



(12) 发明专利申请

(10) 申请公布号 CN 115917002 A

(43) 申请公布日 2023. 04. 04

(21) 申请号 202180040815.0

(22) 申请日 2021.04.05

(30) 优先权数据

63/005,996 2020.04.06 US

63/062,406 2020.08.06 US

63/136,449 2021.01.12 US

63/154,571 2021.02.26 US

(85) PCT国际申请进入国家阶段日

2022.12.06

(86) PCT国际申请的申请数据

PCT/US2021/025806 2021.04.05

(87) PCT国际申请的公布数据

W02021/207091 EN 2021.10.14

(71) 申请人 奥科坦特公司

地址 美国加利福尼亚州

(72) 发明人 斯里拉姆·库苏里

埃里克·马修·琼斯 艾伦·库珀

莫利·珍妮特·加斯佩里尼

(74) 专利代理机构 北京安信方达知识产权代理

有限公司 11262

专利代理师 徐爱文 张奎燕

(51) Int. Cl.

C12Q 1/6853 (2006.01)

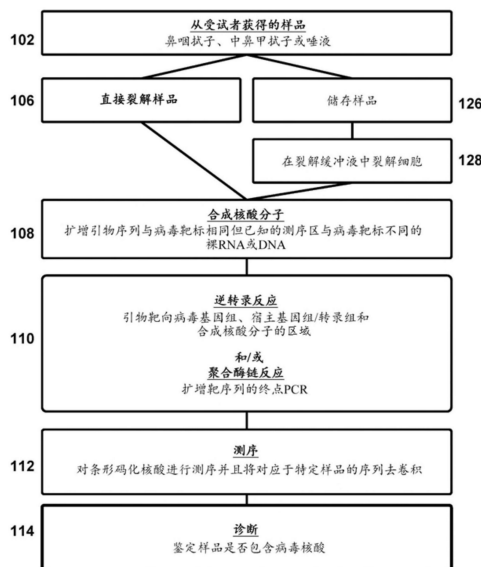
权利要求书7页 说明书108页 附图30页

(54) 发明名称

病原体诊断测试

(57) 摘要

本文描述了可用于使用PCR和测序来检测和诊断病原体感染的方法。



1. 一种检测个体的冠状病毒感染的方法,所述方法包括:
 - (a) 提供来自所述个体的生物样品,其中所述生物样品包含冠状病毒合成RNA,其中所述冠状病毒合成RNA的序列与天然存在的冠状病毒核酸序列不同;
 - (b) 裂解所述生物样品,从而产生裂解的生物样品;
 - (c) 对所述裂解的生物样品进行逆转录反应,以获得裂解的逆转录生物样品;
 - (d) 对所述裂解的逆转录生物样品进行扩增反应,以获得扩增的生物样品,其中对所述裂解的逆转录生物样品的所述扩增反应采用对冠状病毒核酸序列具有特异性的冠状病毒引物集进行,其中所述冠状病毒引物集扩增所述冠状病毒核酸序列和所述冠状病毒合成RNA;以及
 - (e) 使用下一代测序对所述扩增的生物样品进行测序。
2. 根据权利要求1所述的方法,其还包括如果检测到来自所述冠状病毒核酸序列的序列读段,则提供针对冠状病毒感染的阳性诊断。
3. 根据权利要求1或2所述的方法,其中所述冠状病毒感染是SARS-Cov-2感染。
4. 根据权利要求2或3所述的方法,其中如果来自所述冠状病毒核酸序列的所述序列读段与所述冠状病毒核酸RNA的序列读段的比率或其数学等同形式超过约0.1的比率,则提供针对所述冠状病毒感染或SARS-Cov-2感染的所述阳性诊断。
5. 根据权利要求2至4中任一项所述的方法,其中如果来自所述冠状病毒核酸的所述序列读段和冠状病毒合成RNA的所述序列读段超过约100,则提供针对冠状病毒感染的所述阳性诊断。
6. 根据权利要求1至5中任一项所述的方法,其中在进行所述逆转录反应之前未对所述裂解的生物样品进行分离和纯化。
7. 根据权利要求1至5中任一项所述的方法,其中裂解所述生物样品和对所述裂解的生物样品进行所述逆转录反应在同一孔、管或反应容器中发生。
8. 根据权利要求1至7中任一项所述的方法,其中裂解所述生物样品包括热裂解。
9. 根据权利要求8所述的方法,其中所述热裂解包括将所述生物样品加热至至少约50℃的温度。
10. 根据权利要求1至9中任一项所述的方法,其中来自所述个体的所述生物样品包含多个冠状病毒合成RNA序列,其中所述多个冠状病毒合成RNA序列包含至少两个不同的合成冠状病毒RNA序列。
11. 根据权利要求10所述的方法,其中所述多个合成冠状病毒RNA序列包含至少四个不同的合成冠状病毒RNA核酸序列。
12. 根据权利要求1至11中任一项所述的方法,其中所述合成RNA核酸或所述多个合成核酸包含约20%至约30%的量的鸟嘌呤核苷酸、约20%至约30%的量的腺嘌呤核苷酸、约20%至约30%的量的胞嘧啶核苷酸、约20%至约30%的量的尿嘧啶核苷酸。
13. 根据权利要求1至11中任一项所述的方法,其中所述合成冠状病毒RNA核酸或所述多个合成冠状病毒RNA核酸包含大约等比率的鸟嘌呤:胞嘧啶:腺嘌呤:尿嘧啶。
14. 根据权利要求1至13中任一项所述的方法,其中所述合成冠状病毒RNA核酸或所述多个合成冠状病毒RNA核酸包含合成SARS-Cov-2 RNA核酸或多个合成SARS-Cov-2 RNA核酸。

15. 根据权利要求1至14中任一项所述的方法,其中所述方法还包括检测甲型流感病毒感染、乙型流感病毒感染或其组合。

16. 根据权利要求1至15中任一项所述的方法,其中对所述裂解的生物样品的所述扩增反应采用对甲型流感病毒核酸序列具有特异性的甲型流感病毒引物集或对乙型流感病毒核酸序列具有特异性的乙型流感病毒引物集进行。

17. 根据权利要求16所述的方法,其中对所述甲型流感病毒核酸序列具有特异性的所述甲型流感病毒引物集包含SEQ ID NO:24或25和SEQ ID NO:26,或SEQ ID NO:27和SEQ ID NO:28中所示的序列。

18. 根据权利要求16所述的方法,其中对所述乙型流感病毒核酸序列具有特异性的所述乙型流感病毒引物集包含SEQ ID NO:29或30中所示的序列。

19. 根据权利要求1至18中任一项所述的方法,其中对所述裂解的生物样品的所述扩增反应采用对甲型流感病毒核酸序列具有特异性的甲型流感病毒引物集和对乙型流感病毒核酸序列具有特异性的乙型流感病毒引物集进行。

20. 根据权利要求1至19中任一项所述的方法,其中来自所述个体的所述生物样品还包含甲型流感病毒合成RNA、乙型流感病毒合成RNA或其组合,其中所述甲型流感病毒合成RNA、所述乙型流感病毒合成RNA或所述其组合与天然存在的甲型流感病毒或乙型流感病毒核酸序列不同。

21. 根据权利要求1至20中任一项所述的方法,其还包括如果来自甲型流感病毒的序列读段与所述甲型流感病毒合成RNA的序列读段的比率或其数学等同形式超过约0.1的比率,则提供针对甲型流感感染的阳性诊断。

22. 根据权利要求1至21中任一项所述的方法,其还包括如果来自甲型流感病毒的序列读段与所述乙型流感病毒合成RNA的序列读段的比率或其数学等同形式超过约0.1的比率,则提供针对乙型流感感染的阳性诊断。

23. 根据权利要求1至22中任一项所述的方法,其中所述冠状病毒核酸序列是N1序列、S2序列或其组合。

24. 根据权利要求1至23中任一项所述的方法,其中所述冠状病毒核酸序列是SARS-Cov-2 N1序列、SARS-Cov-2 S2序列或其组合。

25. 根据权利要求1至24中任一项所述的方法,其中所述冠状病毒引物集包含SEQ ID NO:13和SEQ ID NO:14,或SEQ ID NO:18和SEQ ID NO:19中所示的序列。

26. 根据权利要求1至24中任一项所述的方法,其中所述冠状病毒引物集包含SEQ ID NO:20和SEQ ID NO:21,或SEQ ID NO:22和SEQ ID NO:23中所示的序列。

27. 根据权利要求1至26中任一项所述的方法,其中所述冠状病毒引物集以约50纳摩尔至约250纳摩尔的浓度存在。

28. 根据权利要求1至26中任一项所述的方法,其中所述冠状病毒引物集以约100纳摩尔的浓度存在。

29. 根据权利要求16至28中任一项所述的方法,其中所述甲型流感病毒核酸序列是甲型流感病毒基质序列、甲型流感病毒非结构蛋白1序列、甲型流感病毒血凝素序列、甲型流感病毒神经氨酸酶序列、甲型流感病毒核蛋白序列及其组合。

30. 根据权利要求16至29中任一项所述的方法,其中所述乙型流感病毒核酸序列是乙

型流感病毒基质序列、乙型流感病毒非结构蛋白1序列、乙型流感病毒血凝素序列、乙型流感病毒神经氨酸酶序列、乙型流感病毒核蛋白序列及其组合。

31. 根据权利要求1至30中任一项所述的方法,其中所述冠状病毒引物集、所述甲型流感病毒引物集或所述乙型流感病毒引物集中的任一种或多种包含允许样品复用的一个或多个索引序列。

32. 根据权利要求1至31中任一项所述的方法,其中所述冠状病毒合成RNA包含与SEQ ID NO:1至12的任一个中所示的序列中的任一个或多个具有至少约90%同源性的核酸序列。

33. 根据权利要求32所述的方法,其中所述冠状病毒合成RNA包含与SEQ ID NO:1至12的任一个中所示的序列中的任一个或多个具有至少约90%同源性的多个不同核酸序列。

34. 根据权利要求20至33中任一项所述的方法,其中所述甲型流感病毒合成RNA包含与SEQ ID NO:31或32的任一个中所示的序列中的任一个或多个具有至少约90%同源性的RNA序列。

35. 根据权利要求20至34中任一项所述的方法,其中所述乙型流感病毒合成RNA包含与SEQ ID NO:33中所示的序列具有至少约90%同源性的RNA序列。

36. 根据权利要求1至35中任一项所述的方法,其中所述冠状病毒合成RNA包含与SEQ ID NO:1至4中所示的四个序列具有至少约90%同源性的多个四个不同核酸序列。

37. 根据权利要求1至36中任一项所述的方法,其中所述冠状病毒合成RNA、所述甲型流感病毒合成RNA或所述乙型流感病毒合成RNA以约10个拷贝/反应至约500个拷贝/反应的浓度存在。

38. 根据权利要求37所述的方法,其中所述冠状病毒合成RNA以约10个拷贝/反应至约500个拷贝/反应的浓度存在。

39. 根据权利要求1至36中任一项所述的方法,其中所述冠状病毒合成RNA、所述甲型流感病毒合成RNA或所述乙型流感病毒合成RNA以约200个拷贝/反应的浓度存在。

40. 根据权利要求39所述的方法,其中所述冠状病毒合成RNA以约200个拷贝/反应的浓度存在。

41. 根据权利要求1至40中任一项所述的方法,其中所述生物样品包括鼻拭子或唾液样品。

42. 根据权利要求41所述的方法,其中所述生物样品包含所述个体的少于约10微升唾液或已用所述个体的鼻拭子接种的少于约10微升缓冲液。

43. 根据权利要求1至42中任一项所述的方法,其中所述生物样品包含已用鼻拭子接种的少于约10微升缓冲液。

44. 根据权利要求1至43中任一项所述的方法,其中对所述裂解的生物样品的所述扩增反应采用对样品对照具有特异性的引物对进行。

45. 根据权利要求44所述的方法,其中所述样品对照是持家基因。

46. 根据权利要求44所述的方法,其中对所述样品对照具有特异性的所述引物对对RPP30具有特异性。

47. 根据权利要求46所述的方法,对所述样品对照具有特异性的所述引物对包含SEQ ID NO:15或16和SEQ ID NO:17中所示的序列。

48. 一种合成核酸,其包含:
- 5'近端区,所述5'近端区包含来自病毒的第一核苷酸序列;
- 3'近端区,所述3'近端区包含来自所述病毒的第二核苷酸序列;和
- 间插核苷酸序列,其中所述间插核苷酸序列包含百分比为约20%至约30%的鸟嘌呤核苷酸、量为约20%至约30%的腺嘌呤核苷酸、量为约20%至约30%的胞嘧啶核苷酸、量为约20%至约30%的尿嘧啶或胸苷核苷酸,并且所述间插序列不同于所述病毒天然存在的序列。
49. 根据权利要求48所述的合成核酸,其中所述合成核酸包含DNA。
50. 根据权利要求49所述的合成核酸,其中所述合成核酸由DNA组成。
51. 根据权利要求48所述的合成核酸,其中所述合成核酸包含RNA。
52. 根据权利要求51所述的合成核酸,其中所述合成核酸由RNA组成。
53. 根据权利要求48或52中任一项所述的方法,其中所述病毒是甲型流感病毒、乙型流感病毒或冠状病毒。
54. 根据权利要求48或52中任一项所述的合成核酸,其中所述病毒是冠状病毒。
55. 根据权利要求54所述的合成核酸,其中所述冠状病毒是SARS-Cov-2。
56. 根据权利要求48或55中任一项所述的合成核酸,其中所述间插核苷酸序列核酸包含大约等比率的鸟嘌呤核苷酸:胞嘧啶核苷酸:腺嘌呤核苷酸:尿嘧啶或胸苷核苷酸。
57. 根据权利要求48至56中任一项所述的合成核酸,其中所述3'近端区和所述5'近端区包含与冠状病毒S2基因序列具有至少90%同源性的核苷酸序列。
58. 根据权利要求48至56中任一项所述的合成核酸,其中所述5'近端区和所述3'近端区包含与冠状病毒S2基因序列具有至少95%同源性的核苷酸序列。
59. 根据权利要求48至56中任一项所述的合成核酸,其中所述5'近端区和所述3'近端区包含与冠状病毒S2基因序列相同的核苷酸序列。
60. 根据权利要求48至56中任一项所述的合成核酸,其中所述5'近端区和所述3'近端区包含与冠状病毒N1基因序列具有至少90%同源性的核苷酸序列。
61. 根据权利要求48至56中任一项所述的合成核酸,其中所述5'近端区和所述3'近端区包含与冠状病毒N1基因序列具有至少95%同源性的核苷酸序列。
62. 根据权利要求48至56中任一项所述的合成核酸,其中所述5'近端区和所述3'近端区包含与冠状病毒N1基因序列相同的核苷酸序列。
63. 根据权利要求57至62中任一项所述的合成核酸,其中所述冠状病毒N1基因序列或所述冠状病毒S2基因序列是SARS-CoV-2基因序列。
64. 根据权利要求46至63中任一项所述的合成核酸,其中所述序列核酸包含与SEQ ID NO:1至12的任一个中所示的任一个或多个序列具有至少90%同源性的序列。
65. 根据权利要求46或63中任一项所述的合成核酸,其中所述序列核酸包含与SEQ ID NO:1至12的任一个中所示的任一个或多个序列具有至少95%同源性的序列。
66. 根据权利要求46或63中任一项所述的合成核酸,其中所述序列核酸包含与SEQ ID NO:1至12的任一个中所示的任一个或多个序列相同的序列。
67. 根据权利要求46至66中任一项所述的多个合成核酸,其中所述多个合成核酸包括包含至少两个不同核苷酸序列的合成核酸。

68. 根据权利要求67中任一项所述的多个合成核酸,其中所述多个合成核酸包括包含至少两个不同核苷酸序列的合成核酸。

69. 根据权利要求67中任一项所述的多个合成核酸,其中所述多个合成核酸包括至少四个不同核苷酸序列。

70. 根据权利要求69所述的多个合成核酸,其中所述四个不同核苷酸序列是SEQ ID NO:1至4中所示的那些。

71. 根据权利要求69所述的多个合成核酸,其中所述四个不同核苷酸序列选自SEQ ID NO:5至12中所示的那些。

72. 一种反应混合物,其用于确定生物样品中是否存在病毒核酸,所述反应混合物包含根据权利要求48至66中任一项所述的合成核酸或根据权利要求67至71中任一项所述的多个合成核酸、所述生物样品的至少一部分和足以在所述病毒核酸存在时扩增所述生物样品中的所述病毒核酸的一种或多种酶或试剂。

73. 根据权利要求72所述的反应混合物,其中所述生物样品是人生物样品。

74. 根据权利要求72或73所述的反应混合物,其中所述生物样品包括唾液、口腔拭子、鼻咽拭子或中鼻甲拭子。

75. 根据权利要求72或73所述的反应混合物,其中所述生物样品包括唾液或鼻咽拭子。

76. 根据权利要求72至75中任一项所述的反应混合物,其中所述病毒核酸是甲型流感病毒、乙型流感病毒或冠状病毒核酸。

77. 根据权利要求76所述的反应混合物,其中所述冠状病毒核酸是Sars-Cov-2核酸。

78. 根据权利要求72至77中任一项所述的反应混合物,其中所述一种或多种试剂选自逆转录酶、dNTP、对所述病毒核苷酸序列具有特异性的引物对、对样品对照核苷酸序列具有特异性的引物对、镁盐及其组合。

79. 根据权利要求78所述的反应混合物,其中对所述样品对照核苷酸序列具有特异性的所述引物对对人核苷酸序列具有特异性。

80. 根据权利要求78或79所述的反应混合物,其中对所述样品对照核苷酸序列具有特异性的所述引物对对持家基因具有特异性。

81. 根据权利要求80所述的反应混合物,其中对所述样品对照具有特异性的所述引物对对RPP30具有特异性。

82. 根据权利要求81所述的反应混合物,其中对所述样品对照具有特异性的所述引物对包含SEQ ID NO:15或16和SEQ ID NO:17中所示的序列。

83. 根据权利要求78至82中任一项所述的反应混合物,其中对所述病毒核苷酸序列具有特异性的所述引物对对甲型流感病毒核苷酸序列、乙型流感病毒核苷酸序列和冠状病毒核苷酸序列具有特异性。

84. 根据权利要求83所述的反应混合物,其中对所述病毒核苷酸序列具有特异性的所述引物对对冠状病毒S1或N2序列具有特异性。

85. 根据权利要求78或84中任一项所述的反应混合物,其中所述冠状病毒S1或N2序列是冠状病毒S1或N2核酸序列。

86. 根据权利要求78至85中任一项所述的反应混合物,其中对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对包含SEQ

ID NO:13至30或100至605中的任一个中所示的序列。

87. 根据权利要求78至85中任一项所述的反应混合物,其中对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对包含SEQ ID NO:13和SEQ ID NO:14;SEQ ID NO:18和SEQ ID NO:19;SEQ ID NO:20和SEQ ID NO:21;SEQ ID NO:24或25和SEQ ID NO:26;SEQ ID NO:29和SEQ IDNO:30中所示的序列。

88. 根据权利要求78至87中任一项所述的反应混合物,其中对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对以约50微摩尔至约250微摩尔的浓度存在。

89. 根据权利要求78至88中任一项所述的反应混合物,其中对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对以约100微摩尔的浓度存在。

90. 根据权利要求78至88中任一项所述的反应混合物,其中对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对以约200微摩尔的浓度存在。

91. 根据权利要求72至90中任一项所述的反应混合物,其中所述冠状病毒合成RNA以约10个拷贝/反应至约500个拷贝/反应混合物的浓度存在。

92. 根据权利要求72至90中任一项所述的反应混合物,其中所述冠状病毒合成RNA以约200个拷贝/反应混合物的浓度存在。

93. 根据权利要求72至92中任一项所述的反应混合物,其中所述反应混合物的体积为约10微升至约100微升。

94. 根据权利要求72至92中任一项所述的反应混合物,其中所述反应混合物的体积为约20微升。

95. 一种试剂盒,其用于确定生物样品中是否存在病毒核酸,所述试剂盒包括根据权利要求46至66中任一项所述的合成核酸或根据权利要求67至71中任一项所述的多个合成核酸和足以扩增来自所述生物样品的所述病毒核酸的一种或多种酶或试剂。

96. 根据权利要求95所述的试剂盒,其中所述病毒核酸是甲型流感病毒、乙型流感病毒或冠状病毒核酸。

97. 根据权利要求95或96所述的试剂盒,其中所述冠状病毒核酸是冠状病毒核酸。

98. 根据权利要求95至97中任一项所述的试剂盒,其中所述一种或多种试剂选自逆转录酶、dNTP、对所述病毒核苷酸序列具有特异性的引物对、对样品对照核苷酸序列具有特异性的引物对、镁盐及其组合。

99. 根据权利要求98所述的试剂盒,其中对所述样品对照核苷酸序列具有特异性的所述引物对对人核苷酸序列具有特异性。

100. 根据权利要求98或99所述的试剂盒,其中对所述样品对照核苷酸序列具有特异性的所述引物对对持家基因具有特异性。

101. 根据权利要求100所述的试剂盒,其中对所述样品对照具有特异性的所述引物对对RPP30具有特异性。

102. 根据权利要求101所述的试剂盒,其中对所述样品对照具有特异性的所述引物对包含SEQ ID NO:15或16和SEQ ID NO:17中所示的序列。

103. 根据权利要求98至102中任一项所述的试剂盒,其中对所述病毒核苷酸序列具有特异性的所述引物对对甲型流感病毒核苷酸序列、乙型流感病毒核苷酸序列和冠状病毒核苷酸序列具有特异性。

104. 根据权利要求103所述的试剂盒,其中对所述病毒核苷酸序列具有特异性的所述引物对对冠状病毒S1或N2序列具有特异性。

105. 根据权利要求103或104中任一项所述的试剂盒,其中所述冠状病毒S1或N2序列是冠状病毒S1或N2核酸序列。

106. 根据权利要求98至105中任一项所述的试剂盒,其中对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对包含SEQ ID NO:13至30或100至605中的任一个中所示的序列。

107. 根据权利要求98至105中任一项所述的试剂盒,其中对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对包含SEQ ID NO:13和SEQ ID NO:14;SEQ ID NO:18和SEQ ID NO:19;SEQ ID NO:20和SEQ ID NO:21;SEQ ID NO:24或25和SEQ ID NO:26;SEQ ID NO:29和SEQ IDNO:30中所示的序列。

病原体诊断测试

相关申请的交叉引用

[0001] 本申请要求2020年4月6日提交的美国临时申请序列号63/005,996;2020年8月6日提交的美国临时申请序列号63/062,406;2021年1月12日提交的美国临时申请序列号63/136,449;2021年2月26日提交的美国临时申请序列号63/154,571的权益,所述临时申请以引用的方式整体并入本文。

背景技术

[0002] 发明人和申请人期望按照2020年3月31日的Covid开放承诺书(Open Covid Pledge),将本文所述的某些主题免费许可给推进结束COVID-19大流行的共同事业的实体。

[0003] 病毒性疾病及其爆发已经困扰了人类数千年。鉴定、监测和指导对病毒性疾病的应答的一个重要部分是有效且准确地测试个体的病毒感染的的能力。新型冠状病毒SARS-Cov-2导致COVID-19。存在对快速、灵敏且具有成本效益的测试的需求。

发明内容

[0004] 本文描述了一种诊断病原体感染的个体的方法。所述方法使用PCR和测序来实现高度特异性且灵敏地检测生物样品的病毒基因组。本文所述的允许此类检测的方法的特征包括:1)直接在裂解剂中或在裂解条件后进行逆转录和/或扩增,无需纯化或分离;2)存在合成核酸,合成核酸能够通过靶向感兴趣的病毒序列但包含不同的可区分间插序列的寡核苷酸引物来扩增,加标至逆转录扩增混合物中;并且允许更准确的定量和更低的检测阈值;3)允许通过下一代测序来复用的索引。在某些实施方案中,病原体是病毒感染(例如,SARS-CoV-2)。本文的方法还可以被复用,以允许检测多于一种病毒病原体(例如,SARS-CoV-2和甲型流感病毒或乙型流感病毒或两者)。

[0005] 在一个方面中,本文描述了一种诊断病原体感染的个体的方法,所述方法包括:a)提供来自所述个体的生物样品;a)将来自所述个体的所述生物样品与裂解剂接触,以获得裂解的生物样品;b)对所述裂解的生物样品进行聚合酶链反应(PCR),以获得PCR扩增的裂解的生物样品,其中对所述裂解的生物样品的所述PCR反应用第一PCR引物集进行,其中所述第一PCR引物集扩增病原体核酸序列;c)使用下一代测序对所述PCR扩增的裂解的生物样品进行测序;以及d)如果通过所述PCR或通过所述测序检测到病原体序列,则提供针对所述病原体感染的阳性诊断,或者如果通过所述PCR或通过所述测序未检测到病原体序列,则提供针对所述个体的阴性诊断。在某些实施方案中,所述个体是人类个体。在某些实施方案中,所述病原体感染包括细菌感染、病毒感染、真菌感染及其组合。在某些实施方案中,所述细菌感染是链球菌属(*Streptococcus*)、假单胞菌属(*Pseudomonas*)、志贺氏菌属(*Shigella*)、弯曲菌属(*Campylobacter*)、沙门氏菌属(*Salmonella*)、梭菌属(*Clostridium*)或埃希氏菌属(*Escherichia*)及其组合的感染。在某些实施方案中,所述真菌感染是念珠菌属(*Candida*)、芽生菌属(*Blastomyces*)、隐球菌属(*Cryptococcus*)、球孢子菌属(*Coccidioides*)、组织胞浆菌属(*Histoplasma*)、副球孢子菌属(*Paracoccidioides*)、孢子丝

菌属 (*Sporothrix*) 或肺囊虫属 (*Pneumocystis*) 及其组合的感染。在某些实施方案中, 病毒感染是DNA病毒的感染。在某些实施方案中, 所述DNA病毒包括甲型肝炎病毒、乙型肝炎病毒、丙型肝炎病毒、乳头瘤病毒、E-B病毒 (*Epstein-Barr virus*)、水痘病毒或天花病毒及其组合。在某些实施方案中, 所述病毒感染是RNA病毒的感染。在某些实施方案中, 所述RNA病毒包括流感病毒、冠状病毒、脊髓灰质炎病毒、麻疹病毒、埃博拉病毒、反转录病毒或正粘病毒。在某些实施方案中, 所述病毒感染是冠状病毒感染。在某些实施方案中, 所述冠状病毒感染是SARS-COV-2感染。在某些实施方案中, 来自所述个体的所述生物样品来自血液样品、血浆样品、血清样品、口腔拭子、尿液样品、精液样品、阴道拭子、粪便样品、鼻咽拭子、中鼻甲拭子或其任何组合。在某些实施方案中, 来自所述个体的所述生物样品来自鼻咽拭子、中鼻甲拭子或其任何组合。在某些实施方案中, 所述方法包括向所述裂解剂或所述裂解的生物样品中添加合成核酸。在某些实施方案中, 所述合成核酸是RNA。在某些实施方案中, 所述合成核酸是DNA。在某些实施方案中, 所述合成核酸包含被配置成由所述第一PCR引物集结合的序列集。在某些实施方案中, 所述第一引物集扩增所述病原体核酸序列和所述合成核酸。在某些实施方案中, 所述合成核酸序列包含与所述病原体核酸序列不相同的核苷酸序列。在某些实施方案中, 所述方法包括对所述裂解的生物样品执行逆转录反应。在某些实施方案中, 所述逆转录反应在所述执行所述聚合酶链反应之前执行。在某些实施方案中, 所述逆转录反应在没有进一步纯化所述裂解的生物样品的情况下执行。在某些实施方案中, 对所述裂解的生物样品的所述逆转录反应产生病毒cDNA。在某些实施方案中, 所述病毒cDNA是冠状病毒cDNA。在某些实施方案中, 所述冠状病毒cDNA是SARS-COV-2 cDNA。在某些实施方案中, 所述逆转录反应和所述PCR是单步骤反应。在某些实施方案中, 所述PCR是终点分析。在某些实施方案中, 所述PCR不是实时PCR反应。在某些实施方案中, 所述第一PCR引物集扩增冠状病毒核酸序列。在某些实施方案中, 所述冠状病毒核酸序列是SARS-COV-2核酸序列。在某些实施方案中, 所述SARS-COV-2核酸序列包含N1或S2基因。在某些实施方案中, 所述方法包括第二引物集。在某些实施方案中, 所述第二PCR引物集扩增人核酸序列。在某些实施方案中, 所述第二PCR引物集扩增选自以下的人核酸序列: GAPDH、ACTB、RPP30及其组合。在某些实施方案中, 所述第二PCR引物集扩增人RPP30。在某些实施方案中, 所述第二PCR引物集包含具有测序衔接子序列的引物和不具有测序衔接子序列的引物的混合物。在某些实施方案中, 具有测序衔接子序列的引物与不具有测序衔接子序列的引物的比率是约1:1、约1:2、约1:3或约1:4。在某些实施方案中, 所述PCR包括30至45个扩增循环。在某些实施方案中, 所述PCR包括35至45个扩增循环。在某些实施方案中, 所述PCR包括39至42个扩增循环。在某些实施方案中, 所述第一PCR引物集、所述第二PCR引物集或所述第一PCR引物集和所述第二PCR引物集包含有包含可变核苷酸序列的核酸序列。在某些实施方案中, 所述可变核苷酸序列是所述个体唯一的样品ID。在某些实施方案中, 所述第一PCR引物集、所述第二PCR引物集或所述第一PCR引物集和所述第二PCR引物集包含用于下一代测序反应的衔接子序列。在某些实施方案中, 所述方法可以检测少于10个病原体基因组拷贝。在某些实施方案中, 所述方法可以检测少于5个病原体基因组拷贝。在某些实施方案中, 所述病原体基因组是冠状病毒基因组。在某些实施方案中, 所述冠状病毒基因组是SARS-COV-2基因组。在某些实施方案中, 当通过所述PCR检测到冠状病毒序列时的所述针对冠状病毒的阳性诊断是SARS-COV-2诊断。在某些实施方案中, 所述方法确定冠状病毒的毒株。在某些实施方案中,

所述方法确定COVID-19毒株。

[0006] 本文还描述了一种诊断病原体感染的个体的方法,所述方法包括:(a)提供来自所述个体的生物样品;(c)对所述生物样品进行聚合酶链反应(PCR),以获得PCR扩增的生物样品,其中对所述生物样品的所述PCR反应用第一PCR引物集进行,其中所述第一PCR引物集扩增病原体核酸序列;(d)使用下一代测序对所述PCR扩增的生物样品进行测序;以及(e)如果通过所述PCR或通过所述测序检测到病原体序列,则提供针对所述病原体感染的阳性诊断,或者如果通过所述PCR或通过所述测序未检测到病原体序列,则提供针对所述个体的阴性诊断。在某些实施方案中,所述个体是人类个体。在某些实施方案中,所述病原体感染包括细菌感染、病毒感染、真菌感染及其组合。在某些实施方案中,所述细菌感染是链球菌属、假单胞菌属、志贺氏菌属、弯曲菌属、沙门氏菌属、梭菌属或埃希氏菌属及其组合的感染。在某些实施方案中,所述真菌感染是念珠菌属、芽生菌属、隐球菌属、球孢子菌属、组织胞浆菌属、副球孢子菌属、孢子丝菌属或肺囊虫属及其组合的感染。在某些实施方案中,病毒感染是DNA病毒的感染。在某些实施方案中,所述DNA病毒包括甲型肝炎病毒、乙型肝炎病毒、丙型肝炎病毒、乳头瘤病毒、E-B病毒、水痘病毒或天花病毒及其组合。在某些实施方案中,所述病毒感染是RNA病毒的感染。在某些实施方案中,所述RNA病毒包括流感病毒、冠状病毒、脊髓灰质炎病毒、麻疹病毒、埃博拉病毒、反转录病毒或正粘病毒。在某些实施方案中,所述病毒感染是冠状病毒感染。在某些实施方案中,所述冠状病毒感染是SARS-COV-2感染。在某些实施方案中,来自所述个体的所述生物样品来自血液样品、血浆样品、血清样品、口腔拭子、尿液样品、精液样品、阴道拭子、粪便样品、鼻咽拭子、中鼻甲拭子或其任何组合。在某些实施方案中,来自所述个体的所述生物样品来自鼻咽拭子、中鼻甲拭子或其任何组合。在某些实施方案中,所述方法包括向裂解剂或所述生物样品中添加合成核酸。在某些实施方案中,所述合成核酸是RNA。在某些实施方案中,所述合成核酸是DNA。在某些实施方案中,所述合成核酸包含被配置成由所述第一PCR引物集结合的序列集。在某些实施方案中,所述第一引物集扩增所述病原体核酸序列和所述合成核酸。在某些实施方案中,所述合成核酸序列包含与所述病原体核酸序列不相同的核苷酸序列。在某些实施方案中,所述方法包括对所述生物样品执行逆转录反应。在某些实施方案中,所述逆转录反应在所述执行所述聚合酶链反应之前执行。在某些实施方案中,所述逆转录反应在没有进一步纯化所述生物样品的情况下执行。在某些实施方案中,对所述生物样品的所述逆转录反应产生病毒cDNA。在某些实施方案中,所述病毒cDNA是冠状病毒cDNA。在某些实施方案中,所述冠状病毒cDNA是SARS-COV-2 cDNA。在某些实施方案中,所述逆转录反应和所述PCR是单步骤反应。在某些实施方案中,所述PCR是终点分析。在某些实施方案中,所述PCR不是实时PCR反应。在某些实施方案中,所述第一PCR引物集扩增冠状病毒核酸序列。在某些实施方案中,所述冠状病毒核酸序列是SARS-COV-2核酸序列。在某些实施方案中,所述SARS-COV-2核酸序列包含N1或S2基因。在某些实施方案中,所述方法包括第二引物集。在某些实施方案中,所述第二PCR引物集扩增人核酸序列。在某些实施方案中,所述第二PCR引物集扩增选自以下的人核酸序列:GAPDH、ACTB、RPP30及其组合。在某些实施方案中,所述第二PCR引物集扩增人RPP30。在某些实施方案中,所述第二PCR引物集包含具有测序衔接子序列的引物和不具有测序衔接子序列的引物的混合物。在某些实施方案中,具有测序衔接子序列的引物与不具有测序衔接子序列的引物的比率是约1:1、约1:2、约1:3或约1:4。在某些实施方案中,所述PCR包括30至45

个扩增循环。在某些实施方案中,所述PCR包括35至45个扩增循环。在某些实施方案中,所述PCR包括39至42个扩增循环。在某些实施方案中,所述第一PCR引物集、所述第二PCR引物集或所述第一PCR引物集和所述第二PCR引物集包含有包含可变核苷酸序列的核酸序列。在某些实施方案中,所述可变核苷酸序列是所述个体唯一的样品ID。在某些实施方案中,所述第一PCR引物集、所述第二PCR引物集或所述第一PCR引物集和所述第二PCR引物集包含用于下一代测序反应的衔接子序列。在某些实施方案中,所述方法可以检测少于10个病原体基因组拷贝。在某些实施方案中,所述方法可以检测少于5个病原体基因组拷贝。在某些实施方案中,所述病原体基因组是冠状病毒基因组。在某些实施方案中,所述冠状病毒基因组是SARS-COV-2基因组。在某些实施方案中,当通过所述PCR检测到冠状病毒序列时的所述针对冠状病毒的阳性诊断是SARS-COV-2诊断。在某些实施方案中,所述方法确定冠状病毒的毒株。在某些实施方案中,所述方法确定COVID-19毒株。

[0007] 在另一方面中,本文描述了一种诊断病原体感染的个体的方法,所述方法包括使用第一PCR引物集扩增来自所述个体的生物样品的核酸,从而获得扩增的核酸,其中所述第一PCR引物集扩增所述生物样品的病原体核酸序列和合成核酸序列,其中所述合成核酸序列与所述病原体核酸序列相差至少一个核苷酸。在某些实施方案中,所述个体是人类个体。在某些实施方案中,所述病原体感染包括细菌感染、病毒感染或真菌感染。在某些实施方案中,所述细菌感染是链球菌属、假单胞菌属、志贺氏菌属、弯曲菌属、沙门氏菌属、梭菌属或埃希氏菌属及其组合的感染。在某些实施方案中,所述真菌感染是念珠菌属、芽生菌属、隐球菌属、球孢子菌属、组织胞浆菌属、副球孢子菌属、孢子丝菌属或肺囊虫属及其组合的感染。在某些实施方案中,病毒感染是DNA病毒的感染。在某些实施方案中,所述DNA病毒包括乙型肝炎病毒、丙型肝炎病毒、乳头瘤病毒、E-B病毒、水痘病毒、天花病毒或它们的任何组合。在某些实施方案中,所述病毒感染是RNA病毒的感染。在某些实施方案中,所述RNA病毒包括流感病毒、冠状病毒、脊髓灰质炎病毒、麻疹病毒、埃博拉病毒、反转录病毒或正粘病毒。在某些实施方案中,所述病毒感染是冠状病毒感染。在某些实施方案中,所述冠状病毒感染是SARS-COV-2感染。在某些实施方案中,来自所述个体的所述生物样品来自血液样品、血浆样品、血清样品、口腔拭子、尿液样品、精液样品、阴道拭子、粪便样品、鼻咽拭子、中鼻甲拭子或其任何组合。在某些实施方案中,来自所述个体的所述生物样品来自鼻咽拭子、中鼻甲拭子或其任何组合。在某些实施方案中,所述合成核酸是RNA。在某些实施方案中,所述合成核酸是DNA。在某些实施方案中,所述合成核酸包含被配置成由所述第一PCR引物集结合的序列集。在某些实施方案中,所述合成核酸序列与所述病原体核酸序列相差至少5个核苷酸。在某些实施方案中,所述病原体感染是基于合成核酸序列与病原体核酸序列的比率诊断的。在某些实施方案中,所述方法包括对所述生物样品的所述核酸执行逆转录反应。在某些实施方案中,对所述生物样品的所述核酸的所述逆转录反应产生冠状病毒cDNA。在某些实施方案中,所述冠状病毒cDNA是SARS-COV-2 cDNA。在某些实施方案中,所述扩增核酸包括PCR反应。在某些实施方案中,所述PCR反应是终点分析。在某些实施方案中,所述PCR反应不是实时PCR反应。在某些实施方案中,所述第一PCR引物集扩增冠状病毒核酸序列。在某些实施方案中,所述冠状病毒核酸序列是SARS-COV-2核酸序列。在某些实施方案中,所述SARS-COV-2核酸序列包含N1或S2基因。在某些实施方案中,所述第一PCR引物集包含有包含可变核苷酸序列的核酸序列。在某些实施方案中,所述可变核苷酸序列是所述个体唯一的

