2011,205-220.;Cooper,Geoffery M.The Cell:A Molecular Approach.第4版Washington D.C.:ASM Press,2007.262-299,所述文献中的每一篇出于此目的以引用方式并入本文。一般来讲,IVT利用特征在于在目标序列上游具有启动子序列的DNA模板。该启动子序列最通常起源于噬菌体(例如,T7、T3或SP6启动子序列),但是包括从头(de novo)设计的启动子序

列在内的许多其他启动子序列都可被允许。对DNA模板的转录通常通过使用对应于特定噬菌体启动子序列的RNA聚合酶来最佳地实现。示例性RNA聚合酶包括但不限于T7 RNA聚合酶、T3 RNA聚合酶或SP6 RNA聚合酶等等。IVT一般在dsDNA处被引发,但可在单链上进行。

[0218] 应当理解,可使用任何适当的合成方法来制备本公开的核酸癌症疫苗(例如,mRNA癌症疫苗),例如编码癌症抗原的mRNA。例如,在一些实施方案中,使用IVT由作为模板的单一底链DNA和充当启动子的互补寡核苷酸来制备本公开的mRNA疫苗。单一底链DNA可充当体外RNA转录的DNA模板,并且可从例如质粒、PCR产物或化学合成获得。在一些实施方案中,单一底链DNA是从环状模板线性化的。单一底链DNA模板一般包括启动子序列,例如噬菌体启动子序列,以有助于IVT。使用单一底链DNA和顶链启动子互补寡核苷酸制备RNA的方法在本领域中是已知的。示例性方法包括但不限于使DNA底链模板与顶链启动子互补寡核苷酸(例如,T7启动子互补寡核苷酸、T3启动子互补寡核苷酸或SP6启动子互补寡核苷酸)退火,随后使用对应于启动子序列的RNA聚合酶,例如T7 RNA聚合酶、T3 RNA聚合酶或SP6 RNA聚合酶进行IVT。

[0219] IVT方法还可使用双链DNA模板来进行。例如,在一些实施方案中,通过使用本领域中可获得的链延伸技术使互补寡核苷酸延伸以产生互补DNA链来制备双链DNA模板。在一些实施方案中,使含有启动子序列和编码一个或多个目标肽表位的序列的单一底链DNA模板退火于顶链启动子互补寡核苷酸,并且经受PCR样过程以使顶链延伸来产生双链DNA模板。替代地或另外,使含有互补于底链启动子序列以及互补于编码一个或多个目标肽表位的序列的顶链DNA退火于底链启动子寡核苷酸,并且经受PCR样过程以使底链延伸来产生双链DNA模板。在一些实施方案中,PCR样循环的数目在1至20个循环,例如3至10个循环的范围内。在一些实施方案中,双链DNA模板完全或部分地通过化学合成方法合成。可使双链DNA模板经受如本文所述的体外转录。

[0220] 在另一方面,包含例如编码肽表位的mRNA的本公开的核酸癌症疫苗可使用在序列的重叠部分上互补的两条DNA链来制备,从而当使互补部分退火时,留下单链突出端(即,粘性末端)。可通过使用另一链作为模板进行延伸来将这些单链突出端制成双链,由此产生双链DNA。在一些情况下,例如相较于通过顶链DNA合成方法获得的并入到模板DNA序列中的大小,此引物延伸方法可容许较大ORF并入到模板DNA序列中。在引物延伸方法中,第一链(呈5' -3' 方向)的3' 端的一部分互补于第二链(呈3' -5' 方向)的3' 端的一部分。在一些此类实施方案中,单一第一链DNA可包括启动子(例如,T7、T3或SP6)的序列,任选地包括5' -UTR,以及ORF的一些或全部(例如,ORF的5' 端的一部分)。在一些实施方案中,单一第二链DNA可包括ORF的一些或全部的互补序列(例如,互补于ORF的3' 端的一部分),以及任选地包括3' -UTR、终止序列和/或聚腺苷酸尾部。使用两条合成DNA链制备RNA的方法可包括使具有重叠互补部分的两条链退火,随后使用一个或多个PCR样循环进行引物延伸以使链延伸来产生双链DNA模板。在一些实施方案中,PCR样循环的数目在1至20个循环,例如3至10个循环的范围内。可使此种双链DNA经受如本文所述的体外转录。

[0221] 在另一方面,可使用作为双链DNA模板的合成双链线性DNA分子诸如gBlocks® (Integrated DNA Technologies, Coralville, Iowa)来制备包含例如编码肽表位的mRNA的本公开的核酸疫苗。此类合成双链线性DNA分子的优势是它们提供产生mRNA所依据的较长

模板。例如，gBlocks[®]的大小可在45-1000个（例如，125-750个核苷酸）的范围内。在一些实施方案中，合成双链线性DNA模板包括全长5' UTR、全长3' UTR或两者。全长5' UTR的长度可多达100个核苷酸，例如约40-60个核苷酸。全长3' UTR的长度可多达300个核苷酸，例如约100-150个核苷酸。

[0222] 为有助于产生较长构建体，被设计在3' 链上具有重叠序列的两个或更多个双链线性DNA分子和/或基因片段可使用本领域中已知的方法组装在一起。例如，可按如下方式进行GibsonAssembly[™] Method (Synthetic Genomics, Inc., La Jolla, CA)：使用从双链DNA片段的5' 端切合碱基的中温核酸外切酶，随后使新形成的互补单链3' 端退火，进行聚合酶依赖性延伸以填充任何单链空位，以及最终通过DNA连接酶来使DNA区段接合。

[0223] 在另一方面，可使用RNA的化学合成来制备包含例如编码肽表位的mRNA的本公开的核酸癌症疫苗。方法例如涉及使包含编码多肽的开放阅读框的第一多核苷酸和包含5' UTR的第二多核苷酸退火于与固体支撑物缀合的互补多核苷酸。接着使第二多核苷酸的3' 末端在适合条件下结扎(ligated)至第一多核苷酸的5' 末端上。适合条件包括使用DNA连接酶。结扎反应产生第一结扎产物。接着使包含3' UTR的第三多核苷酸的5' 末端在适合条件下结扎至第一结扎产物的3' 末端上。第二结扎反应的适合条件包括RNA结扎酶。在第二结扎反应中产生第二结扎产物。使第二结扎产物从固体支撑物释放以产生编码目标多肽的mRNA。在一些实施方案中，该mRNA在30个与1000个核苷酸之间。

[0224] 编码一个或多个肽表位的mRNA也可通过使包含编码核酸的开放阅读框的第一核酸结合至包含3' UTR的第二核酸以及结合至与固体支撑物缀合的互补核酸来制备。使第二核酸的5' 末端在适合条件（包括例如DNA结扎酶）下结扎至第一核酸的3' 末端上。所述方法产生第一结扎产物。使包含5' UTR的第三核酸在适合条件下（包括，例如RNA结扎酶，诸如T4 RNA）结扎至第一结扎产物上以产生第二结扎产物。使第二结扎产物从固体支撑物释放以产生编码一个或多个肽表位的mRNA。

[0225] 在一些实施方案中，第一核酸的特征在于具有5' -三磷酸酯和3' OH。在其他实施方案中，第二核酸包含3' OH。在又一些其他实施方案中，第三核酸包含5' -三磷酸酯和3' OH。第二核酸还可包含5' -帽结构。所述方法还可涉及使包含聚腺苷酸区域的第四核酸结扎至第三核酸的3' 末端上的进一步步骤。第四核酸可包含5' -三磷酸酯。

[0226] 所述方法可包括或不包括反相纯化。所述方法还可包括洗涤步骤，其中洗涤固体支撑物以移除未反应核酸。固体支撑物可以是例如捕集树脂。在一些实施方案中，所述方法涉及dT纯化。

[0227] 根据本公开，编码本公开的核酸（例如，mRNA）癌症疫苗的模板DNA包括编码一个或多个肽表位的开放阅读框(ORF)。在一些实施方案中，模板DNA包括具有多达1000个核苷酸，例如约10-350个、30-300个核苷酸或约50-250个核苷酸的ORF。在一些实施方案中，模板DNA包括具有约150个核苷酸的ORF。在一些实施方案中，模板DNA包括具有约200个核苷酸的ORF。

[0228] 在一些实施方案中，在反应发生之后，从IVT反应混合物的组分中纯化出IVT转录物。例如，可用无RNA酶的DNA酶处理粗制IVT混合物以消化原始模板。核酸（例如，mRNA）可使用本领域中已知的方法纯化，所述方法包括但不限于使用有机溶剂进行沉淀或基于柱的纯

化方法。市售的试剂盒可用于纯化RNA,例如MEGACLEAR™试剂盒(Ambion,Austin,TX)。核酸(例如,mRNA)可使用本领域中已知的方法定量,所述方法包括但不限于可商购获得的仪器例如NanoDrop。纯化的核酸(例如,mRNA)可例如通过琼脂糖凝胶电泳来分析以确认该核酸具有适当大小,以及/或者确认没有发生核酸降解。

[0229] 非翻译区域(UTR)

[0230] 非翻译区域(UTR)是在起始密码子(5' UTR)之前和终止密码子(3' UTR)之后的未翻译的核酸区段。在一些实施方案中,包含编码一个或多个肽表位的开放阅读框(ORF)的本公开的核酸(例如,核糖核酸(RNA),例如信使RNA(mRNA))进一步包含一个或多个UTR(例如,5' UTR或其功能片段、3' UTR或其功能片段,或它们的组合)。

[0231] UTR可与核酸中的编码区域同源或异源。在一些实施方案中,UTR与编码一个或多个肽表位的ORF同源。在一些实施方案中,UTR与编码一个或多个肽表位的ORF异源。在一些实施方案中,所述核酸包含两个或更多个5' UTR或其功能片段,其中的每一个均具有相同或不同的核苷酸序列。在一些实施方案中,所述核酸包含两个或更多个3' UTR或其功能片段,其中的每一个均具有相同或不同的核苷酸序列。

[0232] 在一些实施方案中,5' UTR或其功能片段、3' UTR或其功能片段或它们的任何组合是经序列优化的。

[0233] 在一些实施方案中,5' UTR或其功能片段、3' UTR或其功能片段,或它们的任何组合包含至少一种化学修饰的核碱基,例如5-甲氧基尿嘧啶。

[0234] UTR可具有提供调控作用的特征,例如稳定性、定位和/或翻译效率的增加或降低。可将包含UTR的核酸施用给细胞、组织或生物体,并且可使用常规方法测量一种或多种调控特征。在一些实施方案中,5' UTR或3' UTR的功能片段分别包含全长5' UTR或3' UTR的一种或多种调控特征。

[0235] 天然5' UTR携带在翻译启动中起作用的特征。它们拥有像通常已知参与核糖体启动许多基因的翻译的过程的Kozak序列之类的签名序列(signature)。还已知5' UTR形成参与延伸因子结合的二级结构。

[0236] 通过工程改造通常在特定靶器官的丰富表达的基因中发现的特征,可增强核酸的稳定性和蛋白质产生量。例如,引入肝表达的mRNA(诸如白蛋白、血清淀粉样蛋白A、载脂蛋白A/B/E、转铁蛋白、甲胎蛋白、促红细胞生成素或因子VIII)的5' UTR可增强肝细胞系或肝脏中核酸的表达。同样,使用来自其他组织特异性mRNA的5' UTR来改进在该组织中的表达对于以下各项来说是可能的:肌肉(例如,MyoD、肌球蛋白、肌红蛋白、肌细胞生成素(Myogenin)、Herculin),内皮细胞(例如,Tie-1、CD36),髓样细胞(例如,C/EBP、AML1、G-CSF、GM-CSF、CD11b、MSR、Fr-1、i-NOS),白细胞(例如,CD45、CD18),脂肪组织(例如,CD36、GLUT4、ACRP30、脂联素)以及肺上皮细胞(例如,SP-A/B/C/D)。

[0237] 在一些实施方案中,UTR选自其蛋白质共有共同的功能、结构、特征或特性的转录物的家族。例如,所编码的多肽可属于在特定细胞、组织中或者在发育期间的某些时间表达的蛋白质家族(即,共有至少一种功能、结构、特征、定位、起源或表达模式的蛋白质家族)。可以将来自任何所述基因或mRNA的UTR交换为相同或不同蛋白质家族的任何其他UTR以便形成新的核酸。

[0238] 在一些实施方案中,5' UTR和3' UTR可以是异源的。在一些实施方案中,5' UTR可源

自与3' UTR不同的物种。在一些实施方案中,3' UTR可源自与5' UTR不同的物种。

[0239] 国际专利申请号PCT/US2014/021522 (公布号W0/2014/164253) 提供了可以在本公开的核酸中用作ORF的侧接区域的示例性UTR的列表。该公布出于此目的以引用方式并入本文。

[0240] 可在本公开的核酸中利用的另外的示例性UTR包括但不限于源自以下的核酸序列的一个或多个5' UTR和/或3' UTR:球蛋白,诸如 α -球蛋白或 β -球蛋白(例如,爪蟾(*Xenopus*)、小鼠、兔或人球蛋白);强Kozak翻译启动信号;CYBA(例如,人细胞色素b-245 α 多肽);白蛋白(例如,人白蛋白7);HSD17B4(羟基类固醇(17- β)脱氢酶);病毒(例如,烟草蚀纹病毒(TEV)、委内瑞拉马脑炎病毒(VEEV)、登革热病毒、巨细胞病毒(CMV;例如CMV立即早期1(IE1))、肝炎病毒(例如,乙型肝炎病毒)、辛德毕斯病毒或PAV大麦黄矮病毒);热休克蛋白(例如,hsp70);翻译启动因子(例如,eIF4G);葡萄糖转运蛋白(例如,hGLUT1(人葡萄糖转运蛋白1));肌动蛋白(例如,人 α 或 β 肌动蛋白);GAPDH;微管蛋白;组蛋白;柠檬酸循环酶;拓扑异构酶(例如,缺乏5' TOP基序的TOP基因的5' UTR(寡嘧啶束));核糖体蛋白大32(L32);核糖体蛋白(例如,人或小鼠核糖体蛋白,诸如例如rps9);ATP合酶(例如,ATP5A1或线粒体H⁺-ATP合酶的 β 亚基);生长激素e(例如,牛(bGH)或人(hGH));延伸因子(例如,延伸因子1 α 1(EEF1A1));锰超氧化物歧化酶(MnSOD);肌细胞增强因子2A(MEF2A); β -F1-ATP酶、肌酸激酶、肌红蛋白、粒细胞集落刺激因子(G-CSF);胶原(例如,I型胶原 α 2(Col1A2)、I型胶原 α 1(Col1A1)、VI型胶原 α 2(Col6A2)、VI型胶原 α 1(Col6A1));核糖体结合蛋白(例如,核糖体结合蛋白I(RPNI));低密度脂蛋白受体相关蛋白(例如,LRP1);心肌营养素样细胞因子(例如,Nnt1);钙网蛋白(Calr);原胶原-赖氨酸、2-氧代戊二酸5-双加氧酶1(Plod1);以及核连蛋白(例如,Nucb1)。

[0241] 在一些实施方案中,5' UTR选自由以下组成的组: β -球蛋白5' UTR;含有强Kozak翻译启动信号的5' UTR;细胞色素b-245 α 多肽(CYBA)5' UTR;羟基类固醇(17- β)脱氢酶(HSD17B4)5' UTR;烟草蚀纹病毒(TEV)5' UTR;委内瑞拉马脑炎病毒(VEEV)5' UTR;编码非结构蛋白的风疹病毒(RV)RNA的5'近端开放阅读框;登革热病毒(DEN)5' UTR;热休克蛋白70(Hsp70)5' UTR;eIF4G 5' UTR;GLUT1 5' UTR;其功能片段以及它们的任何组合。

[0242] 在一些实施方案中,3' UTR选自由以下组成的组: β -球蛋白3' UTR;CYBA 3' UTR;白蛋白3' UTR;生长激素(GH)3' UTR;VEEV 3' UTR;乙型肝炎病毒(HBV)3' UTR; α -球蛋白3' UTR;DEN 3' UTR;PAV大麦黄矮病毒(BYDV-PAV)3' UTR;延伸因子1 α 1(EEF1A1)3' UTR;锰超氧化物歧化酶(MnSOD)3' UTR;线粒体H(+)-ATP合酶(β -mRNA)3' UTR的 β 亚基;GLUT1 3' UTR;MEF2A 3' UTR; β -F1-ATP酶3' UTR;其功能片段以及它们的组合。

[0243] 可将源自任何基因或mRNA的野生型UTR并入到本公开的核酸中。在一些实施方案中,可相对于野生型或原生UTR更改UTR以产生变体UTR,例如,通过变换所述UTR相对于ORF的取向或位置来进行;或者通过包含另外的核苷酸、核苷酸的删除、核苷酸的交换或转位来进行。在一些实施方案中,可利用5'或3' UTR的变体,例如野生型UTR的突变体,或其中一个或多个核苷酸被添加至UTR的末端或者从UTR的末端被移除的变体。

[0244] 另外,可将一种或多种合成UTR与一种或多种非合成UTR组合使用。参见例如,Mandal和Rossi,Nat.Protoc.20138(3):568-82,以及可在www.addgene.org/Derrick_Rossi/获得的序列,其中每一者的内容以引用方式整体并入本文。可将UTR或其部分以与它

们所选自的转录物中的取向相同的取向放置,或者可以在取向或位置上对所述UTR或其部分进行更改。因此,可使5' UTR和/或3' UTR反转、缩短、延长,或者与一个或多个其他5' UTR或3' UTR组合。

[0245] 在一些实施方案中,所述核酸可包含多个UTR,例如双重、三重或四重5' UTR或3' UTR。例如,双重UTR包含串联或基本上串联的相同UTR的两个拷贝。例如,可使用双重 β -球蛋白3' UTR(参见例如US2010/0129877,其内容以引用方式整体并入本文)。

[0246] 本公开的核酸可包含特征的组合。例如,ORF可侧接有包含强Kozak翻译启动信号的5' UTR和/或包含用于模板化添加聚腺苷酸尾部的oligo(dT)序列的3' UTR。5' UTR可包含来自相同和/或不同UTR的第一核酸片段和第二核酸片段(参见例如US2010/0293625,其出于此目的以引用方式整体并入本文)。

[0247] 其他非UTR序列可用作本公开的核酸内的区域或亚区域。例如,可将内含子或内含子序列的部分并入到本公开的核酸中。内含子序列的并入可增加蛋白质产生量以及核酸表达水平。在一些实施方案中,本公开的核酸包含内部核糖体进入位点(IRES)而不是UTR或者外加UTR(参见例如Yakubov等人,Biochem.Biophys.Res.Commun.2010394(1):189-193,其内容以引用方式整体并入本文)。在一些实施方案中,所述核酸包含IRES而不是5' UTR序列。在一些实施方案中,所述核酸包含ORF和病毒衣壳序列。在一些实施方案中,所述核酸包含合成5' UTR与非合成3' UTR的组合。

[0248] 在一些实施方案中,UTR还可包含至少一种翻译增强子核酸、一种翻译增强子元件或多种翻译增强子元件(统称为“TEE”,其是指增加由多核苷酸产生的多肽或蛋白质的量的核酸序列)。作为非限制性实例,TEE可包括出于此目的以引用方式整体并入本文的US2009/0226470中描述的那些和本领域中已知的那些。作为非限制性实例,TEE可位于转录启动子与起始密码子之间。在一些实施方案中,5' UTR包含TEE。在一方面,TEE是UTR中的保守元件,其可促进核酸的翻译活性,诸如但不限于帽依赖性或非帽依赖性翻译。在一个非限制性实例中,TEE包含Gtx同源结构域蛋白的5' 前导区中的TEE序列。参见Chappell等人,PNAS 2004 101:9590-9594,其出于此目的以引用方式整体并入本文。

[0249] 术语“翻译增强子多核苷酸”或“翻译增强子多核苷酸序列”是指包含本文所提供和/或本领域中已知的一种或多种TEE的核酸(参见例如US6310197、US6849405、US7456273、US7183395、US2009/0226470、US2007/0048776、US2011/0124100、US2009/0093049、US2013/0177581、W02009/075886、W02007/025008、W02012/009644、W02001/055371、W01999/024595、EP2610341A1和EP2610340A1;其中每一篇的内容出于此目的以引用方式整体并入本文)或它们的变体、同源物或功能衍生物。在一些实施方案中,本公开的核酸包含TEE的一个或多个拷贝。翻译增强子核酸中的TEE可组构在一个或多个序列区段中。序列区段可拥有本文所提供的TEE中的一种或多种,其中每种TEE以一个或多个拷贝存在。当多个序列区段存在于翻译增强子核酸中时,它们可以是同质的(homogenous)或异质的(heterogeneous)。因此,翻译增强子核酸中的多个序列区段可拥有同一或不同类型的本文所提供的TEE、同一或不同数量的每种所述TEE的拷贝,以及/或者同一或不同的在每个序列区段内的TEE组构(organization)。在一个实施方案中,本公开的核酸包含翻译增强子核酸序列。

[0250] 在一些实施方案中,可将包含至少一种本文所述的TEE的5' UTR和/或3' UTR并入到单顺反子序列,诸如但不限于载体系统或核酸载体中。在一些实施方案中,本公开的多核苷

酸的5' UTR和/或3' UTR包含本文所述的TEE或其部分。在一些实施方案中,3' UTR中的TEE可与位于5' UTR中的TEE相同和/或不同。

[0251] 在一些实施方案中,本公开的核酸的5' UTR和/或3' UTR可包含至少1个、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、至少21个、至少22个、至少23个、至少24个、至少25个、至少30个、至少35个、至少40个、至少45个、至少50个、至少55个或超过60个TEE序列。在一个实施方案中,本公开的核酸的5' UTR可包含1-60个、1-55个、1-50个、1-45个、1-40个、1-35个、1-30个、1-25个、1-20个、1-15个、1-10个、9个、8个、7个、6个、5个、4个、3个、2个或1个TEE序列。本公开的核酸的5' UTR中的TEE序列可以是相同或不同的TEE序列。本公开的核酸的5' UTR中的不同TEE序列的组合可包括其中并入有任何不同TEE序列的超过一个拷贝的组合。

[0252] 在一些实施方案中,5' UTR和/或3' UTR包含用于分隔两个TEE序列的间隔子。作为非限制性实例,该间隔子可以是15个核苷酸的间隔子和/或本领域中已知的其他间隔子(例如,为三个核苷酸的倍数)。作为另一个非限制性实例,5' UTR和/或3' UTR分别包含5' UTR和/或3' UTR中重复至少一次、至少两次、至少3次、至少4次、至少5次、至少6次、至少7次、至少8次、至少9次、至少10次或超过10次的TEE序列-间隔子模块。在一些实施方案中,5' UTR和/或3' UTR包含重复1、2、3、4、5、6、7、8、9或10次的TEE序列-间隔子模块。

[0253] 3' UTR和AU富含元件

[0254] 在某些实施方案中,本公开的核酸(例如,编码本公开的肽表位的核酸)进一步包含3' UTR。

[0255] 3' UTR是紧接在翻译终止密码子之后的mRNA区段,并且常常含有转录后影响基因表达的调控区域。3' UTR内的调控区域可影响mRNA的聚腺苷酸化、翻译效率、定位和稳定性。在一个实施方案中,可用于本公开的3' UTR包含调控蛋白或微小RNA的结合位点。在一些实施方案中,3' UTR具有沉默子区域,所述沉默子区域与阻遏蛋白结合并抑制mRNA的表达。在其他实施方案中,3' UTR包含AU富含元件(ARE)。蛋白质结合ARE从而以局部方式影响转录物的稳定性或衰变速率或者影响翻译启动。在其他实施方案中,3' UTR包含序列AAUAAA,其指导称作聚腺苷酸尾部的数百个腺嘌呤残基添加至mRNA转录物的末端。

[0256] 已知天然或野生型3' UTR具有嵌入在它们中的腺苷和尿苷链段。这些AU富含签名序列在具有高周转率的基因中是特别普遍的。AU富含元件(ARE)可基于它们的序列特征和功能特性被分成三个类别(Chen等人,1995):I类ARE在U富含区域内含有AUUUA基序的若干分散拷贝。C-Myc和MyoD含有I类ARE。II类ARE拥有两个或更多个重叠的UUAUUUA(U/A)(U/A)九聚物。含有这种类型的ARE的分子包括GM-CSF和TNF- α 。III类ARE不含有AUUUA基序。c-Jun和肌细胞生成素是这种类别的两个被充分研究的实例。已知大多数结合至ARE的蛋白质会使信使不稳定,而ELAV家族的成员(最显著的是HuR)已被证明能增加mRNA的稳定性。HuR结合所有三个类别的ARE。将HuR特异性结合位点工程改造至核酸分子的3' UTR中将导致HuR结合,并由此导致体内信使的稳定化。

[0257] 3' UTR AU富含元件(ARE)的引入、去除或修饰可用于调节本公开的核酸的稳定性。当工程改造特定核酸时,可引入ARE的一个或多个拷贝来使本公开的核酸不那么稳定,由此缩减翻译并减少所得蛋白质的产生量。同样,可鉴定出ARE并将其去除或者使其突变以增加

细胞内稳定性并因此增加翻译和所得蛋白质的产生量。可使用本公开的核酸在相关细胞系中进行转染实验,并且可在转染后的各种时间点测定蛋白质产生量。例如,可用不同的ARE工程改造分子转染细胞,并且使用针对相关蛋白质的ELISA试剂盒测定转染后第6小时、第12小时、第24小时、第48小时以及第7天产生的蛋白质。

