



(12)发明专利

(10)授权公告号 CN 104685360 B

(45)授权公告日 2018.02.13

(21)申请号 201380043182.4

(73)专利权人 比奥德希克斯股份有限公司

(22)申请日 2013.03.15

地址 美国科罗拉多州

(65)同一申请的已公布的文献号

专利权人 全球免疫股份有限公司

申请公布号 CN 104685360 A

(72)发明人 J.罗德 H.罗德

(43)申请公布日 2015.06.03

(74)专利代理机构 北京市柳沈律师事务所
11105

(30)优先权数据

代理人 闵丹

61/664,329 2012.06.26 US

(51)Int.Cl.

61/664,308 2012.06.26 US

G01N 33/68(2006.01)

(85)PCT国际申请进入国家阶段日

G01N 33/574(2006.01)

2015.02.13

审查员 胡琴明

(86)PCT国际申请的申请数据

PCT/US2013/032024 2013.03.15

(87)PCT国际申请的公布数据

W02014/003853 EN 2014.01.03

权利要求书7页 说明书48页

序列表12页 附图37页

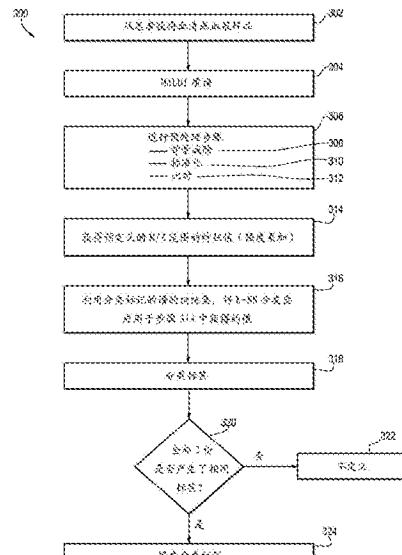
(54)发明名称

用于选择和去选择用产生免疫应答的疗法治疗的癌症患者的质谱方法

(57)摘要

一种用于在治疗前预测癌症患者是否有可能或不可能从施用基于酵母的免疫应答生成型治疗中受益的方法和系统，所述治疗可以是单独或联合另一种抗癌治疗的针对基于突变Ras癌症的基于酵母的免疫治疗。本方法使用血液来源的患者样品的质谱和配置为分类器的计算机，所述分类器使用来自单独或联合另一种抗癌治疗的免疫应答生成型治疗中受益或未受益的其他癌症患者的分类标签谱的训练集。还公开了治疗癌症患者的方法，包括对根据本文公开的预测性质谱方法测出的患者施用基于酵母的免疫应答生成型治疗（其可以是针对基于突变Ras癌症的基于酵母的免疫治疗），其中所述谱的分类标签表明该患者很有可能从该基于酵母的免疫治疗中受益。

CN 104685360 B



CN

1. 质谱仪在制备系统中的用途,所述系统用于执行预测癌症患者是否可能从单独或配合另一抗癌治疗施用的针对癌症的基于酵母的免疫应答生成型治疗中受益的方法,其中所述方法包括如下步骤:

(a) 获得患者血液来源的样品;

(b) 使用所述质谱仪对该样品进行质谱测定并从该样品获得质谱;并且

(c) 在程序计算机中,对来自所述样品的质谱执行一个或多个预定义的预处理步骤,在执行预处理步骤之后,在预定义的m/z范围内获取来自所述样品的质谱中所选特征的累积强度值,并将累积强度值与包括来自其他癌症患者的分类标签质谱(class-labeled mass spectra)的训练集进行比较并用分类标签对来自所述样品的质谱进行分级;

其中该分类标签预测患者是否有可能从单独或配合其它抗癌治疗进行的基于酵母的免疫应答生成型治疗中受益。

2. 权利要求1的用途,其中基于酵母的免疫应答生成型治疗包括针对突变Ras阳性癌症的基于酵母的免疫治疗。

3. 权利要求2的用途,其中针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗包括施用GI-4000或等价物。

4. 权利要求3的用途,其中针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗协同吉西他滨或等价物对患者进行治疗。

5. 权利要求1的用途,其中癌症患者包括胰腺癌患者。

6. 权利要求2的用途,其中癌症患者包括胰腺癌患者。

7. 权利要求1的用途,其中训练集包括来自其他癌症患者的分类标签质谱,所述患者从施用单独或联合其它抗癌治疗的基于酵母的免疫应答生成型治疗或免疫治疗中受益或不受益。

8. 权利要求1的用途,其中选择的特征包括表1、2、3、4、5或6中列举的一种或多种特征。

9. 质谱仪在制备系统中的用途,所述系统用于执行预测癌症患者是否不可能从单独或联合其它抗癌治疗施用的基于酵母的免疫应答生成型治疗中受益的方法,其中所述方法包括如下步骤:

(a) 获得患者血液来源的样品;

(b) 使用所述质谱仪对该样品进行质谱测定并从该样品获得质谱;

(c) 在程序计算机中,对来自所述样品的质谱执行一个或多个预定义的预处理步骤,在执行预处理步骤之后,在预定义的m/z范围内获取来自所述样品的质谱中所选特征的累积强度值,并将累积强度值与包括来自其他癌症患者的分类标签质谱的训练集进行比较,并用分类标签对来自所述样品的质谱进行分级;

其中分类标签预测患者是否不可能从单独或联合其它抗癌治疗施用的基于酵母的免疫应答生成型治疗中受益。

10. 权利要求9的用途,其中针对癌症的基于酵母的免疫应答生成型治疗包括针对突变Ras阳性癌症的基于酵母的免疫治疗。

11. 权利要求10的用途,其中针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗包括施用GI-4000或等价物。

12. 权利要求10的用途,其中预测患者不可能从联合吉西他滨或等价物施用的针对突

变Ras阳性癌症的针对癌症的基于酵母的免疫应答生成型治疗中受益。

13. 权利要求9的用途,其中癌症患者包括胰腺癌患者。

14. 权利要求9的用途,其中训练集包括来自其他癌症患者的分类标签质谱,所述患者从施用单独或联合其它抗癌治疗的免疫治疗中受益或未受益。

15. 权利要求9的用途,其中选择的特征包括表1、2、3、4、5或6中列举的一种或多种特征。

16. 一种用于预测癌症患者是否有可能从单独或联合另一种抗癌治疗施用的针对癌症的基于酵母的免疫应答生成型治疗中受益的系统,包括,组合的:

 质谱仪,从来自癌症患者的血液来源的样品中产生质谱;

 机器可读存储器,存储来自其他癌症患者的分类标签质谱的训练集,该训练集包括来自多个未从单独或与另一种抗癌治疗联合施用的针对癌症的基于酵母的免疫应答生成型治疗中受益的患者的分类标签质谱,和来自多个从单独或与另一种抗癌治疗联合施用的免疫应答生成型治疗中受益的患者的分类标签质谱;和

 计算机系统,配置为对来自血液来源的样品的质谱进行操作并使用所述训练集对来自血液来源的样品的质谱进行分类、生成来自血液来源的样品的质谱的分类标签,其中该分类标签预测患者是否可能从单独或与另一种抗癌治疗联合施用的针对癌症的基于酵母的免疫应答生成型治疗中受益。

17. 权利要求16的系统,其中该分类标签预测患者是否可能从施用针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗中受益。

18. 权利要求16的系统,其中针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗包括施用GI-4000或等价物。

19. 权利要求16的系统,其中该分类标签预测患者是否可能从对患者施用GI-4000和吉西他滨中受益。

20. 权利要求16的系统,其中癌症患者包括胰腺癌患者。

21. 权利要求16的系统,其中所述计算机系统被配置为获得来自血液来源的样品的质谱以m/z范围表示的特征,包括表1、2、3、4、5或6中列举的一种或多种特征的累积强度值。

22. 一种用于预测癌症患者是否不可能从单独或联合其它抗癌治疗施用的针对癌症的基于酵母的免疫应答生成型治疗中受益的系统,包括,组合的:

 质谱仪,从来自癌症患者的血液来源的样品中产生质谱;

 机器可读存储器,存储来自其他癌症患者的分类标签质谱的训练集,该训练集包括来自多个未从单独或与另一种抗癌治疗联合施用的针对癌症的基于酵母的免疫应答生成型治疗中受益的患者的分类标签质谱,和来自多个从单独或与另一种抗癌剂联合施用的针对癌症的基于酵母的免疫应答生成型治疗中受益的患者的分类标签质谱;和

 计算机系统,配置为对来自血液来源的样品的质谱进行操作并使用所述训练集对来自血液来源的样品的质谱进行分类、生成来自血液来源的样品的质谱的分类标签,其中该分类标签预测患者是否不可能从单独施用或与另一种抗癌剂联合施用的针对癌症的基于酵母的免疫应答生成型治疗中受益。

23. 权利要求22的系统,其中针对癌症的基于酵母的免疫应答生成型治疗包括针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗。

24. 权利要求23的系统,其中针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗包括施用GI-4000或等价物。

25. 权利要求24的系统,其中分类标签预测患者是否不可能从施用GI-4000和吉西他滨中受益。

26. 权利要求24的系统,其中癌症患者包括胰腺癌患者。

27. 权利要求22的系统,其中所述计算机系统被配置为获得来自血液来源的样品的质谱以m/z范围表示的特征,包括表1、2、3、4、5或6中列举的一种或多种特征的累积强度值。

28. 权利要求1的用途,其中对样品执行质谱的步骤包括执行MALDITOF质谱。

29. 权利要求28的用途,其中MALDI-TOF质谱包括执行DeepMALDI质谱,其中血液来源的样品接受20,000次以上激光射击。

30. 权利要求1的用途,其中从对癌症患者血液来源的样品的多个质谱的考察中获得预定义的m/z区域,每种质谱由DeepMALDI质谱获得,其中血液来源的样品接受20,000次以上激光射击。

31. 权利要求16的系统,其中质谱仪包括配置为通过使血液来源的样品接受20,000次以上激光射击来获得谱的MALDI-TOF质谱仪。

32. 执行分类算法的计算机系统,其整合了非暂时性计算机可读介质,所述非暂时性计算机可读介质以包括来自多名癌症患者的分类标签质谱的分类器的训练集的形式存储数据,包括来自多名未从单独或联合另一种抗癌治疗的基于酵母的免疫应答生成型治疗中受益的癌症患者的分类标签质谱,和多名从单独或联合另一种抗癌治疗的针对癌症的基于酵母的免疫应答生成型治疗中受益的癌症患者的分类标签质谱,所述分类算法产生表明受益或未受益的状态的分类标签。

33. 权利要求32的计算机系统,其中所述介质以包括多名癌症患者分类标签质谱的分类器的训练集的形式存储数据,所述训练集包括来自多名未从单独或联合另一种抗癌治疗的针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗中受益的癌症患者的分类标签质谱,和多名从单独或联合另一种抗癌治疗的针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗中受益的癌症患者的分类标签质谱。

34. 权利要求33的计算机系统,其中所述分类标签质谱表明该训练集中的患者从GI-4000和吉西他滨组合用药中受益或未受益。

35. 权利要求32的计算机系统,其中训练集中的癌症患者包括胰腺癌患者。

36. 权利要求32的计算机系统,其中所述介质以分类器的形式存储密码以将包括表1、2、3或4中列出的一种或更多特征的训练集中的特征与测试质谱进行比较,从而生成该测试质谱的分类标签。

37. 对质谱进行分类的系统,包括:

权利要求32所述非暂时性计算机可读介质,和

配置为分类器以使用权利要求32的训练集对癌症患者的谱进行分类的通用计算机。

38. 权利要求37的系统,其中的计算机实施的是K最近邻分类器。

39. 创建分类器的方法,所述分类器用于在治疗前预测癌症患者是否有可能从单独或联合其它抗癌治疗的针对癌症的基于酵母的免疫应答生成型治疗中受益,包括如下步骤:

a) 对通过单独或联合其它抗癌治疗的针对癌症的基于酵母的免疫应答生成型治疗进

行治疗之后获得的大量患者样品进行质谱，并产生对应的质谱集；

b) 通过将提供了步骤a) 样品的患者分为能或不能从治疗中受益的两个组，并给每个谱分配分类标签，来创建分类器的训练集；

c) 分析来自两个组的谱，以鉴定两个组中患者特有的谱的区别性特征；和

d) 通过将分类算法应用于来自施用了该治疗的其他患者的测试谱集来测试该区别性特征预测患者从治疗中受益的能力，并评价该分类算法的表现，所述分类算法用分类标签质谱和步骤c) 中鉴定的区别性特征进行。

40. 权利要求39的方法，包括用两组患者的谱中不同的区别性特征重复步骤b) 、c) 和d) 。

41. 权利要求39的方法，其中分类算法包括K-最近邻算法。

42. 权利要求39的方法，其中步骤a) 的质谱包括MALDI-TOF质谱仪中的DeepMALDI质谱，其中样品接受20,000次以上激光射击。

43. 权利要求39的方法，其中区别性特征包括表1、2、3、4、5或6中列举的一种或多种特征。

44. 权利要求39的方法，进一步包括基于训练集中的谱特征进行谱的部分离子流标准化。

45. 权利要求39的方法，其中分类标签通过考察训练集中的患者的临床数据来判断。

46. 分类标签质谱在制备系统中的用途，所述系统用于选择待用针对癌症的基于酵母的免疫应答生成型治疗来治疗的癌症患者，

其中患者将通过质谱和分类测试进行选择，所述测试包括如下步骤：

(a) 对来自患者的血液来源的样品执行质谱测定，并从样品获得质谱；

(b) 在程序计算机中，对来自所述样品的质谱执行一个或多个预定义的预处理步骤，在执行预处理步骤之后，在预定义的m/z范围内获取来自所述样品的质谱中所选特征的累积强度值，并将该累积强度值与包含用针对癌症的基于酵母的免疫应答生成型治疗进行治疗的其他癌症患者的分类标签质谱的训练集进行比较，并用分类标签给来自所述样品的质谱进行分级；

且其中来自所述样品的质谱的分类标签表明患者很有可能从针对癌症的基于酵母的免疫应答生成型治疗中受益。

47. 权利要求46的用途，其中患者在针对癌症的基于酵母的免疫应答生成型治疗之前、同时或之后被另外进行一种或多种额外的抗癌治疗。

48. 权利要求47的用途，其中额外的抗癌治疗选自：手术、化疗、放疗、靶向癌症治疗、姑息治疗、或其任意组合。

49. 权利要求46的用途，其中针对癌症的基于酵母的免疫应答生成型治疗是表达一种或多种来自患者癌症的癌症抗原的完整的、热失活的重组酵母。

50. 权利要求49的用途，其中酵母来自酿酒酵母 (*Saccharomyces cerevisiae*)。

51. 权利要求46的用途，其中患者罹患突变Ras阳性癌症，且其中患者被施用针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗。

52. 权利要求51的用途，其中患者在进行针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗之前、同时或之后被另外进行一种或多种额外的抗癌治疗。

53. 权利要求52的用途,其中额外的抗癌治疗选自:手术、化疗、放疗、靶向癌症治疗、姑息治疗、或其任意组合。

54. 权利要求51的用途,其中突变Ras阳性癌症选自:胰腺癌、非小细胞肺癌(NSCLC)、结肠直肠癌(CRC)、子宫内膜癌、卵巢癌、黑素瘤或多发性骨髓瘤。

55. 权利要求51的用途,其中突变Ras阳性癌症是胰腺癌。

56. 权利要求51的用途,其中针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗是表达突变Ras抗原的完整的、热失活的重组酵母。

57. 权利要求55的用途,其中针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗是针对癌症的基于酵母的免疫应答生成型治疗的GI-4000系列产品中的产品,或其等价物。

58. 权利要求57的用途,其中产品选自GI-4014、GI-4015、GI-4016或GI-4020。

59. 权利要求51的用途,其中针对癌症的基于酵母的免疫应答生成型治疗联合吉西他滨或其等价物施用。

60. 权利要求51的用途,其中癌症患者的肿瘤在针对癌症的基于酵母的免疫应答生成型治疗之前已手术切除。

61. 权利要求49的用途,其中训练集包括从单独或联合其它抗癌治疗施用的针对癌症的基于酵母的免疫应答生成型治疗中受益和未受益的其他癌症患者的分类标签质谱。

62. 权利要求46的用途,其中预定义的特征包括表1、2、3、4、5或6中列举的一种或多种特征。

63. 权利要求49的用途,其中对样品执行的质谱包括执行
MALDI-TOF质谱。

64. 权利要求63的用途,其中MALDI-TOF质谱包括执行DeepMALDI质谱,其中血液来源的样品接受20,000次以上激光射击。

65. 权利要求49的用途,其中对通过DeepMALDI质谱获得的癌症患者的大量质谱进行考察而获得预定义的m/z范围,在所述DeepMALDI质谱中,这些癌症患者的血液来源的样品接受20,000次以上激光射击。

66. 分类标签质谱在制备系统中的用途,所述系统用于执行决定是否用针对癌症的基于酵母的免疫应答生成型治疗来治疗癌症患者的方法,其中所述方法包括:

a) 对来自癌症患者的血液来源的样品执行质谱测定,并从样品获得质谱;

b) 在程序计算机中,对来自所述样品的质谱执行一个或多个预定义的预处理步骤,在执行预处理步骤之后,在预定义的m/z范围内获取来自所述样品的质谱中所选特征的累积强度值,并将该累积强度值与包含用针对癌症的基于酵母的免疫应答生成型治疗进行治疗的其他癌症患者的分类标签质谱的训练集进行比较,并用分类标签对来自所述样品的质谱进行分级;和

c) 当来自所述样品的质谱的分类标签表明患者很有可能从针对癌症的基于酵母的免疫应答生成型治疗中受益时,基于所述决定对癌症患者施用针对癌症的基于酵母的免疫应答生成型治疗。

67. 权利要求66的用途,其中患者在针对癌症的基于酵母的免疫应答生成型治疗之前、同时或之后被另外进行一种或多种额外的抗癌治疗。

68. 权利要求66的用途,其中针对癌症的基于酵母的免疫应答生成型治疗是表达一种

或多种来自患者癌症的癌症抗原的完整的、热失活的来自酿酒酵母的重组酵母。

69. 权利要求66的用途,其中患者罹患突变Ras阳性癌症,且其中患者被施用针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗。

70. 权利要求69的用途,其中产品选自GI-4014、GI-4015、GI-4016或GI-4020。

71. 权利要求66的用途,其中训练集包括从施用单独或联合其它抗癌治疗的针对癌症的基于酵母的免疫应答生成型治疗中受益的和未受益的其他癌症患者的分类标签质谱。

72. 权利要求66的用途,其中预定义的特征包括表1、2、3、4、5或6中列举的一种或多种特征。

73. 权利要求66的用途,其中对样品执行的质谱包括执行MALDI-TOF质谱。

74. 权利要求73的用途,其中MALDI-TOF质谱包括执行多射(multi-shot)质谱,其中血液来源的样品接受20,000次以上激光射击。

75. 权利要求66的用途,其中对通过多射质谱获得的癌症患者的大量质谱进行考察获得预定义的m/z范围,其中该癌症患者的血液来源的样品接受20,000次以上激光射击。

76. 针对癌症的基于酵母的免疫应答生成型治疗在制备用于治疗癌症患者中癌症的药物中的用途,已经通过质谱和分类测试选择所述患者来进行治疗,所述测试包括如下步骤:

(a) 对来自患者的血液来源的样品执行质谱测定,并从样品获得质谱;

(b) 在程序计算机中,对来自所述样品的质谱执行一个或多个预定义的预处理步骤,在执行预处理步骤之后,在预定义的m/z范围内获取来自所述样品的质谱中所选特征的累积强度值,并将该累积强度值与包含用针对癌症的基于酵母的免疫应答生成型治疗进行治疗的其他癌症患者的分类标签质谱的训练集进行比较,并用分类标签给来自所述样品的质谱进行分级;

且其中来自所述样品的质谱的分类标签表明患者很有可能从针对癌症的基于酵母的免疫应答生成型治疗中受益。

77. 权利要求76的用途,其中患者在针对癌症的基于酵母的免疫应答生成型治疗之前、同时或之后被另外进行一种或多种额外的抗癌治疗。

78. 权利要求77的用途,其中额外的抗癌治疗选自:手术、化疗、放疗、靶向癌症治疗、姑息治疗、或其任意组合。

79. 权利要求76的用途,其中针对癌症的基于酵母的免疫应答生成型治疗是表达一种或多种来自患者癌症的癌症抗原的完整的、热失活的重组酵母。

80. 权利要求79的用途,其中酵母来自酿酒酵母(Saccharomyces cerevisiae)。

81. 权利要求76的用途,其中患者罹患突变Ras阳性癌症,且其中针对癌症的基于酵母的免疫应答生成型治疗是针对突变Ras阳性癌症的基于酵母的免疫治疗。

82. 权利要求81的用途,其中患者在进行针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗之前、同时或之后被另外进行一种或多种额外的抗癌治疗。

83. 权利要求82的用途,其中额外的抗癌治疗选自:手术、化疗、放疗、靶向癌症治疗、姑息治疗、或其任意组合。

84. 权利要求81的用途,其中突变Ras阳性癌症选自:胰腺癌、非小细胞肺癌(NSCLC)、结肠直肠癌(CRC)、子宫内膜癌、卵巢癌、黑素瘤或多发性骨髓瘤。

85. 权利要求81的用途,其中突变Ras阳性癌症是胰腺癌。

86. 权利要求81的用途,其中针对突变Ras阳性癌症的基于酵母的免疫应答生成型治疗是表达突变Ras抗原的完整的、热失活的重组酵母。

87. 权利要求85的用途,其中针对突变Ras阳性癌症的基于酵母的免疫治疗是基于酵母的免疫应答生成型治疗的GI-4000系列产品中的产品,或其等价物。

88. 权利要求87的用途,其中产品选自GI-4014、GI-4015、GI-4016或GI-4020。

89. 权利要求81的用途,其中针对癌症的基于酵母的免疫应答生成型治疗联合吉西他滨或其等价物施用。

90. 权利要求81的用途,其中癌症患者的肿瘤在针对癌症的基于酵母的免疫应答生成型治疗之前已手术切除。

91. 权利要求76的用途,其中训练集包括从单独或联合其它抗癌治疗施用的针对癌症的基于酵母的免疫应答生成型治疗中受益和未受益的其他癌症患者的分类标签质谱。

92. 权利要求76的用途,其中预定义的特征包括表1、2、3、4、5或6中列举的一种或多种特征。

93. 权利要求76的用途,其中对样品执行的质谱包括执行MALDI-TOF质谱。

94. 权利要求93的用途,其中MALDI-TOF质谱包括执行DeepMALDI质谱,其中血液来源的样品接受20,000次以上激光射击。

95. 权利要求76的用途,其中对通过DeepMALDI质谱获得的癌症患者的大量质谱进行考察而获得预定义的m/z范围,在所述DeepMALDI质谱中,这些癌症患者的血液来源的样品接受20,000次以上激光射击。

用于选择和去选择用产生免疫应答的疗法治疗的癌症患者的质谱方法

[0001] 对相关申请的交叉参考

[0002] 本申请根据35 U.S.C. §119 (e) 要求2012年6月26日提交的61/664,308号美国临时专利申请和2012年6月26日提交的61/664,329号美国临时专利申请的优先权的权益。美国临时专利申请号61/664,308和美国临时专利申请号61/664,329各自的全部公开内容在此引入作为参考。

[0003] 关于合作研究协议的声明

[0004] 本发明由2012年2月22日生效的合作研究协议的双方或代表双方完成。合作研究协议的双方是Biodesix, Inc. 和GlobeImmune, Inc.

[0005] 对序列表的参考

[0006] 本申请含有一份通过EFS-Web电子提交的作为文本文档的序列表。该命名为"12-621-PRO_ST25"的文本文档的字节数为19KB, 录制于2013年3月15日。该文本文档中包含的信息在此依据37 CFR§1.52 (e) (5) 全文引入作为参考。

发明领域

[0007] 本发明大致涉及引导癌症患者治疗方法的领域。更具体的, 在一方面, 本发明涉及一种在治疗之前预测患者是否是可能从施用产生免疫应答的疗法(例如, 作为细胞免疫疗法试剂)中受益的一类患者中的一员的方法, 所述治疗单独使用或者配合标准抗癌药物和/或用于治疗癌症的其它治疗方案使用。还公开了鉴定那些不太可能从产生免疫应答的疗法和/或使用产生免疫应答的疗法配合标准化疗药物中受益的一类患者的方法。本公开的方法使用从所述患者的血液来源的样品获得的质谱数据、经配置作为对所述质谱数据进行操作的分类器的计算机、和包括来自其他癌症患者的分类标签谱的训练集。

[0008] 发明背景

[0009] 癌症是一大类各种疾病的统称, 其均涉及未受调控的细胞生长。在癌症中, 细胞不受控制地分裂和生长, 形成恶性肿瘤, 并侵入身体临近部分。癌症还可能通过淋巴系统或血液扩散至身体更远的部分。根据美国国家癌症研究所(NCI)跟踪的这些统计数据, 在2012年美国估计新增癌症病例数为1,638,910(不包括非黑素瘤皮肤癌), 而在美国每年死于癌症的数字估计为577,190 (<http://cancer.gov/cancertopics/cancerlibrary/what-is-cancer>)。存在对癌症进行管理和治疗的选择。其主要包括手术(例如, 肿瘤外科手术)、化疗、放疗、靶向癌症治疗(例如, 小分子药物或特异性靶向肿瘤生长和发展中涉及的分子的单克隆抗体疗法)、和姑息治疗, 或其中的一些组合(本文一起称为“抗癌治疗”)。

[0010] 额外的癌症治疗方法包括诱导、增强或抑制免疫应答的治疗策略, 统称为“免疫疗法”或“免疫治疗”(本文通常也会一般地称为“产生免疫应答的疗法”)。最近, 免疫疗法在晚期或转移性实体肿瘤治疗中已变得更相关。用于癌症的免疫疗法通常是为了增强或刺激患者自身的免疫反应, 以更好地控制或消除癌细胞, 并可能另外支持其他治疗如化疗、手术、放疗和靶向癌症疗法的使用。这种用于肿瘤的免疫疗法的一些例子包括:(1)

PROVENGE® (Dendreon) , 其中刺激树突状细胞以激活对抗原的细胞毒性反应, 用于晚期去势难治性前列腺癌; (2) T细胞过继转移以激活对癌症的细胞毒性反应; (3) 通过引入病毒来基因工程化T细胞, 所述病毒引入经设计能识别肿瘤抗原的T细胞受体; (4) Algenpantucel-L, 一种包括被转染表达小鼠 α 1,3-半乳糖基转移酶的具有潜在抗肿瘤活性的辐射同种异体胰腺癌细胞的癌症疫苗; (5) 基于病毒载体的免疫疗法; 和 (6) 基于酵母的免疫治疗。

[0011] 基于酵母的免疫治疗也被称为 TARMOGEN® (GlobeImmune, Inc., Louisville, Colorado) 技术, 通常指的是胞外(在其表面)、胞内(在内部或细胞溶胶中)或胞外和胞内同时表达一种或多种异源靶抗原的酵母载体。基于酵母的免疫治疗技术已有大致描述(参见, 例如, 美国专利号5,830,463)。一些基于酵母的免疫治疗组合物及其制备和一般应用的方法, 也详细描述于例如美国专利号5,830,463, 美国专利号7,083,787, 美国专利号7,736,642, Stubbs等人, Nat. Med. 7:625-629 (2001), Lu等人, Cancer Research 64: 5084-5088 (2004), 以及Bernstein等人, Vaccine 2008 Jan 24; 26 (4): 509-21, 其各自全文引入作为参考。基于酵母的癌症免疫治疗描述于例如美国专利号7,465,454, 美国专利号7,563,447, 美国专利号8,067,559, 美国专利号8,153,136, 美国专利公开号2009-0098154, 以及PCT公开号W0 07/133835, 其各自全文引入作为参考。

[0012] 与其它免疫疗法相比, 基于酵母的免疫治疗具有一种独特的能力, 其能诱导针对靶抗原的先天性免疫应答和大范围的适应性免疫应答, 包括依赖CD4的TH17和TH1 T细胞应答和抗原特异性CD8⁺T细胞应答, 其包括细胞毒性T淋巴细胞(CTL)应答, 而均不需要使用外源性佐剂、细胞因子或其它免疫刺激分子, 这些物质有很多具有毒性问题。此外, 基于酵母的免疫治疗组合物抑制调节性T细胞(Treg)数和/或功能, 从而增强效应T细胞应答, 后者通常可能受例如肿瘤存在的抑制。此外, 与通过产生抗体应答进行免疫的免疫治疗组合物相比, 由基于酵母的免疫疗法引发的抗原特异性、广谱性、和有效的细胞免疫应答被认为在靶向肿瘤细胞时特别有效, 即使面对的可能是抑制性的环境。由于这种类型的免疫疗法利用抗原呈递细胞的天然能力呈现相关的免疫原, 没有必要知道靶抗原的CTL表位或II类MHC表位的确切身份来产生有效的基于酵母的免疫治疗, 也不需要分离患者的任何免疫细胞来产生免疫治疗。事实上, 在一个基于酵母的免疫治疗组合物中可以靶向多个CD4⁺和CD8⁺T细胞表位, 所以使用算法和复杂的公式来确定假定的T细胞表位或T细胞受体就被淘汰了。

[0013] 已开发了一系列基于酵母的免疫治疗产品来针对由患者肿瘤表达的突变Ras蛋白刺激免疫应答, 包括GlobeImmune, Inc. 目前处于临床研发的被称为“GI-4000”的TARMOGEN®产品候选物。“Ras”这个名字是指一个发现于细胞内, 包括人类细胞的相关蛋白质家族。所有Ras蛋白家族成员属于一类称为小GTP酶的蛋白质, 其与细胞中传输信号(细胞信号转导)有关。在美国每年大约有180,000新癌症病例中发现Ras突变, 所属肿瘤的类型包括胰腺癌、非小细胞肺癌(NSCLC)、结肠直肠癌、子宫内膜癌和卵巢癌、以及黑色素瘤和多发性骨髓瘤。研究表明, Ras突变肿瘤对常规化疗和靶向药物的应答通常低于正常Ras肿瘤。对于某些癌症, 如NSCLC、结直肠癌, 针对表皮生长因子受体或EGFR的疗法已经改善了临床结果。然而, 在结直肠癌中尽管使用了EGFR靶向疗法, 肿瘤中Ras突变的存在仍然关联不良预后。同样的, 其他研究已经表明, 与那些接受同样的治疗时生存率改善的Ras正

常的患者相比,Ras突变结直肠肿瘤患者不能从另一种EGFR靶向剂——西妥昔单抗治疗中受益。结果,Ras突变患者有更少的有效治疗选择。包含Ras突变细胞的靶向减少可能导致罹患多种由于突变Ras在肿瘤生长中的作用而导致的人类癌症的患者临床结果改善。然而,目前没有可用的靶向突变Ras治疗的晚期临床试验。

[0014] 免疫治疗领域的进展缓慢,但最近的临床成功对这种癌症治疗方法的潜力给了大力支持。然而,在本领域有必要定义在治疗之前能识别将从免疫治疗中获得对癌症的临床益处的患者和识别临床应答者和非应答者的生物标志物。免疫疗法标志物的例子包括CD54表达和白介素12p70生产,但是它们尚未被完全确认。还使用了几种细胞免疫标志物分析(细胞因子流式细胞术、MHC四聚体、以及酶联免疫吸附测定(ELISPOT))。值得注意的是,预测免疫疗法收益的分析需要是标准化的,以产生可复制的和可比较的结果。这在本领域尚未完成。

[0015] 在本领域,对于在治疗前判断给定的癌症患者是否可能从单独施用或配合其它抗癌药物治疗施用的产生免疫应答的疗法中受益,或者相反的,对于该治疗是否不太可能对给定癌症患者有益处的实用的、有效的检测存在需求。本发明满足了这一需求。

[0016] 与预测癌症患者从一些类型的药物中受益的能力有关的进一步有用的现有技术包括美国专利号7,736,905,7,858,390;7,858,389,7,867,775,8,024,282;7,906,342和7,879,620,以及2012年1月24日提交的未决的美国专利申请系列号13/356,730和2011年2月22日提交的美国专利申请系列号12/932,295,其作为美国2011/0208433公开,所有这些均归属于Biodesix, Inc.。'905专利和2011年2月22日提交的美国专利申请系列号12/932,295在此引入作为参考。除其它方面,'905专利描述了用于判断NSCLC癌症患者是否可能从表皮生长因子受体(EGFR)靶向药物中受益的基于质谱的检测。该检测在其商业版本作为“VeriStrat”而为人所知;在以下讨论中对“VeriStrat”的参考将被理解为参考'905专利中描述的检测。

[0017] 发明简述

[0018] 本发明大致涉及引导癌症患者治疗方法的领域。在一方面,对癌症患者的这种治疗是对于癌症的免疫治疗,且在一方面,该治疗是基于酵母的癌症免疫治疗,且在另一方面,该治疗是针对突变Ras阳性癌症的基于酵母的免疫治疗(即,其中至少一些肿瘤对于突变Ras蛋白是阳性的,通常通过检测ras核苷酸序列突变来检测)。在一方面,该治疗是针对突变Ras阳性的胰腺癌的基于酵母的免疫治疗。

[0019] 更具体的,在一方面,本发明涉及一种在治疗启动之前预测患者是否是可能从施用基于酵母的免疫应答生成型治疗(例如,作为细胞免疫疗法试剂)中受益的一类患者中的一员的方法,所述治疗单独使用或者配合标准抗癌药物治疗和/或用于治疗癌症的其它治疗方案。还公开了鉴定那些不太可能应答基于酵母的免疫治疗和/或在标准化疗试剂中添加免疫治疗的一类患者的方法。还公开了鉴定较少可能或者不可能应答基于酵母的免疫治疗,和/或在标准化疗试剂和/或癌症的其它治疗(例如,手术切除)中添加免疫治疗的患者的方法。

[0020] 在另一方面,本发明涉及一种在治疗启动之前预测患者是否是可能从施用针对突变Ras阳性癌症的基于酵母的免疫治疗(例如,作为细胞免疫疗法试剂)中受益的一类患者中的一员的方法,所述治疗单独使用或者配合标准抗癌药物治疗和/或用于治疗癌症的其

它治疗方案。还公开了鉴定较少可能或者不可能应答针对突变Ras阳性癌症的基于酵母的免疫治疗,和/或在标准化疗试剂和/或癌症的其它治疗(例如,手术切除)中添加针对突变Ras阳性癌症的基于酵母的免疫治疗的患者的方法。在一方面,突变Ras阳性癌症是胰腺癌。作为一个例子,本文件描述了一种预测胰腺癌患者是否可能从施用靶向突变Ras的基于酵母的免疫治疗(例如,作为GI-4000为人所知的一系列产品,本文中有更详细描述)联合施用吉西他滨中受益的方法。

[0021] 本公开内容的方法使用从所述患者的血液来源的样品获得的质谱数据、经配置作为对所述质谱数据进行操作的分类器的计算机、和包括来自其他癌症患者的分类标签谱的训练集。

[0022] 申请人发现了一种在治疗之前预测患者是否是可能或不可能从施用基于酵母的免疫应答生成型治疗(例如,作为细胞免疫疗法试剂)中受益的一类患者中的一员的方法,所述治疗单独使用或者联合其它抗癌治疗使用。该方法是基于血液来源的样品的质谱。血液来源的样品(例如,血清、血浆)的使用是重要的,因为其通过对免疫系统的循环标记物进行考察而增加了检测对免疫治疗总体敏感性的可能。此外,该方法可通过对血液来源的样品进行简单的质谱检测而快速实施,而不必对患者样品进行复杂耗时的分析或从患者中获得肿瘤样品。值得注意的是,申请人已经从获得样品的预处理中演示了其检测的有效性。从而,该检测的现实应用中使用来自患者的预处理的样品并预测患者是否可能,或不可能从基于酵母的免疫应答生成型治疗中受益。

[0023] 本公开的方法表现为实用的、有效的检测的形式,其可以在质谱仪(例如,MALDI TOF设备)和经配置作为分类器起作用的通用计算机的帮助下实现。

[0024] 在一方面,一种预测癌症患者是否可能从单独施用或配合其它抗癌治疗施用的基于酵母的免疫应答生成型治疗中受益的方法被描述为包括以下步骤:

[0025] (a) 获得患者血液来源的样品;

[0026] (b) 对该样品进行质谱测定并从该样品获得质谱;

[0027] (c) 在程序计算机中,对该质谱执行一个或多个预定义的预处理步骤,在执行预处理步骤之后,在预定义的m/z范围内获取所述质谱中所选特征的累积强度值,并将累积强度值与包括来自其他癌症患者的分类标签谱的训练集进行比较并从而将质谱用分类标签进行分类。分类标签预测患者是否可能或不可能从单独进行或配合其它抗癌治疗进行的免疫应答生成型治疗中受益。例如,分类标签可采用“慢”或“快”的形式,“慢”表示患者可能受益且癌症复发或疾病发展的时间相对较慢,而“快”可以表示患者不太可能受益且癌症复发或疾病发展的时间相对较短。当然,可以使用其它相当的分类标签,例如“有益”、“无益”、“好”、“差”等等。

[0028] 在一方面,一种预测癌症患者是否可能从施用单独进行或配合其它抗癌治疗进行的针对突变Ras阳性癌症的基于酵母的免疫治疗中受益的方法被描述为包括以下步骤:

[0029] (a) 获得要单独或联合其它抗癌治疗的针对突变Ras阳性癌症的基于酵母的免疫治疗的患者血液来源的样品;

[0030] (b) 对该样品进行质谱测定并从该样品获得质谱;

[0031] (c) 在程序计算机中,对该质谱执行一个或多个预定义的预处理步骤,在执行预处理步骤之后,在预定义的m/z范围内获取所述质谱中所选特征的累积强度值,在其进行针对

突变Ras阳性癌症的基于酵母的免疫治疗之前将所述累积强度值与包含来自其他癌症患者的分类标签谱的训练集进行比较，并从而将质谱用分类标签进行分类。分类标签预测患者是否可能或不可能从单独进行或配合其它抗癌治疗进行的针对突变Ras阳性癌症的基于酵母的免疫治疗中受益。例如，分类标签可采用“慢”或“快”的形式，“慢”表示患者可能受益且癌症复发或疾病发展的时间相对较慢，而“快”可以表示患者不太可能受益且癌症复发或疾病发展的时间相对较短。同上，可以使用其它相当的分类标签，例如“有益”、“无益”、“好”、“差”等等。在本发明的一个具体实施方式中，进行该测试的癌症患者是胰腺癌患者。在这个实施方式中，针对突变Ras阳性癌症的基于酵母的免疫治疗可以表现为GI-4000(如下详述)或等价物的形式。在一方面，针对突变Ras阳性癌症的基于酵母的免疫治疗协同吉西他滨或等价物对患者进行治疗。在本发明这一实施方式的一个方面，突变Ras阳性癌症可包括但不限于胰腺癌、非小细胞肺癌(NSCLC)、结肠直肠癌、子宫内膜癌、卵巢癌、黑色素瘤和多发性骨髓瘤。

[0032] 在一方面，在前文或本文其它部分描述的上述任一预测方法被认为可用于其他癌症患者，包括例如，但不限于，非小细胞肺癌(NSCLC)患者和结肠直肠癌(CRC)患者，其单独使用或作为其它标准抗癌剂的辅助使用，因为在本公开中对于分类有用的质谱特征被认为除了其它方面之外还与细胞炎症反应调节有关，且对很大范围的肿瘤类型都有预测作用，如2011年2月22日提交的美国专利申请系列号12/932,295和之前引用的Biodesix, Inc的专利中所阐述的那样。

[0033] 在前文或本文其它部分描述的方法中使用的训练集优选采用来自其他癌症患者的分类标签谱的形式，所述患者从单独施用或联合其它抗癌治疗使用的免疫应答生成型治疗中受益或未受益。分类中使用的患者的谱中的特征(m/z 范围)可从形成训练集的患者的质谱分析中考察和选择。我们推测，在从与预测NSCLC患者从表皮生长因子受体抑制剂(EGFR-Is)中受益完全不同的环境中开发的美国专利号7,736,905中使用的和本文表3和4中所列举的一种或多种特征对本方法的使用是一个合适的特征集，因为他们可以关联宿主对肿瘤的免疫学和炎性应答(参见美国专利申请公开2011/0208433)。(用于分类的准确特征值可以例如根据在谱预处理过程中进行的谱对齐(偏移)而不同于‘905专利和表1-4中所列的列表，因为样品类型实质上不同。)与‘905专利(来自应答或不应答EGFR靶向药物的NSCLC患者的谱)中使用的训练集不同，本方法中使用的训练集使用来自单独施用或联合其它抗癌治疗使用的免疫应答生成型治疗中受益和未受益的患者样品谱。应当注意，本申请的样品类型不同于‘905专利中描述的血清和血浆。然而，如下述解释，在谱中有其他特征可被用于对形成本文给出的训练集示例的谱的考察的分类。

[0034] 在本公开的另一方面，一种治疗癌症患者的方法被描述为包括以下步骤：根据前文或本文其它部分描述的任一预测方法进行测试，如果对谱的分类标签表明患者可能从基于酵母的免疫应答生成型治疗中受益，即对该患者单独施用或联合另一种抗癌剂施用基于酵母的免疫应答生成型治疗。

[0035] 在一方面，在基于酵母的癌症免疫治疗之前、同时或之后对患者另外用一种或多种额外的抗癌治疗。在一个实施方式中，该额外的抗癌治疗包括但不限于化疗、放疗、靶向癌症治疗(例如，小分子药物或特异性靶向肿瘤生长和发展中涉及的分子的单克隆抗体疗法)、和姑息治疗，或其任意组合。

[0036] 在本公开的另一方面，描述了一种用基于酵母的癌症免疫疗法治疗癌症患者的方法，包括如下步骤：对癌症患者施用基于酵母的癌症免疫治疗，所述癌症患者是根据前文或本文其它部分描述的任何一种预测方法的测试筛选的，其中对谱的分类标签表明该患者可能从基于酵母的癌症免疫治疗中受益。在一方面，该患者在基于酵母的癌症免疫治疗之前、同时或之后被另外进行一种或多种额外的抗癌治疗。在一个实施方式中，该额外的抗癌治疗包括但不限于手术（例如，手术切除肿瘤）、化疗、放疗、靶向癌症治疗（例如、小分子药物或特异性靶向肿瘤生长和发展中涉及的分子的单克隆抗体疗法）、和姑息治疗，或其任意组合。