样品ID。在某些实施方案中,所述第一PCR引物集包含用于下一代测序反应的衔接子序列。在某些实施方案中,所述方法包括使用第二PCR引物集扩增所述生物样品的核酸,其中所述第二PCR引物集扩增人核酸序列。在某些实施方案中,所述第二PCR引物集扩增选自以下的核酸序列:GAPDH、ACTB、RPP30及其组合。在某些实施方案中,所述第二PCR引物集扩增人RPP30。在某些实施方案中,所述第二PCR引物集包含具有测序衔接子序列的引物和不具有测序衔接子序列的引物的混合物。在某些实施方案中,具有测序衔接子序列的引物与不具有测序衔接子序列的引物的比率是约1:1、约1:2、约1:3或约1:4。在某些实施方案中,所述PCR包括30至45个扩增循环。在某些实施方案中,所述PCR包括35至45个扩增循环。在某些实施方案中,所述PCR包括39至42个扩增循环。在某些实施方案中,所述第二PCR引物集包含有包含可变核苷酸序列的核酸序列。在某些实施方案中,所述可变核苷酸序列是所述个体唯一的样品ID。在某些实施方案中,所述第二PCR引物集包含用于下一代测序反应的衔接子序列。在某些实施方案中,使用下一代测序技术对所述生物样品的所述扩增的核酸进行测序。在某些实施方案中,所述方法可以检测少于10个病原体基因组拷贝。在某些实施方案中,所述方法可以检测少于5个病原体基因组拷贝。在某些实施方案中,所述病原体基因组是冠状病毒基因组。在某些实施方案中,所述病原体基因组是SARS-COV-2基因组。在某些实施方案中,所述方法确定冠状病毒的毒株。在某些实施方案中,所述方法确定SARS-COV-2的毒株。

[0008] 在另一方面中,本文还描述了一种合成核酸,其包含5'近侧区、3'近侧区和间插核酸序列。在某些实施方案中,所述合成核酸包含RNA。在某些实施方案中,所述合成核酸包含DNA。在某些实施方案中,所述5'近侧区包含病毒核酸序列。在某些实施方案中,所述病毒核酸序列包含冠状病毒序列。在某些实施方案中,所述病毒核酸序列包含SARS-COV-2序列。在某些实施方案中,所述3'近侧区包含病毒核酸序列。在某些实施方案中,所述病毒核酸序列包含冠状病毒序列。在某些实施方案中,所述病毒核酸序列包含SARS-COV-2序列。在某些实施方案中,所述5'近侧区、所述3'近侧区或所述5'近侧区和所述3'近侧区的长度少于约30个核苷酸。在某些实施方案中,所述5'近侧区、所述3'近侧区或所述5'近侧区和所述3'近侧区的长度少于约25个核苷酸。在某些实施方案中,所述5'近侧区、所述3'近侧区或所述5'近侧区和所述3'近侧区的长度少于约20个核苷酸。在某些实施方案中,所述5'近侧区在所述合成核酸的5'末端。在某些实施方案中,所述3'近侧区在所述合成核酸的3'末端。在某些实施方案中,所述间插核酸序列与病毒核酸序列具有少于约99%、98%、97%、95%、90%、85%、80%或75%同一性。在某些实施方案中,所述合成核酸序列是冠状病毒序列。在某些实施方案中,所述合成核酸序列是SARS-COV-2序列。在某些实施方案中,所述合成核酸用于检测所述个体的病原体感染的方法中。在某些实施方案中,所述病原体感染是冠状病毒感染。在某些实施方案中,所述病毒感染是SARS-COV-2感染。

[0009] 在另一方面中,本文描述了一种组合物,其包含有包含第一核酸序列和第二核酸序列的合成核酸分子,其中(1)所述第一核酸序列与病原体核酸分子的序列相同,并且(2)所述第二核酸序列与所述病原体核酸分子的序列不相同。在某些实施方案中,所述第一核酸序列位于所述第二核酸序列的3'。在某些实施方案中,所述合成核酸分子还包含第三核酸序列,其中所述第三核酸序列与所述病原体核酸分子的第二序列相同。在某些实施方案中,所述第三核酸序列在所述第二核酸序列的5'。在某些实施方案中,所述第一核酸序列或所述第三核酸序列少于5、10、15、20、25或30个核苷酸。在某些实施方案中,所述第二核酸序

列所包含的核苷酸的总数少于25、50、100、150、200或500个核苷酸。在某些实施方案中,所述第二核酸序列所包含的核苷酸的总数多于25、50、100、150、200或500个核苷酸。在某些实施方案中,所述合成核酸分子是核糖核酸(RNA)分子、脱氧核糖核酸(DNA)分子或RNA-DNA杂交分子。在某些实施方案中,所述第一核酸序列、所述第三核酸序列或所述第一核酸裂和所述第三核酸序列包含引物结合位点。在某些实施方案中,所述组合物还包含病原体核酸分子。在某些实施方案中,所述病原体核酸分子来自病原体,其中所述病原体包括细菌、病毒、真菌或其组合。在某些实施方案中,所述细菌来自链球菌属、假单胞菌属、志贺氏菌属、弯曲菌属、沙门氏菌属、梭菌属或埃希氏菌属及其组合。在某些实施方案中,真菌是念珠菌属、芽生菌属、隐球菌属、球孢子菌属、组织胞浆菌属、副球孢子菌属、孢子丝菌属或肺囊虫属及其组合。在某些实施方案中,病毒是DNA病毒。在某些实施方案中,DNA病毒包括乙型肝炎病毒、丙型肝炎病毒、乳头瘤病毒、E-B病毒、水痘病毒或天花病毒及其组合。在某些实施方案中,病毒是RNA病毒。在某些实施方案中,所述RNA病毒包括流感病毒、冠状病毒、脊髓灰质炎病毒、麻疹病毒、埃博拉病毒、反转录病毒或正粘病毒。在某些实施方案中,病毒是冠状病毒。在某些实施方案中,冠状病毒是SARS-COV-2病毒。在某些实施方案中,组合物还包含多个引物,其中多个引物中的引物被配置成与合成核酸分子的序列或病原体核酸分子的序列杂交。在某些实施方案中,合成核酸分子的序列和病原体核酸分子的序列是相同的。在某些实施方案中,合成核酸分子以与病原体核酸分子相同的效率扩增。在某些实施方案中,合成核酸分子被配置成产生大小与所述病原体核酸分子的扩增产物的大小相同或在其10个碱基对内的扩增产物。

[0010] 在另一方面中,本文描述了一种诊断病原体感染的个体的方法,所述方法包括:(a)提供来自所述个体的生物样品;(b)将来自所述个体的所述生物样品与裂解剂接触,以获得裂解的生物样品;(c)对所述裂解的生物样品进行聚合酶链反应(PCR),以获得PCR扩增的裂解的生物样品,其中对所述裂解的生物样品的所述PCR反应用第一PCR引物集进行,其中所述第一PCR引物集扩增病原体核酸序列;(d)使用下一代测序对所述PCR扩增的裂解的生物样品进行测序;以及(e)如果通过所述PCR或通过所述测序检测到病原体序列,则提供针对所述病原体感染的阳性诊断,或者如果通过所述PCR或通过所述测序未检测到病原体序列,则提供针对所述个体的阴性诊断。

[0011] 在另一方面中,本文描述了一种组合物,其包含多个具有不同核酸序列的合成核酸,具有不同核酸序列的所述多个合成核酸包含与病原体核酸序列相同的共同5'序列、与病原体核酸序列相同的共同3'序列和多个序列不同的间插序列。在某些实施方案中,具有不同核酸序列的所述多个合成核酸是单链的。在某些实施方案中,具有不同核酸序列的所述多个合成核酸是双链的。在某些实施方案中,具有不同核酸序列的所述多个合成核酸由RNA组成或包含RNA。在某些实施方案中,具有不同核酸序列的所述多个合成核酸由DNA组成或包含DNA。在某些实施方案中,与病原体核酸序列相同的所述共同5'序列是30个核苷酸或更少。在某些实施方案中,与病原体核酸序列相同的所述共同3'序列是30个核苷酸或更少。在某些实施方案中,所述间插序列是50个核苷酸或更少。在某些实施方案中,所述间插序列是30个核苷酸或更少。在某些实施方案中,所述病原体核酸序列来自细菌病原体、真菌病原体或病毒病原体。在某些实施方案中,所述细菌病原体是链球菌属、假单胞菌属、志贺氏菌属、弯曲菌属、沙门氏菌属、梭菌属或埃希氏菌属及其组合。在某些实施方案中,所述真菌病

原体是念珠菌属、芽生菌属、隐球菌属、球孢子菌属、组织胞浆菌属、副球孢子菌属、孢子丝菌属或肺囊虫属及其组合。在某些实施方案中,所述病毒病原体是DNA病毒。在某些实施方案中,所述DNA病毒包括乙型肝炎病毒、丙型肝炎病毒、乳头瘤病毒、E-B病毒、水痘病毒或天花病毒及其组合。在某些实施方案中,所述病毒病原体是RNA病毒。在某些实施方案中,所述RNA病毒包括流感病毒、冠状病毒、脊髓灰质炎病毒、麻疹病毒、埃博拉病毒、反转录病毒、正粘病毒或其组合。在某些实施方案中,所述病毒病原体是冠状病毒。在某些实施方案中,所述冠状病毒是SARS-COV-2。在某些实施方案中,所述病原体核酸序列是编码冠状病毒刺突蛋白的核酸序列。在某些实施方案中,具有不同核酸序列的所述多个合成核酸包含选自S2_001、S2_002、S2_003和S2_004中的任一种或多种的序列或由其组成。在某些实施方案中,所述组合物在诊断或检测病原体感染的方法中使用。在某些实施方案中,所述组合物在归一化病原体下一代序列读段(read)的方法中使用。

[0012] 本文描述了一种检测个体的冠状病毒感染的方法,所述方法包括:(a)提供来自所述个体的生物样品,其中所述生物样品包含冠状病毒合成RNA,其中所述冠状病毒合成RNA的序列与天然存在的冠状病毒核酸序列不同;(b)裂解所述生物样品,从而产生裂解的生物样品;(c)对所述裂解的生物样品进行逆转录反应,以获得裂解的逆转录的生物样品;(d)对所述裂解的逆转录的生物样品进行扩增反应,以获得扩增的生物样品,其中对所述裂解的逆转录生物的样品的所述扩增反应用对冠状病毒核酸序列具有特异性的冠状病毒引物集进行,其中所述冠状病毒引物集扩增所述冠状病毒核酸序列和所述冠状病毒合成RNA;以及(e)使用下一代测序对所述扩增的生物样品进行测序。在某些实施方案中,所述方法还包括如果检测到来自所述冠状病毒核酸序列的序列读段,则提供针对冠状病毒感染的阳性诊断。

[0013] 在某些实施方案中,所述冠状病毒感染是SARS-Cov-2感染。在某些实施方案中,如果来自所述冠状病毒核酸序列的所述序列读段与所述冠状病毒核酸RNA的序列读段的比率或其数学等同形式超过约0.1的比率,则提供针对所述冠状病毒感染或SARS-Cov-2感染的所述阳性诊断。在某些实施方案中,如果来自所述冠状病毒核酸的所述序列读段和冠状病毒合成RNA的所述序列读段超过约100,则提供针对冠状病毒感染的所述阳性诊断。在某些实施方案中,在进行所述逆转录反应之前未对所述裂解的生物样品进行分离和纯化。在某些实施方案中,裂解所述生物样品和对所述裂解的生物样品进行所述逆转录反应在同一孔、管或反应容器中发生。在某些实施方案中,裂解所述生物样品包括热裂解。在某些实施方案中,所述热裂解包括将所述生物样品加热至至少约50°C的温度。在某些实施方案中,来自所述个体的所述生物样品包含多个冠状病毒合成RNA序列,其中所述多个冠状病毒合成RNA序列包含至少两个不同的合成冠状病毒RNA序列。在某些实施方案中,所述多个合成冠状病毒RNA序列包含至少四个不同的合成冠状病毒RNA核酸序列。在某些实施方案中,所述合成RNA核酸或所述多个合成核酸包含约20%至约30%的量的鸟嘌呤核苷酸、约20%至约30%的量的腺嘌呤核苷酸、约20%至约30%的量的胞嘧啶核苷酸、约20%至约30%的量的尿嘧啶核苷酸。在某些实施方案中,所述合成冠状病毒RNA核酸或所述多个合成冠状病毒RNA核酸包含大约等比率的鸟嘌呤:胞嘧啶:腺嘌呤:尿嘧啶。在某些实施方案中,所述合成冠状病毒RNA核酸或所述多个合成冠状病毒RNA核酸包含合成SARS-Cov-2 RNA核酸或多个合成SARS-Cov-2 RNA核酸。在某些实施方案中,所述方法还包括检测甲型流感病毒感染、乙

型流感病毒感染或其组合。在某些实施方案中,对所述裂解的生物样品的所述扩增反应采用对甲型流感病毒核酸序列具有特异性的甲型流感病毒引物集或对乙型流感病毒核酸序列具有特异性的乙型流感病毒引物集进行。在某些实施方案中,对所述甲型流感病毒核酸序列具有特异性的所述甲型流感病毒引物集包含SEQ ID NO:24或25和SEQ ID NO:26,或SEQ ID NO:27和SEQ ID NO:28中所示的序列。在某些实施方案中,对所述乙型流感病毒核酸序列具有特异性的所述乙型流感病毒引物集包含SEQ ID NO:29或30中所示的序列。在某些实施方案中,对所述裂解的生物样品的所述扩增反应采用对甲型流感病毒核酸序列具有特异性的甲型流感病毒引物集和对乙型流感病毒核酸序列具有特异性的乙型流感病毒引物集进行。在某些实施方案中,来自所述个体的所述生物样品还包含甲型流感病毒合成RNA、乙型流感病毒合成RNA或其组合,其中所述甲型流感病毒合成RNA、所述乙型流感病毒合成RNA或所述其组合与天然存在的甲型流感病毒或乙型流感病毒核酸序列不同。在某些实施方案中,所述方法还包括如果来自甲型流感病毒的序列读段与所述甲型流感病毒合成RNA的序列读段的比率或其数学等同形式超过约0.1的比率,则提供针对甲型流感病毒感染的阳性诊断。在某些实施方案中,所述方法还包括如果来自甲型流感病毒的序列读段与所述乙型流感病毒合成RNA的序列读段的比率或其数学等同形式超过约0.1的比率,则提供针对乙型流感病毒感染的阳性诊断。在某些实施方案中,所述冠状病毒核酸序列是N1序列、S2序列或其组合。在某些实施方案中,所述冠状病毒核酸序列是SARS-Cov-2 N1序列、SARS-Cov-2 S2序列或其组合。在某些实施方案中,所述冠状病毒引物集包含SEQ ID NO:13和SEQ ID NO:14,或SEQ ID NO:18和SEQ ID NO:19中所示的序列。在某些实施方案中,所述冠状病毒引物集包含SEQ ID NO:20和SEQ ID NO:21,或SEQ ID NO:22和SEQ ID NO:23中所示的序列。在某些实施方案中,所述冠状病毒引物集以约50纳摩尔至约250纳摩尔的浓度存在。在某些实施方案中,所述冠状病毒引物集以约100纳摩尔的浓度存在。在某些实施方案中,所述甲型流感病毒核酸序列是甲型流感病毒基质序列、甲型流感病毒非结构蛋白1序列、甲型流感病毒血凝素序列、甲型流感病毒神经氨酸酶序列、甲型流感病毒核蛋白序列及其组合。在某些实施方案中,所述乙型流感病毒核酸序列是乙型流感病毒基质序列、乙型流感病毒非结构蛋白1序列、乙型流感病毒血凝素序列、乙型流感病毒神经氨酸酶序列、乙型流感病毒核蛋白序列及其组合。在某些实施方案中,所述冠状病毒引物集、所述甲型流感病毒引物集或所述乙型流感病毒引物集中的任一种或多种包含允许样品复用的一个或多个索引序列。在某些实施方案中,所述冠状病毒合成RNA包含与SEQ ID NO:1至12的任一个中所示的序列中的任一个或多个具有至少约90%同源性的核酸序列。在某些实施方案中,所述冠状病毒合成RNA包含与SEQ ID NO:1至12的任一个中所示的序列中的任一个或多个具有至少约90%同源性的多个不同核酸序列。在某些实施方案中,所述甲型流感病毒合成RNA包含与SEQ ID NO:31或32的任一个中所示的序列中的任一个或多个具有至少约90%同源性的RNA序列。在某些实施方案中,所述乙型流感病毒合成RNA包含与SEQ ID NO:33中所示的序列具有至少约90%同源性的RNA序列。在某些实施方案中,所述冠状病毒合成RNA包含与SEQ ID NO:1至4中所示四个序列具有至少约90%同源性的多个四个不同核酸序列。在某些实施方案中,所述冠状病毒合成RNA、所述甲型流感病毒合成RNA或所述乙型流感病毒合成RNA以约10个拷贝/反应至约500个拷贝/反应的浓度存在。在某些实施方案中,所述冠状病毒合成RNA以约10个拷贝/反应至约500个拷贝/反应的浓度存在。在某些实施方案中,所述冠状病

毒合成RNA、所述甲型流感病毒合成RNA或所述乙型流感病毒合成RNA以约200个拷贝/反应的浓度存在。在某些实施方案中,所述冠状病毒合成RNA以约200个拷贝/反应的浓度存在。在某些实施方案中,所述生物样品包括鼻拭子或唾液样品。在某些实施方案中,所述生物样品包含所述个体的少于约10微升唾液或已用所述个体的鼻拭子接种的少于约10微升缓冲液。在某些实施方案中,所述生物样品包含已用鼻拭子接种的少于约10微升缓冲液。在某些实施方案中,对所述裂解的生物样品的所述扩增反应采用对样品对照具有特异性的引物对进行。在某些实施方案中,所述样品对照是持家基因。在某些实施方案中,对所述样品对照具有特异性的所述引物对对RPP30具有特异性。在某些实施方案中,对所述样品对照具有特异性的所述引物对包含SEQ ID NO:15或16和SEQ ID NO:17中所示的序列。

[0014] 本文还描述了一种合成核酸,其包含:5'近端区,其包含来自病毒的第一核苷酸序列;3'近端区,其包含来自所述病毒的第二核苷酸序列;和间插核苷酸序列,其中所述间插核苷酸序列包含百分比为约20%至约30%的鸟嘌呤核苷酸、量为约20%至约30%的腺嘌呤核苷酸、量为约20%至约30%的胞嘧啶核苷酸、量为约20%至约30%的尿嘧啶或胸苷核苷酸,并且所述间插序列不同于所述病毒的天然存在的序列。在某些实施方案中,所述合成核酸包含DNA。在某些实施方案中,所述合成核酸由DNA组成。在某些实施方案中,所述合成核酸包含RNA。在某些实施方案中,所述合成核酸由RNA组成。在某些实施方案中,所述病毒是甲型流感病毒、乙型流感病毒或冠状病毒。在某些实施方案中,所述病毒是冠状病毒。在某些实施方案中,所述冠状病毒是SARS-CoV-2。在某些实施方案中,所述间插核苷酸序列核酸包含大约等比率的鸟嘌呤核苷酸:胞嘧啶核苷酸:腺嘌呤核苷酸:尿嘧啶或胸苷核苷酸。在某些实施方案中,所述3'近端区和所述5'近端区包含与冠状病毒S2基因序列具有至少90%同源性的核苷酸序列。在某些实施方案中,所述5'近端区和所述3'近端区包含与冠状病毒S2基因序列具有至少95%同源性的核苷酸序列。在某些实施方案中,所述5'近端区和所述3'近端区包含与冠状病毒S2基因序列相同的核苷酸序列。在某些实施方案中,所述5'近端区和所述3'近端区包含与冠状病毒N1基因序列具有至少90%同源性的核苷酸序列。在某些实施方案中,所述5'近端区和所述3'近端区包含与冠状病毒N1基因序列具有至少95%同源性的核苷酸序列。在某些实施方案中,所述5'近端区和所述3'近端区包含与冠状病毒N1基因序列相同的核苷酸序列。在某些实施方案中,所述冠状病毒N1基因序列或所述冠状病毒S2基因序列是SARS-CoV-2基因序列。在某些实施方案中,所述序列核酸包含与SEQ ID NO:1至12的任一个中所示的任一个或多个序列具有至少90%同源性的序列。在某些实施方案中,所述序列核酸包含与SEQ ID NO:1至12的任一个中所示的任一个或多个序列具有至少95%同源性的序列。在某些实施方案中,所述序列核酸包含与SEQ ID NO:1至12的任一个中所示的任一个或多个序列相同的序列。在某些实施方案中,本文描述了多个合成核酸,其中所述多个合成核酸包括包含至少两个不同核苷酸序列的合成核酸。在某些实施方案中,所述多个合成核酸包括包含至少两个不同核苷酸序列的合成核酸。在某些实施方案中,所述多个合成核酸包括至少四个不同核苷酸序列。在某些实施方案中,所述四个不同核苷酸序列是SEQ ID NO:1至4中所示的那些。在某些实施方案中,所述四个不同核苷酸序列选自SEQ ID NO:5至12中所示的那些。

[0015] 本文还描述了一种反应混合物,其用于确定生物样品中是否存在病毒核酸,所述反应混合物包含本文所述的合成核酸或本文所述的多个合成核酸、所述生物样品的至少一

部分和足以在所述病毒核酸存在时扩增所述生物样品中的所述病毒核酸的一种或多种酶或试剂。在某些实施方案中,所述生物样品是人生物样品。在某些实施方案中,所述生物样品包括唾液、口腔拭子、鼻咽拭子或中鼻甲拭子。在某些实施方案中,所述生物样品包括唾液或鼻咽拭子。在某些实施方案中,所述病毒核酸是甲型流感病毒、乙型流感病毒或冠状病毒核酸。在某些实施方案中,所述冠状病毒核酸是Sars-Cov-2核酸。在某些实施方案中,在某些实施方案中,所述一种或多种试剂选自逆转录酶、dNTP、对所述病毒核苷酸序列具有特异性的引物对、对样品对照核苷酸序列具有特异性的引物对、镁盐及其组合。在某些实施方案中,对所述样品对照核苷酸序列具有特异性的所述引物对对人核苷酸序列具有特异性。在某些实施方案中,对所述样品对照核苷酸序列具有特异性的所述引物对对持家基因具有特异性。在某些实施方案中,对所述样品对照具有特异性的所述引物对对RPP30具有特异性。在某些实施方案中,对所述样品对照具有特异性的所述引物对包含SEQ ID NO:15或16和SEQ ID NO:17中所示的序列。在某些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对对甲型流感病毒核苷酸序列、乙型流感病毒核苷酸序列和冠状病毒核苷酸序列具有特异性。在某些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对对冠状病毒S1或N2序列具有特异性。在某些实施方案中,所述冠状病毒S1或N2序列是冠状病毒S1或N2核酸序列。在某些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对包含SEQ ID NO:13至30或100至605中的任一个中所示的序列。在某些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对包含SEQ ID NO:13和SEQ ID NO:14;SEQ ID NO:18和SEQ ID NO:19;SEQ ID NO:20和SEQ ID NO:21;SEQ ID NO:24或25和SEQ ID NO:26;SEQ ID NO:29和SEQ ID NO:30中所示的序列。在某些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对以约50微摩尔至约250微摩尔的浓度存在。在某些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对以约100微摩尔的浓度存在。在某些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对以约200微摩尔的浓度存在。在某些实施方案中,所述冠状病毒合成RNA以约10个拷贝/反应至约500个拷贝/反应混合物的浓度存在。在某些实施方案中,所述冠状病毒合成RNA以约200个拷贝/反应混合物的浓度存在。在某些实施方案中,所述反应混合物的体积为约10微升至约100微升。在某些实施方案中,所述反应混合物的体积为约20微升。

[0016] 本文还描述了一种试剂盒,其用于确定生物样品中是否存在病毒核酸,所述试剂盒包括本文所述的合成核酸或多个本文所述的合成核酸的和足以扩增来自所述生物样品的所述病毒核酸的一种或多种酶或试剂。在某些实施方案中,所述病毒核酸是甲型流感病毒、乙型流感病毒或冠状病毒核酸。在某些实施方案中,所述冠状病毒核酸是冠状病毒核酸。在某些实施方案中,所述一种或多种试剂选自逆转录酶、dNTP、对所述病毒核苷酸序列具有特异性的引物对、对样品对照核苷酸序列具有特异性的引物对、镁盐及其组合。在某些实施方案中,对所述样品对照核苷酸序列具有特异性的所述引物对对人核苷酸序列具有特异性。在某些实施方案中,对所述样品对照核苷酸序列具有特异性的所述引物对对持家基因具有特异性。在某些实施方案中,对所述样品对照具有特异性的所述引物对对RPP30具有

特异性。在某些实施方案中,对所述样品对照具有特异性的所述引物对包含SEQ ID NO:15或16和SEQ ID NO:17中所示的序列。在某些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对对甲型流感病毒核苷酸序列、乙型流感病毒核苷酸序列和冠状病毒核苷酸序列具有特异性。在某些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对对冠状病毒S1或N2序列具有特异性。在某些实施方案中,所述冠状病毒S1或N2序列是冠状病毒S1或N2核酸序列。在某些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对包含SEQ ID NO:13至30或100至605中的任一个中所示的序列。在某些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对包含SEQ ID NO:13和SEQ ID NO:14;SEQ ID NO:18和SEQ ID NO:19;SEQ ID NO:20和SEQ ID NO:21;SEQ ID NO:24或25和SEQ ID NO:26;SEQ ID NO:29和SEQ ID NO:30中所示的序列。

附图说明

[0017] 本文所述的新颖特征在所附权利要求中具体阐述。通过参考阐述其中利用了本文所述特征原理的说明性实例的以下具体实施方式和附图将更好地理解本文所述的特征和特征优点,在所述附图中:

[0018] 图1示出了病毒感染的诊断的示例性示意图。

[0019] 图2A和2B示出了说明样品中SARS-CoV-2核酸的检测的数据。

[0020] 图3示出了根据本公开的示例性引发方案。

[0021] 图4示出了针对COVID19的SwabSeq诊断测试平台。

[0022] 图5显示了在临床样本中的验证并且说明了等同于灵敏的RT-QPCR反应的检测限。

[0023] 图6显示了测序文库设计。

[0024] 图7显示了当扩增COVID-19扩增子和合成S2标 (spike) 时,S2引物显示等同的PCR效率。

[0025] 图8显示了在非常高的病毒浓度下,SwabSeq维持线性。

[0026] 图9显示了在MiSeq或NextSeq机器上进行的测序表现出相似的灵敏度。

[0027] 图10显示了使用NextSeq550的RNA纯化样品的初步和确证性检测线数据。

[0028] 图11显示了向传统收集介质和缓冲液中的无提取方案需要稀释以克服RT和PCR抑制的影响。

[0029] 图12显示了实现每天对数千个样品进行可扩展测试的轻量级样品登记、收集和处理系统的示例性开发。

[0030] 图13显示将唾液预加热至95C持续30分钟改进了RT-PCR。

[0031] 图14显示PCR抑制对扩增产物有显著影响。

[0032] 图15显示观察到增加PCR循环数目并且用未纯化或抑制性样品类型 (例如,唾液) 工作的tapestation使文库制备中非特异性峰的大小增加。

[0033] 图16显示了相对于NEB Luna,TaqPath减少了SARS-CoV2阴性样品中S2读段的数目。

[0034] 图17显示了来自MiSeq的模板线的滞留 (carryover) 污染导致交叉污染。

[0035] 图18显示了扩增子读段中的测序错误和潜在的扩增子错误分配。

- [0036] 图19显示了不同索引策略的可视化。
- [0037] 图20显示了使用混合模型对索引错误分配的计算校正。
- [0038] 图21显示了对索引错误分配作为S2读段中的噪声源的作用进行定量。
- [0039] 图22显示了降低的引物浓度对引物二聚体和非特异性扩增产物的影响。
- [0040] 图23显示了多样化合成核酸加标序列。
- [0041] 图24显示了以N1加标(上图)或S2加标(下图)和不同扩增子N1(上图)或S2(下图)的检测所获得的数据。
- [0042] 图25显示了与单个地相比,将N1和S2组合可以提高SARS-CoV2检测的灵敏度。
- [0043] 图26显示了使用不同体积的唾液样品进行检测的结果。
- [0044] 图27显示了使用不同体积的鼻拭子样品进行检测的结果。
- [0045] 图28显示了SwabSeq可以检测甲型流感病毒或乙型流感病毒。
- [0046] 图29显示了用于检出阳性样品和阴性样品的示例性算法。

具体实施方式

[0047] 高传染性或高致病性疾病的爆发频率日益增加。全世界三分之一的死亡归因于感染性疾病,这是全世界死亡和残疾的第二大原因。为了鉴别和诊断使人患病的感染性疾病而获得快速、准确的读出以便是诊断医学,特别是在病毒感染情况下的的重要组成部分,以便实施有效的公共卫生响应并改进人医疗保健的提供。出于临床诊断目的,已经开发了許多用于检测病毒感染的方法,但是这些测试中的大多数未提供足够快速或高通量的读出并且/或者由于与快速测试数目快速增加的受试者相关的资源负担而不可行。

[0048] 此外,本文提供了诊断病毒感染的个体的方法和系统。如本文所述的此类方法和系统利用聚合酶链反应(PCR)、文库制备策略和下一代测序以产生准确鉴别病毒感染的读出或信息,并且可以被有效扩展以满足与对有效测试数目快速增加的受试者的需要相关的挑战。为了促进样品内病毒感染的鉴定和诊断,所述方法提供了在以下方面中的进步或针对以下的解决方案:(1)有效且高效地从病毒分离出核酸序列;(2)有效且高效地处理对应于或来源于病毒的核酸分子;以及(3)有效且高效地复用样品,其实现了平行测试多个样品并且减少测试的总体资源负担。

[0049] 本文描述了一种检测样品中病原体基因组的方法,所述方法包括使用第一PCR引物集扩增来自样品的核酸,获得扩增的核酸,其中所述第一PCR引物集扩增病毒核酸序列和合成核酸序列。与没有合成核酸的扩增相比,合成核酸序列在测定中提供样品和扩增对照,实现检测下限。可以将合成核酸添加(“加标(spiked)”)到已经与生物样品接触的裂解物中,或者它可以在裂解缓冲液与生物样品接触之前存在于裂解缓冲液中。样品可以是生物样品。生物样品可以来自个体。可以在此过程中的任何步骤添加合成核酸。在本文所述的方法中可以使用多于一种合成核酸。例如,一种合成核酸可以用作一种或多种病毒核酸的对照,并且一种合成核酸可以用作一种或多种人核酸的对照(例如,持家对照),或者例如可以添加多个合成核酸以用作多个病原体序列的对照。

[0050] 本文还描述了一种合成核酸,其包含5'近侧区、3'近侧区和间插核酸序列。合成核酸可以在5'末端和3'末端包含允许病原体序列扩增的任何序列,前提是5'末端与3'末端之间的序列可以通过测序进行区分。在某些实施方案中,除与病原体序列不同的1、2、3、4、5、

6、7、8、9或10个核苷酸的序列之外,5'末端与3'末端之间的序列与病原体序列相同。

[0051] 本文还描述了一种合成核酸,其包含5'近侧区、3'近侧区和间插核酸序列。合成核酸可以在5'末端和3'末端包含允许病原体序列扩增的任何序列,前提是5'末端与3'末端之间的序列可以通过测序进行区分。在某些实施方案中,除与病原体序列不同的1、2、3、4、5、6、7、8、9或10个核苷酸的序列之外,5'末端与3'末端之间的序列与病原体序列相同。

[0052] 本文还描述了一种检测个体的冠状病毒感染的方法,所述方法包括:(a)提供来自所述个体的生物样品,其中所述生物样品包含冠状病毒合成RNA,其中所述冠状病毒合成RNA的序列与天然存在的冠状病毒核酸序列不同;(b)裂解所述生物样品,从而产生裂解的生物样品;(c)对所述裂解的生物样品进行逆转录反应,以获得裂解的逆转录的生物样品;(d)对所述裂解的逆转录的生物样品进行扩增反应,以获得扩增的生物样品,其中对所述裂解的逆转录生物的样品的所述扩增反应用对冠状病毒核酸序列具有特异性的冠状病毒引物集进行,其中所述冠状病毒引物集扩增所述冠状病毒核酸序列和所述冠状病毒合成RNA;以及(e)使用下一代测序对所述扩增的生物样品进行测序。在某些实施方案中,所述方法还包括如果检测到来自所述冠状病毒核酸序列的序列读段,则提供针对冠状病毒感染的阳性诊断。在某些实施方案中,所述冠状病毒感染是SARS-Cov-2感染。在某些实施方案中,如果来自所述冠状病毒核酸序列的所述序列读段与所述冠状病毒核酸RNA的序列读段的比率或其数学等同形式超过约0.1的比率,则提供针对所述冠状病毒感染或SARS-Cov-2感染的所述阳性诊断。在某些实施方案中,如果来自所述冠状病毒核酸的所述序列读段和冠状病毒合成RNA的所述序列读段超过约100,则提供针对冠状病毒感染的所述阳性诊断。

[0053] 本文还描述了一种检测个体的冠状病毒感染的方法,所述方法包括:(a)提供来自所述个体的生物样品,其中所述生物样品包含冠状病毒合成RNA,其中所述冠状病毒合成RNA的序列与天然存在的冠状病毒核酸序列不同;(b)裂解所述生物样品,从而产生裂解的生物样品;(c)对所述裂解的生物样品进行逆转录反应,以获得裂解的逆转录的生物样品;(d)对所述裂解的逆转录的生物样品进行扩增反应,以获得扩增的生物样品,其中对所述裂解的逆转录的生物样品的所述扩增反应用对冠状病毒核酸序列具有特异性的冠状病毒引物集进行,其中所述冠状病毒引物集扩增所述冠状病毒核酸序列和所述冠状病毒合成RNA。在某些实施方案中,所述冠状病毒感染是SARS-Cov-2感染。

[0054] 本文还描述了一种检测个体的冠状病毒感染的方法,所述方法包括:(a)提供来自所述个体的生物样品,其中所述生物样品包含冠状病毒合成RNA,其中所述冠状病毒合成RNA的序列不同于天然存在的冠状病毒核酸序列;(b)裂解所述生物样品,从而产生裂解的生物样品;(c)对所述裂解的生物样品进行逆转录反应,以获得裂解的逆转录生物样品。在某些实施方案中,所述冠状病毒感染是SARS-Cov-2感染。

[0055] 还提供了包含合成核酸分子的组合物并且所述组合物可用于本文公开的方法。合成核苷酸通常是体外转录或合成对照RNA,除允许区分对应于合成对照的测序读段和对应于病原体序列的测序读段的短的、改变的链段(stretch)之外,它与目标用于扩增的病毒序列相同。例如,公开了组合物,其包含有包含第一核酸序列和第二核酸序列的合成核酸分子,其中第一核酸序列与病原体核酸分子的序列相同,并且第二核酸序列与病原体核酸分子的序列不相同。在一些实施方案中,第一核酸序列位于合成核酸分子的3'区。在一些实施方案中,合成核酸分子还包含第三核酸序列,其中第一核酸序列与来自病原体核酸分子的

第一序列相同,并且其中第三核酸序列与来自病原体核酸分子的第二序列相同。在一些实施方案中,第一核酸序列位于第二核酸分子的3'区,并且第三核酸序列位于第二核酸分子的5'区。在一些实施方案中,第二核酸序列少于5、10、15、20、25或30个核苷酸。在一些实施方案中,第二合成核酸分子包含总数少于25、50、100、150、200或500个核苷酸的核苷酸。在一些实施方案中,第二合成核酸分子包含总数大于25、50、100、150、200或500个核苷酸的核苷酸。在一些实施方案中,合成核酸分子是核糖核酸(RNA)分子、脱氧核糖核酸(DNA)分子或RNA-DNA杂交分子。在某些实施方案中,合成核酸分子以在对应病原体序列的约10%内的效率进行PCR扩增。在某些实施方案中,合成核酸分子以在对应病原体序列的约5%内的效率进行PCR扩增。在某些实施方案中,合成核酸分子以与对应病原体序列相同的效率进行PCR扩增。

[0056] 根据本文所述的方法,可以使用多个合成核酸以进一步增加灵敏度,减少假阳性或提高核酸序列的定量的准确度和/或精密度。所述多个可以包含2、3、4、5、6、7、8、9或更多个不同序列,所述序列能够与对待检测的病原体序列具有特异性的引物集共同扩增。所述多个可以具有所述多个所需的某些特征。在某些实施方案中,所述多个的不同序列的解链温度可以实质上相同或在所述多个的平均解链温度的约0.5、1、2、3、4或5°C内。在某些实施方案中,所述多个的核苷酸构成应为在约30%与约20%之间的A、在约30%与约20%之间的G、在约30%与约20%之间的C、在约30%与约20%之间的T。在某些实施方案中,所述多个的核苷酸构成应为约25% A、约25% G、约25% C、约25% T。在某些实施方案中,所述多个的核苷酸构成应为在约30%与约20%之间的A、在约30%与约20%之间的G、在约30%与约20%之间的C、在约30%与约20%之间的T中的一种或多种。在某些实施方案中,所述多个的核苷酸构成应为约25% A、约25% G、约25% C、约25% T中的一种或多种。在某些实施方案中,选择或设计多种合成核酸以最小化所述多个的不同序列之间的二级结构或二聚化。