[0258] 具有5'帽的区域

[0259] 本文所述的核酸癌症疫苗可以是包含一种或多种具有编码肽表位的开放阅读框的mRNA的mRNA癌症疫苗。这些mRNA中的每一种均可具有5'帽。

[0260] 天然mRNA的5'帽结构参与核输出,从而增加mRNA稳定性,并且结合mRNA帽结合蛋白(CBP),所述mRNA帽结合蛋白(CBP)通过与聚腺苷酸结合蛋白缔合形成成熟的环状mRNA物质来负责细胞中的mRNA稳定性和翻译能力。所述帽进一步协助在mRNA剪接过程中去除5'近端内含子。

[0261] 内源性mRNA分子可以是5'端加帽的,从而在末端鸟苷帽残基与mRNA分子的5'末端转录的有义核苷酸之间产生5'-ppp-5'-三磷酸酯键联(帽)。这种5'-鸟苷酸盐帽接着可被甲基化以产生N7-甲基-鸟苷酸酯残基(帽0)。mRNA的5'端的末端和/或末端前(anteterminal)的转录的核苷酸的核糖还可任选地被2'-O-甲基化(例如,就第一核糖上的2'-羟基而言(帽1);或者就前两个核糖上的2'-羟基而言(帽2))。通过鸟苷酸酯帽结构的水解和切割进行的5'-脱帽可靶向用于降解的核酸分子,诸如mRNA分子。

[0262] 在一些实施方案中,本公开的核酸(例如,编码肽表位的核酸)并入有帽部分。

[0263] 在一些实施方案中,本公开的核酸(例如,编码肽表位的核酸)包含不可水解的帽结构,从而防止脱帽并由此增加mRNA半衰期。因为帽结构水解需要5'-ppp-5'磷酸二酯键联的切割,所以可在加帽反应过程中使用经修饰的核苷酸。例如,可根据制造商的说明书将来自New England Biolabs(Ipswich,MA)的牛痘加帽酶与 α -硫代-鸟苷核苷酸一起使用以便在5'-ppp-5'帽中创建硫代磷酸酯键联。可使用另外的经修饰的鸟苷核苷酸,诸如 α -甲基-磷酸酯和硒代磷酸酯核苷酸。

[0264] 另外的修饰包括但不限于多核苷酸(如上面所提到的)的5'末端和/或5'末端前核苷酸的核糖的在糖环的2'-羟基上的2'-O-甲基化。多种不同的5'帽结构可用于产生核酸分子(诸如充当mRNA分子的多核苷酸)的5'帽。帽类似物,在本文中也称为合成帽类似物、化学帽、化学帽类似物或者结构或功能帽类似物,在化学结构上与天然(即,内源性、野生型或生理的)5'帽不同,同时保留帽功能。帽类似物可以是化学(即,非酶促)合成或酶促合成的,并且/或者联接至本公开的多核苷酸。

[0265] 例如,抗反向帽类似物(ARCA)帽含有通过5'-5'-三磷酸酯基团联接的两个鸟嘌呤,其中一个鸟嘌呤含有N7甲基以及3'-O-甲基(即,N7,3'-O-二甲基-鸟苷-5'-三磷酸酯-5'-鸟苷(m7G-3' mppp-G);其可等效地称为3'-O-Me-m7G(5') Ppp(5') G)。另一个未修饰鸟嘌呤的3'-O原子变成联接至加帽的多核苷酸的5'末端核苷酸。N7甲基化的鸟嘌呤和3'-O甲基化的鸟嘌呤提供加帽的多核苷酸的末端部分。

[0266] 另一种示例性帽是与ARCA相似但在鸟苷上具有2'-O-甲基的mCAP(即,N7,2'-O-二甲基-鸟苷-5'-三磷酸酯-5'-鸟苷,m7Gm-ppp-G)。

[0267] 在一些实施方案中,该帽是二核苷酸帽类似物。作为非限制性实例,二核苷酸帽类似物可在不同磷酸酯位置处用硼烷磷酸酯基团或硒代磷酸酯(phosphoroselenoate)基团修

饰,如在美国专利号US 8,519,110中描述的二核苷酸帽类似物,所述专利的内容以引用方式整体并入本文。

[0268] 在另一个实施方案中,该帽是帽类似物,所述帽类似物是本领域中已知和/或本文所描述的帽类似物的N7-(4-氯苯氧基乙基)取代的二核苷酸形式。帽类似物的N7-(4-氯苯氧基乙基)取代的二核苷酸形式的非限制性实例包括N7-(4-氯苯氧基乙基)-G(5')ppp(5')G和N7-(4-氯苯氧基乙基)-m3'-OG(5')ppp(5')G帽类似物(参见例如在Kore等人 *Bioorganic&Medicinal Chemistry* 2013 21:4570-4574中描述的各种帽类似物和合成帽类似物的方法;所述文献的内容出于此目的以引用方式整体并入本文)。在另一个实施方案中,本公开的帽类似物是4-氯/溴苯氧基乙基类似物。

[0269] 虽然帽类似物允许多核苷酸或其区域在体外转录反应中的伴随加帽,但高达20%的转录物可仍保持未加帽。这种情况以及帽类似物与由内源性细胞转录机构产生的核酸的内源性5'帽结构的结构差异可导致降低的翻译能力和降低的细胞稳定性。

[0270] 本公开的核酸(例如,编码肽抗原的核酸)还可在制造后使用酶加帽(无论是通过IVT还是化学合成加帽)以便产生更真实的5'帽结构。如本文所用,短语“更真实的”是指在结构或功能上紧密地反映或模拟内源性或野生型特征的特征。也就是说,“更真实的”特征是与合成特征或类似物等相比,内源性、野生型、天然或生理细胞功能和/或结构的更好表示,或者在一个或多个方面胜过相应的内源性、野生型、天然或生理特征的特征。更真实的5'帽结构的非限制性实例是与本领域中已知的合成5'帽结构(或者与野生型、天然或生理5'帽结构)相比,尤其具有增强的帽结合蛋白的结合、增加的半衰期、降低的对5'核酸内切酶的易感性和/或减少的5'脱帽的那些。例如,重组牛痘病毒加帽酶和重组2'-O-甲基转移酶可在多核苷酸的5'末端核苷酸与鸟嘌呤帽核苷酸之间形成规范的5'-5'-三磷酸酯键联,其中所述帽鸟嘌呤含有N7甲基化并且mRNA的5'末端核苷酸含有2'-O-甲基。这种结构被称为帽1结构。这种帽导致与例如本领域中已知的其他5'帽类似物结构相比更高的翻译能力和细胞稳定性以及减少的细胞促炎细胞因子的激活。帽结构包括但不限于,7mG(5')PPP(5')N,pN2p(帽0)、7mG(5')ppp(5')N1mpNp(帽1)和7mG(5')-ppp(5')N1mpN2mp(帽2)。

[0271] 作为非限制性实例,在制造后对嵌合核酸加帽可更有效,因为几乎100%的嵌合核酸都可被加帽。这与在体外转录反应过程中帽类似物被联接至嵌合核酸时的约80%形成对比。

[0272] 根据本公开,5'末端帽可包括内源性帽或帽类似物。根据本公开,5'末端帽可包含鸟嘌呤类似物。有用的鸟嘌呤类似物包括但不限于,肌苷、N1-甲基-鸟苷、2'氟-鸟苷、7-脱氮-鸟苷、8-氧代-鸟苷、2-氨基-鸟苷、LNA-鸟苷以及2-叠氮基-鸟苷。

[0273] 聚腺苷酸尾部

[0274] 在一些实施方案中,本公开的核酸(例如,编码肽表位的核酸)进一步包含聚腺苷酸尾部。在进一步实施方案中,可将聚腺苷酸尾部上的末端基团并入以进行稳定化。在其他实施方案中,聚腺苷酸尾部包含脱-3'羟基尾部。

[0275] 在RNA加工过程中,长的腺嘌呤核苷酸链(聚腺苷酸尾部)可被添加至核酸(诸如mRNA分子)上以便增加稳定性。紧随转录之后,转录物的3'末端可被切割以释放3'羟基。接着聚腺苷酸聚合酶将腺嘌呤核苷酸链添加至RNA上。被称为聚腺苷酸化的过程添加了聚腺苷酸尾部,该聚腺苷酸尾部的长度可在例如近似80个残基至近似250个残基之间,包括近似

80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240或250个残基的长度。在一些实施方案中,所述聚腺苷酸尾部包含约100个核苷酸。

[0276] 聚腺苷酸尾部还可在构建体从细胞核中被排出之后添加。

[0277] 根据本公开,可将聚腺苷酸尾部上的末端基团并入以进行稳定化。本公开的多核苷酸可包含脱-3' 羟基尾部。它们还可包含如Junjie Li等人(*Current Biology*,第15卷,1501-1507,2005年8月23日,该文献的内容出于此目的以引用方式整体并入本文)所教导的结构部分或2' -O甲基修饰。

[0278] 本公开的核酸可被设计成编码具有包括组蛋白mRNA在内的替代聚腺苷酸尾部结构的转录物。根据Norbury,末端尿苷化还已经在人复制依赖性组蛋白mRNA上被检测到。这些mRNA的周转被认为对于在染色体DNA复制完成或受到抑制后预防潜在有毒的组蛋白累积来说是重要的。这些mRNA的特点在于它们缺乏3' 聚腺苷酸尾部,其功能则由稳定的茎-环结构及其同源茎-环结合蛋白(SLBP)承担;后者执行与聚腺苷酸化mRNA上的PABP的功能相同的功能”(Norbury,“Cytoplasmic RNA:a case of the tail wagging the dog,”*Nature Reviews Molecular Cell Biology*;AOP,2013年8月29日在线公布;doi:10.1038/nrm3645),该文献的内容出于此目的以引用方式整体并入本文。

[0279] 独特的聚腺苷酸尾部长度的本公开的核酸提供了某些优点。一般来讲,当聚腺苷酸尾部存在时,其长度大于30个核苷酸。在另一个实施方案中,聚腺苷酸尾部的长度大于35个核苷酸(例如,至少或大于约15个、20个、25个、30个、35个、40个、45个、50个、55个、60个、70个、80个、90个、100个、120个、140个、160个、180个、200个、250个、300个、350个、400个、450个、500个、600个、700个、800个、900个、1,000个、1,100个、1,200个、1,300个、1,400个、1,500个、1,600个、1,700个、1,800个、1,900个、2,000个、2,500个,或3,000个核苷酸)。

[0280] 在一些实施方案中,所述核酸或其区域包含约15至约3,000个核苷酸(例如,15至50个、15至100个、15至200个、15至300个、15至400个、15至500个、15至600个、15至700个、15至800个、15至900个、15至1000个、15至1200个、15至1400个、15至1500个、15至1800个、15至2000个、15至2500个、15至3000个、50至100个、50至200个、50至300个、50至400个、50至500个、50至600个、50至700个、50至800个、50至900个、50至1000个、50至1200个、50至1400个、50至1500个、50至1800个、50至2000个、50至2500个、50至3000个、100至200个、100至300个、100至400个、100至500个、100至600个、100至700个、100至800个、100至900个、100至1000个、100至1200个、100至1400个、100至1500个、100至1800个、100至2000个、100至2500个、100至3000个、200至300个、200至400个、200至500个、200至600个、200至700个、200至800个、200至900个、200至1000个、200至1500个、200至3000个、500至1000个、500至1500个、500至2000个、500至2500个、500至3000个、1000至1500个、1000至2000个、1000至2500个、1000至3000个、1500至3000个、2500至3000个,或2000至3000个核苷酸)。

[0281] 在一些实施方案中,相对于总体核酸的长度或所述核酸的特定区域的长度来设计聚腺苷酸尾部。这种设计可基于编码区域的长度、特定特征或区域的长度或者基于从核酸表达的最终产物的长度。

[0282] 在这种背景下,聚腺苷酸尾部可在长度上比所述核酸或其特征大10%、20%、30%、40%、50%、60%、70%、80%、90%或100%。聚腺苷酸尾部还可被设计为它所属的核酸的一部分。在这种背景下,聚腺苷酸尾部可以是构建体的总长度、构建体区域或者构建体

减去聚腺苷酸尾部的总长度的10%、20%、30%、40%、50%、60%、70%、80%或90%或更高。进一步地,工程改造聚腺苷酸结合蛋白的结合位点和核酸缀合可增强表达。

[0283] 另外,多个不同的核酸可利用在聚腺苷酸尾部的3'末端处的经修饰的核苷酸经由PABP(聚腺苷酸结合蛋白)通过3'端联接在一起。可在相关细胞系中进行转染实验并且可在转染后第12小时、第24小时、第48小时、第72小时和第7天通过ELISA测定蛋白质产生量。

[0284] 在一些实施方案中,本公开的核酸被设计成包括聚腺苷酸-G四联体区域。G四联体是可由DNA和RNA两者中的G富含序列形成的四个鸟嘌呤核苷酸的环状氢键合阵列。在这个实施方案中,G四联体并入在聚腺苷酸尾部的末端。在各种时间点针对稳定性、蛋白质产生量和包括半衰期在内的其他参数来对所得核酸进行测定。已发现,聚腺苷酸-G四联体导致来自mRNA的蛋白质产生量等效于单独使用具有120个核苷酸的聚腺苷酸尾部所观察到的蛋白质产生量的至少75%。

[0285] 起始密码子区域

[0286] 本公开还包括核酸,所述核酸包含起始密码子区域和本文所述的核酸(例如,包含编码肽表位的核苷酸序列的核酸)。在一些实施方案中,本公开的核酸可具有与起始密码子区域类似的或者与起始密码子区域有类似功能的区域。

[0287] 在一些实施方案中,核酸的翻译可启动于不为起始密码子AUG的密码子。核酸的翻译可启动于替代起始密码子,诸如但不限于ACG、AGG、AAG、CTG/CUG、GTG/GUG、ATA/AUA、ATT/AUU、TTG/UUG(参见Touriol等人Biology of the Cell 95(2003)169-178以及Matsuda和Mauro PLoS ONE,2010 5:11;该文献中每一篇的内容出于此目的以引用方式整体并入本文)。

[0288] 作为非限制性实例,核酸的翻译开始于替代起始密码子ACG。作为另一个非限制性实例,核酸翻译开始于替代起始密码子CTG或CUG。作为又一个非限制性实例,核酸的翻译开始于替代起始密码子GTG或GUG。

[0289] 已知侧接启动翻译的密码子(诸如但不限于起始密码子或替代起始密码子)的核苷酸影响核酸的翻译效率、长度和/或结构。(参见例如Matsuda和Mauro PLoS ONE,2010 5:11;该文献的内容出于此目的以引用方式整体并入本文)。掩蔽侧接启动翻译的密码子的任何核苷酸可用于更改多核苷酸的翻译启动位置、翻译效率、长度和/或结构。

[0290] 在一些实施方案中,掩蔽剂可在起始密码子或替代起始密码子附近使用,以便掩蔽或隐蔽所述密码子以降低在所掩蔽的起始密码子或替代起始密码子处翻译启动的可能性。掩蔽剂的非限制性实例包括反义锁核酸(LNA)核酸和外显子连接复合物(EJC)(参见,例如,Matsuda和Mauro描述了掩蔽剂LNA多核苷酸和EJC(PLoS ONE,2010 5:11);该文献的内容出于此目的以引用方式整体并入本文)。

[0291] 在另一个实施方案中,掩蔽剂可用于掩蔽核酸的起始密码子以便增加翻译将启动于替代起始密码子的可能性。在一些实施方案中,掩蔽剂可用于掩蔽第一起始密码子或替代起始密码子,以便增加翻译将启动于所掩蔽的起始密码子或替代起始密码子下游的起始密码子或替代起始密码子的机会。

[0292] 在另一个实施方案中,可从核酸序列中去除核酸的起始密码子,以便使核酸的翻译开始于不为起始密码子的密码子。核酸的翻译可开始于所去除的起始密码子之后的密码子或者开始于下游起始密码子或替代起始密码子。在非限制性实例中,将起始密码子ATG或

AUG作为核酸序列的前3个核苷酸去除,以便使翻译启动于下游起始密码子或替代起始密码子。起始密码子被去除的核酸序列可进一步包含至少一种针对下游起始密码子和/或替代起始密码子的掩蔽剂,以便控制或试图控制翻译的启动、核酸的长度和/或核酸的结构。

[0293] 终止密码子区域

[0294] 本公开还包括核酸,所述核酸包含终止密码子区域和本文所述的核酸(例如,编码肽表位的核酸)。在一些实施方案中,本公开的核酸可在3'非翻译区域(UTR)之前包含至少两个终止密码子。终止密码子可选自TGA、TAA和TAG(在DNA的情况下),或者可选自UGA、UAA和UAG(在RNA的情况下)。在一些实施方案中,本公开的核酸包含终止密码子TGA(在DNA的情况下)或终止密码子UGA(在RNA的情况下)以及一个另外的终止密码子。在进一步实施方案中,所述另外的终止密码子可以是TAA或UAA。在另一个实施方案中,本公开的核酸包含三个连续终止密码子、四个终止密码子或更多。

[0295] 插入和取代

[0296] 本公开还包括本公开的核酸,该核酸进一步包含插入和/或取代。

[0297] 在一些实施方案中,所述核酸的5'UTR可通过插入相同碱基的核苷的至少一个区域和/或串段来替换。核苷酸的区域和/或串段可包括但不限于至少3个、至少4个、至少5个、至少6个、至少7个或至少8个核苷酸并且所述核苷酸可以是天然的和/或非天然的。作为非限制性实例,核苷酸的群可包含5-8个腺嘌呤、胞嘧啶、胸腺嘧啶、本文所公开的任何其他核苷酸的串段和/或它们的组合。

[0298] 在一些实施方案中,所述核酸的5'UTR可通过插入两个不同碱基的核苷酸的至少两个区域和/或串段来替换,所述碱基诸如但不限于腺嘌呤、胞嘧啶、胸腺嘧啶、本文所公开的任何其他核苷酸和/或它们的组合。例如,5'UTR可通过插入5-8个腺嘌呤碱基、接着插入5-8个胞嘧啶碱基来替换。在另一个实例中,5'UTR可通过插入5-8个胞嘧啶碱基、接着插入5-8个腺嘌呤碱基来替换。

[0299] 在一些实施方案中,所述核酸可包括在可由RNA聚合酶辨识的转录起始位点下游的至少一个取代和/或插入。作为非限制性实例,至少一个取代和/或插入可通过取代正好在转录起始位点下游的区域中的至少一个核酸(诸如但不限于+1至+6)而在转录起始位点的下游发生。正好在转录起始位点下游的核苷酸的区域的变化可影响启动速率、增加表观三磷酸核苷酸(NTP)反应恒定值,并且增加短转录物从转录复合物的解离从而消除(curing)初始转录(Briebe等人,Biochemistry(2002)41:5144-5149;出于此目的以引用方式整体并入本文)。至少一个核苷的修饰、取代和/或插入可引起序列的沉默突变或者可引起氨基酸序列中的突变。

[0300] 在一些实施方案中,所述核酸可包括在转录起始位点下游的至少1个、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个或至少13个鸟嘌呤碱基的取代。

[0301] 在一些实施方案中,所述核酸可包括正好在转录起始位点下游的区域中的至少1个、至少2个、至少3个、至少4个、至少5个或至少6个鸟嘌呤碱基的取代。作为非限制性实例,如果所述区域中的核苷酸是GGGAGA,那么鸟嘌呤碱基可被至少1个、至少2个、至少3个或至少4个腺嘌呤核苷酸取代。在另一个非限制性实例中,如果所述区域中的核苷酸是GGGAGA,那么鸟嘌呤碱基可被至少1个、至少2个、至少3个或至少4个胞嘧啶碱基取代。在另一个非限

制性实例中,如果所述区域中的核苷酸是GGGAGA,那么鸟嘌呤碱基可被至少1个、至少2个、至少3个或至少4个胸腺嘧啶和/或本文所述的任何核苷酸取代。

[0302] 在一些实施方案中,所述核酸可包括在起始密码子上游的至少一个取代和/或插入。为清楚起见,本领域技术人员将理解起始密码子是蛋白质编码区域的第一密码子,而转录起始位点是转录开始的位点。所述核酸可包括但不限于至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个或至少8个核苷酸碱基的取代和/或插入。所述核苷酸碱基可在起始密码子上游的1个、至少1个、至少2个、至少3个、至少4个或至少5个位置处被插入或取代。被插入和/或取代的核苷酸可以是相同碱基(例如,全部是A,或者全部是C,或者全部是T,或者全部是G)、两种不同的碱基(例如,A和C、A和T,或C和T)、三种不同的碱基(例如,A、C和T,或A、C和T)或至少四种不同的碱基。

[0303] 作为非限制性实例,核酸中的编码区域上游的鸟嘌呤碱基可被腺嘌呤、胞嘧啶、胸腺嘧啶或本文所述的任何核苷酸取代。在另一个非限制性实例中,核酸中的鸟嘌呤碱基的取代可被设计成在转录起始位点下游和起始密码子之前的区域中留下一个鸟嘌呤碱基(参见Esvelt等人Nature (2011) 472 (7344):499-503;该文献的内容出于此目的以引用方式整体并入本文)。作为非限制性实例,至少5个核苷酸可被插入在转录起始位点下游但在起始密码子上游的1个位置处并且所述至少5个核苷酸可以是相同的碱基类型。

[0304] 根据本公开,嵌合核酸的两个区域或部分可例如使用三磷酸酯化学来接合或结扎。在一些实施方案中,具有100个或更少核苷酸的第一区域或部分被化学合成为具有5'-单磷酸酯和末端3'-脱OH或封闭的OH。如果该区域长于80个核苷酸,那么它可被合成为将随后通过结扎加以化学联接的两条或更多条链。如果使用IVT将所述第一区域或部分合成为非位置修饰的区域或部分,那么可以随后执行向5'-单磷酸酯的转化并随后对3'末端加帽。单磷酸酯保护基团可选自本领域中已知的任何一种。嵌合核酸的第二区域或部分可使用例如如本文所述的化学合成或IVT方法来合成。IVT方法可包括使用可利用具有经修饰的帽的引物的RNA聚合酶。替代地,可化学合成帽并使其偶联至IVT区域或部分。

[0305] 应注意对于结扎方法,用DNA T4结扎酶结扎并随后用DNA酶处理(以消除DNA T4结扎酶活性所需的DNA夹板)应当能容易地防止不期望的多联产物形成。

[0306] 整个嵌合多核苷酸不需要用磷酸酯-糖主链制造。如果所述区域或部分之一编码多肽,那么优选此种区域或部分包含磷酸酯-糖主链。

[0307] 可使用任何适当技术执行结扎,所述技术诸如酶促结扎、点击化学、正点击化学(orthoclick chemistry)、solulink或本领域人员已知的其他生物缀合化学。在一些实施方案中,所述结扎由互补寡核苷酸夹板来引导。在一些实施方案中,所述结扎在不存在互补寡核苷酸夹板的情况下执行。

[0308] 计算机化系统

[0309] 上述实施方案可以众多方式中的任一种来实施。例如,可使用硬件、软件或它们的组合来实施这些实施方案。当在软件中实现时,软件代码可在任何适合的处理器或一批处理器上执行,无论所述处理器是提供在单一计算机中还是分布在多个计算机之中。应当理解,执行上述功能的任何组件或一批组件都可在属类上被视为控制上面所讨论的功能的一个或多个控制器。所述一个或多个控制器可以众多方式实现,诸如采用专用硬件,或者采用使用微代码或软件编程以执行上面叙述的功能的通用硬件(例如,一个或多个处理器)来实

现。

[0310] 在这个方面,应当理解一个实现方式包括至少一种用计算机程序(即,多个指令)编码的计算机可读存储介质(即,至少一个有形非暂时计算机可读介质),诸如计算机存储器(例如,硬盘驱动器、闪存存储器、处理器工作存储器等)、软盘、光盘、磁带或其他有形非暂时计算机可读介质,所述计算机程序当在一个或多个处理器上被执行时会执行上面所讨论的功能。计算机可读存储介质可为可移动的,以使其上存储的程序可被加载到任何计算机资源上以实施本文讨论的技术。另外,应当理解对当被执行时会执行上面所讨论的功能的计算机程序的提及不限于在主计算机上运行的应用程序。更确切来说,术语“计算机程序”在本文中以通用含义用于指代可用于对一个或多个处理器编程以实现以上技术的任何类型的计算机代码(例如,软件或微代码)。

[0311] 作为非限制性实例,在一方面,本公开提供了用于选择用于包含在具有最大长度的核酸癌症疫苗中的核酸的计算机化系统,所述系统包括:通信接口,所述通信接口被配置为接收编码多个肽表位的多个核酸序列,其中所述肽表位中的每一个均是个性化癌症抗原的部分;和至少一个计算机处理器,所述至少一个计算机处理器被编程为:针对所述多个肽表位中的每一个,计算各自包含一个或多个肽表位中的至少一个肽表位的肽中多种核酸中的每一种的得分,其中所述核酸序列中的至少两个具有不同的长度;以及基于所计算的得分,对多种肽中的多个核酸序列进行排名;以及基于疫苗的排名和最大长度,选择用于包含在所述疫苗中的核酸序列。得分可通过本领域中已知的任何方式来计算。作为一组非限制性实例,得分可至少部分地基于选自由以下组成的组的一个或多个因素来计算:基因表达、RNA Seq、转录物丰度、DNA等位基因频率、氨基酸保守性、生理化学相似性、致癌基因、与特定HLA等位基因的预测结合亲和力、克隆性、结合效率以及在indel中的存在。在一些实施方案中,可使用变异等位基因频率(VAF)。在一个实施方案中,将VAF截止值选择为处于避免添加亚克隆突变的水平,因为肿瘤样品被相邻正常组织污染既降低肿瘤纯度,又导致(表观)VAF降低。因此,在肿瘤纯度低的情况下(例如,当平均VAF小于20%时),VAF截止值降低(例如,从10%降至5%)。在一些实施方案中,VAF截止值小于15%、14%、13%、12%、11%、10%、9%、8%、7%、6%、5%、4%、3%、2%、1%或更小。在某些实施方案中,所述一个或多个因素被输入到统计模型中。在一些实施方案中,所述统计模型可以是回归模型(例如,线性回归模型、逻辑回归模型、广义线性模型等)。在一些实施方案中,所述统计模型可以是广义线性模型(例如,逻辑回归模型、机率单位回归模型等)。在一些实施方案中,所述统计模型可以是例如随机森林回归模型、神经网络、支持向量机、高斯混合模型、分层贝叶斯模型和/或任何其他适合的统计模型。