[0037] 在本公开的另一方面，描述了一种用针对突变Ras阳性癌症的基于酵母的免疫疗法治疗癌症患者的方法，包括如下步骤：根据前文或本文其它部分描述的任何一种预测方法进行测试，如果对谱的分类标签表明患者可能从针对突变Ras阳性癌症的基于酵母的免疫治疗中受益，则施用针对突变Ras阳性癌症的基于酵母的免疫治疗。在本发明的这一方面，该患者罹患癌症，其中至少一部分来自该患者的肿瘤细胞中被鉴定有突变Ras。在一方面，在进行针对突变Ras阳性癌症的基于酵母的免疫治疗之前、同时或之后，该患者被另外进行一种或多种额外的抗癌治疗。在一个实施方式中，针对突变Ras阳性癌症的基于酵母的免疫治疗是一系列称为GI-4000或等价物的基于酵母的免疫治疗产品中的一种产品。在本发明这一实施方式的一个方面，突变Ras阳性癌症可包括但不限于胰腺癌、非小细胞肺癌（NSCLC）、结肠直肠癌、子宫内膜癌、卵巢癌、黑色素瘤和多发性骨髓瘤。在一方面，该癌症是胰腺癌。在一个实施方式中，额外的抗癌治疗包括但不限于手术（例如，手术切除肿瘤）、化疗、放疗、靶向癌症治疗（例如，小分子药物或特异性靶向肿瘤生长和发展中涉及的分子的单克隆抗体疗法）、和姑息治疗，或其任意组合。在一方面，针对突变Ras阳性癌症的基于酵母的免疫治疗协同吉西他滨或等价物对患者进行治疗。在一个实施方式中，患者是胰腺癌患者，该治疗包括一系列称为GI-4000或等价物的基于酵母的免疫治疗产品中的一种产品（如下详述），其单独施用或联合吉西他滨或等价物施用。在一方面，癌症患者的肿瘤在用基于酵母的免疫治疗组合物治疗之前已被手术切除。

[0038] 在本公开的另一方面，描述了一种用基于酵母的癌症免疫疗法治疗癌症患者的方法，包括如下步骤：对癌症患者施用针对突变Ras阳性癌症的基于酵母的癌症免疫治疗，所述癌症患者是根据前文或本文其它部分描述的任何一种预测方法的测试筛选的，其中对谱的分类标签表明该患者可能从针对突变Ras阳性癌症的基于酵母的癌症免疫治疗中受益。在本发明的这一方面，该患者罹患癌症，其中来自该患者的肿瘤细胞中被鉴定有突变Ras。在一方面，在进行针对突变Ras阳性癌症的基于酵母的免疫治疗之前、同时或之后，该患者被另外进行一种或多种额外的抗癌治疗。在一个实施方式中，该针对突变Ras阳性癌症的基于酵母的免疫治疗是一系列称为GI-4000或等价物的基于酵母的免疫治疗产品中的一种产品。在本发明这一实施方式的一个方面，突变Ras阳性癌症可包括但不限于胰腺癌、非小细胞肺癌（NSCLC）、结肠直肠癌、子宫内膜癌、卵巢癌、黑色素瘤和多发性骨髓瘤。在一方面，该癌症是胰腺癌。在一个实施方式中，额外的抗癌治疗包括但不限于手术（例如，手术切除肿瘤）、化疗、放疗、靶向癌症治疗（例如，小分子药物或特异性靶向肿瘤生长和发展中涉及的分子的单克隆抗体疗法）、和姑息治疗，或其任意组合。在一方面，针对突变Ras阳性癌症的基于酵母的免疫治疗协同吉西他滨或等价物对患者进行治疗。在一个实施方式中，患者是

胰腺癌患者,治疗包括一系列称为GI-4000或等价物的基于酵母的免疫治疗产品中的一种产品,其单独施用或联合吉西他滨或等价物施用。在一方面,癌症患者的肿瘤在用基于酵母的免疫治疗组合物治疗之前已被手术切除。

[0039] 在使用基于酵母的免疫治疗来治疗癌症患者的方法的任何一个方面,该基于酵母的免疫治疗可包括但不限于一种完整的、热失活的重组酵母,其已经表达至少一种与患者的肿瘤相关的或由其表达的癌症抗原。在一方面,该酵母可以是来自一个属的酵母,包括但不限于,酵母属(*Saccharomyces*)。在一方面,该酵母可以是来自一个种的酵母,包括但不限于,酿酒酵母(*Saccharomyces cerevisiae*)。

[0040] 在另一方面,公开了一种预测癌症患者是否有可能从施用单独或与另一种抗癌剂联合的基于酵母的免疫应答生成型治疗中受益的系统。该系统包括从来自癌症患者的血液来源的样品产生质谱的质谱仪。该系统还包括储存有来自其他癌症患者的分类标签谱的训练集的机器可读的存储器。该训练集包括来自多个未从单独或与另一种抗癌剂联合施用的基于酵母的免疫应答生成型治疗中受益的患者的分类标签谱,和来自多个从单独或与另一种抗癌剂联合使用的细胞免疫治疗中受益的患者的分类标签谱。该系统进一步包括配置为对所述质谱数据进行操作并使用训练集对质谱进行分类、为质谱产生分类标签的计算机系统,其中分类标签被用于预测患者是否可能从施用单独或与另一种抗癌剂联合施用的基于酵母的免疫应答生成型治疗中受益。

[0041] 在另一方面,公开了一种预测癌症患者是否有可能从施用单独或与另一种抗癌治疗联合施用的基于酵母的癌症免疫治疗中受益的系统。该系统包括从来自癌症患者的血液来源的样品产生质谱的质谱仪。该系统还包括储存有来自其他癌症患者的分类标签谱的训练集的机器可读的存储器。该训练集包括来自多个未从单独或与另一种抗癌治疗联合施用的基于酵母的癌症免疫治疗中受益的患者的分类标签谱,和来自多个从单独或与另一种抗癌治疗联合施用的基于酵母的癌症免疫治疗中受益的患者的分类标签谱。该系统进一步包括配置为对所述质谱数据进行操作并使用训练集对质谱进行分类、为质谱产生分类标签的计算机系统,其中分类标签被用于预测患者是否可能从施用单独或与另一种抗癌剂联合施用的基于酵母的癌症免疫治疗中受益。

[0042] 在另一方面,公开了一种预测癌症患者是否有可能从施用单独或与另一种抗癌治疗联合使用的针对突变Ras阳性癌症的基于酵母的免疫治疗中受益的系统。该系统包括从来自癌症患者的血液来源的样品产生质谱的质谱仪。该系统还包括储存有来自其他癌症患者的分类标签谱的训练集的机器可读的存储器。该训练集包括来自多个未从单独或与另一种抗癌治疗联合使用的针对突变Ras阳性癌症的基于酵母的免疫治疗中受益的患者的分类标签谱,和来自多个从单独或与另一种抗癌治疗联合使用的针对突变Ras阳性癌症的基于酵母的免疫治疗中受益的患者的分类标签谱。该系统进一步包括配置为对所述质谱数据进行操作并使用训练集对质谱进行分类、为质谱产生分类标签的计算机系统,其中分类标签被用于预测患者是否可能从施用单独或与另一种抗癌治疗联合使用的针对突变Ras阳性癌症的基于酵母的免疫治疗中受益。在一方面,针对突变Ras阳性癌症的基于酵母的免疫治疗是被称为GI-4000的系列产品中的一种产品。在一方面,突变Ras阳性癌症选自胰腺癌、非小细胞肺癌(NSCLC)、结肠直肠癌、子宫内膜癌、卵巢癌、黑色素瘤和多发性骨髓瘤。在一方面,该癌症是胰腺癌。

[0043] 附图简述

[0044] 以下详述部分将参考附图,其通过实例的方式提供而并非限制,其中:

[0045] 图1是显示胰腺癌二期临床实验设计的示意图,其中使用靶向突变Ras阳性癌症的GlobeImmune的基于酵母的免疫治疗产品系列,称为GI-4000。

[0046] 图2A和2B是用Kaplan-Meier绘图显示的无复发生存率(RFS)和总存活数(OS),其说明了本公开的质谱方法鉴定那些很可能在胰腺癌治疗中从GI-4000和吉西他滨的组合受益的患者的鉴定能力。

[0047] 图3A-3F是在K最近邻分类算法中使用不同K值时,用Kaplan-Meier绘图成对显示的患者的RFS和OS,所述患者是GI-4000和吉西他滨的胰腺癌治疗研究中治疗组和对照组的患者。在图3A-3B中,K值为1,在图3C和3D中,K值为3,在3E和3F中,K值为5。这些图显示了在治疗组中(患者用吉西他滨和GI-4000两者治疗),申请人的分类器区分受益和未受益患者的能力,以及更进一步区分出治疗组中一些患者比对照组中一些患者表现更差的能力。因此,这些图显示了申请人的质谱方法既预测可能从产生免疫应答的疗法进行的治疗中受益的那些患者,也预测不可能从产生免疫应答的疗法进行的治疗中受益的那些患者的能力。

[0048] 图4A-4H是GI-4000和吉西他滨研究中一组RFS和OS的Kaplan-Meier图,是用对血液来源的样品以150,000次射击(shots)进行“DeepMALDI”质谱方法获得的谱的基础上的分类器定义的,其中采用本文所述和美国临时专利申请61/652,394(2012年5月29日提交,其内容在此引入作为参考)中描述的技术。图4A-4B、4C-4D、4E-4F和4G-4H各自分别代表来自四个不同分类器的RFS和OS图,其基于用于分类的质谱的不同峰的集合。

[0049] 图5是对胰腺癌/GI-4000和吉西他滨研究中使用的分类器校验(classifier validation)进行交叉验证过程的流程图。

[0050] 图6是RFS“快”和“慢”组之间即时事件比(Hazard ratio)的分布图,所述RFS来自图3E-3F的“稀释并发射(dilute-and-shoot)”分类器的交叉验证分析。

[0051] 图7A是GI-4000和吉西他滨研究的交叉验证分析的测试组中,“快”和“慢”组中值RFS的分布图。图7B是GI-4000和吉西他滨研究的交叉验证分析的测试组中,“快”和“慢”组之间中值差别的分布图。两个图都使用图3E-3F的“稀释并发射”分类器。

[0052] 图8是在图3E-3F的“稀释并发射”分类器的交叉验证分析中,对照组(Control Arm)和测试集合(Test Set)的“快”和“慢”组中值分布图。

[0053] 图9是在图3E-3F的“稀释并发射”分类器的交叉验证分析中,“慢”组的对照组和测试集合的中值差别分布图。

[0054] 图10A是在图3E-3F的“稀释并发射”分类器的交叉验证分析中使用不同K值时,对照组的慢/快分类比例分布图。图10B是在K-最近邻分类器中使用不同K值时,测试集合的即时事件比图。

[0055] 图11是检测方法的流程图,所述方法预测癌症患者从单独或联合其它抗癌剂使用的产生免疫应答的疗法中受益或不受益。

[0056] 图12是系统示意图,所述系统对血液来源的患者样品进行检测并预测患者是否可从单独或与另一种抗癌剂联合使用的免疫应答生成型治疗中受益。

[0057] 图13A-13L是一套在GI-4000和吉西他滨研究中RFS和OS的Kaplan-Meier图,由基于从使用500,000发射数(shots)的DeepMALDI方法获得的谱的分类器定义,其中使用2012

年5月29日提交的未决的美国临时专利申请系列号61/652,394中的技术,其内容在此引入作为参考。图13A-13B,13C-13D,13E-13F,13G-13H,13I-13J,和13K-13L分别代表来自6个不同分类器的RFS和OS图,其基于用来分类的质谱的不同的峰或特征的集合。

[0058] 图14是来自使用150,000发射数的DeepMALDI方法获得的谱所开发的图4A-4H的分类器的交叉验证分析的RFS的“快”和“慢”组的分布图。

[0059] 图15是来自使用500,000发射数的DeepMALDI方法获得的谱所开发的图13A-13L的分类器的交叉验证分析的RFS的“快”和“慢”组的即时事件分布图。

[0060] 图16A-16C是对选定的质量/电荷范围 (m/z 比7,000-8,000) 内同一样品的3个 MALDI质谱的描述,显示了随着增加的发射数,可检测峰内容的增加。图16A的谱来自2,000发射数,图16B的谱来自100,000发射数,图16C的谱来自500,000发射数。注意来自我们的方法的图16B和16C的谱是如何揭示了丰富的样品谱信息,其在图16A的谱中不存在,图16A的谱基本表现为噪音。

[0061] 图16D和16E是显示从我们的DeepMALDI方法中获得的谱的巨大的动态范围的进一步的质谱实例。在图16D中,一部分 m/z 范围为7140-7890Da的谱以放大的方式显示在图16D的插图中,显示了在约500,000发射数下获得的丰富的谱信息,所述谱以Y轴放大的方式在插图中显示,以显示在 m/z 9520左右的区域内额外的谱信息和峰,其在DeepMALDI方法中显示,但是在典型的~1,000发射数的谱中不可见。

[0062] 图17A是含有384个矩阵排列的样品点或“点”的MALDI-TOF靶板的平面图。这些点由列数1...24和行数A...P确定,例如,左上方的点被确定为A1。图17B是同一样品点P1的放大图,其被显示分为一个5X5的直角坐标网,其具有X/Y位置坐标和位于点中心的原点(0, 0)。直角坐标网和位置坐标被用于自动光栅扫描方法以从如本文详述的点获得来自100,000或更多发射数的谱。

[0063] 图18是储存在图17A的MALDI板的单一点中的生物样品/基质混合物的图。理想的,在该点中包含一个统一的、同质的结晶样品,如图18所示。

[0064] 图19是对用于从图18的点获得100,000或更多发射数的一种可能的光栅扫描模式的描述。该点用光栅扫描多次,例如,25次。图19中显示的每个符号集(三角、方块、X等)描述了一组个体的、离散的X/Y位置,其中该点在单一光栅扫描中被扫描(发射)。在每个位置处,该点可进行多个发射数,例如,700或800发射数。

[0065] 图20是显示图19的光栅扫描模式对图18的样品点叠加的图。

[0066] 图21是来自MALDI-TOF设备用户界面的截图,显示了对来自每个位置/光栅800激光发射数的累加谱进行加和的命令,例如,在图17B或20的光栅扫描中。

[0067] 图22是一个样品点的一部分图像,显示样品/基质混合物未以空间均匀方式结晶的区域。

[0068] 图23是来自MALDI-TOF设备用户界面的截图,显示由设备中的照相机捕捉的一部分点的图像,以及为了对这些点进行自动光栅扫描而对一组点进行的选择。

[0069] 图24是来自MALDI-TOF设备用户界面的另一幅截图,显示了各种工具,它们用于谱的评价、谱的累加、和激光移动以不同方式穿过一个点而发射。

[0070] 图25是在数据采集中用于接受或拒绝瞬态谱的评价页面的截图。

[0071] 图26是显示用于消除背景峰的排除列表的截图。

[0072] 发明详述、工作实施例和实验结果

[0073] 本文描述了用于免疫应答生成疗法的预测性测试、相关的分类器和系统，以及对由这些测试、分类器和系统鉴定的患者的治疗。

[0074] 特别的，本文描述了用于在治疗之前预测癌症患者是否有可能或不可能从单独施用或联合其它抗癌治疗施用的基于酵母的免疫应答生成型治疗中受益的方法。还描述了用于在治疗之前预测癌症患者是否有可能或不可能从针对突变Ras阳性癌症的基于酵母的免疫治疗中受益（例如，本文描述的GI-4000）的方法，包括但不限于，胰腺癌。本发明的方法基于治疗前获得的血液来源的样品（例如，血清或血浆）的质谱，以及基于由质谱揭示的样品中的蛋白质组学签名（signature）的分类。血液来源的样品的使用是重要的，因为其通过对免疫系统的循环标记物进行考察而增加了检测对免疫治疗总体敏感性的可能。此外，该方法可通过对血液来源的样品进行简单的质谱检测而快速实施，而不必对患者样品进行复杂耗时的分析或从患者中获得肿瘤样品。该测试的用处在于，如果患者被预测可能受益，则可以对患者将获得改善的表现抱有一些信心而继续治疗；而如果患者被提前预测他们不可能受益，则可以指导患者进行该患者有可能受益的其它治疗，或可以考虑其它治疗选择。

[0075] 本文还描述了用单独或与另一种抗癌剂联合使用的基于酵母的免疫应答生成疗法治疗患者癌症的方法，其中首先已经通过根据前文或本文其它部分描述的任何一种预测方法的方法或测试选择患者为可能从免疫应答生成型治疗癌症中受益（例如，该方法或测试中产生的对谱的分类标签表明患者有可能从免疫应答生成型治疗癌症中受益）。

[0076] 本文还描述了用单独或与另一种抗癌剂联合使用的针对突变Ras阳性癌症的基于酵母的免疫治疗（例如，本文描述的GI-4000）治疗患者癌症的方法，其中首先已经通过根据前文或本文其它部分描述的任何一种预测方法的方法或测试选择患者为可能从针对突变Ras阳性癌症的基于酵母的免疫治疗中受益（例如，该方法或测试中产生的对谱的分类标签表明患者有可能从针对突变Ras阳性癌症的基于酵母的免疫治疗中受益）。在本发明的这些治疗方法的任意一种中，所述方法包括向患有表达癌症抗原的受试者施用基于酵母的免疫应答生成型治疗（其可包括但不限于针对突变Ras阳性癌症的基于酵母的免疫治疗）的步骤，所述受试者已通过根据如本文描述的发明的任何一种预测方法进行的测试被鉴定或选择为可能从施用组合物中受益。

[0077] 描述GI-4000-02的工作实施例：对GI-4000和吉西他滨的胰腺癌二期临床实验

[0078] GI-4000-02是一种在R0或R1切除的胰腺癌患者中GI-4000加吉西他滨或安慰剂加吉西他滨的全登记（fully-enroll）2b期随机、双盲、安慰剂对照、多中心、辅助性临床实验（参见图1）。R0切除定义为手术切缘不存在显微残留疾病。R1切除定义为手术切缘存在显微残留疾病。R0和R1患者具有不同的期望生存率，R0患者平均生存得更久。在本临床试验中，在筛选过程中从每个受试者获得肿瘤组织样品，且对肿瘤评价了Ras突变的存在。如果受试者具有产品相关突变，则施用与在受试者的肿瘤中的具体Ras突变匹配的GI-4000基于酵母的免疫治疗产品（GI-4000系列如下详述）。

[0079] 研究群体包括在美国和5个国际中心登记的176名具有Ras突变的切除胰腺癌受试者。在切除后，将受试者根据切除状态按预期分为2组，R1和R0组都按照1:1比例随机分配至两个治疗组接受40个Y.U.的GI-4000（“Y.U.”是“酵母单位”或“酵母细胞等价物；1个Y.U.=1千万酵母细胞）加吉西他滨或安慰剂加吉西他滨。登记了39名R1受试者，其中19名分配至

GI-4000加吉西他滨组,20名分配至安慰剂加吉西他滨组。登记了137名R0受试者,其中69名分配至GI-4000加吉西他滨组,68名分配至安慰剂加吉西他滨组。对40个Y.U.剂量的GI-4000按照4支分离的10Y.U.皮下注射给药,每条胳膊和每条大腿各一支。在切除手术和吉西他滨治疗开始之间,对受试者给予3个每周剂量的GI-4000或安慰剂。所有受试者都给药多达6个月的吉西他滨周期,始于切除后6-8周。每个吉西他滨周期后给予每月剂量的GI-4000或安慰剂,以与每月的吉西他滨治疗的预定中断同步。GI-4000或安慰剂的每月给药持续至受试者退出研究、经历疾病复发或死亡。评价了一些疾病特异性基线特征,包括下述已显示对结果有影响的预后因素:(a)淋巴结状态由胰腺癌细胞显微证据的存在或不存在定义。阳性结被认为是差的预后指示物;(b)体力状态,其包括反应患者整体健康的五个等级尺度(0,1,2,3,4),其中0是最好的状态,4是最差的;(c)CA19-9,其是一种胰腺癌细胞的血液生物标志物,作为肿瘤负担的衡量物。较高的CA19-9水平与较差的临床结果有关;(d)肿瘤体积(厘米),较大的体积通常与较差的临床结果有关;和(e)肿瘤阶段,范围从I期至IV期,其定义基于包括原发瘤大小、局部入侵程度、区域淋巴结参与程度和癌症自原发瘤向全身扩散程度的标准化评分系统。

[0080] 本临床试验的主要终点是无复发存活。第二终点包括总存活数、免疫应答和疾病负担的生物标志物,例如CA19-9。迄今为止,GI-4000联合辅助吉西他滨已经表现出对Ras突变阳性的R1胰腺癌受试者的存活率临床有意义的效果,包括:(a)中值OS有2.6个月的改善(17.2个月对14.6个月);18%的相对改善;(b)对GI-4000免疫应答者中值OS有5.0个月的改善(19.6个月对14.6个月);34%的相对改善;(c)一年存活率有16%的优势(72%vs.56%);30%的相对改善;和(d)中值RFS有1个月的改善(GI-4000/吉西他滨9.6个月vs.8.5个月);13%的相对优势。此外,GI-4000在R1受试者中具有免疫原性并良好耐受:(a)GI-4000/gem方案中的7/15(47%)vs.安慰剂/gem方案中的1/12(8%)具有Ras突变特异性T细胞应答;和(b)GI-4000至今被良好耐受而未出现严重的新毒性的迹象。在开发本发明的预测性质谱方法中观察到额外的结果,如下详述。

[0081] 预测性质谱方法

[0082] 对本公开的质谱方法的一个实例结合上述胰腺癌GI-4000+吉西他滨二期临床试验中的样品的研究进行如下详述。与接受吉西他滨和安慰剂的那些患者相比,接受了GI-4000+吉西他滨组合的一些患者,但不是全部,经历了RFS和OS的实质性的改善。开发了一种分类器用于在治疗前预测一名患者是否是有可能受益的一类患者中的一员(在下述讨论中称为“慢”),或相反的不可能受益(在下述讨论中称为“快”)。图2A和2B是蛋白质组学签名阳性患者的无复发生存率(RFS)和总存活数(OS)的Kaplan-Meier图,表明他们可能从GI-4000联合吉西他滨中受益。图2A显示了对具有延迟复发(“慢”)蛋白质组学签名的受试者的治疗组(GI-4000vs.安慰剂)的RFS,而图2B对具有延迟复发蛋白质组学签名的受试者的治疗组的OS。这些图显示了本公开的质谱方法鉴定在胰腺癌治疗中可能从组合的GI-4000和吉西他滨中受益的患者的能力。显示我们的方法鉴定在治疗组中从GI-4000和吉西他滨的组合中受益或不受益的患者的能力的我们研究中进一步的Kaplan-Meier图将配合下图3和4进行讨论。

[0083] 在我们的研究中,适于产生质谱的样品可来自GI-4000-02中登记的90名患者,如上详述。样品来自血液,在这一具体案例中来自血浆,其通过预先的基于密度的分离方法从

完整血液获得。完整血液(在肝素钠玻璃管中)在GlobeImmune实验室从临床试验位置通过过夜的室温运输接受，并在收集30个小时内处理。为了分离外周血单个核细胞(PBMCs)，将血液用Dulbecco磷酸盐缓冲盐水(D-PBS; Gibco/InVitrogen目录号14190-250)约1:1稀释，在Leucosep™管(Greiner)中的Ficoll-Hypaque梯度上分层并在室温1000x g下离心10分钟。在从梯度中收获细胞前，吸走用D-PBS 1:1稀释的血浆并在-80℃下冷冻。在用于本发明的质谱方法之前，将样品解冻、分装并随后在使用前再冷冻一次。

[0084] 使用标准的稀释并发射(DNS)方法和“DeepMALDI”方法从样品中生成质谱，所述方法描述于未决的2012年5月29日提交的美国临时申请系列号61/652,394，在此引入作为参考，并在下文中进一步详述。

[0085] 我们对稀释并发射谱进行了如美国专利号7,736,905和之前引用的Biodesix, Inc.的专利文献中所描述的VeriStrat检测，但是发现该分类器未产生有用的信息。鉴定了几个VeriStrat差的样品，在任一治疗组中的VeriStrat好和差患者之间未发现显著差异。

[0086] 从而，需要定义一种具有新训练集的新分类器。我们推测VeriStrat特征可能有用，因为我们相信这些与患者针对癌症存在的免疫和炎性应答有关(参见美国专利申请公开2011/0208433)。事实上，已发现只要正确定义分类器训练集并改变谱预处理流程，如下所述，VeriStrat测试(参见下表3和4)中使用的谱的质谱特征即可被用于胰腺癌研究中的分类。对分类器设计和训练集的这一发现将在以下部分中描述。

[0087] 分类器设计

[0088] 设计能将患者分为对GI-4000治疗具有更好和更差的预后的分类器的起点是定义更好和更差的无复发生存率(RFS)患者的训练集。基于RFS次数的分布，决定将快速复发(“快”)的患者定义为那些在276天之前复发的患者，将缓慢复发(“慢”)的患者定义为在500天前没有复发事件的患者。这样给出了一个训练集，其中有20名患者在“快”组，14名患者在“慢”组，9名患者具有中间的RFS次数(注：事实上在276天之前有21名出现RFS事件的患者，但是这些患者之一的谱在开始本项目的时候遗失了，因此该名患者在一开始没有包括在“稀释并发射”和150,000发射数的DeepMALDI分析的训练集之内，而仅仅用在分类器应用到整个患者群时。对于500,000发射数的DeepMALDI分析，该患者被包括在训练集中，一名复发时间长的患者被排除在外，因为确认在治疗中该患者的血浆样品已被取走。)通过取不同的截止点来定义“慢”和“快”组，或通过至死亡的快速和缓慢的时间来定义它们，也将可能产生类似的结果。

[0089] 已经定义了“快”和“慢”组的训练集之后，使用美国专利No.7,736,905的方法预处理要被比较的质谱，包括背景剪除、部分离子流标准化和谱对齐。对稀释并发射谱和DeepMALDI谱预处理的细节不相同，虽然大致程序相似。首先估计背景并从谱中减去。将谱标准化至部分离子流。用于计算部分离子流的区域可以多种方法选择，只要其能排除谱中的最强和最可变峰。在下文给出结果的稀释并发射分类器例子中，用于部分离子流计算和标准化的区域是3kDa-11.4kDa、13kDa-15kDa和16.1kDa-30kDa，但是也可进行其它选择。例如对基于500,000发射数的DeepMALDI谱的分类器，用于部分离子流标准化的区域是4.9kDa-6.54kDa、12kDa-13.5kDa和18kDa-27kDa。估计了谱中的噪音。一旦在谱中检测到峰，则使用一系列对齐点将谱进行对齐。一系列对齐点可以通过选择在与大部分待对齐的谱共同的谱中检测到的峰的子集来编辑，或可从之前经验中已知存在于大部分谱的峰中提

前选择。在下述稀释并发射分类器的情况下,根据之前的经验选择了以下对齐点。这些是在下述m/z位置的峰:6434.5,6632.1,11686.9,12864.8,15131.1,15871.5和28102.5。值得注意的是,如果使用从样品中获得质谱数据的DeepMALDI方法(参见下文解释),谱中的其他特征也可被用于部分离子流标准化和谱对齐,可使用更适合DeepMALDI谱背景减除的预处理方法。例如,在下述基于500,000发射数DeepMALDI谱的分类器,使用了下述对齐点:3315,4153,4457,4710,4855,5289,6431,6629,6835,7561,7931,8202,8807,8912,9707,12856,13735,14031,14134,15117,15856,17366,21046,27890,28019,28067和28228。然而,应当注意但是,对齐数量和位置的其他选择对于两种谱获得方法是可能的。

[0090] 谱预处理使得其能够彼此比较,从而基于从外部考虑定义的特征,其可以用于产生分类器,或者可以比较谱的组,以确定各组间差异化表达的特征,然后可以选择这些特征的一个子集,并使用这一系列特征建立分类器。具有下述结果的一个分类器是使用从外部考虑确定的特征构建的。

[0091] 由于已确信VeriStrat分类器中使用的质谱特征(参见下表3和4)与关联宿主对肿瘤存在的应答的炎性过程有关(参见我们之前于2011年2月22日提交的美国专利申请系列号12/932,295,在此引入作为参考),且这与免疫系统对肿瘤的应答有关,因此有兴趣尝试使用具有由GI-4000治疗后患者的复发次数定义的新的参考集的谱的8个VeriStrat特征。结果见下述图3。然而,可以使用在两个参考集的组的比较过程中发现的二者不同的特征来构建其它的分类器。对于稀释并发射谱这些特征包括下述一种或更多种:

[0092] 表1

[0093]

| m/Z特征中心 | m/Z特征左侧 | m/Z特征右侧 |
|----------|----------|----------|
| 5841.168 | 5831.035 | 5851.302 |
| 6433.537 | 6427.831 | 6439.244 |
| 8765.432 | 8757.666 | 8773.197 |
| 9669.883 | 9618.946 | 9720.82 |

[0094]

| | | |
|----------|----------|----------|
| 11442.47 | 11428.45 | 11456.49 |
| 11474.93 | 11460.88 | 11488.99 |
| 11529.59 | 11518.95 | 11540.22 |
| 11696.79 | 11650.43 | 11743.15 |
| 11900.24 | 11878.99 | 11921.49 |
| 12865.48 | 12856.36 | 12874.6 |

[0095] 表1的一种或更多种的特征可组合表2,3和5中的特征用于分类器中。

[0096] 对于150,000发射数的“DeepMALDI”谱,可用于分类的特征包括下述:

[0097] 表2

[0098]

| 峰号 | m/z 中心 | m/z 左侧 | m/z 右侧 |
|----|----------|----------|----------|
| 1 | 3039.529 | 3037.053 | 3042.005 |
| 2 | 3366.615 | 3358.219 | 3375.011 |
| 3 | 3432.149 | 3412.777 | 3451.52 |
| 4 | 3473.153 | 3453.337 | 3492.969 |
| 5 | 3552.474 | 3543.957 | 3560.992 |
| 6 | 3680.091 | 3672.995 | 3687.187 |
| 7 | 3842.836 | 3836.58 | 3849.092 |
| 8 | 4203.838 | 4193.707 | 4213.969 |
| 9 | 5180.126 | 5176.524 | 5183.728 |
| 10 | 5291.206 | 5287.73 | 5294.682 |
| 11 | 5700.687 | 5689.381 | 5711.994 |
| 12 | 5843.129 | 5840 | 5846.258 |
| 13 | 5860.729 | 5851.114 | 5870.344 |
| 14 | 5866.576 | 5862.227 | 5870.926 |
| 15 | 6008.828 | 6004.271 | 6013.384 |
| 16 | 6192.594 | 6182.541 | 6202.648 |
| 17 | 6298.994 | 6294.655 | 6303.333 |
| 18 | 6873.363 | 6870.128 | 6876.598 |
| 19 | 6903.79 | 6894.561 | 6913.018 |
| 20 | 6971.019 | 6967.629 | 6974.408 |
| 21 | 6985.879 | 6982.461 | 6989.296 |
| 22 | 6995.012 | 6991.486 | 6998.538 |
| 23 | 7009.091 | 7002.267 | 7015.914 |
| 24 | 7023.426 | 7019.431 | 7027.422 |
| 25 | 7035.244 | 7031.664 | 7038.825 |
| 26 | 7045.484 | 7041.962 | 7049.005 |
| 27 | 7057.527 | 7050.655 | 7064.399 |

| | | | | |
|--------|----|----------|----------|----------|
| [0099] | 28 | 7074.719 | 7070.399 | 7079.039 |
| | 29 | 7150.539 | 7136.65 | 7164.429 |
| | 30 | 7245.544 | 7237.996 | 7253.093 |
| | 31 | 7301.145 | 7297.011 | 7305.279 |
| | 32 | 7783.506 | 7778.087 | 7788.925 |
| | 33 | 8361.169 | 8355.605 | 8366.734 |
| | 34 | 8476.978 | 8470.468 | 8483.489 |
| | 35 | 8767.557 | 8761.746 | 8773.368 |
| | 36 | 9362.525 | 9353.951 | 9371.098 |
| | 37 | 9671.94 | 9664.199 | 9679.681 |
| | 38 | 9759.032 | 9751.081 | 9766.982 |
| | 39 | 9788.134 | 9772.707 | 9803.561 |
| | 40 | 9871.152 | 9861.387 | 9880.918 |
| | 41 | 11302.64 | 11295.53 | 11309.76 |
| | 42 | 10485.61 | 10471.48 | 10499.73 |
| | 43 | 10776.72 | 10762.23 | 10791.21 |
| | 44 | 11475.86 | 11468.27 | 11483.45 |
| | 45 | 11494.79 | 11487.31 | 11502.27 |
| | 46 | 11529.86 | 11522.63 | 11537.1 |
| | 47 | 11555.31 | 11541.01 | 11569.61 |
| | 48 | 11655.18 | 11616.93 | 11693.44 |
| | 49 | 11709.95 | 11701.41 | 11718.49 |
| | 50 | 11761.94 | 11724.72 | 11799.16 |
| | 51 | 11858.02 | 11814.24 | 11901.79 |
| | 52 | 11909.8 | 11902.01 | 11917.58 |
| | 53 | 11939.01 | 11929.83 | 11948.2 |
| | 54 | 12348.32 | 12339.42 | 12357.22 |
| | 55 | 12866.6 | 12858.57 | 12874.62 |
| | 56 | 13072.62 | 13064.42 | 13080.81 |
| | 57 | 13090.91 | 13082.14 | 13099.68 |
| | 58 | 13360.65 | 13351.7 | 13369.59 |
| | 59 | 13807.33 | 13786.47 | 13828.18 |
| | 60 | 13913.92 | 13897.53 | 13930.32 |
| | 61 | 14043.98 | 14035.87 | 14052.09 |
| | 62 | 14092.28 | 14084.11 | 14100.46 |
| | 63 | 14125.28 | 14117.13 | 14133.43 |
| | 64 | 14148.15 | 14139.99 | 14156.32 |
| | 65 | 14197.84 | 14181.99 | 14213.69 |
| | 66 | 14258.28 | 14249.88 | 14266.69 |

| | | | | |
|--------|----|----------|----------|----------|
| [0100] | 67 | 14429.84 | 14414.32 | 14445.35 |
| | 68 | 14384.69 | 14357.6 | 14411.78 |
| | 69 | 14530.4 | 14520.1 | 14540.7 |
| | 70 | 18801.94 | 18784.71 | 18819.18 |
| | 71 | 18861.88 | 18844.22 | 18879.55 |
| | 72 | 18904.03 | 18884.67 | 18923.4 |
| | 73 | 19860.06 | 19771.09 | 19949.03 |
| | 74 | 21710.89 | 21683.49 | 21738.29 |
| | 75 | 22998.95 | 22839.61 | 23158.29 |
| | 76 | 28314 | 28282.63 | 28345.37 |
| | 77 | 28518.76 | 28486.77 | 28550.75 |

[0101] 对于500,000发射数的“DeepMALDI”谱,可用于分类的特征包括下述:

[0102] 表5

[0103]

| | m/z中心 | m/z左侧 | m/z右侧 |
|----|----------|----------|----------|
| 1 | 3463.350 | 3459.400 | 3467.300 |
| 2 | 3609.960 | 3605.630 | 3614.280 |
| 3 | 3679.430 | 3674.080 | 3684.770 |
| 4 | 3840.800 | 3834.810 | 3846.780 |
| 5 | 3892.750 | 3889.050 | 3896.440 |
| 6 | 4077.700 | 4072.100 | 4083.300 |
| 7 | 4492.020 | 4485.400 | 4498.640 |
| 8 | 4597.040 | 4592.710 | 4601.370 |
| 9 | 5405.660 | 5400.820 | 5410.500 |
| 10 | 5555.430 | 5549.440 | 5561.410 |
| 11 | 5582.800 | 5578.980 | 5586.620 |
| 12 | 5635.590 | 5631.640 | 5639.530 |
| 13 | 5706.610 | 5701.520 | 5711.700 |
| 14 | 5762.000 | 5758.050 | 5765.940 |
| 15 | 5840.660 | 5836.580 | 5844.730 |
| 16 | 5873.770 | 5870.200 | 5877.330 |
| 17 | 5890.570 | 5885.730 | 5895.410 |
| 18 | 6107.260 | 6102.290 | 6112.220 |
| 19 | 6129.670 | 6125.720 | 6133.610 |
| 20 | 6328.570 | 6323.730 | 6333.410 |
| 21 | 6892.580 | 6888.250 | 6896.900 |
| 22 | 6951.990 | 6948.040 | 6955.930 |
| 23 | 6981.660 | 6977.580 | 6985.730 |

[0104]

| | | | |
|----|-----------|-----------|-----------|
| 24 | 7004.170 | 6999.330 | 7009.000 |
| 25 | 7054.210 | 7048.220 | 7060.190 |
| 26 | 7153.620 | 7148.910 | 7158.330 |
| 27 | 7295.310 | 7289.320 | 7301.290 |
| 28 | 7421.920 | 7417.970 | 7425.870 |
| 29 | 8357.210 | 8351.730 | 8362.680 |
| 30 | 8762.290 | 8756.430 | 8768.140 |
| 31 | 8992.440 | 8987.350 | 8997.530 |
| 32 | 9199.780 | 9193.160 | 9206.400 |
| 33 | 9794.720 | 9791.280 | 9798.160 |
| 34 | 10000.020 | 9993.530 | 10006.510 |
| 35 | 10018.610 | 10013.260 | 10023.960 |
| 36 | 10090.630 | 10083.880 | 10097.370 |
| 37 | 10174.760 | 10168.520 | 10181.000 |
| 38 | 10200.610 | 10191.950 | 10209.270 |
| 39 | 10657.440 | 10651.960 | 10662.910 |
| 40 | 10713.840 | 10704.160 | 10723.510 |
| 41 | 10912.750 | 10905.490 | 10920.000 |
| 42 | 11402.140 | 11396.410 | 11407.870 |
| 43 | 11432.410 | 11425.280 | 11439.540 |
| 44 | 11466.160 | 11459.660 | 11472.650 |
| 45 | 11488.190 | 11481.820 | 11494.550 |
| 46 | 11520.580 | 11514.720 | 11526.430 |
| 47 | 11543.880 | 11538.400 | 11549.350 |
| 48 | 11563.080 | 11556.070 | 11570.080 |
| 49 | 11620.250 | 11613.120 | 11627.380 |
| 50 | 11676.220 | 11670.870 | 11681.570 |
| 51 | 11699.260 | 11694.040 | 11704.480 |
| 52 | 11723.200 | 11716.450 | 11729.950 |
| 53 | 11744.210 | 11739.370 | 11749.050 |
| 54 | 11775.740 | 11767.330 | 11784.140 |
| 55 | 11821.600 | 11814.600 | 11828.600 |
| 56 | 11839.680 | 11833.950 | 11845.410 |
| 57 | 11882.560 | 11877.850 | 11887.270 |
| 58 | 11900.770 | 11894.660 | 11906.880 |
| 59 | 12401.800 | 12392.760 | 12410.840 |
| 60 | 12975.740 | 12968.730 | 12982.740 |
| 61 | 13104.010 | 13098.150 | 13109.860 |

| | | | |
|--------|-----------|-----------|-----------|
| 62 | 13145.130 | 13138.380 | 13151.880 |
| [0105] | | | |
| 63 | 13399.430 | 13388.090 | 13410.760 |
| 64 | 13475.490 | 13468.100 | 13482.870 |
| 65 | 13622.780 | 13616.670 | 13628.890 |
| 66 | 13648.070 | 13642.590 | 13653.540 |
| 67 | 13677.100 | 13670.860 | 13683.340 |
| 68 | 13735.490 | 13731.670 | 13739.310 |
| 69 | 13781.970 | 13774.960 | 13788.970 |
| 70 | 13800.220 | 13794.870 | 13805.570 |
| 71 | 14007.100 | 14002.010 | 14012.190 |
| 72 | 14030.530 | 14023.400 | 14037.660 |
| 73 | 14057.140 | 14051.410 | 14062.870 |
| 74 | 14078.920 | 14073.060 | 14084.770 |
| 75 | 14096.180 | 14091.340 | 14101.020 |
| 76 | 14112.220 | 14107.890 | 14116.550 |
| 77 | 14137.690 | 14129.540 | 14145.840 |
| 78 | 14186.020 | 14178.130 | 14193.910 |
| 79 | 14244.340 | 14232.110 | 14256.560 |
| 80 | 14280.730 | 14272.960 | 14288.490 |
| 81 | 14408.640 | 14403.290 | 14413.980 |
| 82 | 14423.030 | 14418.310 | 14427.740 |
| 83 | 14436.400 | 14431.050 | 14441.740 |
| 84 | 14518.120 | 14508.310 | 14527.920 |
| 85 | 14537.850 | 14531.740 | 14543.960 |
| 86 | 14653.490 | 14645.820 | 14661.150 |
| 87 | 14715.700 | 14706.780 | 14724.610 |
| 88 | 15618.160 | 15597.660 | 15638.660 |
| 89 | 17443.590 | 17430.700 | 17456.470 |
| 90 | 18734.930 | 18724.840 | 18745.010 |
| 91 | 21675.550 | 21649.880 | 21701.220 |
| 92 | 22987.930 | 22971.630 | 23004.220 |
| 93 | 23020.930 | 23009.110 | 23032.740 |
| 94 | 28038.090 | 27994.740 | 28081.430 |
| 95 | 28231.950 | 28203.250 | 28260.650 |
| 96 | 28438.910 | 28393.220 | 28484.590 |
| 97 | 28804.560 | 28750.090 | 28859.030 |

[0106] 注意谱标准化的改进可揭示进一步的区别峰,从而上述列表并非被认为是穷尽的。同样,在预处理中精确的m/z位置可根据谱对齐进行轻微偏移。

[0107] A. 基于稀释并发射谱的分类器

[0108] 图3A-3F的图显示了对参考集使用“慢”和“快”定义(“慢”=在500天之前没有复发事件,“快”=在276天前复发)、VeriStrat特征定义(表3和4)、和稀释并发射预处理谱建立的分类器的表现。将分类器用于GI-4000(治疗)方案的谱和来自安慰剂(对照)方案的46个谱。分类器基于K-最近邻分类算法(参见美国专利7,736,905)对谱产生了快或慢的分类标签,图3A-3B显示了K=1的分类,图3C-3D显示了K=3的分类,图3E和3F显示了K=5的分类。