[0057] 本文还描述了一种诊断病毒感染的个体的方法,所述方法包括:(a)提供来自所述个体的生物样品;(b)将来自所述个体的所述生物样品与裂解剂接触,以获得裂解的生物样品;(c)对所述裂解的生物样品进行逆转录反应;(d)对所述裂解的生物样品进行聚合酶链反应(PCR),以获得PCR扩增的裂解的生物样品,其中对所述裂解的生物样品的所述PCR反应用第一PCR引物集和第二PCR引物集进行,其中所述第一PCR引物集扩增病毒核酸序列;其中所述第二引物集扩增所述个体所属的物种的基因组序列;(e)使用下一代测序对所述PCR扩增的裂解的生物样品进行测序;以及(f)任选地,如果通过所述PCR或通过所述测序检测到病毒序列,则提供针对所述病毒感染的阳性诊断,或者如果通过所述PCR或通过所述测序未检测到病毒序列,则提供针对所述人类个体的阴性诊断。在某些实施方案中,所述第一PCR引物集可以扩增裂解的生物样品中存在的合成核酸序列。合成核酸序列包含与病毒核酸序列相同的引物结合位点,除了间插核酸序列不同,使得可以通过测序进行区分。在某些实施方案中,合成核酸序列是RNA。

[0058] 本文还描述了一种诊断病原体感染的个体的方法,所述方法包括:(a)提供来自所述个体的生物样品;(b)将来自所述个体的所述生物样品与裂解剂接触,以获得裂解的生物样品;(c)对所述裂解的生物样品进行聚合酶链反应(PCR),以获得PCR扩增的裂解的生物样品,其中对所述裂解的生物样品的所述PCR反应用第一PCR引物集和第二PCR引物集进行,其中所述第一PCR引物集扩增病原体核酸序列,其中所述第二引物集扩增所述个体的核酸序

列；(d) 使用下一代测序对所述PCR扩增的裂解的生物样品进行测序；以及(d) 任选地，如果通过所述PCR或通过所述测序检测到病原体序列，则提供针对所述病原体感染的阳性诊断，或者如果通过所述PCR或通过所述测序未检测到病原体序列，则提供针对所述人类个体的阴性诊断。在某些实施方案中，所述第一PCR引物集可以扩增裂解的生物样品中存在的合成核酸序列。合成核酸序列包含与病毒核酸序列相同的引物结合位点，除了间插核酸序列不同，使得可以通过测序进行区分。在某些实施方案中，合成核酸序列是RNA。

[0059] 本文所述的方法可用于监测任何数目的样品中病原体的存在，包括非人样品。监视可以包括监测一群驯养或野生动物、野生动物群体或围封的活体动物（例如，动物园、野生动物公园或出售动物用作食物或作为宠物的活体动物市场）。

[0060] 本文还描述了一种监视样品中病原体的存在的方法，所述方法包括：(a) 提供样品；(b) 将所述样品与提取剂接触，以获得提取的样品；(c) 对所述提取的样品进行聚合酶链反应(PCR)，以获得PCR扩增的提取的样品，其中对所述提取的样品的所述PCR反应用第一PCR引物集进行；(d) 使用下一代测序对所述PCR扩增的裂解的生物样品进行测序；以及(e) 任选地，如果通过所述PCR或通过所述测序检测到病原体序列，则提供针对所述病原体的阳性读出，或者如果通过所述PCR或通过所述测序未检测到病原体序列，则提供阴性读出。在某些实施方案中，所述第一PCR引物集可以扩增提取的生物样品中存在的合成核酸序列。合成核酸序列包含与病原体核酸序列相同的引物结合位点，除了间插核酸序列不同，使得可以通过测序进行区分。在某些实施方案中，合成核酸序列是RNA。在某些实施方案中，合成核酸序列是DNA。

[0061] 为了实现对病毒感染的鉴定、检测和/或诊断，本文所公开的方法和系统包括：(a) 提供来自所述个体的生物样品；(b) 将来自所述个体的所述生物样品与裂解剂接触，以获得裂解的生物样品；(c) 对所述裂解的生物样品进行初始核酸延伸反应；(d) 对所述裂解的生物样品进行聚合酶链反应(PCR)，以获得PCR扩增的裂解的生物样品，其中对所述裂解的生物样品的所述PCR反应用第一PCR引物集和第二PCR引物集进行，其中所述第一PCR引物集扩增病毒核酸序列；其中所述第二引物集扩增所述个体所属的物种的基因组序列；(e) 使用下一代测序对所述PCR扩增的裂解的生物样品进行测序；以及(f) 如果通过所述PCR或通过所述测序检测到病毒序列或其衍生物，则提供针对病毒感染的阳性诊断，或者如果通过所述PCR或通过所述测序未检测到冠状病毒序列，则提供针对所述人类个体的阴性诊断。

[0062] 本文还公开了用于检测病毒感染的核酸处理方法。此类方法包括：(a) 提供样品，其包含病毒核酸分子和宿主核酸分子；(b) 通过使用包含条形码序列的第一引物对所述病毒核酸分子进行核酸延伸反应，生成条形码化病毒核酸分子；(c) 通过使用包含所述条形码序列的第二引物对所述宿主核酸分子进行核酸延伸反应，生成条形码化宿主核酸分子；(d) 对所述条形码化病毒核酸分子和所述条形码化宿主核酸分子进行测序，以鉴定(i) 所述条形码序列和(ii) 对应于所述病毒核酸分子或其衍生物和所述宿主核酸分子的序列；以及(e) 如果在(d) 中鉴定出对应于所述病毒核酸分子的所述序列，则提供针对病毒感染的阳性诊断。

[0063] 如本文所用的术语“条形码”通常是指传达或能够传达关于分析物的信息的标记或标识符。条形码可以是分析物的一部分。条形码可以与分析物无关。除了分析物的内源特征（例如，分析物的大小或末端序列）之外，条形码可以是连接至分析物（例如，核酸分子）的

标签或标签的组合。条形码可以是唯一的。条形码可以有多种不同的格式。例如,条形码可包括:多核苷酸条形码;随机核酸和/或氨基酸序列;和合成核酸和/或氨基酸序列。条形码可以以可逆或不可逆的方式连接至分析物。在对样品进行测序之前、期间和/或之后,可以将条形码添加至例如脱氧核糖核酸(DNA)或核糖核酸(RNA)样品的片段。条形码可以允许鉴定和/或定量各个测序读段。

[0064] 如本文所用的术语“实时”可以指小于约1秒、十分之一秒、百分之一秒、毫秒或更小的响应时间。响应时间可能大于1秒。在一些情况下,实时可以指同时或实质上同时的处理、检测或鉴定。

[0065] 如本文所用的术语“基因组”是指来自植物、动物、细菌、真菌或病毒的基因组信息,其可以是例如受试者遗传信息的至少一部分或全部。基因组可以在DNA或RNA中编码。基因组可包含编码区(例如,蛋白质的编码区)以及非编码区。基因组可以包括生物体中所有染色体一起的序列。例如,人基因组普通具有总计46条染色体。所有这些染色体一起的序列可以构成人基因组。

[0066] 术语“衔接子(adaptor)”、“衔接子(adapter)”和“标签”可以同义地使用。衔接子或标签可以通过任何方法,包括连接、杂交或其他方法与待“加标签”的多核苷酸序列联接。

[0067] 如本文所用的术语“测序”通常是指用于确定一种或多种多核苷酸中核苷酸碱基的序列的方法和技术。多核苷酸可以是例如核酸分子,诸如脱氧核糖核酸(DNA)或核糖核酸(RNA),包括其变体或衍生物(例如,单链DNA)。“下一代测序”是指不为桑格(Sanger)测序的高通量测序方法。测序可以通过目前可用的各种系统执行,诸如但不限于通过Illumina®(例如,iSeq 100、MiniSeq、MiSeq或NextSeq系列机器)、Pacific Biosciences (PacBio®)、Oxford Nanopore®或Life Technologies (Ion Torrent®)的测序系统。替代地或另外,可以使用核酸扩增、聚合酶链反应(PCR)(例如,数字PCR、定量PCR或实时PCR)或等温扩增来进行测序。此类系统可以提供对应于受试者(例如,人)的遗传信息的多个原始遗传数据,如从受试者提供的样品由系统所生成。在一些实例中,此类系统提供测序读段(本文中也称为“读段”)。读段可以包括一串核苷酸碱基,其对应于已经测序的核酸分子的序列。在一些情况下,本文所提供的系统和方法可以与蛋白质组信息一起使用。

[0068] 如本文所用的术语“珠子”通常是指颗粒。珠子可以是固体或半固体颗粒。珠子可以是凝胶珠子。凝胶珠子可以包括聚合物基体(例如,通过聚合或交联形成的基体)。聚合物基体可以包括一种或多种聚合物(例如,具有不同官能团或重复单元的聚合物)。聚合物基体中的聚合物可以随机排列,诸如在无规共聚物中,并且/或者具有有序结构,诸如在嵌段共聚物中。交联可以通过共价、离子或诱导、相互作用或物理缠结。珠子可以是高分子。珠子可以由结合在一起的核酸分子形成。珠子可以通过分子(例如,大分子)的共价或非共价组装形成,诸如单体或聚合物。此类聚合物或单体可以是天然的或合成的。此类聚合物或单体可以是或包括例如核酸分子(例如,DNA或RNA)。珠子可以由聚合物材料形成。珠子可以是磁性的或非磁性的。珠子可以是刚性的。珠子可以是柔性的和/或可压缩的。珠子可以是可破坏的或可溶解的。珠子可以用包含一种或多种聚合物的涂层覆盖的固体颗粒(例如,包括但不限于氧化铁、金或银的基于金属的颗粒)。此类涂层可以是可破坏的或可溶解的。

[0069] 如本文所用的术语“样品”被广泛使用并且可以指环境样品(例如,水样品、污水样品)、生的或准备好的食物样品、从非人类个体群体(例如,野生或驯养动物)得到的样品或

来自人或非人动物的生物样品。样品可以包含任何数目的大分子,例如细胞大分子。样品可以是细胞样品。样品可以是细胞系或细胞培养物样品。样品可以包含一个或多个细胞。样品可以包含一种或多种微生物。生物样品可以是核酸样品或蛋白质样品。生物样品还可以是碳水化合物样品或脂质样品。生物样品可以来源于另一样品。样品可以是组织样品,例如活检、核芯活检(core biopsy)、针抽出物或细针抽出物。样品可以是流体样品,诸如血液样品、尿液样品或唾液样品。样品可以是皮肤样品。样品可以是口腔拭子。样品可以是鼻咽拭子。样品可以是血浆或血清样品。样品可以是无细胞(cell-free或cell free)样品。无细胞样品可以包含细胞外多核苷酸。细胞外多核苷酸可以从身体样品中分离,所述身体样品可以选自自由血液、血浆、血清、尿液、唾液、粘膜分泌物、痰液、粪便和泪液组成的组。

[0070] 如本文所用的术语“生物颗粒”通常是指来源于生物样品的离散生物系统。生物颗粒可以是生物大分子。生物颗粒可以是小分子。生物颗粒可以是病毒。生物颗粒可以是细胞或细胞衍生物。生物颗粒可以是细胞器。生物颗粒可以是来自细胞群的稀有细胞。生物颗粒可以是任何类型的细胞,包括但不限于原核细胞、宿主细胞、细菌、真菌、植物、哺乳动物或其他动物细胞型、支原体、正常组织细胞、肿瘤细胞或任何其他细胞型,无论是来源于单细胞还是多细胞生物。生物颗粒可以是细胞的成分。生物颗粒可以是或可以包括DNA、RNA、细胞器、蛋白质或其任何组合。生物颗粒可以是或可以包括基体(例如,凝胶或聚合物基体),其包含细胞或来自细胞的一种或多种组分(例如,细胞珠子),例如来自细胞的DNA、RNA、细胞器、蛋白质或其任何组合。生物颗粒可以从受试者的组织获得。生物颗粒可以是硬化细胞。这种硬化细胞可以包括或不包括细胞壁或细胞膜。生物颗粒可以包括细胞的一种或多种成分,但不包括细胞的其他成分。这些成分的实例是细胞核或细胞器。细胞可以是活细胞。活细胞可以能够被培养,例如,当被包封在凝胶或聚合物基体中时被培养,或者当包含凝胶或聚合物基体时被培养。

[0071] 如本文所用的术语“病原体”包括能够引起个体群体的疾病的任何生物体或病毒,此类群体可以包括动物或植物。这还涵盖携带个体或物种中存在的病原体,其中病原体不引起携带个体或物种的疾病,但可以传播至另一个体或物种以引起疾病。如本文所用,病原体包括但不限于细菌、原生动物、真菌、线虫、类病毒和病毒或其任何组合,其中每种病原体本身或与另一病原体一起能够引发脊椎动物的疾病,脊椎动物包括但不限于哺乳动物,且包括但不限于人。如本文所用,术语“宿主”是指可被病原体感染的生物体,并且包括植物、动物、脊椎动物、哺乳动物、啮齿动物、母牛、马、猪、家禽、雏鸡、鹅、鸭、鱼、贝类等等。

[0072] 如本文所用,术语“细菌”或“真细菌”是指原核生物的领域。细菌包括如下至少11个不同的组:(1)革兰氏阳性细菌,其中有两个主要细分:(i)高G+C组(放线菌(Actinomycetes)、分枝杆菌(Mycobacteria)、微球菌(Micrococcus)等);(ii)低G+C组(芽孢杆菌(Bacillus)、梭菌(Clostridia)、乳杆菌(Lactobacillus)、葡萄球菌(Staphylococci)、链球菌(Streptococci)、支原体(Mycoplasma));(2)变形菌(Proteobacteria),例如紫色光合+非光合革兰氏阴性菌(包括最“常见”的革兰氏阴性细菌);(3)蓝细菌(Cyanobacteria),例如产氧光合细菌(oxygenic phototroph);(4)螺旋体(Spirochete)及相关物种;(5)浮霉菌属(Planctomyces);(6)拟杆菌属(Bacteroides)、黄杆菌属(Flavobacteria);(7)衣原体(Chlamydia);(8)绿色硫细菌(Green sulfur bacteria);(9)绿色非硫细菌(Green non-sulfur bacteria)(还称为厌氧光合细菌);(10)

耐辐射微球菌及其亲属；(11)嗜热热袍菌(Thermotoga)和栖热腔菌(Thermosipho)。“革兰氏阴性细菌”包括球菌、非肠道杆菌和肠杆菌。革兰氏阴性菌的属包括例如奈瑟氏菌属(Neisseria)、螺菌属(Spirillum)、巴斯德菌属(Pasteurella)、布鲁菌属(Brucella)、耶尔森菌属(Yersinia)、弗朗西斯氏菌属(Francisella)、嗜血杆菌属(Haemophilus)、博德特氏菌属(Bordetella)、埃希氏菌属、沙门氏菌属、志贺氏菌属、克雷伯氏菌属(Klebsiella)、变形杆菌属(Proteus)、弧菌属(Vibrio)、假单胞菌属、拟杆菌属(Bacteroides)、醋杆菌属(Acetobacter)、气杆菌属(Aerobacter)、土壤杆菌属(Agrobacterium)、固氮菌属(Azotobacter)、螺旋状菌属(Spirilla)、沙雷氏菌属(Serratia)、弧菌属(Vibrio)、根瘤菌属(Rhizobium)、衣原体属(Chlamydia)、立克次体属(Rickettsia)、密螺旋体属(Treponema)和梭杆菌属(Fusobacterium)。“革兰氏阳性细菌”包括球菌、非产孢子杆菌和产孢子杆菌。革兰氏阳性细菌的属包括例如放线菌属(Actinomyces)、芽孢杆菌属(Bacillus)、梭菌属、棒状杆菌属(Corynebacterium)、丹毒丝菌属(Erysipelothrix)、乳杆菌属(Lactobacillus)、李斯特菌属(Listeria)、分枝杆菌属(Mycobacterium)、粘球菌属(Myxococcus)、诺卡氏菌属(Nocardia)、葡萄球菌属(Staphylococcus)、链球菌属和链霉菌属(Streptomyces)。“致病性细菌(pathogenic bacteria或pathogenic bacterium)”是通过直接感染另一生物体或通过产生在另一生物体中引起疾病的因子(例如,产生病原性毒素等的细菌)来引起另一宿主生物体(例如,动物和植物)的疾病的细菌物种。

[0073] 如本文所用的术语“大分子成分”通常是指包含在生物颗粒内或来自生物颗粒的大分子。大分子成分可以包含核酸。在一些情况下,生物颗粒可以是或包含大分子。大分子成分可以包含DNA。大分子成分可以包含RNA。RNA可以是编码的或非编码的。例如,RNA可以是信使RNA(mRNA)、核糖体RNA(rRNA)或转移RNA(tRNA)。RNA可以是转录物。RNA可以是长度小于200个核苷酸基的小RNA或者长度大于200个核苷酸基的大RNA。小RNA可以包括5.8S核糖体RNA(rRNA)、5S rRNA、转移RNA(tRNA)、微RNA(miRNA)、小干扰RNA(siRNA)、小核仁RNA(snoRNA)、Piwi相互作用RNA(piRNA)、tRNA衍生的小RNA(tsRNA)和小rDNA衍生的RNA(srRNA)。RNA可以是双链RNA或单链RNA。RNA可以是环状RNA。大分子成分可以包含蛋白质。大分子成分可以包含肽。大分子成分可以包含多肽。

[0074] 如本文所用的术语“分子标签”通常是指能够结合大分子成分分子。分子标签可以以高亲和力结合大分子成分。分子标签可以以高特异性结合大分子成分。分子标签可以包含核苷酸序列。分子标签可以包含核酸序列。核酸序列可以是分子标签的至少一部分或全部。分子标签可以是核酸分子或者可以是核酸分子的一部分。分子标签可以是寡核苷酸或多肽。分子标签可以包含DNA适体。分子标签可以是或包含引物。分子标签可以是或包含蛋白质。分子标签可以包含多肽。分子标签可以是条形码。

[0075] 术语“持家基因”、“持家控制”或本文所用的类似术语通常是指在正常和病理生理条件下在生物体中表达或者由不同组织和细胞类型表达的基因。在一些情况下,持家基因是维持基本细胞功能所需的组成型基因。在大多数正常和病理生理条件下,持家基因通常以相对恒定的速率表达。持家基因的具体实例包括但不限于RPP30、 β 肌动蛋白和/或GAPDH。

[0076] 如本文所用的术语“分区(partition)”通常是指适合于包含一种或多种物质或进行一个或多个反应的空间或体积。分区可以是物理隔室,例如液滴或孔。分区可以将空间或体积与另一个空间或体积隔离。液滴可以是与第一相不混溶的第二相(例如,油)中的第

一相(例如,水相)。液滴可以是不与第一相相分离的第二相中的第一相,例如像水相中的胶囊或脂质体。分区可以包括一个或多个其他(内部)分区。在一些情况下,分区可以是虚拟隔室,其可以由在多个和/或远程物理隔室上的索引(例如,索引库)来定义和鉴定。例如,物理隔室可包括多个虚拟隔室。

[0077] 除非上下文另外明确规定,如本文所用的术语“一个”、“一种”和“所述”通常指单数和复数指示物。

[0078] 每当术语“至少”、“大于”或“大于或等于”在一系列两个或更多个数值中的第一数值之前时,术语“至少”、“大于”或“大于或等于”适用于此系列数值中的每个数值。例如,大于或等于1、2或3等同于大于或等于1、大于或等于2或大于或等于3。

[0079] 每当术语“不大于”、“小于”或“小于或等于”在一系列两个或更多个数值中的第一数值之前时,术语“不大于”、“小于”或“小于或等于”适用于此系列数值中的每个数值。例如,小于或等于3、2或1等同于小于或等于3、小于或等于2或小于或等于1。

[0080] 以下描述中,阐述了某些具体细节以提供对各种实施方案的全面理解。然而,本领域技术人员将理解,所提供的实施方案可以在没有这些细节的情况下实践。除非上下文另外要求,否则在整个说明书和随后的权利要求中,词语“包含(comprise)”及其变体如“包含(comprises)”和“包含(comprising)”应以开放、包含性的含义解释,即“包括但不限于”。除非内容另外明确规定,否则如在本说明书和所附权利要求中所用的,单数形式“一个”、“一种”和“所述”包括复数指示物。还应当注意,除非内容另外明确规定,否则术语“或”通常以其包括“和/或”的含义使用。此外,本文提供的标题仅是为了方便,并且不解释所要求保护的实施方案的范围或含义。

[0081] 如本文所用,术语“约”是指在所述量附近10%或更少的量。

[0082] 如本文所用,术语“同源的”、“同源性”或“同源性百分比”当在本文中用于相对于参考序列描述氨基酸序列或核酸序列时可以使用由以下描述的公式来确定:Karlin和Altschul(Proc.Natl.Acad.Sci.USA 87:2264-2268,1990,如在Proc.Natl.Acad.Sci.USA 90:5873-5877,1993中那样修改)。这种公式被并入Altschul等人(J.Mol.Biol.215:403-410,1990)的基本局部比对搜索工具(BLAST)程序中。可以使用截至本申请提交日最新版本的BLAST来确定序列的同源性百分比。

[0083] 如本文所用,术语“个体”、“患者”或“受试者”是指被诊断患有、怀疑患有或有风险患上所述组合物和方法可用于检测的至少一种疾病的个体。在某些实施方案中,个体是哺乳动物。在某些实施方案中,哺乳动物是小鼠、大鼠、兔、狗、猫、马、母牛、绵羊、猪、山羊、美洲驼、羊驼或牦牛。在某些实施方案中,个体是人。

病原性感染

[0084] 本文所述的方法可以允许检测许多不同的病原体,包括细菌、病毒、真菌、原生生物、线虫和类病毒。

[0085] 例如,可以检测或诊断的病原体包括但不限于炭疽杆菌(*Bacillus anthracis*) (炭疽病(anthrax))、肉毒杆菌(*Clostridium botulinum*) 毒素(肉毒杆菌中毒(botulism))、鼠疫耶尔森氏菌(*Yersinia pestis*) (鼠疫(plague))。主天花(*Variola major*) (天花(smallpox))和其他相关痘病毒(pox virus)、土拉弗朗西斯菌(*Francisella tularensis*) (兔热病(tularemia))、病毒性出血热(Viral hemorrhagic fever)、沙粒病毒

(Arenavirus) (例如, 胡宁病毒 (Junin)、马秋波病毒 (Machupo)、瓜纳瑞托病毒 (Guanarito)、查帕雷病毒 (Chapare)、拉萨病毒 (Lassa) 和/或卢约病毒 (Lujo))、布尼亚病毒 (Bunyaviruses) (例如, 引起汉坦病毒肺综合征 (Hanta Pulmonary syndrome)、裂谷热 (Rift Valley Fever) 和/或克里米雅-刚果出血热 (Crimean Congo Hemorrhagic Fever) 的汉坦病毒 (Hantavirus))、黄病毒属 (Flavivirus)、登革病毒 (Dengue)、丝状病毒 (Filovirus) (例如, 埃博拉 (Ebola) 病毒和马尔堡 (Marburg) 病毒)、假鼻疽伯克霍尔德菌 (*Burkholderia pseudomallei*) (类鼻疽 (melioidosis))、伯纳特柯克斯体 (*Coxiella burnetii*) (Q热 (Q fever))、布鲁氏菌 (*Brucella*) 物种 (布鲁氏菌病 (brucellosis))、鼻疽伯克霍尔德菌 (*Burkholderia mallei*) (鼻疽病 (glanders))、鹦鹉热衣原体 (*Chlamydia psittaci*) (鹦鹉热 (Psittacosis))、蓖麻毒素 (Ricin toxin) (蓖麻 (*Ricinus communis*))、ε毒素 (Epsilon toxin) (产气夹膜梭状芽胞杆菌 (*Clostridium perfringens*))、葡萄球菌肠毒素B (*Staphylococcus enterotoxin B*, SEB)、斑疹伤寒 (Typhus fever) (普氏立克次体 (*Rickettsia prowazekii*))、食物传播和水传播病原体、致腹泻性大肠杆菌 (Diarrheagenic *E. coli*)、致病性弧菌 (Pathogenic *Vibrios*)、志贺氏菌属物种 (*Shigella species*)、沙门氏菌属、单核细胞增多性李斯特菌 (*Listeria monocytogenes*)、空肠弯曲菌 (*Campylobacter jejuni*)、耶尔森结肠炎菌 (*Yersinia enterocolitica*)、杯状病毒 (Caliciviruses)、甲型肝炎病毒、微小隐孢子虫 (*Cryptosporidium parvum*)、卡晏环孢子虫 (*Cyclospora cayatanensis*)、蓝氏贾第鞭毛虫 (*Giardia lamblia*)、溶组织内阿米巴 (*Entamoeba histolytica*)、刚地弓形虫 (*Toxoplasma gondii*)、福氏耐格里变形虫 (*Naegleria fowleri*)、狒狒巴拉姆希阿米巴 (*Balamuthia mandrillaris*)、真菌、微孢子虫 (Microsporidia)、蚊媒病毒 (Mosquito-borne virus) (例如, 西尼罗河病毒 (West Nile virus, WNV)、LaCrosse脑炎 (LACV)、加利福尼亚脑炎 (California encephalitis) 病毒、委内瑞拉马脑炎 (Venezuelan equine encephalitis, VEE) 病毒、东方马脑炎 (Eastern equine encephalitis, EEE) 病毒、西方马脑炎 (Western equine encephalitis, WEE) 病毒、日本脑炎病毒 (Japanese encephalitis virus, JE)、圣路易脑炎病毒 (St. Louis encephalitis virus, SLEV)、黄热病病毒 (Yellow fever virus, YFV)、基肯尼亚病毒 (Chikungunya virus)、寨卡病毒 (Zika virus)、尼帕病毒和亨德拉病毒 (Nipah and Hendra virus)、另外的汉坦病毒 (hantavirus)、蜱传出血热病毒 (Tickborne hemorrhagic fever virus)、布尼亚病毒 (Bunyavirus)、发热伴血小板减少综合征病毒 (Severe Fever with Thrombocytopenia Syndrome virus, SFTSV)、腹地病毒 (Heartland virus)、黄病毒属 (Flavivirus) (例如, 鄂木斯克出血热病毒 (Omsk Hemorrhagic Fever virus)、阿科赫玛病毒 (Alkhurma virus)、卡萨诺尔森林病病毒 (Kysanur Forest virus))、蜱传脑炎复合黄病毒 (Tickborne encephalitis complex flavivirus)、蜱传脑炎病毒 (Tickborne encephalitis virus)、波瓦桑/鹿蜱虫病毒 (Powassan/Deer Tick virus)、结核病 (Tuberculosis) 病毒 (包括耐药性TB)、流感病毒、狂犬病病毒 (Rabies virus)、朊病毒 (Prion)、链球菌属、假单胞菌属、志贺氏菌属、弯曲菌属、沙门氏菌属、梭菌属、埃希氏菌属、乙型肝炎病毒、丙型肝炎病毒、乳头瘤病毒、E-B病毒、水痘病毒、天花病毒、正粘病毒、严重急性呼吸综合征相关冠状病毒 (SARS-CoV)、SARS-CoV-2 (COVID-19)、MERS-CoV、其他高致病性人冠状病毒或它们的任何组合。

[0086] 本文所述的方法还可以用于检测诊断病毒。在某些实施方案中,病毒包括DNA病毒。在某些实施方案中,病毒包括RNA病毒。

[0087] 病毒核酸的检测是本文所述的检测和诊断病毒感染的方法的基础。对于病毒感染的检测和诊断,病毒核酸分子是正测试的病毒的基因组的一部分。在一些实施方案中,所述病毒是冠状病毒。在一些实施方案中,所述冠状病毒选自严重急性呼吸综合征冠状病毒2 (COVID-19)、严重急性呼吸综合征冠状病毒 (SARS-CoV) 和中东呼吸综合征冠状病毒 (MERS-CoV)。在一些实施方案中,所述冠状病毒是COVID-19。所公开的方法可用于其他RNA和DNA基因组病毒的分类。在一些实施方案中,所述病毒是RNA病毒。在一些实施方案中,所述RNA病毒包含双链RNA基因组。在一些实施方案中,所述RNA病毒包含单链RNA基因组。在一些实施方案中,所述RNA病毒选自冠状病毒、流感病毒、人免疫缺陷病毒和埃博拉病毒。在一些实施方案中,所述病毒是DNA病毒。在一些实施方案中,所述病毒感染是COVID-19。

[0088] 例如,本文的方法可用于检测冠状病毒。用于诊断冠状病毒感染的方法适用于其他病毒。冠状病毒是冠状病毒科 (Coronaviridae) 中的冠状病毒亚科 (Coronavirinae) 和网巢病毒目 (Nidovirales) 的成员 (国际病毒分类委员会 (International Committee on Taxonomy of Viruses))。冠状病毒亚科由四个属组成:α冠状病毒、β冠状病毒、γ 冠状病毒和δ冠状病毒。这些属基于系统发育关系和基因组结构进行区分。通常,α冠状病毒和β冠状病毒感染哺乳动物。另外,一般来说,γ 冠状病毒和δ冠状病毒感染鸟类,也可以感染哺乳动物。α冠状病毒和β冠状病毒可能与人呼吸道疾病和肠胃炎有关,并且是其病因。高致病性病毒(例如,严重急性呼吸综合征冠状病毒 (SARS-CoV) 和中东呼吸综合征冠状病毒 (MERS-CoV)) 可导致人严重呼吸综合征。其他四种冠状病毒(例如HCoV-NL63、HCoV-229E、HCoV-OC43和HKU1) 通常在免疫活性宿主、婴儿、幼儿和老年个体中诱导轻度上呼吸道疾病。α冠状病毒和β冠状病毒不仅可能给人带来沉重的疾病负担,还可能给牲畜带来沉重的疾病负担;这些病毒包括猪传染性胃肠炎病毒、猪肠道腹泻病毒 (PEDV) 和猪急性腹泻综合征冠状病毒 (SADS-CoV)。基于目前的序列数据库,所有人冠状病毒都具有人畜共患病起源。例如,SARS-CoV、MERS-CoV、HCoV-NL63和HCoV-229E被认为起源于蝙蝠。家畜可能作为中间宿主而发挥重要作用,它们使病毒能够从自然宿主传播给人。此外,家畜本身也可能罹患冠状病毒引起的疾病。

生物样品

[0089] 本文提供了使用和处理来自个体的生物样品以诊断病毒感染的方法。此类生物样品可以来自先前暴露于病毒、先前被诊断为对病毒呈阳性的个体或被认为处于病毒暴露的风险中的个体。本文所述的方法还可用于群体监视,即,使用由多个个体提供的样品,目的是提供关于给定群体中病毒感染的量的信息。生物样品可以是口腔或鼻粘膜。样品可以是血液、血清或血浆。在某些实施方案中,生物样品包括鼻拭子、鼻咽拭子、口腔拭子、口腔液拭子、中鼻甲拭子或其任何组合,其中所述拭子已用于从个体收集样品。

[0090] 在实践中,使用大量的收集装置收集生物样品,其全部都可以与本发明的设备一起使用。收集装置通常将是可商购获得的,但也可以是为给定应用专门设计和制造的。对于临床样品,多种商业拭子类型可供使用,包括鼻拭子、鼻咽拭子、口腔拭子、口腔液、粪便拭子、扁桃体拭子、阴道拭子、宫颈拭子和伤口拭子。样品收集装置的尺寸和材料各不相同,并且所述装置可含有专门的手柄、盖子、便于和引导破坏的刻痕以及收集基体。

[0091] 血液样品收集在多个体积不同的可商购获得的管中,其中一些含有添加剂(包括抗凝剂,诸如肝素、柠檬酸盐和EDTA)、便于样品进入的真空、便于针插入的塞子和保护操作者免于暴露于样品的覆盖物。组织和体液(例如,痰、脓性物质、吸出物)也收集在管中,通常与采血管不同。这些临床样品收集装置通常被送到水平高超的医院或商业临床实验室进行测试(但是某些测试可以在护理点进行,诸如评估咽喉/扁桃体拭子以进行快速链球菌测试)。环境样品可能呈过滤器或滤筒(例如来自空气呼吸器、气溶胶或水过滤装置)、拭子、粉末或流体而存在。

样品的处理

[0092] 收集后,将生物样品与裂解剂接触,以释放从样品获得的细胞中存在的病毒核酸。此类裂解还释放测试个体的基因组DNA。此类基因组DNA可以用作样品/扩增对照。样品收集和裂解可以在本文所述的用于病毒序列的扩增的方法之前进行,或者在病毒序列的扩增和测序的场所进行。

[0093] 根据所公开的方法和系统,样品可以与裂解试剂一起被收集或分隔,以释放分区内样品的内容物。分区包括小瓶、管和/或板中的孔。在此类实施方案中,裂解剂可以在添加用于核酸分子的延伸和扩增的试剂的同时或之前与样品悬浮液接触。在一些实施方案中,样品内核酸的处理(例如扩增、引物延伸、逆转录酶等)在用于细胞裂解的相同条件下发生。离散的分区可以包括单个样品和/或一种或多种试剂。在一些实施方案中,产生的离散分区可以包括用于核酸的扩增的引物和酶(例如逆转录酶和/或聚合酶)。在一些实施方案中,所产生的离散分区可以包含条形码化寡核苷酸(例如包含条形码序列的引物)。在一些实施方案中,所产生的离散分区可以包含携带珠子的条形码。在一些实施方案中,离散分区可能未被占用(例如,没有试剂,没有样品)。

[0094] 有利地,当裂解试剂和样品共同分区时,裂解试剂可以促进分区内样品的内容物的释放。分区中释放的内容物可能保持与其他分区的内容物离散。

[0095] 裂解剂的实例包括生物活性试剂,诸如用于裂解不同细胞类型(例如,革兰氏阳性或阴性细菌、植物、酵母、哺乳动物等)的裂解酶,诸如溶菌酶、无色肽酶、溶葡萄球菌素、labiase、kitalase、lytiembodiment和可从例如Sigma-Aldrich公司(St Louis, Mo.)购得的多种其他裂解酶以及其他可商购获得的裂解酶。其他裂解剂可以另外地或替代地与样品共同分区以引起细胞内容物向分区中的释放。例如,在一些实施方案中,基于表面活性剂的裂解溶液可用于裂解细胞,但是这些可能不太适用于基于乳液的系统,在基于乳液的系统中表面活性剂可干扰稳定的乳液。在一些实施方案中,裂解溶液可包括非离子表面活性剂,例如像TritonX-100和/或Tween 20。在一些实施方案中,裂解溶液可包括离子表面活性剂,例如像十二烷基肌氨酸钠和十二烷基硫酸钠(SDS)。本文所述的裂解剂可以包括一种或多种蛋白酶(例如,蛋白酶K)。在某些实施方案中也可以使用电穿孔、热、声学或机械细胞破裂,例如,基于非乳液的分配,诸如可以补充或代替分区分配的样品的包封,其中包封物的任何孔径足够小以在细胞破裂后保留给定大小的核酸片段。

[0096] 裂解剂还可以包括在添加样品之前或之后加热的高温水或反应缓冲液。例如,裂解剂可以是加热至至少约50、55、60、65、70、75、80、85、90或95℃持续约1、2、3、4、5、6、7、8、9、10分钟或更长的时间的加热的PCR或逆转录酶反应混合物。

[0097] 替代地或除了上文所述与样品共同分区的裂解剂之外,其他试剂也可以与样品共

同分区,包括例如DNA酶和RNA酶灭活剂或抑制剂(诸如蛋白酶K)、螯合剂(诸如EDTA)和用于消除或以其他方式降低不同细胞裂解物组分对后续核酸处理的负面活性或影响的其他试剂。此外,在包封的样品的实施方案中,可以将样品暴露于适当的刺激以从共同分区的微胶囊释放样品或其内容物。例如,在一些实施方案中,化学刺激可以连同包封的样品共同分区,以允许微胶囊的降解和细胞或其内容物向更大分区中的释放。在一些实施方案中,此刺激可以与本文其他地方描述的用于从其各自的微胶囊(例如珠子)中释放核酸分子(例如,寡核苷酸)的刺激相同。在替代方面中,这可以是不同且不重叠的刺激,以允许包封的样品在与核酸分子释放到相同分区中不同的时间释放到分区中。

[0098] 分区可以包含用于进行一个或多个反应的物质(例如试剂)。物质可以包括例如用于核酸扩增反应的试剂(例如,引物、聚合酶、dNTP、辅因子(例如,离子辅因子)、缓冲液),包括本文所述的那些;用于酶促反应的试剂(例如,酶、辅因子、底物、缓冲液);用于逆转录的试剂(例如,逆转录酶);用于核酸修饰反应(诸如聚合、连接或消化)的试剂;和/或用于模板制备的试剂。在一些情况下,可以将引物连接至前体。引物可用于逆转录。引物可包含poly-T序列或对靶核酸分子具有特异性(即,与靶核酸的序列互补)。在一些实施方案中,引物与特定靶核酸序列(例如病毒核酸序列)杂交。

[0099] 另外的试剂也可以与样品共同分区,例如用于扩增样品核酸片段并将条形码分子标签连接至扩增的片段的片段化样品DNA的内切核酸酶、DNA聚合酶和dNTP。其他酶可以共同分区,包括但不限于聚合酶、转座酶、连接酶、蛋白酶K、DNA酶等。如果裂解的样品包含基于RNA的病毒,则另外的试剂也可以包括逆转录酶。