[0312] 治疗方法

[0313] 本文提供了用于预防 and/或治疗人(例如,受试者或患者)和其他哺乳动物的癌症的组合物(例如,药物组合物)、方法、试剂盒和试剂。核酸癌症疫苗可用作医学上的治疗剂或防治剂,以预防和/或治疗癌症。在示例性方面,本公开的癌症疫苗用于提供免遭癌症的防治性保护。免遭癌症的防治性保护可在施用本公开的癌症疫苗之后实现。疫苗可施用一次、两次、三次、四次或更多次,但施用疫苗一次(任选地继之以单次加强)可能就足够。也可能合乎需要的是向患有癌症的个体施用疫苗以实现治疗响应。可能需要相应调整给药。

[0314] 一旦合成出癌症疫苗(例如,核酸癌症疫苗),就将其施用于患者。在一些实施方案

中,根据时间表施用疫苗长达两个月、长达三个月、长达四个月、长达五个月、长达六个月、长达七个月、长达八个月、长达九个月、长达十个月、长达十一个月、长达1年、长达1年半、长达两年、长达三年,或长达四年。时间表可相同或不同。在一些实施方案中,时间表是在前3周内每周施用,之后每月施用。该时间表可由本领域技术人员(例如,医师)根据个体患者或受试者的标准(例如,体重、年龄、癌症类型等)来确定或改变。

[0315] 疫苗可通过任何途径来施用。在一些实施方案中,疫苗通过真皮内、肌内、血管内、肿瘤内和/或皮下途径来施用。

[0316] 在一些实施方案中,核酸癌症疫苗也可与抗癌治疗剂一起施用。核酸癌症疫苗和其他治疗剂可同时或相继施用。当同时施用其他治疗剂时,它们可以以同一制剂或分开的制剂施用,但是在同一时间施用。当其他治疗剂和核酸癌症疫苗的施用在时间上分开时,所述其他治疗剂彼此相继施用并且与核酸癌症疫苗相继施用。介于这些化合物的施用之间的时间间隔可以是大约几分钟,或者可以是更长时间,例如,几小时、几天、几周、几个月。其他治疗剂包括但不限于抗癌治疗剂、佐剂、细胞因子、抗体、抗原等。

[0317] 在治疗的任何时候,都可对患者进行检查以确定疫苗中的突变是否仍然适当。基于该分析,可调整或重新配置疫苗,以使其包含一种或多种不同的突变或者以移除一种或多种突变。

[0318] 在示例性实施方案中,可向受试者(例如,哺乳动物受试者,诸如人受试者)施用如本文所述的含有RNA多核苷酸的癌症疫苗,并且所述RNA多核苷酸在体内被翻译以产生抗原性多肽。

[0319] 所述癌症疫苗可被诱导用于在细胞、组织或生物体中翻译多肽(例如,抗原或免疫原)。在示例性实施方案中,此种翻译在体内发生,但可设想其中此种翻译离体、在培养物中或在体外发生的实施方案。在示例性实施方案中,使细胞、组织或生物体与有效量的含有癌症疫苗的组合物接触,所述疫苗含有具有至少一个编码抗原性多肽的可翻译区域的多核苷酸。

[0320] 癌症RNA疫苗的“有效量”可至少部分地基于靶组织、靶细胞类型、施用手段、多核苷酸的物理特征(例如,经修饰的核苷的大小和程度)和癌症疫苗的其他组分以及其他决定因素来提供。一般来讲,癌症疫苗组合物的有效量提供随细胞中的抗原产生量而变化的诱导或加强的免疫应答,优选地比含有编码相同抗原或肽抗原的相应未修饰多核苷酸的组合物更高效。抗原产生量增加可通过以下来证明:细胞转染增加(用癌症疫苗转染的细胞的百分比)、从多核苷酸进行的蛋白质翻译增加、核酸降解降低(如例如由从经修饰的多核苷酸进行的蛋白质翻译的持续时间增加所证明),或宿主细胞的抗原特异性免疫应答更改。

[0321] 癌症疫苗可作为主动免疫方案的一部分向健康个体或者在癌症早期或在症状发作之后的活动性癌症(active cancer)期间防治性或治疗性地施用。在一些实施方案中,对细胞、组织或受试者提供的本公开的RNA疫苗的量可以是有效用于免疫防治的量。

[0322] 癌症疫苗可与其他防治性或治疗性化合物一起施用。作为非限制性实例,防治性或治疗性化合物可以是免疫增强剂或加强剂。如本文在提及组合物诸如疫苗时所用,术语“加强剂”是指额外施用防治性(疫苗)组合物。加强剂(或加强疫苗)可在较早地施用防治性组合物之后给予。在防治性组合物的初始施用与加强剂之间的施用时间可以是但不限于1分钟、2分钟、3分钟、4分钟、5分钟、6分钟、7分钟、8分钟、9分钟、10分钟、15分钟、20分钟、35

分钟、40分钟、45分钟、50分钟、55分钟、1小时、2小时、3小时、4小时、5小时、6小时、7小时、8小时、9小时、10小时、11小时、12小时、13小时、14小时、15小时、16小时、17小时、18小时、19小时、20小时、21小时、22小时、23小时、1天、36小时、2天、3天、4天、5天、6天、1周、10天、2周、3周、1个月、2个月、3个月、4个月、5个月、6个月、7个月、8个月、9个月、10个月、11个月、1年、18个月、2年、3年、4年、5年、6年、7年、8年、9年、10年、11年、12年、13年、14年、15年、16年、17年、18年、19年、20年、25年、30年、35年、40年、45年、50年、55年、60年、65年、70年、75年、80年、85年、90年、95年或超过99年。在示例性实施方案中,在防治性组合物的初始施用与加强剂之间的施用时间可以是但不限于1周、2周、3周、1个月、2个月、3个月、6个月或1年。

[0323] 所述癌症疫苗可用于各种环境中,这取决于癌症的严重性或未满足医学需求的程度或水平。作为非限制性实例,所述癌症疫苗可用于治疗任何阶段的癌症。

[0324] 下面给出所述癌症疫苗可治疗的癌症的非限制性清单。肽表位或抗原可源自这些癌症或肿瘤的任何抗原。此类表位可被称为癌症或肿瘤抗原。癌细胞可在肿瘤进展的不同时期期间差异性地表达细胞表面分子。例如,癌细胞可以良性状态表达细胞表面抗原,但在转移后使那个特定细胞表面抗原下调。因此,设想肿瘤抗原或癌症抗原可涵盖在癌症进展的任何阶段期间产生的抗原。可调整本公开的方法以适应这些变化。例如,可针对特定患者制取若干不同癌症疫苗。例如,可在治疗开始时使用第一疫苗。在稍后的时间点,可制取新的癌症疫苗,并且将其施用给患者以虑及所表达的不同抗原。

[0325] 在一些实施方案中,肿瘤抗原是以下抗原中的一种:CD2、CD19、CD20、CD22、CD27、CD33、CD37、CD38、CD40、CD44、CD47、CD52、CD56、CD70、CD79、CD137、4-1BB、5T4、AGS-5、AGS-16、血管生成素2、B7.1、B7.2、B7DC、B7H1、B7H2、B7H3、BT-062、BTLA、CAIX、癌胚抗原、CTLA4、Cripto、ED-B、ErbB1、ErbB2、ErbB3、ErbB4、EGFL7、EpCAM、EphA2、EphA3、EphB2、FAP、纤连蛋白、叶酸受体、神经节苷脂GM3、GD2、糖皮质激素诱导的肿瘤坏死因子受体(GITR)、gp100、gpA33、GPNMB、ICOS、IGF1R、整合素 α v、整合素 α v β 、LAG-3、Lewis Y、间皮素、c-MET、MN碳酸酐酶IX、MUC1、MUC16、结合素-4、NKGD2、NOTCH、OX40、OX40L、PD-1、PDL1、PSCA、PSMA、RANKL、ROR1、ROR2、SLC44A4、多配体聚糖-1、TACI、TAG-72、腱生蛋白、TIM3、TRAILR1、TRAILR2、VEGFR-1、VEGFR-2、VEGFR-3,以及它们的变体。

[0326] 癌症或肿瘤包括但不限于赘瘤、恶性肿瘤、转移癌,或特征在于细胞生长不受控制以致它将被视为具有癌性的任何疾病或病症。癌症可以是原发性或转移性癌症。可根据本公开治疗的特定癌症包括但不限于下面列出的那些(对于此类病症的综述,参见Fishman等人,1985,Medicine,第2版,J.B.Lippincott Co.,Philadelphia)。用于与本发明描述的方法和组合物一起使用的癌症可包括但不限于胆道癌;膀胱癌;脑癌,包括胶质母细胞瘤和神经管母细胞瘤;乳腺癌;子宫颈癌;绒毛膜癌;结肠癌;子宫内膜癌;食道癌;胃癌;血液学赘瘤,包括急性淋巴细胞性和骨髓性白血病;多发性骨髓瘤;AIDS相关的白血病和成人T细胞白血病淋巴瘤;上皮内赘瘤,包括鲍恩氏病(Bowen's disease)和佩吉特氏病(Paget's disease);肝癌;肺癌;淋巴瘤,包括霍奇金氏病(Hodgkin's disease)和淋巴细胞性淋巴瘤;神经母细胞瘤;口腔癌,包括鳞状细胞癌;卵巢癌,包括由上皮细胞、基质细胞、生殖细胞和间质细胞引起的那些;胰腺癌;前列腺癌;直肠癌;肉瘤,包括平滑肌肉瘤、横纹肌肉瘤、脂肪肉瘤、纤维肉瘤和骨肉瘤;皮肤癌,包括黑素瘤、卡波济氏肉瘤(Kaposi's sarcoma)、基底细胞癌(basocellular cancer)和鳞状细胞癌;睾丸癌,包括生发肿瘤(germinal tumor),诸

如精原细胞瘤、非精原细胞瘤、畸胎瘤；高肿瘤突变负荷肿瘤；绒毛膜癌；基质肿瘤和生殖细胞肿瘤；甲状腺癌，包括甲状腺腺癌和髓癌；以及肾癌，包括腺癌和威尔姆斯肿瘤 (Wilms' tumor)。在一些实施方案中，所述癌症是黑素瘤、膀胱癌、HPV阴性HNSCC、NSCLC、SCLC、高MSI肿瘤或高TMB(肿瘤突变负荷)癌症中的任一种。

[0327] 在一些实施方案中，所述癌症选自由以下组成的组：非小细胞肺癌 (NSCLC)、小细胞肺癌、黑素瘤、膀胱尿路上皮癌、HPV阴性头颈鳞状细胞癌 (HNSCC) 以及有高度微卫星不稳定性 (MSI H) / 错配修复 (MMR) 缺陷的实体恶性肿瘤。在一些实施方案中，NSCLC缺乏EGFR敏化突变和/或ALK易位。在一些实施方案中，有高度微卫星不稳定性 (MSI H) / 错配修复 (MMR) 缺陷的实体恶性肿瘤选自由以下组成的组：结肠直肠癌、胃腺癌、食道腺癌和子宫内膜癌。

[0328] 本文提供了药物组合物，所述药物组合物包含任选地与一种或多种药学上可接受的赋形剂组合在一起的癌症疫苗和RNA疫苗组合物和/或复合物。所述癌症疫苗可单独地或者与一种或多种如本文所述的其他组分结合配制或施用。

[0329] 在其他实施方案中，本文所述的癌症疫苗可与可用于治疗患者的任何其他疗法组合。例如，可用癌症疫苗和抗癌剂治疗患者。因此，在一个实施方案中，本公开的方法可与一种或多种癌症治疗结合使用，诸如与抗癌剂、传统癌症疫苗、化学疗法、放射疗法等结合使用(例如，同时使用，或作为总体治疗程序的一部分来使用)。可变化的癌症治疗参数包括但不限于剂量、施用时间安排或持续时间或疗法；并且癌症治疗可在剂量、时间安排或持续时间方面变化。另一种癌症治疗是手术，其可单独地或与任何先前治疗方法组合使用。已知可用于或者已用于或者当前正用于预防或治疗癌症的任何剂或疗法(例如，传统癌症疫苗、化学疗法、放射疗法、手术、激素疗法和/或生物疗法/免疫疗法)可与本文所述的根据本公开的本公开组合物组合使用。医学领域中的普通技术人员可确定对受试者的适当治疗。

[0330] 此类剂(即，抗癌剂)的实例包括但不限于DNA相互作用剂，包括但不限于烷基化剂(例如，氮芥(nitrogen mustard)，例如苯丁酸氮芥(Chlorambucil)、环磷酰胺(Cyclophosphamide)、异环磷酰胺(Isofamide)、二氯甲基二乙胺(Mechlorethamine)、美法仑(Melphalan)、尿嘧啶氮芥(Uracil mustard)；氮丙啶(Aziridine)，诸如噻替派(Thiotepa)；甲烷磺酸酯，诸如白消安(Busulfan)；亚硝基脲，诸如卡莫司汀(Carmustine)、洛莫司汀(Lomustine)、链脲霉素(Streptozocin)；铂络合物，诸如顺铂(Cisplatin)、卡铂(Carboplatin)；生物还原性烷基化剂，诸如丝裂霉素(Mitomycin)以及丙卡巴肼(Procarbazine)、达卡巴嗪(Dacarbazine)和六甲蜜胺(Altretamine))；DNA链断裂剂，例如博来霉素(Bleomycin)；嵌入性拓扑异构酶II抑制剂，例如嵌入剂，诸如安吖啶(Amsacrine)、更生霉素(Dactinomycin)、柔红霉素(Daunorubicin)、多柔比星(Doxorubicin)、伊达比星(Idarubicin)、米托蒽醌(Mitoxantrone)，以及非嵌入剂，诸如依托泊苷(Etoposide)和替尼泊苷(Teniposide)；非嵌入性拓扑异构酶II抑制剂，例如依托泊苷和替尼泊苷；以及DNA小沟结合物，例如普卡霉素(Plicamycin)；抗代谢剂，包括但不限于叶酸拮抗剂，诸如甲氨蝶呤(Methotrexate)和三甲曲沙(trimetrexate)；嘧啶拮抗剂，诸如氟尿嘧啶(Fluorouracil)、氟脱氧尿苷(Fluorodeoxyuridine)、CB3717、阿扎胞苷(Azacitidine)和氟尿苷(Floxuridine)；嘌呤拮抗剂，诸如巯基嘌呤(Mercaptopurine)、6-硫鸟嘌呤(6-Thioguanine)、喷司他汀(Pentostatin)；经糖修饰的类似物，诸如阿糖胞苷(Cytarabine)和氟达拉滨(Fludarabine)；以及核糖核苷酸还原酶抑制剂，诸如羟基脲

(hydroxyurea);微管蛋白(tubulin)相互作用剂,包括但不限于秋水仙素(colchicine)、长春新碱(Vincristine)和长春花碱(Vinblastine)(两者均是生物碱),以及紫杉醇(Paclitaxel)和癌得星(cytoxan);激素剂,包括但不限于雌激素(estrogen)、缀合雌激素以及乙炔雌二醇(Ethinyl Estradiol)和己烯雌酚(Diethylstilbesterol)、氯烯雌醚(Chlortrianisen)和双烯雌酚(Idenestrol);孕激素(progestin),诸如己酸羟孕酮(Hydroxyprogesterone caproate)、甲羟孕酮(Medroxyprogesterone)和甲地孕酮(Megestrol);以及雄激素(androgen),诸如睾酮(testosterone)、丙酸睾酮(testosterone propionate);氟羟甲睾酮(fluxymesterone)、甲基睾酮(methyltestosterone);肾上腺皮质类固醇,例如泼尼松(Prednisone)、地塞米松(Dexamethasone)、甲基泼尼松龙(Methylprednisolone)和泼尼松龙(Prednisolone);促黄体生成激素释放激素剂或促性腺激素释放激素拮抗剂,例如乙酸亮丙瑞林(leuprolide acetate)和乙酸戈舍瑞林(goserelin acetate);抗激素抗原,包括但不限于抗雌激素剂诸如他莫昔芬(Tamoxifen)、抗雄激素剂诸如氟他胺(Flutamide);以及抗肾上腺素能剂,诸如米托坦(Mitotane)和胺鲁米特(Aminoglutethimide);细胞因子,包括但不限于IL-1 α 、IL-1 β 、IL-2、IL-3、IL-4、IL-5、IL-6、IL-7、IL-8、IL-9、IL-10、IL-11、IL-12、IL-13、IL-18、TGF- β 、GM-CSF、M-CSF、G-CSF、TNF- α 、TNF- β 、LAF、TCGF、BCGF、TRF、BAF、BDG、MP、LIF、OSM、TMF、PDGF、IFN- α 、IFN- β 、IFN- γ 和子宫球蛋白(Uteroglobulin)(美国专利号5,696,092);抗血管生成剂,包括但不限于抑制VEGF的剂(例如其他中和抗体)、可溶性受体构建体、酪氨酸激酶抑制剂、反义策略、针对VEGF或VEGF受体的RNA适体和核酶、免疫毒素和共凝集配体(coaguligand)、肿瘤疫苗和抗体。

[0331] 可根据本公开的方法使用的抗癌剂的特定实例包括但不限于:阿西维辛(acivicin);阿柔比星(aclarubicin);盐酸阿考达唑(acodazole hydrochloride);阿克罗宁(acronine);阿多来新(adozelesin);阿地白介素(aldesleukin);六甲蜜胺(altretamine);安波霉素(ambomycin);乙酸阿美蒽醌(ametanttrone acetate);胺鲁米特(aminoglutethimide);安吡啶(amsacrine);阿那曲唑(anastrozole);安曲霉素(anthramycin);天冬酰胺酶(asparaginase);曲林菌素(asperlin);阿扎胞苷(azacitidine);阿替派(azetepa);阿佐霉素(azotomycin);巴马司他(batimastat);苯佐替派(benzodepa);比卡鲁胺(bicalutamide);盐酸比生群(bisantrene hydrochloride);二甲磺酸双奈法德(bisnafide dimesylate);比折来新(bizelesin);硫酸博来霉素(bleomycin sulfate);布喹那钠(brequinar sodium);溴匹立明(bropirimine);白消安(busulfan);放线菌素(cactinomycin);卡鲁睾酮(calusterone);卡醋胺(caracemide);卡贝替姆(carbetimer);卡铂(carboplatin);卡莫司汀(carmustine);盐酸卡柔比星(carubicin hydrochloride);卡折来新(carzelesin);西地芬戈(cedefingol);苯丁酸氮芥(chlorambucil);西罗霉素(cirolemycin);顺铂(cisplatin);克拉屈滨(cladribine);甲磺酸克立那托(crisnatol mesylate);环磷酰胺(cyclophosphamide);阿糖胞苷(cytarabine);达卡巴嗪(dacarbazine);更生霉素(dactinomycin);盐酸柔红霉素(daunorubicin hydrochloride);地西他滨(decitabine);右奥马铂(dexormaplatin);地扎胍宁(dezaguanine);甲磺酸地扎胍宁(dezaguanine mesylate);地吡醌(diaziquone);多西他赛(docetaxel);多柔比星(doxorubicin);盐酸多柔比星(doxorubicin

hydrochloride); 屈洛昔芬(droloxifene); 柠檬酸屈洛昔芬(droloxifene citrate); 丙酸屈他雄酮(dromostanolone propionate); 达佐霉素(duazomycin); 依达曲沙(edatrexate); 盐酸依氟鸟氨酸(eflomithine hydrochloride); 依沙芦星(elsamitrucin); 恩洛铂(enloplatin); 恩普氨酯(enpromate); 依匹哌啉(epipropidine); 盐酸表柔比星(epirubicin hydrochloride); 厄布洛唑(erbulozole); 盐酸依索比星(esorubicin hydrochloride); 雌氮芥(estramustine); 雌氮芥磷酸钠(estramustine phosphate sodium); 依他硝唑(etanidazole); 依托泊苷(etoposide); 磷酸依托泊苷(etoposide phosphate); 艾托卜宁(etoprine); 盐酸法倔唑(fadrozole hydrochloride); 法扎拉滨(fazarabine); 芬维A胺(fenretinide); 氟尿苷(floxuridine); 磷酸氟达拉滨(fludarabine phosphate); 氟尿嘧啶(flourouracil); 氟西他滨(flurocitabine); 磷喹酮(fosquidone); 福司曲星钠(fostriecin sodium); 吉西他滨(gemcitabine); 盐酸吉西他滨(gemcitabine hydrochloride); 羟基脲(hydroxyurea); 盐酸伊达比星(idarubicin hydrochloride); 异环磷酰胺(ifosfamide); 依莫佛新(ilmofofosine); 白介素II(包括重组白介素II或rIL2)、干扰素 α -2a; 干扰素 α -2b; 干扰素 α -n1; 干扰素 α -n3; 干扰素 β -1a; 干扰素 γ -1b; 异丙铂(iproplatin); 盐酸伊立替康(irinotecan hydrochloride); 乙酸兰瑞肽(lanreotide acetate); 来曲唑(letrozole); 乙酸亮丙瑞林(leuprolide acetate; 盐酸利阿唑(liarozole hydrochloride); 洛美曲索钠(lometrexol sodium); 洛莫司汀(lomustine); 盐酸洛索蒽醌(losoxantrone hydrochloride); 马丙考(masoprocol); 美登素(maytansine); 盐酸二氯甲基二乙胺(mechlorethamine hydrochloride); 乙酸甲地孕酮(megestrol acetate); 乙酸美仑孕酮(melengestrol acetate); 美法仑(melphalan); 美诺立尔(menogaril); 巯基嘌呤(mercaptapurine); 甲氨蝶呤(methotrexate); 甲氨蝶呤钠(methotrexate sodium); 氯苯氨啉(metoprine); 美乌替派(meturedopa); 米丁度胺(mitindomide); 米托克星(mitocarcin); 丝裂红素(mitocromin); 丝裂吉霉素(mitogillin); 丝裂马菌素(mitomalcin); 丝裂霉素(mitomycin); 米托司培(mitosper); 米托坦(mitotane); 盐酸米托蒽醌(mitoxantrone hydrochloride); 霉酚酸(mycophenolic acid); 诺考达唑(nocodazole); 诺加霉素(nogalamycin); 奥马铂(ormaplatin); 奥昔舒仑(oxisuran); 紫杉醇(paclitaxel); 培加帕酶(pegaspargase); 佩里霉素(peliomycin); 戊氮芥(pentamustine); 硫酸培洛霉素(peplomycin sulfate); 培磷酰胺(perfosfamide); 哌泊溴烷(pipobroman); 保释芬(piposulfan); 盐酸吡罗蒽醌(piroxantrone hydrochloride); 光神霉素(plicamycin); 普洛美坦(plomestane); 吡吩姆钠(porfimer sodium); 波福霉素(porfiromycin); 松龙苯芥(prednimustine); 盐酸丙卡巴肼(procarbazine hydrochloride); 嘌呤霉素(puromycin); 盐酸嘌呤霉素(puromycin hydrochloride); 吡唑呋喃菌素(pyrazofurin); 利波腺苷(riboprime); 洛太米特(roglitimide); 沙芬戈(safingol); 盐酸沙芬戈(safingol hydrochloride); 司莫司汀(semustine); 辛曲秦(simtrazene); 磷乙酰天冬氨酸钠(sparfosate sodium); 稀疏霉素(sparsomycin); 盐酸螺旋锗(spirogermanium hydrochloride); 螺莫司汀(spiromustine); 螺铂(spiroplatin); 链黑菌素(streptonigrin); 链脲霉素(streptozocin); 磺氯苯脲(sulofenur); 他利霉素(talisomycin); 替可加兰钠(tecogalan sodium); 替加氟(tegafur); 盐酸替洛蒽醌(teloxantrone hydrochloride); 替莫泊芬

(temoporfin);替尼泊苷(teniposide);替罗昔隆(teroxirone);睾内酯(testolactone);硫咪嘌呤(thiamiprine);硫鸟嘌呤(thioguanine);噻替派(thiotepa);噻唑呋林(tiazofurin);替拉扎明(tirapazamine);柠檬酸托瑞米芬(toremifene citrate);乙酸曲托龙(trestolone acetate);磷酸曲西立滨(triciribine phosphate);三甲曲沙(trimetrexate);葡萄糖醛酸三甲曲沙(trimetrexate glucuronate);曲普瑞林(triptorelin);盐酸妥布氯唑(tubulozole hydrochloride);尿嘧啶氮芥(uracil mustard);乌瑞替派(uredepa);伐普肽(vapreotide);维替泊芬(verteporfin);硫酸长春花碱(vinblastine sulfate);硫酸长春新碱(vincristine sulfate);长春地辛(vindesine);硫酸长春地辛(vindesine sulfate);硫酸长春匹定(vinepidine sulfate);硫酸长春甘酯(vinglycinate sulfate);硫酸长春罗辛(vinleurosine sulfate);酒石酸长春瑞滨(vinorelbine tartrate);硫酸长春罗定(vinrosidine sulfate);硫酸长春利定(vinzolidine sulfate);伏氯唑(vorozole);折尼铂(zeniplatin);净司他汀(zinostatin);以及盐酸佐柔比星(zorubicin hydrochloride)。