[0109] 从图3A-3F中显示的这些结果,我们发现有可能将治疗组(GI-4000+吉西他滨)分为两个组,“快”和“慢”,其中“快”组在RFS和OS两个方面较之“慢”组具有显著更差的表现。相反,对照组(吉西他滨+安慰剂)在“快”和“慢”组中具有相似的RFS。在“慢”组中支持GI-4000有RFS的治疗益处。OS中治疗效果的情况难于破解,可能是在“慢”组中来自添加GI-4000的治疗益处被复发后接受的治疗所稀释了。OS的分析还受到患者群中38%的事件审查而变得复杂。

[0110] 注意‘快’治疗组低于对照组,表明‘快’患者未从GI-4000和吉西他滨的治疗中受益。因此,分类器提供了预测那些不可能从刺激免疫应答治疗中受益的患者的能力。

[0111] B. 150,000发射数的基于DeepMALDI的分类器

[0112] 图4A和4B中显示了基于预处理DeepMALDI谱(参见下文详述)和选自如上定义的训练集组“快”和“慢”比较的特征的分类器的表现,图4A显示了分类器在治疗组中但不是在对照组中将快和慢患者按照RFS分离的能力,图4B显示了分类器在治疗组中但不是在对照组中将快和慢患者按照OS分离的能力。分类器中使用的特征被选择为接近表4中的8个特征。对于RFS和OS两者均在治疗组中观察到对“快”和“慢”组的显著分离,但不是在对照组中。虽然没有显著差异,但是“慢”组对GI-4000治疗较之安慰剂的对照组显示了更好结果的趋势,特别是RFS。

[0113] 图4C-4D、4E-4F和4G-4H是使用上表2所列出的用于DeepMALDI的77个特征的样品和子集的DeepMALDI质谱的3个另外的分类器的RFS和OS的Kaplan-Meier图。所有3个分类器使用了来自复发事件在276天之前(“快”)的患者的20个谱和来自在500天前没有复发事件或审查(“慢”)的患者的14个谱的参考集、相同的为DeepMALDI谱优化的预处理、以及K-最近邻分类算法中K=5。

[0114] 图4C和4D的分类器使用了来自上表2的77名候选特征的特征子集。我们将该77个特征通过p值进行分类,并使用了具有描述其相对表达差异的最低p值的20个特征:

[0115]

| m/Z特征中心 | m/Z特征左侧 | m/Z特征右侧 |
|-----------|-----------|-----------|
| 3842.836 | 3836.5801 | 3849.0918 |
| 5860.7288 | 5851.1136 | 5870.3441 |
| 6903.7895 | 6894.5605 | 6913.0184 |
| 7023.4264 | 7019.4307 | 7027.422 |
| 7074.7189 | 7070.3989 | 7079.0389 |
| 7301.1449 | 7297.0108 | 7305.279 |
| 9362.5245 | 9353.9508 | 9371.0983 |
| 9671.9403 | 9664.1994 | 9679.6812 |

| | | |
|------------|------------|------------|
| 9759.0315 | 9751.0808 | 9766.9822 |
| 10776.7214 | 10762.229 | 10791.2138 |
| 11709.9483 | 11701.4091 | 11718.4875 |
| 11761.9383 | 11724.7165 | 11799.1602 |
| 11858.0166 | 11814.2429 | 11901.7903 |
| 13090.909 | 13082.1414 | 13099.6766 |
| 14043.9806 | 14035.8701 | 14052.091 |
| 14125.2798 | 14117.1271 | 14133.4325 |
| 14148.1548 | 14139.9896 | 14156.3199 |
| 14197.8412 | 14181.9925 | 14213.6898 |
| 14258.2834 | 14249.8779 | 14266.689 |
| 18801.9437 | 18784.7057 | 18819.1817 |

[0116] 图4E和4F显示了第三DeepMALDI分类器的RFS和OS的Kaplan-Meier图。在这一实例中,我们使用了来自谱(表2)列表的特征,其位于VeriStrat特征(表3和4)的区域内或与其密切相关。

[0117]

| m/Z特征中心 | m/Z特征左侧 | m/Z特征右侧 |
|------------|------------|------------|
| 5843.1286 | 5839.9995 | 5846.2576 |
| 5860.7288 | 5851.1136 | 5870.3441 |
| 11475.8583 | 11468.2665 | 11483.4501 |

[0118]

| | | |
|------------|------------|------------|
| 11494.7897 | 11487.3141 | 11502.2653 |
| 11529.8615 | 11522.6267 | 11537.0962 |
| 11555.3129 | 11541.0141 | 11569.6118 |
| 11655.1818 | 11616.9251 | 11693.4385 |
| 11709.9483 | 11701.4091 | 11718.4875 |
| 11761.9383 | 11724.7165 | 11799.1602 |
| 11858.0166 | 11814.2429 | 11901.7903 |
| 11909.7985 | 11902.0132 | 11917.5837 |
| 11939.0138 | 11929.8302 | 11948.1974 |
| 22998.948 | 22839.6075 | 23158.2885 |

[0119] 图4G-4H中显示了第四DeepMALDI分类器的结果。对于这一分类器,我们使用了(表2)特征的子集,其组间表达水平以与图4E-4F的分类器相反的方式相关,且其与VeriStrat特征不相关。

[0120]

| m/Z特征中心 | m/Z特征左侧 | m/Z特征右侧 |
|-----------|-----------|-----------|
| 7009.0908 | 7002.2674 | 7015.9142 |
| 7023.4264 | 7019.4307 | 7027.422 |
| 7035.244 | 7031.6636 | 7038.8245 |

| | | |
|------------|------------|------------|
| 7074.7189 | 7070.3989 | 7079.0389 |
| 14043.9806 | 14035.8701 | 14052.091 |
| 14092.2825 | 14084.1062 | 14100.4588 |
| 14125.2798 | 14117.1271 | 14133.4325 |
| 14148.1548 | 14139.9896 | 14156.3199 |
| 14197.8412 | 14181.9925 | 14213.6898 |

[0121] 注意图4A-4H的DeepMALDI分类器在治疗组中清楚地分离了快和慢患者而在对照组中显示了很少的或不能分离,从而与图3A-3F的“稀释和发射”分类器表现相似。

[0122] 在解读图3A-3F和4A-4H的结果中,应当注意上述结果趋向于高估了对“快”和“慢”组的分离,因为在分析中使用了分类器的参考集。不幸的是,还不存在用来以独立性方式检测分类器表现的校验集。从而,为了提高可选的分类器表现评估,如下“分类器的交叉验证”节所述进行了交叉验证分析。

[0123] C. 500,000发射数的基于DeepMALDI谱的分类器

[0124] 能够在治疗组中根据RFS良好分离“快”和“慢”患者、但不能在对照组中分离的分类器还可以通过使用500,000发射数的DeepMALDI谱进行构建。对该DeepMALDI谱进行与本节呈现的所有分类器相同的预处理,对这些分类器的训练集仍然是如前定义的“快”和“慢”复发组。在这一DeepMALDI分析的时候能够获得更新的存活数据,从而这些分类器的表现分析使用了较之A和B节中呈现的那些而言更新的数据。在各分类器中使用的区别特征集合是上表5中97个特征的子集。对于每个分类器优化了为K-最近邻分类算法选择的K近邻。图13A-13L是RFS和OS的Kaplan-Meier图,其显示了使用500,000发射数的DeepMALDI谱和表5中列出的97个特征的子集建立的6个分类器的表现。

[0125] 图13A和13B的分类器使用了表5中列举的42个特征的子集。选择它们以包括VeriStrat特征(表3和4)的m/z区域中所含的特征和基于对训练集中‘慢’和‘快’组之间差异的低单变量p值所选择的额外的特征。对于这一分类器,发现K=3是最优的,所使用特征的中心(m/z)列于表6的第一列。

[0126] 表6

[0127]

| 图13A-13B | 图13C-13D | 图13E-13F | 图13G-13H | 图13I-13J | 图13K-13L |
|----------|----------|----------|----------|----------|----------|
| | 4492 | | | 4492 | |
| | | | | | 5636 |
| | 5762 | | | | |
| 5841 | 5841 | | | | |
| | | | | | 5891 |
| | 6893 | 6893 | | 6893 | |
| | 7004 | | | 7004 | |
| 7295 | | 7295 | | | |
| 7422 | 7422 | | | 7422 | |
| 8357 | | 8357 | | | 8357 |
| | 8762 | 8762 | 8762 | 8762 | |

| | | | | | |
|-------|-------|-------|-------|-------|-------|
| | 8992 | | | 8992 | |
| 9795 | | | | | 9795 |
| 10000 | | 10000 | | | |
| | 10175 | | | 10175 | |
| 11402 | | | | | |
| 11432 | | | | | |
| 11466 | 11466 | 11466 | 11466 | | 11466 |

[0128]

| | | | | | |
|-------|-------|-------|-------|-------|-------|
| 11488 | | | | | |
| 11521 | 11521 | 11521 | 11521 | | |
| 11544 | | | | | |
| 11563 | | | | | |
| 11620 | 11620 | 11620 | 11620 | | |
| 11676 | 11676 | | | | |
| 11699 | 11699 | | | | |
| 11723 | 11723 | 11723 | | | |
| 11744 | | | | | |
| 11776 | | | | | |
| | | | | | 11822 |
| 11883 | 11883 | | | | |
| 11901 | | | | | |
| | 12976 | | | | |
| | | | | | 13104 |
| 13145 | | 13145 | | | |
| 13623 | | 13623 | | | |
| 13648 | 13648 | 13648 | 13648 | 13648 | 13648 |
| | 13677 | | | 13677 | |
| 13735 | 13735 | 13735 | 13735 | 13735 | |
| 13782 | 13782 | 13782 | 13782 | 13782 | |
| | 14007 | | | 14007 | |
| 14057 | 14057 | 14057 | 14057 | 14057 | |
| 14096 | | 14096 | | | |
| 14112 | | 14112 | | | |
| 14138 | | 14138 | | | |
| 14244 | | 14244 | | | |
| 14281 | | 14281 | | | |
| 14423 | 14423 | 14423 | 14423 | 14423 | |
| 14436 | 14436 | | | 14436 | |
| 14538 | 14538 | 14538 | 14538 | 14538 | |

| | | | | | |
|-------|-------|-------|-------|-------|--|
| | 14716 | | | 14716 | |
| 17444 | | | | | |
| 18735 | | | | | |
| 22988 | 22988 | 22988 | 22988 | 22988 | |
| 23021 | | | | | |
| 28038 | 28038 | 28038 | 28038 | 28038 | |

[0129]

| | | | | | |
|-------|-------|-------|-------|-------|--|
| 28232 | 28232 | 28232 | 28232 | 28232 | |
| | 28439 | | | 28439 | |
| 28805 | 28805 | | | 28805 | |

[0130] 图13C和13D的分类器使用表5所列的32个特征的子集(给出了特征m/z中心),其选择基于对训练集中'慢'和'快'组之间差异的单变量p值和'慢'和'快'组间特征值的振幅比例。这一分类器使用K=3和表6第2列所列特征。虽然本分类器和前一分类器有19个特征相同,但是本分类器还包含前一分类器中未使用的13个特征,而在Kaplan-Meier图方面给出了相似的表现。

[0131] 图13E和13F的分类器使用与前述分类器相似标准进行了开发,使用K=5。所使用的特征列于表6第3列。该分类器的25个特征中有23个与图13A和13B的分类器相同,有一半多与图13C和13D的分类器相同。虽然有分类器特征选择的相似性,但是Kaplan-Meier图中的表现显示本分类器的结果更适合预后(prognostic)而不是GI-4000的治疗效果预测。

[0132] 图13G和13H的分类器使用了图13E和13F的分类器的13个特征的子集,列于表6第4列,且K=5;Kaplan-Meier图显示与前两种DeepMALDI分类器更相似的表现(图13A-D)。

[0133] 图13I和13J的分类器的构建没有使用表3的VeriStrat特征的任何m/z区域中的特征。然而,可以看出其在Kaplan-Meier图方面与那些包括来自VeriStrat特征的m/z区域的特征的分类器具有相似表现。所使用的特征列于表6的第5列,且本分类器的K=7。

[0134] 图13K和13L的分类器只使用8个特征,列于表6的最后一列,且K=3。8个特征中的4个在任何其他分类器中都未使用。由Kaplan-Meier图评价的表现与测试的其它分类器仍然没有显著差异。

[0135] 注意这些DeepMALDI分类器的大部分还清楚地在治疗组中分离了'快'和'慢'分类的患者,而在对照组中表现出很少的分离,从而与"稀释并发射"分类器和那些基于具有较少发射数DeepMALDI的分类器表现相似。

分类器的交叉验证

A. 基本构架及其在"稀释并发射"分类器中的应用

[0138] 图3A-3F中所示和如上所述的"稀释和发射"分类器的交叉验证通过根据图5所概括的流程进行。

[0139] 在图5的流程中,在步骤100固定了用于分类和预处理步骤(背景减除、标准化和对齐)的特征(峰或m/z范围)。然后,以重复方式按循环108所示执行步骤102、104、105和106。在步骤102,随机选择省去10个谱用于检测分类器表现。在步骤104,使用相同的复发时间(TTR)标准从剩余的34个谱中选择谱的参考集,即将"快"定义为在276天前复发,"慢"如上定义(500天前没有复发)。在步骤105,选择了K最近邻分类算法中的K的值。在步骤106,依据

谱的“测试集”的风险比 (HR) 和中值评价了分类器的表现。谱的测试集是不包含在参考集中的治疗组中的谱 (步骤104), 即在步骤102中省去的10个谱加上来自TTR在276和500天之间的患者的任何其它的谱。该过程 (108) 重复多次 (在本实例中为70次)。

[0140] 应理解图5的流程是具有普适性的, 只要恰当选择参考和测试集即可给出要考虑的特定研究。

[0141] 图5的这一交叉验证趋向于提供分类器表现的下限, 因为其在以下方面低估了表现:

[0142] 1. 交叉验证分类器的参考集小于原始分类器的那些, 其可能影响表现。

[0143] 2. 不包括在参考集中、但是大于之前的10个中间复发谱的样品的测试集仍然较小。这可能导致计算统计中的较大变化性, 且在一些情况下, 当组体积非常小的时候, 这些统计可能是无意义的。

[0144] 3. 测试集不能代表整个治疗组群。其包含较高比例的具有中间复发时间的患者。这一不平衡使得难以比较测试集结果和测试集外的其它组 (例如对照组)。

[0145] 虽然有这些限制, 但是对使用“快”和“慢”组和VeriStrat特征的定义构建的分类器的交叉验证分析产生了一些有用的观察。图6显示了在对于对照组、完整治疗组和测试集 (排除了分类器参考集的治疗组) 的交叉验证分析中为70个实践 (realizations) 计算的“快”和“慢”组之间的风险比。虽然所有分类器都产生了对于对照组接近1的HR, 但是对于测试集的中值HR为3.1, 接近对于完整治疗组的中值HR。

[0146] 图7A和7B显示了交叉验证分析中测试集中的“快”和“慢”组中值RFS的分布。在图7A中, “慢”组的中值以约382天为中心, 而“快”组的中值以约274天为中心。图7B显示了测试集中的“快”和“慢”组之间中值差别的分布。中值的差别以约111天为中心, 这是一个临床有意义的差别, 非常少的实践显示“慢”组的中值小于“快”组。虽然测试集与对照组的比较被两组之间复发时间分布的不平衡变得复杂化, 但是对于对照组和测试集中“快”和“慢”组的中值的观察显示“慢”测试集的中值显著位于“快”和“慢”对照组的中值之上, 后者反过来位于“快”测试集的中值之上。参见图8, 其是交叉验证分析中对照组和测试集的“快”和“慢”组的中值分布图。

[0147] 此外, “慢”组的测试集和对照组之间中值差异的分布显示在几乎所有的实践中, 测试集的中值RFS都大于对照组的, 虽然存在群体的不平衡, 参见图9, 其是“慢”组的测试集和对照组之间中值差异的分布图。中值RFS的中值差异约为60天, 同样是一个有意义的临床差异。

[0148] OS的交叉验证分析受到初始可获得的临床数据中超过三分之一的整体群组有受审查的数据的阻碍。

[0149] 交叉验证分析的另一结果是确定慢与快分类比例对在K-最近邻分类器中使用的K的选择的依赖性。参见图10A, 其是分类器中使用的不同K值在对照组中慢对快分类比例的分布图。图10B是对不同K值的交叉验证分析的测试集的计算风险比图。当我们进行交叉验证时, 我们对70个重复的每一个选择了K为3、5或7 (图5, 步骤105)。我们每一种约有1/3, 从而K=3、5或7的每个分布至少有20值。为了考察K=1时发生了什么, 我们还使用K=1重新运行了19个重复的交叉验证。发现K=5的测试集中慢/快风险比最大。图10B还显示了交叉验证分析的所有70个重复的对照组获得的风险比。图10B表明对照组中使用的K值没有造成很

大的不同，且分布非常窄。当与图10A一起考虑时，图10B表明K-最近邻分类算法中K的多重选择是可能的，但是K=5可能是一个优选的选择。

[0150] B. 150,000发射数的基于DeepMALDI谱的分类器的交叉验证

[0151] 对前述4个150,000发射数的基于DeepMALDI谱的分类器进行了相同的交叉验证方法，显示其中2个在预测从添加GI-4000得到的相对治疗受益的能力方面具有出众的表现。图14显示了交叉验证的对照组和测试集中来自4A-4H的4个分类器的RFS分类为“快”和“慢”的患者的风险比分布。虽然对来自图4C-4F的分类器观察到的中值风险比与测试集和对照组中相似，但是对另两个分类器观察到的测试集中的中值风险比大于对照组中的，支持了Kaplan-Meier图中的表现评价，其显示在治疗组中RFS在‘快’和‘慢’组之间的分离大于对照组，并支持了分类器对添加GI-4000至吉西他滨对照方案的预测能力。

[0152] C. 500,000发射数的基于DeepMALDI谱的分类器的交叉验证

[0153] 还对进行了前述6个500,000发射数的基于DeepMALDI谱的分类器相同的交叉验证方法，并在图13A-M中进行了评价。图15显示了交叉验证的对照组和测试集中每个分类器的风险比分布。这显示当通过交叉验证方法评价表现时，展示完整的治疗和对照组的Kaplan-Meier图中的显著差异并不总是维持。此外，其是即使在交叉验证中也可能使用不同的特征集(其具有相似的表现特征)构建多个不同分类器的进一步证据。

[0154] 注意训练集可以具有不相等的“快”和“慢”分类的成员数量。为了处理不相等组大小的问题，我们选择的相关参考组大小部分为我们在治疗组中具有的复发时间的分布结果。恰好有很多复发时间在250和275天之间的患者，可能是由于其与计划的MRI/CT评价相一致。从而，看上去不存在合适的位置以将所述组在那些时间中分开。然而，我们决定将一个更大的早期复发组分开更好，从而我们优选在该组中有更多。K-最近邻分类算法可被调整以考虑参考集中不同的组大小，从而原则上在训练集中有不同的组大小不是问题。

[0155] 实际的、有用的测试

[0156] 正如本公开中一直指出的，在本公开的发现后进行了实际的有用的测试。一个方面是本发明的测试方法鉴定了特定的癌症患者是否是可能或不可能从施用单独或配合其它治疗的基于酵母的免疫应答生成型治疗中受益的一组癌症患者中的一员。另一方面是本发明的测试方法鉴定了特定的癌症患者是否是可能或不可能从施用单独或配合其它治疗的针对突变Ras阳性癌症的基于酵母的免疫治疗(例如GI-4000)中受益的一组癌症患者中的一员。这种鉴定可以在治疗前进行。

[0157] 在一个实例中该方法包括以下步骤：a) 从患者获得血液来源的样品；b) 在质谱仪的辅助下获得该血液来源的样品的质谱；c) 在程序计算机中，对该质谱执行预定义的预处理步骤，在执行预处理步骤之后，在预定义的m/z范围内获取所述质谱中所选特征的累积强度值，并将累积强度值与包括来自其他癌症患者的分类标签谱的训练集进行比较并用分类标签对质谱进行分级。对谱分配的分类标签被用于预测患者是否有可能或不可能以施用单独或配合其它抗癌治疗的基于酵母的免疫应答生成型治疗的形式从治疗中受益。

[0158] 在另一实例中该方法包括以下步骤：a) 从患者获得血液来源的样品；b) 在质谱仪的辅助下获得该血液来源的样品的质谱；c) 在程序计算机中，对该质谱执行预定义的预处理步骤，在执行预处理步骤之后，在预定义的m/z范围内获取所述质谱中所选特征的累积强度值，并将累积强度值与包括来自其他癌症患者的分类标签谱的训练集进行比较并用分类

标签对质谱进行分级。对谱分配的分类标签被用于预测患者是否有可能或不可能以施用单独或配合其它抗癌治疗的针对突变Ras阳性癌症的基于酵母的免疫治疗的形式从治疗中受益。

[0159] 该测试在图11中作为过程300以流程图的形式显示。

[0160] 在步骤302中,从患者获得血液来源的样品。在本实施例中所述样品是在对样品进行一些处理步骤后的血浆(例如,以预先的密度分离方法从完整血液获得的血浆)。在一个实施方式中,将该血液来源的样品分成三等份,且质谱分析法和随后的步骤304、306(包括子步骤308、310和312)、314、316和318对每一等份独立地进行。该等份的数量可以变化,例如可以有4、5或10等份,并且每一等份都进行随后的处理步骤。

[0161] 在步骤304,对该样品(等份)进行质谱分析。优选的质谱分析方法是基质辅助激光解吸电离(MALDI)飞行时间(TOF)质谱分析,但是其他方法也是可行的,包括2012年5月29日提交的未决美国专利临时申请61/652,394(其内容在此引入作为参考)中公开的所述“DeepMALDI”质谱方法(参见下文详述)。如本领域常规的,质谱分析产生代表大量质/荷(m/z)值处的强度值的数据点。在一个示例实施例中,将样品解冻,并在4摄氏度下以1500rpm对样品离心5分钟。进一步的,可以在超纯水(MilliQ water)中以1:10或1:5稀释该样品。可以将稀释后的样品一式三份点样(spot)在MALDI板上分配的位置(即在三个不同的MALDI目标或“点”上,如本领域公知的)。在将0.75ul稀释后的样品点样到MALDI板之后,加入0.75ul的35mg/ml芥子酸(用50%乙腈和0.1%的三氟乙酸(TFA)配制),然后通过移液管吹吸混合五次。可以允许在室温下干燥这些板。应当理解,可以根据本发明的原理使用其他技术和过程来制备和处理样品。

[0162] 可以使用自动或手动采集谱的Voyager DE-PRO或DE-STR MALDI TOF质谱仪以线性模式获得阳离子的质谱(当然,可以使用其它MALDI TOF设备,例如,Bruker Corporation的设备)。从每个MALDI点内7或5个位置上收集75或100个谱,以针对每份样品产生平均2,000个谱。使用蛋白质标准品(胰岛素(牛)、硫氧还蛋白(大肠杆菌)以及脱辅基红蛋白(Apomyoglobin)(马))的混合物来从外部校正该谱。

[0163] 注意DeepMALDI方法可被用于步骤304的一个MALDI板点或几个MALDI板点,参见下文详述。

[0164] 在步骤306,对步骤304中获取的谱进行一个或多个预定义的预处理步骤。使用在步骤304中获得的质谱数据上运行的软件指令来在通用计算机中实施预处理步骤306。该预处理步骤306包括背景减除(步骤308)、标准化(步骤310)以及对齐(步骤312)。背景减除的步骤优选地涉及在谱中生成强的非对称背景估计,并从该谱中减去该背景。步骤308使用在US7,736,905中描述的背景减除技术,其在此引入作为参考。标准化步骤310涉及背景减除后的谱的标准化。如美国专利7,736,905中描述的那样,该标准化可以采用部分离子流标准化的形式,或者总离子流标准化的形式。如US7,736,905中描述的那样,步骤312将标准化后的背景减除谱与预定义的质量标度对齐,可以从对分类器所使用的训练集中的谱的考察来获取该预定义的质量标度。预处理步骤还在上述GI-4000+吉西他滨临床研究的讨论中有一些详细描述。然而,预处理的细节,例如,用于部分离子流标准化和对齐的特征或谱区域,可以改变。

[0165] 一旦执行了预处理步骤306,则该处理300前进至获取预定义的 m/z 范围内该谱中

的累计强度的步骤314。可以将标准化后的背景减除强度值在这些m/z范围内积分。将该积分值(即,在对应的预定义的m/z内的强度之和)分配给特征。对于在该m/z范围内没有检测到峰值的谱来说,可以将预定义的m/z范围定义为具有与当前m/z位置的峰值宽度对应的宽度的该特征的平均m/z位置周围的间隔。在美国专利7,736,905中进一步详细公开了该步骤。

[0166] 在步骤314,在一个可能的实施方式中在如下m/z范围中的一个或多个处获得了该谱的强度的积分值:

[0167] 表3

[0168] 5732至5795

[0169] 5811至5875

[0170] 6398至6469

[0171] 11376至11515

[0172] 11459至11599

[0173] 11614至11756

[0174] 11687至11831

[0175] 11830至11976

[0176] 12375至12529

[0177] 12502至12656

[0178] 23183至23525

[0179] 23279至23622和

[0180] 65902至67502。

[0181] 在一个实施方式中,在以如下表4中示出的峰为中心或涵盖其的多达8个m/z范围内获得了值。在美国专利7,736,905和作为US 2011/0208433公开的2011年2月22日提交的美国申请系列号12/932,295(其内容在此引入作为参考)中解释了发现这些峰的重要性以及方法。在实践中上述表3和4中的宽度(范围)或峰位置可例如根据如何进行谱对齐的变化而轻微变化。已进一步指出使用“DeepMALDI”技术(参见下文详述),在谱中揭示了很多可用于分类的特征。对于稀释和发射质谱,一种或更多种表1的峰,或表1和表3和4的特征组合可被用于分类。已进一步指出使用“DeepMALDI”技术,在谱中揭示了很多特征,其组合可用于分类,参见表2、5和6和上述图4A-4H的例子。

[0182] 在步骤316,将步骤314中获得的值提供给分类器,该分类器在所示出的实施例中是K最近邻(KNN)分类器。该分类器使用来自大批其他患者的分类标记谱的训练集。训练集将包括来自从产生免疫应答的疗法中受益或未受益的患者的分类标签谱,例如基于酵母的癌症免疫治疗,其可以是单独或联合其它抗癌治疗的针对突变Ras阳性癌症的基于酵母的免疫治疗。例如,上述工作实施例中描述的GI-4000研究中的训练集包括“快”和“慢”复发时间组中的分类标签谱。对该谱分配的分类标签将表现为“快”、“慢”、或等价物的形式,例如“受益”、“无应答”、“好”、“差”等等。KNN分类算法对314和训练集的值的应用实质上是一种来自在多维特征空间中预定义谱特征的累积强度值比较的距离计算和多数票决算法,如美国专利7,736,905所述。可以使用其他的分类器,包括概率KNN分类器、支撑向量机或其他分类器。

[0183] 在步骤318,分类器为该谱生成标签,例如,”快”或”慢”。该方法可以对单一等份执行,或对分成三等份的样品执行,其中对给来自定患者样品的三份分开的等份(或者使用不管什么数量的等份)并行地执行步骤304-318。在步骤320,进行核对来确定是否所有的等份都产生了相同的分类标签。如果不是,则如步骤322指示的那样,返回未定义(或不确定)结果。如果所有的等份都产生相同的标签,则如步骤324指示的那样报告该标签。

[0184] 如本文所描述的,随后将步骤324报告的分类标签用于引导患者的治疗。例如,根据本公开内容,根据分类步骤被标记“快”的那些胰腺癌患者被预测为不可能从产生免疫应答的疗法中受益,例如基于酵母的癌症免疫治疗,其可以是单独或联合其它抗癌剂包括吉西他滨的针对突变Ras阳性癌症的基于酵母的免疫治疗。作为另一个例子,如果分类后胰腺癌患者谱的分类标签根据测试被鉴定为“慢”,则患者被预测为可能从产生免疫应答的疗法中受益,例如基于酵母的癌症免疫治疗,其可以是单独或联合吉西他滨的针对突变Ras阳性癌症的基于酵母的免疫治疗,患者继续通过施用免疫应答生成疗法进行治疗,例如基于酵母的癌症免疫治疗,其可以是单独或联合吉西他滨的针对突变Ras阳性癌症的基于酵母的免疫治疗。

[0185] 应当理解,步骤306、314、316和318典型地是在被编程的通用计算机中执行,该通用计算机使用软件编码预处理步骤306、步骤314中累积强度值的获取、步骤316中KNN分类算法的应用以及步骤318中分类标签的生成。步骤316中使用的分类标签谱的训练集被存储在计算机的存储器中或存储在计算机可以访问的存储器中。

[0186] 如美国专利7,736,905中描述的那样,该方法和被编程的计算机可以有利地在实验室测试处理中心实施。

[0187] 表4:分类中使用的峰。

峰号 m/z

[0188] 1 5843

2 11445

3 11529

4 11685

5 11759

[0189] 6 11903

7 12452

8 12579

[0190] 注:表4中鉴定的峰的m/z值可根据谱对齐过程进行轻微偏移至更高或更低的m/z值,所述谱对齐过程用于对齐在训练集中使用的、和在预处理中对齐测试谱中的所有谱。

[0191] 图12是可被用于实践本公开的方法的实验室测试处理系统400的系统方框图。该系统400可在作为针对大批患者样品的实验室测试处理中心的实验室中实施,例如,在测试服务提供者商业中。该系统400接收血液来源的样品402,其可以是全血液、血浆、血清、或执行其它处理步骤后的血浆,例如,通过预先的基于密度的分离方法从全血液中获得的血浆。

使用前述流程将样品稀释并等分至MALDI-TOF板404的一个或多个点,然后将其插入MALDI-TOF质谱仪406。该质谱仪产生质谱408,其如常规的为数据对的形式(m/z 位置、强度)。然后将质谱数据以数字形式储存在数据库或机器可读存储器410中。存储器410可例如通过局域网连接通用计算机414,其细节不重要。存储器410进一步存储训练集412的分类标签谱。计算机410实施软件指令以对谱408进行预处理步骤(背景减除,标准化和对齐)并编码用于对预处理后的谱执行分类算法(例如K-最近邻),并使用训练集数据412。然后,计算机如图11所示对谱408产生分类标签,其如本文所公开的那样被用于指导治疗。

[0192] 测试系统400可通过计算机网络422从远端MALDI-TOF设备420接收质谱数据并执行图11的步骤304-324。MALDI-TOF设备可关联远程的临床、医院或实验室,其可以附属于实施分类器计算机414的实体,也可不附属,但是为了保证标准化和可再现性,通常执行分类和产生分类标签的相同实体也将执行患者样品的质谱。

[0193] 用于获得质谱的“DeepMALDI”方法

[0194] 在MALDI(基质辅助激光解吸电离)TOF(飞行时间)质谱中,将样品/基质混合物置于金属板上的定义位置(在此为“点”或“样品点”),称为MALDI板。将激光束以一个非常短的瞬间引导至点上的一个位置(称为一个“发射(shot)”),导致样品的分子或其它成份的解吸和电离。样品成分“飞行”至离子检测器。该设备以质谱的形式检测样品中的成份(分子)的质荷比(m/z)和相对强度。

[0195] 一般,在MALDI-TOF检测中,对MALDI板上的每个点施加几百个发射,将得到的谱(每个发射一份)进行累加或平均以对每个点产生一份完整的质谱。

[0196] 按常规的智慧,至少在复杂生物样品例如血清和血浆等的MALDI-TOF质谱领域而言,是没有必要对样品施加多于大致1,000发射数的,否则蛋白质的含量被减少,设备中的激光和探测器承受不必要的负担,而额外的发射数不会揭示关于样品的显著数量的额外信息。从而,当从复杂生物样品获得质谱数据时,例如在生物标记物发现研究中,通常使用500-1000发射数每样品点。

[0197] 在最近的探索性研究中,我们发现从相同的MALDI点或组合的来自同一样品的多个点的累加谱收集并平均很多发射数(大于20,000,一般100,000至500,000),会导致相对噪音/信号水平的减少,且能揭示来自复杂生物样品质谱的显著数量的额外信息。此外,很多使用MALDI TOF MS的标准范式看上去是明显错误的。首先,在点上的蛋白含量被彻底清除前在单一点上运行几十万发射数是可能的。其次,通过将很多发射数进行平均导致噪音的减少会导致出现之前不可见的峰(即,在1,000发射数下未出现的峰)。第三,当样品接受非常大数量的发射数(远大于1,000)时,甚至是之前可见的峰也变得更好地被定义并允许更可靠的峰强度测量和样品间比较。

[0198] 作为一个例子,已发现在MALDI-TOF质谱中对复杂生物样品例如基于血液的样品在单一点上进行大数量的发射数(>20,000及甚至100,000和500,000发射数)导致噪音水平的降低,并揭示了之前不可见的峰(即,在2,000发射数下不出现的峰)。此外,这可以在不减少样品蛋白含量的情况下完成。此外,之前可见的峰变得更好地被定义,并允许更可靠的样品间比较。在基于血液的样品的标准谱(~1,000发射数)中,通常有60-80个峰是可见的,而使用200,000发射数通常有~200-220个峰是可见的,使用500,000发射数通常~450-480个峰是可见的,且使用2,800,000发射数通常~760个峰是可见的。应理解在此报告的峰数与

MALDI-TOF设备设置有关,且这些数量仅为一个粗略的指导;根据设备设置以及还根据特定的峰检测算法(当然还有实际样品),更多或更少的峰将是可见的。还必须注意峰的质量和强度的定量(与丰度相关)也在至少一些测量中变得更好,如下讨论的图16A-16D中所示。

[0199] 在例如200,000发射数下揭示的峰被认为对应于血清样品中存在的极少量完整的(未消化的)蛋白。使用本文描述和所提及为“DeepMALDI”方法(即,来自同一点或来自多个点组合的大于20,000发射数每个点和优选的约250,000至750,000或更多发射数)的技术,我们相信血清样品中存在的非常大数量的蛋白及可能至少一半的全蛋白可以以半定量和可重复的方式被检测。以半定量方式的检测表示强度的检测(峰高、峰下面积)与绝对丰度或样品中的蛋白浓度有关,以可重复方式的检测表示可以对同一样品检测多次并在一些可接受的变异系数范围内获得相同的结果。

[0200] 从单一MALDI点获得大于20,000发射数可能超过现代MALDI-TOF机器的参数;然而我们在本文中描述了几种在此限制左右进行工作的方法。理想的,MALDI-TOF设备被设计为适应本文中描述的“DeepMALDI”方法,且如下描述中提供了对于此种机器的几个具体建议,包括自动光栅扫描特征和在单一点上进行极大的更多发射数的能力。

[0201] 从一个MALDI样品点使用几十万发射数的最迫切的问题是在常规的点制备中点内只有一些发射位置产生足够的离子流以对组合谱中产生信号有实质上的贡献。当已经使用劳力密集的人工过程靠视觉为激光发射选择MALDI板上给定点的高离子产出位置且可能用这一方法继续时,为激光发射选择位置过程的自动化成为可能并作为本发明高通量实施的优选(如果不是出于不浪费太多激光发射和大幅降低激光寿命的简单原因的话)。变换的方法是以使得大多数随机选择的位置都产生高离子流的方式来改进MALDI点的质量。两种方法都可用于产生DeepMALDI谱。

[0202] 本文的这一部分描述了几种用于谱获得自动化的方法。获得的自动化可包括以光栅模式定义点激光扫描的最优移动方式和在一个点内离散的X/Y坐标位置对多个光栅扫描产生特异性序列,以从一个或多个点产生如750,000或3,000,000发射数。例如,从250,000发射数每4个样品点的每一个获得的谱可被组合至1,000,000发射数的谱。如前所述,在多个含有相同样品的点上收集的几十万发射数至几百万发射数可一起平均,以产生一个谱。一种自动化的办法涉及对一个样品点的非连续X/Y光栅扫描产生光栅文件。另一种方法涉及将点分为子点的网格(例如,3X3或5X5网格)并对子点的离散X/Y坐标位置处的光栅扫描产生光栅文件。公开了第三种方法,其使用图像分析技术以鉴别感兴趣的区域,所述区域含有相对高的样品材料浓度用于谱获得(多发射数),和/或鉴别那些其中蛋白浓度相对较低的区域,并在具有相对高蛋白浓度的区域进行谱的获得。

[0203] 下文描述了优化样品应用至MALDI板(“点样”)的过程以在单一点内产生样品/基质的均一、同质的晶体。这一过程有利于使用自动化方法从MALDI板上的单一点获得几十万发射数。

[0204] 本公开的这一发现和方法具有多种应用,包括生物标志物发现、测试进展、底物测试、现存测试的验证和产生假设,例如,在发现生物标志物的努力中。特别关注的是该方法可用于本文其它部分描述的预测性测试。与现有方法相比,该方法通过其在复杂样品中以高通量方式可重现地定量更多蛋白的能力进一步增强了“稀释和发射”方法在质谱研究中的潜力。

[0205] 本文这一部分使用的术语：

[0206] 1. 术语“瞬态谱”指从指向MALDI点中的单一位置或x/y位置的单一包(packet)的激光发射获得的谱(每个包由定义数量的发射组成,例如,100、500、800发射数,等)。

[0207] 2. 术语“位置谱”指激光在MALDI点中的同一位置发射x次时,一种或多种瞬态谱的累积和。

[0208] 3. 术语“点谱”指在对一个完整的、单一的MALDI点发射的过程中获得的所有位置谱之和。点谱可使用单独的求和操作以对位置谱累加来获得,或在对位置谱进行对齐和/或标准化(例如,总离子流标准化)操作后使用求和操作获得。点谱可通常获得自MALDI点上的100,000至500,000发射数。获得点的其它选项是可能的,包括a)对位置谱进行背景减除和标准化并随后累加;b)对位置谱进行背景减除和对齐on the位置谱并随后累加;c)进行位置谱的背景减除、对齐和标准化并随后累加。我们发现通过位置谱的总离子流标准化(细节参见美国专利7,736,905)并随后累加获得了最佳的动态范围;在点谱中将进行任何背景减除。

[0209] 4. 术语“发射位置”指一个给定的位置,其中激光束拦截MALDI点用于发射。为了获得200,000或500,000发射数每MALDI点,激光束在MALDI点上指向多个(例如,数百个)个体发射位置,例如,通过手动或更优选的以自动化的方式使用激光束对所述点的光栅扫描。如下所述,光栅模式设计是重要的,因为通常不希望连续在紧邻点位置处发射。因此,光栅模式设计顺序选择具有一定空间分离的发射位置,并以空间移动的方式重复对整个MALDI点的扫描,以避免连续在紧邻点位置处发射。

[0210] 5. 术语“瞬态谱过滤”指一个过滤或选择过程,其用于接受或拒绝瞬态谱。作为示例,在瞬态谱过滤中,为了接受一个瞬态谱,在瞬态谱中必须存在预定义m/z范围内的峰的最低数量(例如,5),且瞬态谱中信号/噪音比必须超过一个特定的阈值。也可以采用其它过滤标准,例如谱的总离子流需要超过某一预定义的阈值,或通过使用如下所示的排除列表或包含列表。谱过滤从整体上接受或拒绝瞬态谱。

[0211] 6. 本文使用的术语“复杂生物样品”定义为含有几百或几千分析物的样品,例如完整蛋白,其丰度覆盖一个很大的动态范围,通常为许多数量级。该复杂生物样品的实例包括血液或其成份(血清或血浆)、淋巴液、导管液、脑脊液和表达的前列腺浆液。该复杂生物样品也可包含环境或食物样品。

[0212] “DeepMALDI”方法中揭示的谱信息的一个实例见图16A-16E。图16A-16C是选定的质/荷范围(m/z比7,000-8,000)内显示同一样品(血清)的3个谱的图,显示了随着增加的发射数,可检测峰内容增加。图16A的谱来自2,000发射数,图16B的谱、图16C的谱来自500,000发射数。特别注意图16A的谱是如何基本表现为噪音和表现得含有很少或没有可分辨的目标谱信息的。对比图16A和图16B,其中图16B的谱(来自100,000发射数的谱)含有很多个体的峰,例如在10)处鉴定的峰,其在图16A的谱中不存在。在图16C的谱中,在谱中显示有很多在其它谱中不显示的峰。将图16C和16B与图16A进行比较,显然在100,000发射数和500,000发射数揭示了大量的谱信息,其在图16A(2,000发射数)的谱中不存在,且如在图16B和16C中所示的,通过DeepMALDI方法降低了噪音水平。

[0213] 图16B和16C的谱增加了谱对动态范围的敏感度,所述范围可以指定并可以允许将峰强度与丰度关联。使用峰强度来分析复杂生物样品在给定浓度下的一个分子的存在是可

能的。例如,在本方法中人们将定义样品中的目标分子(具有已知质量),将样品涂(dope)至目标丰度水平(摩尔浓度或ppm)并用于MALDI板;在板上进行若干发射(例如,大于100,000)直至该分子以特定丰度(强度)可靠地呈现于谱中(在已知m/z位置的峰)并记录发射数量("x")。产生被称为“参考谱”的这一过程将有常规限制条件和标准化方法以确保可靠性,其对本领域技术人员而言是显而易见的。然后,待测试的目标样品将进行MALDI-TOF和x数量的发射。如果得到的谱显示对应目标分子的已知位置处的峰强度小于参考谱中的峰强度,那么样品中的目标分子浓度小于产生参考谱中使用的样品中的分子浓度。该方法可同时被用于多个分析物。此外,可以针对一系列已知浓度范围的目标分子以x发射数获得多个参考谱,且可以将测试谱与参考谱进行比较以确定测试样品中目标分子的大致浓度。该方法可用于多种目的,例如,如运动员的药检、代谢物浓度检测、环境样品检测等。目标分子可以是蛋白质,例如代谢物、癌症抗原(CA)125、前列腺特异性抗原(PSA)、C-反应蛋白等,其质量范围约1K道尔顿至50K道尔顿。