[0100] 本文所述方法的一个优点是能够使用唾液样品检测甲型/乙型流感病毒、冠状病毒或其他病原体序列中的一种或多种。在本文所述的方法中使用的唾液样品可以包含少于约100、50、25、10、9、8、7、6、5、4、3或2微升唾液。在某些实施方案中,唾液样品小于约20微升。在某些实施方案中,唾液样品小于约10微升。在某些实施方案中,唾液样品小于约8微升。在某些实施方案中,唾液样品小于约7微升。在某些实施方案中,唾液样品小于约6微升。在某些实施方案中,唾液样品小于约5微升。

[0101] 本文所述方法的一个优点是能够使用少量鼻拭子样品检测甲型/乙型流感病毒、冠状病毒或其他病原体序列中的一种或多种。可以将鼻拭子可以接种至约1毫升缓冲液中,诸如PBS或盐水,然后可以使用少量所述样品来检测病毒感染。在本文所述的方法中使用的鼻拭子样品可以包含少于约100、50、25、10、9、8、7、6、5、4、3或2微升接种的鼻拭子。在某些实施方案中,接种的鼻拭子样品小于约20微升。在某些实施方案中,接种的鼻拭子样品小于约10微升。在某些实施方案中,接种的鼻拭子样品小于约8微升。在某些实施方案中,接种的鼻拭子样品小于约7微升。在某些实施方案中,接种的鼻拭子样品小于约6微升。在某些实施方案中,接种的鼻拭子样品小于约5微升。

病原体序列的检测

[0102] 病毒通常包含有包含脱氧核糖核酸(DNA)或核糖核酸(RNA)的基因组结构。此类基因组可以由单链或双链核酸组成。病毒的基因组通常还包含不同于它们感染的宿主的核酸序列。因此,病毒基因组内的差异核酸序列为使用分子和基因扩增和测序技术进行检测提供了靶标。此外,病毒之间遗传序列和结构的差异允许区分病毒相互之间的类别、亚类和单个毒株。因此,鉴定对病毒具有特异性的核酸序列、元件、模板或基因座的引物(例如探针)

可用于病毒感染的鉴定和检测。

[0103] 本文所述的病毒序列的检测包括使用PCR反应扩增病毒核酸以及对PCR扩增的结果进行测序。在被检测的病毒是基于RNA的病毒的情况下,扩增还包括逆转录反应。逆转录反应可以是PCT扩增之前的分立步骤,也可以是在PCR酶和引物的存在下发生的一步反应。PCR扩增通过寡核苷酸引物对进行。除了包含靶特异性部分之外,此类病毒还可以包含索引序列和/或测序衔接子序列,如图3中所示。

[0104] 病毒和宿主核酸分子的扩增可以通过使用延伸或扩增与病毒和宿主核酸分子杂交的引物的酶来实现。特别地,逆转录酶用于生成对应于宿主和病毒核酸的cDNA。例如,对于具有包含RNA的基因组的病毒,所述核酸延伸反应是逆转录反应。宿主细胞中DNA转录的RNA核酸产物也可以使用逆转录酶进行处理。已知的逆转录酶很容易获得以供使用。禽成髓细胞性白血病毒(AMV)逆转录酶和莫洛尼鼠白血病毒(M-MuLV, MMLV)逆转录酶是分子生物学工作流程中常用的RT。M-MuLV逆转录酶缺乏3'→5'核酸外切酶活性。ProtoScript II逆转录酶是一种重组M-MuLV逆转录酶,具有RNA酶H活性降低并且热稳定性增加。它可用于在比野生型M-MuLV高的温度下合成第一链cDNA。所述酶在高达50°C下具有活性,提供更高的特异性、更高的cDNA产量和更多全长cDNA产物,长度多达12kb。工程化RT的使用提高了全长产物形成的效率,确保mRNA转录本5'末端的拷贝完成,并且实现RNA序列的忠实DNA拷贝的繁殖和表征。在处理含有大量二级结构的RNA时,使用热稳定性更高的RT(反应在较高温度下进行)可能有所帮助。在一些实施方案中,所述核酸延伸反应包括逆转录反酶反应、聚合酶链反应或其组合。在某些实施方案中,所述核酸延伸反应包括(i)将引物与病毒核酸分子杂交;以及(ii)使用逆转录酶;延伸所述引物。

[0105] 在逆转录(如果适用)之后或期间,对样品进行扩增反应。在某些实施方案中,扩增反应是PCR反应。在某些实施方案中,PCR反应不是实时PCR反应。在某些实施方案中,PCR进行一定数量的足以扩增病毒核酸序列的循环。裂解和逆转录的样品可以扩增N次循环。在某些实施方案中,N大于30、35、40或45次循环。在某些实施方案中,N在30与50次循环之间、在40与50次循环之间、在35与45次循环之间、在36与44次循环之间、在37与43次循环之间、在38与42次循环之间、在39与41次循环之间或在40与45次循环之间。在某些实施方案中,裂解和逆转录的样品可以扩增40次循环。

[0106] 可以以优化的浓度将引物对添加至扩增PCR反应。在某些实施方案中,扩增病毒核酸/合成核酸和/或宿主对照的引物集的浓度是可以以低于1微摩尔的浓度将引物对添加至PCR反应。在某些实施方案中,浓度为约800nM、400nM、200nM、100nM、150nM或50nM。

[0107] 聚合酶链反应扩增可用于进一步并入功能序列,诸如可变核苷酸序列(条形码)。在一些实施方案中,所述聚合酶链反应将一个或多个额外序列掺入到所述条形码化病毒核酸分子和所述条形码化宿主核酸分子中的一者或两者中,所述额外序列选自:样品索引序列、衔接子序列、引物序列、引物结合序列、被配置成与测序仪的流式细胞联接的序列和额外条形码序列。

[0108] 本文所述的方法包括与病毒核酸共同逆转录和/或扩增的合成核酸。在某个实施方案中,扩增病毒序列的寡核苷酸引物集还扩增合成核酸。可以将样品“加标”合成核酸分子,所述核酸分子提供关于样品中核酸的处理的信息。在某些实施方案中,在处理或扩增之前将合成核酸分子添加至生物样品中。在某些实施方案中,以已知浓度或量将合成核酸加

标至生物样品中。在某些实施方案中,抑制量为 1×10^4 、 1×10^3 、 1×10^2 、10、5、4、3、2或1个合成核酸拷贝。在某些实施方案中,合成对照核酸是RNA。理想的是,合成核酸的扩增部分的长度与待扩增的病毒核酸靶标的长度相匹配,并且产生长度与病毒核酸靶标相同或在约1、2、3、4、5、6、7、8、9、10、15、20、30、40或50个核苷酸内的扩增子。另外,合成核酸的序列应与病毒序列高度同源,但有至少一个核苷酸不同,使得可以通过测序来区分合成核酸。在某些实施方案中,与待扩增的病毒核酸的序列相比,合成核酸有至少约1、2、3、4、5、6、7、8、9、10、11、12、13、14或15个核苷酸不同。

[0109] 合成核酸序列可用于归一化序列读段并解释在样品制备期间的损失或在扩增或测序期间引入的低效率/偏差。

[0110] 合成核酸可以包含多个具有不同序列的核酸,以进一步提高“加标”合成核酸用于下游处理和分析的性能。所述多个可以包含2、3、4、5、6、7、8、9、10或更多不同序列。所述多个可以包含2、3、4、5个不同序列。所述多个可以包含4个不同序列。所述多个可以用于与扩增S2标序列、N1核蛋白序列或其组合的引物一起使用。示例性序列示出于下表中(因为合成核酸是RNA,所以在RNA版本的序列中,所有T都是U)。合成核酸中的任一种或多种或者多个合成核酸可以与下文所示的SEQ ID NO.1至12中的任一个具有至少约90%、95%、97%、98%、99%或相同的同源性。

| COVID 19 S2 和 N1 合成标序列的示例性序列 | |
|---------------------------------------|---|
| S2_001 SEQ ID NO: 1 | GCUGGUGCUGCAGCUUAUUAUGGGGUGUGUAUCU CACGAAGCGACCCUUUGGAAAAUAUAAUGAAAUG GAACCAUUACAGAUGCUGUAGACUGUGCACUUGA CCU |
| S2_002 SEQ ID NO: 2 | GCUGGUGCUGCAGCUUAUUAUGUGGGUCCUCGCU AGGACGUGCUAUgacgccAAAAUAAUGAAAUGGAA CCAUUACAGAUGCUGUAGACUGUGCACUUGACCC U |
| S2_003 SEQ ID NO: 3 | GCUGGUGCUGCAGCUUAUUAUGGGUAGCACGACU UGAUCUAACUgacacuaAAAAUAUAAUGAAAUGGAA CCAUUACAGAUGCUGUAGACUGUGCACUUGACCU |
| S2_004 SEQ ID NO: 4 | GCUGGUGCUGCAGCUUAUUAUGUGGGUUAAGUAG GACUUCGAUUggaUggaauAAAAUAUAAUGAAAUG GAACCAUUACAGAUGCUGUAGACUGUGCACUUGA CCCU |
| N1_001 SEQ ID NO: 5 | UCUGGUUACUGCCAGUUGAAUCUGAGGGUCCGGA CGGAUAUCGCACUAAGUGUACCUGGUGCAUUUCG CUGAUUUUGGGGUC |
| N1_002 SEQ ID NO: 6 | UCUGGUUACUGCCAGUUGAAUCUGAGGGUCCAU GACCAUGUCACUGGCUACACUGAGGUGCAUUUCG CUGAUUUUGGGGUC |
| N1_003 SEQ ID NO: 7 | UCUGGUUACUGCCAGUUGAAUCUGAGGGUCCUUG ACAUGGCAUGUGACUCCACUGUCGGUGCAUUUCG CUGAUUUUGGGGUC |
| N1_004 SEQ ID NO: 8 | UCUGGUUACUGCCAGUUGAAUCUGAGGGUCCACC UUUGCCAGAUGACUGAGUGGAAGGGUGCAUUUCG CUGAUUUUGGGGUC |
| N1_005 SEQ ID NO: 9 | UCUGGUUACUGCCAGUUGAAUCUGAGGGUCCAGC UUGAAGCGUUCGCGACAAGUGUCGGUGCAUUUCG CUGAUUUUGGGGUC |
| N1_006 SEQ ID NO: 10 | UCUGGUUACUGCCAGUUGAAUCUGAGGGUCCUCA CGUCCUGAGAUAACUGCUACAUGGUGCAUUUCG CUGAUUUUGGGGUC |
| N1_007 SEQ ID NO: 11 | UCUGGUUACUGCCAGUUGAAUCUGAGGGUCCGUU GACUGAUCACAUGCUGCUCCACGGGUGCAUUUCG UGAUUUUGGGGUC |
| N1_008 SEQ ID NO: 12 | UCUGGUUACUGCCAGUUGAAUCUGAGGGUCCAG ACAGUCAUCGGAUUGAUGAGUGAGGUGCAUUUCG CUGAUUUUGGGGUC |

[0111] 合成核酸包含5'和3'近端序列,所述序列可以被扩增测试的病毒核酸序列的相同引物结合,但是包含可以通过测序与测试的病毒核酸序列区分开的不同间插序列。在一些实施方案中,本文所述的方法还包括提供合成核酸分子。在一些实施方案中,本文所述的方法还包括提供合成核酸分子,其包含可被扩增病毒核酸的相同引物结合的5'和3'近端序列。在某些实施方案中,所述5'近侧区、所述3'近侧区或所述5'近侧区和所述3'近侧区的长度少于约30个核苷酸。在某些实施方案中,所述5'近侧区、所述3'近侧区或所述5'近侧区和

所述3'近侧区的长度少于约25个核苷酸。在某些实施方案中,所述5'近侧区、所述3'近侧区或所述5'近侧区和所述3'近侧区的长度少于约20个核苷酸。在某些实施方案中,所述5'近侧区在所述合成核酸的5'末端。在某些实施方案中,所述3'近侧区在所述合成核酸的3'末端。在某些实施方案中,所述间插核酸序列与病毒核酸序列具有少于约99%、98%、97%、95%、90%、85%、80%或75%同一性。在一些实施方案中,合成核酸分子包含与所述病毒核酸分子和所述人核酸分子不同的合成序列。在一些实施方案中,合成核酸分子包含与宿主和/或病毒序列的同一性不大于10%、25%、50%、75%、90%、95%或98%的核苷酸序列。

[0112] 在某些实施方案中,将合成核酸添加至裂解物中以在所述方法中实现约1个拷贝/孔至约1,000,000拷贝/孔的浓度。在某些实施方案中,以约1个拷贝/孔至约50个拷贝/孔、约1个拷贝/孔至约100个拷贝/孔、约1个拷贝/孔至约500个拷贝/孔、约1个拷贝/孔至约1,000个拷贝/孔、约1个拷贝/孔至约5,000个拷贝/孔、约1个拷贝/孔至约10,000个拷贝/孔、约1个拷贝/孔至约100,000个拷贝/孔、约1个拷贝/孔至约1,000,000个拷贝/孔、约50个拷贝/孔至约100个拷贝/孔、约50个拷贝/孔至约500个拷贝/孔、约50个拷贝/孔至约1,000个拷贝/孔、约50个拷贝/孔至约5,000个拷贝/孔、约50个拷贝/孔至约10,000个拷贝/孔、约50个拷贝/孔至约100,000个拷贝/孔、约50个拷贝/孔至约1,000,000个拷贝/孔、约100个拷贝/孔至约500个拷贝/孔、约100个拷贝/孔至约1,000个拷贝/孔、约100个拷贝/孔至约5,000个拷贝/孔、约100个拷贝/孔至约10,000个拷贝/孔、约100个拷贝/孔至约100,000个拷贝/孔、约100个拷贝/孔至约1,000,000个拷贝/孔、约500个拷贝/孔至约1,000个拷贝/孔、约500个拷贝/孔至约5,000个拷贝/孔、约500个拷贝/孔至约10,000个拷贝/孔、约500个拷贝/孔至约100,000个拷贝/孔、约500个拷贝/孔至约1,000,000个拷贝/孔、约1,000个拷贝/孔至约5,000个拷贝/孔、约1,000个拷贝/孔至约10,000个拷贝/孔、约1,000个拷贝/孔至约100,000个拷贝/孔、约1,000个拷贝/孔至约1,000,000个拷贝/孔、约5,000个拷贝/孔至约10,000个拷贝/孔、约5,000个拷贝/孔至约100,000个拷贝/孔、约5,000个拷贝/孔至约1,000,000个拷贝/孔、约10,000个拷贝/孔至约100,000个拷贝/孔、约10,000个拷贝/孔至约1,000,000个拷贝/孔或约100,000个拷贝/孔至约1,000,000个拷贝/孔添加合成核酸。在某些实施方案中,以约1个拷贝/孔、约50个拷贝/孔、约100个拷贝/孔、约500个拷贝/孔、约1,000个拷贝/孔、约5,000个拷贝/孔、约10,000个拷贝/孔、约100,000个拷贝/孔或约1,000,000个拷贝/孔添加合成核酸。在某些实施方案中,以至少约1个拷贝/孔、约50个拷贝/孔、约100个拷贝/孔、约500个拷贝/孔、约1,000个拷贝/孔、约5,000个拷贝/孔、约10,000个拷贝/孔或约100,000个拷贝/孔添加合成核酸。在某些实施方案中,以至多约50个拷贝/孔、约100个拷贝/孔、约500个拷贝/孔、约1,000个拷贝/孔、约5,000个拷贝/孔、约10,000个拷贝/孔、约100,000个拷贝/孔或约1,000,000个拷贝/孔添加合成核酸。

[0113] 在扩增和/或测序后,所得的检测到的病原体核酸分子和合成寡核苷酸的比率可用于检测和诊断病原体感染。用于检测和诊断传染性病原体的比率的组分可以基于特定序列或序列组的读段数目、特定序列或序列组的拷贝数(例如,与特定序列缔合的不同唯一分子标识符序列的数目)、得自测序信息的数目或量或其任何组合。在一些实施方案中,病原体核酸与合成核酸的比率大于1。在一些实施方案中,病原体核酸与合成核酸的比率等于或是约1。在一些实施方案中,病原体核酸与合成核酸的比率小于1。

[0114] 在某些实施方案中,病原体核酸与合成核酸的比率是约0.00001:1至约0.5:1。在

某些实施方案中,病原体核酸与合成核酸的比率是约0.00001:1至约0.00005:1、约0.00001:1至约0.0001:1、约0.00001:1至约0.0002:1、约0.00001:1至约0.0003:1、约0.00001:1至约0.0004:1、约0.00001:1至约0.0005:1、约0.00001:1至约0.001:1、约0.00001:1至约0.005:1、约0.00001:1至约0.01:1、约0.00001:1至约0.1:1、约0.00001:1至约0.5:1、约0.00005:1至约0.0001:1、约0.00005:1至约0.0002:1、约0.00005:1至约0.0003:1、约0.00005:1至约0.0004:1、约0.00005:1至约0.0005:1、约0.00005:1至约0.001:1、约0.00005:1至约0.005:1、约0.00005:1至约0.01:1、约0.00005:1至约0.1:1、约0.00005:1至约0.5:1、约0.0001:1至约0.0002:1、约0.0001:1至约0.0003:1、约0.0001:1至约0.0004:1、约0.0001:1至约0.0005:1、约0.0001:1至约0.001:1、约0.0001:1至约0.005:1、约0.0001:1至约0.01:1、约0.0001:1至约0.1:1、约0.0001:1至约0.5:1、约0.0002:1至约0.0003:1、约0.0002:1至约0.0004:1、约0.0002:1至约0.0005:1、约0.0002:1至约0.001:1、约0.0002:1至约0.005:1、约0.0002:1至约0.01:1、约0.0002:1至约0.1:1、约0.0002:1至约0.5:1、约0.0003:1至约0.0004:1、约0.0003:1至约0.0005:1、约0.0003:1至约0.001:1、约0.0003:1至约0.005:1、约0.0003:1至约0.01:1、约0.0003:1至约0.1:1、约0.0003:1至约0.5:1、约0.0004:1至约0.0005:1、约0.0004:1至约0.001:1、约0.0004:1至约0.005:1、约0.0004:1至约0.01:1、约0.0004:1至约0.1:1、约0.0004:1至约0.5:1、约0.0005:1至约0.001:1、约0.0005:1至约0.005:1、约0.0005:1至约0.01:1、约0.0005:1至约0.1:1、约0.0005:1至约0.5:1、约0.001:1至约0.005:1、约0.001:1至约0.01:1、约0.001:1至约0.1:1、约0.001:1至约0.5:1、约0.005:1至约0.01:1、约0.005:1至约0.1:1、约0.005:1至约0.5:1、约0.01:1至约0.1:1、约0.01:1至约0.5:1或约0.1:1至约0.5:1。在某些实施方案中,病原体核酸与合成核酸的比率是约0.00001:1、约0.00005:1、约0.0001:1、约0.0002:1、约0.0003:1、约0.0004:1、约0.0005:1、约0.001:1、约0.005:1、约0.01:1、约0.1:1或约0.5:1。在某些实施方案中,病原体核酸与合成核酸的比率是至少约0.00001:1、约0.00005:1、约0.0001:1、约0.0002:1、约0.0003:1、约0.0004:1、约0.0005:1、约0.001:1、约0.005:1、约0.01:1或约0.1:1。

[0115] 在某些实施方案中,病原体核酸与合成核酸的比率是约1:1至约1:100。在某些实施方案中,病原体核酸与合成核酸的比率是约1:100至约1:50、约1:100至约1:25、约1:100至约1:10、约1:100至约1:5、约1:100至约1:4、约1:100至约1:3、约1:100至约1:2、约1:100至约1:1、约1:50至约1:25、约1:50至约1:10、约1:50至约1:5、约1:50至约1:4、约1:50至约1:3、约1:50至约1:2、约1:50至约1:1、约1:25至约1:10、约1:25至约1:5、约1:25至约1:4、约1:25至约1:3、约1:25至约1:2、约1:25至约1:1、约1:25至约1:1、约1:10至约1:5、约1:10至约1:4、约1:10至约1:3、约1:10至约1:2、约1:10至约1:1、约1:5至约1:4、约1:5至约1:3、约1:5至约1:2、约1:5至约1:1、约1:5至约1:1、约1:4至约1:3、约1:4至约1:2、约1:4至约1:1、约1:4至约1:1、约1:3至约1:2、约1:3至约1:1、约1:3至约1:1、约1:2至约1:1、约1:2至约1:1或约1:1至约1:1。在某些实施方案中,病原体核酸与合成核酸的比率是约1:100、约1:50、约1:25、约1:10、约1:5、约1:4、约1:3、约1:2、约1:1或约1:1。在某些实施方案中,病原体核酸与合成核酸的比率是至少约1:100、约1:50、约1:25、约1:10、约1:5、约1:4、约1:3、约1:2或约1:1。在某些实施方案中,病原体核酸与合成核酸的比率是至多约1:50、约1:25、约1:10、约1:5、约1:4、约1:3、约1:2、约1:1或约1:

1。

[0116] 在某些实施方案中,合成核酸与病原体核酸的比率是约1:100至约1:50、约1:100至约1:25、约1:100至约1:10、约1:100至约1:5、约1:100至约1:4、约1:100至约1:3、约1:100至约1:2、约1:100至约1:1、约1:100至约1:1、约1:50至约1:25、约1:50至约1:10、约1:50至约1:5、约1:50至约1:4、约1:50至约1:3、约1:50至约1:2、约1:50至约1:1、约1:50至约1:1、约1:25至约1:10、约1:25至约1:5、约1:25至约1:4、约1:25至约1:3、约1:25至约1:2、约1:25至约1:1、约1:25至约1:1、约1:10至约1:5、约1:10至约1:4、约1:10至约1:3、约1:10至约1:2、约1:10至约1:1、约1:10至约1:1、约1:5至约1:4、约1:5至约1:3、约1:5至约1:2、约1:5至约1:1、约1:5至约1:1、约1:4至约1:3、约1:4至约1:2、约1:4至约1:1、约1:4至约1:1、约1:3至约1:2、约1:3至约1:1、约1:3至约1:1、约1:2至约1:1、约1:2至约1:1或约1:1至约1:1。在某些实施方案中,病原体核酸与合成核酸的比率是约1:100、约1:50、约1:25、约1:10、约1:5、约1:4、约1:3、约1:2、约1:1或约1:1。在某些实施方案中,病原体核酸与合成核酸的比率是至少约1:100、约1:50、约1:25、约1:10、约1:5、约1:4、约1:3、约1:2或约1:1。在某些实施方案中,病原体核酸与合成核酸的比率是至多约1:50、约1:25、约1:10、约1:5、约1:4、约1:3、约1:2、约1:1或约1:1。

[0117] 病原体读段与病原体合成核酸(序列可与病原体读段区分)的比率可用于指示特定病原体(例如冠状病毒、甲型流感病毒、乙型流感病毒)的阳性诊断。替代地,如果比率不超过阳性阈值,则作出阴性诊断。在一些实施方案中,如果从病原体核酸到合成核酸的序列读段超过约0.01至约0.5的比率,则作出针对病原体的阳性诊断。在一些实施方案中,如果从病原体核酸至合成核酸的序列读段超过约0.01至约0.4、约0.01至约0.3、约0.01至约0.2、约0.02至约0.5、约0.02至约0.2、约0.03至约0.5、约0.03至约0.2、约0.04至约0.5、约0.03至约0.2、约0.05至约0.5、约0.05至约0.2、约0.06至约0.5、约0.06至约0.2、约0.07至约0.5、约0.07至约0.2、约0.08至约0.5、约0.08至约0.2的比率,则作出针对病原体的阳性诊断。在一些实施方案中,如果从病原体核酸至合成核酸的序列读段超过约0.01、约0.02、约0.03、约0.04、约0.05、约0.06、约0.07、约0.08、约0.09或约0.1的比率,则作出针对病原体的阳性诊断。

[0118] 病原体读段加病原体合成核酸(序列可与病原体读段区分)的总量可用于指示是否存在足够的序列数据来尝试针对病原体(例如冠状病毒、甲型流感病毒、乙型流感病毒)的存在作出阳性或阴性诊断。在一些实施方案中,如果病原体核酸与合成核酸的序列读段的组合数目超过最小阈值,则仅作出阳性或阴性诊断,否则结果是不确定的。在一些实施方案中,最小阈值是至少约10个读段、至少约20个读段、至少约30个读段、至少约40个读段、至少约50个读段、至少约60个读段、至少约70个读段、至少约80个读段、至少约90个读段、至少约100个读段、至少约150个读段、至少约200个读段或至少约250个读段。

[0119] 在某些情况下,样品对照可用于验证样品中是否存在足够的遗传物质,以可靠地检出阳性或阴性诊断。可以对样品对照的读段的量进行计数。样品对照通常是进行测试的个体物种的持家基因。在某些实施方案中,当测试的个体是人时,样品对照是 β 肌动蛋白、GAPDH或RPP30。在某些实施方案中,当测试的个体是人时,样品对照是RPP30。在某些实施方案中,来自存在以传递阳性或阴性诊断的样品对照的读段的量超过5、6、7、8、9、10、11、12、13、14、15、16、17、128、19、20、25或30个读段。在某些实施方案中,来自存在以传递阳性或阴

性诊断的样品对照的读段的量等于或超过10个读段。

[0120] 所公开的方法还可以与复用策略组合,以有效且高效地实现并行测试多个样品。如本文所用,“复用”是指同时对一个或多个样品进行多个测定。复用还可以包括在多个单独样品中的每一个中同时进行多个测定。例如,分析的反应混合物的数目可以基于分区的数目,并且在每个分区中进行的测定的数目可以基于接触每个孔的内容物的探针的数目。在所公开的方法的应用中可以有效地采用复用设计和策略。例如,可以将一个或多个唯一条形码或衔接核酸分子与靶核酸分子联接,其中唯一条形码或衔接核酸分子鉴定或条形码化靶核酸分子。一旦进行了条形码化,可以将样品组合或汇集至单个测序文库中,其中可以将条形码化靶核酸(例如病原体核酸分子和/或合成核酸分子)与其他条形码化样品区分开。因此,样品复用和核酸条形码的使用可用于鉴定一组或多个个体的第一个体的测序读段。在一些实施方案中,样品在测序之前汇集。在某些实施方案中,大于10、25、50、75、100或500个样品在单个测序文库中汇集和分析。

[0121] 此外,可以在单一测定中测试一种或多种病原体。例如,可以在一次测定中分析多种呼吸道病毒(或病毒亚型)。因此,针对单一病原体所公开的方法可以组合用于在单一反应中检测和分析多种病原体。(例如使用对多种病原体核酸分子具有特异性的多种引物,并且可以将多种合成核酸分子添加至样品中)。以另一个实例的方式,在一些实施方案中,第一种病原体是冠状病毒,并且第二种病原体是流感病毒。多种病原体的实例还包括不同的病毒亚型或进化枝(例如甲型流感病毒H1N1和甲型流感病毒H3N2)。

[0122] 本文所述的方法和反应混合物可以利用多个冠状病毒序列的扩增和测序。在某些实施方案中,被扩增和测序的多个冠状病毒序列包含S2序列。在某些实施方案中,被扩增和测序的多个冠状病毒序列包含N1序列。在某些实施方案中,被扩增和测序的多个冠状病毒序列包含S2和N1序列。对于每个病毒靶标,此类测试可以利用多个合成核酸加标对照。

[0123] 本文所述的方法、反应混合物和试剂盒可用于测试冠状病毒(例如,SARS-CoV2)和甲型流感病毒和/或乙型流感病毒。此类测试将可用于医疗保健提供者和卫生机构同时监测不同常见呼吸道病原体的爆发。本文所述的方法可用于一式三份的测试以同时检测冠状病毒和甲型流感病毒和乙型流感病毒。此类测试可以利用三种不同的合成核酸加标对照,每种病毒靶标一种。

[0124] 本文所述的方法还可以包括第二寡核苷酸引物集,其靶向病毒宿主(例如,所测试的个体)的核酸序列,这种引物集为样品和扩增提供了阳性对照。在某些实施方案中,所述第二PCR引物集扩增选自以下的人核酸序列:GAPDH、ACTB、RPP30及其组合。在某些实施方案中,所述第二PCR引物集扩增人RPP30序列。在某些实施方案中,所述第二PCR引物集包含具有测序衔接子序列的引物和不具有测序衔接子序列的引物的混合物。在某些实施方案中,具有测序衔接子序列的引物与不具有测序衔接子序列的引物的比率是约1:1、约1:2、约1:3或约1:4。具有或不具有衔接子的引物的混合物允许检测病毒宿主核酸,但允许更多的测序读段用于病毒序列。

[0125] 在一些实施方案中,如果未检测到最小预定数目的人核酸序列的读段,则病原体感染的诊断结果是不确定的。在一些实施方案中,人核酸序列的读段的最小预定数目必须超过至少约2、至少约3、至少约4、至少约5、至少约6、至少约7、至少约8、至少约9、至少约10、至少约15、至少约20或至少约25个人核酸序列的读段。在一些实施方案中,人核酸序列的读

段的最小预定数目必须超过至少约10个读段。

[0126] 冠状病毒形成直径为100-160nm的有包膜的病毒颗粒。冠状病毒是单链核糖核酸(ssRNA)病毒,其包含大小为27-32kb的正义RNA基因组。冠状病毒基因组的5'区编码多聚蛋白pp1ab,其被进一步裂解成16个参与基因组转录和复制的非结构蛋白。3'区编码结构蛋白,其包括包膜糖蛋白刺突(例如病毒刺突)、包膜、膜和核衣壳。编码结构蛋白的基因也可以用作物种特异性的辅助基因,并且对于病毒复制来说可为可有可无的。

[0127] COVID-19(还称为HCoV-19或SARS-CoV-2)是β冠状病毒并且是已知的第七种感染人类的冠状病毒。COVID-19冠状病毒可以造成严重疾病(在SARS-CoV、MERS-CoV的情况下也观察到)。COVID-19证明了对人受体ACE2的有效靶向。刺突蛋白中的受体结合结构域(RBD)是冠状病毒基因组中变化最大的部分。六个受体结合结构域(RBD)氨基酸已显示对于与ACE2受体结合以及确定SARS-CoV样病毒的宿主范围来说至关重要。基于SARS-CoV蛋白序列,参与ACE2结合并且有助于受体靶向的残基是Y442、L472、N479、D480、T487和Y4911(这些残基还对应于SARS-CoV-2中的L455、F486、Q493、S494、N501和Y505)。在COVID-19中,这六个残基中有五个在COVID-19与SARS-CoV之间存在差异,其对应于COVID-19受体结合结构域(RBD)蛋白的L455、F486、Q493、S494和N501。COVID-19的此类序列特性可用于鉴定、检测和/或诊断COVID-19。例如,COVID-19受体结合结构域(RBD)的序列的唯一序列可用于鉴定、检测或诊断COVID-19。

[0128] COVID-19还通过插入12个核苷酸在病毒刺突蛋白边界的S1亚基和S2亚基内包含功能性多碱基(弗林蛋白酶)裂解位点。插入的序列还加强了三个O-连接聚糖的获取。多碱基裂解位点和三个相邻的预测O-连接聚糖是COVID-19唯一的,并且先前未在α或β冠状病毒中鉴定到。因此,COVID-19的病毒刺突序列特性可用于鉴定、检测和/或诊断COVID-19。例如,COVID-19病毒刺突插入的序列的唯一序列可用于鉴定、检测或诊断COVID-19。

[0129] 本文所述的方法可用于检测和诊断个体的病毒感染。在某些实施方案中,病毒感染是流感病毒感染。在某些实施方案中,病毒感染是冠状病毒感染。在某些实施方案中,病毒感染是COVID-19感染。在某些实施方案中,病毒感染是MERS-CoV感染。在某些实施方案中,病毒感染是SARS-CoV-2感染。

[0130] 原则上,任何合适的病毒核酸序列都可以被寡核苷酸引物对靶向。此类靶标包括编码病毒核衣壳蛋白、病毒刺突蛋白、病毒包膜蛋白或病毒膜蛋白的核酸序列。此类靶标包括非结构蛋白。

[0131] 如果待检测的病毒感染是SARS-CoV-2感染,则检测的病毒核酸可以是编码病毒核衣壳蛋白(N1)、病毒刺突蛋白(S2)、病毒包膜蛋白或病毒膜蛋白中的一种或多种的核酸,在某些实施方案中,检测的核酸是编码16种SARS-CoV-2非结构蛋白(NSP1、NSP2、NSP3、NSP4、NSP5、NSP6、NSP7、NSP8、NSP9、NSP10、NSP11、NSP12、NSP13、NSP14、NSP15或NSP16)中任一种或多种的核酸。

[0132] 可用于本文所述方法的示例性引物对可结合本公开之后的序列附录中所述的序列。在某些实施方案中,附录中列出的任何引物可用于本文所述的方法中。

[0133] 包含条形码化寡核苷酸的引物可用于样品分析(例如样品鉴定、追踪、定量等)。例如,包含条形码化寡核苷酸的分区内的引物还包含共同条形码序列和唯一分子标识符序列,可用于(1)通过使用共同条形码鉴定属于分区的序列和(2)通过使用唯一分子标识符鉴

定转录本拷贝数目。在一些实施方案中,所述第一引物集还包含一个或多个选自以下的额外功能序列:引物序列、衔接子序列、引物退火序列、唯一分子标识符序列和捕获序列。在一些实施方案中,所述第二引物还包含一个或多个选自以下的额外功能序列:引物序列、衔接子序列、引物退火序列、唯一分子标识符序列和捕获序列。一般来说,引物对病毒、宿主或合成核酸分子的靶序列具有特异性(即与靶序列互补或包含与靶序列互补的序列)。

[0134] 图4-图21进一步展示和/或举例说明本文公开的方法和组合物以及它们的使用和/或应用的优点。图4举例说明针对COVID19的SwabSeq诊断测试平台。在(A)中,Swab Seq的工作流程是一个五步过程,从开始到结束需要大约12个小时。在(B)每个孔中,对临床样品进行RT-PCR。每个孔具有两个索引引物集,从而生成SARS-CoV-2S2基因和人RPP30基因的cDNA和扩增子。每个引物以用于Illumina测序的P5和P7衔接子、唯一i7和i5分子条形码以及唯一引物对合成。重要的是,每个孔具有合成体外S2标准物,这是允许所述方法大规模工作的关键。在(C)中,体外S2标准物(缩写为S2-标)与病毒S2基因有6个互补的碱基对(以下划线表示)不同。在(D)中,读取在各种病毒浓度下的计数。在(E)中,比率计量归一化允许对每个扩增子进行孔内归一化。在(F)中,每个孔具有两个用于扩增的内部孔对照,体外S2标准物和人RPP30。RPP30扩增子用作样本收集的对照。体外S2标准物对于SwabSeq区分真阴性的能力来说是至关重要的。

[0135] 图5显示了在临床样本中的验证说明了等同于灵敏的RT-QPCR反应的检测限。在(A)中,汇集了不具有SARS-CoV2的鼻拭子样品的检测限(LOD),并以不同浓度添加ATCC灭活病毒。将鼻拭子样品RNA纯化,并且使用SwabSeq,显示每毫升250个基因组拷贝当量(GCE)的检测限。在(B)中,通过UCLA健康临床微生物学实验室获得的RNA纯化的临床鼻拭子样本使用FDA授权平台基于临床协议进行测试,然后还使用SwabSeq进行测试。显示与测试为对SARS-CoV-2呈阳性($n=31$)并且对SARS-CoV-2呈阴性($n=35$)的样品100%符合率。在测试的(C)中,对来自无提取鼻咽拭子的RNA纯化的样品进行测试,并且显示检测限为558GCE/mL,并且在(D)临床样品中,显示在UCLA健康临床微生物学实验室进行的测试之间的100%符合率,阴性($n=20$)和阳性($n=20$)。在(E)中,唾液样本的无提取处理显示检测限降至每毫升1000GCE。

[0136] 图6显示了测序文库设计。显示了S2(上图)和RPP30(下图)扩增子的扩增子设计。扩增子被设计成使得i5和i7分子索引唯一鉴定每个样品。SwabSeq被设计成与所有Illumina平台兼容。图7显示了当扩增COVID-19扩增子和合成S2标时,S2引物显示等同的PCR效率。用S2_标或SARS-CoV-2病毒(以绿色标记为C19gRNA)输入的引物的PCR效率斜率如下:S2_标斜率 $=-6.68e-6$,并且C19gRNA(Twist对照)斜率 $=-6.74e-6$ 。如果引物未显示S2标RNA与C19gRNA的优先扩增,则预计斜率相等(平行)。这表明,使用S2引物对,S2标和C19gRNA具有相等的扩增效率。条带代表预测值的95%置信区间,由于截距不同而不重叠,并且与此斜率分析无关。

[0137] 图8显示了在非常高的病毒浓度下,SwabSeq维持线性。包括了内部孔控制,S2标,以即使在存在异质样品类型和PCR抑制的情况下,也能够检出阴性样品。在(A)中,随着病毒浓度增加,观察到归因于S2的读段增加,并且在(B)中,观察到归因于S2标的读段降低。在(C)中,S2与S2标之间的比率提供了额外水平的比率计量归一化,并且表现出高达至少200万个拷贝/mL裂解物的线性。需注意,两个轴上的刻度以 \log_{10} 标度间隔。

[0138] 图9显示了在MiSeq或NextSeq机器上进行的测序表现出相似的灵敏度。在MiSeq和NextSeq上运行的复用文库在纯化的RNA背景中在广泛的SARS-CoV2病毒拷贝中显示线性。图10显示了使用NextSeq550的RNA纯化样品的初步和确证性检测线数据。在(A)中,初步LOD数据确定了250个拷贝/mL的LOD,并且在(B)中,确证性研究显示了250个拷贝/mL的LOD。在(C)中,显示了纯化的RNA的示例性结果解释指南。