[0332] 可与本发明的组合物和方法一起使用的其他抗癌药物包括但不限于:20-表-1,25二羟基维生素D₃;5-乙炔基尿嘧啶;血管生成抑制剂;抗背侧化形态发生蛋白-1;ara-CDP-DL-PTBA;BCR/ABL拮抗剂;CaRest M3;CARN 700;酪蛋白激酶抑制剂(ICOS);克霉唑(clotrimazole);碰撞霉素A(collismycin A);碰撞霉素B;考布他汀A4(combretastatin A4);甘蓝海绵素816(crambescidin 816);念珠藻素8(cryptophycin 8);库瑞辛A(curacin A);脱氢膜海鞘素B(dehydrodidemnin B);膜海鞘素B(didemnin B);二氢-5-氮杂胞苷(dihydro-5-azacytidine);二氢泰素(dihydrotaxol)、倍癌霉素SA(duocarmycin SA);卡哈拉肽F(kahalalide F);三乙酸片螺素-N(lamellarin-N triacetate);亮脯利特(leuprolide)+雌激素+孕酮;利索纳得7(lissoclinamide 7);单磷酸基脂质A+分枝杆菌细胞壁sk;N-乙酰基地那林(N-acetyldinaline);N-取代的苯甲酰胺;06-苯甲基鸟嘌呤;普拉色汀A(placetin A);普拉色汀B;铂络合物;铂化合物;铂-三胺络合物;依替膦酸铼Re 186(rhenium Re 186etidronate);RII维甲酰胺(RII retinamide);鲁比津酮B1(rubiginone B1);SarCNU;肌肉叶绿醇A(sarcophytol A);沙莫司亭(sargramostim);衰老源性抑制剂1;斯卡霉素D(spicamycin D);他莫司汀(tallimustine);5-氟尿嘧啶;血小板生成素(thrombopoietin);胸腺曲南(thymotrinan);甲状腺刺激激素;瓦立奥林B(variolin B);沙利度胺(thalidomide);维拉雷琐(velaresol);藜芦明(veramine);瓦尔丁(verdins);维替泊芬(verteporfin);长春瑞滨(vinorelbine);威科萨汀(vinxaltine);维塔辛(vitaxin);扎诺特隆(zanoterone);折尼铂(zeniplatin);以及亚苾维C(zilasorb)。

[0333] 本公开还涵盖包含癌症疫苗的组合物与包括使用x射线、γ射线和其他放射源来破坏癌细胞的放射疗法的组合施用。在某些实施方案中,放射治疗以外部射束放射或远距疗法形式施用,其中放射是从远程源(remote source)引导的。在其他实施方案中,放射治疗以内部疗法或近距疗法形式施用,其中放射源被置于身体内靠近癌细胞或肿瘤块的位置。

[0334] 在特定实施方案中,(例如,由医师)根据癌症类型选择适当的抗癌方案。例如,可将防治或治疗有效量的包含癌症疫苗的组合物与防治或治疗有效量的一种或多种其他可用于卵巢癌疗法的剂组合施用给患有卵巢癌的患者,所述其他可用于卵巢癌疗法的剂包括

但不限于腹膜内放射疗法诸如P32疗法、总体腹部和骨盆放射疗法、顺铂、紫杉醇(泰索)或多西他赛(泰索帝(Taxotere))与顺铂或卡铂的组合、环磷酰胺与顺铂的组合、环磷酰胺与卡铂的组合、5-FU与甲酰四氢叶酸的组合、依托泊苷、脂质体多柔比星、吉西他滨或拓扑替康(topotecan)。癌症疗法和它们的剂量、施用途和推荐用法在本领域中是已知的,并且已描述于诸如Physician's Desk Reference(第56版,2002)的文献中。

[0335] 在本公开的一些实施方案中,所述癌症疫苗与T细胞活化剂诸如免疫检查点调节剂一起施用。免疫检查点调节剂包括刺激性检查点分子与抑制性检查点分子两者(例如,抗CTLA4抗体和/或抗PD1抗体)。

[0336] 刺激性检查点抑制剂通过促进检查点过程来起作用。数种刺激性检查点分子是肿瘤坏死因子(TNF)受体超家族的成员(例如,CD27、CD40、OX40、GITR或CD137),而其他属于B7-CD28超家族(例如,CD28或ICOS0)。OX40(CD134)参与效应T细胞和记忆T细胞的扩增。抗OX40单克隆抗体已被证明在治疗晚期癌症中有效。MEDI0562是人源化的OX40激动剂。糖皮质激素诱导的TNFR家族相关基因GITR参与T细胞扩增。针对GITR的数种抗体已被证明能促进抗肿瘤应答。诱导型T细胞共刺激物ICOS在T细胞效应子功能中是重要的。CD27支持初始T细胞的抗原特异性扩增,并且参与T细胞和B细胞记忆的产生。数种激动性抗CD27抗体处于开发中。CD122是白介素2受体 β 亚基。NKTR-214是CD122偏向型免疫刺激性细胞因子。

[0337] 抑制性检查点分子包括但不限于PD-1、TIM-3、VISTA、A2AR、B7-H3、B7-H4、BTLA、CTLA-4、IDO、KIR和LAG3。CTLA-4、PD-1和它的配体是遍及T细胞功能和其他细胞功能的所有阶段起重要作用的共同信号传导分子的CD28-B7家族的成员。细胞毒性T淋巴细胞相关蛋白4CTLA-4(CD152)参与控制T细胞增殖。

[0338] PD-1受体在活化的T细胞(和B细胞)的表面上表达,并且在正常情况下结合至它的在诸如树突状细胞或巨噬细胞的抗原呈递细胞的表面上表达的配体(PD-L1和PD-L2)。这个相互作用将信号传送到T细胞中,并且抑制T细胞。癌细胞通过驱动PD-L1在它们的表面上高水平表达来利用这个系统。这使它们能够获得对PD-1路径的控制,并且切断可进入肿瘤微环境中的表达PD-1的T细胞,由此抑制抗癌免疫应答。派姆单抗(Pembrolizumab)(先前是MK-3475和兰罗利珠单抗(lambrolizumab),商品名是Keytruda)是一种用于癌症免疫疗法中的人抗体,并且靶向PD-1受体。

[0339] 检查点抑制剂是诸如单克隆抗体、人源化抗体、完全人抗体、融合蛋白或它们的组合或小分子的分子。例如,检查点抑制剂抑制检查点蛋白,其可以是CTLA-4、PDL1、PDL2、PD1、B7-H3、B7-H4、BTLA、HVEM、TIM3、GAL9、LAG3、VISTA、KIR、2B4、CD160、CGEN-15049、CHK1、CHK2、A2aR、B-7家族配体或它们的组合。检查点蛋白的配体包括但不限于CTLA-4、PDL1、PDL2、PD1、B7-H3、B7-H4、BTLA、HVEM、TIM3、GAL9、LAG3、VISTA、KIR、2B4、CD160、CGEN-15049、CHK 1、CHK2、A2aR和B-7家族配体。在一些实施方案中,抗PD-1抗体是BMS-936558(纳武单抗(nivolumab))。在其他实施方案中,抗CTLA-4抗体是易普利单抗(ipilimumab)(商品名是Yervoy,先前称为MDX-010和MDX-101)。

[0340] 在一些实施方案中,包括检查点调节剂在内的癌症治疗剂以编码癌症治疗剂的mRNA形式递送。

[0341] 在一些实施方案中,癌症治疗剂是靶向疗法。靶向疗法可以是BRAF抑制剂,诸如维罗非尼(vemurafenib)(PLX4032)或达拉菲尼(dabrafenib)。BRAF抑制剂可以是PLX 4032、

PLX 4720、PLX 4734、GDC-0879、PLX 4032、PLX-4720、PLX 4734和甲苯磺酸索拉非尼。BRAF是产生称为B-Raf的蛋白质的人基因,也被称为原致癌基因B-Raf和v-Raf鼠肉瘤病毒致癌基因同源物B1。B-Raf蛋白参与在细胞内部传送信号,所述信号参与指导细胞生长。BRAF抑制剂维罗菲尼由FDA批准用于治疗晚期黑素瘤。

[0342] 在其他实施方案中,癌症治疗剂是细胞因子。在又一些其他实施方案中,癌症治疗剂是包含基于群体的肿瘤特异性抗原的疫苗。在又一些其他实施方案中,癌症治疗剂是这样的疫苗,其含有一种或多种由癌症种系基因表达的传统抗原(在多名患者中发现的肿瘤所共同具有的抗原,也称为“共有癌症抗原”)。在一些实施方案中,传统抗原是已知普遍在癌症或肿瘤中被发现,或者在特定类型的癌症或肿瘤中被发现的抗原。在一些实施方案中,传统癌症抗原是非突变的肿瘤抗原。在一些实施方案中,传统癌症抗原是突变的肿瘤抗原。

[0343] 在人癌症的情况下,相比于任何其他基因,p53基因(官方符号TP53)更频繁地突变。大量群组研究已证明对于大多数p53突变,基因组位置是一名患者或仅少许患者所特有的,并且该突变不能用作针对特定患者群体设计的治疗性疫苗的频发新抗原。然而,p53基因座的小型子集确实表现出“热点”样式(在本文别处描述),其中基因中的若干位置以相对较高频率发生突变。引人注目的是,这些频发突变区域中的大部分发生在外显子-内含子边界附近,从而破坏由mRNA剪接机构辨识的规范的核苷酸序列基序。

[0344] 剪接基序的突变可更改最终mRNA序列,即使预测到局部氨基酸序列无变化(即,对于同义突变或内含子突变来说)。因此,这些突变常常被常见的注释工具注释为“非编码”,并且在进一步分析时被忽略,即使它们可以不可预测方式更改mRNA剪接并且对所翻译的蛋白质造成严重功能性影响。如果选择性剪接同种型产生同框序列变化(即,不产生提前终止密码子(PTC)),那么它可逃脱由无义介导的mRNA降解(NMD)所致的耗竭,并且容易被HLA系统表达、加工,以及呈递在细胞表面上。进一步地,突变源性选择性剪接通常是“隐蔽的”,即不在正常组织中表达,因此可被T细胞辨识为非自身新抗原。

[0345] 在一些情况下,癌症治疗剂是疫苗,其包含一种或多种新抗原,该新抗原为频发多态性(“热点突变”)。例如,除别的之外,本公开提供了由p53中的某些频发体细胞癌症突变产生的新抗原肽序列。

[0346] 制剂

[0347] 癌症疫苗(例如,核酸癌症疫苗诸如mRNA癌症疫苗)可与一种或多种药学上可接受的赋形剂组合配制或施用。作为一组非限制性实例,可使用一种或多种赋形剂来配制癌症疫苗,以:(1)增大稳定性;(2)增加细胞转染;(3)允许持续释放或延迟释放(例如,从贮库制剂中);(4)更改生物分布(例如,针对特定的组织或细胞类型);(5)增加体内编码的蛋白质的翻译;和/或(6)更改体内编码的蛋白质(抗原)的释放特征。除传统的赋形剂诸如任何和所有溶剂、分散介质、稀释剂或其他液体媒介物、分散助剂或悬浮助剂、表面活性剂、等渗剂、增稠剂或乳化剂、防腐剂之外,赋形剂还可包括但不限于类脂质、脂质体、脂质纳米粒子、聚合物、脂质体复合物、核-壳纳米粒子、肽、蛋白质、用癌症疫苗转染的细胞(例如,用于移植到受试者中的细胞)、透明质酸酶、纳米粒子模拟物,以及它们的组合。

[0348] 在一些实施方案中,疫苗组合物包含至少一种另外的活性物质,诸如例如治疗活性物质、防治活性物质或两者的组合。疫苗组合物可以是无菌的,无热原的,或者无菌的且无热原的。在配制和/或制造药物制剂诸如疫苗组合物时的一般考虑因素可例如见于

Remington: The Science and Practice of Pharmacy第21版, Lippincott Williams & Wilkins, 2005 (出于此目的以引用方式整体并入本文) 中。

[0349] 在一些实施方案中, 向人、人患者或受试者施用癌症疫苗。出于本公开的目的, 短语“活性成分”一般是指癌症疫苗或其中含有的核酸, 例如编码抗原性多肽的RNA (例如, mRNA)。

[0350] 本文所述的疫苗组合物的制剂可通过药理学领域中已知或今后开发的任何方法来制备。一般来讲, 此类制备方法包括以下步骤: 将活性成分 (例如, 核酸诸如mRNA) 与赋形剂和/或者一种或多种其他辅助成分联合在一起, 然后如果必要和/或合乎需要, 将产品划分、成形和/或包装成期望的单次或多次剂量单位。

[0351] 本文所公开的任何组合物的制剂除包含上述组分外还可包含一种或多种组分。例如, 脂质组合物可包含一种或多种渗透性增强剂分子、碳水化合物、聚合物、表面更改剂 (例如, 表面活性剂), 或其他组分。例如, 渗透性增强剂分子可以是美国专利申请公布号2005/0222064所描述的分子。碳水化合物可包括单糖 (例如, 葡萄糖) 和多糖 (例如, 糖原及其衍生物和类似物)。

[0352] 聚合物可包含在本文所公开的药物组合物 (例如, 脂质纳米粒子形式的药物组合物) 中, 以及/或者用来包封或部分包封本文所公开的药物组合物 (例如, 脂质纳米粒子形式的药物组合物)。聚合物可以是可生物降解的和/或生物相容的。聚合物可选自但不限于聚胺、聚醚、聚酰胺、聚酯、聚氨基甲酸酯、聚脲、聚碳酸酯、聚苯乙烯、聚酰亚胺、聚砜、聚氨基甲酸乙酯、聚乙炔、聚乙烯、聚乙烯亚胺、聚异氰酸酯、聚丙烯酸酯、聚甲基丙烯酸酯、聚丙烯腈和聚芳酯。

[0353] 在一些实施方案中, 本文所公开的组合物可被配制成脂质纳米粒子 (LNP)。因此, 本公开还提供了纳米粒子组合物, 所述纳米粒子组合物包含: (i) 包含递送剂的脂质组合物, 和 (ii) 编码一个或多个肽表位的核酸。在此种纳米粒子组合物中, 本文所公开的脂质组合物可包封编码一个或多个肽表位的核酸。

[0354] 纳米粒子组合物的尺寸通常为微米级或更小, 并且可包括脂质双层。纳米粒子组合物涵盖脂质纳米粒子 (LNP)、脂质体 (例如, 脂质囊泡) 和脂质体复合物。例如, 纳米粒子组合物可以是具有直径为500nm或更小的脂质双层的脂质体。

[0355] 纳米粒子组合物包括例如脂质纳米粒子 (LNP)、脂质体和脂质体复合物。在一些实施方案中, 纳米粒子组合物是包含一个或多个脂质双层的囊泡。在某些实施方案中, 纳米粒子组合物包含由水性隔室隔开的两个或更多个同心双层。脂质双层可被官能化和/或彼此交联。脂质双层可包含一种或多种配体、蛋白质或通道。

[0356] 在一个实施方案中, 脂质纳米粒子包含可电离脂质、结构脂质、磷脂和mRNA。在一些实施方案中, LNP包含可电离脂质、经PEG修饰的脂质、磷脂和结构脂质。

[0357] 脂质组合物与癌症疫苗之间的比率可以是约10:1至约60:1 (wt/wt)。在一些实施方案中, 脂质组合物与核酸之间的比率可以是约10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、21:1、22:1、23:1、24:1、25:1、26:1、27:1、28:1、29:1、30:1、31:1、32:1、33:1、34:1、35:1、36:1、37:1、38:1、39:1、40:1、41:1、42:1、43:1、44:1、45:1、46:1、47:1、48:1、49:1、50:1、51:1、52:1、53:1、54:1、55:1、56:1、57:1、58:1、59:1或60:1 (wt/wt)。在一些实施方案中, 脂质组合物与癌症疫苗的wt/wt比率是约20:1或约15:1。

[0358] 在一个实施方案中,癌症疫苗(例如,核酸癌症疫苗)可包含在脂质纳米粒子中,使得脂质:多核苷酸重量比为5:1、10:1、15:1、20:1、25:1、30:1、35:1、40:1、45:1、50:1、55:1、60:1或70:1,或这些比率中的范围或任一者,诸如但不限于5:1至约10:1、约5:1至约15:1、约5:1至约20:1、约5:1至约25:1、约5:1至约30:1、约5:1至约35:1、约5:1至约40:1、约5:1至约45:1、约5:1至约50:1、约5:1至约55:1、约5:1至约60:1、约5:1至约70:1、约10:1至约15:1、约10:1至约20:1、约10:1至约25:1、约10:1至约30:1、约10:1至约35:1、约10:1至约40:1、约10:1至约45:1、约10:1至约50:1、约10:1至约55:1、约10:1至约60:1、约10:1至约70:1、约15:1至约20:1、约15:1至约25:1、约15:1至约30:1、约15:1至约35:1、约15:1至约40:1、约15:1至约45:1、约15:1至约50:1、约15:1至约55:1、约15:1至约60:1,或约15:1至约70:1。

[0359] 在一个实施方案中,癌症疫苗(例如,核酸癌症疫苗)可以近似0.1mg/ml至2mg/ml,诸如但不限于0.1mg/ml、0.2mg/ml、0.3mg/ml、0.4mg/ml、0.5mg/ml、0.6mg/ml、0.7mg/ml、0.8mg/ml、0.9mg/ml、1.0mg/ml、1.1mg/ml、1.2mg/ml、1.3mg/ml、1.4mg/ml、1.5mg/ml、1.6mg/ml、1.7mg/ml、1.8mg/ml、1.9mg/ml、2.0mg/ml或大于2.0mg/ml的浓度包含在脂质纳米粒子中。

[0360] 如本文一般定义的,术语“脂质”是指具有疏水特性或两亲特性的小分子。脂质可以是天然存在的或合成的。脂质类别的实例包括但不限于脂肪、蜡、含甾醇的代谢物、维生素、脂肪酸、甘油脂质、甘油磷脂、鞘脂、糖脂和聚酮化合物,以及孕烯醇酮脂质(prenol lipid)。在一些情况下,一些脂质的两亲特性导致它们在水性介质中形成脂质体、囊泡或膜。

[0361] 在一些实施方案中,脂质纳米粒子(LNP)可包含可电离脂质。如本文所用,术语“可电离脂质”具有其在本领域中的普通含义,并且可指包含一个或多个带电荷部分的脂质。在一些实施方案中,可电离脂质可带正电荷或带负电荷。可电离脂质可带正电荷,在这种情况下,它可被称为“阳离子脂质”。在某些实施方案中,可电离脂质分子可包含胺基团,并且可被称为可电离氨基脂质。如本文所用,“带电荷部分”是携带形式电子电荷的化学部分,例如单价(+1或-1)、二价(+2或-2)、三价(+3或-3)等。带电荷部分可以是阴离子型的(即,带负电荷)或阳离子型的(即,带正电荷)。带正电荷部分的实例包括胺基团(例如,伯胺、仲胺和/或叔胺)、铵基团、吡啶鎓基团、胍基和咪唑鎓基团。在特定实施方案中,带电荷部分包含胺基团。带负电荷基团或其前体的实例包括羧酸根基团、磺酸根基团、硫酸根基团、磷酸根基团、磷酸根基团、羟基等。在一些情况下,带电荷部分的电荷可随环境条件而变化,例如,pH的变化可更改所述部分的电荷,以及/或者使所述部分变得带电荷或不带电荷。一般来讲,可根据需要选择分子的电荷密度。可电离脂质也可以是公开于国际公布号:W02017075531、W02015199952、W02013086354或W02013116126中的化合物,或者选自美国专利号7,404,969的式CLI-CLXXXXII;所述文献中的每一篇出于此目的以引用方式整体并入本文。

[0362] 应当理解,术语“带电荷”或“带电荷部分”不是指分子上的“部分负电荷(partial negative charge)”或“部分正电荷”。术语“部分负电荷”和“部分正电荷”以其在本领域中的普通含义给出。“部分负电荷”可能产生于当官能团包含被极化的键从而使得电子密度被拉向该键的一个原子,从而在该原子上产生部分负电荷时。一般来讲,本领域普通技术人员将会辨识可以这种方式变得极化的键。

[0363] 在一些实施方案中,可电离脂质是可电离氨基脂质,其在本领域中有时称为“可电

离阳离子脂质”。在一个实施方案中,可电离氨基脂质可具有经由接头结构连接的带正电荷的亲水性头部和疏水性尾部。除这些之外,可电离脂质也可以是包含环状胺基的脂质。

[0364] 本公开的疫苗通常被配制成脂质纳米粒子。在一些实施方案中,脂质纳米粒子包含至少一种可电离氨基脂质、至少一种非阳离子脂质、至少一种甾醇和/或至少一种经聚乙二醇 (PEG) 修饰的脂质。

[0365] 在一些实施方案中,脂质纳米粒子包含所占摩尔比为20%-60%的可电离氨基脂质。例如,脂质纳米粒子可包含所占摩尔比为20%-50%、20%-40%、20%-30%、30%-60%、30%-50%、30%-40%、40%-60%、40%-50%或50%-60%的可电离氨基脂质。在一些实施方案中,脂质纳米粒子包含所占摩尔比为20%、30%、40%、50%或60%的可电离氨基脂质。

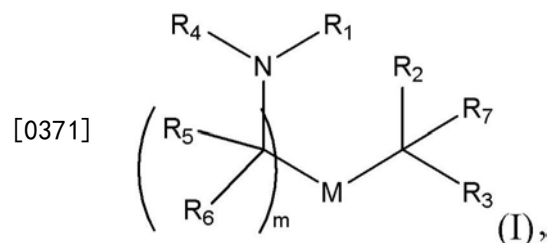
[0366] 在一些实施方案中,脂质纳米粒子包含所占摩尔比为5%-25%的非阳离子脂质。例如,脂质纳米粒子可包含所占摩尔比为5%-20%、5%-15%、5%-10%、10%-25%、10%-20%、10%-25%、15%-25%、15%-20%,或20%-25%的非阳离子脂质。在一些实施方案中,脂质纳米粒子包含所占摩尔比为5%、10%、15%、20%或25%的非阳离子脂质。

[0367] 在一些实施方案中,脂质纳米粒子包含所占摩尔比为25%-55%的甾醇。例如,脂质纳米粒子可包含所占摩尔比为25%-50%、25%-45%、25%-40%、25%-35%、25%-30%、30%-55%、30%-50%、30%-45%、30%-40%、30%-35%、35%-55%、35%-50%、35%-45%、35%-40%、40%-55%、40%-50%、40%-45%、45%-55%、45%-50%或50%-55%的甾醇。在一些实施方案中,脂质纳米粒子包含所占摩尔比为25%、30%、35%、40%、45%、50%或55%的甾醇。

[0368] 在一些实施方案中,脂质纳米粒子包含所占摩尔比为0.5%-15%的经PEG修饰的脂质。例如,脂质纳米粒子可包含所占摩尔比为0.5%-10%、0.5%-5%、1%-15%、1%-10%、1%-5%、2%-15%、2%-10%、2%-5%、5%-15%、5%-10%或10%-15%。在一些实施方案中,脂质纳米粒子包含所占摩尔比为0.5%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%或15%的经PEG修饰的脂质。

[0369] 在一些实施方案中,脂质纳米粒子包含所占摩尔比为20%-60%的可电离氨基脂质、所占摩尔比为5%-25%的非阳离子脂质、所占摩尔比为25%-55%的甾醇和所占摩尔比为0.5%-15%的经PEG修饰的脂质。

[0370] 在一些实施方案中,本公开的可电离氨基脂质包含式 (I) 的化合物:



[0372] 或其盐或异构体,其中:

[0373] R_1 选自自由以下组成的组: C_{5-30} 烷基、 C_{5-20} 烯基、 $-R^*YR''$ 、 $-YR''$ 和 $-R''M'R'$;

[0374] R_2 和 R_3 独立地选自自由以下组成的组: H 、 C_{1-14} 烷基、 C_{2-14} 烯基、 $-R^*YR''$ 、 $-YR''$ 和 $-R^*OR''$, 或者 R_2 和 R_3 与它们所连接的原子一起形成杂环或碳环;

[0375] R_4 选自自由以下组成的组: C_{3-6} 碳环、 $-(CH_2)_nQ$ 、 $-(CH_2)_nCHQR$ 、 $-CHQR$ 、 $-CQ(R)_2$ 和未取代

的 C_{1-6} 烷基,其中Q选自碳环、杂环、 $-OR$ 、 $-O(CH_2)_nN(R)_2$ 、 $-C(O)OR$ 、 $-OC(O)R$ 、 $-CX_3$ 、 $-CX_2H$ 、 $-CXH_2$ 、 $-CN$ 、 $-N(R)_2$ 、 $-C(O)N(R)_2$ 、 $-N(R)C(O)R$ 、 $-N(R)S(O)_2R$ 、 $-N(R)C(O)N(R)_2$ 、 $-N(R)C(S)N(R)_2$ 、 $-N(R)R_8$ 、 $-O(CH_2)_nOR$ 、 $-N(R)C(=NR_9)N(R)_2$ 、 $-N(R)C(=CHR_9)N(R)_2$ 、 $-OC(O)N(R)_2$ 、 $-N(R)C(O)OR$ 、 $-N(OR)C(O)R$ 、 $-N(OR)S(O)_2R$ 、 $-N(OR)C(O)OR$ 、 $-N(OR)C(O)N(R)_2$ 、 $-N(OR)C(S)N(R)_2$ 、 $-N(OR)C(=NR_9)N(R)_2$ 、 $-N(OR)C(=CHR_9)N(R)_2$ 、 $-C(=NR_9)N(R)_2$ 、 $-C(=NR_9)R$ 、 $-C(O)N(R)OR$ 和 $-C(R)N(R)_2C(O)OR$,并且每个n独立地选自1、2、3、4和5;