[0214] 图16D是对在DeepMALDI方法中呈现的谱中大量动态范围的描述。图16D的插图是m/z范围为7140kDa至7890kDa的谱的一部分,其表现了在约~500,000发射数下获得的所述谱和多个峰10。将背景估值(虚线)叠加在谱上,其可被减除以产生背景减除的谱。注意插图和具体的很多峰10中的谱信息在图16D的主要部分中不可见。在图16E中,谱在插图中表现,其Y轴放大以显示额外的谱信息和约9520的m/z区域的具体的峰强度信息,其在DeepMALDI方法中有显示,但是在典型的~1,000发射数的谱中不可见。

[0215] 图16A是含有384个矩阵排列的样品点或“点”的MALDI-TOF靶板的平面图。这些点由列数1...24和行数A...P确定,例如,左上方的点被确定为A1。图16B是一个个体样品点P1(14)的放大图,其上被叠加了一个具有原点(0,0)的X/Y坐标系16。样品点14被显示分为一个5X5的直角坐标网25个体子点(sub-spots)18。直角坐标网18和位置坐标系16被用于自动光栅扫描方法以从这些点获得100,000或更多发射数,如下详述。

[0216] 最初注意到大数量的发射(>20,000)的自动产生并非绝对必要的,目前可获得的MALDI-TOF设备中存在的特征可以使用。一般而言,在DeepMALDI技术中,暴露于激光发射时,在能产生高蛋白得率的MALDI点上选择位置是重要的。现存质谱设备中的标准软件允许使用规则的预定义途径在点上移动,即方形模式、六边形模式、螺旋模式(从一个点的中心)。MALDI板上的发射位置在一个称为“教学(teaching)”的过程中定义,其是现存的BrukerCorporation的MALDI-TOF设备中的FlexControlTM(Bruker)质谱仪控制软件的一部分。(虽然本文偶尔提及一个Bruker Corporation设备的特征,但是本发明的方法显然不会限于任何特定设备或特定生产商的设备。)

[0217] 图18中显示了含有均匀分布在点中的样品/基质混合物的一个MALDI点的示例。来自Bruker Corporation的质谱设备包括嵌入式摄像机,其显示了一个MALDI点的区域;人们可以人工选择挑选亮的位置30以用激光对准。暗的位置32应当被避免。有时候亮的位置不产生好的得率,其可能与盐晶体的存在有关。在发射过程中,一个点中的区域可能变得耗尽(depleted);从而暗的区域(具有低得率的耗尽的区域)需要被避免。手工方法将继续在发射过程中获得并展示点的图像。

[0218] 在我们预备实验的过程中,我们发现随着越来越多发射数的使用,要发现好的位置会变得越来越难。当重复使用同一个点的时候也发现了这种效果,例如在之前的50万发

射数后增加第二个50万发射数。第二轮没有获得如期望的质谱噪音水平的减少。事实上,得到的平均化的谱可能总体质量更差,可能是由于对来自太多空位置的发射数进行平均。这可能导致对早期位置的获得偏差(acquisition bias),如果只使用眼睛选择发射位置并接受或拒绝谱且不使用瞬态谱过滤的话,这种偏差需要被控制。如果人们使用自动光栅扫描和位置谱过滤的话,则这种偏差被消除。

[0219] 然而,为了增加通量,需要自动化位置选择过程并从给定点获得高数量的发射。以下部分描述了几种方法。下述方法能够在13–15分钟内从MALDI板中位于3个点的一个样品获得750,000发射数(250,000发射数每个点),其样品需要3微升血清。

[0220] 谱收集的自动化

[0221] 虽然已使用劳力密集的人工过程来在MALDI板上的给定点视觉选择位置产生每个点100,000或500,000发射数的多发射数获得了结果,且可能用这种方法继续,但是为激光发射数选择位置过程的自动化是可能的,且本文中描述了几种方法。获得的自动化可包括定义以光栅模式进行激光扫描点的最优移动模式和在一个点内离散的X/Y位置处多个光栅扫描的产生序列以从该样品点产生例如100,000、250,000或500,000发射数。一种自动化方法涉及对样品点非连续X/Y光栅扫描的光栅文件的产生。光栅模式设计是重要的,因为通常不希望连续在紧邻点位置处发射。因此,光栅模式设计顺序选择具有一定空间分离的发射位置,并以空间移动的方式重复对整个MALDI点的扫描,以避免连续在紧邻点的位置处发射,和选择新的发射位置。

[0222] 另一种方法涉及将点分为子点的网格(例如,3X3或5X5网格)(参见图17B)并对子点的离散X/Y坐标位置处的光栅扫描产生光栅文件。

[0223] 公开了第三种方法,其使用图像分析技术以鉴别感兴趣的区域,所述区域含有相对高的样品材料浓度用于谱获得(多发射数),和/或鉴别那些其中样品(例如蛋白)浓度相对较低的区域,并避免在相对低样品(例如蛋白)浓度的区域的谱获得。

[0224] A. 非连续X-Y坐标的光栅扫描

[0225] 从一个点获得大数量发射数的过程的自动化方法涉及对样品点非连续X/Y光栅扫描的光栅文件的产生。这将结合图19和20进行描述。

[0226] 图19是对用于从图18的点14获得100,000或更多发射数的光栅扫描模式500的描述。该点14用光栅扫描多次,例如,25次。图19中显示的符号集502描述了个体的、离散的X/Y位置,其中该点在单一光栅扫描中被扫描(发射)。根据图中显示的在中心(位置0,0)具有原点的坐标系定义X/Y位置。在扫描过程中,当激光被引导至每个位置,该位置的点可进行多个发射数,例如,700或800发射数每位置(position)/位置(location)。从图19中显示的模式人们可以注意到每个光栅扫描由在点内个体的、离散的位置的发射组成。顺序实施个体光栅扫描从而避免在点紧邻位置发射。图20显示了图19的光栅扫描模式对图18的点的叠加。

[0227] 为如图19所示的光栅扫描产生具有非连续X/Y坐标的25个光栅文件的过程描述于2012年5月29日提交的美国临时申请61/652,394的附件中,感兴趣的读者被指向该文件进行进一步参考。

[0228] B. 使用网格将点分离为子点和子点的光栅扫描

[0229] 本方法的一个目的是使得手动在样品点(即点A1,点A2,等)上选择位置/光栅的过

程自动化,其在数据获得过程中得到“可接受的”谱,并进行这一过程直至几十万的谱被添加至累加缓冲器中。累加/平均几十万的谱增加了信号/噪音比,从而允许检测显著更多的峰,如前所述。

[0230] 如上述非连续光栅扫描的情况,在如图18所示的样品/基质混合物是基本上均匀和同质分布于整个点时,本部分描述的网格的使用效果最好。为达到这个目的的目前优选的方法在本文后面描述,用以稀释并发射血清和芥子酸(基质)。由于这种均匀分布,我们从而可以从实质上在样品点上所有的位置/光栅获得谱,其消除了预先对“可接受的”谱评价所有的位置/光栅的必要性。

[0231] 在一个样品点上收集几十万个谱可通过定义一个将点14细分为子点或网格元素18的网格(图17B)来完成,其覆盖了样品点并从子点18中的每个位置/网格点/光栅收集定义数量的谱,直至期望数量的谱已被加至累加缓冲器中。之前版本的Bruker软件只允许以自动模式每样品点累加最大20,000的总谱(图21)。

[0232] 为了绕开这一限制,我们最初定义了一个5乘5的网格区(图17B,16),其将每个样品点分为25个8x 8网格或子点18(图17B)。对每个网格或子点18产生了分离的光栅文件。该设备被命令在网格18的每个位置/光栅处获得800个谱(发射)直至20,000谱被加至(谱)累加缓冲器。那时,自动化方法将命令设备移动至下一个网格或子点18并使用下一个光栅文件和产生另一个20,000谱。在实践中,人们设计25个光栅文件,每个子点18一个,其每一个被附加至分开的自动ExecuteTM(Bruker)方法,其根据该方法中建立的评价标准获得数据。

[0233] 该过程允许分20,000发射数每批来获得500,000发射谱(20,000发射谱每格x 25格),每批使用Bruker的flexcontrolTM软件工具而不必使用成像应用例如flexImagingTM(Bruker)。该过程的结果是每个样品点25个谱文件,每个含有一个由20,000发射谱组成的累加谱。这些25个谱文件可以随后被累加,以产生MALDI板上一个点从500,000发射数获得的总谱,例如,如图16C、16D和16E所示。

[0234] 最新版本的flexcontrol TM(Bruker)允许人们累积一个来自最多500,000发射数的累加谱。例如,在图21中自动ExecuteTM(Bruker)方法编辑器允许在800发射数的步骤中累加20,000发射数(800发射数每位置/光栅)。

[0235] 然而,人们仅能对每个样品点收集一个累加谱(x瞬态谱之和)。为了从一个单一样品点得到几批累加的谱,我们需要对现有的MS设备中的软件进行调整。随着这些调整我们可以从一个或几个形成网格的光栅获得谱,例如上述的那些,并单独存储每个瞬态或位置谱。例如,设备可被命令为收集并存储图17B的网格或子点18中每个光栅(x,y位置)获得的每个800发射位置谱,而不必加至累加缓冲器。对在样品点A1、A2、A3等的所有子点重复相同过程(例如可从250光栅每样品点获得800发射谱=200,000发射数每样品点)。可将谱过滤应用于自动Execute TM(Bruker)或不应用来获得位置谱。

[0236] C. 图像分析

[0237] 谱获得的自动化的一个选择是图像处理技术,用以鉴定具有高蛋白得率/高样品浓度的点上的空间位置,特别是在样品没有在点的空间上均匀分布、而是集中于离散区域的情况下。在一个可能的实施方式中,设备中包括的摄像机被用于获得训练点的光学图像。然后,从训练点上一个光栅的位置获得质谱。得到的质谱与点的光学图像组合被用于产生分类机制以从光学图像检测从给定样品制备物制备的进一步的点的高得率位置。然后将该

分类用于实际的样品点。虽然这是一种优质解,但我们遇到过捕捉摄像机供给(feed)和从摄像机图像至激光发射位置的位置可重复校正问题。

[0238] 一种可选的方法是使用直接为质谱成像方法的形式的质谱仪来考察点。该想法是首先允许初步扫描并在点上每个精密标度(方形)的位置发射低数量的发射数(几十个)。对这些光栅位置的每一个收集谱,且对每个位置将记录总离子流或一些预定义m/z范围内的离子流。基于来自初步扫描过程中的N最高强度位置产生光栅文件并用于最终获得质谱。该方法利用了Bruker FlexImagingTM软件作为最可行解以在质谱仪成像过程产生多个谱。软件分析这些谱并产生最终光栅扫描模式。虽然本方法看上去可用于使用芥子酸作为基质的标准稀释和发射过程,但是其对于其它基质和预先分级的样品集(例如CLCCA,参见Leszyk, J.D. Evaluation of the new MALDI Matrix 4-Chloro-a-Cyanocinnamic Acid, J.Biomolecular Techniques, 21:81-91 (2010))以及如NOG沉淀的其它方法(Zhang N.等, Effects of common surfactants on protein digestion and matrix-assisted laser desorption/ionization mass spectrometric analysis of the digested peptides using two-layer sample preparation. Rapid Commun. Mass Spectrom. 18:889-896 (2004))可能是次优的。该可选方法的一个重要方面是在MS成像部分发现获得设定从而不产生太大的文件。标准的获得文件是兆字节的数量级,以400乘400光栅扫描(400个位置,400发射数每位置),我们产生了16,000谱。由于这些谱的需求一点都不大,且我们只需要估计总离子流,因此我们可以以低分辨率设定进行工作。可能从自动化谱获得设定来直接获得一系列可用位置,即,得到一系列成功或失败的获得。从我们的考察来看,看上去使用质量过滤作为MS成像包的一部分产生一系列通过某些标准的位置(通过文件列表验证)是可能的。虽然这将很大地帮助产生原型工作流,其还需要通过专门的软件进行优化以避免半人工过程。

[0239] 图22显示了使用CLCCA作为基质的一个MALDI点,其中高得率区域包括线结构,低得率区域作为黑暗区域显示。对于这些其中基质样品非常不均匀结晶的情形,如图20中所示,成像分析方法似乎最敏感。成像分析鉴定了相对高得率的区域(120,122)。相对低得率的区域,例如左下方的区域124和基质区126通过成像分析软件进行鉴定,并在发射过程中被忽略。

[0240] 鉴定一个点上的高得率和低得率区域的图像分析软件可以多种形式,并可以由本领域技术人员开发。例如,点(图19)的黑色和白色图像包括像素数组,每个具有8比特的量化值,0为黑色(无信号),255为白色(饱和的)。可以使用过滤来鉴定相对高得率的区域,例如将像素值大于如100的像素鉴定为“高得率”,将像素值低于40的像素鉴定为相对“低得率”。然后可以继续对那些对应像素具有100或更高的值的样品点区域进行扫描。还可以将其中像素值为240-255的点位置滤除,因为这些区域可被判定为具有盐结晶或其它导致低得率的性质。再参见图22,晶体结构120,122的像素具有落入100-240范围内的像素值,从而其将被扫描,而黑色区域124和126将不会被扫描。还可以使用形态学处理技术以鉴定如图22的晶体120的结构。图像分析软件可同时包括形态学处理和过滤以确定待扫描的区域。此外,在扫描过程中可以改变点(由于样品的耗尽),且在扫描过程中可以运行图像处理以使得从一个点产生100,000或更多发射数的过程中的发射最优化,而那些低样品浓度的位置在发射过程中被避免。

[0241] 图23是来自MALDI-TOF设备的截图，显示设备工作站130的展示，包括点14的图像132，在此情况下为板上的点F17。板的布局见于12'，其中点F17在14'指明。选择了一组点134(D9-F20)来使用上述图像分析方法以自动方式运行。

[0242] 图24是来自该设备的另一截图。目前的设备运行用户设定评价区以接受或拒绝瞬态谱(使用评价标签)，设定每个点累加多少谱(使用累加标签)和跨点“移动”从而激光可以某种模式发射(使用显示的“移动”标签)。选项包括以例如六边形或螺旋模式随机步进或移动。软件还允许用户保持发射激光并根据该参数获得和增加总谱，直至从一个发射位置收集了来自750发射数的谱为止，并随后移动至下一个发射位置。人们可以在发射位置被认为是一个失败的点之前设定尝试次数。鉴定可能的低得率区域并避免在这些区域中发射的图像分析方法有助于大量降低或消除那些失败的判断。

[0243] 图25显示了一个评价页面，其中选择了接受或拒绝瞬态谱的质量范围，如在150处指出的。在捕获过程中，如果瞬态谱不具有超过阈值集(基于分辨率、信号强度或其它因素)的预定义范围内的峰——在此情形下为5,000–18,000Da，那么其将被拒绝。也就是说，该瞬态谱将不被累加至累加缓冲器以形成位置谱(累加来自所有发射的谱)。

[0244] 图26显示了一个评价页面，其中如果有人们不想要包括在评价中的特别的峰，人们可以制作一个排除列表，并将这些峰标记为“背景峰”。软件已经对基质预定义了“对照列表”，其定义了背景峰，或人们可以导入峰列表。

[0245] 从多个点收集谱

[0246] 一般而言，人们可以将DeepMALDI技术延伸至来自多个点的组合谱。例如，人们可以从标准MALDI板上的点A1、A2、A3、A4和A5的每一个获得一个样品的500,000发射数(参见图17A)，并组合(累加)得到的谱至一个总谱，其由2,500,000个谱(发射)的和组成。可推理，没有理由相信人们不能从多个点组合谱以达到极其高数量的发射，即，100个点x 1百万发射数每点可以给我们来自1亿发射的结果。对此过程可能有实践方面的限制，例如，激光可能因太频繁而失败。

[0247] 从DeepMALDI中的多个点收集谱的实施例

[0248] 在本方法的一个实施例中，可能从MALDI板上同一血清的多个点收集来自5百万发射数的谱，其中使用手动或自动产生的光栅用来使用前述技术扫描多个点。在本方法中，优选获得MALDI板上单一样品的可再产生的均质点。这可以使用在此描述的方法获得。

[0249] 1. 将稀释的血清点样至MALDI目标板。

[0250] 流程：

[0251] 用HPLC级水1:10稀释血清并漩涡振荡。在0.5ml微量离心管中将样品与基质(20mg/ml芥子酸，用50%ACN/0.1%TFA配制)1:1(v/v)混合并漩涡振荡。在MALDI目标的一个或多个点上点样4μl的基质/样品混合物。

[0252] 在本实施例中使用了MALDI板中的36个点(位置)：

[0253] 管1：点样在MALDI板的位置E13,E14和E15(参见图2A)

[0254] 管2：点样在位置E16,E17和E18

[0255] 管3：点样在位置E19,E20和E21

[0256] 管4：点样在位置E22,E23和E24

[0257] 管5：点样在位置F1,F2和F3

- [0258] 管6:点样在位置F4,F5和F6
- [0259] 管7:点样在位置F7,F8和F9
- [0260] 管8:点样在位置F10,F11和F12
- [0261] 管9:点样在位置F13,F14和F15
- [0262] 管10:点样在位置F16,F17和F18
- [0263] 管11:点样在位置F19,F20和F21
- [0264] 管12:点样在位置F22,F23和F24
- [0265] 使用相同的吸管头振荡3次(3x每个管中15 μ l的4 μ l)后,样品点E13-F18(管1-10)被直接使用;而最后6个样品点F19-F24(管11和12)如点E13-F18中使用,但是还要在板上上下吹吸。
- [0266] 通过将目标板置于台子上使MALDI板上的点在环境温度下干燥。
- [0267] 结果:
- [0268] 对于点E13-F17(其直接应用至板,没有进一步在板上混合),每个管的第3个点显然比前两个更均匀。均匀性通过视觉评价:第3个点最好,第2个次之,第1个最不均匀,除了E23之外,其来自管4的3个点的第2个,但是看上去比第2个点更像是每个管的第3个点。
- [0269] 在管中漩涡振荡并在板上上下吹吸进行混合的样品点F18、F19、F20、F21、F23和F24非常相似并具有与来自E13-F17的集合的第3个点相同的均匀性表现。F22看上去与E23相同。
- [0270] 2.从5百万发射数获得谱
- [0271] 在进行了上述流程后,从16个MALDI点获得了来自约312,500发射数每点的质谱数据:
- [0272] E15、E18、E21、E23、E24、F3、F6、F9、F12、F15、F18、F19、F20、F21、F23和F24。
- [0273] 使用如上所述的光栅扫描文件,来自每个点的谱被累加以产生一个从约5,000,000发射数获得的样品总谱。
- [0274] 样品应用至MALDI板的优化(点样)
- [0275] 样品应用至MALDI板被优化以对MALDI板上的每个样品点提供同质和均匀分布的结晶样品,图15中显示了其的一个例子。如下所述进行了几个实验来发现将样品混合物提供至MALDI板上的点(“点样”)的最优流程。这些实验在本部分描述。
- [0276] 最初,制备了几个不同的血清制备物。除非另有说明,否则点样2 μ l的基质。除非另有说明,否则在样品制备管中混合稀释的样品和基质介质。除非另有说明,否则我们从一个单一的制备管中没有点样多于1个点,因为从样品制备管中取出多个等份会影响结晶。
- [0277] 进行了基础钢板(Ground Steel Plate)实验,其产生了均匀的点。流程如下:
- [0278] 1.1:10稀释样品(2 μ l样品+18 μ l水),然后与50%ACN/0.1%TFA的基质(芥子酸25mg/ml)1:1(v/v)混合,点样2 μ l的基质。该流程未产生良好均质的晶体。
- [0279] 2.装填(prime)基质头。在点样头中吸取2 μ l的基质并静置30秒。1:10稀释样品(2 μ l样品+18 μ l水),然后与50%ACN/0.1%TFA的基质(芥子酸25mg/ml)1:1(v/v)混合。从吸管头射出过量基质。将吸管头放入样品基质混合物中并吹吸3次。不更换点样头的情况下点样2 μ l的样品基质混合物。这一流程形成了良好的均质晶体。因为这是基础钢板,因此样品基质混合物的展开没有像抛光钢板上那样多。留在吸管头的干燥的晶体可通过作为进一步晶

体形成的种子来改进结晶。

[0280] 3. 研究了温度对结晶的影响。1:10稀释样品 ($2\mu\text{l}$ 样品 + $18\mu\text{l}$ 水), 然后与 50% ACN / 0.1% TFA 的基质 (芥子酸 25mg/ml) 1:1 (v/v) 混合。将样品在 37°C 水浴下放置 5 分钟。从水浴中移出样品并立即点样。该过程没有产生良好均质的晶体。

[0281] 4. 重复如上的实验 2 但是点样 $4\mu\text{l}$ 的样品而非 $2\mu\text{l}$ 。该流程形成了良好的均质晶体。点样 $4\mu\text{l}$ 完整地覆盖了点的直径, 并产生了良好的晶体和数据。这是目前被认为最优的流程。

[0282] 在此点样的流程通过实例的方式提供而非限制, 来自公开方法的变形当然是可能的。例如, 人们可以在管中混合基质和样品材料并在点样前使其静置几分钟。已经注意到使用相同吸管头从相同管中制备的点越多则人们得到更均质的晶体。例如, 人们可以使用同一吸管头从同一管点样 10 个点, 并只收集最后的 5 个点左右的数据; 或可选的人们可以在 MALDI 板上点样前弃去来自管的第一轮的 5 个 $4\mu\text{l}$ 等份。

[0283] 我们还发现按照 1 中的流程但是使用同一吸管头点样同一样品管 10 次 ($2.5\mu\text{l}$ 每个点) 至抛光的目标钢板上产生了相似结果 (谱质量)。

[0284] 进一步的考虑

[0285] 工艺再现性

[0286] 可通过例如以每天 100 批的方式运行 1,000 次工艺重复来进行工艺再现性研究。人们可以研究对样品 (点) 制备 (板上或板下) 的依赖性, 特别要考察是否存在产生更均一的离子流得率的制备方法, 例如在样品稀释方面的变化。人们也可以监控高得率位置的数量是如何从点到点之间变化以及如何最小化其中的改变。在高水平的粒度下监控和记录所有获得和制备是好的实践。

[0287] 样品到样品再现性

[0288] 样品到样品再现性的相似问题可针对样品到样品的变化进行研究。新的现象可能发生: 可能一些样品富含蛋白质并得到具有更高得率位置的点。可能从样品属性 (光密度和颜色) 的一些模式获得检测结果, 或使样品获得设备 (例如用于血清的) 标准化以产生更可再现的流程。人们可以使用具有尽可能多样的来源的组合的样品集合来尝试覆盖最多变量。应当从研究现存集合中获得这样的一个集合, 并根据已知样品收集和条件进行匹配, 其使得能更好地使用现存的样品数据库。

[0289] 灵敏度

[0290] 观察谱中更多的峰导致了一个问题, 即在本方法中我们可以发现多大的丰度范围, 以及什么蛋白种类是事实上可见的。这涉及“常规智慧”, 即在复杂样品的 MALDI MS 中由于“离子抑制”人们不能观察较低丰度的离子——一种来自更丰富蛋白的离子抑制来自较丰富蛋白的离子信号的观念——从而使得较低丰度蛋白不能检测。该观点看上去完全基于对较低丰度离子观察的缺失。事实上, 我们对峰含量增加的观察 (参见例如, 图 16C) 对此解释提出了一些疑问。相反, 看上去人们要认真对待 MALDI MS 的 (半) 定量属性。如果人们同意蛋白丰度以很多数量级跨过一个大的范围, 那么人们可以预期对应的质谱将通过展示峰高度 (或者峰下面积) 的很大不同来模拟这一表现。人们将不会预期在 MALDI 谱中观察到低丰度蛋白, 这不是因为其没有离子化, 而是因为对应低丰度蛋白的峰的幅度应该是非常低的。由于关注大的峰是质谱中的常规实践, 且因为较低丰度的峰的将是几个数量级地更少,

因此不奇怪这些峰从未被观察过。这不是说像离子抑制等的现象不会发生,或离子化概率不发挥作用,而是说这些现象没有完全抑制来自低丰度蛋白的峰,且如果人们在谱的低强度区域寻找低丰度蛋白峰,它们的确会变得可观察到。从而可以将覆盖显著比例的血清蛋白质组的寻求视为扩大质谱动态范围的寻求。正如任何一种其它基于计数的技术,对此问题的简单解决方案是增加检测离子(每飞行时间箱)的数量。

[0291] 为了对这种与常规智慧相抵触的简单解释获得更多信心,人们可能希望建立质谱的动态范围并将其与蛋白丰度关联。这应当同时从建立灵敏度曲线(作为m/z的函数)的分析化学视角和通过鉴定对应一些峰的对应蛋白质和通过正交技术如ELISA的这些蛋白的比较性丰度测定来完成。

[0292] 使用预分级的样品

[0293] 本公开的方法可联合用于分级样品的沉淀方法例如NOG沉淀、脱液化等使用。该方法还可与其它基质如CLCCA一起使用。有可能这些方法也可以从DeepMALDI方法中获益很多。我们使用样品预分级的初步数据表明人们的确可以看到不同的峰,但是峰含量远远不到最优。这可能被预期为一个目的是避免高丰度蛋白。

[0294] 过去我们试图使用消耗和/或质量过滤来降低不想要的蛋白的含量,如白蛋白和血红蛋白,但是这些方法都不能导致整体去除,这些峰的残余仍然是可见的。使用在此描述的DeepMALDI方法对于消耗和质量过滤的样品应该产生更好的结果,因为减少大的峰将同时减少看到较低丰度的蛋白所需要的动态范围。

[0295] 得到谱获得设定的感性选择

[0296] 在自动ExecuteTM(Bruker)方法中,可能定义过滤设定以只收集通过某种标准的瞬态谱;在我们的情形下我们期望仅累加那些具有总离子流大于外部定义阈值的瞬态谱(来自<xx>数量的发射)。虽然这看上去以简单方式是不可能的,但是在处理方法标签中有可能被用于简单目的的过滤标准。可选的,在峰评价方法中可能存在我们能够调节用于该目的的参数。虽然这样不能减少发射的数量,但是其可以克服相对较早的发射数的发射偏差的问题,即不获得仅包括噪音的瞬态谱。自动化过滤操作在累加瞬态谱以产生位置谱中的使用避免了偏差问题。

[0297] 增加点的大小

[0298] 由于存在由激光照度大小和由预光栅化步骤的最小网格大小所产生的限制,很有可能一个标准点上没有具有足够离子得率的足够的发射位置。解决这个问题的一个简单方法是增加点的大小。FlexImagingTM(Bruker)软件能很容易地支持这一点。还有在MS成像应用中使用的矩形点样区域的选项,其可能适合这个目的。使用更大的点的另一个好处是人们不必担心是否能定位类似数量的合适的(decent)发射位置并产生从一个点到另一个点的相似质量的谱。看上去样品体积不成问题。如果更大的点是可能的,其将减少对同一获得物处理多个点的物流(logistics),其对高数量的发射可能是必须的。

[0299] 对“DeepMALDI”方法更进一步的考虑描述于2012年5月29日提交的美国临时专利申请系列号61/652,394,感兴趣的读者被指向该文件。

[0300] 基于酵母的免疫治疗和本发明的治疗方法

[0301] 本公开的一个实施方式指向设计为刺激针对由罹患癌症的患者的肿瘤细胞表达的癌症抗原的治疗性免疫应答的基于酵母的免疫治疗组合物的使用。该方法包括对患有表

达癌症抗原的癌症的对象和已通过根据如本文描述的发明的质谱预测方法被鉴定或选择作为可能从施用组合物中受益的对象施用针对癌症(即,包括癌症抗原)的基于酵母的免疫治疗组合物的步骤。

[0302] 更具体的,在一方面,描述了一种用基于酵母的癌症免疫疗法治疗癌症患者的方法。该方法包括以下步骤:(a)执行按照本文描述的任一预测方法的测试,如果对谱的分类标签表明患者可能从基于酵母的癌症免疫治疗中受益,则(b)施用基于酵母的癌症免疫治疗。在一方面,患者在基于酵母的癌症免疫治疗之前、同时或之后被另外进行一种或多种额外的抗癌治疗。在一个实施方式中,额外的抗癌治疗包括但不限于手术(例如,手术切除肿瘤)、化疗、放疗、靶向癌症治疗(例如,小分子药物或特异性靶向肿瘤生长和发展中涉及的分子的单克隆抗体疗法)、和姑息治疗,或其任意组合。

[0303] 在另一方面,本公开涉及使用基于酵母的癌症免疫疗法治疗癌症患者的方法。该方法包括对根据任一如本文描述的发明的预测方法的测试选择的癌症患者施用基于酵母的癌症免疫治疗的步骤,其中谱的分类标签显示患者有可能从基于酵母的癌症免疫治疗中受益。在一方面,患者在基于酵母的癌症免疫治疗之前、同时或之后被另外进行一种或多种额外的抗癌治疗。在一个实施方式中,额外的抗癌治疗包括但不限于手术(例如,手术切除肿瘤)、化疗、放疗、靶向癌症治疗(例如,小分子药物或特异性靶向肿瘤生长和发展中涉及的分子的单克隆抗体疗法)、和姑息治疗,或其任意组合。

[0304] 在本公开的另一方面,描述了一种用针对突变Ras阳性癌症的基于酵母的免疫疗法治疗癌症患者的方法。该方法包括以下步骤:(a)执行根据上述任何一种预测方法的测试,如果对谱的分类标签表明患者可能从针对突变Ras阳性癌症的基于酵母的免疫治疗中受益,则(b)施用针对突变Ras阳性癌症的基于酵母的免疫治疗。在一方面,在进行针对突变Ras阳性癌症的基于酵母的免疫治疗之前、同时或之后,患者被另外进行一种或多种额外的抗癌治疗。在一个实施方式中,针对突变Ras阳性癌症的基于酵母的免疫治疗是一系列称为GI-4000或等价物的基于酵母的免疫治疗产品中的一种产品。在本发明这一实施方式的一个方面,突变Ras阳性癌症可包括但不限于胰腺癌、非小细胞肺癌(NSCLC)、结肠直肠癌、子宫内膜癌、卵巢癌、黑色素瘤和多发性骨髓瘤。在一方面,该癌症是胰腺癌。在一个实施方式中,额外的抗癌治疗包括但不限于手术(例如,手术切除肿瘤)、化疗、放疗、靶向癌症治疗(例如,小分子药物或特异性靶向肿瘤生长和发展中涉及的分子的单克隆抗体疗法)、和姑息治疗,或其任意组合。在一方面,针对突变Ras阳性癌症的基于酵母的免疫治疗协同吉西他滨或等价物对患者进行治疗。在一个实施方式中,患者是胰腺癌患者,该治疗包括单独或联合吉西他滨的一系列称为GI-4000或等价物的基于酵母的免疫治疗产品中的一种产品。在一方面,癌症患者的肿瘤在用基于酵母的免疫治疗组合物治疗之前已被手术切除。

[0305] 在本公开的另一方面,描述了一种用基于酵母的癌症免疫疗法治疗癌症患者的方法。该方法包括对根据任一如本文描述的发明的预测方法的测试选择的癌症患者施用针对突变Ras阳性癌症的基于酵母的癌症免疫治疗的步骤,其中谱的分类标签表明患者有可能从针对突变Ras阳性癌症的基于酵母的免疫治疗中受益。在一方面,在进行针对突变Ras阳性癌症的基于酵母的免疫治疗之前、同时或之后,患者被另外进行一种或多种额外的抗癌治疗。在一个实施方式中,针对突变Ras阳性癌症的基于酵母的免疫治疗是一系列称为GI-4000或等价物的基于酵母的免疫治疗产品中的一种产品。在本发明这一实施方式的一个方

面,突变Ras阳性癌症可包括但不限于胰腺癌、非小细胞肺癌 (NSCLC)、结肠直肠癌、子宫内膜癌、卵巢癌、黑色素瘤和多发性骨髓瘤。在一方面,该癌症是胰腺癌。在一个实施方式中,额外的抗癌治疗包括但不限于手术(例如,手术切除肿瘤)、化疗、放疗、靶向癌症治疗(例如,小分子药物或特异性靶向肿瘤生长和发展中涉及的分子的单克隆抗体疗法)、和姑息治疗,或其任意组合。在一方面,针对突变Ras阳性癌症的基于酵母的免疫治疗协同吉西他滨或等价物对患者进行治疗。在一个实施方式中,患者是胰腺癌患者,该治疗包括单独或联合吉西他滨的一系列称为GI-4000或等价物的基于酵母的免疫治疗产品中的一种产品。在一方面,癌症患者的肿瘤在用基于酵母的免疫治疗组合物治疗之前已被手术切除。

[0306] 本公开使用术语“基于酵母的免疫治疗”,其短语可被可交换的与“基于酵母的免疫治疗性组合物”、“基于酵母的免疫治疗产品”、“基于酵母的免疫治疗组合物”、“基于酵母的组合物”、“基于酵母的免疫疗法”、“基于酵母的疫苗”、或这些短语的衍生物一起使用。如本文使用的基于酵母的免疫治疗组合物指一种组合物,其包括酵母载体组份和靶向对象中的疾病或状况的抗原组分(即,针对癌症的基于酵母的免疫治疗组合物包括酵母载体组份和靶向患者癌症的癌症抗原组分)。包括在本发明中使用的突变Ras抗原的基于酵母的免疫治疗组合物靶向患者的突变Ras阳性肿瘤。该组合物可被称为“酵母-Ras免疫治疗组合物”、或“表达Ras抗原的基于酵母的免疫治疗组合物”、或“针对突变Ras阳性癌症的基于酵母的免疫治疗”。针对突变Ras阳性癌症的基于酵母的免疫治疗包括但不限于一系列基于酵母的免疫治疗产品,称为“GI-4000”。

[0307] 连接酵母载体的在基于酵母的免疫治疗产品中包括的癌症抗原(例如,突变Ras抗原)最通常作为重组蛋白通过酵母载体表达(例如,通过完整酵母或酵母原生质体,其可选的可被进一步处理为酵母胞质体、酵母鬼(yeastghost)或其酵母膜提取物或碎片),虽然以下也是本发明的实施方式:一种或多种该癌症抗原被上载至酵母载体或以其他方式与酵母载体杂合、吸附、混合或与其一起施用以形成本发明使用的组合物。

[0308] 在本发明的一个方面,用于本发明的一种或多种基于酵母的免疫治疗组合物的抗原包括任何癌症或肿瘤相关抗原。在一方面,抗原包括与肿瘤出现前或增生状态相关的抗原。抗原还可以与癌症相关或称为癌症的原因。这一抗原可以是肿瘤特异性抗原、肿瘤相关抗原(TAA)或组织特异性抗原,其表位或其表位激动剂。癌症抗原包括但不限于来自任何肿瘤或癌症的抗原,包括但不限于,黑素瘤、鳞状细胞癌、乳腺癌、头颈部癌、甲状腺癌、软组织肉瘤、骨肉瘤、睾丸癌、前列腺癌、卵巢癌、膀胱癌、皮肤癌、脑癌、血管肉瘤、血管内皮瘤、肥大细胞肿瘤、白血病、淋巴瘤、主肝癌症、肺癌、胰腺癌、胃肠道癌(包括结肠直肠癌)、肾细胞癌、造血瘤及其转移性癌症。

[0309] 合适的癌症抗原包括但不限于突变Ras癌蛋白(参见,例如,美国专利号7,465,454和7,563,447)、癌胚抗原(CEA)及其表位例如CAP-1、CAP-1-6D(GenBank登录号M29540或Zaremba等,1997,Cancer Research57:4570-4577)、MART-1(Kawakami等,J.Exp.Med.180:347-352,1994)、MAGE-1(美国专利号5,750,395)、MAGE-3、GAGE(美国专利号5,648,226)、GP-100(Kawakami等,Proc.Nat'l Acad.Sci.USA91:6458-6462,1992)、MUC-1(例如,Jerome等,J.Immunol.,151:1654-1662(1993))、MUC-2、正常和突变p53癌蛋白(Hollstein等,Nucleic Acids Res.22:3551-3555,1994)、PSMA(前列腺特异性膜抗原;Israeli等,Cancer Res.53:227-230,1993)、酪氨酸酶(Kwon等,PNAS 84:7473-7477,1987)、TRP-1(gp75)

(Cohen等,Nucleic Acid Res.18:2807–2808,1990;美国专利号5,840,839)、NY-ESO-1(Chen等,PNAS 94:1914–1918,1997)、TRP-2(Jackson等,EMBO J,11:527–535,1992)、TAG72、KSA、CA-125、PSA(前列腺特异性抗原;Xue等,The Prostate,30:73–78(1997))、HER-2/neu/c-erb/B2、(美国专利号5,550,214)、EGFR(表皮生长因子受体;Harris等,Breast Cancer Res.Treat,29:1–2(1994))、hTERT、p73、B-RAF(B-Raf原癌基因丝氨酸/苏氨酸蛋白激酶;Sithanandam等,(1990),Oncogene 5(12):1775–80)、腺瘤性结肠息肉病(APC)、Myc、von Hippel-Lindau蛋白(VHL)、Rb-1、Rb-2、雄激素受体(AR)、Smad4、MDR1(也称为P-糖蛋白)、Flt-3、BRCA-1(乳腺癌1;美国专利号5,747,282)、BRCA-2(乳腺癌2;美国专利号5,747,282)、Bcr-Abl、pax3-fkhr、ews-fli-1、Brachyury(GenBank登录号NP_003172.1或NM_003181.2;Edwards等,1996,Genome Res.6:226–233)、HERV-H(人内源性逆转录病毒H)、HERV-K(人内源性逆转录病毒K)、TWIST(GenBank登录号NM_000474和NP_000465)、Mesothelin(Kojima等,1995,J.Biol.Chem.270(37):21984–90;Chang和Pastan,1996,Proc.Natl.Acad.Sci.U.S.A.93(1):136–40)、NGEP(前列腺中表达的新基因;Bera等,2004,Proc.Natl.Acad.Sci.U.S.A.101(9):3059–3064;Cereda等,2010,Cancer Immunol.Immunother.59(1):63–71;GenBank登录号AAT40139或AAT40140)、这些抗原和组织特异性抗原的修饰、这些抗原的剪接变体、和/或这些抗原的表位激动剂。其它癌症抗原是本领域已知的。其它癌症抗原也可通过本领域已知的方法进行鉴定、分离和克隆,例如那些描述于美国专利号4,514,506中的。癌症抗原也可以包括一种或多种生长因子和每个的剪接变体。

[0310] 在本发明的一个方面,癌症抗原是癌胚抗原(CEA)、包括其表位或由其表位组成的多肽例如CAP-1、CAP-1-6D(GenBank登录号M29540或Zaremba等,1997,Cancer Research 57:4570–4577)、修饰的CEA、CEA的剪接变体、这些CEA蛋白的表位激动剂、和/或包括CEA至少一种免疫原性域的融合蛋白或其表位激动剂。在一方面,CEA是修饰的CEA,其对应具有由2007年3月1日出版的美国专利公开号US 2007_0048860中的SEQ ID NO:46表示的氨基酸序列的修饰的CEA,其由该出版物中SEQ ID NO:45的核酸序列编码。

[0311] 在本发明的一个方面,该基于酵母的免疫治疗组合物靶向人鼠短尾突变体表型(Brachyury)。鼠短尾突变体表型抗原和靶向鼠短尾突变体表型的基于酵母的免疫治疗组合物已被描述于2012年9月20日出版的PCT公开号2012/125998。

[0312] 在本发明的一个方面,该基于酵母的免疫治疗组合物靶向粘蛋白-1(MUC-1)。MUC-1抗原和靶向MUC-1的基于酵母的免疫治疗组合物已被描述于2013年2月21日出版的PCT公开号2013/025972。

[0313] 在本发明的一个方面,基于酵母的免疫治疗组合物靶向突变Ras阳性癌症。Ras是一个原癌基因,其中已知几个突变发生在特定位置处且与一种或多种类型的癌症的发展有关。因此,针对突变Ras阳性癌症的基于酵母的免疫治疗产品包括至少一个Ras的免疫原区,其包含一个已知在某种癌症中突变的氨基酸残基。这些癌症包括但不限于胰腺癌、NSCLC、结肠直肠癌、子宫内膜和卵巢癌以及黑素瘤和多发性骨髓瘤。在一方面,针对突变Ras阳性癌症的基于酵母的免疫治疗产品含有两个、三个或更多Ras的免疫原区,其中每个区包含一种或多种已知在某种癌症中发生的不同的Ras突变,以覆盖几个或所有已知的Ras蛋白中的突变。例如,在本发明的一个方面,基于酵母的免疫治疗组合物中使用的Ras抗原至少包括

野生型Ras蛋白的5-9个连续氨基酸残基,其包含相对于野生型Ras蛋白的氨基酸位点12、13、59、61、73、74、75、76、77和/或78,其中位点12、13、59、61、73、74、75、76、77和/或78的氨基酸残基相对于野生型Ras蛋白是突变的。在一方面,癌症抗原包括:(a)包括野生型Ras蛋白的至少来自位点4-20或至少来自位点8-16的蛋白,除了位点12处的氨基酸残基相对于野生型Ras蛋白是突变的;(b)包括野生型Ras蛋白的至少来自位点5-21或至少来自位点9-17的蛋白,除了位点13处的氨基酸残基相对于野生型Ras蛋白是突变的;(c)包括野生型Ras蛋白的至少来自位点51-67或至少来自位点55-63的蛋白,除了位点59处的氨基酸残基相对于野生型Ras蛋白是突变的;(d)包括野生型Ras蛋白的至少来自位点53-69或至少来自位点57-65的蛋白,除了位点61处的氨基酸残基相对于野生型Ras蛋白是突变的;(e)包括野生型Ras蛋白的至少来自位点65-81或至少来自位点69-77的蛋白,除了位点73处的氨基酸残基相对于野生型Ras蛋白是突变的;(f)包括野生型Ras蛋白的至少来自位点66-82或至少来自位点70-78的蛋白,除了位点74处的氨基酸残基相对于野生型Ras蛋白是突变的;(g)包括野生型Ras蛋白的至少来自位点67-83或至少来自位点71-79的蛋白,除了位点75处的氨基酸残基相对于野生型Ras蛋白是突变的;(h)包括野生型Ras蛋白的至少来自位点69-84或至少来自位点73-81的蛋白,除了位点77处的氨基酸残基相对于野生型Ras蛋白是突变的;(i)包括野生型Ras蛋白的至少来自位点70-85或至少来自位点74-82的蛋白,除了位点78处的氨基酸残基相对于野生型Ras蛋白是突变的;和/(j)包括野生型Ras蛋白的至少来自位点68-84或至少来自位点72-80的蛋白,除了位点76处的氨基酸残基相对于野生型Ras蛋白是突变的。注意这些位点大致对应于K-Ras、N-Ras和H-Ras蛋白和人和小鼠序列,以及其它,因为人和小鼠的序列在Ras蛋白的这些区域是相同的,且因为K-Ras、N-Ras和H-Ras在Ras蛋白的这些区域是相同的。