[0139] 图11显示了向传统收集介质和缓冲液中的无提取方案需要稀释以克服RT和PCR抑制的影响。在(A)中测试了用于放置在病毒运送培养基(VTM)中的鼻咽(NP)拭子的无提取方案。以不同浓度将ATCC活灭活病毒加标至汇集的VTM中,然后用水1:4稀释样品,之后添加至RT-PCR反应。观察到每毫升5714个拷贝的检测限。在(B)中,测试的鼻咽(NP)拭子被收集在生理盐水(NS)中,汇集然后以不同浓度加标ATCC活灭活病毒。人为的样品在水中1:4稀释。在这里,早期研究显示每毫升2857与5714个拷贝之间的相似检测限。在(C)中,测试的自然临床样品被收集至Amies缓冲液(ESwab)中。将阳性样品的S基因Ct计数(x轴)与SwabSeq S2与S2标的比率(y轴)进行比较。样品一式三份(颜色)运行。观察到Ct计数为27的高度一致性以及Ct计数大于27的较低但更大的变异性,表明RT和PCR抑制正影响检测限。

[0140] 图12显示了实现每天对数千个样品进行可扩展测试的轻量级样品登记、收集和处理的开发。在(A)中,为了解决样品收集的挑战,开发了轻量级收集方法,将样品直接收集到可自动化的管中。在这里,对于个人使用漏斗以存放少量唾液样品(将0.25mL存放至漏斗和管中)。此设置可以容纳多种样品类型。在(B)中,为了有利于样品登记和收集,开发了基于网络的应用程序,供个人使用条形码阅读器注册他们的样品管,并将他们的识别信息发送到Qualtrics的安全实例中。然后个人收集他们的样品,然后将管放在架子上。这种低接触的预分析过程允许每天处理数千个样品,而没有繁重的管理负担。在(C)中,总体工作流程使实验室中的处理简单化。首先,个人将样品收集到可自动化的管中,然后将它们放入96管架中。样品以96架形式到达实验室,允许对样品进行高效灭活和处理,大大增加样品通过平台的流量。

[0141] 图13显示将唾液预加热至95°C持续30分钟大大改进了RT-PCR。检测病毒基因组并且显示对对照的检测的稳健性提高。在(A)中,在没有预加热的情况下,S2标的检测是最小的,并且对照扩增子的计数较低。在(B)中,在95°C预加热30分钟的步骤的情况下,观察到S2扩增子和合成S2标的稳健检测。图14显示,PCR抑制对扩增产物有显著影响,对来自Rt-PCR反应的一小部分孔进行了2%琼脂糖基因测试。观察到与纯化的RNA(A9-A12)相比,未纯化的裂解物(A1-A8)中拭子的RT-PCR抑制。在表示S2或S2标(177bp)和RPP30(133bp)引物对的2个扩增子的这一小部分孔中观察到两个条带。图15显示了看到增加PCR循环数目并且用未纯化或抑制性样品类型(例如,唾液)工作的tapestation使文库制备中非特异性峰的大小增加。来自Agilent TapeStation的针对纯化的扩增子文库的代表性结果。观察到,在两个文库迹线中有略高于100bp的非特异性峰(箭头),但在未纯化的样品和数目增加的PCR循环的情况下,此峰的大小增加。重要的是,观察到,此非特异性峰大小的增加导致文库定量不准确。因此,为了优化Illumina测序仪上的簇密度,建议基于所需峰(RPP30和S2)的比例定量最终文库的加载浓度。

[0142] 图16显示了相对于NEB Luna,TaqPath减少了SARS-CoV2阴性样品中S2读段的数目。在(A)和(B)中,比较了Luna One Step RT-PCR Mix(New England Biosciences)和

aqPath™1-Step RT-qPCR Master Mix (ThermoFisher Scientific)。TaqPath Mastermix中UNG的存在可能显著减少SARS-CoV-2阴性样品中S2读段的数目,实现更准确的SARS-CoV-2阳性和SARS-CoV-2阴性样品的诊断。

[0143] 图17显示了来自MiSeq的模板线的滞留污染导致交叉污染。在此实验中,对四个384孔板进行RT-PCR,但仅汇集了三个板。左侧是未包括在运行中的384孔板的每个样品的每个扩增子的观察计数(但在前一次运行中使用了索引)。前一次运行中使用的索引的扩增子读段以低水平(0-150个读段)存在。在后续运行之前,除了常规洗涤外还进行了漂洗。在此后续运行中,将三个不同的板汇集并且省略了第四个384孔板。右侧是对应于省略板的样品索引的每个扩增子的观察计数(同样,在上一次运行中使用了索引)。观察到滞留污染的量显著减少,每个样品的滞留读段 <10 。替代地,在一些实施方案中,为了减少由此产生污染和结果混淆,在同一仪器上对后续运行使用反向和正向引物的不同组合。旋转序列上的引物允许确定读段是来自样品的真实读段还是来自前一次运行的交叉污染。图18显示了扩增子读段中的测序错误和潜在的扩增子错误分配。在实验v18中,加载比平时(11%)少的PhiX,并且读段的总体质量较低。这里注意到的趋势在其他运行中仍存在,但这次运行更清楚地突出了由于测序错误和过度容忍的纠错而可能发生的问题。在(A)中,读段1中每个位置碱基质量分数小于12的读段的百分比。需注意,读段1的前6个碱基将S2与S2标区分开,并且具有最高百分比的低质量碱基检出(base call)。在(B)中,每个读段1序列与预期S2序列(行)或S2标序列(列)之间的汉明距离,黄色是完美匹配和编辑距离1序列,其可以被清楚地鉴定为S2或S2标。红色是可能被错误分配的具有错误的序列(分配为S2的S2标对于此测定来说是最有问题的。)

[0144] 图19显示了不同索引策略的可视化。这里i5索引被描绘为水平线,i7索引被描绘为垂直线,并且颜色表示唯一索引。在组合(或完全组合)索引中,将i5和i7索引组合以得到唯一组合,但每个i5和i7索引可以在一个板中多次使用,并且是所有可能的i5和i7。对于唯一双重索引,每个i5和i7索引在每个板中仅使用1次。这需要合成更多寡核苷酸。对于半组合索引,使用的组合更加有限,使得仅重复一小部分孔的索引并且许多可能的组合不被使用。在实践中(此处未描述),使用了一种设计,在所述设计中,i7索引是唯一的,但是i5索引可以在384孔板中重复多达四次。对于大部分Swabseq开发,使用了允许1536个组合或样品运行的半组合索引(384x96)或唯一双重索引(384UDI)。

[0145] 图20显示了使用混合模型对索引错误分配的计算校正。为了扩大能够测试的样品的数目,可以使用组合索引策略。在此实验中,i5上的单个索引用于唯一鉴定板,并且96个i7索引用于鉴定孔。在(A)中,基于RT-qPCR是否检测到Covid(x轴),绘制了临床样品的S2与S2标的比率(y轴)。过滤SARS-CoV-2阳性样品,得到 $Ct < 32$ 。可以观察到在各板内索引错误分配的影响,因为在所有样品内具有高S2和S2标的总和的i7索引在各板中共享相同i7条形码(颜色)。在(B)中,在通过将i7条形码的身份处理为随机效应来对 $\log_{10}(S2+1/S2_{\text{标}}+1)$ 比率进行计算校正后,为A中的数据绘制了最佳线性无变差预测残差(y轴)。

[0146] 图21显示了定量索引错误分配作为S2读段中的噪声源的作用。在(A)中,来自使用唯一双重索引设计的运行v19的每对i5和i7索引对的病毒S2+S2标计数的匹配矩阵。沿对角线的索引对对应于实验中存在的样品的预期索引对(预期匹配索引),并且对角线以外的索引对对应于索引错误分配事件。在(B)中,已知病毒RNA的量为零的样品的病毒S计数与标计

数的比率分布。平均比率为0.00028。在(C)中,每个样品的i7错误分配事件的数目相对于病毒S2+S2标计数的数目。在(D)中,每个样品的i5错误分配事件的数目相对于病毒S2+S2标计数的数目。

条形码

[0147] 用作索引的可变核苷酸序列(条形码)可以被包括在本文所述的第一或第二PCR引物集中的任一个上。此外,可以在单独的文库制备反应中添加条形码。本文所述的可变核苷酸序列可用作样品索引,以将从本文使用的测序反应获得的结果去卷积。

[0148] 一旦细胞的内容物通过裂解剂被释放到它们各自的分区中,其中包含的大分子组分(例如,样品的大分子成分,诸如RNA、DNA或蛋白质)可以在分区内被进一步处理。根据本文所述的方法和系统,单独的样品的大分子组分内容物可以具有唯一标识符,使得在表征那些大分子组分时,它们可以归属为来源于相同的样品或颗粒。通过将唯一标识符特定地分配给单独的样品或样品组来提供将特征归属于单独的样品或样品组的能力。可以分配例如以核酸条形码形式的唯一标识符或将其与单独的样品或样品群组缔合,以使用唯一标识符对样品的大分子组分(并且因此其特征)加标签或进行标记。然后,可以使用这些唯一标识符将样品的组分和特征归属于单独的样品或样品组。

[0149] 在一些方面中,这通过将单独的样品或样品组与唯一标识符或包含唯一分子标识符序列(UMI)的条形码共同分区来进行。在一些方面中,唯一标识符以核酸分子(例如,寡核苷酸)的形式提供,所述核酸分子包含核酸条形码序列,所述核酸条形码序列可以连接至单独的样品的核酸内容物或以其他方式与其缔合,或者连接至样品的其他组分,并且特别是那些核酸的片段。核酸分子被分配,使得在给定分区中的核酸分子之间,其中包含的核酸条形码序列是相同的,但是在不同分区之间,核酸分子可以并且确实具有不同的条形码序列,或者至少表示在给定分析中所有分区中的大量不同条形码序列。在一些方面中,只有一个核酸条形码序列可以与给定的分区缔合,但是在一些实施方案中,可以存在两个或更多个不同的条形码序列。

[0150] 核酸条形码序列可以在核酸分子(例如,寡核苷酸)的序列内包括约6至约20个或更多个核苷酸。核酸条形码序列可包括约6至约20、30、40、50、60、70、80、90、100或更多个核苷酸。在一些实施方案中,条形码序列的长度可以是约6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个核苷酸或更长。在一些实施方案中,条形码序列的长度可以是至少约6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个核苷酸或更长。在一些实施方案中,条形码序列的长度可以是至多约6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个核苷酸或更短。这些核苷酸可以是完全连续的,即在相邻核苷酸的单个链段中,或者它们可以被分成两个或更多个由1个或多个核苷酸分开的单独子序列。在一些实施方案中,分开的条形码子序列的长度可为约4至约16个核苷酸。在一些实施方案中,条形码子序列可以是约4、5、6、7、8、9、10、11、12、13、14、15、16个核苷酸或更长。在一些实施方案中,条形码子序列可以是至少约4、5、6、7、8、9、10、11、12、13、14、15、16个核苷酸或更长。在一些实施方案中,条形码子序列可以是至多约4、5、6、7、8、9、10、11、12、13、14、15、16个核苷酸或更短。

[0151] 共同分区的核酸分子还可以包含可用于处理来自共分配的样品的核酸的其他功能序列。这些序列包括例如靶向或随机/通用扩增引物序列,其用于扩增分区内单独的样品的基因组DNA,同时连接相关的条形码序列、测序引物或引物鉴定位点、杂交或探测序列,例

如用于鉴定序列的存在或用于拉下条形码化核酸,或许多其他潜在的功能序列中任一种。还可以使用共分区寡核苷酸的其他机制,包括例如两个或更多个分区的聚结,其中一个分区包含寡核苷酸,或寡核苷酸微分配到分区中,例如微流体系统内的分区。在一些实施方案中,引物包含条形码寡核苷酸。在一些实施方案中,引物序列是与模板核酸分子中的序列互补的靶向引物序列。在一些实施方案中,第一核酸分子还包含一个或多个功能序列,并且其中第二核酸分子包含一个或多个功能序列。在一些实施方案中,一个或多个功能序列选自衔接子序列、附加引物序列、引物退火序列、测序引物序列、被配置为连接至测序仪的流动池的序列和唯一分子标识符序列。

[0152] 例如,将上文所述的条形码化核酸分子(例如,条形码化寡核苷酸)添加至样品中。在一些实施方案中,分区包含具有相同条形码序列的条形码化寡核苷酸。在一些实施方案中,多个分区中的分区包含具有相同条形码序列的条形码化寡核苷酸,其中多个分区内的每个分区包含唯一条形码序列。在一些实施方案中,条形码化寡核苷酸群体提供各种条形码序列文库,其包括至少约1,000种不同的条形码序列、至少约5,000种不同的条形码序列、至少约10,000种不同的条形码序列、至少约50,000种不同的条形码序列、至少约100,000种不同的条形码序列、至少约1,000,000种不同的条形码序列、至少约5,000,000种不同的条形码序列或至少约10,000,000种不同的条形码序列或更多。另外,每个条形码化寡核苷酸可以具有大量连接的核酸(例如,寡核苷酸)分子。特别地,在单独的条形码化寡核苷酸上包括条形码序列的核酸分子的分子数可以是至少约1,000个核酸分子、至少约5,000个核酸分子、至少约10,000个核酸分子、至少约50,000个核酸分子、至少约100,000个核酸分子、至少约500,000个核酸、至少约1,000,000个核酸分子、至少约5,000,000个核酸分子、至少约10,000,000个核酸分子、至少约50,000,000个核酸分子、至少约100,000,000个核酸分子、至少约250,000,000个核酸分子并且在一些实施方案中至少约10亿个核酸分子或更多。给定条形码化寡核苷酸的核酸分子可包括相同(或共同)的条形码序列、不同条形码序列或两者的组合。给定条形码化寡核苷酸的核酸分子可包括多个核酸分子组。给定组的核酸分子可以包括相同的条形码序列。相同的条形码序列可以与另一组的核酸分子的条形码序列不同。

[0153] 此外,当分配条形码化寡核苷酸群体时,所得到的分区群体还可以包括各种条形码文库,其包括至少约1,000种不同的条形码序列、至少约5,000种不同的条形码序列、至少约10,000种不同的条形码序列、至少约50,000种不同的条形码序列、至少约100,000种不同的条形码序列、至少约1,000,000种不同的条形码序列、至少约5,000,000种不同的条形码序列、或至少约10,000,000种不同的条形码序列。另外,群体的每个分区可以包括至少约1,000个核酸分子、至少约5,000个核酸分子、至少约10,000个核酸分子、至少约50,000个核酸分子、至少约100,000个核酸分子、至少约500,000个核酸、至少约1,000,000个核酸分子、至少约5,000,000个核酸分子、至少约10,000,000个核酸分子、至少约50,000,000个核酸分子、至少约100,000,000个核酸分子、至少约250,000,000个核酸分子并且在一些实施方案中至少约10亿个核酸分子。

[0154] 在一些实施例中,可能希望在给定分区内并入多个不同的条形码。例如,在一些实施方案中,分区内的条形码化寡核苷酸可包含(1)在分区内所有条形码化寡核苷酸共享的共同条形码序列和(2)在每个条形码化寡核苷酸内不同的唯一分子标识符或附加条形码序列。共同条形码序列可以在后续处理提供较大鉴定保证,例如,通过向给定分区提供较强的

条形码地址或归属,作为对给定分区的输出的重复或独立的确认。

[0155] 在一些实施方案中,条形码化寡核苷酸连接至珠子,其中连接至特定珠子的所有核酸分子将包括相同的核酸条形码序列,但是在所使用的珠子群体中表示大量不同的条形码序列。在一些实施方案中,水凝胶珠粒(例如包含聚丙烯酰胺聚合物基质)用作核酸分子进入分区的固体支持物和递送载体,因为它们能够携带大量核酸分子,并且可以配置为在暴露于特定刺激时释放那些核酸分子,如本文其他地方所述。

[0156] 核酸分子(例如,寡核苷酸)在对珠子施加特定刺激时可从珠子释放。在一些实施方案中,刺激可以是光刺激,例如通过裂解光不稳定键联以释放核酸分子。在其他实施方案中,可以使用热刺激,其中珠子环境的温度升高将导致键联裂解或核酸分子从珠子的其他释放。在其他实施方案中,可以使用化学刺激,其裂解核酸分子与珠子的键联,或者以其他方式导致核酸分子从珠子释放。在一个实施方案中,此类组合物包含上文所述的用于包封样品的聚丙烯酰胺基体,并且可以通过暴露于还原剂诸如DTT而降解以释放连接的核酸分子。

[0157] 可以考虑在本公开的方法中使用支持物,例如孔、基体、棒、容器或珠子。支持物可以具有任何可用的特征部和特性,诸如任何可用的大小、表面化学、流动性、坚固性、密度、孔隙率和组成。在一些实施方案中,支持物是板上孔的表面。在一些实施方案中,支持物可以是珠子,诸如凝胶珠子。珠子可以是固体或半固体。珠子的其他细节在本文别处提供。

[0158] 支持物(例如,珠子)可以包含对其进行功能化的锚定序列(例如,如本文所述的)。锚定序列可以通过例如二硫键连接至支持物。锚定序列可以包括部分读段序列和/或流动池功能序列。这种序列可以允许通过测序仪(例如,Illumina测序仪)对连接到序列的核酸分子进行测序。不同的锚定序列可能可用于不同的测序应用。锚定序列可以包括例如TruSeq或Nextera序列。锚定序列可以具有任何可用的特征,诸如任何可用的长度和核苷酸组成。例如,锚定序列可以包含2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或更多个核苷酸。在一些实施方案中,锚定序列可以包含15个核苷酸。锚定序列的核苷酸可以是天然存在的或非天然存在的(例如,如本文所述的)。珠子可以包含与其连接的多个锚定序列。例如,珠子可以包含与其连接的多个第一锚定序列。在一些实施方案中,珠子可以包含两个或更多个与其连接的不同锚定序列。例如,珠子可以包含与其连接的多个第一锚定序列(例如,Nextera序列)和多个第二锚定序列(例如,TruSeq序列)。对于包含两个或更多个与其连接的不同锚定序列的珠子,每个不同锚定序列的序列可以与在珠子远侧末端的每个其他锚定序列的序列区分开。例如,不同锚定序列可以在离珠子最远的2、3、4、5、6、7、8、9、10或更多个核苷酸中包含一个或多个核苷酸差异。

[0159] 在一些实施方案中,可以在同一支持物(例如,珠子)上生成多个不同条形码分子(例如,核酸条形码分子)。例如,可以在同一支持物上生成两个不同条形码分子。替代地,可以在同一支持物上生成三个或更多个不同条形码分子。连接至同一支持物的不同条形码分子可以包含一个或多个不同序列。例如,不同条形码分子可以包含一个或多个不同条形码序列和/或其他序列(例如,起始序列)。在一些实施方案中,连接至同一支持物的不同条形码分子可以包含相同条形码序列。连接至同一支持物的不同条形码分子可以包含相同或不同的条形码序列。类似地,不同条形码分子可以包含相同或不同的唯一分子标识符(UMI)。

下一代测序

[0160] 如本文所公开的方法中所述,使用核酸分子的测序,并且其可用于检测和诊断病原体感染。一般来讲,测序是指用于确定一种或多种多核苷酸中核苷酸碱基的序列的方法和技术。多核苷酸可以是例如脱氧核糖核酸(DNA)或核糖核酸(RNA),包括其变体或衍生物(例如,单链DNA)。测序可以通过目前可用的各种系统进行,诸如但不限于通过Illumina、Pacific Biosciences、Oxford Nanopore或Life Technologies(Ion Torrent)的测序系统。此类装置可以提供对应于受试者(例如,人)的遗传信息的多个原始遗传数据,如从受试者提供的样品由装置所生成。在一些情况下,本文所提供的系统和方法可以与蛋白质组信息一起使用。替代地或另外,可以使用核酸扩增、聚合酶链反应(PCR)(例如,数字PCR、定量PCR或实时PCR)或等温扩增来进行测序。此类系统可以提供对应于受试者(例如,人)的遗传信息的多个原始遗传数据,如从受试者提供的样品由系统所生成。在一些实例中,此类系统提供测序读段(本文中也称为“读段”)。读段可以包括一串核苷酸碱基,其对应于已经测序的核酸分子的序列。在一些情况下,本文所提供的系统和方法可以与蛋白质组信息一起使用。

[0161] 下一代测序包括许多能够生成大量序列信息的技术,但不包括桑格测序或克萨姆-吉尔伯特测序。一般来讲,下一代测序涵盖单分子实时测序、边合成边测序、离子半导体测序等等。示例性下一代测序的机器可以包括来自Illumina, Inc的MiniSeq、iSeq100、NextSeq 1000、NextSeq 2000、NovaSeq 6000、NextSeq 550系列等等;来自Thermo Fisher Scientific的Ion Torrent机器;或来自Pacific Biosciences的Sequel系统。

[0162] 与本文方法一起使用的下一代测序机器可以在24小时的时期内从单个机器生成至少1、5、10、15、25、50、75、100、200、300千兆碱基的数据或更多。

[0163] 与本文方法一起使用的下一代测序机器可以在24小时的时期内从单个机器生成至少100、100、400、1,000、1,500、2,500、5,000、7,500、10,000、20,000、30,000、50,000或100,000万个序列读段的数据或更多。

[0164] 还包括计算机程序、计算装置或分析平台/系统,其接收和分析测序数据并输出一个或多个报告,所述报告可以通过服务器、分析门户或电子邮件以电子方式传输或访问。计算装置或分析平台可以根据本文所述的算法和方法进行操作。

反应混合物

[0165] 本文还提供了用于确定生物样品中是否存在病毒核酸的反应混合物。在一些实施方案中,反应混合物包含本文提供的合成核酸、所述生物样品的至少一部分和足以扩增所述生物样品中的所述病毒核酸(如果存在)的一种或多种酶或试剂。

[0166] 生物样品可以是本文提供的任何生物样品。在一些实施方案中,生物样品是人生物样品。在一些实施方案中,生物样品包括唾液、口腔拭子、鼻咽拭子或中鼻甲拭子。在一些实施方案中,所述生物样品包括唾液或鼻咽拭子。在一些实施方案中,所述生物样品包括唾液。在一些实施方案中,所述生物样品包括鼻咽拭子。

[0167] 合成核酸可以是本文提供的任何合成核酸。在一些实施方案中,合成核酸是本文提供的SARS-CoV-2合成RNA核酸。在一些实施方案中,所述SARS-CoV-2合成RNA核酸以约10个拷贝/反应混合物至约500个拷贝/反应混合物的浓度存在。在一些实施方案中,所述SARS-CoV-2合成RNA核酸以约10、20、30、40、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290、300、320、340、360、380、400、420、440、460、480或500个拷贝/反应混合物的浓度存在。

[0168] 病毒核酸可以是本文提供的任何病毒核酸。在一些实施方案中,所述病毒核酸是甲型流感病毒、乙型流感病毒或冠状病毒核酸。在一些实施方案中,所述冠状病毒核酸是SARS-COV-2核酸。

[0169] 在一些实施方案中,酶或试剂包括逆转录酶、dNTP、对所述病毒核苷酸序列具有特异性的引物对、对样品对照核苷酸序列具有特异性的引物对、镁盐或其组合。

[0170] 在一些实施方案中,反应混合物包含一种或多种可用于扩增病毒核酸的酶。在一些实施方案中,酶是逆转录酶。逆转录酶的非限制性实例包括禽成髓细胞性白血病毒(AMV)逆转录酶和莫洛尼鼠白血病毒(M-MuLV, MMLV)逆转录酶及其变体。

[0171] 在一些实施方案中,反应包括脱氧核苷酸三磷酸(dNTP)。在一些实施方案中,试剂盒包含扩增病毒核酸所必需的dNTP中的每一种以及任何其他所需核酸(例如,dATG、dCTP、dTTP、dGTP)的混合物。

[0172] 在一些实施方案中,反应混合物包含对病毒核苷酸序列具有特异性的引物对。对病毒核苷酸序列具有特异性的引物对可以是本文提供的任何引物对。对所述病毒核苷酸序列具有特异性的所述引物对对甲型流感病毒核苷酸序列、乙型流感病毒核苷酸序列和冠状病毒核苷酸序列具有特异性。在一些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对对冠状病毒S1或N2序列具有特异性。

[0173] 在一些实施方案中,反应混合物包含对样品对照核苷酸序列具有特异性的引物对。在一些实施方案中,对样品对照核苷酸序列具有特异性的引物对生物样品所来源的生物体所表达的内源核苷酸序列具有特异性。在一些实施方案中,对样品对照核苷酸具有特异性的引物对对持家基因具有特异性。在一些实施方案中,对样品对照核苷酸序列具有特异性的引物对对GAPDH、RPP30或ACTB具有特异性。在一些实施方案中,对样品对照核苷酸具有特异性的引物对对RPP30具有特异性。

[0174] 在一些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对以约50微摩尔至约250微摩尔的浓度存在。在一些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对以约50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240或250微摩尔的浓度存在。在一些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对以约100微摩尔的浓度存在。在一些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对或对所述样品对照核苷酸序列具有特异性的所述引物对以约200微摩尔的浓度存在。

[0175] 在一些实施方案中,反应混合物包含镁盐。在一些实施方案中,镁盐的含量足以使反应的酶(例如,逆转录酶)发挥作用并扩增靶核酸。在一些实施方案中,镁盐是氯化镁。在一些实施方案中,反应混合物包含浓度约0.1mM至约50mM的镁离子。在一些实施方案中,镁离子的浓度为约1mM至约10mM。

[0176] 在一些实施方案中,所述反应混合物的体积为约10微升至约100微升。在一些实施方案中,所述反应混合物的体积为约20微升至约90微升、约30微升至约80微升或约40微升至约60微升。在一些实施方案中,所述反应混合物的体积为约10、20、30、40、50、60、70、80、90或100微升。

试剂盒

[0177] 本文还描述了一种试剂盒,其包括本文所述的寡核苷酸引物对。在某些实施方案中,是一种试剂盒,其包括本文所述的寡核苷酸引物对和根据本公开的合成核酸。合成核酸可用于S1、N2扩增。试剂盒还可以包括足以用于扩增的试剂:诸如用于逆转录、PCR扩增和/或测序反应的缓冲液(以10x、5x、2x、或1x浓度提供)、dNTP、逆转录酶、PCR扩增酶(例如,Taq聚合酶和/或其变体)。这些组分可以适当地单独地或以组合方式包装在小瓶或容器中。

[0178] 本文还描述了一种试剂盒,其用于确定生物样品中是否存在病毒核酸。在一些实施方案中,试剂盒包括本文所提供的合成核酸和足以扩增来自生物样品的病毒核酸的一种或多种酶或试剂。

[0179] 在一些实施方案中,酶或试剂包括逆转录酶、dNTP、对所述病毒核苷酸序列具有特异性的引物对、对样品对照核苷酸序列具有特异性的引物对、镁盐或其组合。

[0180] 在一些实施方案中,试剂盒包括一种或多种可用于扩增病毒核酸的酶。在一些实施方案中,酶是逆转录酶。逆转录酶的非限制性实例包括禽成髓细胞性白血病病毒(AMV)逆转录酶和莫洛尼鼠白血病病毒(M-MuLV,MMLV)逆转录酶及其变体。

[0181] 在一些实施方案中,试剂盒包括脱氧核苷酸三磷酸(dNTP)。在一些实施方案中,试剂盒包含扩增病毒核酸所必需的dNTP中的每一种以及任何其他所需核酸(例如,dATG、dCTP、dTTP、dGTP)的混合物。

[0182] 在一些实施方案中,试剂盒包括对病毒核苷酸序列具有特异性的引物对。对病毒核苷酸序列具有特异性的引物对可以是本文提供的任何引物对。对所述病毒核苷酸序列具有特异性的所述引物对对甲型流感病毒核苷酸序列、乙型流感病毒核苷酸序列和冠状病毒核苷酸序列具有特异性。在一些实施方案中,对所述病毒核苷酸序列具有特异性的所述引物对对冠状病毒S1或N2序列具有特异性。

[0183] 在一些实施方案中,试剂盒包括对样品对照核苷酸序列具有特异性的引物对。在一些实施方案中,对样品对照核苷酸序列具有特异性的引物对对生物样品所来源的生物体所表达的内源核苷酸序列具有特异性。在一些实施方案中,对样品对照核苷酸具有特异性的引物对对持家基因具有特异性。在一些实施方案中,对样品对照核苷酸序列具有特异性的引物对对GAPDH、RPP30或ACTB具有特异性。在一些实施方案中,对样品对照核苷酸具有特异性的引物对对RPP30具有特异性。

[0184] 在一些实施方案中,试剂盒包括镁盐。在一些实施方案中,镁盐的含量足以使试剂盒的酶(例如,逆转录酶)发挥作用。在一些实施方案中,镁盐是氯化镁。

[0185] 以下实施方案列举了本文所公开的特征组合的非限制性排列。还设想了特征组合的其他排列。特别地,这些编号的实施方案中的每一个都被设想为依赖于每个先前或后续编号的实施方案或与其相关,与它们列出的顺序无关。

[0186] 1. 一种诊断病原体感染的个体的方法,所述方法包括:(a)提供来自所述个体的生物样品;(b)将来自所述个体的所述生物样品与裂解剂接触,以获得裂解的生物样品;(c)对所述裂解的生物样品进行聚合酶链反应(PCR),以获得PCR扩增的裂解的生物样品,其中对所述裂解的生物样品的所述PCR反应用第一PCR引物集进行,其中所述第一PCR引物集扩增病原体核酸序列;(d)使用下一代测序对所述PCR扩增的裂解的生物样品进行测序;以及(e)如果通过所述PCR或通过所述测序检测到病原体序列,则提供针对所述病原体感染的阳性

诊断,或者如果通过所述PCR或通过所述测序未检测到病原体序列,则提供针对所述个体的阴性诊断。2.根据实施方案1所述的方法,其中所述个体是人类个体。3.根据实施方案1或2所述的方法,其中所述病原体感染包括细菌感染、病毒感染或真菌感染及其组合。4.根据实施方案1至3中任一项所述的方法,其中所述裂解剂包括加热到至少50℃的水或PCR反应缓冲液。5.根据实施方案1至3中任一项所述的方法,其中所述裂解剂包括加热到至少90℃的水或PCR反应缓冲液。6.根据实施方案1至5中任一项所述的方法,其中所述细菌感染是链球菌属、假单胞菌属、志贺氏菌属、弯曲菌属、沙门氏菌属、梭菌属或埃希氏菌属及其组合的感染。7.根据实施方案1至5中任一项所述的方法,其中所述感染是念珠菌属、芽生菌属、隐球菌属、球孢子菌属、组织胞浆菌属、副球孢子菌属、孢子丝菌属或肺囊虫属及其组合的感染。8.根据实施方案1至5中任一项所述的方法,其中所述病毒感染是DNA病毒的感染。9.根据实施方案8所述的方法,其中所述DNA病毒包括乙型肝炎病毒、丙型肝炎病毒、乳头瘤病毒、E-B病毒、水痘病毒或天花病毒及其组合。10.根据实施方案1至5中任一项所述的方法,其中所述病毒感染是RNA病毒的感染。11.根据实施方案10所述的方法,其中所述RNA病毒包括流感病毒、冠状病毒、脊髓灰质炎病毒、麻疹病毒、埃博拉病毒、反转录病毒或正粘病毒。12.根据实施方案11所述的方法,其中所述病毒感染是冠状病毒感染。13.根据实施方案12所述的方法,其中所述冠状病毒感染是SARS-COV-2感染。14.根据实施方案1至13中任一项所述的方法,其中来自所述个体的所述生物样品来自血液样品、血浆样品、血清样品、口腔拭子、尿液样品、精液样品、阴道拭子、粪便样品、鼻咽拭子、中鼻甲拭子或其任何组合。15.根据实施方案14所述的方法,其中来自所述个体的所述生物样品来自鼻咽拭子、中鼻甲拭子或其任何组合。16.根据实施方案1至15中任一项所述的方法,其包括向所述裂解剂或所述裂解的生物样品中添加合成核酸。17.根据实施方案1至15中任一项所述的方法,其包括所述裂解剂或所述裂解的生物样品中添加多个合成核酸,所述多个包括具有不同序列的合成核酸。18.根据实施方案17所述的方法,其中所述多个包括四种。19.根据实施方案1至18中任一项所述的方法,其包括向所述方法中添加合成核酸。20.根据实施方案1至18中任一项所述的方法,其包括向所述方法中添加多个合成核酸,所述多个包括具有不同序列的合成核酸。21.根据实施方案20所述的方法,其中所述多个包括四种。22.根据实施方案19所述的方法,其中所述合成核酸是RNA。23.根据实施方案19所述的方法,其中所述合成核酸是DNA。24.根据实施方案19或23中任一项所述的方法,其中所述合成核酸包含被配置成由所述第一PCR引物集结合的序列集。25.根据实施方案19至24中任一项所述的方法,其中所述第一引物集扩增所述病原体核酸序列和所述合成核酸。26.根据实施方案19至24中任一项所述的方法,其中所述合成核酸序列包含与所述病原体核酸序列不相同的核苷酸序列。27.根据实施方案1至26中任一项所述的方法,其包括对所述裂解的生物样品执行逆转录反应。28.根据实施方案27所述的方法,其中所述逆转录反应在所述执行所述聚合酶链反应之前执行。29.根据实施方案1至26中任一项所述的方法,其中所述逆转录反应在没有进一步纯化所述裂解的生物样品的情况下执行。30.根据实施方案1至26中任一项所述的方法,其中对所述裂解的生物样品的所述逆转录反应产生病毒cDNA。31.根据实施方案30所述的方法,其中所述病毒cDNA是冠状病毒cDNA。32.根据实施方案31所述的方法,其中所述冠状病毒cDNA是SARS-COV-2 cDNA。33.根据实施方案1至31中任一项所述的方法,其中所述逆转录反应和所述PCR是单步骤反应。34.根据实施方案1至33中任一项所述的方法,其中所述PCR是终点分析。35.