[0376] 每个 R_5 独立地选自由以下组成的组: C_{1-3} 烷基、 C_{2-3} 烯基和H;

[0377] 每个 R_6 独立地选自由以下组成的组: C_{1-3} 烷基、 C_{2-3} 烯基和H;

[0378] M和M'独立地选自 $-C(O)O-$ 、 $-OC(O)-$ 、 $-C(O)N(R')-$ 、 $-N(R')C(O)-$ 、 $-C(O)-$ 、 $-C(S)-$ 、 $-C(S)S-$ 、 $-SC(S)-$ 、 $-CH(OH)-$ 、 $-P(O)(OR')O-$ 、 $-S(O)_2-$ 、 $-S-S-$ 、芳基和杂芳基;

[0379] R_7 选自由以下组成的组: C_{1-3} 烷基、 C_{2-3} 烯基和H;

[0380] R_8 选自由以下组成的组: C_{3-6} 碳环和杂环;

[0381] R_9 选自由以下组成的组:H、CN、 NO_2 、 C_{1-6} 烷基、 $-OR$ 、 $-S(O)_2R$ 、 $-S(O)_2N(R)_2$ 、 C_{2-6} 烯基、 C_{3-6} 碳环和杂环;

[0382] $_2N(R)_2$ 、 C_{2-6} 烯基、 C_{3-6} 碳环和杂环;

[0383] 每个R独立地选自由以下组成的组: C_{1-3} 烷基、 C_{2-3} 烯基和H;

[0384] 每个R'独立地选自由以下组成的组: C_{1-18} 烷基、 C_{2-18} 烯基、 $-R^*YR^*$ 、 $-YR^*$ 和H;

[0385] 每个R''独立地选自由以下组成的组: C_{3-14} 烷基和

[0386] C_{3-14} 烯基;

[0387] 每个 R^* 独立地选自由以下组成的组: C_{1-12} 烷基和

[0388] C_{2-12} 烯基;

[0389] 每个Y独立地是 C_{3-6} 碳环;

[0390] 每个X独立地选自由以下组成的组:F、Cl、Br和I;并且

[0391] m选自5、6、7、8、9、10、11、12和13。

[0392] 在一些实施方案中,式(I)的化合物的子集包括这样的子集,其中:当 R_4 是 $-(CH_2)_nQ$ 、 $-(CH_2)_nCHQR$ 、 $-CHQR$ 或 $-CQ(R)_2$ 时,那么(i)当n是1、2、3、4或5时,Q不是 $-N(R)_2$;或者(ii)当n是1或2时,Q不是5、6或7元杂环烷基。

[0393] 在一些实施方案中,式(I)的化合物的另一个子集包括这样的子集,其中:

[0394] R_1 选自由以下组成的组: C_{5-30} 烷基、 C_{5-20} 烯基、 $-R^*YR^*$ 、 $-YR^*$ 和 $-R^*M'R^*$;

[0395] R_2 和 R_3 独立地选自由以下组成的组:H、 C_{1-14} 烷基、 C_{2-14} 烯基、 $-R^*YR^*$ 、 $-YR^*$ 和 $-R^*OR^*$,或者 R_2 和 R_3 与它们所连接的原子一起形成杂环或碳环;

[0396] R_4 选自由以下组成的组: C_{3-6} 碳环、 $-(CH_2)_nQ$ 、 $-(CH_2)_nCHQR$ 、 $-CHQR$ 、 $-CQ(R)_2$ 和未取代的 C_{1-6} 烷基,其中Q选自 C_{3-6} 碳环,具有一个或多个选自N、O和S的杂原子的5至14元杂芳基、 $-OR$ 、 $-O(CH_2)_nN(R)_2$ 、 $-C(O)OR$ 、 $-OC(O)R$ 、 $-CX_3$ 、 $-CX_2H$ 、 $-CXH_2$ 、 $-CN$ 、 $-C(O)N(R)_2$ 、 $-N(R)C(O)R$ 、 $-N(R)S(O)_2R$ 、 $-N(R)C(O)N(R)_2$ 、 $-N(R)C(S)N(R)_2$ 、 $-CRN(R)_2C(O)OR$ 、 $-N(R)R_8$ 、 $-O(CH_2)_nOR$ 、 $-N(R)C(=NR_9)N(R)_2$ 、 $-N(R)C(=CHR_9)N(R)_2$ 、 $-OC(O)N(R)_2$ 、 $-N(R)C(O)OR$ 、 $-N(OR)C(O)R$ 、 $-N(OR)S(O)_2R$ 、 $-N(OR)C(O)OR$ 、 $-N(OR)C(O)N(R)_2$ 、 $-N(OR)C(S)N(R)_2$ 、 $-N(OR)C(=NR_9)N(R)_2$ 、 $-N(OR)C(=CHR_9)N(R)_2$ 、 $-C(=NR_9)N(R)_2$ 、 $-C(=NR_9)R$ 、 $-C(O)N(R)OR$ 以及具有一个或多个选自N、O和S的杂原子的5至14元杂环烷基,该5至14元杂环烷基基团被一个或多个选自氧代($=O$)、OH、氨基、单烷基氨基或二烷基氨基,以及 C_{1-3} 烷基的取代基取代,并且每个n独立地选自1、2、3、

4和5;

[0397] 每个 R_5 独立地选自自由以下组成的组: C_{1-3} 烷基、 C_{2-3} 烯基和H;

[0398] 每个 R_6 独立地选自自由以下组成的组: C_{1-3} 烷基、 C_{2-3} 烯基和H;

[0399] M和M'独立地选自 $-C(O)O-$ 、 $-OC(O)-$ 、 $-C(O)N(R')$ 、 $-N(R')C(O)-$ 、 $-C(O)-$ 、 $-C(S)-$ 、 $-C(S)S-$ 、 $-SC(S)-$ 、 $-CH(OH)-$ 、 $-P(O)(OR')O-$ 、 $-S(O)_2-$ 、 $-S-S-$ 、芳基和杂芳基;

[0400] R_7 选自自由以下组成的组: C_{1-3} 烷基、 C_{2-3} 烯基和H;

[0401] R_8 选自自由以下组成的组: C_{3-6} 碳环和杂环;

[0402] R_9 选自自由以下组成的组:H、CN、 NO_2 、 C_{1-6} 烷基、 $-OR$ 、 $-S(O)_2R$ 、 $-S(O)$

$_2N(R)_2$ 、 C_{2-6} 烯基、 C_{3-6} 碳环和杂环;

[0404] 每个R独立地选自自由以下组成的组: C_{1-3} 烷基、 C_{2-3} 烯基和H;

[0405] 每个R'独立地选自自由以下组成的组: C_{1-18} 烷基、 C_{2-18} 烯基、 $-R^*YR''$ 、 $-YR''$ 和H;

[0406] 每个R''独立地选自自由以下组成的组: C_{3-14} 烷基和 C_{3-14} 烯基;

[0407] 每个R*独立地选自自由以下组成的组: C_{1-12} 烷基和 C_{2-12} 烯基;

[0408] 每个Y独立地是 C_{3-6} 碳环;

[0409] 每个X独立地选自自由以下组成的组:F、Cl、Br和I;并且

[0410] m选自5、6、7、8、9、10、11、12和13,

[0411] 或其盐或异构体。

[0412] 在一些实施方案中,式(I)的化合物的另一个子集包括这样的子集,其中:

[0413] R_1 选自自由以下组成的组: C_{5-30} 烷基、 C_{5-20} 烯基、 $-R^*YR''$ 、 $-YR''$ 和 $-R''M'R'$;

[0414] R_2 和 R_3 独立地选自自由以下组成的组:H、 C_{1-14} 烷基、 C_{2-14} 烯基、 $-R^*YR''$ 、 $-YR''$ 和 $-R^*OR''$,或者 R_2 和 R_3 与它们所连接的原子一起形成杂环或碳环;

[0415] R_4 选自自由以下组成的组: C_{3-6} 碳环、 $-(CH_2)_nQ$ 、 $-(CH_2)_nCHQR$ 、 $-CHQR$ 、 $-CQ(R)_2$ 和未取代的 C_{1-6} 烷基,其中Q选自 C_{3-6} 碳环,具有一个或多个选自N、O和S的杂原子的5至14元杂环, $-OR$ 、 $-O(CH_2)_nN(R)_2$ 、 $-C(O)OR$ 、 $-OC(O)R$ 、 $-CX_3$ 、 $-CX_2H$ 、 $-CXH_2$ 、 $-CN$ 、 $-C(O)N(R)_2$ 、 $-N(R)C(O)R$ 、 $-N(R)S(O)_2R$ 、 $-N(R)C(O)N(R)_2$ 、 $-N(R)C(S)N(R)_2$ 、 $-CRN(R)_2C(O)OR$ 、 $-N(R)R_8$ 、 $-O(CH_2)_nOR$ 、 $-N(R)C(=NR_9)N(R)_2$ 、 $-N(R)C(=CHR_9)N(R)_2$ 、 $-OC(O)N(R)_2$ 、 $-N(R)C(O)OR$ 、 $-N(OR)C(O)R$ 、 $-N(OR)S(O)_2R$ 、 $-N(OR)C(O)OR$ 、 $-N(OR)C(O)N(R)_2$ 、 $-N(OR)C(S)N(R)_2$ 、 $-N(OR)C(=NR_9)N(R)_2$ 、 $-N(OR)C(=CHR_9)N(R)_2$ 、 $-C(=NR_9)R$ 、 $-C(O)N(R)OR$ 和 $-C(=NR_9)N(R)_2$,并且每个n独立地选自1、2、3、4和5;并且当Q是5至14元杂环且(i) R_4 是 $-(CH_2)_nQ$,其中n是1或2,或者(ii) R_4 是 $-(CH_2)_nCHQR$,其中n是1,或者(iii) R_4 是 $-CHQR$ 和 $-CQ(R)_2$,那么Q是5至14元杂芳基或8至14元杂环烷基;

[0416] 每个 R_5 独立地选自自由以下组成的组: C_{1-3} 烷基、 C_{2-3} 烯基和H;

[0417] 每个 R_6 独立地选自自由以下组成的组: C_{1-3} 烷基、 C_{2-3} 烯基和H;

[0418] M和M'独立地选自

[0419] $-C(O)O-$ 、 $-OC(O)-$ 、 $-C(O)N(R')$ 、 $-N(R')C(O)-$ 、 $-C(O)-$ 、 $-C(S)-$ 、 $-C(S)S-$ 、 $-SC(S)-$ 、 $-CH(OH)-$ 、 $-P(O)(OR')O-$ 、 $-S(O)_2-$ 、 $-S-S-$ 、芳基和杂芳基;

[0420] R_7 选自自由以下组成的组: C_{1-3} 烷基、 C_{2-3} 烯基和H;

[0421] R_8 选自自由以下组成的组: C_{3-6} 碳环和杂环;

[0422] R_9 选自自由以下组成的组:H、CN、 NO_2 、 C_{1-6} 烷基、 $-OR$ 、 $-S(O)_2R$ 、 $-S(O)$

[0423] $_2N(R)_2$ 、 C_{2-6} 烯基、 C_{3-6} 碳环和杂环;

- [0424] 每个R独立地选自由以下组成的组： C_{1-3} 烷基、 C_{2-3} 烯基和H；
- [0425] 每个R'独立地选自由以下组成的组： C_{1-18} 烷基、 C_{2-18} 烯基、-R*YR''、-YR''和H；
- [0426] 每个R''独立地选自由以下组成的组： C_{3-14} 烷基和 C_{3-14} 烯基；
- [0427] 每个R*独立地选自由以下组成的组： C_{1-12} 烷基和 C_{2-12} 烯基；
- [0428] 每个Y独立地是 C_{3-6} 碳环；
- [0429] 每个X独立地选自由以下组成的组：F、Cl、Br和I；并且
- [0430] m选自5、6、7、8、9、10、11、12和13，
- [0431] 或其盐或异构体。
- [0432] 在一些实施方案中，式(I)的化合物的另一个子集包括这样的子集，其中：
- [0433] R_1 选自由以下组成的组： C_{5-30} 烷基、 C_{5-20} 烯基、-R*YR''、-YR''和-R''M'R'；
- [0434] R_2 和 R_3 独立地选自由以下组成的组：H、 C_{1-14} 烷基、 C_{2-14} 烯基、-R*YR''、-YR''和-R*OR''，或者 R_2 和 R_3 与它们所连接的原子一起形成杂环或碳环；
- [0435] R_4 选自由以下组成的组： C_{3-6} 碳环、 $-(CH_2)_nQ$ 、 $-(CH_2)_nCHQR$ 、 $-CHQR$ 、 $-CQ(R)_2$ 和未取代的 C_{1-6} 烷基，其中Q选自 C_{3-6} 碳环，具有一个或多个选自N、O和S的杂原子的5至14元杂芳基，-OR，-O $(CH_2)_nN(R)_2$ ，-C(O)OR，-OC(O)R，-CX₃，-CX₂H，-CXH₂，-CN，-C(O)N(R)₂，-N(R)C(O)R，-N(R)S(O)₂R，-N(R)C(O)N(R)₂，-N(R)C(S)N(R)₂，-CRN(R)₂C(O)OR，-N(R)R₈，-O $(CH_2)_nOR$ ，-N(R)C(=NR₉)N(R)₂，-N(R)C(=CHR₉)N(R)₂，-OC(O)N(R)₂，-N(R)C(O)OR，-N(OR)C(O)R，-N(OR)S(O)₂R，-N(OR)C(O)OR，-N(OR)C(O)N(R)₂，-N(OR)C(S)N(R)₂，-N(OR)C(=NR₉)N(R)₂，-N(OR)C(=CHR₉)N(R)₂，-C(=NR₉)R，-C(O)N(R)OR和-C(=NR₉)N(R)₂，并且每个n独立地选自1、2、3、4和5；
- [0436] 每个 R_5 独立地选自由以下组成的组： C_{1-3} 烷基、 C_{2-3} 烯基和H；
- [0437] 每个 R_6 独立地选自由以下组成的组： C_{1-3} 烷基、 C_{2-3} 烯基和H；
- [0438] M和M'独立地选自-C(O)O-、-OC(O)-、-C(O)N(R')-、-N(R')C(O)-、-C(O)-、-C(S)-、-C(S)S-、-SC(S)-、-CH(OH)-、-P(O)(OR')O-、-S(O)₂-、-S-S-、芳基和杂芳基；
- [0439] R_7 选自由以下组成的组： C_{1-3} 烷基、 C_{2-3} 烯基和H；
- [0440] R_8 选自由以下组成的组： C_{3-6} 碳环和杂环；
- [0441] R_9 选自由以下组成的组：H、CN、NO₂、 C_{1-6} 烷基、-OR、-S(O)₂R、-S(O)₂N(R)₂、 C_{2-6} 烯基、 C_{3-6} 碳环和杂环；
- [0442] 每个R独立地选自由以下组成的组： C_{1-3} 烷基、 C_{2-3} 烯基和H；
- [0443] 每个R'独立地选自由以下组成的组： C_{1-18} 烷基、 C_{2-18} 烯基、-R*YR''、-YR''和H；
- [0444] 每个R''独立地选自由以下组成的组： C_{3-14} 烷基和 C_{3-14} 烯基；
- [0445] 每个R*独立地选自由以下组成的组： C_{1-12} 烷基和 C_{2-12} 烯基；
- [0446] 每个Y独立地是 C_{3-6} 碳环；
- [0447] 每个X独立地选自由以下组成的组：F、Cl、Br和I；并且
- [0448] m选自5、6、7、8、9、10、11、12和13，
- [0449] 或其盐或异构体。
- [0450] 在一些实施方案中，式(I)的化合物的另一个子集包括这样的子集，其中：
- [0451] R_1 选自由以下组成的组： C_{5-30} 烷基、 C_{5-20} 烯基、-R*YR''、-YR''和-R''M'R'；
- [0452] R_2 和 R_3 独立地选自由以下组成的组：H、 C_{2-14} 烷基、 C_{2-14} 烯基、-R*YR''、-YR''和-R*

OR”，或者 R_2 和 R_3 与它们所连接的原子一起形成杂环或碳环；

[0454] R_4 是 $-(CH_2)_nQ$ 或 $-(CH_2)_nCHQR$ ，其中 Q 是 $-N(R)_2$ ，并且 n 选自3、4和5；

[0455] 每个 R_5 独立地选自由以下组成的组： C_{1-3} 烷基、 C_{2-3} 烯基和H；

[0456] 每个 R_6 独立地选自由以下组成的组： C_{1-3} 烷基、 C_{2-3} 烯基和H；

[0457] M 和 M' 独立地选自 $-C(O)O-$ 、 $-OC(O)-$ 、 $-C(O)N(R')$ 、 $-N(R')C(O)-$ 、 $-C(O)-$ 、 $-C(S)-$ 、 $-C(S)S-$ 、 $-SC(S)-$ 、 $-CH(OH)-$ 、 $-P(O)(OR')O-$ 、 $-S(O)_2-$ 、 $-S-S-$ 、芳基和杂芳基；

[0458] R_7 选自由以下组成的组： C_{1-3} 烷基、 C_{2-3} 烯基和H；

[0459] 每个 R 独立地选自由以下组成的组： C_{1-3} 烷基、 C_{2-3} 烯基和H；

[0460] 每个 R' 独立地选自由以下组成的组： C_{1-18} 烷基、 C_{2-18} 烯基、 $-R^*YR''$ 、 $-YR''$ 和H；

[0461] 每个 R'' 独立地选自由以下组成的组： C_{3-14} 烷基和 C_{3-14} 烯基；

[0462] 每个 R^* 独立地选自由以下组成的组： C_{1-12} 烷基和 C_{1-12} 烯基；

[0463] 每个 Y 独立地是 C_{3-6} 碳环；

[0464] 每个 X 独立地选自由以下组成的组：F、Cl、Br和I；并且

[0465] m 选自5、6、7、8、9、10、11、12和13，

[0466] 或其盐或异构体。

[0467] 在一些实施方案中，式(I)的化合物的另一个子集包括这样的子集，其中：

[0468] R_1 选自由以下组成的组： C_{5-30} 烷基、 C_{5-20} 烯基、 $-R^*YR''$ 、 $-YR''$ 和 $-R''M'R'$ ；

[0469] R_2 和 R_3 独立地选自由以下组成的组： C_{1-14} 烷基、 C_{2-14} 烯基、 $-R^*YR''$ 、 $-YR''$ 和 $-R^*OR''$ ，或者 R_2 和 R_3 与它们所连接的原子一起形成杂环或碳环；

[0470] R_4 选自由以下组成的组： $-(CH_2)_nQ$ 、 $-(CH_2)_nCHQR$ 、 $-CHQR$ 和 $-CQ(R)_2$ ，其中 Q 是 $-N(R)_2$ ，并且 n 选自1、2、3、4和5；

[0471] 每个 R_5 独立地选自由以下组成的组： C_{1-3} 烷基、 C_{2-3} 烯基和H；

[0472] 每个 R_6 独立地选自由以下组成的组： C_{1-3} 烷基、 C_{2-3} 烯基和H；

[0473] M 和 M' 独立地选自 $-C(O)O-$ 、 $-OC(O)-$ 、 $-C(O)N(R')$ 、 $-N(R')C(O)-$ 、 $-C(O)-$ 、 $-C(S)-$ 、 $-C(S)S-$ 、 $-SC(S)-$ 、 $-CH(OH)-$ 、 $-P(O)(OR')O-$ 、 $-S(O)_2-$ 、 $-S-S-$ 、芳基和杂芳基；

[0474] R_7 选自由以下组成的组： C_{1-3} 烷基、 C_{2-3} 烯基和H；

[0475] 每个 R 独立地选自由以下组成的组： C_{1-3} 烷基、 C_{2-3} 烯基和H；

[0476] 每个 R' 独立地选自由以下组成的组： C_{1-18} 烷基、 C_{2-18} 烯基、 $-R^*YR''$ 、 $-YR''$ 和H；

[0477] 每个 R'' 独立地选自由以下组成的组： C_{3-14} 烷基和 C_{3-14} 烯基；

[0478] 每个 R^* 独立地选自由以下组成的组： C_{1-12} 烷基和 C_{1-12} 烯基；

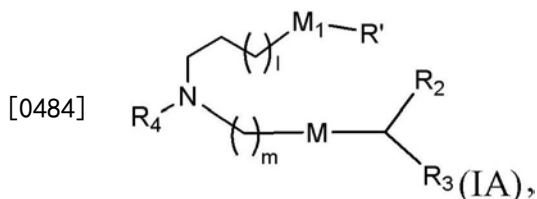
[0479] 每个 Y 独立地是 C_{3-6} 碳环；

[0480] 每个 X 独立地选自由以下组成的组：F、Cl、Br和I；并且

[0481] m 选自5、6、7、8、9、10、11、12和13，

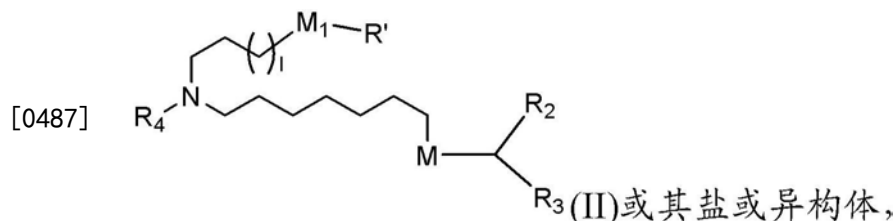
[0482] 或其盐或异构体。

[0483] 在一些实施方案中，式(I)的化合物的子集包括式(IA)的子集：



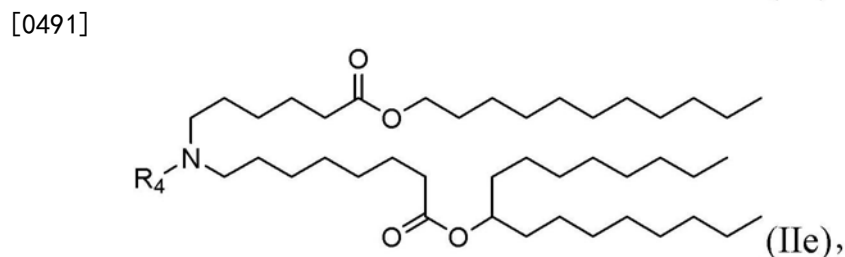
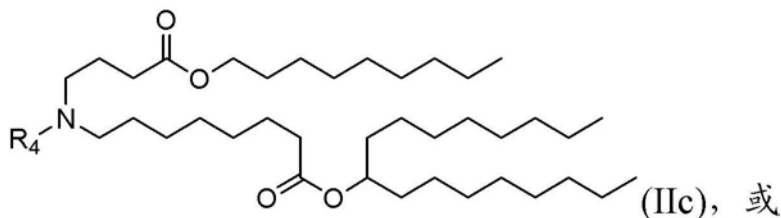
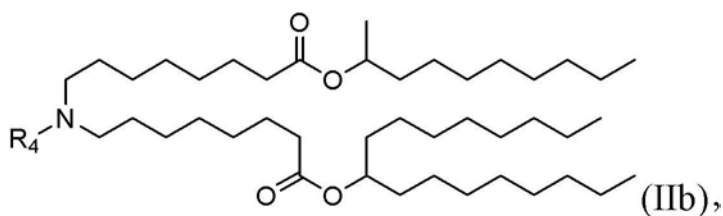
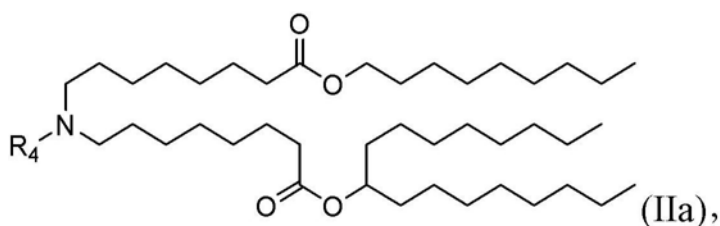
[0485] 或其盐或异构体,其中1选自1、2、3、4和5;m选自5、6、7、8和9; M_1 是键或 M' ; R_4 是未取代的 C_{1-3} 烷基或 $-(CH_2)_n Q$,其中Q是OH、 $-NHC(S)N(R)_2$ 、 $-NHC(O)N(R)_2$ 、 $-N(R)C(O)R$ 、 $-N(R)S(O)_2R$ 、 $-N(R)R_8$ 、 $-NHC(=NR_9)N(R)_2$ 、 $-NHC(=CHR_9)N(R)_2$ 、 $-OC(O)N(R)_2$ 、 $-N(R)C(O)OR$ 、杂芳基或杂环烷基;M和 M' 独立地选自 $-C(O)O-$ 、 $-OC(O)-$ 、 $-C(O)N(R')$ 、 $-P(O)(OR')O-$ 、 $-S-S-$ 、芳基和杂芳基;并且 R_2 和 R_3 独立地选自由以下组成的组:H、 C_{1-14} 烷基和 C_{2-14} 烯基。

[0486] 在一些实施方案中,式(I)的化合物的子集包括式(II)的子集:



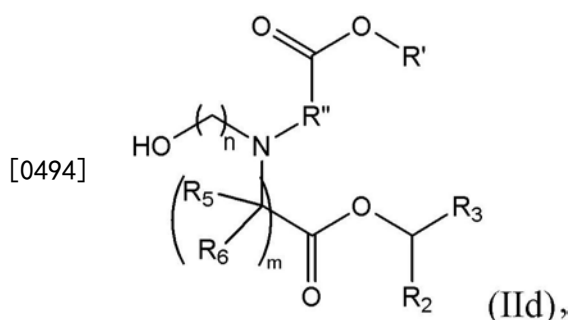
[0488] 其中1选自1、2、3、4和5; M_1 是键或 M' ; R_4 是未取代的 C_{1-3} 烷基或 $(CH_2)_n Q$,其中n是2、3或4,并且Q是OH、 $-NHC(S)N(R)_2$ 、 $-NHC(O)N(R)_2$ 、 $-N(R)C(O)R$ 、 $-N(R)S(O)_2R$ 、 $-N(R)R_8$ 、 $-NHC(=NR_9)N(R)_2$ 、 $-NHC(=CHR_9)N(R)_2$ 、 $-OC(O)N(R)_2$ 、 $-N(R)C(O)OR$ 、杂芳基或杂环烷基;M和 M' 独立地选自 $-C(O)O-$ 、 $-OC(O)-$ 、 $-C(O)N(R')$ 、 $-P(O)(OR')O-$ 、 $-S-S-$ 、芳基和杂芳基;并且 R_2 和 R_3 独立地选自由以下组成的组:H、 C_{1-14} 烷基和 C_{2-14} 烯基。

[0489] 在一些实施方案中,式(I)的化合物的子集包括式(IIa)、(IIb)、(IIc)或(IIe)的子集:



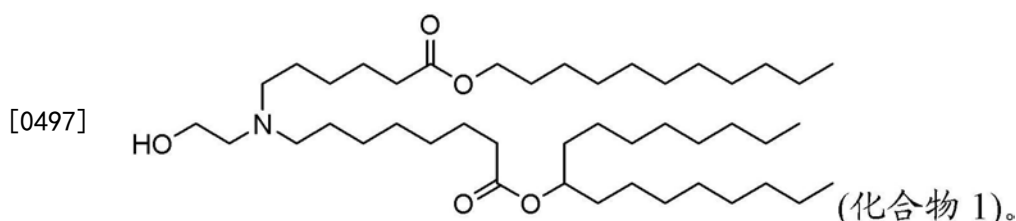
[0492] 或其盐或异构体,其中 R_4 如本文所述。

[0493] 在一些实施方案中,式(I)的化合物的子集包括式(IIId)的子集:



[0495] 或其盐或异构体,其中 n 是2、3或4;并且 m 、 R' 、 R'' 以及 R_2 至 R_6 如本文所述。例如, R_2 和 R_3 中的每一者均可独立地选自由以下组成的组: C_{5-14} 烷基和 C_{5-14} 烯基。

[0496] 在一些实施方案中,本公开的可电离阳离子脂质包含具有下面的结构的化合物:



[0498] 在一些实施方案中,本公开的非阳离子脂质包含1,2-二硬脂酰基-sn-甘油-3-磷酸胆碱(DSPC)、1,2-二油酰基-sn-甘油-3-磷酸乙醇胺(DOPE)、1,2-二亚油酰基-sn-甘油-3-磷酸胆碱(DLPC)、1,2-二肉豆蔻酰基-sn-甘油-3-磷酸胆碱(DMPC)、1,2-二油酰基-sn-甘油-3-磷酸胆碱(DOPC)、1,2-二棕榈酰基-sn-甘油-3-磷酸胆碱(DPPC)、1,2-双十一酰基-sn-甘油-3-磷酸胆碱(DUPC)、1-棕榈酰基-2-油酰基-sn-甘油-3-磷酸胆碱(POPC)、1,2-二-0-十八烯基-sn-甘油-3-磷酸胆碱(18:0Diether PC)、1-油酰基-2-胆甾醇基半琥珀酰基-sn-甘油-3-磷酸胆碱(0ChemsPC)、1-十六烷基-sn-甘油-3-磷酸胆碱(C16 Lyso PC)、1,2-二亚麻酰基-sn-甘油-3-磷酸胆碱、1,2-二花生四烯酰基-sn-甘油-3-磷酸胆碱、1,2-双二十二碳六烯酰基-sn-甘油-3-磷酸胆碱、1,2-二植烷酰基-sn-甘油-3-磷酸乙醇胺(ME 16.0PE)、1,2-二硬脂酰基-sn-甘油-3-磷酸乙醇胺、1,2-二亚油酰基-sn-甘油-3-磷酸乙醇胺、1,2-二亚麻酰基-sn-甘油-3-磷酸乙醇胺、1,2-二花生四烯酰基-sn-甘油-3-磷酸乙醇胺、1,2-双二十二碳六烯酰基-sn-甘油-3-磷酸乙醇胺、1,2-二油酰基-sn-甘油-3-磷酸-rac-(1-甘油)钠盐(DOPG)、鞘磷脂以及它们的混合物。

[0499] 在一些实施方案中,本公开的经PEG修饰的脂质包含经PEG修饰的磷脂酰乙醇胺、经PEG修饰的磷脂酸、经PEG修饰的神经酰胺、经PEG修饰的二烷基胺、经PEG修饰的二酰基甘油、经PEG修饰的二烷基甘油以及它们的混合物。在一些实施方案中,经PEG修饰的脂质是PEG-DMG、PEG-c-DOMG(也称为PEG-DOMG)、PEG-DSG和/或PEG-DPG。

[0500] 在一些实施方案中,本公开的甾醇包括胆固醇、粪甾醇(fecosterol)、谷甾醇、麦角甾醇、菜油甾醇、豆甾醇、菜子甾醇、番茄碱、熊果酸、 α -生育酚以及它们的混合物。

[0501] 在一些实施方案中,本公开的LNP包含化合物1的可电离氨基脂质,其中非阳离子脂质是DSPC,结构脂质是胆固醇,并且PEG脂质是PEG-DMG。

[0502] 在一些实施方案中,本公开的LNP包含约2:1至约30:1的N:P比。

[0503] 在一些实施方案中,本公开的LNP包含约6:1的N:P比。

[0504] 在一些实施方案中,本公开的LNP包含约3:1的N:P比。

[0505] 在一些实施方案中,本公开的LNP包含约10:1至约100:1的可电离阳离子脂质组分与RNA的wt/wt比。

[0506] 在一些实施方案中,本公开的LNP包含约20:1的可电离阳离子脂质组分与RNA的wt/wt比。

[0507] 在一些实施方案中,本公开的LNP包含约10:1的可电离阳离子脂质组分与RNA的wt/wt比。

[0508] 在一些实施方案中,本公开的LNP具有约50nm至约150nm的平均直径。

[0509] 在一些实施方案中,本公开的LNP具有约70nm至约120nm的平均直径。

[0510] 在一个实施方案中,所述脂质可以是可切割脂质,诸如国际公布号W02012170889中所述的那些,该国际公布出于此目的以引用方式整体并入本文。在一个实施方案中,所述脂质可通过本领域中已知的方法合成以及/或者如国际公布号W02013086354中所述的那样合成;该国际公布的内容出于此目的以引用方式整体并入本文。

[0511] 纳米粒子组合物可通过多种方法来表征。例如,显微术(例如,透射电子显微术或扫描电子显微术)可用于检查纳米粒子组合物的形态和尺寸分布。动态光散射或电位测定法(例如,电位测定滴定)可用于测量 ζ 电位。动态光散射还可用于测定粒度。还可使用诸如Zetasizer Nano ZS (Malvern Instruments Ltd, Malvern, Worcestershire, UK)的仪器来测量纳米粒子组合物的多种特征,诸如粒度、多分散性指数和 ζ 电位。

[0512] 纳米粒子组合物可通过多种方法来表征。例如,显微术(例如,透射电子显微术或扫描电子显微术)可用于检查纳米粒子组合物的形态和尺寸分布。动态光散射或电位测定法(例如,电位测定滴定)可用于测量 ζ 电位。动态光散射还可用于测定粒度。还可使用诸如Zetasizer Nano ZS (Malvern Instruments Ltd, Malvern, Worcestershire, UK)的仪器来测量纳米粒子组合物的多种特征,诸如粒度、多分散性指数和 ζ 电位。

[0513] 纳米粒子的尺寸可帮助抵抗生物反应,诸如但不限于炎症,或者可增大多核苷酸的生物学效应。如本文所用,在纳米粒子组合物的背景下的“尺寸”或“平均尺寸”是指纳米粒子组合物的平均直径。

[0514] 试剂盒

[0515] 在本公开的其他方面,还提供了用于实现这些方法的试剂盒。试剂盒包括容纳制剂的容器、容纳疫苗制剂的容器以及关于以下操作的说明书:将癌症疫苗添加至疫苗制剂中以产生癌症疫苗制剂,在向受试者施用的24小时内对所述癌症疫苗制剂进行混合。在一些实施方案中,试剂盒包括具有编码3-200(例如,3-130)种癌症抗原的开放阅读框的mRNA。

[0516] 制品包括在一个或多个容器中的医药级或诊断级本公开化合物。制品可包括宣传或描述本公开化合物的用途的说明书或标签。

[0517] 如本文所用,“宣传”包括所有营商方法,包括教育、医院和其他临床教导、医药工业活动(包括医药销售),以及任何做广告或其他促销活动(包括与本公开组合物关联的与癌症治疗相关的任何形式的书面、口头和电子沟通)的方法。

[0518] “说明书”可详细说明宣传组分,并且通常涉及在本公开组合物的包装上或者与本公开组合物的包装相伴的书面说明书。说明书还可包括以任何方式提供的任何口头或电子

说明书。

[0519] 因此,在一些实施方案中,可将本文所述的剂组装到医药或诊断或研究试剂盒中以有助于它们用于治疗、诊断或研究应用中。试剂盒可包括一个或多个容纳本公开的组分的容器和使用说明书。具体来说,此类试剂盒可包括一种或多种本文所述的剂,以及描述这些剂的预期治疗应用和适当施用的说明书。在某些实施方案中,试剂盒中的剂可呈适合于所述剂的特定应用和施用方法的药物制剂和剂量的形式。

[0520] 试剂盒可被设计成使医师容易使用本文所述的方法,并且可采用许多形式。试剂盒的组合物中的每一种在适用情况下可以液体形式(例如,以溶液形式)或以固体形式(例如,干燥粉末)提供。在某些情况下,所述组合物中的一些可以是可例如通过添加可与或不与试剂盒一起提供的适合溶剂或其他物质(例如,水或细胞培养基)来构配(constitutable)或以其他方式加工(例如,制成活性形式)的。如本文所用,“说明书”可详细说明教导和/或宣传的组分,并且通常涉及在本公开的包装上或者与本公开的包装相伴的书面说明书。说明书还可包括以任何方式提供以使得使用者将明确认识到说明书将与试剂盒相伴的任何口头或电子说明书,例如视听材料(例如,录像带、DVD等)、因特网和/或基于网络的通信等。书面说明书可呈由监管药物或生物产品的制造、使用或销售的政府机构指定的形式,所述说明书也可反映出所述机构批准了制造、使用或销售用于人施用。

[0521] 在某些方面,本公开涉及用于通过IVT方法来制备核酸癌症疫苗(例如,RNA癌症疫苗)的试剂盒。在个性化癌症疫苗中,重要的是鉴定出特定于患者的突变并为患者接种一种或多种新表位。在此类疫苗中,由此种核酸的ORF编码的一种或多种抗原将是特定于患者的。编码一种或多种抗原的核酸(例如,RNA)的5'端和3'端可能适用于更广的范围,因为它们包括许多核酸(例如,RNA)所共同具有的非翻译区域和稳定区域。除别的之外,本公开提供了试剂盒,所述试剂盒包括嵌合核酸中的一种或部分,诸如RNA的一个或多个5'区域和/或3'区域,其可与编码特定于患者的表位的ORF组合。例如,试剂盒可包括含有5'-ORF、3'-ORF和聚腺苷酸尾部中的一者或多者的核酸。在一些实施方案中,每种核酸组分处于单独的容器中。在其他实施方案中,超过一种核酸组分一起存在于单个容器中。在一些实施方案中,试剂盒包括结扎酶。在一些实施方案中,所提供的试剂盒包括使用说明书。在一些实施方案中,说明书包括关于将编码肽表位的ORF结扎至试剂盒中的一种或多种其他组分(例如,5'-ORF、3'-ORF和/或聚腺苷酸尾部)上的说明。

[0522] 试剂盒可含有在一个或多个容器中的任何一种或多种本文所述的组分。作为实例,在一个实施方案中,试剂盒可包括关于混合试剂盒的一种或多种组分以及/或者分离和混合样品以及向受试者进行施加的说明书。试剂盒可包括容纳本文所述的剂的容器。所述剂可被无菌制备,包装在注射器中,并且冷冻运送。替代地,可将它容纳在小瓶或其他容器中以进行储存。第二容器可具有无菌制备的其他剂。替代地,试剂盒可包括预先混合并在注射器、小瓶、管或其他容器加以运送的活性剂。

[0523] 试剂盒可具有多种形式,诸如泡罩小袋、收缩包装小袋、真空可密封小袋、可密封热成形托盘,或其中附件松散地包装在小袋内的类似小袋或托盘形式、一个或多个管、容器、盒子或袋子。试剂盒可在添加附件之后进行灭菌,从而允许容器中的各个附件以其他方式拆开。试剂盒可使用任何适当灭菌技术进行灭菌,所述技术诸如辐射灭菌、加热灭菌或本领域中已知的其他灭菌方法。取决于特定应用,试剂盒还可包括其他组分,例如容器、细胞

培养基、盐、缓冲剂、试剂、注射器、针、用于施加或移除消毒剂的织物诸如纱布、一次性手套、在施用之前用于剂的支撑物等。

[0524] 试剂盒的组合物可以任何适合形式,例如以液体溶液形式或者以干燥粉末形式提供。当提供的组合物是干燥粉末时,所述粉末可通过添加适合的溶剂(其也可被提供)来复原。在使用组合物的液体形式的实施方案中,液体形式可以是浓缩的或即用型的。溶剂将取决于化合物以及使用或施用方式。用于药物组合物的适合溶剂是众所周知的,并且可在文献中获得。溶剂将取决于化合物以及使用或施用方式。

[0525] 在一组实施方案中,试剂盒可包括被分隔开来以紧密限制方式容纳一个或多个容器部件(诸如小瓶、管等)的托架部件,所述容器部件中的每一个均包含待用于所述方法中的单独元件中的一个。例如,所述容器中的一个可包含用于测定的阳性对照。另外,试剂盒可包括用于其他组分,例如可用于测定中的缓冲剂的容器。

[0526] 本公开还涵盖一种成品包装且贴有标签的药物产品。这种制造物品包括在适当的器皿或容器(诸如玻璃小瓶或其他气密封容器)中的适当的单位剂型。在适于肠胃外施用的剂型的情况下,活性成分是无菌的,并且适于以无粒子溶液形式施用。换句话说,本公开涵盖肠胃外溶液与冻干粉末两者,它们各自是无菌的,并且后者适于在注射之前进行复原。替代地,单位剂型可以是适于口服、经真皮、局部或经粘膜递送的固体。

[0527] 在优选实施方案中,单位剂型适于静脉内、肌内或皮下递送。因此,本公开涵盖适合于每个递送途径的优选无菌的溶液。

[0528] 在另一个优选实施方案中,将本公开的组合物与生物相容性清洁剂(包括但不限于卵磷脂、牛磺胆酸和胆固醇)一起;或者与其他蛋白质(包括但不限于 γ 球蛋白和血清白蛋白)一起储存在容器中。更优选地,将本公开的组合物与人血清白蛋白一起储存以供人使用,以及与牛血清白蛋白一起储存以供兽医学使用。

[0529] 正如任何药物产品一样,包装材料和容器被设计来在储存和装运期间保护产品的稳定性。进一步地,本公开的产品包括建议医师、技术人员或患者如何适当地预防或治疗所论述的疾病或病症的使用说明书或其他信息材料。换句话说,制造物品包括指示或建议包括但不限于实际剂量、监测程序(诸如用于监测平均绝对淋巴细胞计数、肿瘤细胞计数和肿瘤尺寸的方法)和其他监测信息的给药方案的说明书部件。

[0530] 更具体来说,本公开提供了一种制造物品,所述制造物品包括包装材料诸如盒子、瓶、管、小瓶、容器、喷雾器、吹入器、静脉内(i.v.)袋、封袋等;以及在所述包装材料内含有的药物制剂的至少一种单位剂型。本公开还提供了一种制造物品,所述制造物品包括包装材料诸如盒子、瓶、管、小瓶、容器、喷雾器、吹入器、静脉内(i.v.)袋、封袋等;以及在所述包装材料内含有的每种药物制剂的至少一种单位剂型。本公开进一步提供了一种制造物品,所述制造物品包括包装材料诸如盒子、瓶、管、小瓶、容器、喷雾器、吹入器、静脉内(i.v.)袋、封袋等;以及在所述包装材料内含有的每种药物制剂的至少一种单位剂型。本公开进一步提供了一种制造物品,所述制造物品包括用于注射制剂的优选以无菌形式包装的针或注射器,和/或包装酒精片。

[0531] 疫苗组合物中的活性成分(例如,核酸癌症疫苗)、药学上可接受的赋形剂和/或任何另外的成分的相对量可根据所治疗受试者的身份、身材和/或状况而变化,并且进一步根据组合物的施用途径而变化。例如,所述组合物可包含介于0.1%与99% (w/w) 之间的活性

成分。举例来说,所述组合物可包含介于0.1%与100%之间,例如介于.5%与50%之间、介于1%-30%之间、介于5%-80%之间、至少80% (w/w) 的活性成分。

[0532] 在一些实施方案中,含有药物产品的包装含有0.1mg至1mg的核酸(例如,mRNA)。在一些实施方案中,含有药物产品的包装含有0.35mg的核酸(例如,mRNA)。在一些实施方案中,所述核酸(例如,mRNA)的浓度为1mg/mL。

[0533] 在一些实施方案中,所述核酸(例如,mRNA)疫苗组合物可以以足以递送0.0001mg/kg至100mg/kg、0.001mg/kg至0.05mg/kg、0.005mg/kg至0.05mg/kg、0.001mg/kg至0.005mg/kg、0.05mg/kg至0.5mg/kg、0.01mg/kg至50mg/kg、0.1mg/kg至40mg/kg、0.5mg/kg至30mg/kg、0.01mg/kg至10mg/kg、0.1mg/kg至10mg/kg,或1mg/kg至25mg/kg受试者体重的剂量水平每天、一天一次或多次、每周、每个月等施用以获得期望的治疗、诊断、防治或成像作用(参见例如国际公布号W02013078199中所述的单位剂量的范围,所述国际公布以引用方式整体并入本文)。在一些实施方案中,所述核酸(例如,mRNA)疫苗以足以递送0.0100mg、0.025mg、0.050mg、0.075mg、0.100mg、0.125mg、0.150mg、0.175mg、0.200mg、0.225mg、0.250mg、0.275mg、0.300mg、0.325mg、0.350mg、0.375mg、0.400mg、0.425mg、0.450mg、0.475mg、0.500mg、0.525mg、0.550mg、0.575mg、0.600mg、0.625mg、0.650mg、0.675mg、0.700mg、0.725mg、0.750mg、0.775mg、0.800mg、0.825mg、0.850mg、0.875mg、0.900mg、0.925mg、0.950mg、0.975mg或1.0mg的剂量水平施用。在一些实施方案中,所述核酸(例如,mRNA)疫苗以足以向受试者递送介于10 μ g与400 μ g之间的mRNA疫苗的剂量水平施用。在一些实施方案中,所述核酸(例如,mRNA)疫苗以足以向受试者递送0.033mg、0.1mg、0.2mg或0.4mg的剂量水平施用。

[0534] 期望的剂量可一天三次、一天两次、一天一次、每隔一天、每三天、每周、每两周、每三周、每四周、每2个月、每3个月、每6个月等加以递送。在某些实施方案中,可使用多次施用(例如,两次、三次、四次、五次、六次、七次、八次、九次、十次、十一次、十二次、十三次、十四次或更多次施用)来递送期望的剂量。当采用多次施用时,可使用分次给药方案,诸如本文所述的那些。在一些实施方案中,所述核酸(例如,mRNA)疫苗组合物可以以足以递送0.0005mg/kg至0.01mg/kg,例如约0.0005mg/kg至约0.0075mg/kg,例如约0.0005mg/kg、约0.001mg/kg、约0.002mg/kg、约0.003mg/kg、约0.004mg/kg或约0.005mg/kg的剂量水平施用。在一些实施方案中,所述核酸(例如,mRNA)疫苗组合物可以以足以递送0.025mg/kg至0.250mg/kg、0.025mg/kg至0.500mg/kg、0.025mg/kg至0.750mg/kg,或0.025mg/kg至1.0mg/kg的剂量水平施用一次或两次(或更多次)。

[0535] 在一些实施方案中,所述核酸(例如,mRNA)疫苗组合物可以以下总剂量或者以足以递送以下总剂量的剂量水平施用两次(例如,第0天和第7天、第0天和第14天、第0天和第21天、第0天和第28天、第0天和第60天、第0天和第90天、第0天和第120天、第0天和第150天、第0天和第180天、第0天和3个月后、第0天和6个月后、第0天和9个月后、第0天和12个月后、第0天和18个月后、第0天和2年后、第0天和5年后,或第0天和10年后):0.0100mg、0.025mg、0.050mg、0.075mg、0.100mg、0.125mg、0.150mg、0.175mg、0.200mg、0.225mg、0.250mg、0.275mg、0.300mg、0.325mg、0.350mg、0.375mg、0.400mg、0.425mg、0.450mg、0.475mg、0.500mg、0.525mg、0.550mg、0.575mg、0.600mg、0.625mg、0.650mg、0.675mg、0.700mg、0.725mg、0.750mg、0.775mg、0.800mg、0.825mg、0.850mg、0.875mg、0.900mg、0.925mg、

0.950mg、0.975mg或1.0mg。本公开涵盖更高和更低的施用剂量和频率。例如，可施用核酸（例如，mRNA）疫苗组合物三次或四次或更多次。在一些实施方案中，mRNA疫苗组合物每三周一天一次地施用。

[0536] 在一些实施方案中，核酸（例如，mRNA）疫苗组合物可以以下总剂量或者以足以递送以下总剂量的剂量水平施用两次（例如，第0天和第7天、第0天和第14天、第0天和第21天、第0天和第28天、第0天和第60天、第0天和第90天、第0天和第120天、第0天和第150天、第0天和第180天、第0天和3个月后、第0天和6个月后、第0天和9个月后、第0天和12个月后、第0天和18个月后、第0天和2年后、第0天和5年后，或第0天和10年后）：0.010mg、0.025mg、0.100mg或0.400mg。

[0537] 在一些实施方案中，用于对受试者接种疫苗的方法中的核酸（例如，mRNA）疫苗以用于对所述受试者接种疫苗的有效量，以在10 μ g/kg与400 μ g/kg之间的单次剂量的核酸疫苗向所述受试者施用。在一些实施方案中，用于对受试者接种疫苗的方法中的RNA疫苗以用于对所述受试者接种疫苗的有效量，以在10 μ g与400 μ g之间的单次剂量的核酸疫苗向所述受试者施用。

[0538] 本文所述的方法和组合物在其应用上不限于以下描述中阐述的构建细节和组分排列。本文所述的方法和组合物能够具有其他实施方案，并且能够以各种方式实施或进行。同样，本文所使用的措辞和术语是用于描述的目的，而不应当被视为是限制性的。本文中“包括”、“包含”或“具有”、“含有”、“涉及”及其变化形式的使用意图涵盖此后所列的条目及其等效物以及另外的条目。

[0539] 实施例

[0540] 实施例1.多核苷酸的制造

[0541] 根据本公开，对核酸和/或其部分或区域的制造可利用本领域中教导的方法来实现，所述方法包括题为“Manufacturing Methods for Production of RNA Transcripts”的国际申请W02014/152027中详述的那些方法，该国际申请的内容出于此目的以引用方式整体并入本文。

[0542] 纯化方法可包括国际申请号W02014/152030和W02014/152031中教导的那些，该国际申请中的每一篇均出于此目的以引用方式整体并入本文。

[0543] 用于与所述核酸一起使用的检测和表征方法可使用本领域中已知的任何方法来执行，所述方法包括W02014/144039中教导的那些方法，该文献出于此目的以引用方式整体并入本文。

[0544] 对本公开的多核苷酸的表征可使用例如选自由多核苷酸作图、逆转录酶测序、电荷分布分析和RNA杂质检测组成的组的程序来实现，其中表征包括测定RNA转录物序列、测定RNA转录物的纯度，或测定RNA转录物的电荷异质性。此类方法教导于例如W02014/144711和W02014/144767中，所述文献中每一篇的内容出于此目的以引用方式整体并入本文。

[0545] 实施例2嵌合多核苷酸合成

[0546] 引言

[0547] 根据本公开，嵌合核酸的两个区域或部分可使用三磷酸酯化学来接合或结扎。

[0548] 根据此方法，具有100个或更少核苷酸的第一区域或部分被化学合成为具有5'-单磷酸酯和末端3'脱OH或封闭的OH。如果该区域长于80个核苷酸，那么它可被合成为用于结

扎的两条链。

[0549] 如果使用体外转录 (IVT) 将第一区域或部分合成为非位置修饰的区域或部分,那么可以随后执行向5' 单磷酸酯的转化并随后对3' 末端加帽。

[0550] 单磷酸酯保护基可选自本领域中已知的任何保护基。

[0551] 嵌合多核苷酸的第二区域或部分可使用化学合成或IVT方法来合成。IVT方法可包括可利用具有经修饰的帽的引物的RNA聚合酶。替代地,可化学合成多达130个核苷酸的帽并且使其偶联至IVT区域或部分。