[0314] 本文使用的术语“GI-4000”一般指一系列基于酵母的免疫治疗组合物(TARMOGEN®产品),其中每种基于酵母的免疫治疗组合物表达一种或多种Ras突变,其靶向人癌症中观察到的Ras突变。这些突变与肿瘤发展有关。作为在临床使用和本文实施例中描述的GI-4000目前包括一系列的4种基于酵母的免疫治疗产品版本,分别表示为GI-4014、GI-4015、GI-4016和GI-4020。每个版本是一种热失活的、完全的(完整的)酿酒酵母,其重组表达含有独特组合的3种Ras突变(一个突变位于相对天然Ras蛋白的位点12处,两个不同的突变位于相对天然Ras蛋白的位点61处)的融合蛋白,其一起靶向7种人类癌症中观察到的最常见的Ras突变(4种不同的位点12突变和3种不同的位点61突变)。在GI-4000临床研究中,每名患者的肿瘤被进行测序以鉴定该患者的肿瘤中包含的具体Ras突变,然后施用含有鉴定的突变蛋白的对应酵母-Ras免疫治疗产品。GI-4000系列的每种产品被单独生产和装瓶。

[0315] 更具体的,由目前临幊上使用的GI-4000产品系列(分别为GI-4014、GI-4015、GI-4016和GI-4020)中的基于酵母的免疫治疗组合物所表达的每种融合蛋白都具有大致有以下整体结构的氨基酸序列,从N端到C端:(1)两个氨基酸的N端肽序列(M-V),(2)对应Ras位点56-67的氨基酸序列,其具有对应天然Ras蛋白的位点61处的单氨基酸取代,和(3)对应Ras位点2-165的氨基酸序列,其具有分别对应天然Ras蛋白的2个位点12和61处的2个单氨基酸取代。每种融合蛋白中3个氨基酸取代的具体组合将融合蛋白彼此区分。靶向突变Ras阳性癌症的基于酵母的免疫治疗产品中的其他结构和免疫原区的组织是可能的,并详述

于,例如,美国专利号7,465,454。

[0316] 编码GI-4014的构建体的核苷酸序列和翻译的氨基酸序列分别表示为SEQ ID Nos:1和2。GI-4014包括以下Ras突变:Q61L-G12V-Q61R。编码GI-4015的构建体的核苷酸序列和翻译的氨基酸序列分别表示为SEQ ID Nos:3和4。GI-4015包括以下Ras突变:Q61L-G12C-Q61R。编码GI-4016的构建体的核苷酸序列和翻译的氨基酸序列分别表示为SEQ ID Nos:5和6。GI-4016包括以下Ras突变:Q61L-G12D-Q61R。编码GI-4020的构建体的核苷酸序列和翻译的氨基酸序列分别表示为SEQ ID Nos:7和8。GI-4020包括以下Ras突变:Q61L-G12R-Q61H。

[0317] 本发明还包括使用任一上述Ras抗原的同源物。在一方面,本发明包括使用Ras抗原,其具有与本文描述的任一Ras抗原至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%同一性的氨基酸序列,包括通过本文的具体序列标识符引用的任一Ras抗原,其全长蛋白或融合蛋白,或关于融合蛋白或定义的蛋白或其结构域(免疫原区或功能区(即具有至少一种生物活性的结构域))中一个定义的片段,其形成融合蛋白的一部分。除非另有说明,否则如本文使用的对百分比(%)同一性的引用指的是同源性的评价,其使用以下方式进行:(1)BLAST 2.0基础BLAST同源性搜索,以标准的缺省参数使用blastp进行氨基酸检索,blastn进行核酸检索,其中查询序列默认被过滤用于低复杂度区域(描述于Altschul,S.F.,Madden,T.L.,Schäffer,A.A.,Zhang,J.,Zhang,Z.,Miller,W.&Lipman,D.J.(1997)"Gapped BLAST and PSI-BLAST:a new generation of protein database search programs."Nucleic Acids Res.25:3389-3402,在此全文引入作为参考);(2)BLAST 2对齐(使用下述参数);(3)和/或使用标准缺省参数的PSI-BLAST(位点特异性迭代BLAST)。

[0318] 在本发明使用的任一基于酵母的免疫治疗组合物中,以下涉及酵母载体的方面被包括在本发明中。根据本发明,酵母载体是任何酵母细胞(例如,全细胞或完整细胞)或其衍生物(参见下文),其能被用于在本发明使用的治疗组合物中配合一种或多种抗原、其免疫原区或其表位。从而酵母载体可包括但不限于活的完整(全)酵母微生物(即,具有所有元件的酵母细胞,包括细胞壁),被杀死的(死的)或失活的完整酵母微生物,或完整/全酵母衍生物,包括:酵母原生质体(即,缺少细胞壁的酵母细胞)、酵母细胞质(即,缺少细胞壁和细胞核的酵母细胞)、酵母鬼(ghost)(即,缺少细胞壁、细胞核和细胞质的酵母细胞)、亚细胞酵母膜提取物或其片段(也称为酵母膜颗粒,之前称为亚细胞酵母颗粒)、任何其他酵母颗粒或酵母细胞壁制备物。这些酵母载体详述于例如,美国专利号5,830,463、美国专利号7,083,787和美国专利号7,736,642,其公开内容在此引入作为参考。

[0319] 可以使用任何酵母菌株来产生本发明使用的基于酵母的免疫治疗产品。可用于本发明的酵母属包括但不限于酿酒酵母(*Saccharomyces*)、假丝酵母(*Candida*)、隐球酵母(*Cryptococcus*)、汉逊酵母(*Hansenula*)、克鲁维酵母(*Kluyveromyces*)、毕赤酵母(*Pichia*)、红酵母(*Rhodotorula*)、裂殖酵母(*Schizosaccharomyces*)和耶氏酵母(*Yarrowia*)。可用于本发明的酵母种包括但不限于酿酒酵母(*Saccharomyces cerevisiae*)、卡尔酵母(*Saccharomyces carlsbergensis*)、白假丝酵母(*Candida albicans*)、乳酒假丝酵母(*Candida kefyr*)、热带假丝酵母(*Candida tropicalis*)、罗伦隐球酵母(*Cryptococcus laurentii*)、新型隐球菌(*Cryptococcus neoformans*)、异常汉逊酵

母 (*Hansenula anomala*)、多形汉逊酵母 (*Hansenula polymorpha*)、脆壁克鲁维酵母 (*Kluyveromyces fragilis*)、乳酸克鲁维酵母 (*Kluyveromyces lactis*)、变种乳酸马克思克鲁维酵母 (*Kluyveromyces marxianus* var. *lactis*)、毕赤酵母 (*Pichia pastoris*)、深红酵母 (*Rhodotorula rubra*)、粟酒裂殖酵母 (*Schizosaccharomyces pombe*) 和解脂耶氏酵母 (*Yarrowia lipolytica*)。要理解很多这些种包括大量亚种、分型、亚型等等，其也确定为被包括在前述各种之内。

[0320] 用于生产本发明使用的基于酵母的免疫治疗产品的方法之前已被描述于，例如美国专利号 5,830,463、美国专利号 7,083,787 和美国专利号 7,736,642，其公开内容在此引入作为参考。一般的，本发明使用的基于酵母的免疫治疗产品已被杀死或失活。杀死或失活酵母可通过任何一种本领域已知的多种合适的方法来实现。例如，酵母热失活是一种标准的失活酵母的方式，本领域技术人员可以通过本领域已知的标准方法来监控目标抗原的结构变化，如果需要的话。可选的，可以使用其它失活酵母的方法。例如化学法、电击法、辐射法或紫外线法。

[0321] 可使用多种本领域技术人员已知的技术将酵母载体配制入本发明的基于酵母的免疫治疗组合物或产品，包括待给对象施用的制备物。例如，可通过冻干来干燥酵母载体。也可以通过将酵母填充在蛋糕或片剂中来制备包括酵母载体的剂型，例如对用于烘焙或酿造操作的酵母所做的那样。此外，酵母载体可以与药学可接受的赋形剂混合，例如宿主或宿主细胞耐受的等渗缓冲液。这样的赋形剂的例子包括水、盐水、林格氏溶液、葡萄糖溶液、汉克溶液和其他水性质的生理平衡盐溶液。也可以使用非水载体，如固定油、芝麻油、油酸乙酯、甘油三酯。其他有用的剂型包括包含增粘剂的悬浮液，如羧甲基纤维素钠、山梨糖醇、甘油或右旋糖酐。赋形剂也可以包含少量的添加剂，如提高等渗性和化学稳定性的物质。标准的剂型可以是液体注射剂或固体，后者可以被置于一种合适的液体中作为用于注射的悬液或溶液。因此，在非液体剂型中，赋形剂可包括例如葡萄糖、人血清白蛋白和/或防腐剂，在给药前可向其中添加无菌水或盐水。组合物应该被配制为适合对人类对象给药（例如，生产条件应适于在人类中使用，且用于结束组合物和/或制备待给药的免疫治疗剂的任一赋形剂或剂型应当适于对人类使用）。在本发明的一个方面，基于酵母的免疫治疗组合物被配制用于患者或对象的注射给药，例如通过非肠道途径（如通过皮下、腹腔内、肌肉内或皮内注射，或另一种合适的非肠道途径）。

[0322] 治疗方法包括将基于酵母的免疫治疗组合物递送（给药、免疫）至对象或个体。给药过程可以体外或体内进行，但通常是体内进行。基于酵母的免疫治疗组合物的给药可以是全身性的、粘膜性的和/或邻近目标位点的位置（例如，靠近肿瘤位点）。合适的给药途径对于本领域技术人员而言是显而易见的，其取决于待预防或治疗的癌症和/或目标细胞群或组织的类型。各种可接受的给药方法包括，但不限于，静脉内给药、腹膜内给药、肌内给药、结内给药、冠状动脉内给药、动脉内给药（例如，至颈动脉内）、皮下给药、经皮递送、气管内给药、关节内给药、心室内给药、吸入（例如，气雾剂）、颅内、脊柱内、眼内、耳内、鼻内、口服、肺部给药、导管注入、和直接注射入组织。在一个方面，给药途径包括：静脉内、腹膜内、皮下、皮内、结内、肌肉内、经皮、吸入、鼻内、口服、眼内、关节内、颅内、和脊柱内。肠胃外递送可以包括皮内、肌内、腹膜内、胸膜内、肺内、静脉内、皮下、心房导管和肾 (venal) 导管的途径。耳部给药可包括滴耳液，鼻腔给药可以包括滴鼻剂或鼻内注射，眼部给药可包括眼药

水。也可以使用本领域的标准方法进行气溶胶(吸入)递送(参见,例如,Stribling等,Proc.Natl.Acad.Sci.USA189:11277-11281,1992)。在一个方面,本发明的基于酵母的免疫治疗组合物从皮下给药。在一个方面,基于酵母的免疫治疗组合物直接给药至肿瘤周围。

[0323] 通常,合适的单一剂量的基于酵母的免疫治疗组合物是在一个合适的时间段内给药一次或多次时,能够对给定的细胞类型、组织或患者身体区域以有效引发针对一种或多种癌症抗原或表位的抗原特异性免疫应答的数量有效提供酵母载体和癌症抗原的剂量。例如,在一个实施方式中,可用于本发明的单一剂量的基于酵母的免疫治疗剂为约 1×10^5 至约 5×10^7 酵母细胞当量每公斤被给药该组合物的生物体体重。在一方面,可用于本发明的单一剂量的基于酵母的免疫治疗剂为从约0.1Y.U. (1×10^6 细胞) 至约100Y.U. (1×10^9 细胞) 每剂量(即,每生物体),包括任意中间剂量,以 0.1×10^6 细胞的增量(即, 1.1×10^6 、 1.2×10^6 、 1.3×10^6 ...) 递增。本文使用的术语“Y.U.”是一个“酵母单位”或“酵母细胞当量”,其中1个Y.U.=1千万酵母细胞。在一个实施方式中,剂量包括1Y.U.-40Y.U.的剂量、1Y.U.-50Y.U.的剂量、1Y.U.-60Y.U.的剂量、1Y.U.-70Y.U.的剂量、或1Y.U.-80Y.U.的剂量,在一方面,为10Y.U.-40Y.U.、50Y.U.、60Y.U.、70Y.U.、或80Y.U.。在一个实施方式中,该剂量在个体的不同位置但在同一时间段给药。例如,40Y.U.的剂量可通过在一个给药阶段注射10Y.U.的剂量至个体的4个不同位置处进行给药,或20Y.U.的剂量可通过在同一个给药阶段注射5Y.U.的剂量至个体的4个不同位置处、或注射10Y.U.的剂量至个体的2个不同位置处进行给药。本发明包括在个体的1、2、3、4、5、6、7、8、9、10或更多个不同位置给药一个数量的基于酵母的免疫治疗组合物(例如,1、2、3、4、5、6、7、8、9、10或更多)来形成一个单一剂量。

[0324] 当例如针对抗原的免疫应答减弱时或根据需要时,施用基于酵母的免疫治疗组合物的“增强剂(Boosters)”或“增强”,以针对特定抗原或抗原组提供免疫应答或记忆应答。增强剂可在初始给药后间隔约1、2、3、4、5、6、7或8个周、至每月、至双月、至每季、至每年、至几年进行给药。在一个实施方式中,给药时间表是其中单一剂量在从几周、至几个月、至几年的时间段内给药至少1、2、3、4、5、6、7、8、9、10或更多次。在一个实施方式中,药剂每周给药1、2、3、4、5、6、7、8、9、10或更多剂量,然后根据获得期望的治疗效果的需要而每月给药。即使患者的肿瘤复发或在患者被相信已缓和之后,也可以给药额外的剂量。

[0325] 在一方面,个体用至少一种其它可用于癌症治疗(抗癌治疗)的治疗化合物或治疗方案进行治疗。可用于癌症治疗的额外的试剂、组合物或方案(例如,治疗方案)包括但不限于化疗、手术切除肿瘤、放疗、同种异体或自体干细胞移植、细胞因子疗法、继承性T细胞转移、和/或给药第二免疫治疗组合物(例如,额外的基于酵母的免疫疗法、基于重组病毒的免疫疗法(病毒性载体)、细胞因子治疗、免疫刺激剂治疗(包括具有免疫刺激属性的化疗)、DNA疫苗和其它免疫治疗组合物和/或靶向癌症治疗(例如,小分子药物、生物制剂或特异性针对参与肿瘤的生长和发展的分子的单克隆抗体治疗,包括但不限于,选择性雌激素受体调节剂(SERMs)、芳香酶抑制剂、酪氨酸激酶抑制剂、丝氨酸/苏氨酸激酶抑制剂、组蛋白去乙酰酶(HDAC)抑制剂、类维生素A受体活化剂、凋亡刺激剂、血管生成抑制剂、聚(ADP-核糖)聚合酶(PARP)抑制剂或免疫刺激剂)。可以在本发明的免疫治疗组合物之前、同时、交替或之后,或在不同时间点施用这些额外的治疗剂和/或治疗方案。例如,当配合化疗或靶向癌症治疗对一名个体给药时,可以在化疗或靶向癌症治疗的多个剂量之间的“假期”中施用基

于酵母的免疫治疗组合物,以最大化免疫治疗组合物的效果。手术切除肿瘤可经常先于施用基于酵母的免疫治疗组合物,但是额外的或初级的手术也可以在施用基于酵母的免疫治疗组合物过程中或之后发生。

[0326] 例如,在本文描述的发明的治疗方法的任一实施方式中,在一方面,当个体患有癌症时,该个体在进行治疗或已被另一种癌症疗法治疗过。这些治疗可包括本文之前描述的任何治疗方案或使用任何治疗化合物或试剂,包括但不限于,化疗、放疗、靶向癌症治疗、手术切除肿瘤、干细胞移植、细胞因子疗法、继承性T细胞转移、和/或给药第二免疫治疗组合物。在给药第二免疫治疗组合物的情形中,该组合物可包括但不限于额外的基于酵母的免疫治疗、基于重组病毒的免疫疗法(病毒性载体)、免疫刺激剂治疗(包括具有免疫刺激属性的化疗)、DNA疫苗和其它免疫治疗组合物)。

[0327] 如本文使用的,“治疗”癌症或其任意排列(例如,“癌症治疗”,等)一般指一旦癌症已经发生(例如,一旦癌症在个体中被诊断或检测到),则施用本发明的基于酵母的免疫治疗组合物,其具有治疗的至少一种治疗目的(与不存在该治疗相比),包括:减少肿瘤负担;抑制肿瘤的生长;增加个体的无复发生存率;增加个体的总存活数;推迟、抑制、阻止或预防转移性癌症的发生或发展(例如通过延迟、抑制、阻止或预防肿瘤转移的发展的发生和/或肿瘤入侵初级癌症以外的组织和/或其他与转移性癌症的进展有关的过程);延迟或阻止癌症的发展;提高对肿瘤的免疫应答;改善针对肿瘤抗原的长期记忆免疫应答;和/或改善个体的整体健康。“预防”或“防止”癌症,或其任意排列(例如,“癌症预防”等)一般指在癌症发生前或癌症中癌症或肿瘤抗原表达的具体阶段发生前施用本发明的组合物,其具有治疗的至少一种目的(与不存在该治疗相比),包括:预防或延缓癌症的发生或发展,或者如果在治疗后发生癌症,至少减轻癌症的严重程度(例如,减少肿瘤的生长水平、阻止癌症进展、改善对癌症的免疫应答、抑制转移过程)或改善个体的结果(例如,改善无复发生存率和/或总存活数)。

[0328] 在本发明的治疗方法中,基于酵母的免疫治疗组合物和其它抗癌疗法可对任何动物施用,包括任何脊椎动物,特别是任何脊椎动物类、哺乳纲的成员,包括但不限于灵长类、啮齿类、家畜和宠物。“个体”是脊椎动物,例如哺乳动物,包括但不限于人类。术语“个体”可以与术语“动物”、“对象”或“患者”替换使用。

[0329] 根据本发明,术语“抗原”在本文中的使用指的是:蛋白(肽、部分蛋白、全长蛋白)的任何部分,其中蛋白是天然产生或合成来源的;细胞组合物(全细胞、细胞裂解物或破碎的细胞);生物体(完整生物体、裂解物或破碎的细胞);或碳水化合物或其它分子或其部分。抗原可诱发针对同一抗原或免疫系统的一个元件(例如,T细胞、抗体)遇到过的相似抗原的抗原特异性免疫应答(例如体液和/或细胞介导的免疫应答)。

[0330] 抗原可以小至一个单一表位、一个单一免疫原区或更大,并可包括多个表位或免疫原区。从而,抗原的大小可以小至约8-12个氨基酸(即,一个肽),和大至:全长蛋白、多聚体、融合蛋白、嵌合蛋白、全细胞、全微生物、或其任意部分(例如,全细胞裂解物或微生物提取物)。此外,抗原可包括碳水化合物,其可被加载至酵母载体或本发明的组合物中。将理解在一些实施方式中(例如当抗原通过酵母载体从重组核酸分子中表达时),抗原是蛋白、融合蛋白、嵌合蛋白或其片段,而非整细胞或微生物。

[0331] 当要在酵母中表达抗原时,抗原具有能在酵母中重组表达的最小大小,通常为至

少或多于25个氨基酸的长度,或至少或多于26、至少或多于27、至少或多于28、至少或多于29、至少或多于30、至少或多于31、至少或多于32、至少或多于33、至少或多于34、至少或多于35、至少或多于36、至少或多于37、至少或多于38、至少或多于39、至少或多于40、至少或多于41、至少或多于42、至少或多于43、至少或多于44、至少或多于45、至少或多于46、至少或多于47、至少或多于48、至少或多于49、或至少或多于50个氨基酸的长度,或至少25–50个氨基酸的长度,至少30–50个氨基酸的长度,或至少35–50个氨基酸的长度,或至少40–50个氨基酸的长度,或至少45–50个氨基酸的长度。可以表达较小的蛋白,也可以表达相当大的蛋白(例如,几百个氨基酸的长度或甚至数千个氨基酸的长度)。在一方面,可以表达全长蛋白,或其结构或功能域,或其免疫原区,其从N端和/或C端缺少一个或多个氨基酸(例如,从N端和/或C端缺少约1个至约20个氨基酸)。融合蛋白和嵌合蛋白也是可在本发明中表达的抗原。“靶抗原”是由本发明的免疫治疗组合物特异性靶向的抗原(即,期望针对其引发免疫应答的抗原)。例如,“Ras抗原”是一种从一种或多种Ras蛋白来源、设计或生产的抗原,从而靶向该抗原还靶向对应的由肿瘤表达的Ras蛋白。“突变Ras抗原”特别指含有一种或多种氨基酸突变的Ras抗原。为了用于本发明,这些突变对应于在肿瘤中的Ras蛋白中发现的突变,且关联肿瘤发展和/或肿瘤进展。

[0332] 当提及刺激免疫应答时,术语“免疫原”是术语“抗原”的子集,从而,在一些情况下,其可被与术语“抗原”可交换的使用。本文使用的免疫原描述了一种抗原,其引发体液和/或细胞介导的免疫应答(即,是免疫原性的),从而对个体施用免疫原诱发(mount)针对同一抗原或个体免疫系统遇到过的相似抗原的抗原特异性免疫应答。在一个实施方式中,基于酵母的免疫治疗组合物中包含的免疫原诱发细胞介导的免疫应答,包括CD4+T细胞应答(例如,TH1、TH2和/或TH17)和/或CD8+T细胞应答(例如,CTL应答)。

[0333] 给定抗原的“免疫原区”可以是抗原的任意部分、片段或表位(例如,肽片段或亚单位或抗体表位或其它构像表位),其含有至少一个在对动物给药时作为免疫原的表位。因此,免疫原区大于一个单一氨基酸,且至少具有足以含有一个可作为免疫原的表位的大小。例如,单一蛋白可含有多个不同免疫原区。免疫原区不必是蛋白中的线性序列,例如在体液免疫应答的情形中,其中考虑构象区。

[0334] 本文将表位定义为给定抗原内的单一免疫原性位点,其在合适的共刺激信号和/或免疫系统的活化细胞的背景下对免疫系统提供时足以诱发免疫应答。换而言之,表位是抗原的一部分,其事实上被免疫系统的元件所识别,还可被称为抗原决定簇。本领域技术人员将认可T细胞表位的大小不同,且来自B细胞或抗体表位和通过I型MHC途径呈递的表位的组合物在大小和结构属性方面不同于通过II型MHC途径呈递的表位。例如,通过I型MHC分子呈递的T细胞表位通常为8–11个氨基酸的长度,而通过II型MHC分子呈递的表位在长度方面较少限制,可以是从8个氨基酸一直到25个氨基酸或更长。此外,T细胞表位根据该表位结合的具体MHC分子具有预测的结构特性。表位可以是线性序列表位或构像表位(保守结合区)。大多数抗体识别构像表位。

[0335] 虽然已经详细描述了本发明的各种实施方式,但是显然从这些实施方式的修饰和改编会被本领域技术人员所想到。然而,将清楚地理解,这些修饰和改编是在本发明的范围之内的,如附加的权利要求中所示。

序列表

<110> 比奥德希克斯股份有限公司 (Biodesix, Inc.)
 全球免疫股份有限公司 (GlobalImmune, Inc.)
 Roder, Joanna
 Roder, Heinrich

<120> 用于选择和不选择以产生免疫应答的疗法治疗的癌症患者的质谱方法

<130> 12-621-PRO

<160> 8

<170> PatentIn version 3.5

<210> 1

<211> 537

<212> DNA

<213> 人工的

<220>

[0001] <223> 合成的构建体

<220>

<221> CDS

<222> (1)..(537)

<400> 1

atg gtc ctc gac aca gca ggt ttg gag gag tac agt gca atg act gag
 Met Val Leu Asp Thr Ala Gly Leu Glu Glu Tyr Ser Ala Met Thr Glu
 1 5 10 15

48

tat aaa ctt gtg gtg gtt gga gct gtt ggc gta ggc aag agc gcc ttg
 Tyr Lys Leu Val Val Val Gly Ala Val Gly Val Gly Lys Ser Ala Leu
 20 25 30

96

acg ata cag cta att cag aat cac ttt gtg gat gag tac gac cct acg
 Thr Ile Gln Leu Ile Gln Asn His Phe Val Asp Glu Tyr Asp Pro Thr
 35 40 45

144

ata gag gac tcc tac agg aaa caa gta gta att gat gga gaa acc tgt
 Ile Glu Asp Ser Tyr Arg Lys Gln Val Val Ile Asp Gly Glu Thr Cys
 50 55 60

192

| | | | | |
|--|-----|-----|-----|-----|
| ctc ttg gat att ctc gac aca gca ggt cga gag gag tac agt gca atg Leu Leu Asp Ile Leu Asp Thr Ala Gly Arg Glu Glu Tyr Ser Ala Met 65 | 70 | 75 | 80 | 240 |
| agg gac cag tac atg aga act ggg gag ggc ttt ctt tgt gta ttt gcc Arg Asp Gln Tyr Met Arg Thr Gly Glu Gly Phe Leu Cys Val Phe Ala 85 | 90 | 95 | | 288 |
| ata aat aat act aaa tca ttt gaa gat att cac cat tat aga gaa caa Ile Asn Asn Thr Lys Ser Phe Glu Asp Ile His His Tyr Arg Glu Gln 100 | 105 | | 110 | 336 |
| att aaa aga gta aag gac tct gaa gat gtg cct atg gtc ctg gta ggg Ile Lys Arg Val Lys Asp Ser Glu Asp Val Pro Met Val Leu Val Gly 115 | 120 | 125 | | 384 |
| aat aag tgt gat ttg cct tct aga aca gta gac acg aaa cag gct cag Asn Lys Cys Asp Leu Pro Ser Arg Thr Val Asp Thr Lys Gln Ala Gln 130 | 135 | 140 | | 432 |
| [0002] gag tta gca agg agt tac ggg att ccg ttc att gag acc tca gca aag Glu Leu Ala Arg Ser Tyr Gly Ile Pro Phe Ile Glu Thr Ser Ala Lys 145 | 150 | 155 | 160 | 480 |
| aca aga cag ggt gtt gac gat gcc ttc tat aca tta gtc cga gaa att Thr Arg Gln Gly Val Asp Asp Ala Phe Tyr Thr Leu Val Arg Glu Ile 165 | 170 | 175 | | 528 |
| cga aaa tag Arg Lys | | | | 537 |

<210> 2
<211> 178
<212> PRT
<213> 人工的

<220>
<223> 合成的构建体

<400> 2

Met Val Leu Asp Thr Ala Gly Leu Glu Glu Tyr Ser Ala Met Thr Glu
 1 5 10 15

Tyr Lys Leu Val Val Val Gly Ala Val Gly Val Gly Lys Ser Ala Leu
 20 25 30

Thr Ile Gln Leu Ile Gln Asn His Phe Val Asp Glu Tyr Asp Pro Thr
 35 40 45

Ile Glu Asp Ser Tyr Arg Lys Gln Val Val Ile Asp Gly Glu Thr Cys
 50 55 60

Leu Leu Asp Ile Leu Asp Thr Ala Gly Arg Glu Glu Tyr Ser Ala Met
 65 70 75 80

[0003] Arg Asp Gln Tyr Met Arg Thr Gly Glu Gly Phe Leu Cys Val Phe Ala
 85 90 95

Ile Asn Asn Thr Lys Ser Phe Glu Asp Ile His His Tyr Arg Glu Gln
 100 105 110

Ile Lys Arg Val Lys Asp Ser Glu Asp Val Pro Met Val Leu Val Gly
 115 120 125

Asn Lys Cys Asp Leu Pro Ser Arg Thr Val Asp Thr Lys Gln Ala Gln
 130 135 140

Glu Leu Ala Arg Ser Tyr Gly Ile Pro Phe Ile Glu Thr Ser Ala Lys
 145 150 155 160

Thr Arg Gln Gly Val Asp Asp Ala Phe Tyr Thr Leu Val Arg Glu Ile
 165 170 175

Arg Lys

<210> 3

<211> 537

<212> DNA

<213> 人工的

<220>

<223> 合成的构建体

<220>

<221> CDS

<222> (1),, (\$37)

<400> 3

| | |
|---|----|
| atg gtc ctc gac aca gca ggt ttg gag gag tac agt gca atg act gag | 48 |
| Met Val Leu Asp Thr Ala Gly Leu Glu Glu Tyr Ser Ala Met Thr Glu | |
| 1 5 10 15 | |

[0004]

| | |
|---|----|
| tat aaa ctt gtg gtg gtt gga gct tgt ggc gta ggc aag agc gcc ttg | 96 |
| Tyr Lys Leu Val Val Val Gly Ala Cys Gly Val Gly Lys Ser Ala Leu | |
| 20 25 30 | |

| | |
|---|-----|
| acg ata cag cta att cag aat cac ttt gtg gat gag tac gac cct acg | 144 |
| Thr Ile Gin Leu Ile Gln Asn His Phe Val Asp Glu Tyr Asp Pro Thr | |
| 35 40 45 | |

| | |
|---|-----|
| ata gag gac tcc tac agg aaa caa gta gta att gat gga gaa acc tgt | 192 |
| Ile Glu Asp Ser Tyr Arg Lys Gln Val Val Ile Asp Gly Glu Thr Cys | |
| 50 55 60 | |

| | |
|---|-----|
| ctc ttg gat att ctc gac aca gca ggt cga gag gag tac agt gca atg | 240 |
| Leu Leu Asp Ile Leu Asp Thr Ala Gly Arg Glu Glu Tyr Ser Ala Met | |
| 65 70 75 80 | |

| | |
|---|-----|
| agg gac cag tac atg aga act ggg gag ggc ttc ctt tgt gta ttt gcc | 288 |
| Arg Asp Gln Tyr Met Arg Thr Gly Glu Gly Phe Leu Cys Val Phe Ala | |
| 85 90 95 | |

| | |
|---|-----|
| ata aat aat act aaa tca ttt gaa gat att cac cat tat aga gaa caa | 336 |
| Ile Asn Asn Thr Lys Ser Phe Glu Asp Ile His His Tyr Arg Glu Gln | |

| 100 | 105 | 110 | |
|--|-----|-----|-----|
| att aaa aga gta aag gac tct gaa gat gtg cct atg gtc ctg gta ggg Ile Lys Arg Val Lys Asp Ser Glu Asp Val Pro Met Val Leu Val Gly | | | 384 |
| 115 | 120 | 125 | |
| aat aag tgt gat ttg cct tct aga aca gta gac acg aaa cag get cag Asn Lys Cys Asp Leu Pro Ser Arg Thr Val Asp Thr Lys Gln Ala Gln | | | 432 |
| 130 | 135 | 140 | |
| gag tta gca agg agt tac ggg att ccg ttc att gag acc tca gca aag Glu Leu Ala Arg Ser Tyr Gly Ile Pro Phe Ile Glu Thr Ser Ala Lys | | | 480 |
| 145 | 150 | 155 | 160 |
| aca aga cag ggt gtt gac gat gcc ttc tat aca tta gtc cga gaa att Thr Arg Gln Gly Val Asp Asp Ala Phe Tyr Thr Leu Val Arg Glu Ile | | | 528 |
| 165 | 170 | 175 | |
| cga aaa tag Arg Lys | | | 537 |

[0005]

<210> 4
<211> 178
<212> PRT
<213> 人工的

<220>
<223> 合成的构建体

<400> 4

| | | | |
|---|---|----|----|
| Met Val Leu Asp Thr Ala Gly Leu Glu Glu Tyr Ser Ala Met Thr Glu | | | |
| 1 | 5 | 10 | 15 |

| | | | |
|---|----|----|--|
| Tyr Lys Leu Val Val Val Gly Ala Cys Gly Val Gly Lys Ser Ala Leu | | | |
| 20 | 25 | 30 | |

| | | | |
|---|----|----|--|
| Thr Ile Gln Leu Ile Gln Asn His Phe Val Asp Glu Tyr Asp Pro Thr | | | |
| 35 | 40 | 45 | |

Ile Glu Asp Ser Tyr Arg Lys Gln Val Val Ile Asp Gly Glu Thr Cys
 50 55 60

Leu Leu Asp Ile Leu Asp Thr Ala Gly Arg Glu Glu Tyr Ser Ala Met
 65 70 75 80

Arg Asp Gln Tyr Met Arg Thr Gly Glu Gly Phe Leu Cys Val Phe Ala
 85 90 95

Ile Asn Asn Thr Lys Ser Phe Glu Asp Ile His His Tyr Arg Glu Gln
 100 105 110

Ile Lys Arg Val Lys Asp Ser Glu Asp Val Pro Met Val Leu Val Gly
 115 120 125

[0006] Asn Lys Cys Asp Leu Pro Ser Arg Thr Val Asp Thr Lys Gln Ala Gln
 130 135 140

Glu Leu Ala Arg Ser Tyr Gly Ile Pro Phe Ile Glu Thr Ser Ala Lys
 145 150 155 160

Thr Arg Gln Gly Val Asp Asp Ala Phe Tyr Thr Leu Val Arg Glu Ile
 165 170 175

Arg Lys

<210> 5

<211> 537

<212> DNA

<213> 人工的

<220>

<223> 合成的构建体

<220>
<221> CDS
<222> (1).. (537)

| | |
|--|--|
| <pre> <400> 5 atg gtc ctc gac aca gca ggt ttg gag gag tac agt gca atg act gag Met Val Leu Asp Thr Ala Gly Leu Glu Glu Tyr Ser Ala Met Thr Glu 1 5 10 15 tat aaa ctt gtg gtg gtt gga gct gat ggc gta ggc aag agc gcc ttg Tyr Lys Leu Val Val Val Gly Ala Asp Gly Val Gly Lys Ser Ala Leu 20 25 30 acg ata cag cta att cag aat cac ttt gtg gat gag tac gac cct acg Thr Ile Gln Leu Ile Gln Asn His Phe Val Asp Glu Tyr Asp Pro Thr 35 40 45 ata gag gac tcc tac agg aaa caa gta gta att gat gga gaa acc tgt Ile Glu Asp Ser Tyr Arg Lys Gln Val Val Ile Asp Gly Glu Thr Cys 50 55 60 [0007] ctc ttg gat att ctc gac aca gca ggt cga gag gag tac agt gca atg Leu Leu Asp Ile Leu Asp Thr Ala Gly Arg Glu Glu Tyr Ser Ala Met 65 70 75 80 agg gac cag tac atg aga act ggg gag ggc ttt ctt tgt gta ttt gcc Arg Asp Gln Tyr Met Arg Thr Gly Glu Gly Phe Leu Cys Val Phe Ala 85 90 95 ata aat aat act aaa tca ttt gaa gat att cac cat tat aga gaa caa Ile Asn Asn Thr Lys Ser Phe Glu Asp Ile His His Tyr Arg Glu Gln 100 105 110 att aaa aga gta aag gac tct gaa gat gtg cct atg gtc ctg gta ggg Ile Lys Arg Val Lys Asp Ser Glu Asp Val Pro Met Val Leu Val Gly 115 120 125 aat aag tgt gat ttg cct tct aga aca gta gac acg aaa cag gct cag Asn Lys Cys Asp Leu Pro Ser Arg Thr Val Asp Thr Lys Gln Ala Gln 130 135 140 gag tta gca agg agt tac ggg att ccg ttc att gag acc tca gca aag </pre> | 48 96 144 192 240 288 336 384 432 480 |
|--|--|

Glu Leu Ala Arg Ser Tyr Gly Ile Pro Phe Ile Glu Thr Ser Ala Lys
 145 150 155 160
 aca aga cag ggt gtt gac gat gcc ttc tat aca tta gtc cga gaa att 528
 Thr Arg Gln Gly Val Asp Asp Ala Phe Tyr Thr Leu Val Arg Glu Ile
 165 170 175
 cga aaa tag 537
 Arg Lys

<210> 6
 <211> 178
 <212> PRT
 <213> 人工的

<220>
 <223> 合成的构建体

<400> 6

[0008] Met Val Leu Asp Thr Ala Gly Leu Glu Glu Tyr Ser Ala Met Thr Glu
 1 5 10 15

Tyr Lys Leu Val Val Val Gly Ala Asp Gly Val Gly Lys Ser Ala Leu
 20 25 30

Thr Ile Gln Leu Ile Gln Asn His Phe Val Asp Glu Tyr Asp Pro Thr
 35 40 45

Ile Glu Asp Ser Tyr Arg Lys Gln Val Val Ile Asp Gly Glu Thr Cys
 50 55 60

Leu Leu Asp Ile Leu Asp Thr Ala Gly Arg Glu Glu Tyr Ser Ala Met
 65 70 75 80

Arg Asp Gln Tyr Met Arg Thr Gly Glu Gly Phe Leu Cys Val Phe Ala
 85 90 95

Ile Asn Asn Thr Lys Ser Phe Glu Asp Ile His His Tyr Arg Glu Gln
 100 105 110

Ile Lys Arg Val Lys Asp Ser Glu Asp Val Pro Met Val Leu Val Gly
 115 120 125

Asn Lys Cys Asp Leu Pro Ser Arg Thr Val Asp Thr Lys Gln Ala Gln
 130 135 140

Glu Leu Ala Arg Ser Tyr Gly Ile Pro Phe Ile Glu Thr Ser Ala Lys
 145 150 155 160

Thr Arg Gln Gly Val Asp Asp Ala Phe Tyr Thr Leu Val Arg Glu Ile
 165 170 175

[0009] Arg Lys

<210> 7

<211> 537

<212> DNA

<213> 人工的

<220>

<223> 合成的构建体

<220>

<221> CDS

<222> (1)..(537)

<400> 7

| | |
|---|----|
| atg gtc ctc gac aca gca ggt ttg gag gag tac agt gca atg act gag | 48 |
| Met Val Leu Asp Thr Ala Gly Leu Glu Glu Tyr Ser Ala Met Thr Glu | |
| 1 5 10 15 | |

| | | | | |
|---|-----|-----|-----|-----|
| tat aaa ctt gtg gtc gtt gga gct cgt ggc gta ggc aag agc gcc ttg Tyr Lys Leu Val Val Val Gly Ala Arg Gly Val Gly Lys Ser Ala Leu | 20 | 25 | 30 | 96 |
| acg ata cag cta att cag aat cac ttt gtg gat gag tac gac ect acg Thr Ile Gln Leu Ile Gln Asn His Phe Val Asp Glu Tyr Asp Pro Thr | 35 | 40 | 45 | 144 |
| ata gag gac tcc tac agg aaa caa gta gta att gat gga gaa acc tgt Ile Glu Asp Ser Tyr Arg Lys Gln Val Val Ile Asp Gly Glu Thr Cys | 50 | 55 | 60 | 192 |
| ctc ttg gat att ctc gac aca gca ggt cac gag gag tac agt gca atg Leu Leu Asp Ile Leu Asp Thr Ala Gly His Glu Glu Tyr Ser Ala Met | 65 | 70 | 75 | 240 |
| agg gac cag tac atg aga act ggg gag ggc ttt ctt tgt gta ttt gcc Arg Asp Gln Tyr Met Arg Thr Gly Glu Gly Phe Leu Cys Val Phe Ala | 85 | 90 | 95 | 288 |
| [0010] ata aat aat act aaa tca ttt gaa gat att cac cat tat aga gaa caa Ile Asn Asn Thr Lys Ser Phe Glu Asp Ile His His Tyr Arg Glu Gln | 100 | 105 | 110 | 336 |
| att aaa aga gta aag gac tct gaa gat gtg cct atg gtc ctg gta ggg Ile Lys Arg Val Lys Asp Ser Glu Asp Val Pro Met Val Leu Val Gly | 115 | 120 | 125 | 384 |
| aat aag tgt gat ttg cct tct aga aca gta gac acg aaa cag gct cag Asn Lys Cys Asp Leu Pro Ser Arg Thr Val Asp Thr Lys Gln Ala Gln | 130 | 135 | 140 | 432 |
| gag tta gca agg agt tac ggg att ccg ttc att gag acc tca gca aag Glu Leu Ala Arg Ser Tyr Gly Ile Pro Phe Ile Glu Thr Ser Ala Lys | 145 | 150 | 155 | 480 |
| aca aga cag ggt gtt gac gat gcc ttc tat aca tta gtc cga gaa att Thr Arg Gln Gly Val Asp Asp Ala Phe Tyr Thr Leu Val Arg Glu Ile | 165 | 170 | 175 | 528 |
| cga aaa tag Arg Lys | | | | 537 |

<210> 8

<211> 178

<212> PRT

<213> 人工的

<220>

<223> 合成的构建体

<400> 8

Met Val Leu Asp Thr Ala Gly Leu Glu Glu Tyr Ser Ala Met Thr Glu

1 5 10 15

Tyr Lys Leu Val Val Gly Ala Arg Gly Val Gly Lys Ser Ala Leu

20 25 30

Thr Ile Gin Leu Ile Gin Asn His Phe Val Asp Glu Tyr Asp Pro Thr

35 40 45

[0011]

Ile Glu Asp Ser Tyr Arg Lys Gln Val Val Ile Asp Gly Glu Thr Cys

50 55 60

Leu Leu Asp Ile Leu Asp Thr Ala Gly His Glu Glu Tyr Ser Ala Met

65 70 75 80

Arg Asp Gln Tyr Met Arg Thr Gly Glu Gly Phe Leu Cys Val Phe Ala

85 90 95

Ile Asn Asn Thr Lys Ser Phe Glu Asp Ile His His Tyr Arg Glu Gln

100 105 110

Ile Lys Arg Val Lys Asp Ser Glu Asp Val Pro Met Val Leu Val Gly

115 120 125

Asn Lys Cys Asp Leu Pro Ser Arg Thr Val Asp Thr Lys Gln Ala Gln

130 135 140

Glu Leu Ala Arg Ser Tyr Gly Ile Pro Phe Ile Glu Thr Ser Ala Lys
145 150 155 160

[0012]