根据实施方案1至34中任一项所述的方法,其中所述PCR不是实时PCR反应。36.根据实施方案1至35中任一项所述的方法,其中所述第一PCR引物集扩增冠状病毒核酸序列。37.根据实施方案36所述的方法,其中所述冠状病毒核酸序列是SARS-COV-2核酸序列。38.根据实施方案37所述的方法,其中所述SARS-COV-2核酸序列包含N1或S2基因。39.根据实施方案1至38中任一项所述的方法,其包含第二PCR引物集,其中所述第二引物集扩增所述个体的核酸序列。40.根据实施方案39所述的方法,其中所述第二PCR引物集扩增人核酸序列。41.根据实施方案40所述的方法,其中所述第二PCR引物集扩增选自以下的人核酸序列:GAPDH、ACTB、RPP30及其组合。42.根据实施方案41所述的方法,其中所述第二PCR引物集扩增人RPP30。43.根据实施方案40至42中任一项所述的方法,其中所述第二PCR引物集包含具有测序衔接子序列的引物和不具有测序衔接子序列的引物的混合物。44.根据实施方案43所述的方法,其中具有测序衔接子序列的引物与不具有测序衔接子序列的引物的比率是约1:1、约1:2、约1:3或约1:4。45.根据实施方案1至44中任一项所述的方法,其中所述PCR包括30至45个扩增循环。46.根据实施方案45所述的方法,其中所述PCR包括35至45个扩增循环。47.根据实施方案45所述的方法,其中所述PCR包括39至42个扩增循环。48.根据实施方案1至47中任一项所述的方法,其中所述第一PCR引物集、所述第二PCR引物集或所述第一PCR引物集和所述第二PCR引物集包含有包含可变核苷酸序列的核酸序列。49.根据实施方案48所述的方法,其中所述可变核苷酸序列是所述个体唯一的样品ID。50.根据实施方案1至49中任一项所述的方法,其中所述第一PCR引物集、所述第二PCR引物集或所述第一PCR引物集和所述第二PCR引物集包含用于下一代测序反应的衔接子序列。51.根据实施方案39至50中任一项所述的方法,其包括向所述方法中添加第二合成核酸。52.根据实施方案51所述的方法,其中所述第二合成核酸是RNA。53.根据实施方案51所述的方法,其中所述第二合成核酸是DNA。54.根据实施方案51至53中任一项所述的方法,其中所述第二合成核酸包含被配置成由所述第二PCR引物集结合的序列集。55.根据实施方案51至54中任一项所述的方法,其中所述第二引物集扩增所述人核酸和所述合成核酸。56.根据实施方案51至54中任一项所述的方法,其中所述合成核酸序列包含与所述人核酸序列不相同的核苷酸序列。57.根据实施方案1至56中任一项所述的方法,其中所述方法可以检测少于10个病原体基因组拷贝。58.根据实施方案1至56中任一项所述的方法,其中所述方法可以检测少于5个病原体基因组拷贝。59.根据实施方案57或58所述的方法,其中所述病原体基因组是冠状病毒基因组。60.根据实施方案57或58所述的方法,其中所述冠状病毒基因组是SARS-COV-2基因组。61.根据实施方案1至60中任一项所述的方法,其中所述对冠状病毒呈阳性的诊断,如果通过所述PCR检测到冠状病毒序列,则是SARS-COV-2诊断。62.根据实施方案1至61中任一项所述的方法,其中所述方法确定冠状病毒的毒株。63.根据实施方案1至61中任一项所述的方法,其中所述方法确定COVID-19的毒株。

[0187] 64.一种诊断病原体感染的个体的方法,所述方法包括使用第一PCR引物集扩增来自所述个体的生物样品的核酸,从而获得扩增的核酸,其中所述第一PCR引物集扩增所述生物样品的病原体核酸序列和合成核酸序列,其中所述合成核酸序列与所述病原体核酸序列相差至少一个核苷酸。65.根据实施方案64所述的方法,其中所述个体是人类个体。66.根据实施方案64或65所述的方法,其中所述病原体感染包括细菌感染、病毒感染或真菌感染。67.根据实施方案64至66中任一项所述的方法,其中所述细菌感染是链球菌属、假单胞菌

属、志贺氏菌属、弯曲菌属、沙门氏菌属、梭菌属或埃希氏菌属及其组合的感染。68. 根据实施方案64至66中任一项所述的方法,其中所述感染是念珠菌属、芽生菌属、隐球菌属、球孢子菌属、组织胞浆菌属、副球孢子菌属、孢子丝菌属或肺囊虫属及其组合的感染。69. 根据实施方案64至66中任一项所述的方法,其中所述病毒感染是DNA病毒的感染。70. 根据实施方案69所述的方法,其中所述DNA病毒包括乙型肝炎病毒、丙型肝炎病毒、乳头瘤病毒、E-B病毒、水痘病毒、天花病毒或它们的任何组合。71. 根据实施方案64至66中任一项所述的方法,其中所述病毒感染是RNA病毒的感染。72. 根据实施方案71所述的方法,其中所述RNA病毒包括流感病毒、冠状病毒、脊髓灰质炎病毒、麻疹病毒、埃博拉病毒、反转录病毒或正粘病毒。73. 根据实施方案71所述的方法,其中所述病毒感染是冠状病毒感染。74. 根据实施方案73所述的方法,其中所述冠状病毒感染是SARS-COV-2感染。75. 根据实施方案64至74中任一项所述的方法,其中来自所述个体的所述生物样品来自血液样品、血浆样品、血清样品、口腔拭子、尿液样品、精液样品、阴道拭子、粪便样品、鼻咽拭子、中鼻甲拭子或其任何组合。76. 根据实施方案75所述的方法,其中来自所述个体的所述生物样品来自鼻咽拭子、中鼻甲拭子或其任何组合。77. 根据实施方案64至76中任一项所述的方法,其中所述合成核酸是RNA。78. 根据实施方案64至76中任一项所述的方法,其中所述合成核酸是DNA。79. 根据实施方案64至78中任一项所述的方法,其中所述合成核酸包含被配置成由所述第一PCR引物集结合的序列集。80. 根据实施方案64至79中任一项所述的方法,其中所述合成核酸序列与所述病原体核酸序列相差至少5个核苷酸。81. 根据实施方案64至79中任一项所述的方法,其中所述病原体感染是基于合成核酸序列与病原体核酸序列的比率诊断的。82. 根据实施方案64至81中任一项所述的方法,其包括对所述生物样品的所述核酸执行逆转录反应。83. 根据实施方案64至81中任一项所述的方法,其中对所述生物样品的所述核酸的所述逆转录反应产生冠状病毒cDNA。84. 根据实施方案83所述的方法,其中所述冠状病毒cDNA是SARS-COV-2 cDNA。85. 根据实施方案64至84中任一项所述的方法,其中所述扩增核酸构成PCR反应。86. 根据实施方案85所述的方法,其中所述PCR反应是终点分析。87. 根据实施方案64至86中任一项所述的方法,其中所述PCR反应不是实时PCR反应。88. 根据实施方案64至87中任一项所述的方法,其中所述第一PCR引物集扩增冠状病毒核酸序列。89. 根据实施方案88所述的方法,其中所述冠状病毒核酸序列是SARS-COV-2核酸序列。90. 根据实施方案89所述的方法,其中所述SARS-COV-2核酸序列包含N1或S2基因。91. 根据实施方案64至90中任一项所述的方法,其中所述第一PCR引物集包含有包含可变核苷酸序列的核酸序列。92. 根据实施方案91所述的方法,其中所述可变核苷酸序列是所述个体唯一的样品ID。93. 根据实施方案64至92中任一项所述的方法,其中所述第一PCR引物集包含用于下一代测序反应的衔接子序列。94. 根据实施方案64至92中任一项所述的方法,其包括使用第二PCR引物集扩增所述生物样品的核酸,其中所述第二PCR引物集扩增人核酸序列。95. 根据实施方案94所述的方法,其中所述第二PCR引物集扩增选自以下的核酸序列:GAPDH、ACTB、RPP30及其组合。96. 根据实施方案95所述的方法,其中所述第二PCR引物集扩增人RPP30。97. 根据实施方案94至96中任一项所述的方法,其中所述第二PCR引物集包含具有测序衔接子序列的引物和不具有测序衔接子序列的引物的混合物。98. 根据实施方案97所述的方法,其中具有测序衔接子序列的引物与不具有测序衔接子序列的引物的比率是约1:1、约1:2、约1:3或约1:4。99. 根据实施方案64至98中任一项所述的方法,其中所述PCR包括30至45个扩增循环。100. 根据实施方案99

所述的方法,其中所述PCR包括35至45个扩增循环。101.根据实施方案99所述的方法,其中所述PCR包括39至42个扩增循环。102.根据实施方案94至101中任一项所述的方法,其中所述第二PCR引物集包含有包含可变核苷酸序列的核酸序列。103.根据实施方案102所述的方法,其中所述可变核苷酸序列是所述个体唯一的样品ID。104.根据实施方案94至103中任一项所述的方法,其中所述第二PCR引物集包含用于下一代测序反应的衔接子序列。105.根据实施方案64至104中任一项所述的方法,其包括使用下一代测序技术,对所述生物样品的所述扩增的核酸进行测序。106.根据实施方案64至105中任一项所述的方法,其中所述方法可以检测少于10个病原体基因组拷贝。107.根据实施方案64至104中任一项所述的方法,其中所述方法可以检测少于5个病原体基因组拷贝。108.根据实施方案106或107所述的方法,其中所述病原体基因组是冠状病毒基因组。109.根据实施方案106或107所述的方法,其中所述病原体基因组是SARS-COV-2基因组。110.根据实施方案64至109中任一项所述的方法,其中所述方法确定冠状病毒的毒株。111.根据实施方案110所述的方法,其中所述方法确定SARS-COV-2的毒株。

[0188] 112.一种合成核酸,其包含5'近侧区、3'近侧区和间插核酸序列。113.根据实施方案112所述的合成核酸,其中所述合成核酸包含RNA。114.根据实施方案112所述的合成核酸,其中所述合成核酸包含DNA。115.根据实施方案112至114中任一项所述的合成核酸,其中所述5'近侧区包含病毒核酸序列。116.根据实施方案115所述的方法,其中所述病毒核酸序列包含冠状病毒序列。117.根据实施方案116所述的方法,其中所述病毒核酸序列包含SARS-COV-2序列。118.根据实施方案112至114中任一项所述的合成核酸,其中所述3'近侧区包含病毒核酸序列。119.根据实施方案118所述的方法,其中所述病毒核酸序列包含冠状病毒序列。120.根据实施方案119所述的方法,其中所述病毒核酸序列包含SARS-COV-2序列。121.根据实施方案112至120中任一项所述的合成核酸,其中所述5'近侧区、所述3'近侧区或所述5'近侧区和所述3'近侧区的长度少于约30个核苷酸。122.根据实施方案112至120中任一项所述的合成核酸,其中所述5'近侧区、所述3'近侧区或所述5'近侧区和所述3'近侧区的长度少于约25个核苷酸。123.根据实施方案112至120中任一项所述的合成核酸,其中所述5'近侧区、所述3'近侧区或所述5'近侧区和所述3'近侧区的长度少于约20个核苷酸。124.根据实施方案112至123中任一项所述的合成核酸,其中所述5'近侧区在所述合成核酸的5'末端处。125.根据实施方案112至123中任一项所述的合成核酸,其中所述3'近侧区在所述合成核酸的3'末端处。126.根据实施方案112至125中任一项所述的合成核酸,其中所述间插核酸序列与病毒核酸序列具有少于约99%、98%、97%、95%、90%、85%、80%或75%同一性。127.根据实施方案126所述的合成核酸,其中所述合成核酸序列是冠状病毒序列。128.根据实施方案126所述的合成核酸,其中所述合成核酸序列是SARS-COV-2序列。129.根据实施方案112至128中任一项所述的合成核酸在检测所述个体的病原体感染的方法中的用途。130.根据实施方案129所述的用途,其中所述病原体感染是冠状病毒感染。131.根据实施方案130所述的用途其中所述病毒感染是SARS-COV-2感染。

[0189] 132.一种用于检测病毒感染的核酸处理的方法,所述方法包括:(a)提供样品,其包含病毒核酸分子和宿主核酸分子;(b)通过使用包含第一条形码序列的第一引物对所述病毒核酸分子进行核酸延伸反应,生成条形码化病毒核酸分子;(c)通过使用包含第二条形码序列的第二引物对所述宿主核酸分子进行核酸衍生反应,生成条形码化宿主核酸分子;

(d) 对所述条形码化病毒核酸分子和所述条形码化宿主核酸分子进行测序,以鉴定(i)所述条形码序列和(ii)对应于所述病毒核酸分子或其衍生物和所述宿主核酸分子的序列;以及(e)如果在(d)中鉴定出对应于所述病毒核酸分子的所述序列,则提供针对病毒感染的阳性诊断。133. 根据实施方案132所述的方法,且其中(b)和(c)同时进行。134. 根据实施方案132至133中任一项所述的方法,其中所述第一引物还包含一个或多个选自以下的额外功能序列:引物序列、衔接子序列、引物退火序列、唯一分子标识符序列和捕获序列。135. 根据实施方案132至134中任一项所述的方法,其中所述第二引物还包含一个或多个选自以下的额外功能序列:引物序列、衔接子序列、引物退火序列、唯一分子标识符序列和捕获序列。136. 根据实施方案132至135中任一项所述的方法,其中(a)还包括合成核酸分子。137. 根据实施方案136所述的方法,其中所述合成核酸分子包含与所述病毒核酸分子和所述人核酸分子不同的合成序列。138. 根据实施方案137所述的方法,其中(b)通过使用所述第一引物进行所述核酸延伸,生成条形码化合成核酸分子。139. 根据实施方案137所述的方法,其中(c)通过使用所述第二引物进行所述核酸延伸,生成条形码化合成核酸分子。140. 根据实施方案132至139中任一项所述的方法,其中所述核酸延伸反应是逆转录反应。141. 根据实施方案132至139中任一项所述的方法,其中所述核酸延伸反应是聚合酶链反应。142. 根据实施方案132至139中任一项所述的方法,其中所述核酸延伸反应包括逆转录反酶反应、聚合酶链反应或其组合。143. 根据实施方案142所述的方法,其中(b),所述核酸延伸反应包括:(i)将所述第一引物与所述病毒核酸分子杂交;以及(ii)使用逆转录酶以延伸所述引物。144. 根据实施方案142所述的方法,其中(b),所述核酸延伸反应包括:(i)将所述第一引物与所述病毒核酸分子杂交;以及(ii)使用逆转录酶以延伸所述引物。145. 根据实施方案142所述的方法,其中(b),所述核酸延伸反应包括:(i)将所述第一引物与所述病毒核酸分子杂交;以及(ii)使用聚合酶以延伸所述引物。146. 根据实施方案142所述的方法,其中所述聚合酶链反应是终点聚合物链反应。147. 根据实施方案132至146中任一项所述的方法,其中所述方法还包括在(d)之前,扩增所述条形码化病毒核酸分子和所述条形码化宿主核酸分子。148. 根据实施方案132至147中任一项所述的方法,其中所述方法还包括在(d)之前,对所述条形码化病毒核酸分子和所述条形码化宿主核酸分子进行N次循环的聚合酶链反应。149. 根据实施方案148所述的方法,其中N大于35次循环。150. 根据实施方案148所述的方法,其中N大于40次循环。151. 根据实施方案148所述的方法,其中N大于45次循环。152. 根据实施方案148所述的方法,其中所述聚合酶链反应将一个或多个额外序列掺入到所述条形码化病毒核酸分子和所述条形码化宿主核酸分子中的一者或两者中,所述额外序列选自:样品索引序列、衔接子序列、引物序列、引物结合序列、被配置成与测序仪的流式细胞联接的序列和额外条形码序列。153. 根据实施方案132至152中任一项所述的方法,其中在(b)和(c)之后,将所述样品与在(b)和(c)中进行至少一个核酸延伸反应之后的一种或多种样品组合。154. 根据实施方案153所述的方法,其中将所述样品与在(b)和(c)中进行单轮所述核酸延伸反应之后的一种或多种样品组合。155. 根据实施方案132至154中任一项所述的方法,其中所述样品包含一个或多个细胞。156. 根据实施方案155所述的方法,其还包括在(b)之前,从所述一个或多个细胞释放所述病毒核酸分子和所述宿主核酸分子。157. 根据实施方案155所述的方法,其还包括在(b)之前,使所述样品经历足以从所述一个或多个细胞释放所述病毒核酸分子和所述宿主核酸分子的条件。158. 根据实施方案157所述的方法,其中在足以从所述一个

或多个细胞释放所述病毒核酸分子和所述宿主核酸分子的所述条件下进行 (b) 和 (c)。159. 根据实施方案157所述的方法,其中所述病毒核酸分子和所述宿主核酸分子在 (b) 和 (c) 之前未从所述样品纯化。160. 根据实施方案132至159中任一项所述的方法,其还包括在 (d) 中,在与所述样品缔合时,使用所述条形码序列以缔合所述病毒核酸分子或其衍生物或所述宿主核酸分子或其衍生物。161. 根据实施方案132至160中任一项所述的方法,其还包括在 (e) 中,使用所述条形码化病毒核酸分子和所述条形码化宿主核酸分子的所述条形码序列来鉴定所述样品,其中所述样品对应于测试所述病毒感染的受试者。162. 根据实施方案132至161中任一项所述的方法,其中所述样品获自受试者。163. 根据实施方案162所述的方法,其中所述受试者是人类。164. 根据实施方案162所述的方法,其中所述受试者是动物。165. 根据实施方案162所述的方法,其中所述宿主核酸分子是所述受试者的基因组的一部分。166. 根据实施方案162所述的方法,其中所述宿主核酸分子是所述受试者的转录组的一部分。167. 根据实施方案132至166中任一项所述的方法,其中所述宿主核酸分子编码遍在表达的蛋白质。168. 根据实施方案132至166中任一项所述的方法,其中所述宿主核酸分子是基因组DNA分子。169. 根据实施方案132至166中任一项所述的方法,其中所述宿主核酸分子是遍在转录的RNA分子。170. 根据实施方案132至169中任一项所述的方法,其中所述宿主核酸是GAPDH、ACTB、RPP30、及其组合。171. 根据实施方案132至170中任一项所述的方法,其中所述病毒核酸分子是病毒基因组的一部分。172. 根据实施方案171所述的方法,其中所述病毒是冠状病毒。173. 根据实施方案172所述的方法,其中所述冠状病毒选自严重急性呼吸综合征冠状病毒2 (SARS-COV-2)、严重急性呼吸综合征冠状病毒 (SARS-CoV) 和中东呼吸综合征冠状病毒 (MERS-CoV)。174. 根据实施方案173所述的方法,其中所述冠状病毒是SARS-COV-2。175. 根据实施方案174所述的方法,其中所述病毒是RNA病毒。176. 根据实施方案175所述的方法,其中所述RNA病毒包含双链RNA基因组。177. 根据实施方案175所述的方法,其中所述RNA病毒包含单链RNA基因组。178. 根据实施方案175所述的方法,其中所述RNA病毒选自冠状病毒、流感病毒、人免疫缺陷病毒和埃博拉病毒。179. 根据实施方案170所述的方法,其中所述病毒是DNA病毒。180. 根据实施方案34至169中任一项所述的方法,其还包括在 (a) 之前,分配所述样品。181. 根据实施方案180所述的方法,其中所述分区是孔。182. 根据实施方案180所述的方法,其中所述分区是多个孔中的孔。183. 根据实施方案182所述的方法,其中所述条形码序列是所述多个孔中的所述孔唯一的。184. 根据实施方案132至183中任一项所述的方法,其中 (b) 和 (c) 同时进行。185. 根据实施方案132至184中任一项所述的方法,其中所述病毒感染是SARS-COV-2。186. 根据实施方案132至183中任一项所述的方法,其中所述样品通过鼻咽拭子或中鼻甲拭子获自受试者。

[0190] 187. 一种组合物,其包含有包含第一核酸序列和第二核酸序列的合成核酸分子,其中 (1) 所述第一核酸序列与病原体核酸分子的序列相同,并且 (2) 所述第二核酸序列与所述病原体核酸分子的序列不相同。188. 根据实施方案187所述的组合物,其中所述第一核酸序列位于所述第二核酸序列的3'。189. 根据实施方案187至188中任一项所述的组合物,其中所述合成核酸分子还包含第三核酸序列,其中所述第三核酸序列与所述病原体核酸分子的第二序列相同。190. 根据实施方案176所述的组合物,其中所述第三核酸序列在所述第二核酸序列的5'。191. 根据实施方案187至190中任一项所述的组合物,其中所述第一核酸序列或所述第三核酸序列少于5、10、15、20、25或30个核苷酸。192. 根据实施方案187至191中

任一项所述的组合物,其中所述第二核酸序列所包含的核苷酸的总数少于25、50、100、150、200或500个核苷酸。193.根据实施方案187至191中任一项所述的组合物,其中所述第二核酸序列所包含的核苷酸的总数少于25、50、100、150、200或500个核苷酸。194.根据实施方案187至193中任一项所述的组合物,其中所述合成核酸分子是核糖核酸(RNA)分子、脱氧核糖核酸(DNA)分子或RNA-DNA杂交分子。195.根据实施方案187至194中任一项所述的组合物,其中所述第一核酸序列、所述第三核酸序列或所述第一核酸序列和所述第三核酸序列包含引物结合位点。196.根据实施方案187至194中任一项所述的组合物,其中所述组合物还包含病原体核酸分子。197.根据实施方案196所述的组合物,其中所述病原体核酸分子来自病原体,其中所述病原体包括细菌、病毒、真菌或其组合。198.根据实施方案197所述的组合物,其中所述细菌来自链球菌属、假单胞菌属、志贺氏菌属、弯曲菌属、沙门氏菌属、梭菌属或埃希氏菌属及其组合。199.根据实施方案197所述的组合物,其中真菌是念珠菌属、芽生菌属、隐球菌属、球孢子菌属、组织胞浆菌属、副球孢子菌属、孢子丝菌属或肺囊虫属及其组合。200.根据实施方案197所述的组合物,其中所述病毒是DNA病毒。201.根据实施方案200所述的组合物,其中所述DNA病毒包括乙型肝炎病毒、丙型肝炎病毒、乳头瘤病毒、E-B病毒、水痘病毒或天花病毒及其组合。202.根据实施方案197所述的组合物,其中所述病毒是RNA病毒。203.根据实施方案202所述的组合物,其中所述RNA病毒包括流感病毒、冠状病毒、脊髓灰质炎病毒、麻疹病毒、埃博拉病毒、反转录病毒或正粘病毒。204.根据实施方案189所述的组合物,其中所述病毒是冠状病毒。205.根据实施方案190所述的组合物,其中所述冠状病毒是SARS-COV-2病毒。206.根据实施方案196至205中任一项所述的组合物,其中所述组合物还包含多个引物,其中所述多个引物中的引物被配置成与所述合成核酸分子的序列或所述病原体核酸分子的序列杂交。207.根据实施方案206所述的组合物,其中所述合成核酸分子的序列和所述病原体核酸分子的所述序列是相同的。208.根据实施方案187至207中任一项所述的组合物,其中所述合成核酸以与所述病原体核酸分子相同的效率扩增。209.根据实施方案187至207中任一项所述的组合物,其中所述合成核酸分子被配置成产生大小与所述病原体核酸分子的扩增产物的大小相同或在其10个碱基对内的扩增产物。

[0191] 210.一种诊断病原体感染的个体的方法,所述方法包括:(a)提供来自所述个体的生物样品;(c)对所述生物样品进行聚合酶链反应(PCR),以获得PCR扩增的生物样品,其中对所述生物样品的所述PCR反应用第一PCR引物集进行,其中所述第一PCR引物集扩增病原体核酸序列;(d)使用下一代测序对所述PCR扩增的裂解的生物样品进行测序;以及(e)如果通过所述PCR或通过所述测序检测到病原体序列,则提供针对所述病原体感染的阳性诊断,或者如果通过所述PCR或通过所述测序未检测到病原体序列,则提供针对所述个体的阴性诊断。211.根据实施方案210所述的方法,其中所述个体是人类个体。212.根据实施方案210或211所述的方法,其中所述病原体感染包括细菌感染、病毒感染或真菌感染及其组合。213.根据实施方案210至212中任一项所述的方法,其中所述细菌感染是链球菌属、假单胞菌属、志贺氏菌属、弯曲菌属、沙门氏菌属、梭菌属或埃希氏菌属及其组合的感染。214.根据实施方案210至212中任一项所述的方法,其中所述感染是念珠菌属、芽生菌属、隐球菌属、球孢子菌属、组织胞浆菌属、副球孢子菌属、孢子丝菌属或肺囊虫属及其组合的感染。215.根据实施方案210至212中任一项所述的方法,其中所述病毒感染是DNA病毒的感染。216.根据实施方案215所述的方法,其中所述DNA病毒包括乙型肝炎病毒、丙型肝炎病毒、乳头瘤病

毒、E-B病毒、水痘病毒或天花病毒及其组合。217. 根据实施方案210至216中任一项所述的方法,其中所述病毒感染是RNA病毒的感染。218. 根据实施方案217所述的方法,其中所述RNA病毒包括流感病毒、冠状病毒、脊髓灰质炎病毒、麻疹病毒、埃博拉病毒、反转录病毒或正粘病毒。219. 根据实施方案218所述的方法,其中所述病毒感染是冠状病毒感染。220. 根据实施方案219所述的方法,其中所述冠状病毒感染是SARS-COV-2感染。221. 根据实施方案210至220中任一项所述的方法,其中来自所述个体的所述生物样品来自血液样品、血浆样品、血清样品、口腔拭子、尿液样品、精液样品、阴道拭子、粪便样品、鼻咽拭子、中鼻甲拭子或其任何组合。222. 根据实施方案221所述的方法,其中来自所述个体的所述生物样品来自鼻咽拭子、中鼻甲拭子或其任何组合。223. 根据实施方案210至222中任一项所述的方法,其包括向所述裂解剂或所述裂解的生物样品中添加合成核酸。224. 根据实施方案210至222中任一项所述的方法,其包括向所述方法中添加合成核酸。225. 根据实施方案224所述的方法,其中所述合成核酸是RNA。226. 根据实施方案224所述的方法,其中所述合成核酸是DNA。227. 根据实施方案224至226中任一项所述的方法,其中所述合成核酸包含被配置成由所述第一PCR引物集结合的序列集。228. 根据实施方案224至227中任一项所述的方法,其中所述第一引物集扩增所述病原体核酸序列和所述合成核酸。229. 根据实施方案224至228中任一项所述的方法,其中所述合成核酸序列包含与所述病原体核酸序列不相同的核苷酸序列。230. 根据实施方案210至229中任一项所述的方法,其包括对所述裂解的生物样品执行逆转录反应。231. 根据实施方案230所述的方法,其中所述逆转录反应在所述执行所述聚合酶链反应之前执行。232. 根据实施方案210至231中任一项所述的方法,其中所述逆转录反应在没有进一步纯化所述裂解的生物样品的情况下执行。233. 根据实施方案210至231中任一项所述的方法,其中对所述裂解的生物样品的所述逆转录反应产生病毒cDNA。234. 根据实施方案233所述的方法,其中所述病毒cDNA是冠状病毒cDNA。235. 根据实施方案233所述的方法,其中所述冠状病毒cDNA是SARS-COV-2 cDNA。236. 根据实施方案210至235中任一项所述的方法,其中所述逆转录反应和所述PCR是单步骤反应。237. 根据实施方案210至236中任一项所述的方法,其中所述PCR是终点分析。238. 根据实施方案210至237中任一项所述的方法,其中所述PCR不是实时PCR反应。239. 根据实施方案210至238中任一项所述的方法,其中所述第一PCR引物集扩增冠状病毒核酸序列。240. 根据实施方案239所述的方法,其中所述冠状病毒核酸序列是SARS-COV-2核酸序列。241. 根据实施方案240所述的方法,其中所述SARS-COV-2核酸序列包含N1或S2基因。242. 根据实施方案210至241中任一项所述的方法,其包含第二PCR引物集,其中所述第二引物集扩增所述个体的核酸序列。243. 根据实施方案242所述的方法,其中所述第二PCR引物集扩增人核酸序列。244. 根据实施方案243所述的方法,其中所述第二PCR引物集扩增选自以下的人核酸序列:GAPDH、ACTB、RPP30及其组合。245. 根据实施方案244所述的方法,其中所述第二PCR引物集扩增人RPP30。246. 根据实施方案242至245中任一项所述的方法,其中所述第二PCR引物集包含具有测序衔接子序列的引物和不具有测序衔接子序列的引物的混合物。247. 根据实施方案246所述的方法,其中具有测序衔接子序列的引物与不具有测序衔接子序列的引物的比率是约1:1、约1:2、约1:3或约1:4。248. 根据实施方案210至247中任一项所述的方法,其中所述PCR包括30至45个扩增循环。249. 根据实施方案248所述的方法,其中所述PCR包括35至45个扩增循环。250. 根据实施方案248所述的方法,其中所述PCR包括39至42个扩增循环。251. 根据实施方案210至250中

任一项所述的方法,其中所述第一PCR引物集、所述第二PCR引物集或所述第一PCR引物集和所述第二PCR引物集包含有包含可变核苷酸序列的核酸序列。252.根据实施方案251所述的方法,其中所述可变核苷酸序列是所述个体唯一的样品ID。253.根据实施方案210至252中任一项所述的方法,其中所述第一PCR引物集、所述第二PCR引物集或所述第一PCR引物集和所述第二PCR引物集包含用于下一代测序反应的衔接子序列。254.根据实施方案210至253中任一项所述的方法,其中所述方法可以检测少于10个病原体基因组拷贝。255.根据实施方案210至253中任一项所述的方法,其中所述方法可以检测少于5个病原体基因组拷贝。256.根据实施方案254或255所述的方法,其中所述病原体基因组是冠状病毒基因组。257.根据实施方案254或255所述的方法,其中所述冠状病毒基因组是SARS-COV-2基因组。258.根据实施方案210至257中任一项所述的方法,其中所述对冠状病毒呈阳性的诊断,如果通过所述PCR检测到冠状病毒序列,则是SARS-COV-2诊断。259.根据实施方案210至258中任一项所述的方法,其中所述方法确定冠状病毒的毒株。260.根据实施方案210至258中任一项所述的方法,其中所述方法确定COVID-19的毒株。261.一种组合物,其包含多个具有不同核酸序列的合成核酸,所述多个具有不同核酸序列的合成核酸包含与病原体核酸序列相同的共同5'序列、与病原体核酸序列相同的共同3'序列和在所述多个序列中不同的间插序列。262.根据实施方案261所述的组合物,其中所述多个具有不同核酸序列的合成核酸是单链的。263.根据实施方案261所述的组合物,其中所述多个具有不同核酸序列的合成核酸是双链的。264.根据实施方案261至263中任一项所述的组合物,其中所述多个具有不同核酸序列的合成核酸由RNA组成或包含RNA。265.根据实施方案261至263中任一项所述的组合物,其中具有不同核酸序列的所述多个合成核酸由DNA组成或包含DNA。266.根据实施方案261至265中任一项所述的组合物,其中与病原体核酸序列相同的所述共同5'序列是30个核苷酸或更少。267.根据实施方案261至266中任一项所述的组合物,其中与病原体核酸序列相同的所述共同3'序列是30个核苷酸或更少。268.根据实施方案261至267中任一项所述的组合物,其中所述间插序列是50个核苷酸或更少。269.根据实施方案268所述的组合物,其中所述间插序列30个核苷酸或更少。270.根据实施方案261至269中任一项所述的组合物,其中所述病原体核酸序列来自细菌病原体、真菌病原体或病毒病原体。271.根据实施方案270所述的组合物,其中所述细菌病原体是链球菌属、假单胞菌属、志贺氏菌属、弯曲菌属、沙门氏菌属、梭菌属或埃希氏菌属及其组合。272.根据实施方案270所述的组合物,其中所述真菌病原体是念珠菌属、芽生菌属、隐球菌属、球孢子菌属、组织胞浆菌属、副球孢子菌属、孢子丝菌属或肺囊虫属及其组合。273.根据实施方案270所述的组合物,其中所述病毒病原体是DNA病毒。274.根据实施方案273所述的组合物,其中所述DNA病毒包括乙型肝炎病毒、丙型肝炎病毒、乳头瘤病毒、E-B病毒、水痘病毒或天花病毒及其组合。275.根据实施方案1至3中任一项所述的组合物,其中所述病毒病原体是RNA病毒。276.根据实施方案270所述的组合物,其中所述RNA病毒包括流感病毒、冠状病毒、脊髓灰质炎病毒、麻疹病毒、埃博拉病毒、反转录病毒、正粘病毒或其组合。277.根据实施方案276所述的组合物,其中所述病毒病原体是冠状病毒。278.根据实施方案277所述的组合物,其中所述冠状病毒是SARS-COV-2。279.根据实施方案277或278所述的组合物,其中所述病原体核酸序列是编码冠状病毒刺突蛋白的核酸序列。280.根据实施方案277至279中任一项所述的组合物,其中具有不同核酸序列的所述多个合成核酸包含选自S2_001、S2_002、S2_003和S2_004中任一种或多种的序

列或由其组成。281. 根据实施方案261至280中任一项所述的组合物,其在诊断或检测病原体感染的方法中使用。282. 根据实施方案261至280中任一项所述的组合物,其在归一化病原体下一代序列读段的方法中使用。

实施例

[0192] 以下说明性实施例代表本文所述的组合物和方法的实施方案,并且不意味着以任何方式进行限制。

实施例1-病毒感染的SwabSeq诊断

[0193] 图1显示了用于收集和处理样品的示例工作流程。样品可以使用拭子(例如,唾液、鼻咽或中鼻甲)从人类受试者获得,如102中所例示。可将拭子直接放置于足以裂解样品内的细胞的条件下,如106中所例示。不需要RNA分离,因此(1)减少了进行测定所需的时间,(2)提高了灵敏度,并且(3)降低了测定材料的成本。样品可以可选地储存在缓冲液(例如,TE缓冲液)中,并且稍后在获得样品的数天内进行裂解,如步骤126,128中。为了便于准确并且灵敏地鉴定病毒核酸,可以在裂解的样品中加标合成核酸,从而可以对后续核酸处理和所得序列信息进行基准测试(108中所例示)。合成核酸包含能够通过病毒特异性引物扩增的序列,但具有可以在序列输出中容易地鉴定为与病毒靶标不同的已知序列。将样品在微量滴定板中分区和处理,其实现了在1与32个384孔板之间的缩放,从而使得能够测试至少384-12,288个样品。

[0194] 然后在分区/孔内,通过逆转录和PCR进行核酸处理和文库制备,如110中所例示。用于检测的引物包含与以下互补的序列:(a)感兴趣的靶标(例如来自冠状病毒或流感病毒的小区);(b)人靶标(确保有一定人序列的对照物;例如,RPP30(RNA酶P),CDC在其基于标准qPCR的测试中使用它);和(c)内部合成RNA分子(如上文所说明),确保靶标测定放大并且定量靶标,例如允许对来自病原体序列的读段进行归一化。使通过PCR的扩增进行至终点,然后进行测序。端点PCR和合成模板添加了归一化水平,以避免昂贵的归一化步骤。用于靶标扩增的引物的设计允许添加允许样品去卷积的索引以及鉴定哪个样品包含病毒序列的阳性测序读段。

[0195] 随后对样品进行测序以鉴定条形码化序列和病毒靶标读段的存在。对读段进行生物信息处理,以衡量对病毒载体和病毒负载呈阳性或阴性的测序数据。有可能进一步开发对多态性区进行测序并确定病原体菌株的测定。这可以通过比较靶病原体扩增子与人靶扩增子与扩增子中的标的比率来实现。

实施例2-样品中SARS-CoV RNA的检测

[0196] 使用本文所公开的方法分析了包含SARS-CoV-2 RNA的样品。样品包含确定拷贝数的SARS-CoV-2 RNA分子。将靶向N1区(核衣壳)和亚基S2区的引物用于检测SARS-CoV-2 RNA分子。将所得测序产物与人RRP30 RNA进行比较。分析每个靶区206,216的随拷贝数变化的SARS-CoV:RRP30(图2A-B)。令人惊讶的是,在包含3个SARS-CoV-2 RNA分子拷贝的样品内,靶向亚基S2得到可区分检测。靶向N1在是S2的2-5倍的拷贝数下显示可区分检测,但灵敏度可接受。图2B展示了对与N1基因座和S2基因座互补的不同引物序列210的分析。检测依赖于N1和S2靶标的引物序列。在测试的所有引物中,靶向S2在是N1的1/2-1/5的拷贝数下显示可区分检测,表明对于样品中SARS-CoV RNA的检测,靶向S2提供了更高分辨率的分析。

实施例3-通过扩增和测序的病毒诊断

[0197] 本文所述的与本文所述的方法一起使用的扩增和测序方案的一个实施例。在添加QRT-PCR反应之前,可以加热裂解缓冲液中的样品(例如,75°C持续10分钟)。

[0198] QRT-PCR反应设置,单独的反应在下列中运行:20 μ L总体积;7 μ L样品裂解物;10 μ L Luna[®]通用一步反应混合物;1 μ L酶混合物;400nM病毒核酸靶标引物;持家引物(例如,RPP3 100nM无衔接子,50nM有衔接子);100个合成核酸拷贝(与病毒靶标相同的启动区)。通过用测序衔接子降低RPP3引物的浓度,同时保持总体RPP3引物浓度高,仍然检测到少量RPP3,同时限制专用于宿主RNA的测序读段的数目。此外,这也减弱了引物二聚体形成。

[0199] 循环方案55°C持续30min(逆转录反应);95°C持续1min,然后是40次循环(95°C持续10秒和60°C持续30秒)。

[0200] 每个孔汇集5 μ L以便使用[AxyPrep PCR纯化试剂盒(Clean-up Kit)](www.fishersci.com/shop/products/axygen-axyprep-mag-pcr-clean-up-kits/14223152)进行测序和纯化。用[deNovix dsDNA高灵敏度荧光测定试剂盒](www.denovix.com/denovix-dsdna-assays/)定量文库。可添加至多50% PhiX对照文库DNA。在NextSeq和MiSeq上使用Illumina双重索引单读测序对库进行测序。

[0201] 下表(表1)描述了可用于本文所述的检测病毒感染的方法中的引物对。引物公开为例如o325_octN1_F和o326_octN1_R,描述了被设计成产生病毒核酸的扩增的引物对。下文所述的序列也可以包含测序衔接子和/或索引序列。