[0552] 整个嵌合多核苷酸不需要用磷酸酯-糖主链制造。如果所述区域或部分之一编码多肽,那么优选此种区域或部分包含磷酸酯-糖主链。

[0553] 然后使用任何已知的点击化学、正点击化学、solulink或本领域技术人员已知的其他生物缀合物化学执行结扎。

[0554] 合成路线

[0555] 使用一系列起始区段制备嵌合核酸。此类区段包括:

[0556] (a) 经加帽和经保护的5' 区段,其包含正常3' OH (SEG.1)

[0557] (b) 5' 三磷酸酯区段,其可包括多肽的编码区域以及包含正常3' OH (SEG.2)

[0558] (c) 嵌合多核苷酸的3' 端 (例如,尾部) 的5' 单磷酸酯区段,其包含虫草素或者不包含3' OH (SEG.3)

[0559] 在合成 (化学或IVT合成) 之后,将区段3 (SEG.3) 用虫草素处理,接着用焦磷酸酶处理从而产生5' 单磷酸酯。

[0560] 然后使用RNA结扎酶将区段2 (SEG.2) 结扎至SEG.3上。然后纯化经结扎的多核苷酸并且将其用焦磷酸酶处理以切割二磷酸酯。然后纯化经处理的SEG.2-SEG.3构建体并且将SEG.1结扎至5' 末端上。可执行嵌合多核苷酸的进一步纯化步骤。

[0561] 每个步骤的收率可高达90%-95%。

[0562] 实施例3:用于cDNA产生的PCR

[0563] 使用Kapa Biosystems (Woburn,MA) 的2x KAPA HIFI™ HotStart ReadyMix来执行用于制备cDNA的PCR程序。此体系包括12.5μl 2x KAPA ReadyMix;0.75μl正向引物 (10μM);0.75μl反向引物 (10μM);100ng模板cDNA;和dH₂O,稀释至25.0μl。反应条件是在95℃下持续5分钟,以及25个循环的98℃持续20秒,接着58℃持续15秒,接着72℃持续45秒;接着72℃持续5分钟,接着4℃直至终止。

[0564] 使用Invitrogen的PURELINK™PCR Micro试剂盒 (Carlsbad,CA) 根据制造商的说明书清洗反应物 (多达5μg)。更大的反应物将需要使用具有更大容量的产品来清洗。清洗后,使用NANODROP™对cDNA定量,并通过琼脂糖凝胶电泳对cDNA进行分析,以确认cDNA为预期大小。接着提交cDNA以进行测序分析,随后继续进行体外转录反应。

[0565] 实施例4. 体外转录 (IVT)

[0566] 体外转录反应产生含有经均匀修饰的核酸的核酸。此类经均匀修饰的核酸可包含本公开的核酸的区域或部分。使用天然NTP和非天然NTP内部制备输入三磷酸核苷酸 (NTP) 混合物。

[0567] 典型的体外转录反应包括以下:

[0568] 1. 模板cDNA 1.0μg

[0569] 2.10x转录缓冲液(400mM Tris-HCl pH8.0,190mM MgCl₂,50mM DTT,10mM亚精胺)
2.0μl

[0570] 3.自定义NTPs(各25mM) 7.2μl

[0571] 4.RNA酶抑制剂 20U

[0572] 5.T7 RNA聚合酶 3000U

[0573] 6.dH₂O ,直至20.0μl.以及

[0574] 7.在37℃下孵育3小时至5小时。

[0575] 可将粗制IVT混合物在4℃下储存过夜以在第二天进行清洗。接着使用1U不含RNA酶的DNA酶来消化原始模板。在37℃下孵育15分钟之后,使用Ambion的MEGACLEAR™试剂盒(Austin,TX)按照制造商的说明书纯化mRNA。此试剂盒可纯化多达500μg的RNA。清洗后,使用NanoDrop对RNA定量,并且通过琼脂糖凝胶电泳对RNA进行分析,以确认RNA为适当大小并且RNA未发生降解。

[0576] 实施例5:构建体和侧翼长度的体内研究

[0577] 执行体内免疫原性研究,以检查具有不同表位数目和侧翼长度的疫苗的效果。使用如下表所示的三种构建体执行研究。如通过生物信息学算法所确定,鼠疫苗编码在小鼠结肠(MC38)肿瘤细胞系中存在的预测新表位(单核苷酸变体)。MC38S-1a含有15个I类表位和5个II类表位,MC38S-2b含有26个I类表位和8个II类表位,并且MC38S-3b含有30个I类表位和10个II类表位。在下表中,制备了三种不同的疫苗,在1a中-所有表位被侧接氨基酸包围,总长度为31个氨基酸;对于2b,所述表位被氨基酸包围,表位+侧翼总共25个氨基酸;接着对于3b,所述表位被氨基酸包围,每个表位+侧翼的总长度等于21aa。所述表位的长度根据其预测要结合的MHC分子可能会稍有不同,但是在此实施例中,考虑到这种轻微变化对总长度进行了调整以使总长度保持在31、25或21。

[0578]	mRNA	MC38S-1a	MC38S-2b	MC38S-3b
	表位数目	20	34	40
	侧翼长度	31	25	21
	总nt	1993	2680	2662

[0579] 在第1天(d1;初免)和第8天(d8;加强)给小鼠施予3μg或10μg的测试mRNA疫苗。在第15天收获脾细胞用于ELIspot分析。简而言之,将每孔400,000个细胞与1μg/mL肽一起孵育16-18小时,接着对算IFN γ 点形成单位(SFU)进行计数。将对应于所有三种疫苗中所含的表位的最小肽用于再刺激。对不同组的统计比较显示在下表中:

[0580]

		MC38S-1a		MC38S-2b		MC38S-3b	
剂量		3 ug	10 ug	3 ug	10 ug	3 ug	10 ug
I 类	8	2.90	203.60	5.40	79.60	1.90	2.60
	10	1298.4***	2000***	1123.3***	1640***	274.4***	1131.5***
	12	11.50	2.10	77.40	231.23	4.10	152.90
	13	18.30	2.00	89.60	135.60	1.20	10.00
	15	4.70	13.30	1.20	16.50	3.10	0.40
	19	10.40	26.70	1.00	49.40	1.80	5.90
II 类	37	8.90	4.50	0.90	8.20	2.50	0.60

[0581] 再刺激

		MC38S-1a		MC38S-2b		MC38S-3b	
剂量		3 ug	10 ug	3 ug	10 ug	3 ug	10 ug
I 类	8	1.30	58.40	0.80	64.10	3.10	0.30

[0583]		10	1394.7*	1232.2***	1034.7*	1589.9***	537.6*	347.6***
		12	211.90	148.50	211.60	422.6**	14.10	3.8**
		13	3.30	4.40	1.00	129.70	0.70	0.60
		15						
		19	1.80	2.40	0.90	54.80	2.00	1.40
II 类		37	13.50	24.40	1.20	1.90	0.70	6.30

[0584] 注意：*=与彼此相比都显著的；**=34聚体与40聚体相比；***=(20聚体和34聚体)与40聚体相比

[0585] 如图4A至图4C所示，检测到20聚体/31侧翼与34聚体/25侧翼疫苗对I类表位具有相当的免疫应答，但对30聚体/21侧翼在3μg剂量和10μg剂量下均未检测到。34聚体构建体对所述再刺激中的一些展示出唯一检测到的应答。

[0586] 实施例6：表位选择

[0587] mRNA表位选择过程可能涉及以下：

[0588] 1) 新抗原预测步骤产生在肿瘤而非正常组织中特异性地表达的突变源性肽的列表，并基于所预测到的新抗原被患者的HLA分子呈递的能力以及它们在肿瘤转录组中的丰度和频率来选择最有可能产生稳健的肿瘤特异性T细胞应答的新抗原的子集。

[0589] 2) 可使用自我性分析，通过排除与患者正常组织中可能表达的其他肽匹配的肽来使新抗原与患者基因组中其他序列之间的分子模拟风险减到最低。在多联体中排列新抗原，以使新抗原接合点处假表位的产生减到最少。

[0590] 3) 疫苗设计涉及将所选择的新抗原设计成多联构建体，所述多联构建体可产生为易于合成而优化的核酸序列。

[0591] 新抗原预测

[0592] 用于新抗原预测和选择的核心算法确定了变体的mRNA丰度和频率以及所预测到的变体与患者HLA靶标的结合。通过将体细胞DNA变体的位置对来自高置信度人基因组注释

GENCODE的氨基酸(AA)序列作图来生成肽。RNA-Seq数据用于支持单核苷酸变体水平下的突变识别,并且用于确定基因组和转录组中的变异频率。

[0593] mRNA中的大多数新抗原可由这样的肽组成,所述肽的中心有一个单突变AA,在C末端和N末端处有12个侧接AA,导致每种新抗原的长度为25个氨基酸(mRNA序列中有75个核苷酸)。具有多种突变AA的Indel将由长度为25个AA的AA序列组成,所述AA序列含有至少1种或多种突变AA直至整个25聚体都为突变AA。在突变发生在距蛋白质末端<12AA的位置处的情况下,肽和相应核苷酸的长度可能会更短。在一些实施方案中,优选肽长度将为13AA,基于对所有肿瘤类型的突变组的广泛分析,这将是罕见的。

[0594] 针对每种新抗原评估与抗肿瘤T细胞应答相关的数个特征,这些特征包括以下:1)得自WES和RNA-Seq数据的变体识别置信度;2)得自RNA-Seq数据的mRNA转录物丰度;3)得自WES和RNA-Seq数据的变异等位基因频率;4)得自NetMHCpan和NetMHCIIpan的预测HLA结合亲和力。

[0595] 可靶向患者的HLA同种异型,因为它们将新抗原呈递给患者的T细胞。HLA基因是人基因组中最具多态性的,并且共显性表达导致大多数个体在一些基因座处是杂合的。HLA-A、HLA-B和HLA-C基因座编码I类同种异型,而HLA-DR、HLA-DP和HLA-DQ编码II类同种异型。在一些实施方案中,可将更大的权重分配给HLA-A、HLA-B和HLA-DR(核心靶标)的预测结合物,并且将更低的(但非零)权重分配给患者的其他HLA同种异型(补充靶标)。几乎所有个体都具有至少一种HLA-A、HLA-B和HLA-DR功能同种异型(即核心MHC等位基因),这些是所有已知人表位中的约90%人表位的限制因子(图5)。极少观察到HLA-C限制性或同种异体反应性T细胞,并且HLA-C的细胞表面表达是针对HLA-A和HLA-B观察到的细胞表面表达的10%。其余补充靶标编码II类分子,并且个体可能不具有编码它们的基因。此外,即使使用最先进的NGS和其他基于序列的分型方法,对这些补充II类靶标的4位精度分型也常常是不明确的。如果对核心或补充HLA靶标的基于NGS的等位基因分型是不明确的,那么在对新抗原进行排名时可能不考虑所述一种或多种等位基因。

[0596] 自我性检查

[0597] 可对每种新抗原执行自我性检查。通过根据患者自身的种系蛋白质编码变体集合定制序列,使用来自参照人基因组注释的蛋白质编码转录物氨基酸序列来创建特定于患者的转录物集合。这种特定于患者的外显子组(不包括含有新抗原的基因)可用于检查每个I类HLA结合新抗原表位(8至11聚体)是否具有100%精确的自我匹配。可将使用此工具在基因组和/或转录组中其他地方鉴定为100%自我匹配的任何新抗原从mRNA构建体中排除。

[0598] 新抗原选择

[0599] 可对所有未被自我性检查排除的变体进行评估以确定是否将其包含于特定于患者的mRNA构建体设计中。基于MHC结合预测不完善并且RNA-Seq灵敏度可能受活检物肿瘤含量和测序深度限制的知识,可使用预定义权重而不是硬过滤器。

[0600] 在一些实施方案中,可将每种mRNA构建体设计成具有多达34种新抗原(含有长度达25个氨基酸/75个核苷酸的肽)或13至34种的任选范围的新抗原。此范围对应于mRNA序列长度的1,235-2,924个核苷酸。在包含34种新抗原的构建体的示例性实施方案中,可通过首先选择前29种I类HLA新抗原,并接着选择前5种II类HLA新抗原来确定组成。如果特定新抗原作同时被选为I类和II类新抗原,则可将其作为5种II类新抗原之一来计数。下一个得分

最高的I类新抗原将自动填充在由这些双重I类和II类预测结合物产生的所得新抗原槽中。

[0601] 低突变负荷肿瘤

[0602] 考虑到肿瘤突变组的固有变异性,可用本文所述的癌症疫苗治疗具有低突变负荷的罕见肿瘤病例。在这些实施方案中,可能期望使用少于34种新抗原来创建个体mRNA构建体。例如,可使用少至7种肿瘤新抗原。对于鉴定出少于最优34种抗原但大于或等于13种新抗原的情况,可制取这样的构建体,其中每种新抗原都将被包含在mRNA构建体中一次。在肿瘤突变组中发现少于13种新抗原的实施方案中,可复制新抗原以满足期望的13种新抗原槽。

[0603] 假表位

[0604] 可使新抗原在多联体中有序化,以使新抗原接合点处假表位的产生减到最少。替代地,可使用间隔子(诸如单氨基酸间隔子)来破坏表位并降低预测的HLA结合亲和力。

[0605] 群体和端对端测试

[0606] 对从数个生物库(biobanking repositories)获得的15个肿瘤和血液样品执行NGS。样品来自不同样式的(例如福尔马林固定的石蜡包埋的[FFPE]和新鲜冷冻的)多种肿瘤类型。作为完整鉴定方案的一部分,针对这些代表性样品中的每一个在NGS数据上执行本文所述的方法。另外,使用4个相关的肿瘤样品执行测试。对源自单名患者的三个肿瘤细胞系和原发性肿瘤样品进行WES和RNA-Seq处理,并对结果进行分析。

[0607] 当比较四个独立的输出时,在所识别变体、新抗原排名与被选择包含在疫苗中的那些之间观察到强一致性(图7A至图7D)。发现了差异,但这些差异可通过在体外从原发性肿瘤繁殖的细胞系和彼此的趋异来解释。在四个样品中鉴定出的369种变体中,90.5%是所有样品所共同具有的。在使用原始新抗原得分的情况下,与肿瘤相比并且当该得分出现明显的趋异时,所有细胞系之间都具有很强的相关性,这是由于细胞系或肿瘤中缺乏RNA-Seq数据。当通过本文所述的分析方法从每个肿瘤样品中独立地选择新抗原时,34种是所有4种疫苗设计所共同具有的,而超过5种是3种疫苗设计所共同具有的。总体而言,该分析证明NGS过程、变体识别和mRNA分析系统是稳健的、可重现的,并且产生合理的输出。