Thr Arg Gln Gly Val Asp Asp Ala Phe Tyr Thr Leu Val Arg Glu Ile
165 170 175

Arg Lys

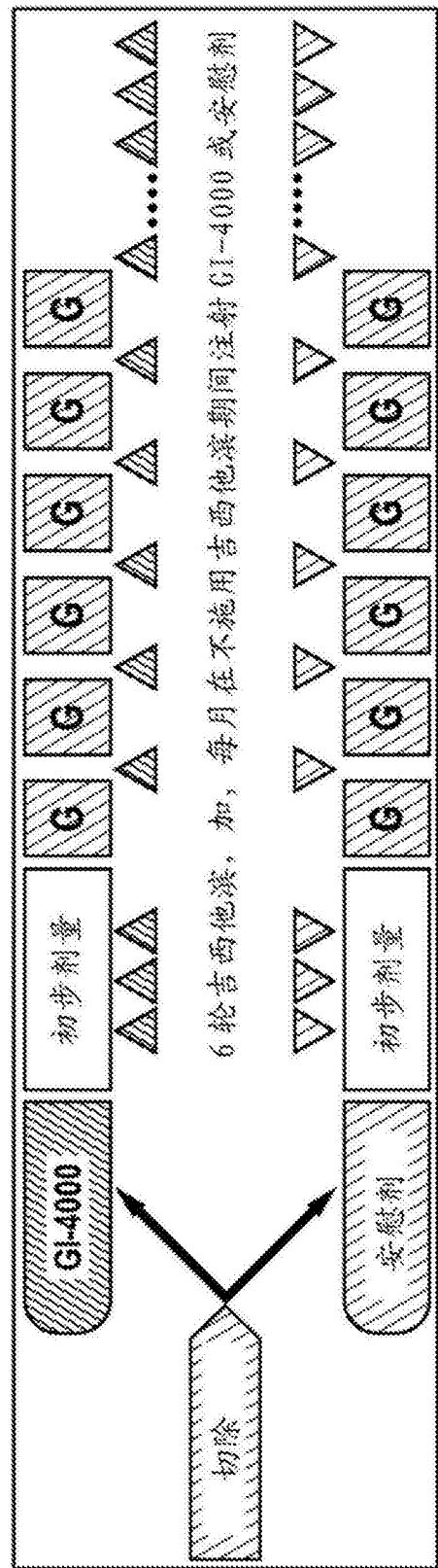


图1

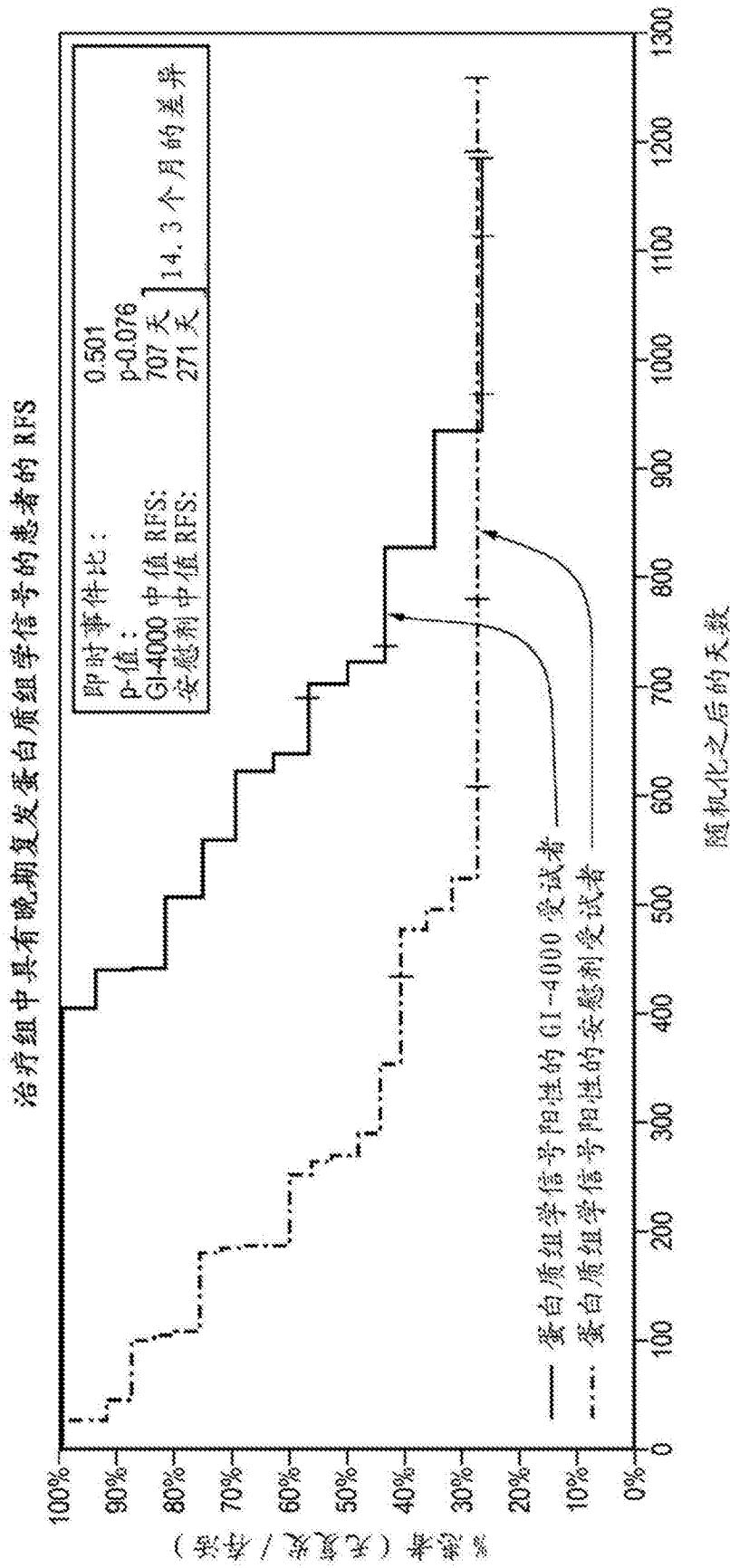


图 2A

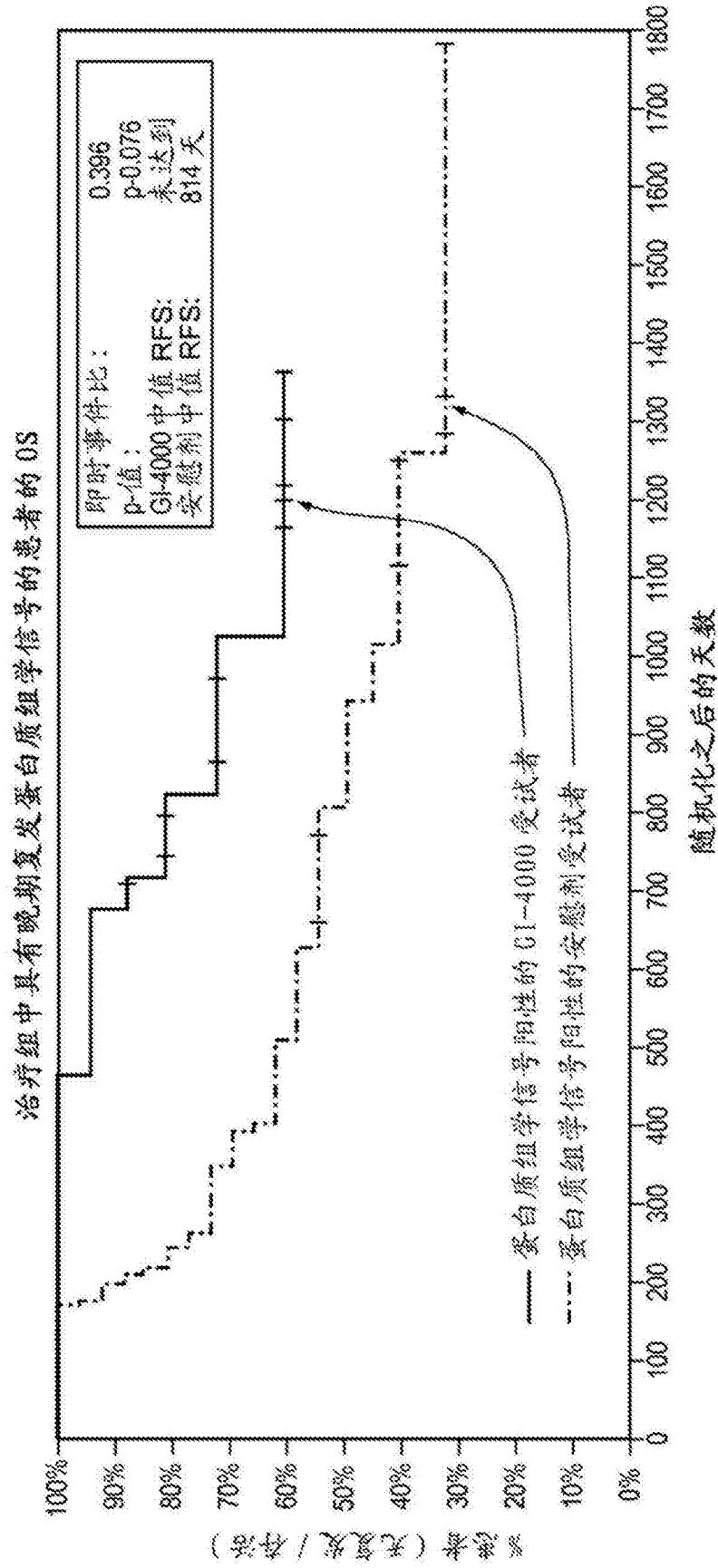


图 2B

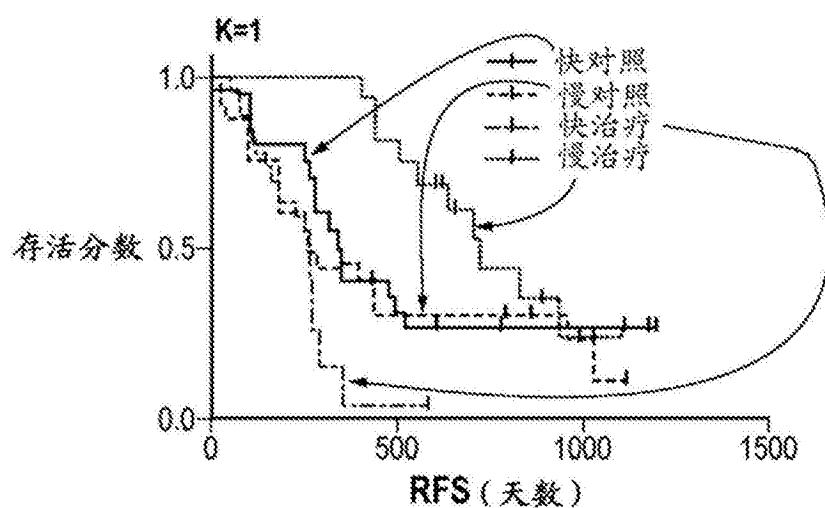


图3A

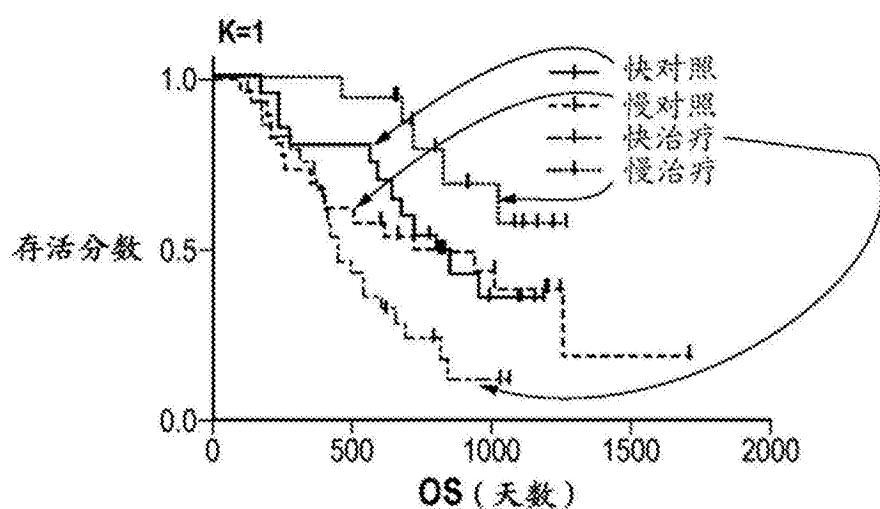


图3B

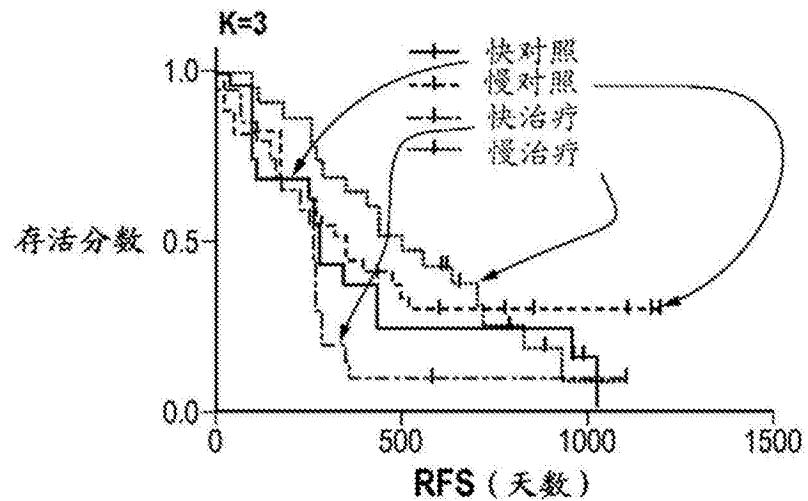


图3C

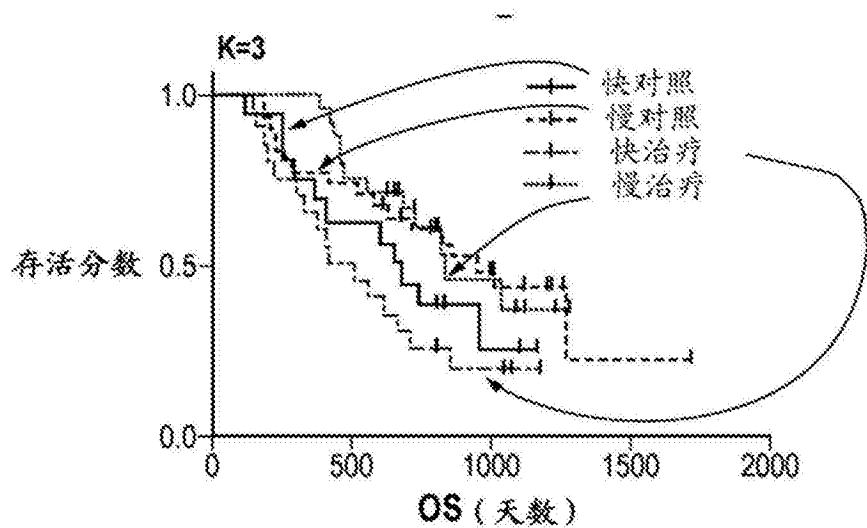


图3D

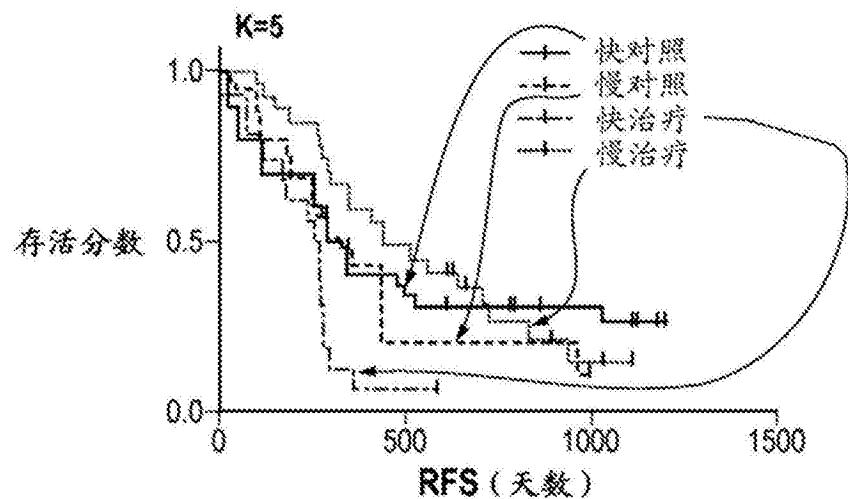


图3E

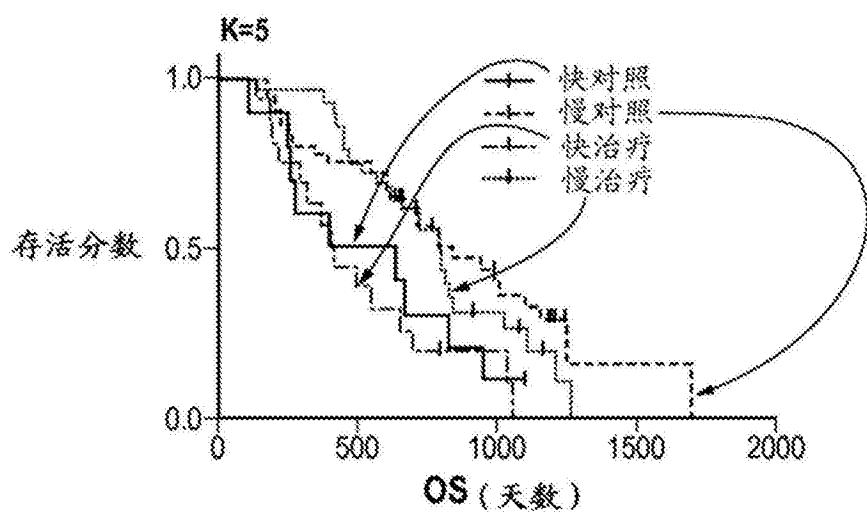


图3F

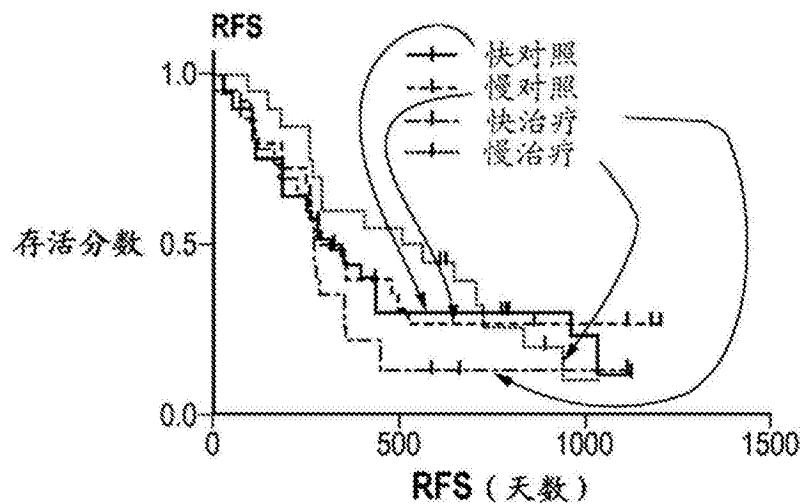


图4A

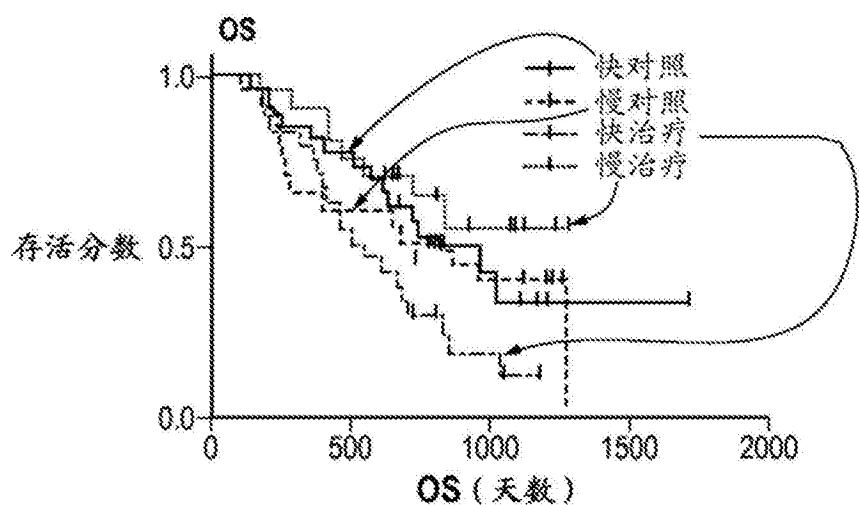


图4B

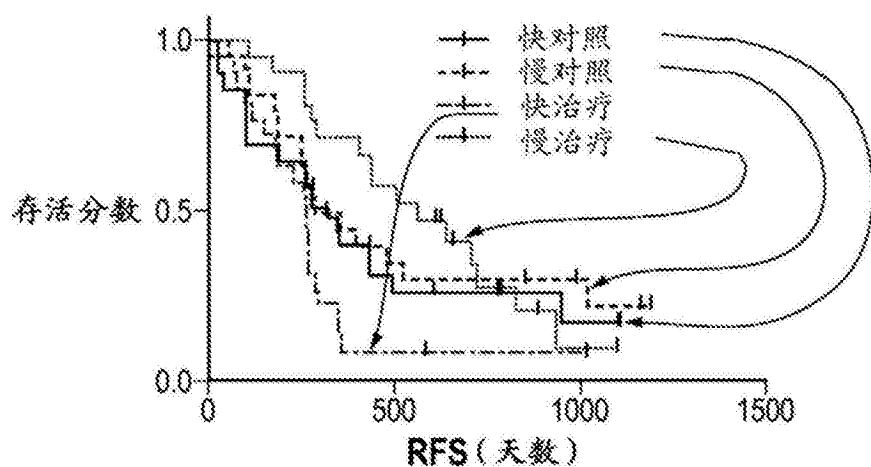


图4C

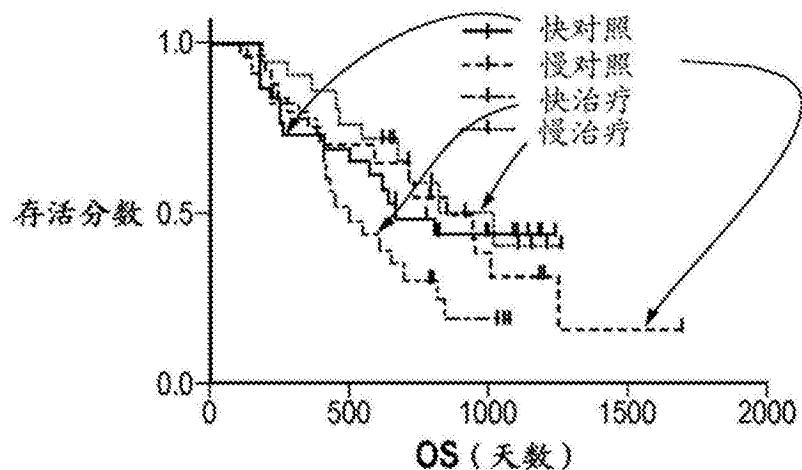


图4D

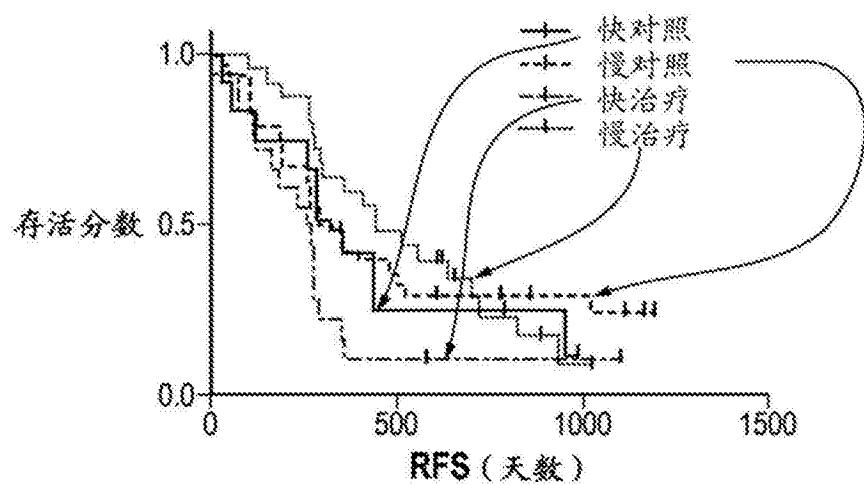


图4E

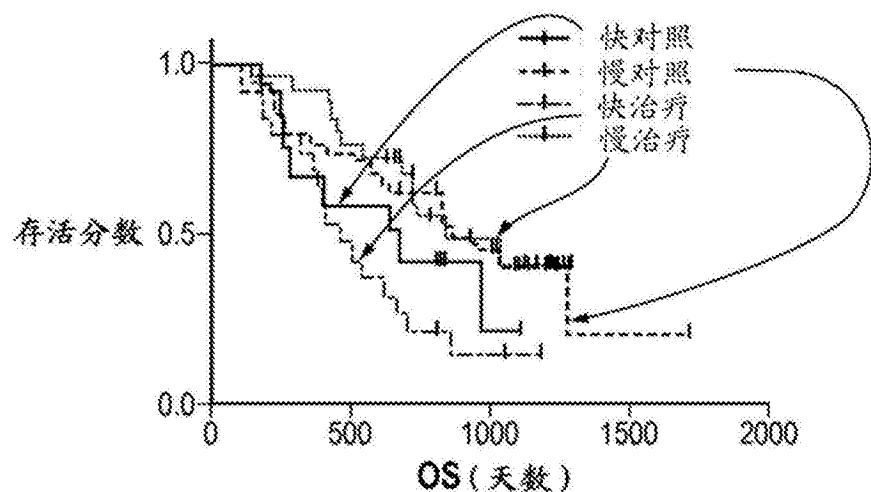


图4F

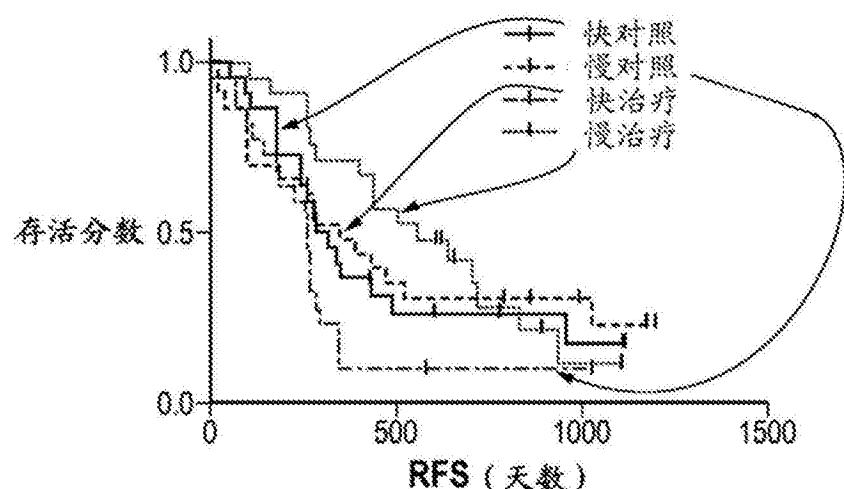


图4G

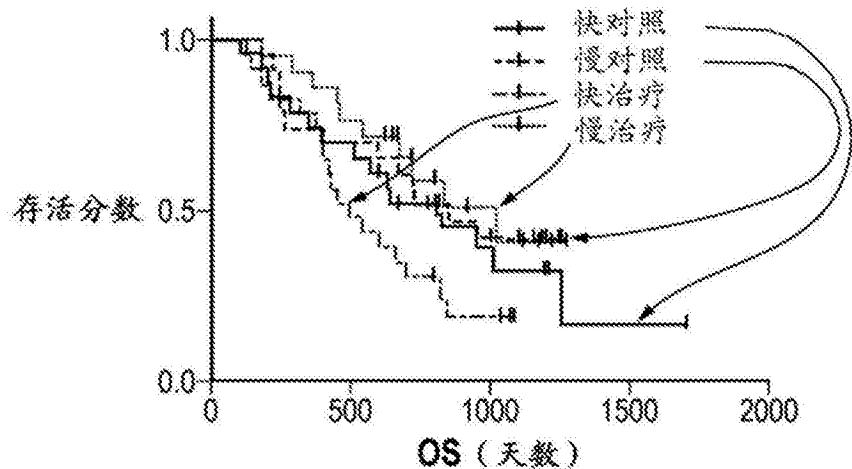


图4H

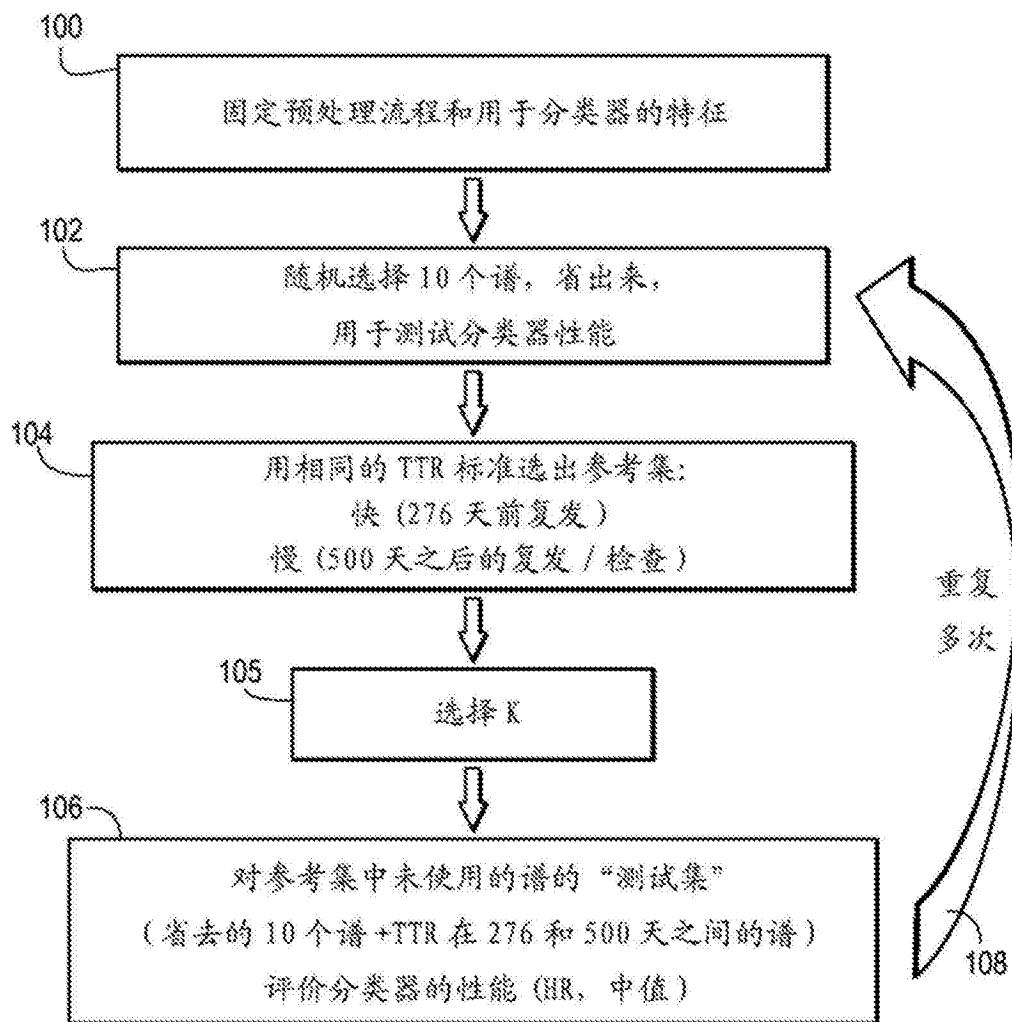


图5

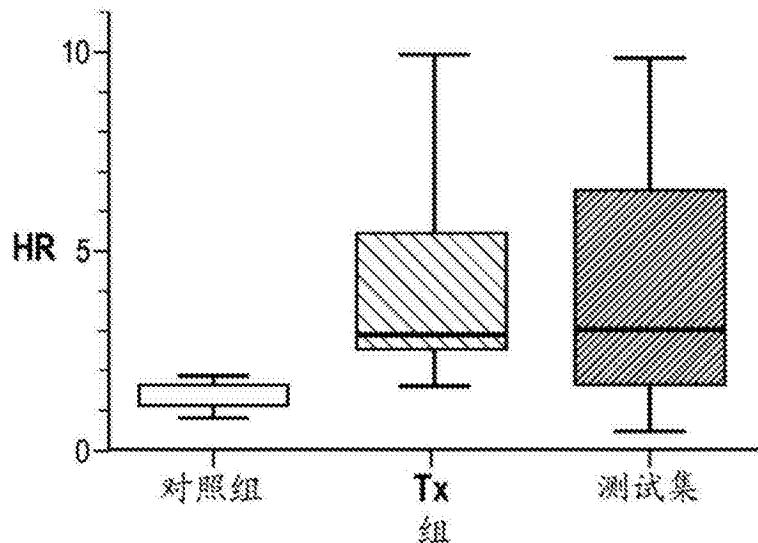


图6

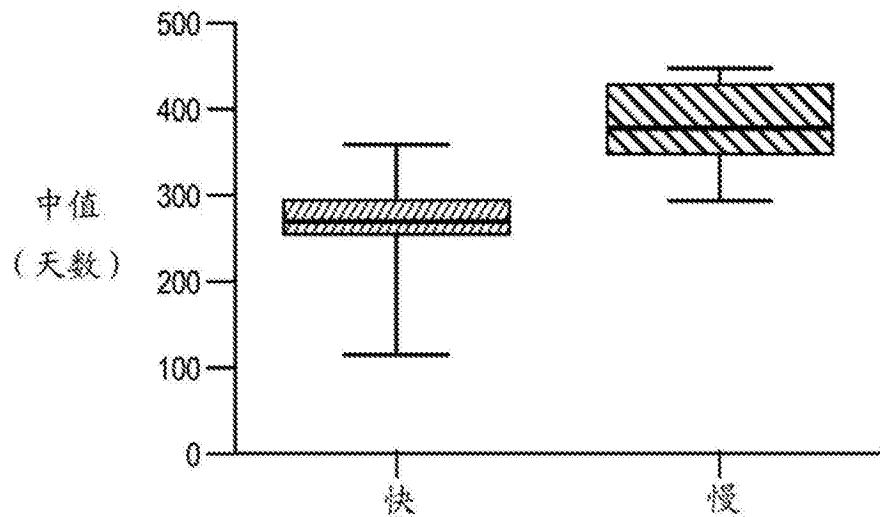


图7A

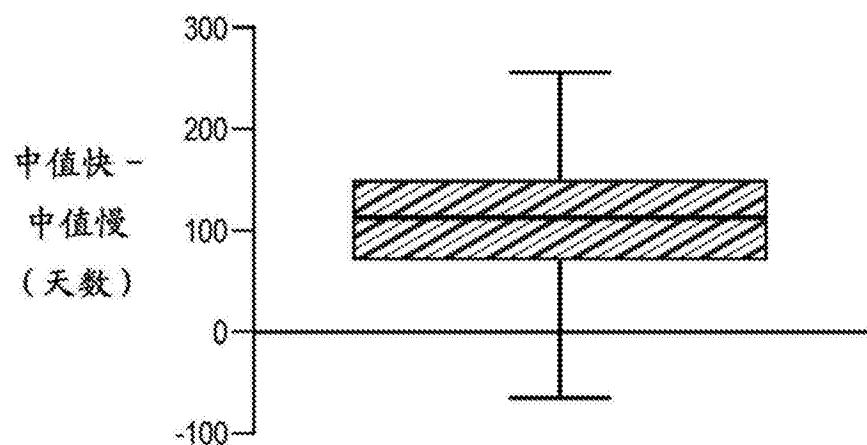


图7B

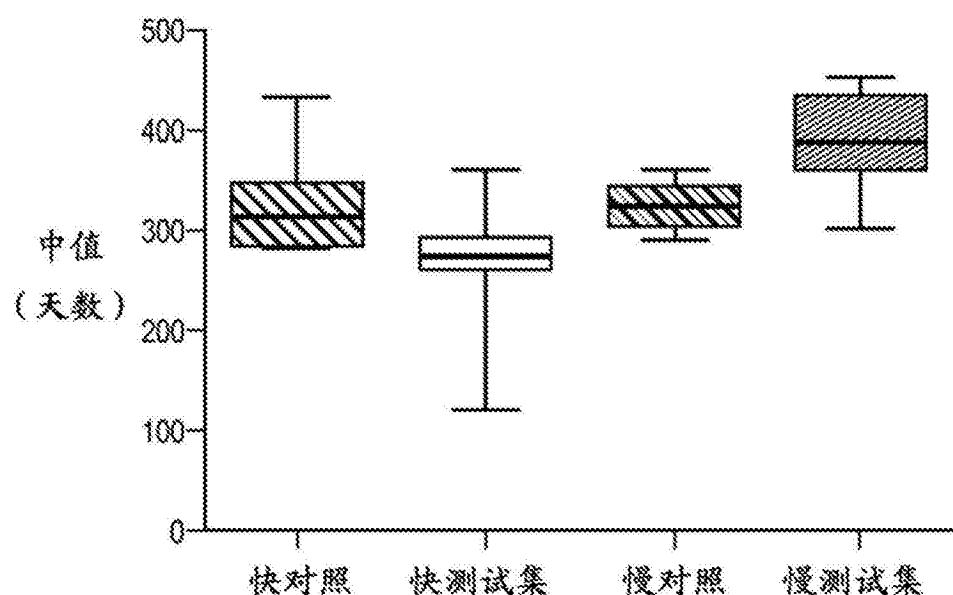


图8

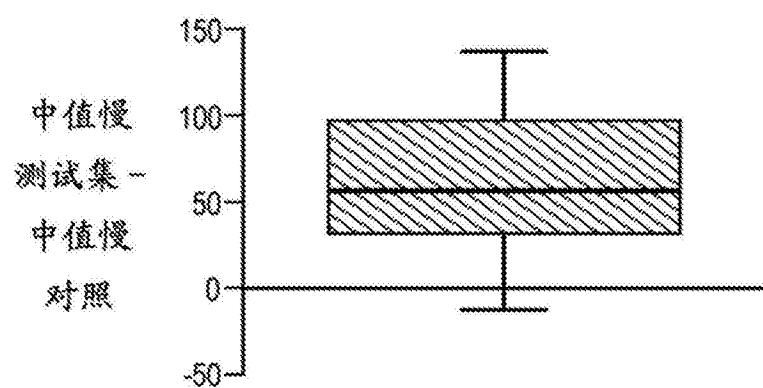


图9

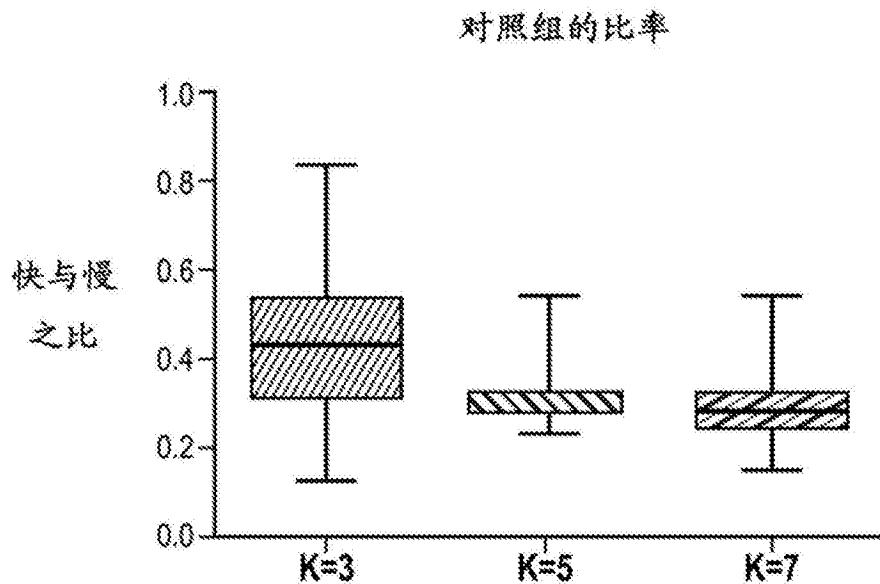
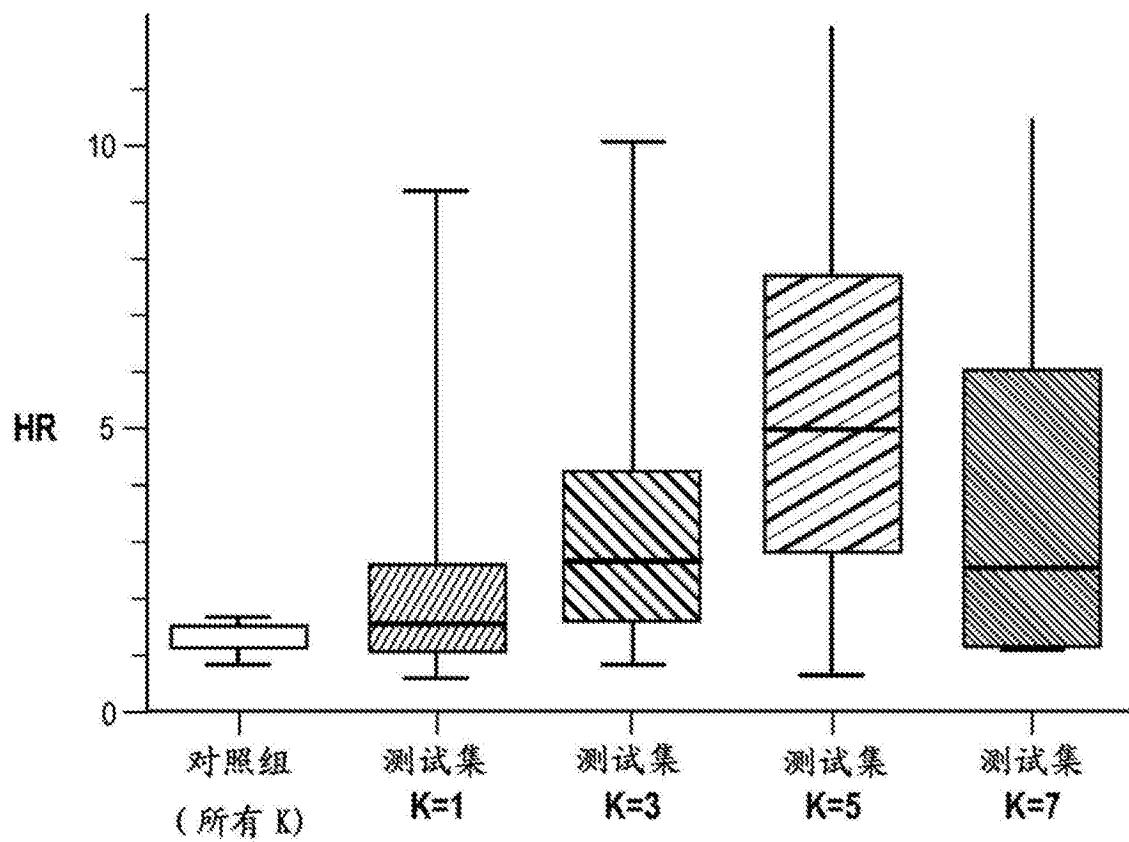