表1

| SEQUENCE ID NO. | 名称 | seq |
|-----------------|---|------------------------------|
| SEQ ID NO: 100 | o258_o258 SARS-CoV-2 N1 luc 索引 1 引物 | tggggtcTTACACGGCGATCTTGCC |
| SEQ ID NO: 101 | o259_o259 SARS-CoV-2 N1 oct 索引 1 引物 | gggtcAACGTGTCGGCATGGATTCT |
| SEQ ID NO: 102 | o260_o260 SARS-CoV-2 N1 luc | AGTTACATTCACGCCAGTTGTGtctggt |

| | | | |
|----------------|----|---|--|
| | | 读段 1 | |
| SEQ NO: 103 | ID | o261_o261 SARS-CoV-2 N2 luc 读段 1 | AGTTACATTCACGCCAGTT GTGgcg |
| SEQ NO: 104 | ID | o262_o262 SARS-CoV-2 N2 luc 索引 1 引物 | gtttgtaaTTACACGGCGATCTT GCC |
| SEQ NO: 105 | ID | o263_o263 RP luc 索 引 1 引物 | GTCCAAATCTTTACACGGC GATCTTGCC |
| SEQ NO: 106 | ID | o264_o264 RP luc 读 段 1 | AGTTACATTCACGCCAGTT GTGGAGC |
| SEQ NO: 107 | ID | o265_o265 T7prom-SARS-CoV-2 N | TAATACGACTCACTATAGgg tctgataatggaccccaaatca |
| SEQ NO: 108 | ID | o266_o266 SARS-CoV-2 N R | ttaggcctgagttgagtcagc |
| SEQ NO: 109 | ID | o267_索引 A01 N1_F | CAAGCAGAAGACGGCATA CGAGAT GTTCTATC GACCCCAAATCAGCGAAA T |
| SEQ NO: 110 | ID | o268_索引板 1 N1_R | AATGATACGGCGACCACCG AGATCTACAC AAGATCTG TCTGGTTACTGCCAGTTGA ATCTG |
| SEQ NO: 111 | ID | o269_索引 A01 N2_F | CAAGCAGAAGACGGCATA CGAGAT GTTCTATC TTACAAACATTGGCCGCAA A |
| SEQ NO: 112 | ID | o270_索引板 1 N2_R | AATGATACGGCGACCACCG AGATCTACAC AAGATCTG GCGCGACATTCCGAAGAA |
| SEQ NO: 113 | ID | o271_索引 A01 RP_F | CAAGCAGAAGACGGCATA CGAGAT GTTCTATC AGATTTGGACCTGCGAGCG |
| SEQ NO: 114 | ID | o272_索引板 1 RP_R | AATGATACGGCGACCACCG AGATCTACAC AAGATCTG GAGCGGCTGTCTCCACAAG T |
| SEQ | ID | o273_索引 A01 | CAAGCAGAAGACGGCATA |

| | | |
|-------------------|------------------------|---|
| NO: 115 | HKU_N_F | CGAGAT GTTCTATC TAATCAGACAAGGAACTGA TTA |
| SEQ ID NO: 116 | o274_索引板 1 HKU_N_R | AATGATACGGCGACCACCG AGATCTACAC AAGATCTG CGAAGGTGTGACTTCCATG |
| SEQ ID NO: 117 | o275_索引 A01 octN1_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTGGATCAAAACAA CGTCGGCCC |
| SEQ ID NO: 118 | o276_索引板 1 octN1_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATCCCATGTTGAGT GAGAGCGGT |
| SEQ ID NO: 119 | o277_索引 A01 octN2_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTGTGGACCCCAA ATCAGCGAA |
| SEQ ID NO: 120 | o278_索引板 1 octN2_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATCACTGCGTTCTCC ATTCTGGTT |
| SEQ ID NO: 121 | o279_索引 A01 octN3_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTGCAGCGTTCTTC GGAATGTCG |
| SEQ ID NO: 122 | o280_索引板 1 octN3_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATCGCACCTGTGTA GGTCAACCA |
| SEQ ID NO: 123 | o281_索引 A01 octN4_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTGGAAATGCACCC CGCATTACG |
| SEQ ID NO: 124 | o282_索引板 1 octN4_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATCCCCACTGCGTTC TCCATTCT |
| SEQ ID NO: 125 | o283_索引 A01 octN5_F | AATGATACGGCGACCACCG AGATCTACAC |

| | | |
|-------------------|-------------------------|--|
| | | AAGATCTGGTCTTGGTTCA CCGCTCTCA |
| SEQ ID NO: 126 | o284_索引板 1 octN5_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATCTTGGAACGCCTT GTCCTCG |
| SEQ ID NO: 127 | o285_索引 A01 octN6_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTGGCAGTCAAGCC TCTTCTCGT |
| SEQ ID NO: 128 | o286_索引板 1 octN6_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATCGAAGTTCCCCT ACTGCTGCC |
| SEQ ID NO: 129 | o287_索引 A01 octN7_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTGCGTTTGGTGGGA CCCTCAGAT |
| SEQ ID NO: 130 | o288_索引板 1 octN7_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATCGACGTTGTTTTG ATCGCGCC |
| SEQ ID NO: 131 | o289_索引 A01 octN8_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTGAAGGCCAACAA CAACAAGGC |
| SEQ ID NO: 132 | o290_索引板 1 octN8_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATCGGCAGTACGTT TTTGCCGAG |
| SEQ ID NO: 133 | o291_索引 A01 octN9_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTGACCAGAATGGA GAACGCAGT |
| SEQ ID NO: 134 | o292_索引板 1 octN9_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATCCGGTGAACCAA GACGCAGTA |
| SEQ ID NO: 135 | o293_索引 A01 octN10_F | AATGATACGGCGACCACCG AGATCTACAC |

| | | |
|-------------------|---------------------------------|--|
| | | AAGATCTGCCGCATTACGT TTGGTGGAC |
| SEQ ID NO: 136 | o294_索引板 1 octN10_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATCGGCCGACGTTG TTTTGATCG |
| SEQ ID NO: 137 | o295_索引 A01 octN11_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTGGCCTCGGCAAA AACGTA CTG |
| SEQ ID NO: 138 | o296_索引板 1 octN11_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATCTTGTTCTGGACC ACGTCTGC |
| SEQ ID NO: 139 | o297_索引 A01 octN12_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTGAATTCCTCGA GGACAAGGC |
| SEQ ID NO: 140 | o298_索引板 1 octN12_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATCTCGTCTGGTAGC TCTTCGGT |
| SEQ ID NO: 141 | o299_索引 A01 octN13_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTGGCTTCAGCGTT CTTCGGAAT |
| SEQ ID NO: 142 | o300_索引板 1 octN13_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATCTGGCACCTGTGT AGGTCAAC |
| SEQ ID NO: 143 | o301_SARS-CoV-2_I BS RdRP2_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTG AGAATAGAGCTCGCACCGT A |
| SEQ ID NO: 144 | o302_SARS-CoV-2_I BS RdRP2_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATC CTCCTCTAGTGGCGGCTAT T |
| SEQ ID NO: 145 | o303_SARS-CoV-2_I BS_S2_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTG |

| | | |
|-------------------|---------------------------------|--|
| | | GCTGGTGCTGCAGCTTATT A |
| SEQ ID NO: 146 | o304_SARS-CoV-2_I BS_S2_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATC AGGGTCAAGTGCACAGTCT A |
| SEQ ID NO: 147 | o305_SARS-CoV-2_I BS_E2_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTG TTCGGAAGAGACAGGTACG TTA |
| SEQ ID NO: 148 | o306_SARS-CoV-2_I BS_E2_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATC AGCAGTACGCACACAATCG |
| SEQ ID NO: 149 | o307_SARS-CoV-2_I BS_N1_F | AATGATACGGCGACCACCG AGATCTACAC AAGATCTG CAATGCTGCAATCGTGCTA C |
| SEQ ID NO: 150 | o308_SARS-CoV-2_I BS_N1_R | CAAGCAGAAGACGGCATA CGAGAT GTTCTATC GTTGCGACTACGTGATGAG G |
| SEQ ID NO: 151 | o309_RNA 酶 P 正向 引物 | AGATTTGGACCTGCGAGCG |
| SEQ ID NO: 152 | o310_RNA 酶 P 反向 引物 | GAGCGGCTGTCTCCACAAG T |
| SEQ ID NO: 153 | o311_SARS-CoV-2_I BS_RdRP2_F | AGAATAGAGCTCGCACCGT A |
| SEQ ID NO: 154 | o312_SARS-CoV-2_I BS_RdRP2_R | CTCCTCTAGTGGCGGCTAT T |
| SEQ ID NO: 155 | o313_SARS-CoV-2_I BS_S2_F | GCTGGTGCTGCAGCTTATT A |
| SEQ ID NO: 156 | o314_SARS-CoV-2_I BS_S2_R | AGGGTCAAGTGCACAGTCT A |
| SEQ ID NO: 157 | o315_SARS-CoV-2_I BS_E2_F | TTCGGAAGAGACAGGTACG TTA |
| SEQ ID NO: 158 | o316_SARS-CoV-2_I BS_E2_R | AGCAGTACGCACACAATCG |
| SEQ ID NO: 159 | o317_SARS-CoV-2_I BS_N1_F | CAATGCTGCAATCGTGCTA C |
| SEQ ID | o318_SARS-CoV-2_I | GTTGCGACTACGTGATGAG |

| | | |
|-------------------|---------------------------|--------------------------|
| NO: 160 | BS_N1_R | G |
| SEQ ID NO: 161 | o319_2019-nCoV_N1 正向引物 | GACCCCAAATCAGCGAAAT |
| SEQ ID NO: 162 | o320_2019-nCoV_N1 反向引物 | TCTGGTTACTGCCAGTTGATCTG |
| SEQ ID NO: 163 | o321_2019-nCoV_N2 正向引物 | TTACAAACATTGGCCGCAA A |
| SEQ ID NO: 164 | o322_2019-nCoV_N2 反向引物 | GCGCGACATTCCGAAGAA |
| SEQ ID NO: 165 | o323_HKU-NF | TAATCAGACAAGGAACTGATTA |
| SEQ ID NO: 166 | o324_HKU-NR | CGAAGGTGTGACTTCCATG |
| SEQ ID NO: 167 | o325_octN1_F | GATCAAACAACGTCGGCC C |
| SEQ ID NO: 168 | o326_octN1_R | CCATGTTGAGTGAGAGCGGT |
| SEQ ID NO: 169 | o327_octN2_F | TGGACCCCAAATCAGCGAA |
| SEQ ID NO: 170 | o328_octN2_R | ACTGCGTTCTCCATTCTGGTT |
| SEQ ID NO: 171 | o329_octN3_F | CAGCGTTCTTCGGAATGTC G |
| SEQ ID NO: 172 | o330_octN3_R | GCACCTGTGTAGGTCAACCA |
| SEQ ID NO: 173 | o331_octN4_F | GAAATGCACCCCGCATTAC G |
| SEQ ID NO: 174 | o332_octN4_R | CCCCTGCGTTCTCCATTCT |
| SEQ ID NO: 175 | o333_octN5_F | GTCTTGGTTCACCGCTCTCA |
| SEQ ID NO: 176 | o334_octN5_R | TTGGAACGCCTTGTCCCTCG |
| SEQ ID NO: 177 | o335_octN6_F | GCAGTCAAGCCTCTTCTCG T |
| SEQ ID NO: 178 | o336_octN6_R | GAAGTTCCCCTACTGCTGC C |
| SEQ ID NO: 179 | o337_octN7_F | CGTTTGGTGGACCCTCAGAT |

| | | |
|-------------------|---------------------------|---|
| SEQ ID NO: 180 | o338_octN7_R | GACGTTGTTTTGATCGCGC C |
| SEQ ID NO: 181 | o339_octN8_F | AAGGCCAACAACAACAAG GC |
| SEQ ID NO: 182 | o340_octN8_R | GGCAGTACGTTTTTGCCGA G |
| SEQ ID NO: 183 | o341_octN9_F | ACCAGAATGGAGAACGCA GT |
| SEQ ID NO: 184 | o342_octN9_R | CGGTGAACCAAGACGCAGT A |
| SEQ ID NO: 185 | o343_octN10_F | CCGCATTACGTTTGGTGGA C |
| SEQ ID NO: 186 | o344_octN10_R | GGCCGACGTTGTTTTGATC G |
| SEQ ID NO: 187 | o345_octN11_F | GCCTCGGCAAAAACGTACT G |
| SEQ ID NO: 188 | o346_octN11_R | TTGTTCTGGACCACGTCTG C |
| SEQ ID NO: 189 | o347_octN12_F | AATTCCTCGAGGACAAGG C |
| SEQ ID NO: 190 | o348_octN12_R | TCGTCTGGTAGCTCTTCGGT |
| SEQ ID NO: 191 | o349_octN13_F | GCTTCAGCGTTCTTCGGAA T |
| SEQ ID NO: 192 | o350_octN13_R | TGGCACCTGTGTAGGTCAA C |
| SEQ ID NO: 193 | o351_skpp15-1-F_RPP 30 | GGGTCACGCGTAGGA GTTCTATC AGATTTGGACCTGCGAGCG |
| SEQ ID NO: 194 | o352_skpp15-1-R_RP P30 | GTTCCGCAGCCACAC AAGATCTG GAGCGGCTGTCTCCACAAG T |
| SEQ ID NO: 195 | o353_skpp15-2-F_RPP 30 | CGCGTCGAGTAGGGT GTTCTATC AGATTTGGACCTGCGAGCG |
| SEQ ID NO: 196 | o354_skpp15-2-R_RP P30 | GCCGTGTGAAGCTGG AAGATCTG GAGCGGCTGTCTCCACAAG T |

| | | |
|----------------|---------------------------------|---|
| SEQ NO: 197 | ID o355_skpp15-3-F_RPP 30 | CGATCGCCCTTGGTG GTTCTATC AGATTTGGACCTGCGAGCG |
| SEQ NO: 198 | ID o356_skpp15-3-R_RP P30 | GGTTTAGCCGGCGTG AAGATCTG GAGCGGCTGTCTCCACAAG T |
| SEQ NO: 199 | ID o357_skpp15-4-F_RPP 30 | GGTCGAGCCGGA ACT GTTCTATC AGATTTGGACCTGCGAGCG |
| SEQ NO: 200 | ID o358_skpp15-4-R_RP P30 | GGATGCGCACCCAGA AAGATCTG GAGCGGCTGTCTCCACAAG T |
| SEQ NO: 201 | ID o359_skpp15-5-F_RPP 30 | TCCCGGCGTTGTCCT GTTCTATC AGATTTGGACCTGCGAGCG |
| SEQ NO: 202 | ID o360_skpp15-5-R_RP P30 | GCTCCGTCACTGCC AAGATCTG GAGCGGCTGTCTCCACAAG T |
| SEQ NO: 203 | ID o361_skpp15-6-F_RPP 30 | CGCAGGGTCCAGAGT GTTCTATC AGATTTGGACCTGCGAGCG |
| SEQ NO: 204 | ID o362_skpp15-6-R_RP P30 | GTTCGCGCGAAGGAA AAGATCTG GAGCGGCTGTCTCCACAAG T |
| SEQ NO: 205 | ID o363_skpp-1-F s_RPP30 | ATATAGATGCCGTCCTAGC G GTTCTATC AGATTTGGACCTGCGAGCG |
| SEQ NO: 206 | ID o364_skpp-1-R s_RPP30 | AAGTATCTTTCCTGTGCC A AAGATCTG GAGCGGCTGTCTCCACAAG T |
| SEQ NO: 207 | ID o365_skpp-2-F s_RPP30 | CCCTTTAATCAGATGCGTC G GTTCTATC AGATTTGGACCTGCGAGCG |
| SEQ NO: 208 | ID o366_skpp-2-R s_RPP30 | TGGTAGTAATAAGGGCGAC C AAGATCTG |

| | | |
|-------------------|------------------------------|---|
| | | GAGCGGCTGTCTCCACAAG T |
| SEQ ID NO: 209 | o367_skpp-3-F s_RPP30 | TTGGTCATGTGCTTTTCGTT GTTCTATC AGATTTGGACCTGCGAGCG |
| SEQ ID NO: 210 | o368_skpp-3-R s_RPP30 | AGGGGTATCGGATACTCAG A AAGATCTG GAGCGGCTGTCTCCACAAG T |
| SEQ ID NO: 211 | o369_skpp-4-F s_RPP30 | GGGTGGGTAAATGGTAATG C GTTCTATC AGATTTGGACCTGCGAGCG |
| SEQ ID NO: 212 | o370_skpp-4-R s_RPP30 | ATCGATTCCCCGGATATAG C AAGATCTG GAGCGGCTGTCTCCACAAG T |
| SEQ ID NO: 213 | o371_skpp-5-F s_RPP30 | TCCGACGGGGAGTATATAC T GTTCTATC AGATTTGGACCTGCGAGCG |
| SEQ ID NO: 214 | o372_skpp-5-R s_RPP30 | TACTAACTGCTTCAGGCCA A AAGATCTG GAGCGGCTGTCTCCACAAG T |
| SEQ ID NO: 215 | o373_skpp-6-F s_RPP30 | CATGTTTAGGAACGCTACC G GTTCTATC AGATTTGGACCTGCGAGCG |
| SEQ ID NO: 216 | o374_skpp-6-R s_RPP30 | AATAATCTCCGTTCCCTCCC AAGATCTG GAGCGGCTGTCTCCACAAG T |
| SEQ ID NO: 217 | o375_T7prom o267_o268 标 F | TAATACGACTCACTATAGgg accccaaatcagcgaatgcaccccgcat tacgAAACCAggaccctcagattcaac tg |
| SEQ ID NO: 218 | o376_T7prom o267_o268 标 R | cgcagtattattgggtaaacct |
| SEQ ID NO: 219 | o377_T7prom o303_o304 标 F | TAATACGACTCACTATAGgg ctggtgctgcagcttattatgtgggATA GAACAacctaggacttttctattaa |
| SEQ ID | o378_T7prom | aacgtacactttgtttctgagagagg |

| | | |
|-------------------|---------------|---|
| NO: 220 | o303_o304 标 R | |
| SEQ ID NO: 221 | o379_A1_N1_F | CAAGCAGAAGACGGCATA CGAGATGAGTCTTCGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 222 | o380_A2_N1_F | CAAGCAGAAGACGGCATA CGAGATGTTCTATCGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 223 | o381_A3_N1_F | CAAGCAGAAGACGGCATA CGAGATTGGGCCAAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 224 | o382_A4_N1_F | CAAGCAGAAGACGGCATA CGAGATATTGTTGGGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 225 | o383_A5_N1_F | CAAGCAGAAGACGGCATA CGAGATTCCTGTTGGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 226 | o384_A6_N1_F | CAAGCAGAAGACGGCATA CGAGATACACACTTGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 227 | o385_A7_N1_F | CAAGCAGAAGACGGCATA CGAGATCCATTCCAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 228 | o386_A8_N1_F | CAAGCAGAAGACGGCATA CGAGATCTAACGGGGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 229 | o387_A9_N1_F | CAAGCAGAAGACGGCATA CGAGATCCATAGGAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 230 | o388_A10_N1_F | CAAGCAGAAGACGGCATA CGAGATCAGTGTAGGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 231 | o389_A11_N1_F | CAAGCAGAAGACGGCATA CGAGATGATTCTCAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 232 | o390_A12_N1_F | CAAGCAGAAGACGGCATA CGAGATCCTTCTTAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 233 | o391_B1_N1_F | CAAGCAGAAGACGGCATA CGAGATTCTAAGACGACCC CAAATCAGCGAAAT |

| | | |
|-------------------|---------------|---|
| SEQ ID NO: 234 | o392_B2_N1_F | CAAGCAGAAGACGGCATA CGAGATGCGGCATAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 235 | o393_B3_N1_F | CAAGCAGAAGACGGCATA CGAGATATTGACGAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 236 | o394_B4_N1_F | CAAGCAGAAGACGGCATA CGAGATGCTCCTGAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 237 | o395_B5_N1_F | CAAGCAGAAGACGGCATA CGAGATGCAATCCTGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 238 | o396_B6_N1_F | CAAGCAGAAGACGGCATA CGAGATATGTCGTTGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 239 | o397_B7_N1_F | CAAGCAGAAGACGGCATA CGAGATTTCTCGGCGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 240 | o398_B8_N1_F | CAAGCAGAAGACGGCATA CGAGATCAGGGCTAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 241 | o399_B9_N1_F | CAAGCAGAAGACGGCATA CGAGATAGCCAAGCGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 242 | o400_B10_N1_F | CAAGCAGAAGACGGCATA CGAGATAAGCCTGAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 243 | o401_B11_N1_F | CAAGCAGAAGACGGCATA CGAGATCTACAGAGGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 244 | o402_B12_N1_F | CAAGCAGAAGACGGCATA CGAGATCGTAGTCGGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 245 | o403_C1_N1_F | CAAGCAGAAGACGGCATA CGAGATTTCTGCTCGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 246 | o404_C2_N1_F | CAAGCAGAAGACGGCATA CGAGATGTGCACACGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 247 | o405_C3_N1_F | CAAGCAGAAGACGGCATA |

| | | |
|-------------------|---------------|--|
| NO: 247 | | CGAGATAAAGCTCAGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 248 | o406_C4_N1_F | CAAGCAGAAGACGGCATA CGAGATGACCTCAGGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 249 | o407_C5_N1_F | CAAGCAGAAGACGGCATA CGAGATCTTTCCAAGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 250 | o408_C6_N1_F | CAAGCAGAAGACGGCATA CGAGATTCTTGGCTGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 251 | o409_C7_N1_F | CAAGCAGAAGACGGCATA CGAGATCGCGTCTAGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 252 | o410_C8_N1_F | CAAGCAGAAGACGGCATA CGAGATTCGCGCTAGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 253 | o411_C9_N1_F | CAAGCAGAAGACGGCATA CGAGATATCCATTCGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 254 | o412_C10_N1_F | CAAGCAGAAGACGGCATA CGAGATGCCAGTAGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 255 | o413_C11_N1_F | CAAGCAGAAGACGGCATA CGAGATTACCGACGGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 256 | o414_C12_N1_F | CAAGCAGAAGACGGCATA CGAGATTCCATACGGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 257 | o415_D1_N1_F | CAAGCAGAAGACGGCATA CGAGATAACATGTCGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 258 | o416_D2_N1_F | CAAGCAGAAGACGGCATA CGAGATCGACTATAGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 259 | o417_D3_N1_F | CAAGCAGAAGACGGCATA CGAGATACCCAAGGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 260 | o418_D4_N1_F | CAAGCAGAAGACGGCATA CGAGATATCGATCGGACCC |

| | | |
|-------------------|---------------|---|
| | | CAAAATCAGCGAAAT |
| SEQ ID NO: 261 | o419_D5_N1_F | CAAGCAGAAGACGGCATA CGAGATGTTGGATGGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 262 | o420_D6_N1_F | CAAGCAGAAGACGGCATA CGAGATCTATGTGAGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 263 | o421_D7_N1_F | CAAGCAGAAGACGGCATA CGAGATTATTTTCGCGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 264 | o422_D8_N1_F | CAAGCAGAAGACGGCATA CGAGATCCATGTATGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 265 | o423_D9_N1_F | CAAGCAGAAGACGGCATA CGAGATGCCACGTTGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 266 | o424_D10_N1_F | CAAGCAGAAGACGGCATA CGAGATGTCGTGTAGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 267 | o425_D11_N1_F | CAAGCAGAAGACGGCATA CGAGATTAAAGTCGGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 268 | o426_D12_N1_F | CAAGCAGAAGACGGCATA CGAGATCTTCGGACGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 269 | o427_E1_N1_F | CAAGCAGAAGACGGCATA CGAGATGCACTCTCGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 270 | o428_E2_N1_F | CAAGCAGAAGACGGCATA CGAGATTCAGATACGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 271 | o429_E3_N1_F | CAAGCAGAAGACGGCATA CGAGATCAGTCCCTGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 272 | o430_E4_N1_F | CAAGCAGAAGACGGCATA CGAGATGCCCTAACGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 273 | o431_E5_N1_F | CAAGCAGAAGACGGCATA CGAGATCTGCATCAGACCC CAAAATCAGCGAAAT |

| | | |
|-------------------|---------------|---|
| SEQ ID NO: 274 | o432_E6_N1_F | CAAGCAGAAGACGGCATA CGAGATCGGTATCGGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 275 | o433_E7_N1_F | CAAGCAGAAGACGGCATA CGAGATAAGTATGGGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 276 | o434_E8_N1_F | CAAGCAGAAGACGGCATA CGAGATATTCGCGCGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 277 | o435_E9_N1_F | CAAGCAGAAGACGGCATA CGAGATATCAAGGTGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 278 | o436_E10_N1_F | CAAGCAGAAGACGGCATA CGAGATTTGTGCATGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 279 | o437_E11_N1_F | CAAGCAGAAGACGGCATA CGAGATCTGTGCTGGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 280 | o438_E12_N1_F | CAAGCAGAAGACGGCATA CGAGATGTCCGTAGGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 281 | o439_F1_N1_F | CAAGCAGAAGACGGCATA CGAGATGTTCAAGAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 282 | o440_F2_N1_F | CAAGCAGAAGACGGCATA CGAGATCACCGTTCGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 283 | o441_F3_N1_F | CAAGCAGAAGACGGCATA CGAGATCGAGTTGAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 284 | o442_F4_N1_F | CAAGCAGAAGACGGCATA CGAGATGAGCACGAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 285 | o443_F5_N1_F | CAAGCAGAAGACGGCATA CGAGATAGTTCGTGGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 286 | o444_F6_N1_F | CAAGCAGAAGACGGCATA CGAGATCATCACTGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 287 | o445_F7_N1_F | CAAGCAGAAGACGGCATA |

| | | |
|-------------------|---------------|---|
| NO: 287 | | CGAGATCGAGATCTGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 288 | o446_F8_N1_F | CAAGCAGAAGACGGCATA CGAGATTGGCCAGAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 289 | o447_F9_N1_F | CAAGCAGAAGACGGCATA CGAGATTTACCATGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 290 | o448_F10_N1_F | CAAGCAGAAGACGGCATA CGAGATGAATGCATGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 291 | o449_F11_N1_F | CAAGCAGAAGACGGCATA CGAGATTGGACCCTGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 292 | o450_F12_N1_F | CAAGCAGAAGACGGCATA CGAGATGATAGCACGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 293 | o451_G1_N1_F | CAAGCAGAAGACGGCATA CGAGATACGACGACGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 294 | o452_G2_N1_F | CAAGCAGAAGACGGCATA CGAGATCTCAGTATGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 295 | o453_G3_N1_F | CAAGCAGAAGACGGCATA CGAGATCTTAGCTAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 296 | o454_G4_N1_F | CAAGCAGAAGACGGCATA CGAGATCTGTTTACGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 297 | o455_G5_N1_F | CAAGCAGAAGACGGCATA CGAGATTGTCCCACGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 298 | o456_G6_N1_F | CAAGCAGAAGACGGCATA CGAGATTCCTGAGGGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 299 | o457_G7_N1_F | CAAGCAGAAGACGGCATA CGAGATTAGTTCCAGACCC CAAATCAGCGAAAT |
| SEQ ID NO: 300 | o458_G8_N1_F | CAAGCAGAAGACGGCATA CGAGATCATGACTCGACCC |

| | | |
|-------------------|---------------|--|
| | | CAAAATCAGCGAAAT |
| SEQ ID NO: 301 | o459_G9_N1_F | CAAGCAGAAGACGGCATA CGAGATGTAAGCGCGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 302 | o460_G10_N1_F | CAAGCAGAAGACGGCATA CGAGATAACCCAGTGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 303 | o461_G11_N1_F | CAAGCAGAAGACGGCATA CGAGATTTTGAGGGGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 304 | o462_G12_N1_F | CAAGCAGAAGACGGCATA CGAGATAGCCGACAGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 305 | o463_H1_N1_F | CAAGCAGAAGACGGCATA CGAGATAAACCCGCGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 306 | o464_H2_N1_F | CAAGCAGAAGACGGCATA CGAGATGTAGGGCTGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 307 | o465_H3_N1_F | CAAGCAGAAGACGGCATA CGAGATAGACGATTGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 308 | o466_H4_N1_F | CAAGCAGAAGACGGCATA CGAGATAGGATGATGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 309 | o467_H5_N1_F | CAAGCAGAAGACGGCATA CGAGATATAATGGCGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 310 | o468_H6_N1_F | CAAGCAGAAGACGGCATA CGAGATCTTGGCGTGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 311 | o469_H7_N1_F | CAAGCAGAAGACGGCATA CGAGATAGCTGTGCGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 312 | o470_H8_N1_F | CAAGCAGAAGACGGCATA CGAGATGAGTCCAAGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 313 | o471_H9_N1_F | CAAGCAGAAGACGGCATA CGAGATGAATACCAGACCC CAAAATCAGCGAAAT |

| | | |
|-------------------|---------------|---|
| SEQ ID NO: 314 | o472_H10_N1_F | CAAGCAGAAGACGGCATA CGAGATAGGAGCTTGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 315 | o473_H11_N1_F | CAAGCAGAAGACGGCATA CGAGATGTGACTTAGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 316 | o474_H12_N1_F | CAAGCAGAAGACGGCATA CGAGATTTTGG AACGACCC CAAAATCAGCGAAAT |
| SEQ ID NO: 317 | o475_A1_S2_R | CAAGCAGAAGACGGCATA CGAGATGAGTCTTCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 318 | o476_A2_S2_R | CAAGCAGAAGACGGCATA CGAGATGTTCTATCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 319 | o477_A3_S2_R | CAAGCAGAAGACGGCATA CGAGATTGGGCCAAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 320 | o478_A4_S2_R | CAAGCAGAAGACGGCATA CGAGATATTGTTGGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 321 | o479_A5_S2_R | CAAGCAGAAGACGGCATA CGAGATTCCCGTTGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 322 | o480_A6_S2_R | CAAGCAGAAGACGGCATA CGAGATACACACTTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 323 | o481_A7_S2_R | CAAGCAGAAGACGGCATA CGAGATCCATTCCAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 324 | o482_A8_S2_R | CAAGCAGAAGACGGCATA CGAGATCTAACGGGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 325 | o483_A9_S2_R | CAAGCAGAAGACGGCATA CGAGATCCATAGGAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 326 | o484_A10_S2_R | CAAGCAGAAGACGGCATA CGAGATCAGTGTAGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 326 | o485_A11_S2_R | CAAGCAGAAGACGGCATA |

| | | |
|-------------------|---------------|---|
| NO: 327 | | CGAGATGATTCTCAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 328 | o486_A12_S2_R | CAAGCAGAAGACGGCATA CGAGATCCTTCTTAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 329 | o487_B1_S2_R | CAAGCAGAAGACGGCATA CGAGATTCTAAGACAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 330 | o488_B2_S2_R | CAAGCAGAAGACGGCATA CGAGATGCGGCATAAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 331 | o489_B3_S2_R | CAAGCAGAAGACGGCATA CGAGATATTGACGAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 332 | o490_B4_S2_R | CAAGCAGAAGACGGCATA CGAGATGCTCCTGAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 333 | o491_B5_S2_R | CAAGCAGAAGACGGCATA CGAGATGCAATCCTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 334 | o492_B6_S2_R | CAAGCAGAAGACGGCATA CGAGATATGTTCGTTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 335 | o493_B7_S2_R | CAAGCAGAAGACGGCATA CGAGATTTCTCGGCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 336 | o494_B8_S2_R | CAAGCAGAAGACGGCATA CGAGATCAGGGCTAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 337 | o495_B9_S2_R | CAAGCAGAAGACGGCATA CGAGATAGCCAAGCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 338 | o496_B10_S2_R | CAAGCAGAAGACGGCATA CGAGATAAGCCTGAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 339 | o497_B11_S2_R | CAAGCAGAAGACGGCATA CGAGATCTACAGAGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 340 | o498_B12_S2_R | CAAGCAGAAGACGGCATA CGAGATCGTAGTCGAGGGT |

| | | |
|-------------------|---------------|--|
| | | CAAGTGCACAGTCTA |
| SEQ ID NO: 341 | o499_C1_S2_R | CAAGCAGAAGACGGCATA CGAGATTTCTGCTCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 342 | o500_C2_S2_R | CAAGCAGAAGACGGCATA CGAGATGTGCACACAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 343 | o501_C3_S2_R | CAAGCAGAAGACGGCATA CGAGATAAAGCTCAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 344 | o502_C4_S2_R | CAAGCAGAAGACGGCATA CGAGATGACCTCAGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 345 | o503_C5_S2_R | CAAGCAGAAGACGGCATA CGAGATCTTTCCAAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 346 | o504_C6_S2_R | CAAGCAGAAGACGGCATA CGAGATTCTTGGCTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 347 | o505_C7_S2_R | CAAGCAGAAGACGGCATA CGAGATCGCGTCTAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 348 | o506_C8_S2_R | CAAGCAGAAGACGGCATA CGAGATTCGCGCTAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 349 | o507_C9_S2_R | CAAGCAGAAGACGGCATA CGAGATATCCATTCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 350 | o508_C10_S2_R | CAAGCAGAAGACGGCATA CGAGATGCCAGTAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 351 | o509_C11_S2_R | CAAGCAGAAGACGGCATA CGAGATTACCGACGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 352 | o510_C12_S2_R | CAAGCAGAAGACGGCATA CGAGATTCCATACGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 353 | o511_D1_S2_R | CAAGCAGAAGACGGCATA CGAGATAACATGTCAGGGT CAAGTGCACAGTCTA |

| | | |
|-------------------|---------------|---|
| SEQ ID NO: 354 | o512_D2_S2_R | CAAGCAGAAGACGGCATA CGAGATCGACTATAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 355 | o513_D3_S2_R | CAAGCAGAAGACGGCATA CGAGATACCCAAAGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 356 | o514_D4_S2_R | CAAGCAGAAGACGGCATA CGAGATATCGATCGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 357 | o515_D5_S2_R | CAAGCAGAAGACGGCATA CGAGATGTTGGATGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 358 | o516_D6_S2_R | CAAGCAGAAGACGGCATA CGAGATCTATGTGAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 359 | o517_D7_S2_R | CAAGCAGAAGACGGCATA CGAGATTATTTTCGCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 360 | o518_D8_S2_R | CAAGCAGAAGACGGCATA CGAGATCCATGTATAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 361 | o519_D9_S2_R | CAAGCAGAAGACGGCATA CGAGATGCCACGTTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 362 | o520_D10_S2_R | CAAGCAGAAGACGGCATA CGAGATGTCGTGTAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 363 | o521_D11_S2_R | CAAGCAGAAGACGGCATA CGAGATTAAAGTCGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 364 | o522_D12_S2_R | CAAGCAGAAGACGGCATA CGAGATCTTCGGACAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 365 | o523_E1_S2_R | CAAGCAGAAGACGGCATA CGAGATGCACTCTCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 366 | o524_E2_S2_R | CAAGCAGAAGACGGCATA CGAGATTCAGATACAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 367 | o525_E3_S2_R | CAAGCAGAAGACGGCATA |

| | | |
|-------------------|---------------|--|
| NO: 367 | | CGAGATCAGTCCCTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 368 | o526_E4_S2_R | CAAGCAGAAGACGGCATA CGAGATGCCCTAACAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 369 | o527_E5_S2_R | CAAGCAGAAGACGGCATA CGAGATCTGCATCAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 370 | o528_E6_S2_R | CAAGCAGAAGACGGCATA CGAGATCGGTATCGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 371 | o529_E7_S2_R | CAAGCAGAAGACGGCATA CGAGATAAGTATGGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 372 | o530_E8_S2_R | CAAGCAGAAGACGGCATA CGAGATATTCGCGCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 373 | o531_E9_S2_R | CAAGCAGAAGACGGCATA CGAGATATCAAGGTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 374 | o532_E10_S2_R | CAAGCAGAAGACGGCATA CGAGATTTGTGCATAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 375 | o533_E11_S2_R | CAAGCAGAAGACGGCATA CGAGATCTGTGCTGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 376 | o534_E12_S2_R | CAAGCAGAAGACGGCATA CGAGATGTCCGTAGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 377 | o535_F1_S2_R | CAAGCAGAAGACGGCATA CGAGATGTTCAAGAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 378 | o536_F2_S2_R | CAAGCAGAAGACGGCATA CGAGATCACCGTTCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 379 | o537_F3_S2_R | CAAGCAGAAGACGGCATA CGAGATCGAGTTGAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 380 | o538_F4_S2_R | CAAGCAGAAGACGGCATA CGAGATGAGCACGAAGGG |

| | | |
|-------------------|---------------|---|
| | | TCAAGTGCACAGTCTA |
| SEQ ID NO: 381 | o539_F5_S2_R | CAAGCAGAAGACGGCATA CGAGATAGTTTCGTGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 382 | o540_F6_S2_R | CAAGCAGAAGACGGCATA CGAGATCATCAACTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 383 | o541_F7_S2_R | CAAGCAGAAGACGGCATA CGAGATCGAGATCTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 384 | o542_F8_S2_R | CAAGCAGAAGACGGCATA CGAGATTGGCCAGAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 385 | o543_F9_S2_R | CAAGCAGAAGACGGCATA CGAGATTTCCACATAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 386 | o544_F10_S2_R | CAAGCAGAAGACGGCATA CGAGATGAATGCATAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 387 | o545_F11_S2_R | CAAGCAGAAGACGGCATA CGAGATTGGACCCTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 388 | o546_F12_S2_R | CAAGCAGAAGACGGCATA CGAGATGATAGCACAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 389 | o547_G1_S2_R | CAAGCAGAAGACGGCATA CGAGATACGACGACAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 390 | o548_G2_S2_R | CAAGCAGAAGACGGCATA CGAGATCTCAGTATAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 391 | o549_G3_S2_R | CAAGCAGAAGACGGCATA CGAGATCTTAGCTAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 392 | o550_G4_S2_R | CAAGCAGAAGACGGCATA CGAGATCTGTTTACAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 393 | o551_G5_S2_R | CAAGCAGAAGACGGCATA CGAGATTGTCCCACAGGGT CAAGTGCACAGTCTA |

| | | |
|-------------------|---------------|--|
| SEQ ID NO: 394 | o552_G6_S2_R | CAAGCAGAAGACGGCATA CGAGATTCCTGAGGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 395 | o553_G7_S2_R | CAAGCAGAAGACGGCATA CGAGATTAGTTCCAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 396 | o554_G8_S2_R | CAAGCAGAAGACGGCATA CGAGATCATGACTCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 397 | o555_G9_S2_R | CAAGCAGAAGACGGCATA CGAGATGTAAGCGCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 398 | o556_G10_S2_R | CAAGCAGAAGACGGCATA CGAGATAACCCAGTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 399 | o557_G11_S2_R | CAAGCAGAAGACGGCATA CGAGATTTTGAGGGAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 400 | o558_G12_S2_R | CAAGCAGAAGACGGCATA CGAGATAGCCGACAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 401 | o559_H1_S2_R | CAAGCAGAAGACGGCATA CGAGATAAACCCGCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 402 | o560_H2_S2_R | CAAGCAGAAGACGGCATA CGAGATGTAGGGCTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 403 | o561_H3_S2_R | CAAGCAGAAGACGGCATA CGAGATAGACGATTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 404 | o562_H4_S2_R | CAAGCAGAAGACGGCATA CGAGATAGGATGATAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 405 | o563_H5_S2_R | CAAGCAGAAGACGGCATA CGAGATATAATGGCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 406 | o564_H6_S2_R | CAAGCAGAAGACGGCATA CGAGATCTTGGCGTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 406 | o565_H7_S2_R | CAAGCAGAAGACGGCATA |

| | | |
|-------------------|-----------------------|--|
| NO: 407 | | CGAGATAGCTGTGCAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 408 | o566_H8_S2_R | CAAGCAGAAGACGGCATA CGAGATGAGTCCAAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 409 | o567_H9_S2_R | CAAGCAGAAGACGGCATA CGAGATGAATACCAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 410 | o568_H10_S2_R | CAAGCAGAAGACGGCATA CGAGATAGGAGCTTAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 411 | o569_H11_S2_R | CAAGCAGAAGACGGCATA CGAGATGTGACTTAAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 412 | o570_H12_S2_R | CAAGCAGAAGACGGCATA CGAGATTTTGGAACAGGGT CAAGTGCACAGTCTA |
| SEQ ID NO: 413 | o571_板 1_N1_R | AATGATACGGCGACCACCG AGATCTACACAAGATCTGT CTGGTACTGCCAGTTGAA TCTG |
| SEQ ID NO: 414 | o572_板 2_N1_R | AATGATACGGCGACCACCG AGATCTACACTGTCATGAT CTGGTACTGCCAGTTGAA TCTG |
| SEQ ID NO: 415 | o573_板 3_N1_R | AATGATACGGCGACCACCG AGATCTACACTGTAACAGT CTGGTACTGCCAGTTGAA TCTG |
| SEQ ID NO: 416 | o574_板 4_N1_R | AATGATACGGCGACCACCG AGATCTACACGCGCAACTT CTGGTACTGCCAGTTGAA TCTG |
| SEQ ID NO: 417 | o575_板 1_N1_R_短 P5 | AATGATACGGCGACCACCG ATCGATGGCTCTGGTACT GCCAGTTGAATCTG |
| SEQ ID NO: 418 | o576_板 2_N1_R_短 P5 | AATGATACGGCGACCACCG ACATGGTTTTCTGGTACTG CCAGTTGAATCTG |
| SEQ ID | o577_板 3_N1_R_短 | AATGATACGGCGACCACCG |