[0608] 等价方案

[0609] 本领域技术人员将会认识到或者仅使用常规实验就能够探知本文所述的本公开的具体实施方案的许多等效方案。此类等效方案意图由所附权利要求书涵盖。

[0610] 本文所公开的包括专利文件在内的所有参考文献都以引用方式整体并入本文。

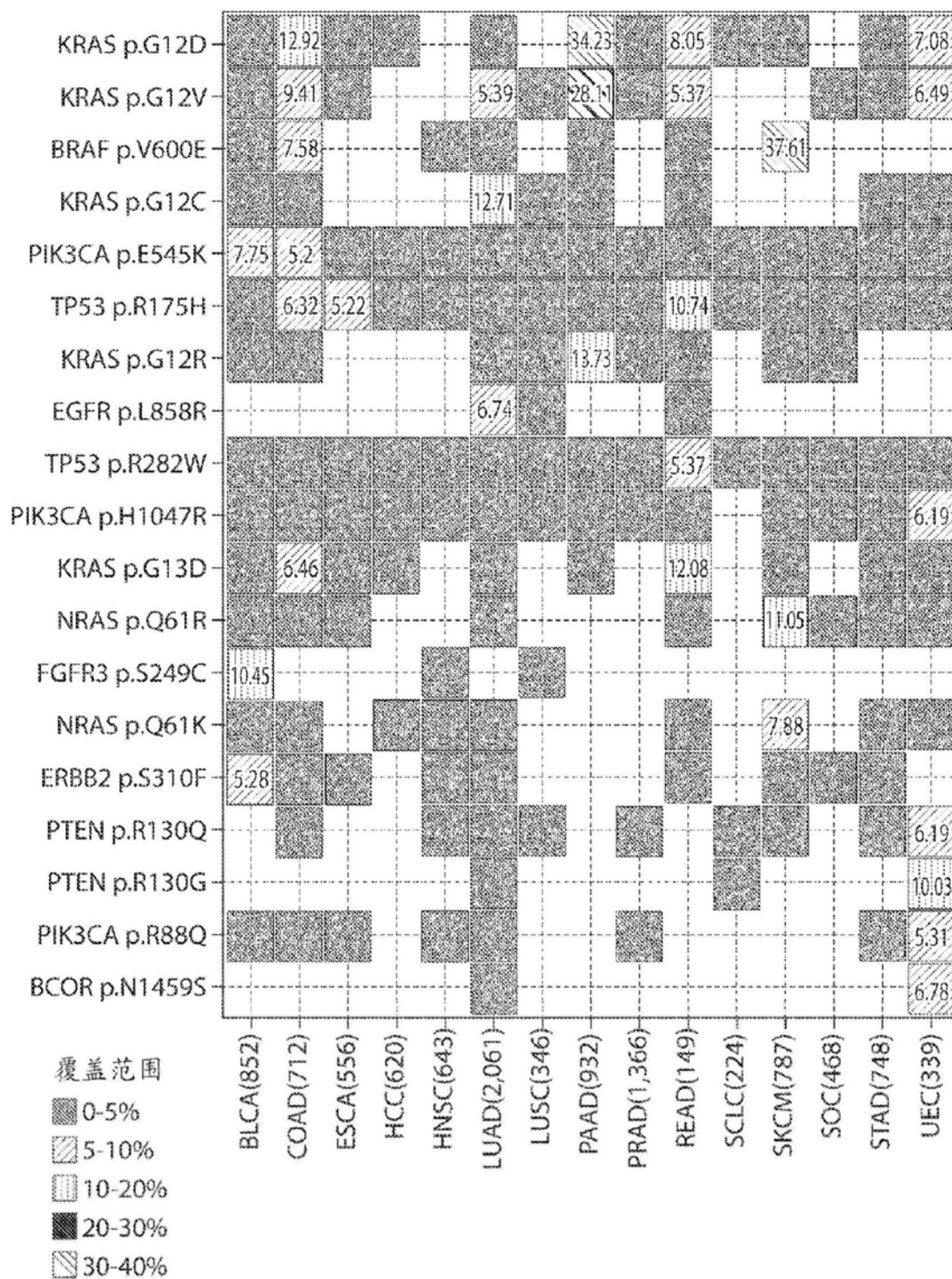


图1

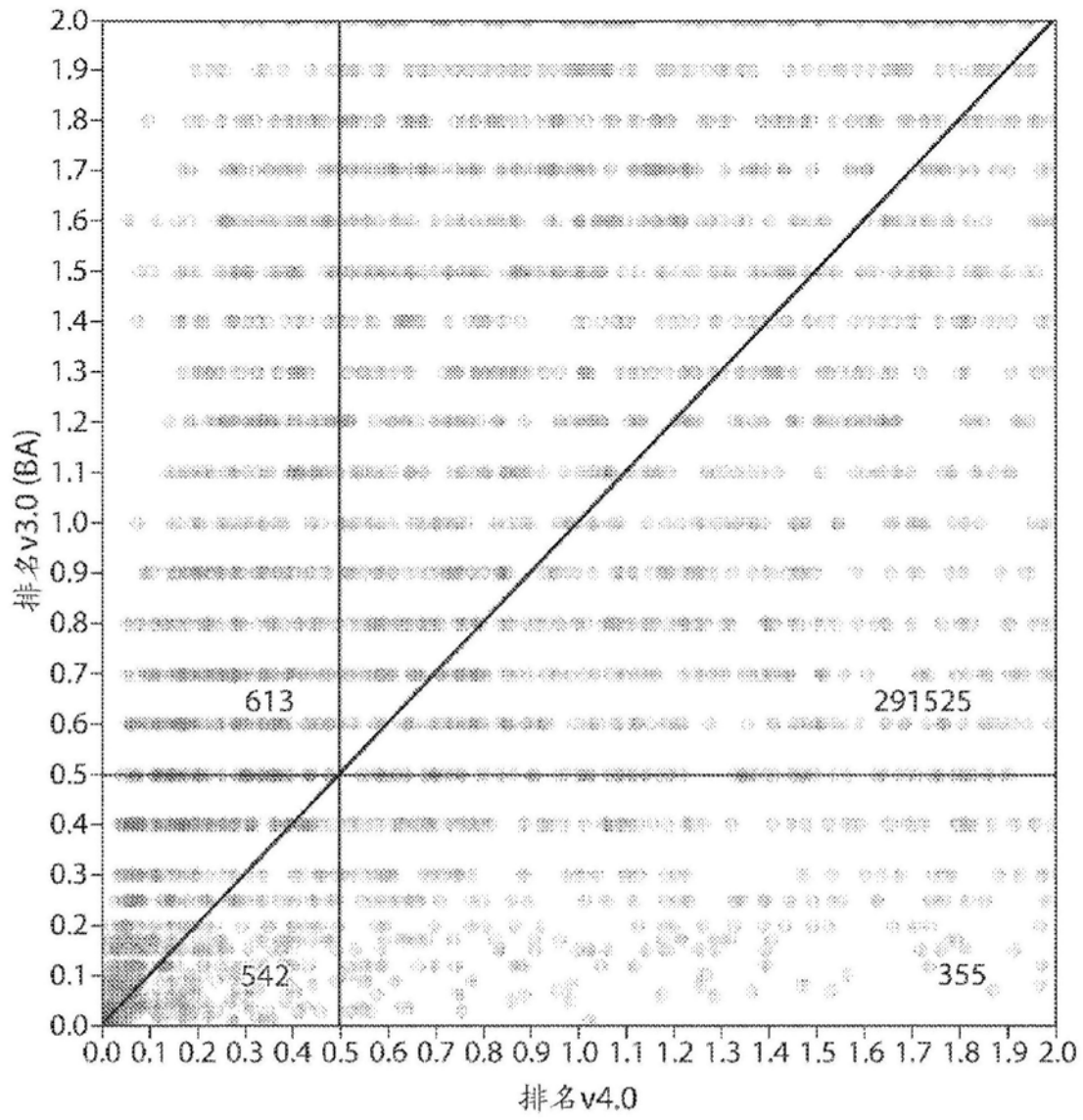


图2

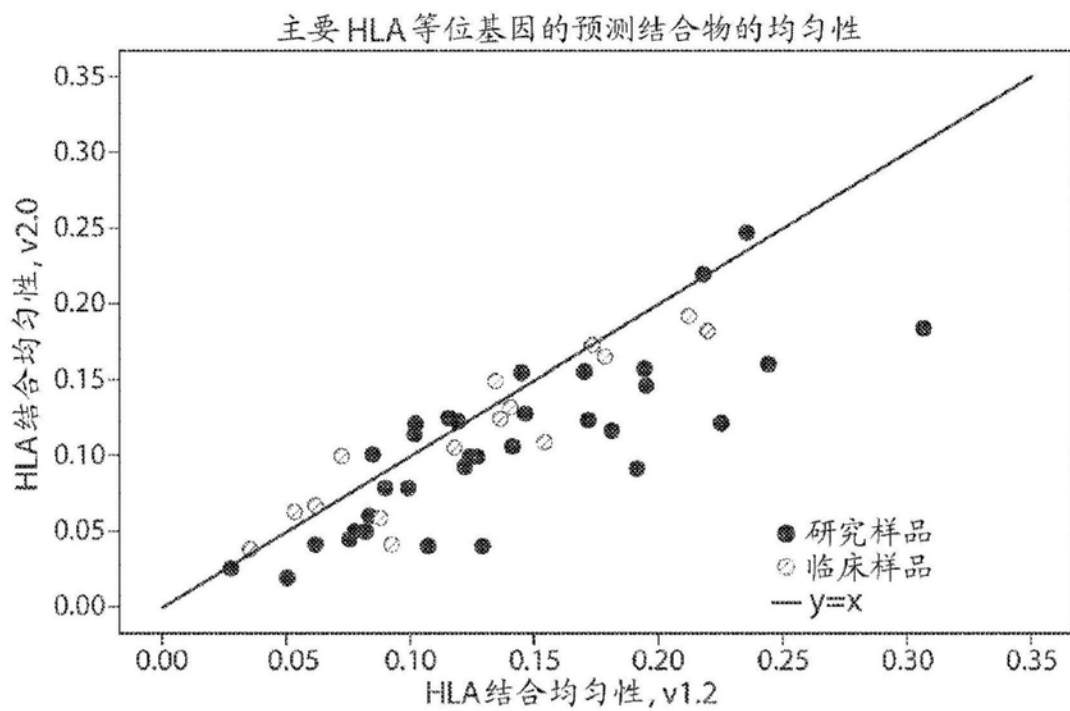


图3A

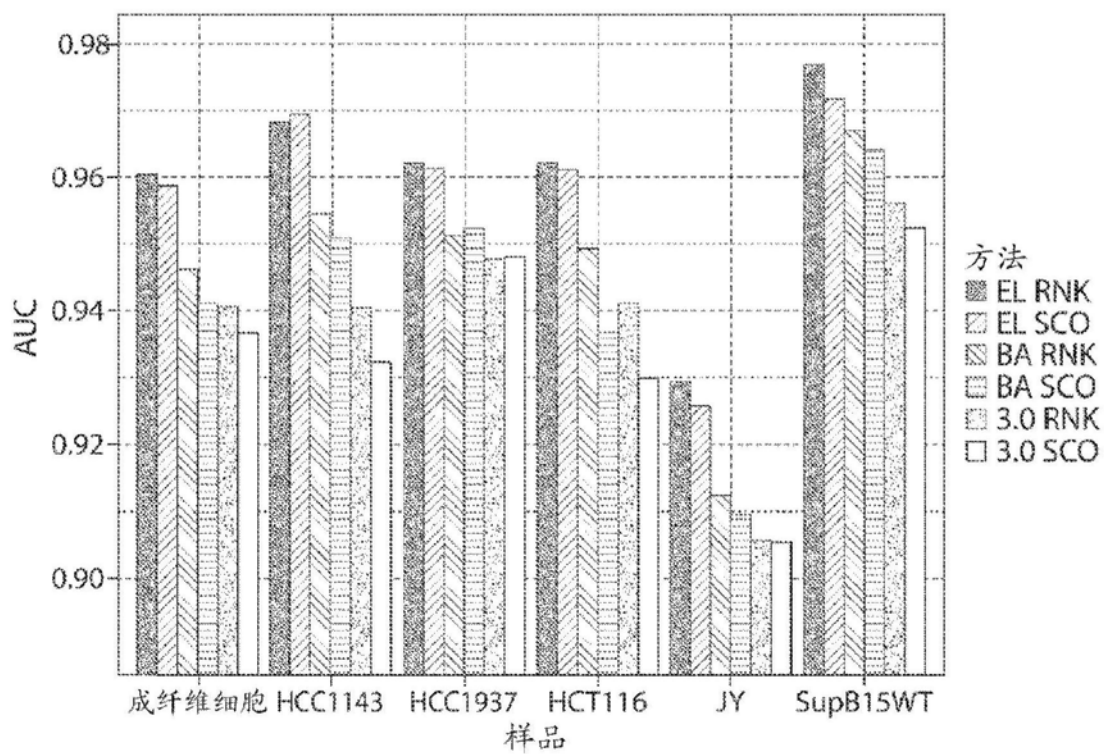


图3B

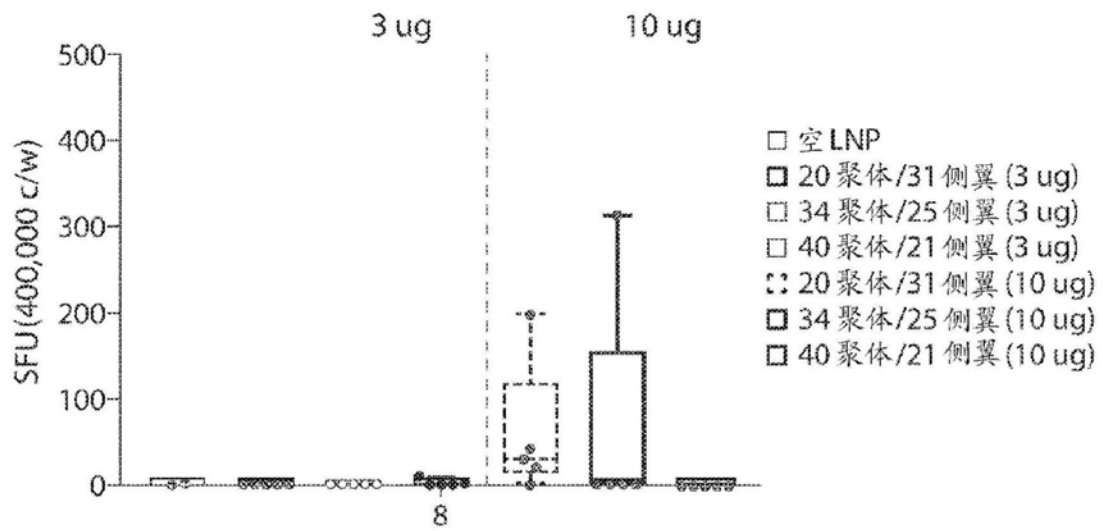


图4A

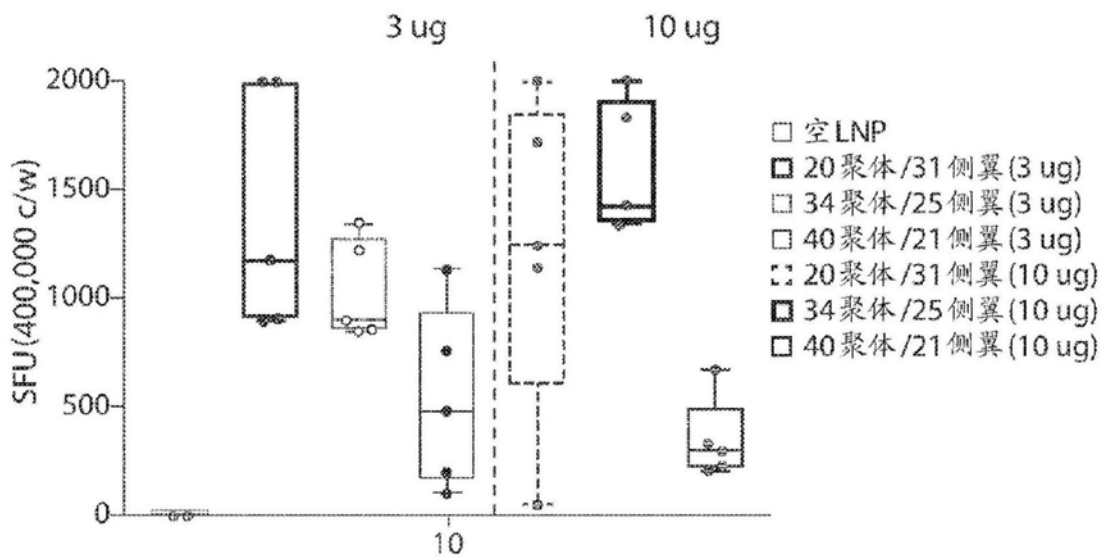


图4B

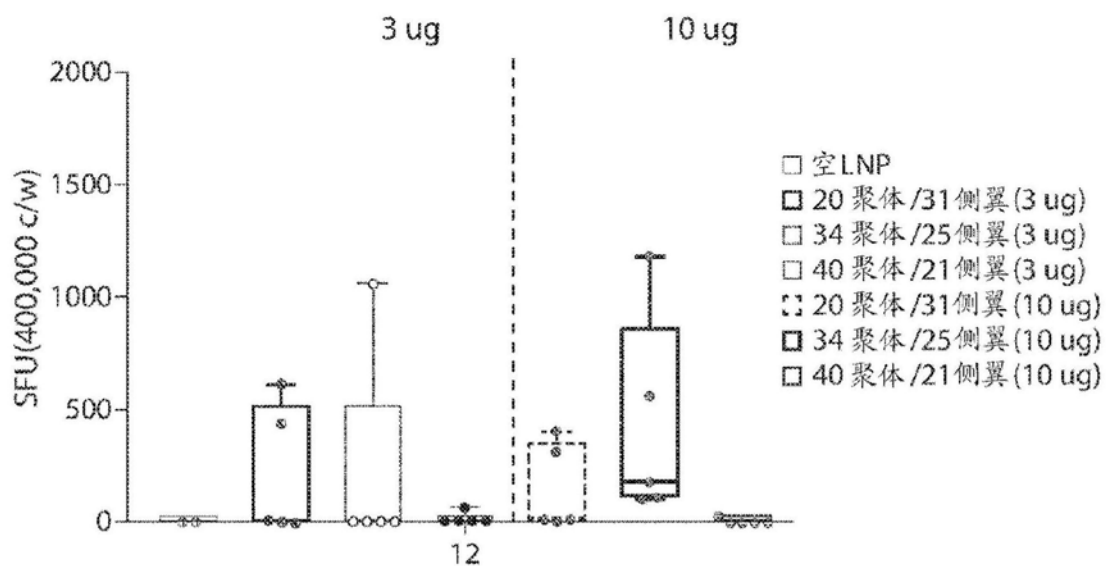


图4C

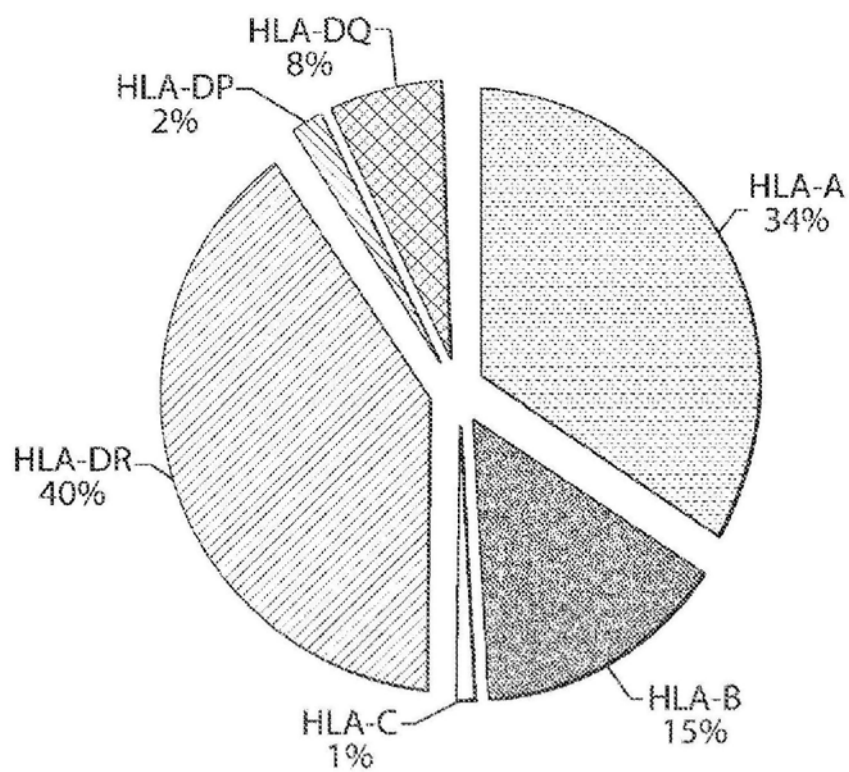


图5A

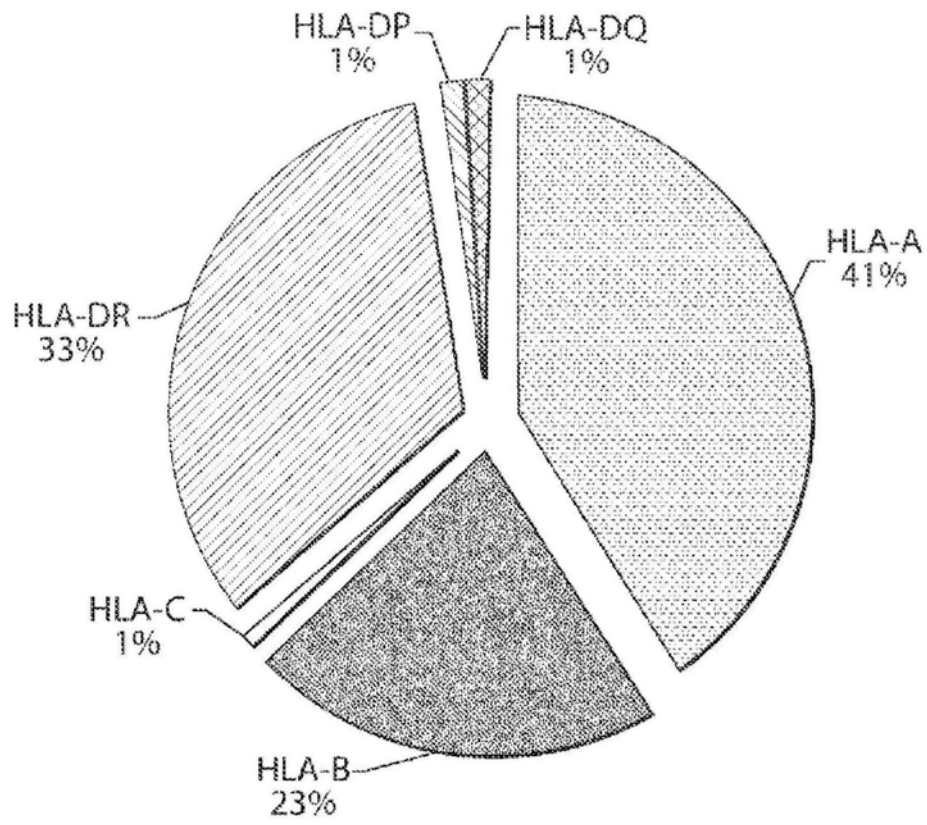


图5B

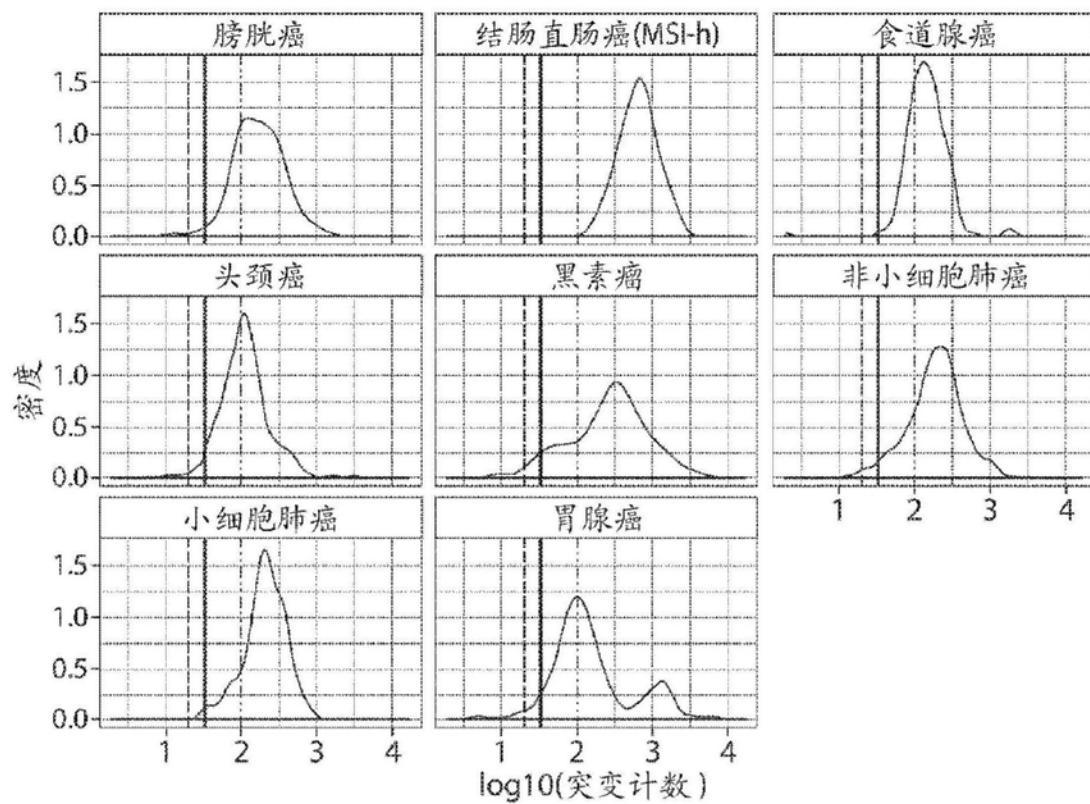


图6

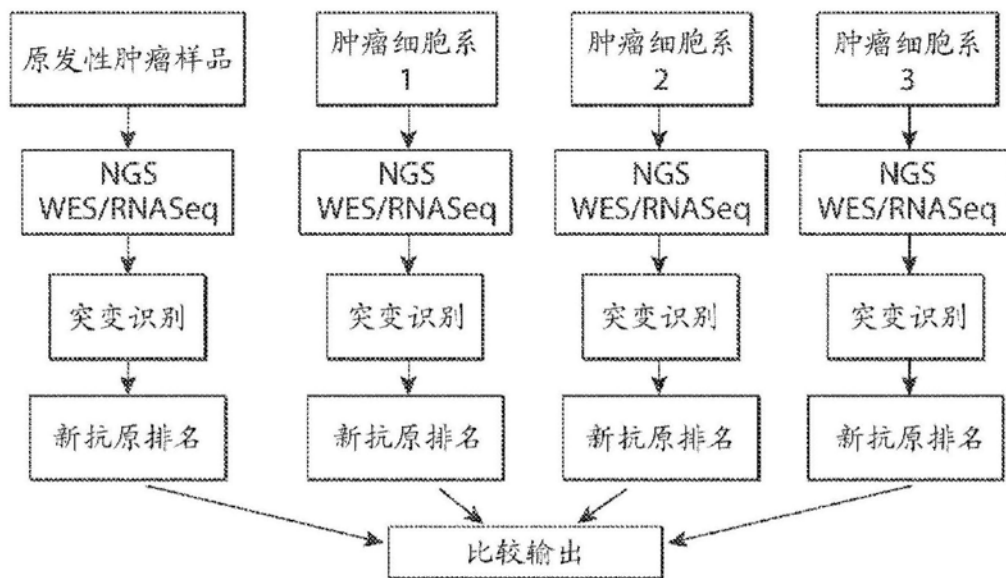


图7A

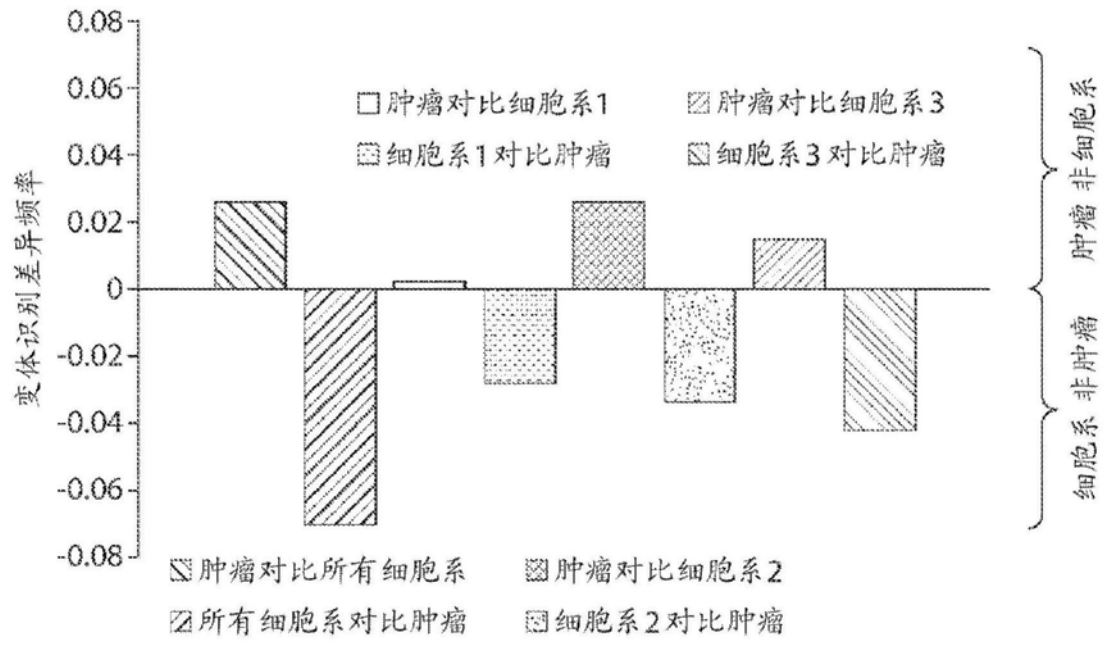


图7B

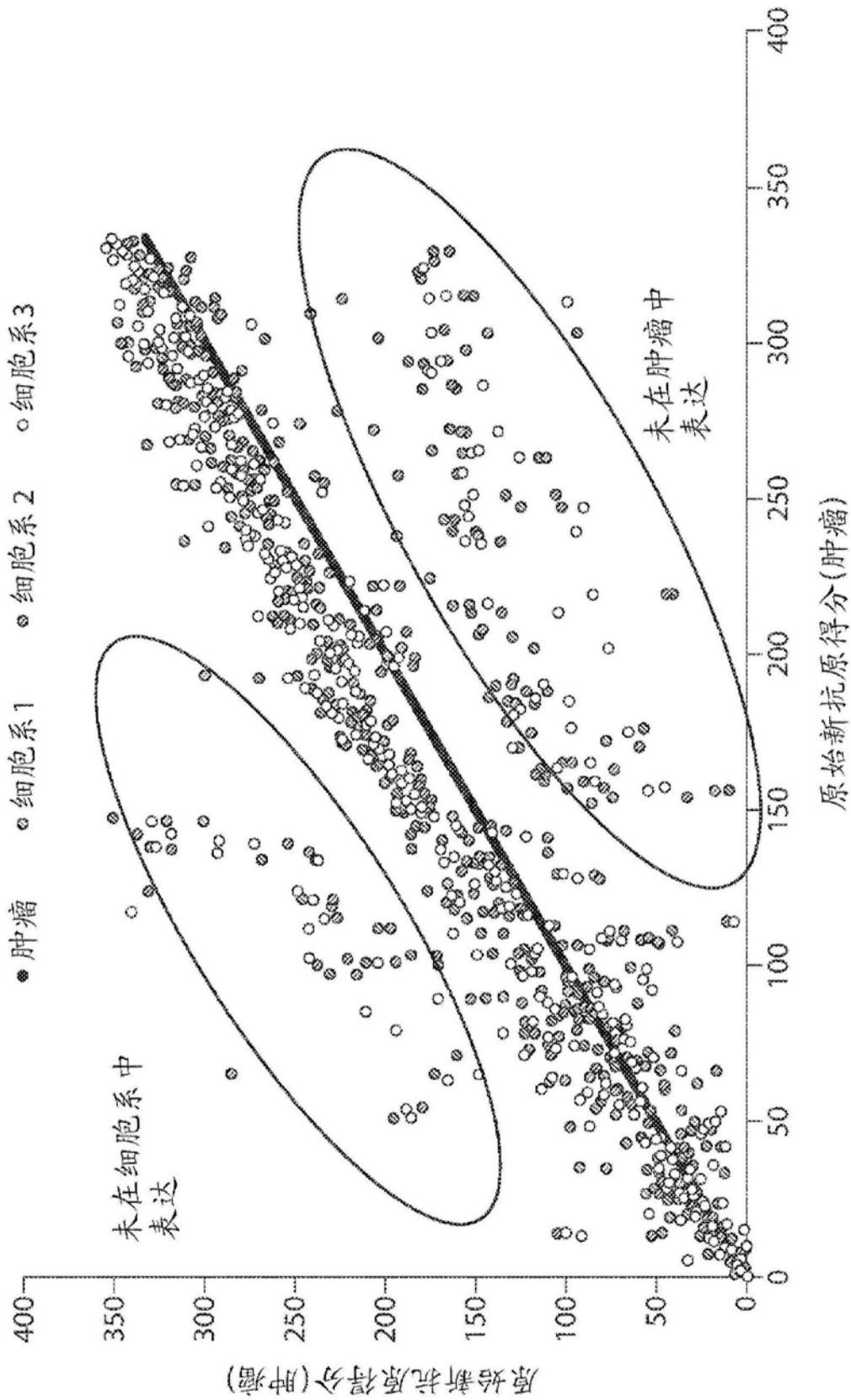


图7C

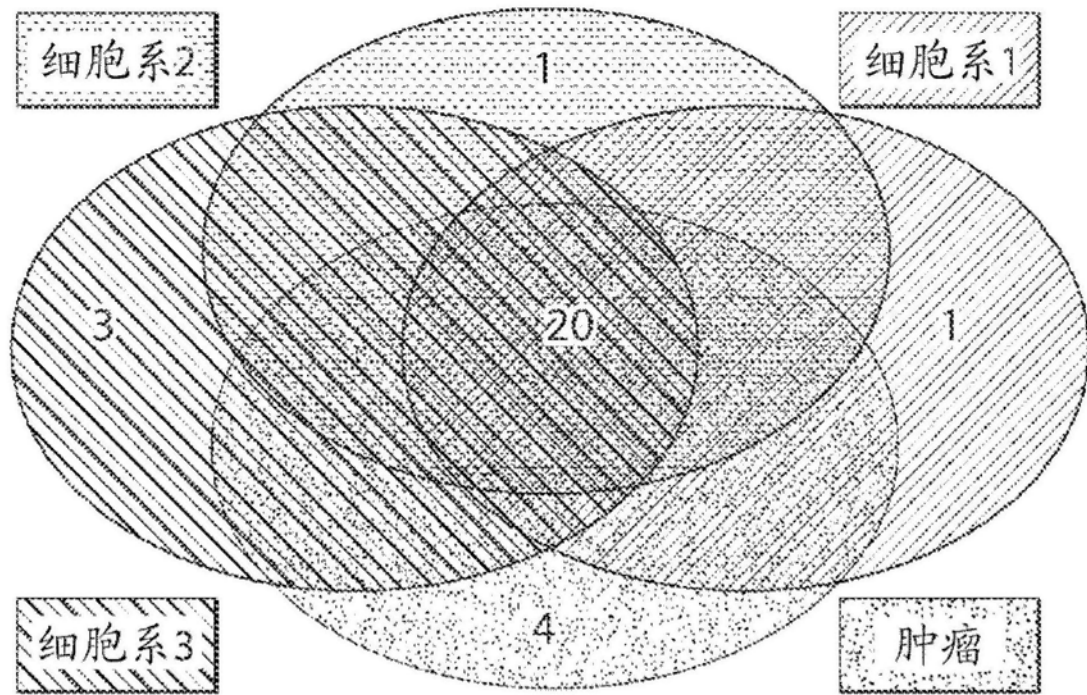


图7D