图10A



交叉验证分析的测试集的即时事件比，
不同的 K 值

图10B

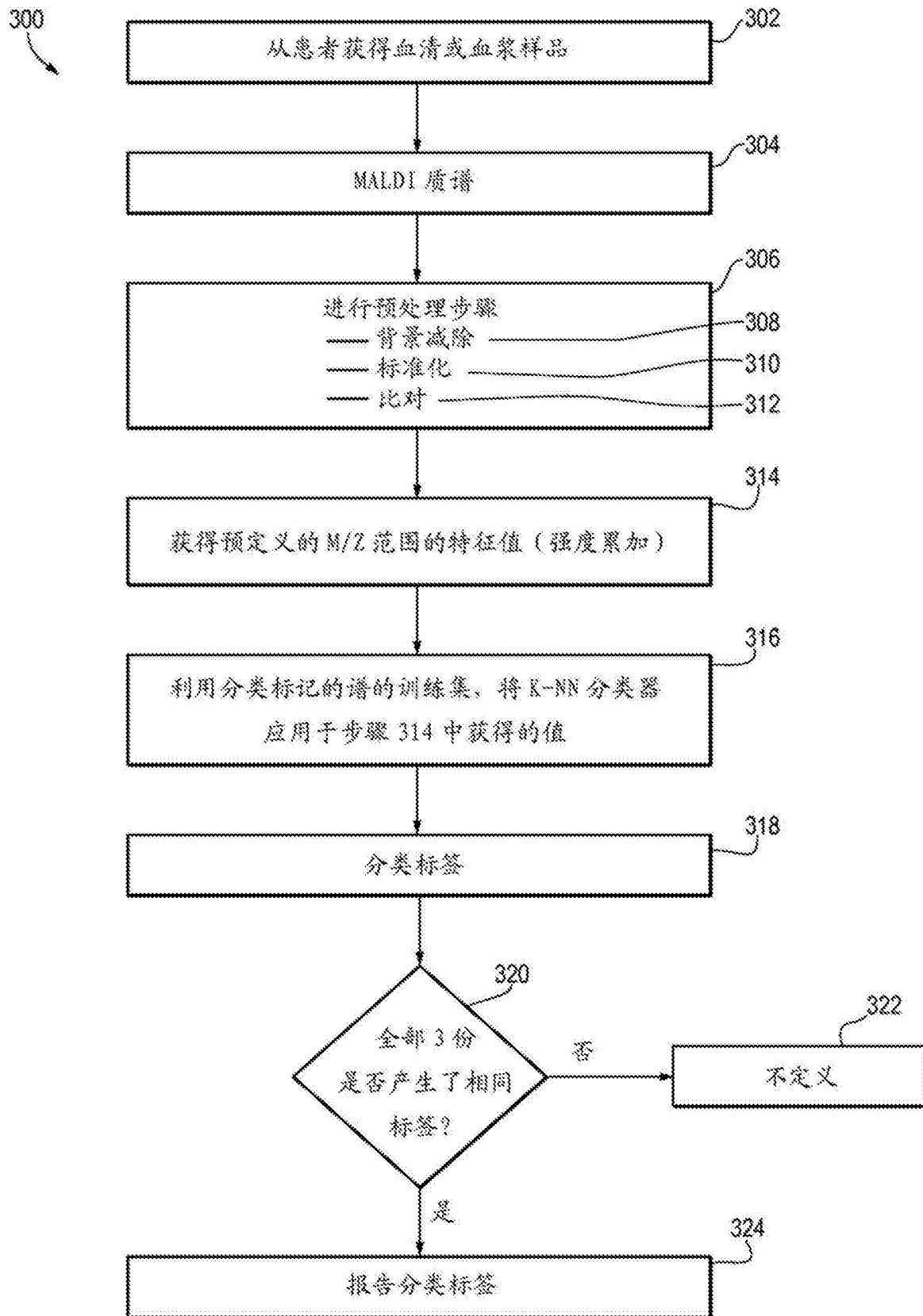


图11

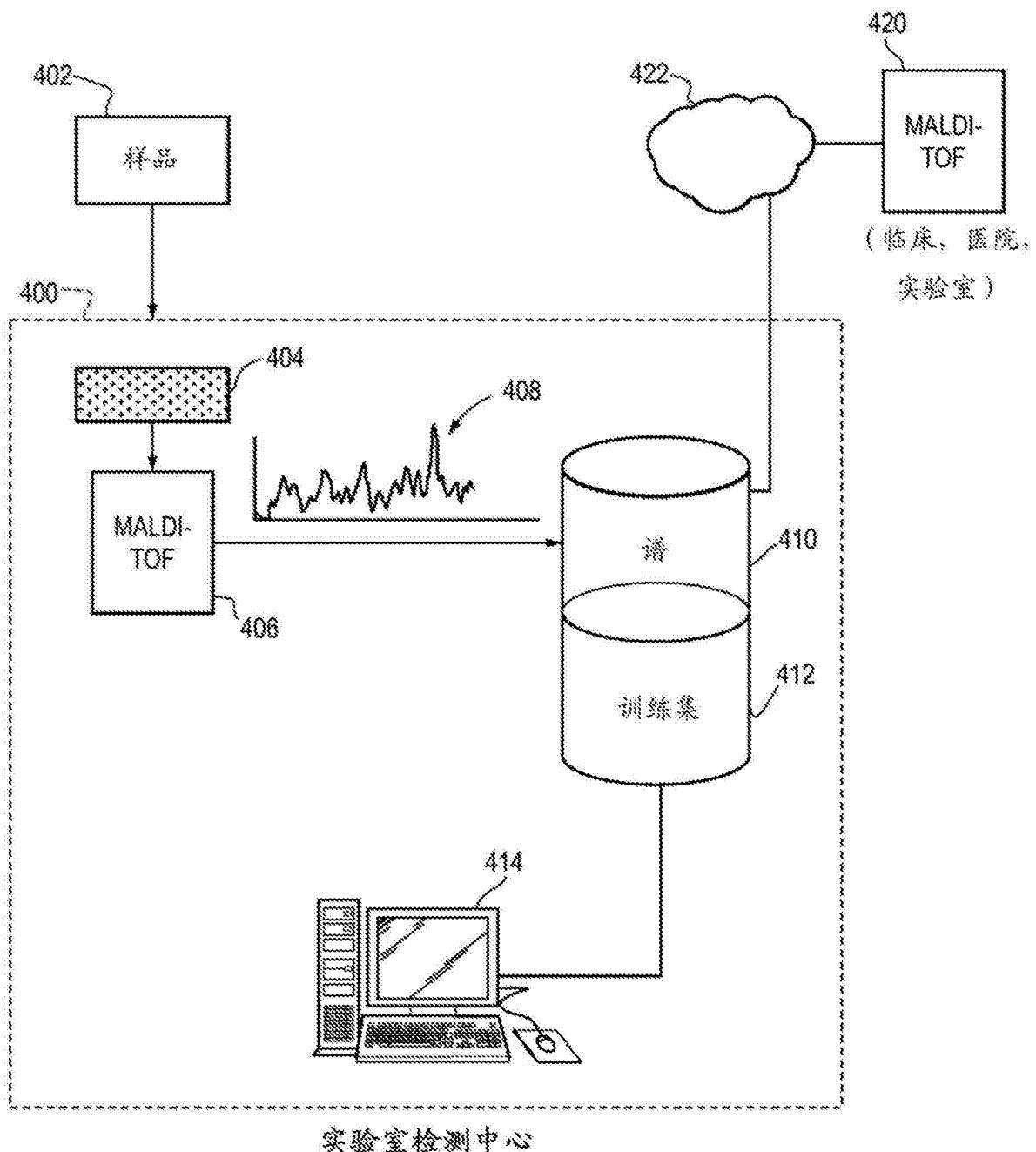


图12

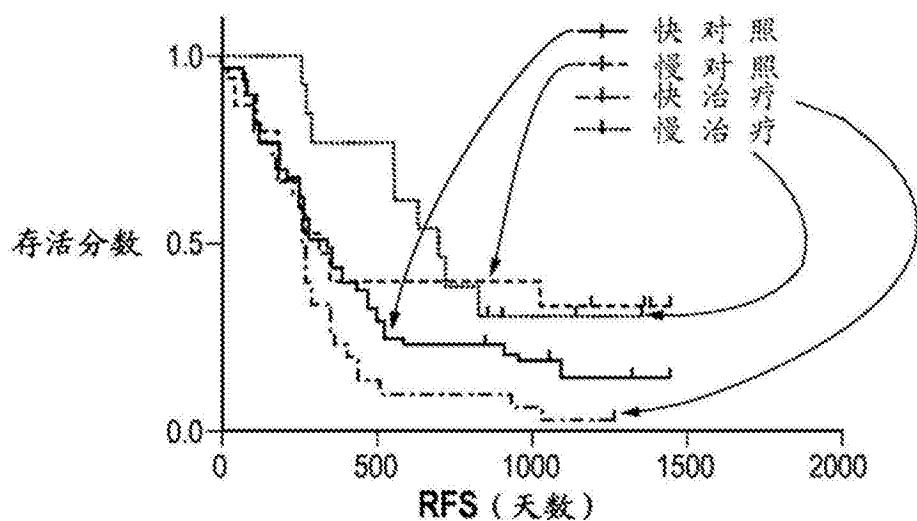


图13A

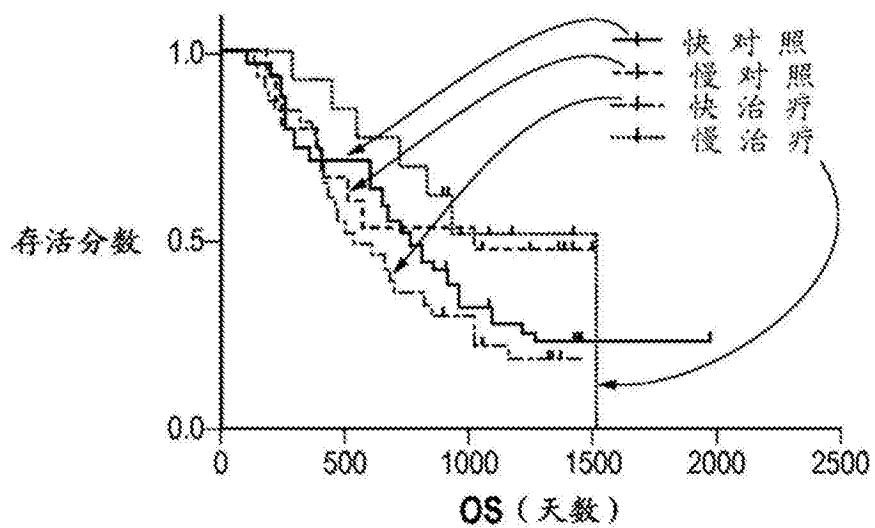


图13B

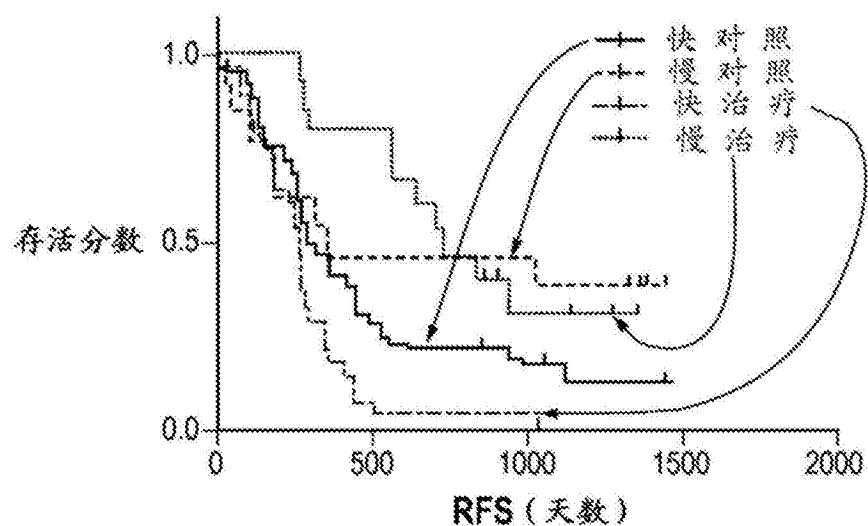


图13C

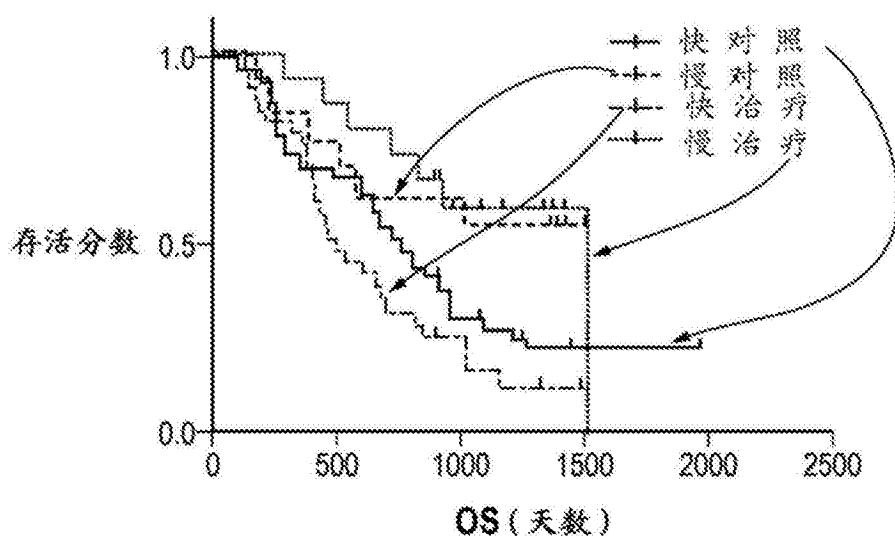


图13D

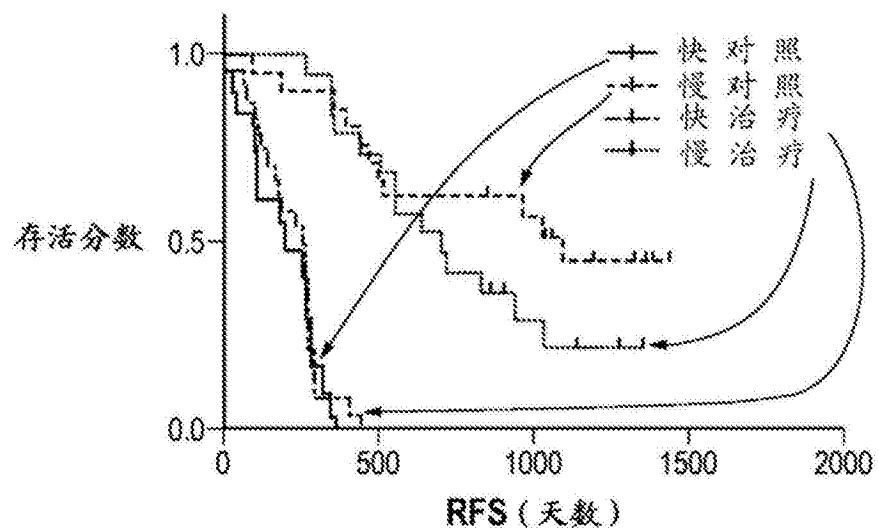


图13E

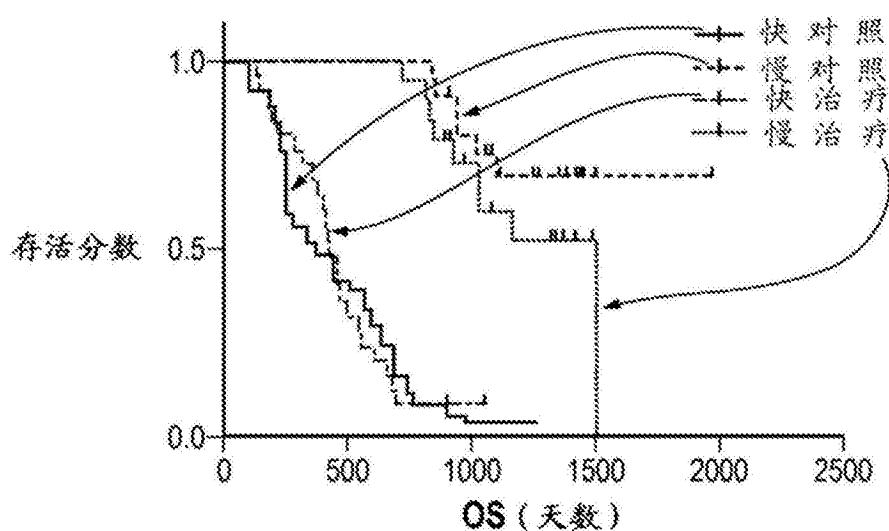


图13F

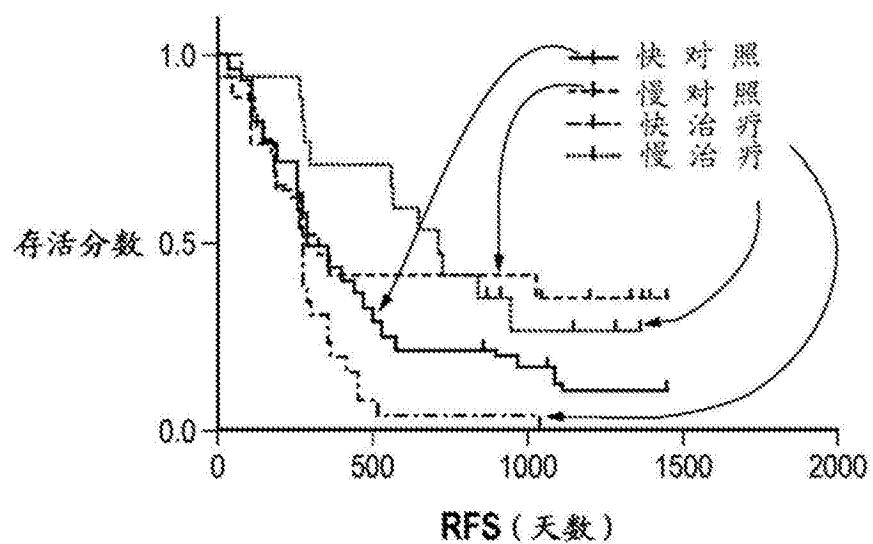


图13G

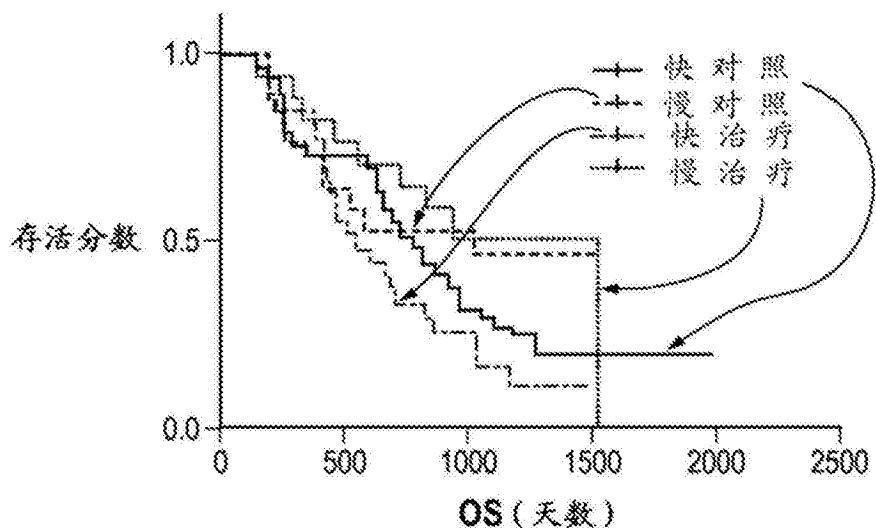


图13H

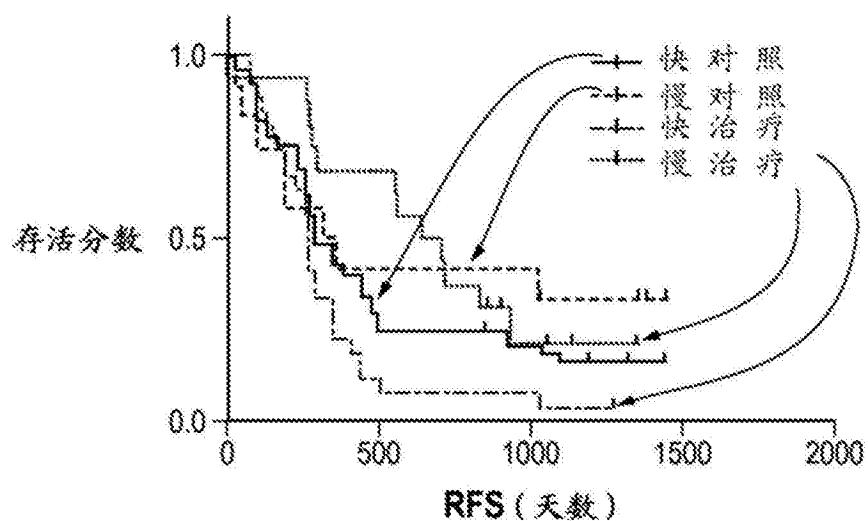


图13I

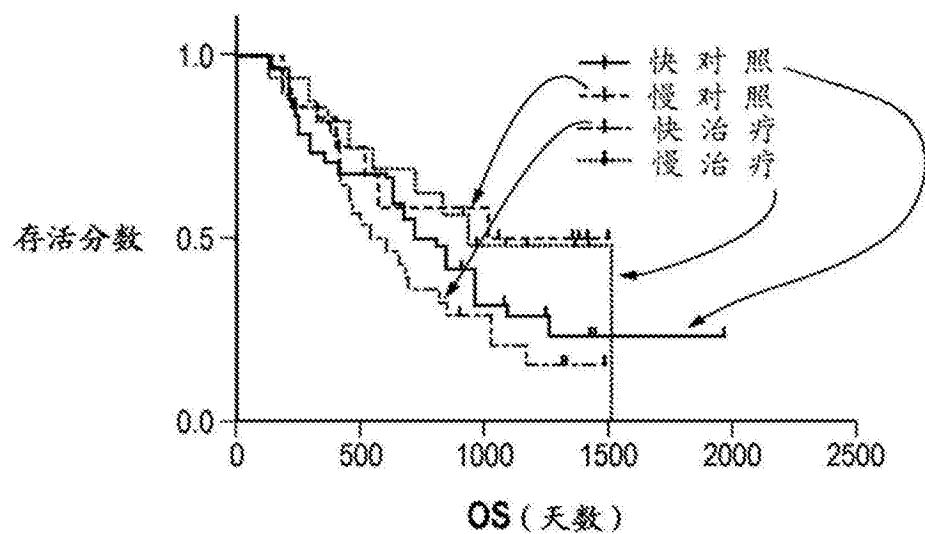


图13J

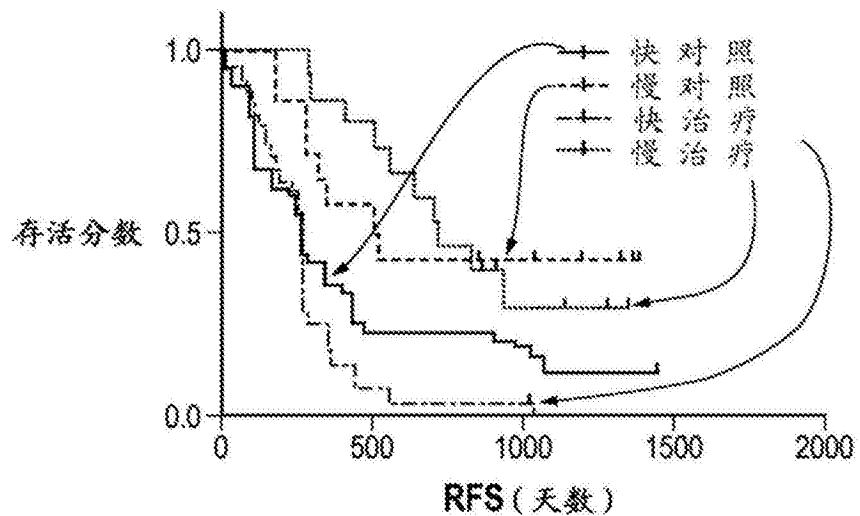


图13K

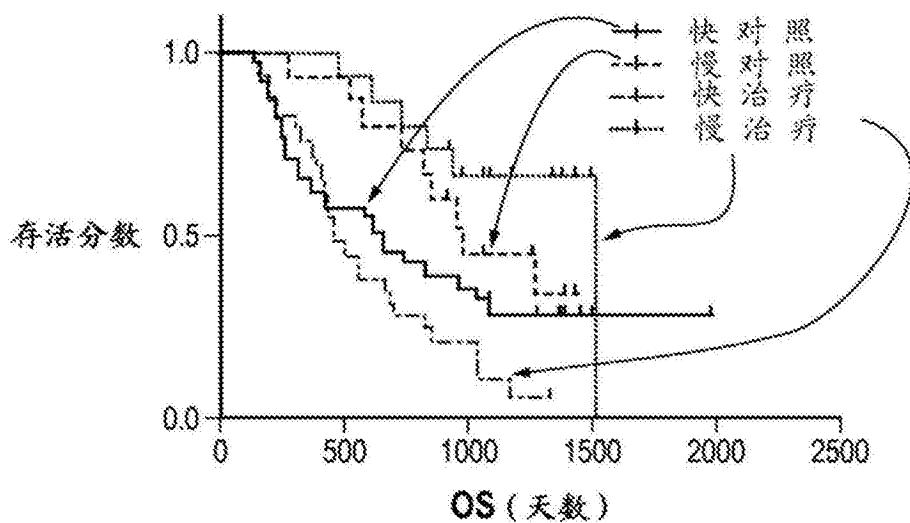


图13L

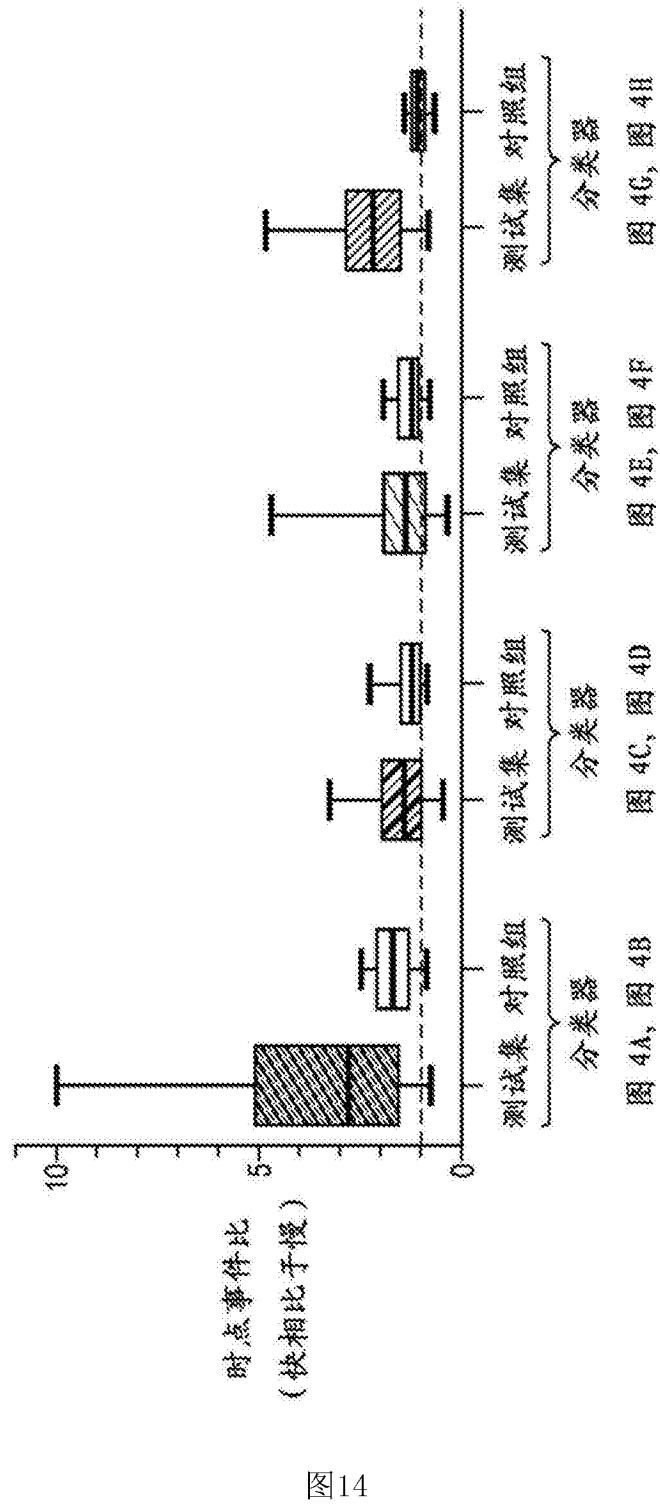


图14

图 4A, 图 4B 图 4C, 图 4D 图 4E, 图 4F 图 4G, 图 4H

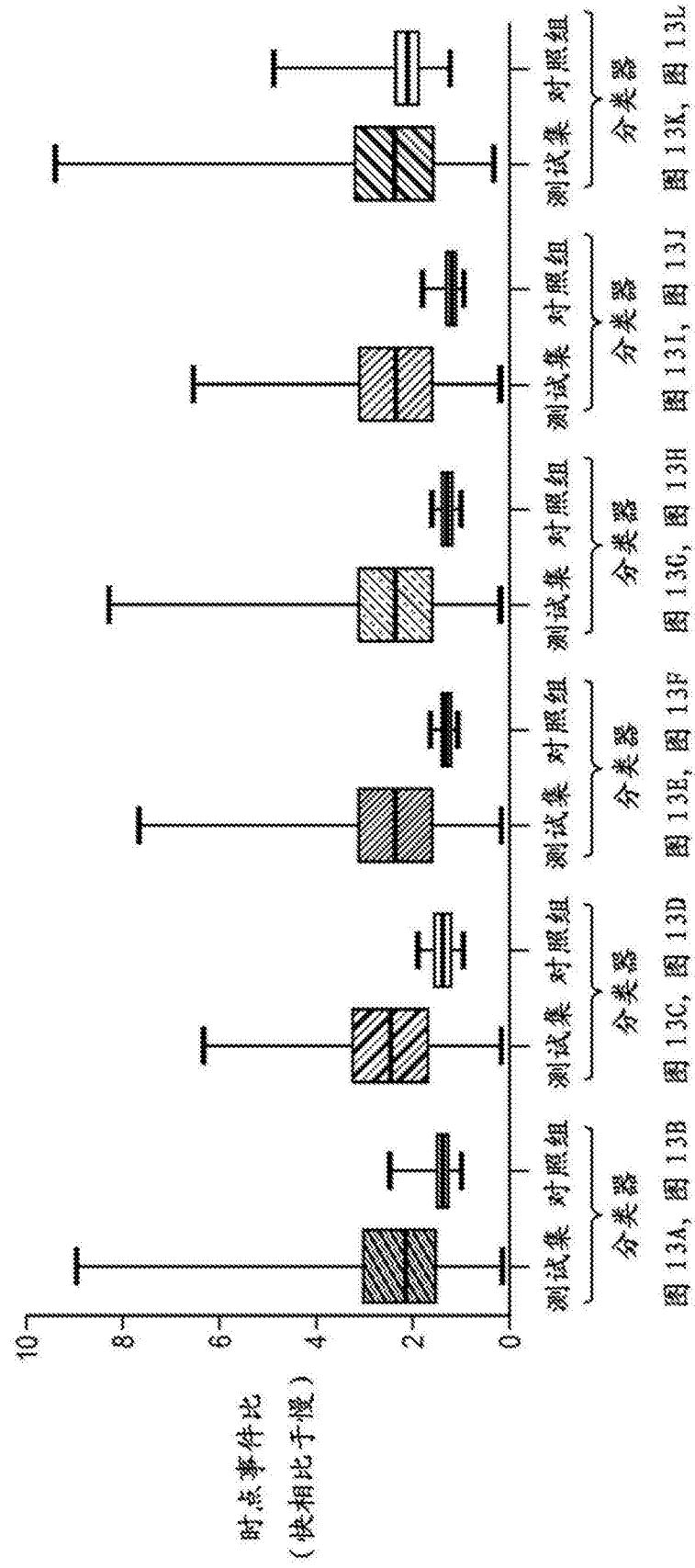


图 15

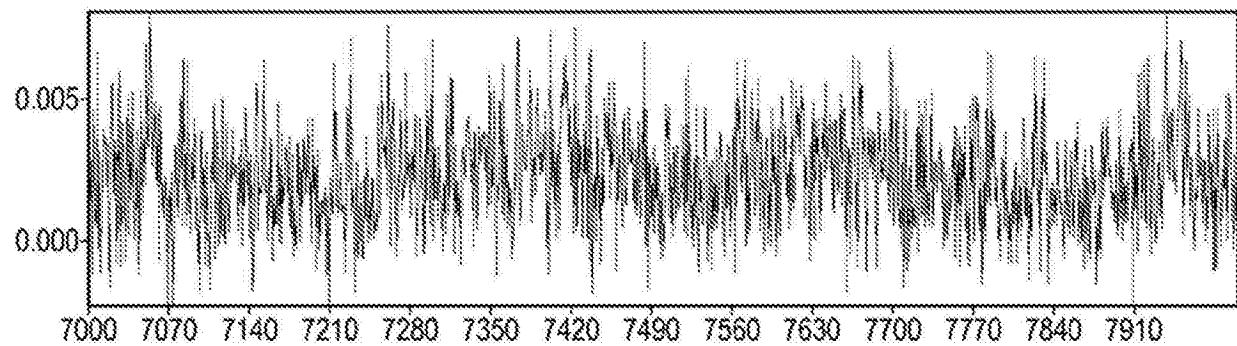


图16A

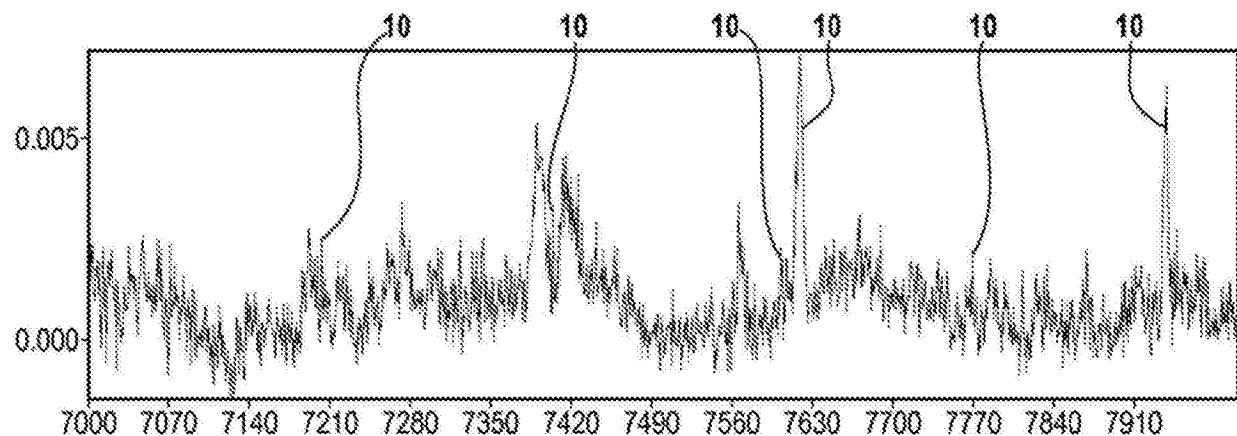


图16B

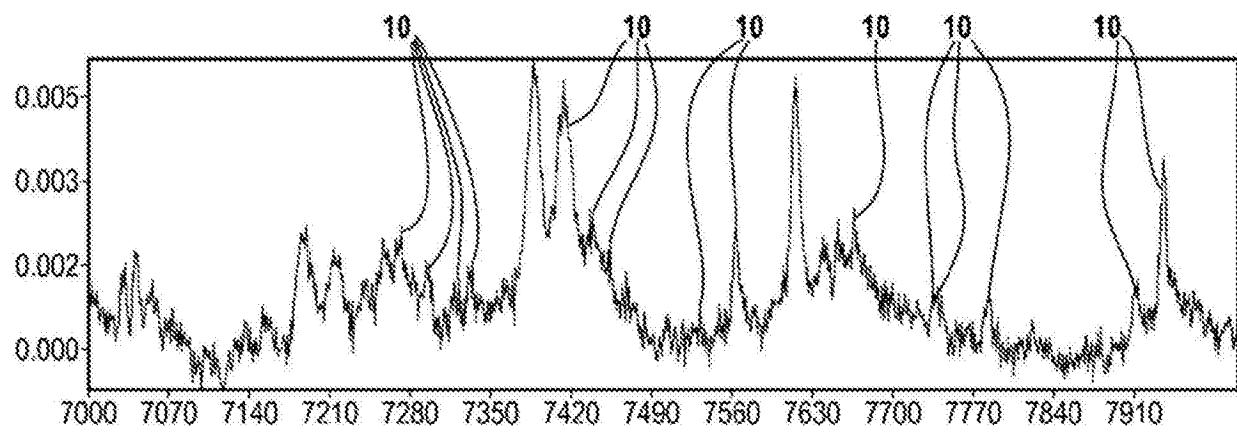


图16C

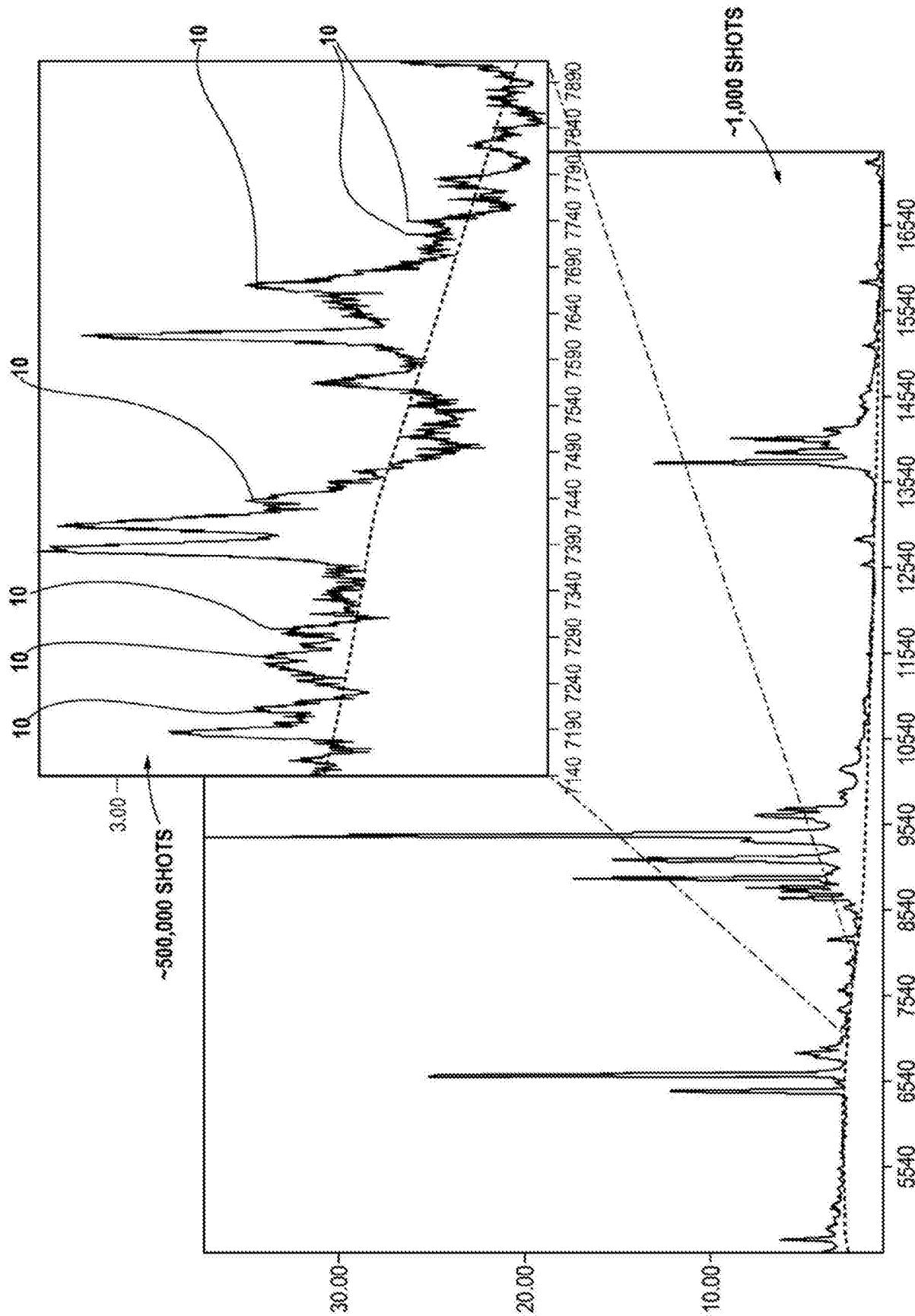


图16D

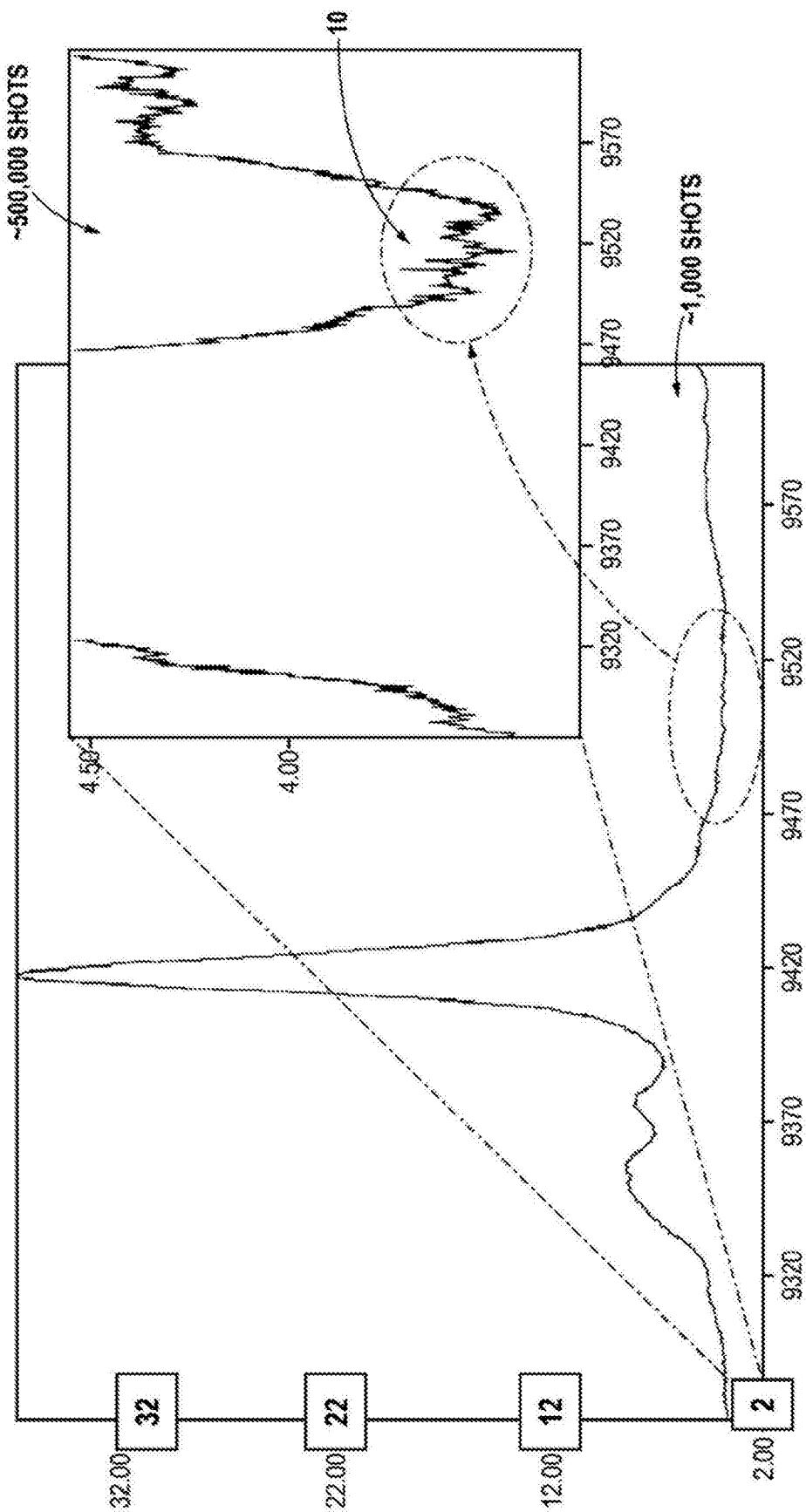


图16E

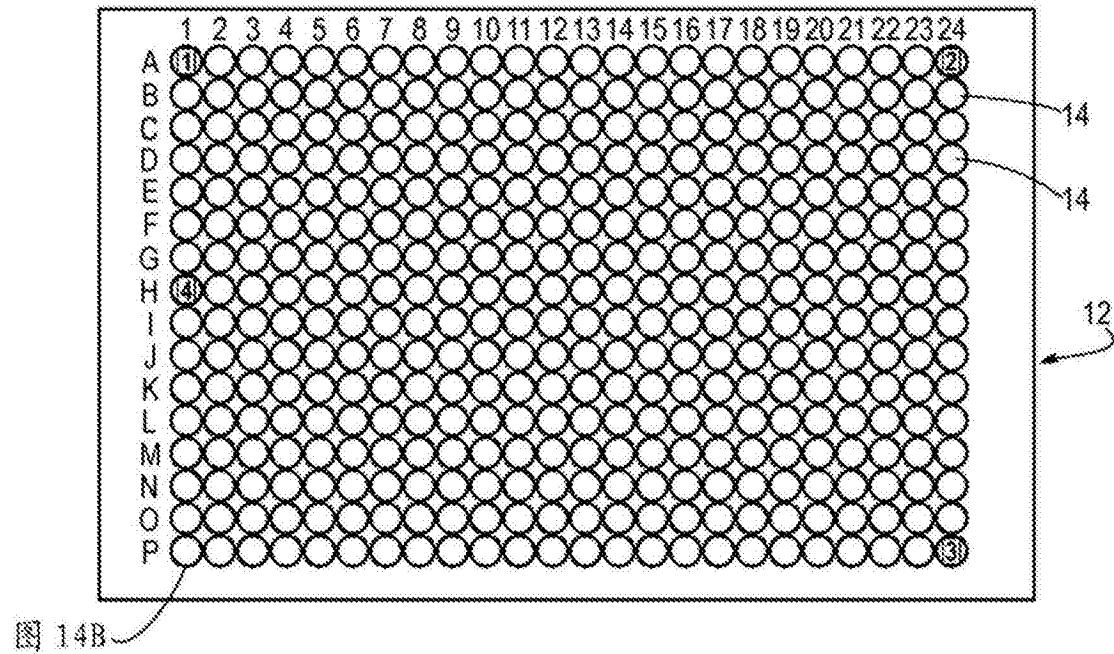


图17A

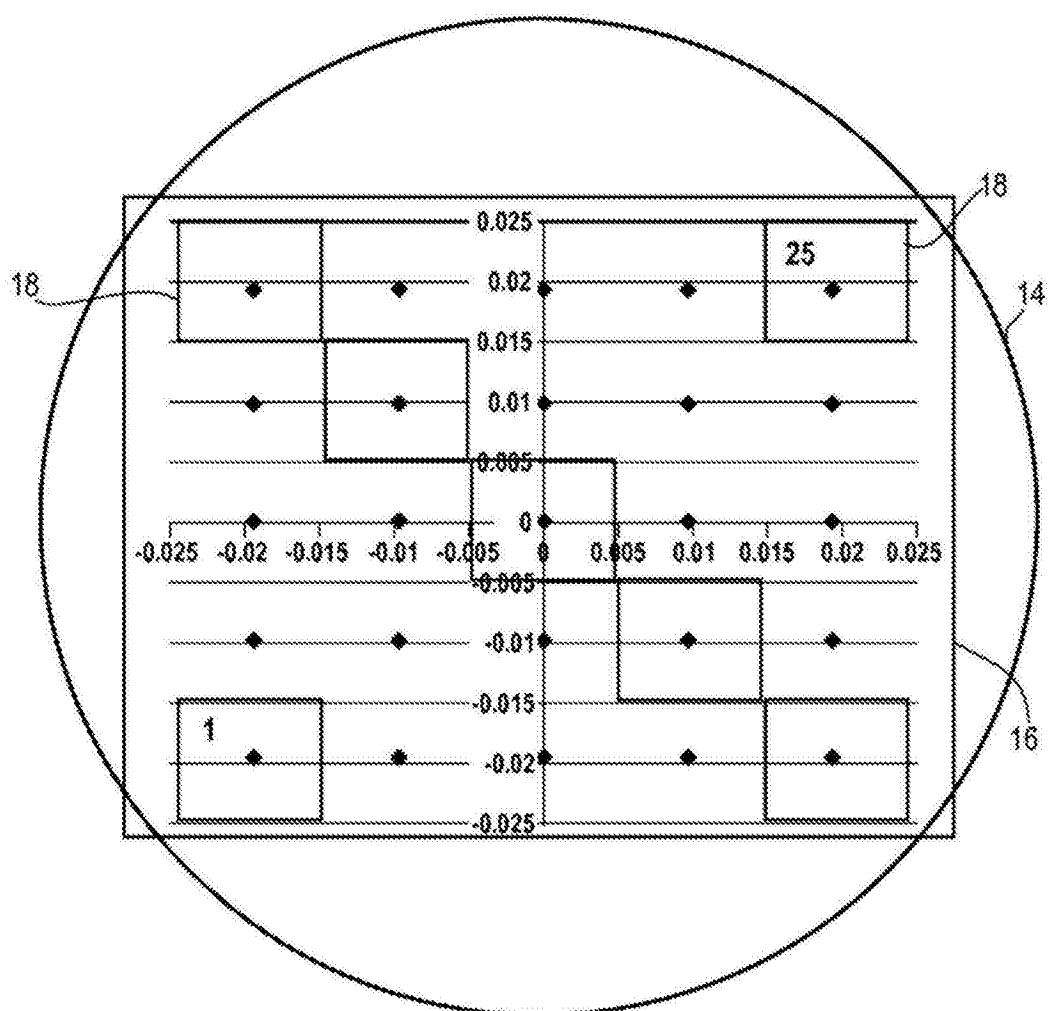


图17B

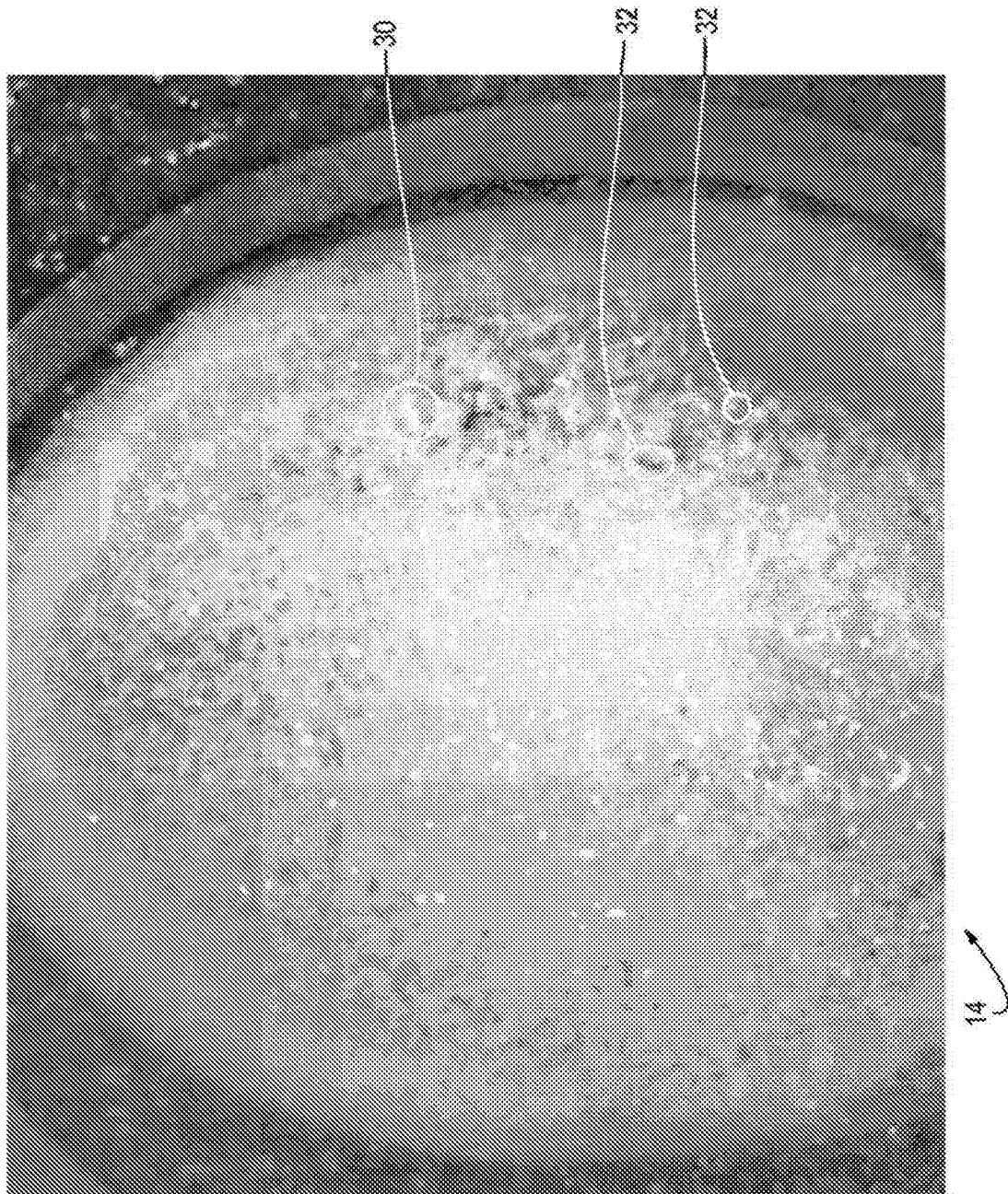


图18

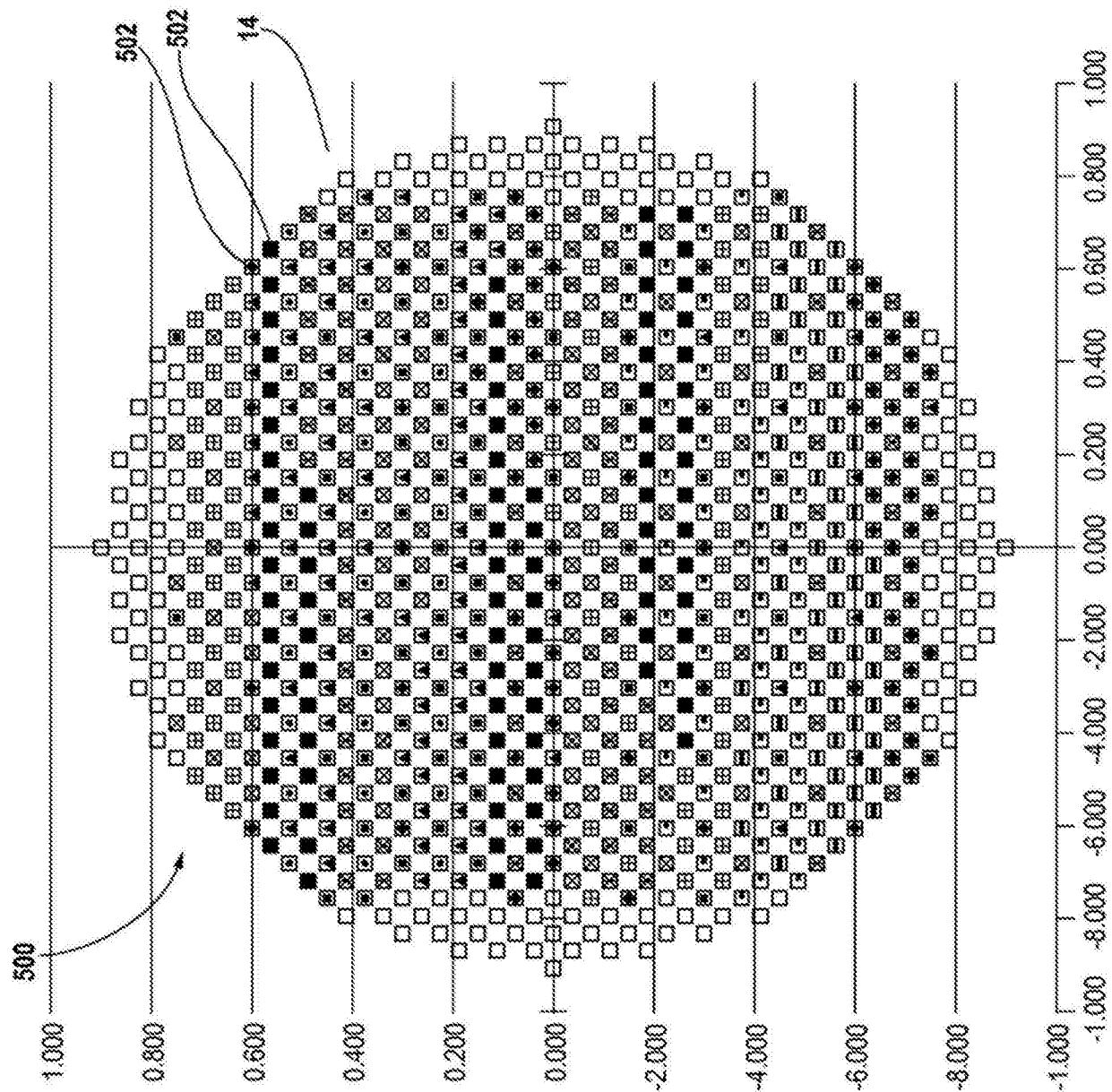


图19

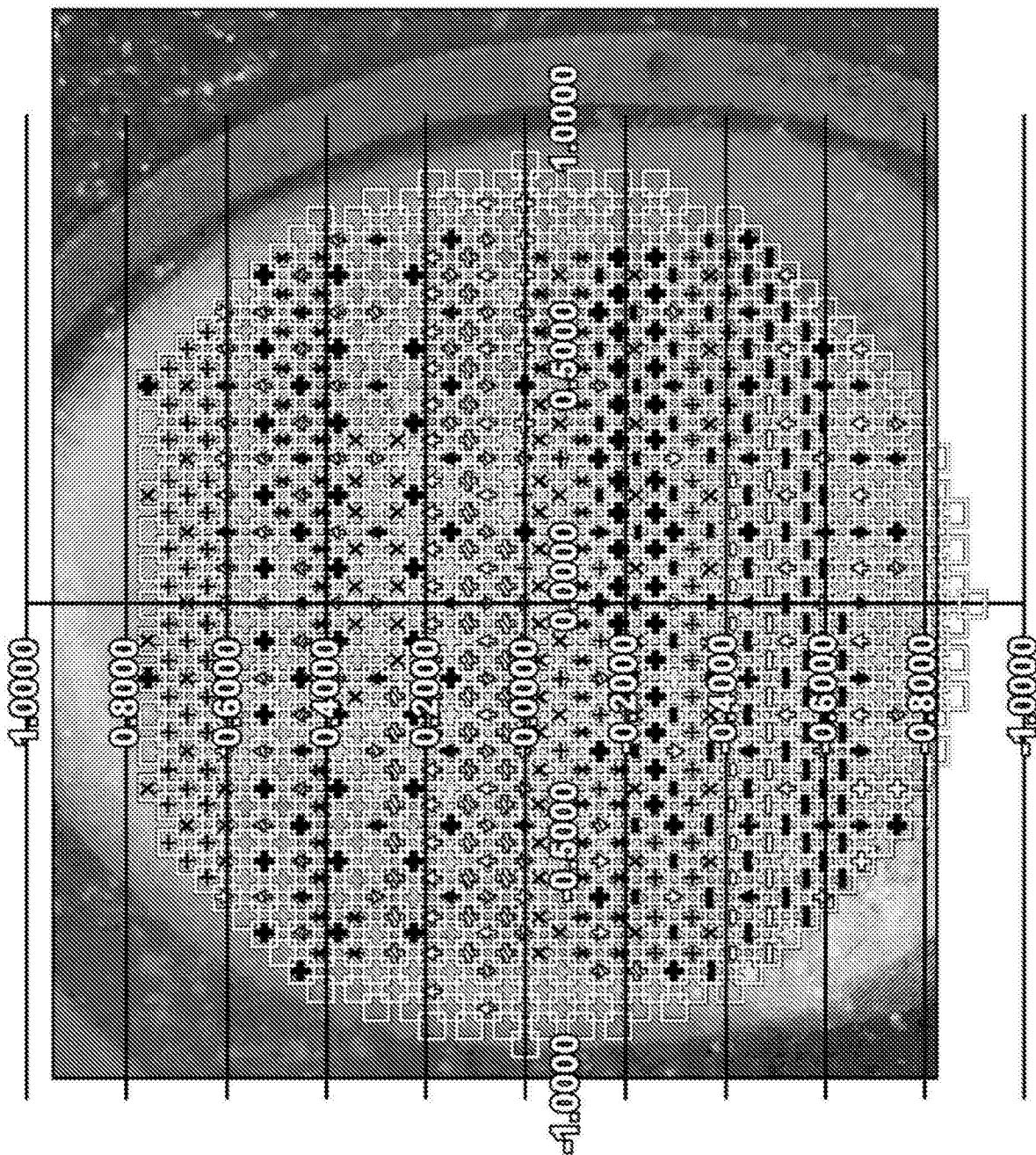


图20

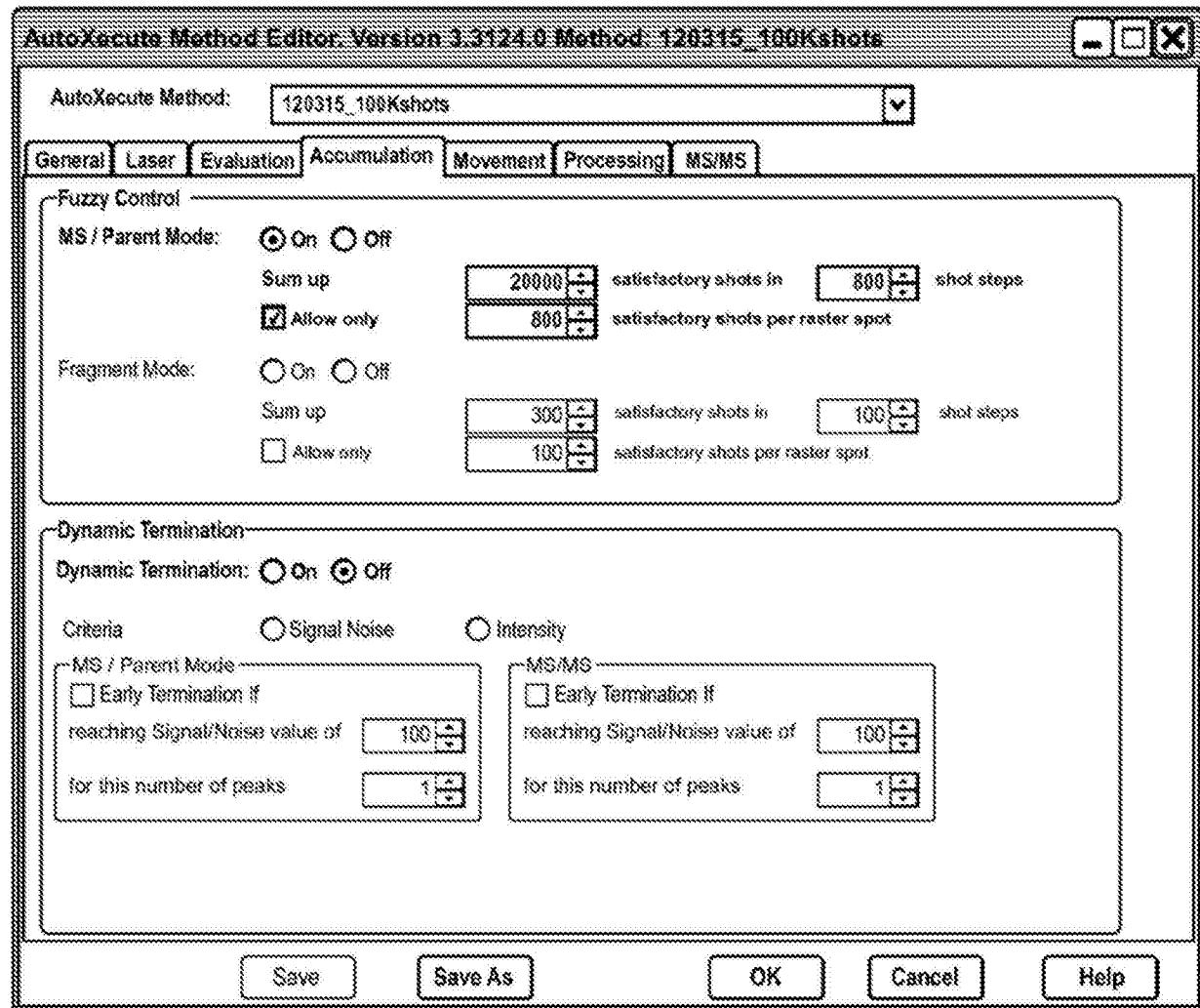


图21

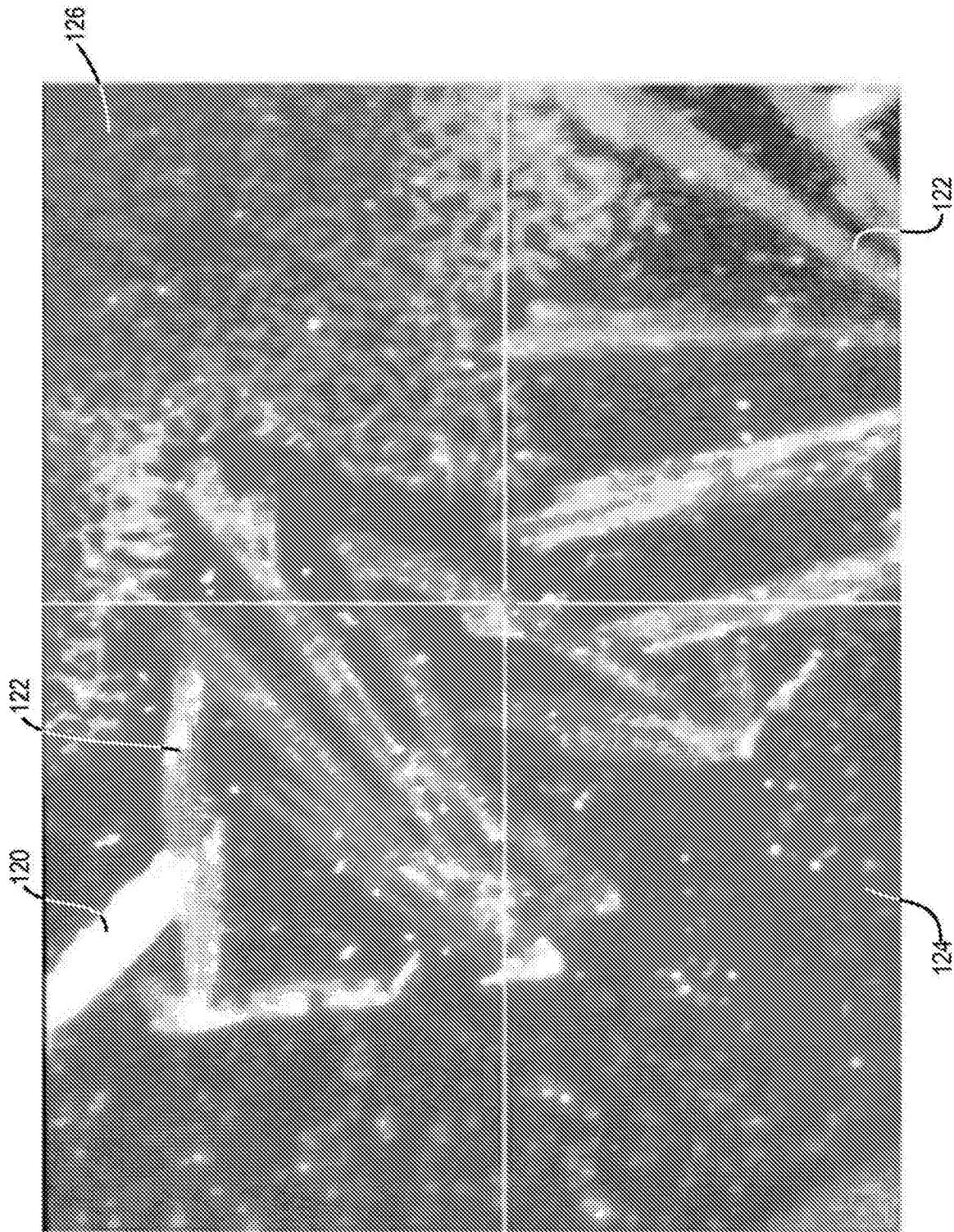


图22

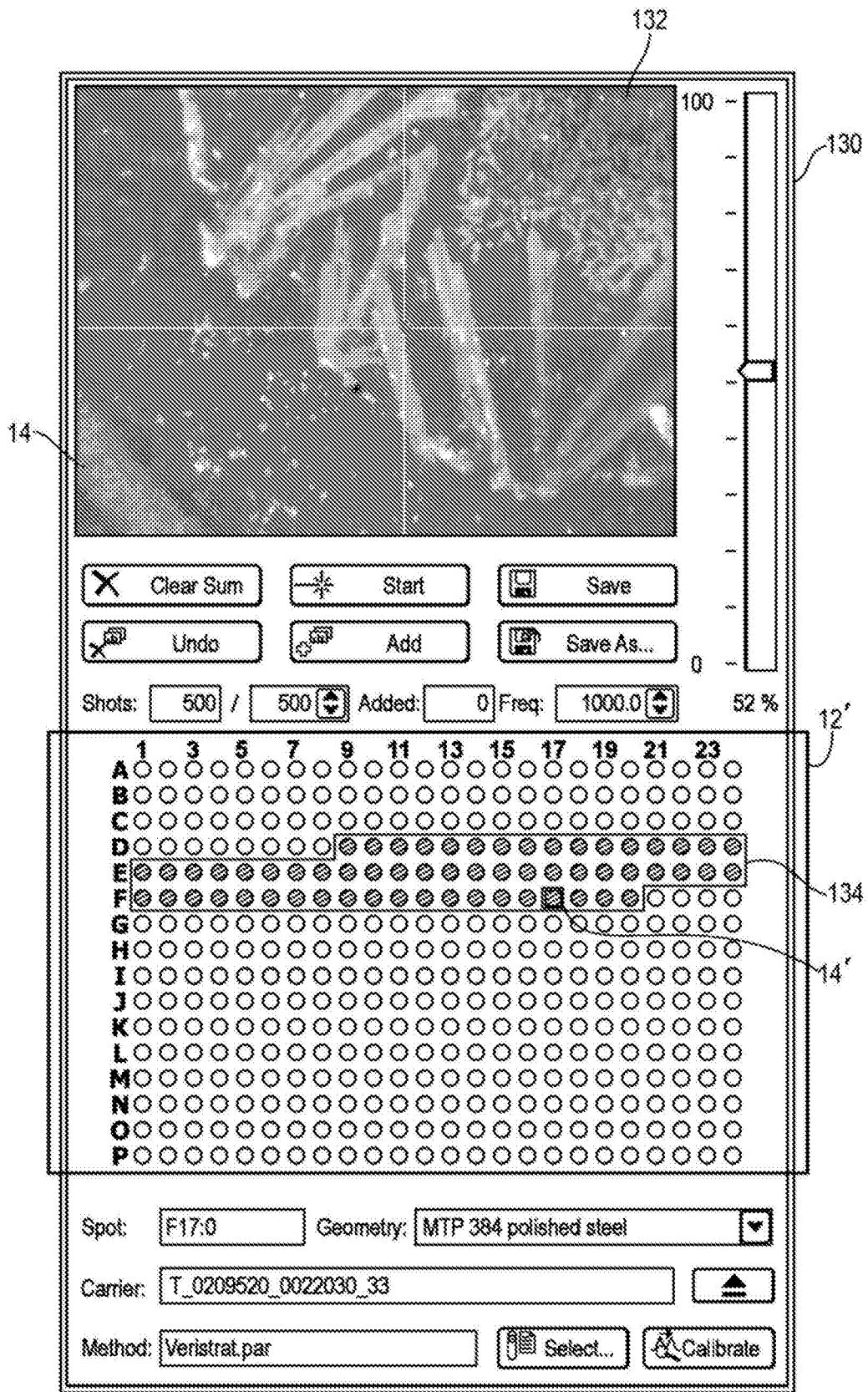


图23

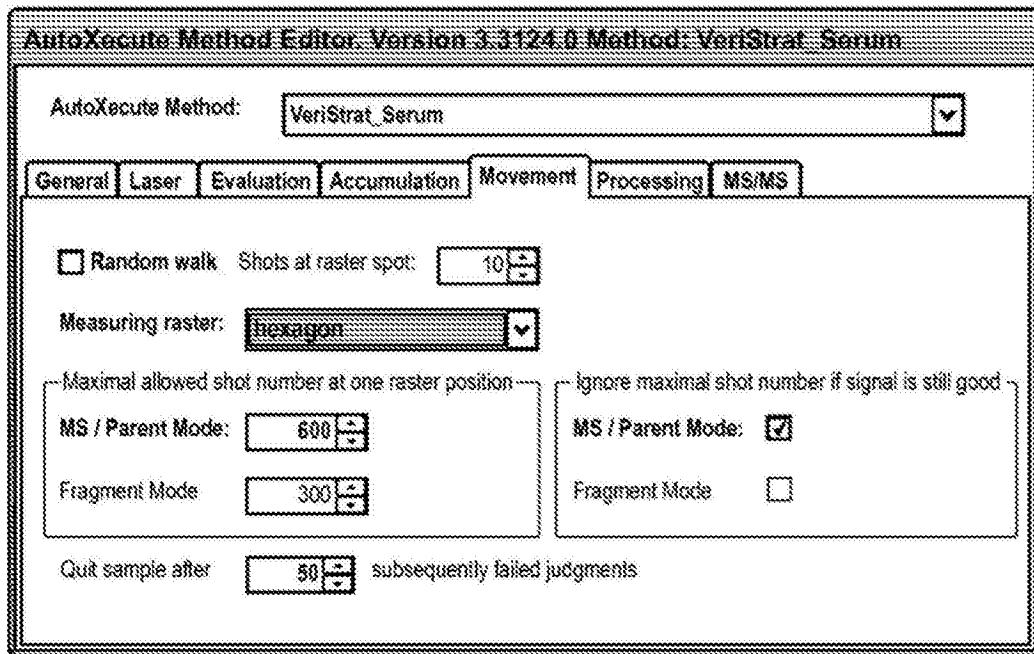
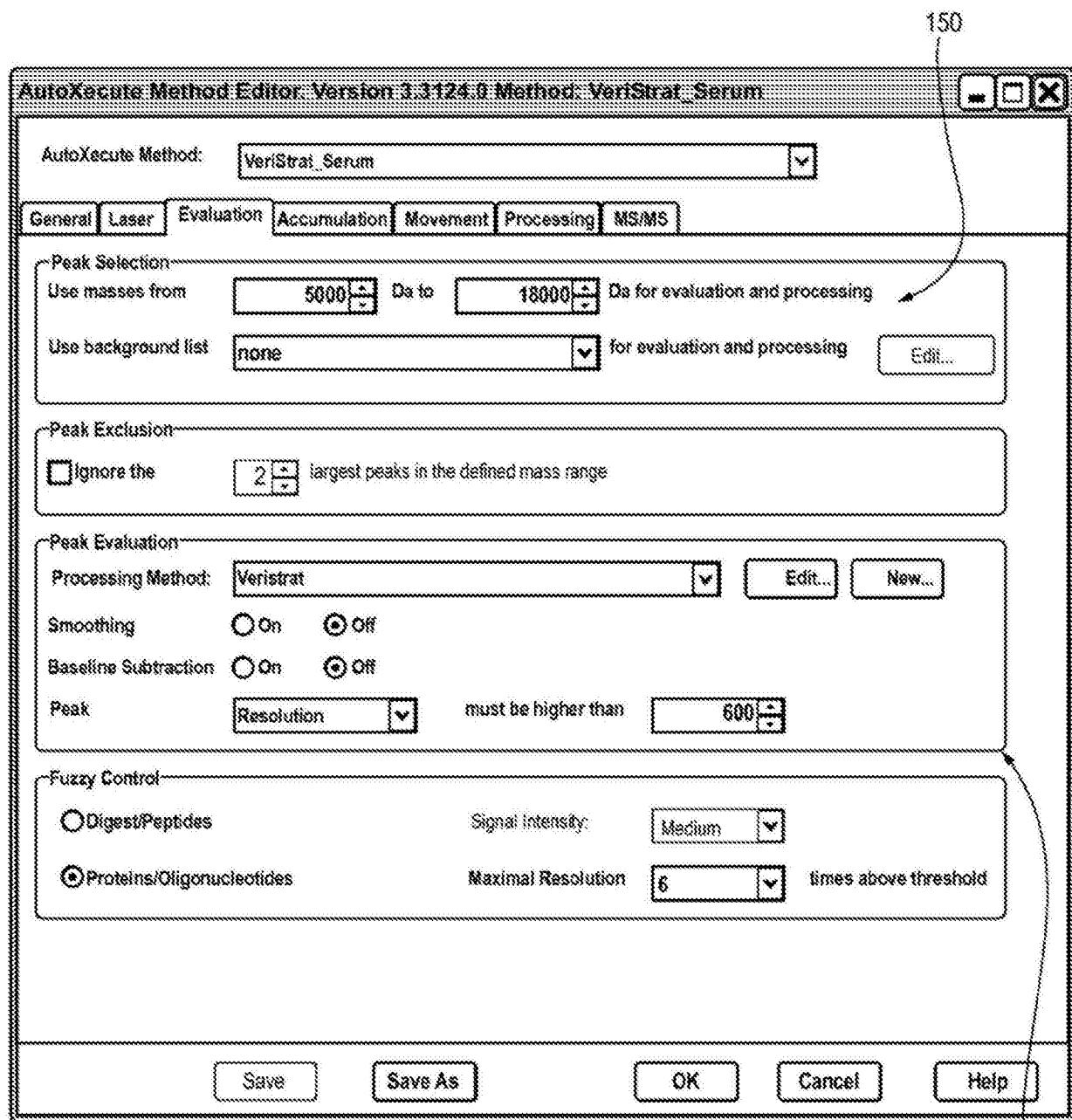


图24



152

图 25

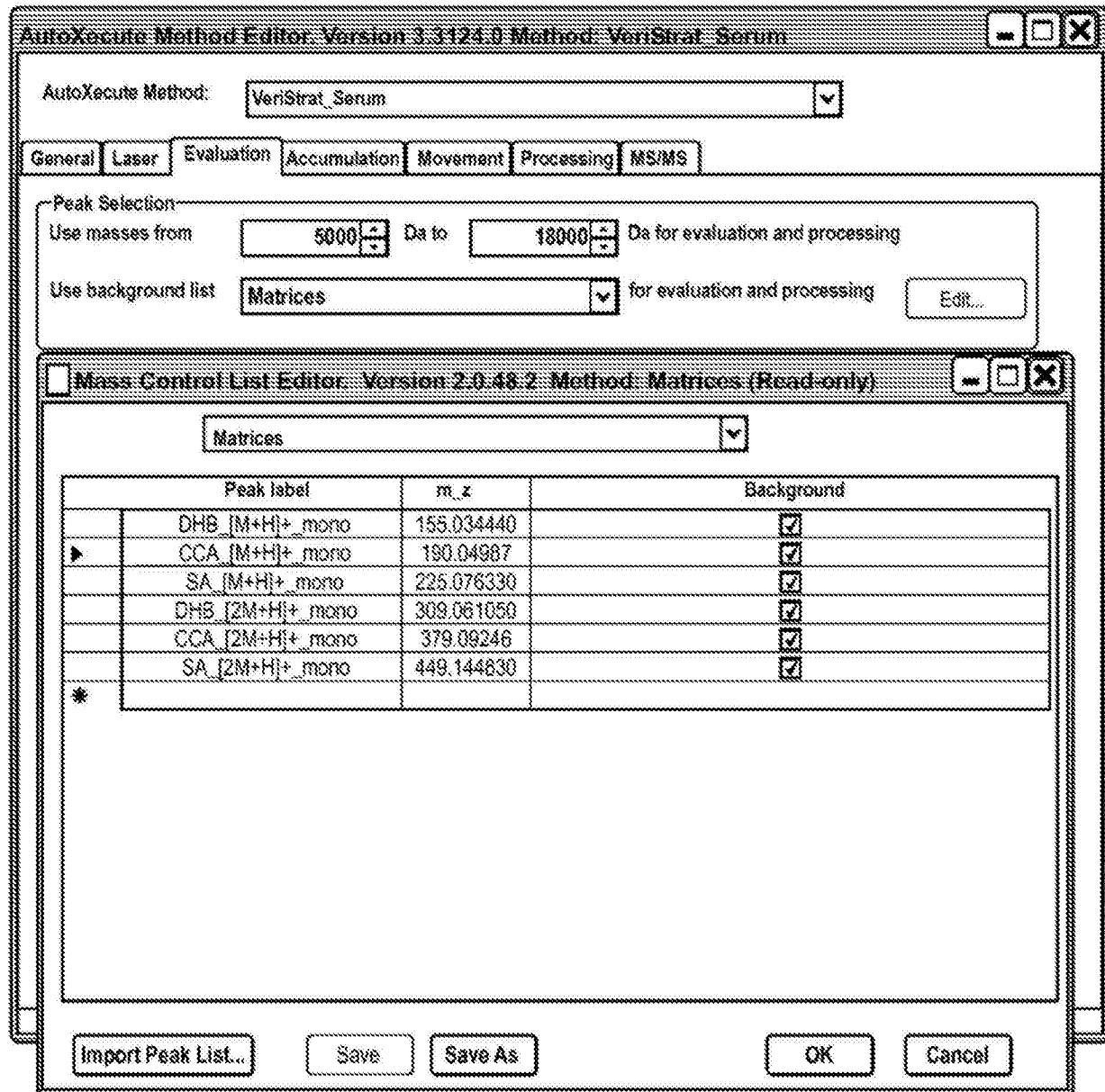


图26