| | | |
|-------------------|-----------------------|---|
| NO: 419 | P5 | AACGAAAGCTCTGGTACT GCCAGTTGAATCTG |
| SEQ ID NO: 420 | o578_板 4_N1_R_短 P5 | AATGATACGGCGACCACCG AGCTCTGATTCTGGTACT GCCAGTTGAATCTG |
| SEQ ID NO: 421 | o579_板 1_S2_F | AATGATACGGCGACCACCG AGATCTACACATGCCCTCG CTGGTGCTGCAGCTTATTA |
| SEQ ID NO: 422 | o580_板 2_S2_F | AATGATACGGCGACCACCG AGATCTACACCTCAGATGG CTGGTGCTGCAGCTTATTA |
| SEQ ID NO: 423 | o581_板 3_S2_F | AATGATACGGCGACCACCG AGATCTACACGTAATCTGG CTGGTGCTGCAGCTTATTA |
| SEQ ID NO: 424 | o582_板 4_S2_F | AATGATACGGCGACCACCG AGATCTACACGCAAGATTG CTGGTGCTGCAGCTTATTA |
| SEQ ID NO: 425 | o583_板 1_S2_F_短 P5 | AATGATACGGCGACCACCG AATACGCCAGCTGGTGCTG CAGCTTATTA |
| SEQ ID NO: 426 | o584_板 2_S2_F_短 P5 | AATGATACGGCGACCACCG AACCAGTCGGCTGGTGCTG CAGCTTATTA |
| SEQ ID NO: 427 | o585_板 3_S2_F_短 P5 | AATGATACGGCGACCACCG AAACACGGAGCTGGTGCTG CAGCTTATTA |
| SEQ ID NO: 428 | o586_板 4_S2_F_短 P5 | AATGATACGGCGACCACCG ACTGCCTAGGCTGGTGCTG CAGCTTATTA |
| SEQ ID NO: 429 | o587_N1_read1 | tctggttactgccagttgaatctgagggtcc |
| SEQ ID NO: 430 | o588_N1_i7 | ggtgcatttcgctgattttggggtc |
| SEQ ID NO: 431 | o589_N1_i5 | ggaccctcagattcaactggcagtaaccag a |
| SEQ ID NO: 432 | o590_S2_read1 | gctggtgctgcagcttattatgtgggt |
| SEQ ID NO: 433 | o591_S2_i7 | agatgctgtagactgtgcacttgaccct |
| SEQ ID | o592_S2_i5 | accacataataagctgcagcaccagc |

| | | |
|-------------------|---------------|---|
| NO: 434 | | |
| SEQ ID NO: 435 | o593_RP_read1 | gagcggctgtctccacaagtccg |
| SEQ ID NO: 436 | o594_RP_i7 | acccgctcgcaggtccaaatctg |
| SEQ ID NO: 437 | o595_RP_i5 | cggacttgtggagacagccgctc |
| SEQ ID NO: 438 | o596_A1_RP_F | CAAGCAGAAGACGGCATA CGAGATGAGTCTTCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 439 | o597_A2_RP_F | CAAGCAGAAGACGGCATA CGAGATGTTCTATCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 440 | o598_A3_RP_F | CAAGCAGAAGACGGCATA CGAGATTGGGCCAAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 441 | o599_A4_RP_F | CAAGCAGAAGACGGCATA CGAGATATTGTTGGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 442 | o600_A5_RP_F | CAAGCAGAAGACGGCATA CGAGATTCCTGTTGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 443 | o601_A6_RP_F | CAAGCAGAAGACGGCATA CGAGATACACACTTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 444 | o602_A7_RP_F | CAAGCAGAAGACGGCATA CGAGATCCATTCCAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 445 | o603_A8_RP_F | CAAGCAGAAGACGGCATA CGAGATCTAACGGGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 446 | o604_A9_RP_F | CAAGCAGAAGACGGCATA CGAGATCCATAGGAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 447 | o605_A10_RP_F | CAAGCAGAAGACGGCATA CGAGATCAGTGTAGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 448 | o606_A11_RP_F | CAAGCAGAAGACGGCATA CGAGATGATTCTCAAGATT TGGACCTGCGAGCG |

| | | |
|-------------------|---------------|---|
| SEQ ID NO: 449 | o607_A12_RP_F | CAAGCAGAAGACGGCATA CGAGATCCTTCTTAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 450 | o608_B1_RP_F | CAAGCAGAAGACGGCATA CGAGATTCTAAGACAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 451 | o609_B2_RP_F | CAAGCAGAAGACGGCATA CGAGATGCGGCATAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 452 | o610_B3_RP_F | CAAGCAGAAGACGGCATA CGAGATATTGACGAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 453 | o611_B4_RP_F | CAAGCAGAAGACGGCATA CGAGATGCTCCTGAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 454 | o612_B5_RP_F | CAAGCAGAAGACGGCATA CGAGATGCAATCCTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 455 | o613_B6_RP_F | CAAGCAGAAGACGGCATA CGAGATATGTCGTTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 456 | o614_B7_RP_F | CAAGCAGAAGACGGCATA CGAGATTTCTCGGCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 457 | o615_B8_RP_F | CAAGCAGAAGACGGCATA CGAGATCAGGGCTAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 458 | o616_B9_RP_F | CAAGCAGAAGACGGCATA CGAGATAGCCAAGCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 459 | o617_B10_RP_F | CAAGCAGAAGACGGCATA CGAGATAAGCCTGAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 460 | o618_B11_RP_F | CAAGCAGAAGACGGCATA CGAGATCTACAGAGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 461 | o619_B12_RP_F | CAAGCAGAAGACGGCATA CGAGATCGTAGTCGAGATT TGGACCTGCGAGCG |
| SEQ ID | o620_C1_RP_F | CAAGCAGAAGACGGCATA |

| | | |
|-------------------|---------------|---|
| NO: 462 | | CGAGATTTCTGCTCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 463 | o621_C2_RP_F | CAAGCAGAAGACGGCATA CGAGATGTGCACACAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 464 | o622_C3_RP_F | CAAGCAGAAGACGGCATA CGAGATAAAGCTCAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 465 | o623_C4_RP_F | CAAGCAGAAGACGGCATA CGAGATGACCTCAGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 466 | o624_C5_RP_F | CAAGCAGAAGACGGCATA CGAGATCTTTCCAAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 467 | o625_C6_RP_F | CAAGCAGAAGACGGCATA CGAGATTCTTGGCTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 468 | o626_C7_RP_F | CAAGCAGAAGACGGCATA CGAGATCGCGTCTAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 469 | o627_C8_RP_F | CAAGCAGAAGACGGCATA CGAGATTCGCGCTAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 470 | o628_C9_RP_F | CAAGCAGAAGACGGCATA CGAGATATCCATTCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 471 | o629_C10_RP_F | CAAGCAGAAGACGGCATA CGAGATGCCAGTAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 472 | o630_C11_RP_F | CAAGCAGAAGACGGCATA CGAGATTACCGACGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 473 | o631_C12_RP_F | CAAGCAGAAGACGGCATA CGAGATTCCATACGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 474 | o632_D1_RP_F | CAAGCAGAAGACGGCATA CGAGATAACATGTCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 475 | o633_D2_RP_F | CAAGCAGAAGACGGCATA CGAGATCGACTATAAGATT |

| | | |
|-------------------|---------------|--|
| | | TGGACCTGCGAGCG |
| SEQ ID NO: 476 | o634_D3_RP_F | CAAGCAGAAGACGGCATA CGAGATACCCAAAGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 477 | o635_D4_RP_F | CAAGCAGAAGACGGCATA CGAGATATCGATCGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 478 | o636_D5_RP_F | CAAGCAGAAGACGGCATA CGAGATGTTGGATGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 479 | o637_D6_RP_F | CAAGCAGAAGACGGCATA CGAGATCTATGTGAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 480 | o638_D7_RP_F | CAAGCAGAAGACGGCATA CGAGATTATTTTCGCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 481 | o639_D8_RP_F | CAAGCAGAAGACGGCATA CGAGATCCATGTATAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 482 | o640_D9_RP_F | CAAGCAGAAGACGGCATA CGAGATGCCACGTTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 483 | o641_D10_RP_F | CAAGCAGAAGACGGCATA CGAGATGTCGTGTAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 484 | o642_D11_RP_F | CAAGCAGAAGACGGCATA CGAGATTAAAGTCGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 485 | o643_D12_RP_F | CAAGCAGAAGACGGCATA CGAGATCTTCGGACAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 486 | o644_E1_RP_F | CAAGCAGAAGACGGCATA CGAGATGCACTCTCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 487 | o645_E2_RP_F | CAAGCAGAAGACGGCATA CGAGATTCAGATACAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 488 | o646_E3_RP_F | CAAGCAGAAGACGGCATA CGAGATCAGTCCCTAGATT TGGACCTGCGAGCG |

| | | |
|-------------------|---------------|---|
| SEQ ID NO: 489 | o647_E4_RP_F | CAAGCAGAAGACGGCATA CGAGATGCCCTAACAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 490 | o648_E5_RP_F | CAAGCAGAAGACGGCATA CGAGATCTGCATCAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 491 | o649_E6_RP_F | CAAGCAGAAGACGGCATA CGAGATCGGTATCGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 492 | o650_E7_RP_F | CAAGCAGAAGACGGCATA CGAGATAAGTATGGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 493 | o651_E8_RP_F | CAAGCAGAAGACGGCATA CGAGATATTCGCGCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 494 | o652_E9_RP_F | CAAGCAGAAGACGGCATA CGAGATATCAAGGTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 495 | o653_E10_RP_F | CAAGCAGAAGACGGCATA CGAGATTTGTGCATAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 496 | o654_E11_RP_F | CAAGCAGAAGACGGCATA CGAGATCTGTGCTGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 497 | o655_E12_RP_F | CAAGCAGAAGACGGCATA CGAGATGTCCGTAGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 498 | o656_F1_RP_F | CAAGCAGAAGACGGCATA CGAGATGTTCAAGAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 499 | o657_F2_RP_F | CAAGCAGAAGACGGCATA CGAGATCACCGTTCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 500 | o658_F3_RP_F | CAAGCAGAAGACGGCATA CGAGATCGAGTTGAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 501 | o659_F4_RP_F | CAAGCAGAAGACGGCATA CGAGATGAGCACGAAGATT TGGACCTGCGAGCG |
| SEQ ID | o660_F5_RP_F | CAAGCAGAAGACGGCATA |

| | | |
|-------------------|---------------|--|
| NO: 502 | | CGAGATAGTTTCGTGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 503 | o661_F6_RP_F | CAAGCAGAAGACGGCATA CGAGATCATCAACTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 504 | o662_F7_RP_F | CAAGCAGAAGACGGCATA CGAGATCGAGATCTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 505 | o663_F8_RP_F | CAAGCAGAAGACGGCATA CGAGATTGGCCAGAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 506 | o664_F9_RP_F | CAAGCAGAAGACGGCATA CGAGATTTCCACCATAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 507 | o665_F10_RP_F | CAAGCAGAAGACGGCATA CGAGATGAATGCATAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 508 | o666_F11_RP_F | CAAGCAGAAGACGGCATA CGAGATTGGACCCTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 509 | o667_F12_RP_F | CAAGCAGAAGACGGCATA CGAGATGATAGCACAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 510 | o668_G1_RP_F | CAAGCAGAAGACGGCATA CGAGATACGACGACAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 511 | o669_G2_RP_F | CAAGCAGAAGACGGCATA CGAGATCTCAGTATAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 512 | o670_G3_RP_F | CAAGCAGAAGACGGCATA CGAGATCTTAGCTAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 513 | o671_G4_RP_F | CAAGCAGAAGACGGCATA CGAGATCTGTTTACAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 514 | o672_G5_RP_F | CAAGCAGAAGACGGCATA CGAGATTGTCCCACAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 515 | o673_G6_RP_F | CAAGCAGAAGACGGCATA CGAGATTCCTGAGGAGATT |

| | | |
|-------------------|---------------|---|
| | | TGGACCTGCGAGCG |
| SEQ ID NO: 516 | o674_G7_RP_F | CAAGCAGAAGACGGCATA CGAGATTAGTTCCAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 517 | o675_G8_RP_F | CAAGCAGAAGACGGCATA CGAGATCATGACTCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 518 | o676_G9_RP_F | CAAGCAGAAGACGGCATA CGAGATGTAAGCGCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 519 | o677_G10_RP_F | CAAGCAGAAGACGGCATA CGAGATAACCCAGTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 520 | o678_G11_RP_F | CAAGCAGAAGACGGCATA CGAGATTTTGAGGGAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 521 | o679_G12_RP_F | CAAGCAGAAGACGGCATA CGAGATAGCCGACAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 522 | o680_H1_RP_F | CAAGCAGAAGACGGCATA CGAGATAAACCCGCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 523 | o681_H2_RP_F | CAAGCAGAAGACGGCATA CGAGATGTAGGGCTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 524 | o682_H3_RP_F | CAAGCAGAAGACGGCATA CGAGATAGACGATTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 525 | o683_H4_RP_F | CAAGCAGAAGACGGCATA CGAGATAGGATGATAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 526 | o684_H5_RP_F | CAAGCAGAAGACGGCATA CGAGATATAATGGCAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 527 | o685_H6_RP_F | CAAGCAGAAGACGGCATA CGAGATCTTGCGGTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 528 | o686_H7_RP_F | CAAGCAGAAGACGGCATA CGAGATAGCTGTGCAGATT TGGACCTGCGAGCG |

| | | |
|-------------------|-----------------------|---|
| SEQ ID NO: 529 | o687_H8_RP_F | CAAGCAGAAGACGGCATA CGAGATGAGTCCAAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 530 | o688_H9_RP_F | CAAGCAGAAGACGGCATA CGAGATGAATACCAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 531 | o689_H10_RP_F | CAAGCAGAAGACGGCATA CGAGATAGGAGCTTAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 532 | o690_H11_RP_F | CAAGCAGAAGACGGCATA CGAGATGTGACTTAAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 533 | o691_H12_RP_F | CAAGCAGAAGACGGCATA CGAGATTTTGGAACAGATT TGGACCTGCGAGCG |
| SEQ ID NO: 534 | o692_板 1_RP_R | AATGATACGGCGACCACCG AGATCTACACCTCTCTATG AGCGGCTGTCTCCACAAGT |
| SEQ ID NO: 535 | o693_板 2_RP_R | AATGATACGGCGACCACCG AGATCTACACTATCCTCTG AGCGGCTGTCTCCACAAGT |
| SEQ ID NO: 536 | o694_板 3_RP_R | AATGATACGGCGACCACCG AGATCTACACGTAAGGAGG AGCGGCTGTCTCCACAAGT |
| SEQ ID NO: 537 | o695_板 4_RP_R | AATGATACGGCGACCACCG AGATCTACACACTGCATAG AGCGGCTGTCTCCACAAGT |
| SEQ ID NO: 538 | o696_板 1_RP_R_短 P5 | AATGATACGGCGACCACCG AAAGGAGTAGAGCGGCTGT CTCCACAAGT |
| SEQ ID NO: 539 | o697_板 2_RP_R_短 P5 | AATGATACGGCGACCACCG ACTAAGCCTGAGCGGCTGT CTCCACAAGT |
| SEQ ID NO: 540 | o698_板 3_RP_R_短 P5 | AATGATACGGCGACCACCG ACGTCTAATGAGCGGCTGT CTCCACAAGT |
| SEQ ID NO: 541 | o699_板 4_RP_R_短 P5 | AATGATACGGCGACCACCG ATCTCTCCGGAGCGGCTGT CTCCACAAGT |
| SEQ ID | o700_板 1_RP_R | AATGATACGGCGACCACCG |

| | | |
|-------------------|-----------------------|---|
| NO: 542 | | AGATCTACACAAGATCTGG AGCGGCTGTCTCCACAAGT |
| SEQ ID NO: 543 | o701_板 2_RP_R | AATGATACGGCGACCACCG AGATCTACACTGTCATGAG AGCGGCTGTCTCCACAAGT |
| SEQ ID NO: 544 | o702_板 3_RP_R | AATGATACGGCGACCACCG AGATCTACACTGTAACAGG AGCGGCTGTCTCCACAAGT |
| SEQ ID NO: 545 | o703_板 4_RP_R | AATGATACGGCGACCACCG AGATCTACACGCGCAACTG AGCGGCTGTCTCCACAAGT |
| SEQ ID NO: 546 | o704_板 1_RP_R_短 P5 | AATGATACGGCGACCACCG ATCGATGGCGAGCGGCTGT CTCCACAAGT |
| SEQ ID NO: 547 | o705_板 2_RP_R_短 P5 | AATGATACGGCGACCACCG ACATGGTTTGAGCGGCTGT CTCCACAAGT |
| SEQ ID NO: 548 | o706_板 3_RP_R_短 P5 | AATGATACGGCGACCACCG AACGAAAGCGAGCGGCTGT CTCCACAAGT |
| SEQ ID NO: 549 | o707_板 4_RP_R_短 P5 | AATGATACGGCGACCACCG AGCTCTGATGAGCGGCTGT CTCCACAAGT |
| SEQ ID NO: 550 | o708_板 1_RP_R | AATGATACGGCGACCACCG AGATCTACACATGCCCTCG AGCGGCTGTCTCCACAAGT |
| SEQ ID NO: 551 | o709_板 2_RP_R | AATGATACGGCGACCACCG AGATCTACACCTCAGATGG AGCGGCTGTCTCCACAAGT |
| SEQ ID NO: 552 | o710_板 3_RP_R | AATGATACGGCGACCACCG AGATCTACACGTAATCTGG AGCGGCTGTCTCCACAAGT |
| SEQ ID NO: 553 | o711_板 4_RP_R | AATGATACGGCGACCACCG AGATCTACACGCAAGATTG AGCGGCTGTCTCCACAAGT |
| SEQ ID NO: 554 | o712_板 1_RP_R_短 P5 | AATGATACGGCGACCACCG AATACGCCAGAGCGGCTGT CTCCACAAGT |
| SEQ ID NO: 555 | o713_板 2_RP_R_短 P5 | AATGATACGGCGACCACCG AACCAGTCGGAGCGGCTGT |

| | | |
|-------------------|-----------------------|--|
| | | CTCCACAAGT |
| SEQ ID NO: 556 | o714_板 3_RP_R_短 P5 | AATGATACGGCGACCACCG AAACACGGAGAGCGGCTGT CTCCACAAGT |
| SEQ ID NO: 557 | o715_板 4_RP_R_短 P5 | AATGATACGGCGACCACCG ACTGCCTAGGAGCGGCTGT CTCCACAAGT |
| SEQ ID NO: 558 | o716_skpp15-1-F_N1 | GGGTCACGCGTAGGA GTTCTATCGACCCCAAAT CAGCGAAAT |
| SEQ ID NO: 559 | o717_skpp15-1-R_N1 | GTTCCGCAGCCACAC AAGATCTGTCTGGTACTG CCAGTTGAATCTG |
| SEQ ID NO: 560 | o718_skpp15-2-F_N1 | CGCGTCGAGTAGGGT GTTCTATCGACCCCAAAT CAGCGAAAT |
| SEQ ID NO: 561 | o719_skpp15-2-R_N1 | GCCGTGTGAAGCTGG AAGATCTGTCTGGTACTG CCAGTTGAATCTG |
| SEQ ID NO: 562 | o720_skpp15-3-F_N1 | CGATCGCCCTTGGTG GTTCTATCGACCCCAAAT CAGCGAAAT |
| SEQ ID NO: 563 | o721_skpp15-3-R_N1 | GGTTTAGCCGGCGTG AAGATCTGTCTGGTACTG CCAGTTGAATCTG |
| SEQ ID NO: 564 | o722_skpp15-4-F_N1 | GGTCGAGCCGGA ACT GTTCTATCGACCCCAAAT CAGCGAAAT |
| SEQ ID NO: 565 | o723_skpp15-4-R_N1 | GGATGCGCACCCAGA AAGATCTGTCTGGTACTG CCAGTTGAATCTG |
| SEQ ID NO: 566 | o724_skpp15-5-F_N1 | TCCCGGCGTTGTCCT GTTCTATCGACCCCAAAT CAGCGAAAT |
| SEQ ID NO: 567 | o725_skpp15-5-R_N1 | GCTCCGTC ACTGCC AAGATCTGTCTGGTACTG CCAGTTGAATCTG |
| SEQ ID NO: 568 | o726_skpp15-6-F_N1 | CGCAGGGTCCAGAGT GTTCTATCGACCCCAAAT CAGCGAAAT |

| | | |
|-------------------|--------------------|---|
| SEQ ID NO: 569 | o727_skpp15-6-R_N1 | GTTCGCGCGAAGGAA AAGATCTGTCTGGTACTG CCAGTTGAATCTG |
| SEQ ID NO: 570 | o728_skpp-1-F_N1 | ATATAGATGCCGTCCTAGC G GTTCTATCGACCCCAAAT CAGCGAAAT |
| SEQ ID NO: 571 | o729_skpp-1-R_N1 | AAGTATCTTTCCTGTGCC A AAGATCTGTCTGGTACTG CCAGTTGAATCTG |
| SEQ ID NO: 572 | o730_skpp-2-F_N1 | CCCTTTAATCAGATGCGTC G GTTCTATCGACCCCAAAT CAGCGAAAT |
| SEQ ID NO: 573 | o731_skpp-2-R_N1 | TGGTAGTAATAAGGGCGAC C AAGATCTGTCTGGTACTG CCAGTTGAATCTG |
| SEQ ID NO: 574 | o732_skpp-3-F_N1 | TTGGTCATGTGCTTTTCGTT GTTCTATCGACCCCAAAT CAGCGAAAT |
| SEQ ID NO: 575 | o733_skpp-3-R_N1 | AGGGGTATCGGATACTCAG A AAGATCTGTCTGGTACTG CCAGTTGAATCTG |
| SEQ ID NO: 576 | o734_skpp-4-F_N1 | GGGTGGGTAAATGGTAATG C GTTCTATCGACCCCAAAT CAGCGAAAT |
| SEQ ID NO: 577 | o735_skpp-4-R_N1 | ATCGATTCCCCGGATATAG C AAGATCTGTCTGGTACTG CCAGTTGAATCTG |
| SEQ ID NO: 578 | o736_skpp-5-F_N1 | TCCGACGGGGAGTATATAC T GTTCTATCGACCCCAAAT CAGCGAAAT |
| SEQ ID NO: 579 | o737_skpp-5-R_N1 | TACTAACTGCTTCAGGCCA A |

| | | |
|-------------------|--------------------|---|
| | | AAGATCTGTCTGGTACTG CCAGTTGAATCTG |
| SEQ ID NO: 580 | o738_skpp-6-F_N1 | CATGTTTAGGAACGCTACC G GTTCTATCGACCCCAAAT CAGCGAAAT |
| SEQ ID NO: 581 | o739_skpp-6-R_N1 | AATAATCTCCGTTCCCTCCC AAGATCTGTCTGGTACTG CCAGTTGAATCTG |
| SEQ ID NO: 582 | o740_skpp15-1-F_S2 | GGGTCACGCGTAGGA GTTCTATCGCTGGTGCTGC AGCTTATTA |
| SEQ ID NO: 583 | o741_skpp15-1-R_S2 | GTTCCGCAGCCACAC AAGATCTGAGGGTCAAGTG CACAGTCTA |
| SEQ ID NO: 584 | o742_skpp15-2-F_S2 | CGCGTCGAGTAGGGT GTTCTATCGCTGGTGCTGC AGCTTATTA |
| SEQ ID NO: 585 | o743_skpp15-2-R_S2 | GCCGTGTGAAGCTGG AAGATCTGAGGGTCAAGTG CACAGTCTA |
| SEQ ID NO: 586 | o744_skpp15-3-F_S2 | CGATCGCCCTTGGTG GTTCTATCGCTGGTGCTGC AGCTTATTA |
| SEQ ID NO: 587 | o745_skpp15-3-R_S2 | GGTTTAGCCGGCGTG AAGATCTGAGGGTCAAGTG CACAGTCTA |
| SEQ ID NO: 588 | o746_skpp15-4-F_S2 | GGTCGAGCCGGA GTTCTATCGCTGGTGCTGC AGCTTATTA |
| SEQ ID NO: 589 | o747_skpp15-4-R_S2 | GGATGCGCACCCAGA AAGATCTGAGGGTCAAGTG CACAGTCTA |
| SEQ ID NO: 590 | o748_skpp15-5-F_S2 | TCCCGGCGTTGTCCT GTTCTATCGCTGGTGCTGC AGCTTATTA |
| SEQ ID NO: 591 | o749_skpp15-5-R_S2 | GCTCCGTC AAGATCTGAGGGTCAAGTG CACAGTCTA |
| SEQ ID | o750_skpp15-6-F_S2 | CGCAGGGTCCAGAGT |

| | | |
|-------------------|--------------------|--|
| NO: 592 | | GTTCTATCGCTGGTGCTGC AGCTTATTA |
| SEQ ID NO: 593 | o751_skpp15-6-R_S2 | GTTTCGCGCGAAGGAA AAGATCTGAGGGTCAAGTG CACAGTCTA |
| SEQ ID NO: 594 | o752_skpp-1-F s_S2 | ATATAGATGCCGTCCTAGC G GTTCTATCGCTGGTGCTGC AGCTTATTA |
| SEQ ID NO: 595 | o753_skpp-1-R s_S2 | AAGTATCTTTCCTGTGCC A AAGATCTGAGGGTCAAGTG CACAGTCTA |
| SEQ ID NO: 596 | o754_skpp-2-F s_S2 | CCCTTTAATCAGATGCGTC G GTTCTATCGCTGGTGCTGC AGCTTATTA |
| SEQ ID NO: 597 | o755_skpp-2-R s_S2 | TGGTAGTAATAAGGGCGAC C AAGATCTGAGGGTCAAGTG CACAGTCTA |
| SEQ ID NO: 598 | o756_skpp-3-F s_S2 | TTGGTCATGTGCTTTTCGTT GTTCTATCGCTGGTGCTGC AGCTTATTA |
| SEQ ID NO: 599 | o757_skpp-3-R s_S2 | AGGGGTATCGGATACTCAG A AAGATCTGAGGGTCAAGTG CACAGTCTA |
| SEQ ID NO: 600 | o758_skpp-4-F s_S2 | GGGTGGGTAAATGGTAATG C GTTCTATCGCTGGTGCTGC AGCTTATTA |
| SEQ ID NO: 601 | o759_skpp-4-R s_S2 | ATCGATTCCCCGGATATAG C AAGATCTGAGGGTCAAGTG CACAGTCTA |
| SEQ ID NO: 602 | o760_skpp-5-F s_S2 | TCCGACGGGGAGTATATAC T GTTCTATCGCTGGTGCTGC AGCTTATTA |

| | | |
|-------------------|--------------------|--|
| SEQ ID NO: 603 | o761_skpp-5-R s S2 | TACTAACTGCTTCAGGCCA A AAGATCTGAGGGTCAAGTG CACAGTCTA |
| SEQ ID NO: 604 | o762_skpp-6-F s S2 | CATGTTTAGGAACGCTACC G GTTCTATCGCTGGTGCTGC AGCTTATTA |
| SEQ ID NO: 605 | o763_skpp-6-R s S2 | AATAATCTCCGTTCCCTCCC AAGATCTGAGGGTCAAGTG CACAGTCTA |

实施例4-SwabSeq:用于大规模SARS-CoV-2测试的高通量平台

[0202] 在不存在有效疫苗的情况下,公共卫生策略仍然是控制造成COVID-19的SARS-CoV-2蔓延的唯一工具。与传染性与症状相关联的SARS-CoV-1相比,SARS-CoV-2在无症状/症状前期中的传染性很高。因此,仅仅基于症状来限制传播是不可能的,这使得对SARS-CoV-2的筛查对于控制大流行来说至关重要。

[0203] 随着区域封锁解除,人们重返工作岗位并恢复其他活动,感染率再次开始上升。在美国的许多地区,病例的升高已经超出了定量RT-PCR测试的容量,所述测试构成了FDA批准针对COVID-19的测试的大部分。即使在美国的那些地区中已经扩大了测试容量,但是每次检测的成本为大约\$100,妨碍了针对群体筛查的临床测试的部署。此外,返回结果的长时间延迟是由于测试容量不足,而不是因为测定时间要数小时,这使得测试在防止病毒传播和抑制局部爆发方面毫无用处。频繁、廉价的大规模测试,以及接触者追踪和感染个体的隔离,提供了阻止病毒蔓延的最佳机会。本文描述了SARS-CoV-2测试平台SwabSeq,其利用下一代测序以大规模扩大测试容量。

[0204] SwabSeq在几个关键领域中改进了一步逆转录和聚合酶链反应(RT-PCR)方法。SwabSeq使用分子条形码,其被嵌入在RT和PCR引物中以唯一标记每个样品。在RT-PCR步骤之后,将条形码化样品合并至单个测序文库中,实现在Illumina测序仪上多达成千上万个样品的复用。测试的读出不是荧光(因此不需要昂贵的qPCR机器),而是病毒测序读段的计数。短读段长度(26个碱基对)足以鉴定靶序列,并将总测序时间缩短至大约5小时。

[0205] SwabSeq鉴定SARS-CoV-2的稳健能力的关键是包括体外RNA内标,其允许每个孔内的病毒读段计数的归一化。此内标将基于计数的测定转变为比率计量测定。这很重要,因为尽管临床样品中的RT和PCR抑制存在异质性,它仍实现准确检出阳性和阴性样品的能力。依靠测序作为读出,连同体外RNA标准一起,使SwabSeq使用进行多达50次循环的扩增,直到引物耗尽。这种终点PCR克服了无提取样品中存在的酶抑制,使得进一步简化了工作流程,同时使分析灵敏度的损失降至最低。与目前的RT-定量PCR(RT-qPCR)诊断测试相比,这些特征实现了以高得多的通量进行测试,目前的诊断测试需要RNA纯化和RT-qPCR,成本高、劳动和试剂密集型步骤,受必要仪器的可用性的限制。相比之下,SwabSeq仅依赖于易得且未充分利用的热循环仪和DNA测序仪。SwabSeq可以在每天成千上万个样品的规模上进行,而无需进行超过每天数百个样品规模通常需要的广泛自动化。

[0206] 如本文所证明,SwabSeq对于检测纯化样品中的病毒RNA具有极高的灵敏度和特异

性。本文的公开内容和实施例进一步证明了来自中鼻拭子和口腔液的无提取裂解物的检测限低。这些结果证明了SwabSeq用于以前所未有的规模进行的SARS-CoV-2测试的潜力。

结果

[0207] SwabSeq是一个简单且可扩展的方案,由5个步骤组成(图4,A图):(1)样品收集,(2)使用在i7和i5位置包含唯一分子索引的引物进行逆转录和PCR(图4,B图,图6),(3)汇集唯一条形码化样品以用于文库制备,(4)对汇集文库进行测序,以及(5)将条形码化测序读段计算分配至每个样品以用于计数和病毒检测。

[0208] 测定由两个引物集组成,所述引物集扩增两种基因:SARS-CoV-2的S2基因和人核糖核酸酶P/MRP亚基P30(RPP30)。测定包括合成体外转录对照RNA,其与目标用于扩增的病毒序列相同,除了允许区分对应于合成对照的测序读段和对应于天然安序列的测序读段的短改变链段(图4,C图)。引物以相等的效率扩增天然序列和合成序列(图7)。

[0209] S2内部对照提供两个目的。首先,PCR和逆转录的抑制剂以及其他孔间变异的来源可能以相同方式影响对照和样品的扩增。与单独天然读段计数相比,天然读段的数目与对照读段的数目的比率提供了更准确且定量的病毒负载的量度(图4,E和D图)。其次,尽管使用了终点PCR,内部对照仍允许在大范围的病毒输入中保持线性(图8)。使用这种方法,每个孔中的最终DNA量在很大程度上由总引物浓度而不是由病毒负载限定,预期阴性样品的S2读段总数与阳性样品相似。不同之处在于,在阴性样品中,所有S2扩增子读段都映射至对照合成序列。除了病毒S2之外,所述测定/方法还包括逆转录和扩增人持家基因以控制样本质量和收集效率(图4,F图)。

[0210] 在RT-PCR之后,将样品等体积合并、纯化并用于生成一个测序文库。Illumina MiSeq和/或Illumina NextSeq均可用于对这些文库进行测序(图9)。通过仅对扩增子的26个碱基对进行测序(方法),可以最大限度地缩短仪器测序时间。每个读段都基于其序列被分类为源自天然还是对照S2或RPP30,并且基于相关索引序列(条形码)分配给样品。为了最大限度地提高特异性并避免因不正确的分类或分配而产生的假阳性信号,保守的编辑距离阈值用于此匹配操作(方法和补充结果)。如果与预期序列中的一个不匹配,则丢弃测序读段。获得每个样品的天然和对照S2和RPP30读段的计数,并将其用于下游分析(方法,下文)。

[0211] 据观察,几千个读段足以检测并定量样品中病毒RNA的存在(保守地,10,000个)。这意味着在MiSeq v3流动池上每次运行1,500个样品,在NextSeq上每次运行20,000个样品,并且在NovaSeq S2流动池上每次运行至多150,000个样品。计算分析每次运行只需几分钟。SwabSeq是一种用于COVID-19测试的简化且可扩展的方案。SwabSeq方案通过鉴定并消除多个噪声源来进一步优化(图21,A、B、C和D图)。

作为诊断平台的SwabSeq的验证

[0212] SwabSeq首次在纯化的RNA样品上进行了验证,所述样品先前由UCLA临床实验室用标准RT-qPCR分析进行了测试。为了确定分析检测限,用汇集的残留临床鼻咽拭子样本稀释灭活病毒。残留样品全部被证实对SARS-CoV-2呈阴性。在这些残留样品中,对热灭活的SARS-CoV-2(ATCC® VR-1986HK)进行了2倍连续稀释,范围为每毫升8000至125个基因组拷贝当量(GCE)。在低至250GCE/mL的所有样品中检测到SARS-CoV-2,并且在125GCE/mL的大多数样品中检测到(图2A)。这些结果表明,SwabSeq具有高灵敏度,分析检测限(LOD)为250GCE/mL。此检测限低于许多目前已批准且高度灵敏的RT-qPCR测定。这种比较证明

SwabSeq作为临床诊断测试可能非常有效。

[0213] SwabSeq以高灵敏度和特异性检测病毒。对来自UCLA临床微生物学实验室的确认阳性 (n=31) 和阴性 (n=33) 样品进行了重新测试。观察到所有样品与RT-qPCR结果100%符合率 (图5)。在MiSeq和NextSeq 550上对文库进行测序 (图9), 不同测序仪器之间具有100%的一致性。

[0214] 扩大RT-qPCR诊断测试的主要瓶颈之一是RNA纯化步骤。RNA提取难以自动化, 并且供应链不能满足大流行过程中对必要试剂的需求。SwabSeq检测直接来自多种无提取样品类型的SARS-CoV-2的能力被探索为一种规避RNA纯化瓶颈的方式。

[0215] CDC针对鼻拭子收集推荐了几种类型的培养基: 病毒运送培养基 (VTM)、Amies运送培养基和生理盐水。主要的技术挑战起因于收集缓冲液中成分对RT或PCR的抑制。用水稀释样本克服了RT和PCR抑制, 并且使得能够检测人为和阳性临床患者样品中的病毒RNA (图11)。对直接收集至Tris-EDTA (TE) 缓冲液中并用水1:1稀释的鼻拭子进行了测试。这种方法得到检测限558GCE/mL (图5, C图)。对收集至生理盐水中的鼻咽样品的无提取方案与UCLA临床微生物学实验室进行的RT-qPCR之间的比较显示, 所有样品100%符合率 (图2D)。

[0216] 还测试了无提取唾液方案, 在所述方案中, 使用漏斗样收集装置, 将唾液直接收集至基体管中 (图12)。主要的技术挑战是防止唾液中的各种组分降解病毒RNA以及确保准确移取这种异质且粘性的样品类型。与不预加热相比, 将唾液样品加热至95°C持续30分钟减少了PCR抑制并改善了S2扩增子的检测 (图13)。在加热步骤之后, 将样品以1:1体积的2xTBE与0.1% Tween-20稀释。使用此方法, 获得LoD 2000GCE/mL (图5, E图)。

讨论

[0217] Swabseq有可能缓解诊断临床测试中的现有瓶颈。此外, SwabSeq具有更大的潜力, 能够通过群体监视实现大流行抑制所需的规模的测试。所述技术代表了大规模并行下一代测序用于感染性疾病监视和诊断的新用途。SwabSeq可以检测来自纯化的RNA和无提取裂解物的临床样本中的SARS-CoV-2RNA, 并且维持与临床诊断实验室中进行的RT-qPCR相当的高临床和分析灵敏度以及特异性。还对SwabSeq进行了进一步优化, 以优先规模和低成本, 因为这些是当前COVID-19诊断测试缺少的关键因素。

[0218] SwabSeq应该被评估为监视测试的工具, 而不是临床测试的工具。临床测试为临床决策提供信息, 因此需要高灵敏度和特异性, 而对于监视测试来说, 最重要的因素是测试的广度和频率以及周转时间。返回结果快速的足够广泛且频繁的测试可以通过对感染个体的选择性隔离而不是全面的居家命令来有效地抑制病毒爆发。对大学校园的监视测试的流行病学模型显示, 即使诊断测试具有70%的灵敏度, 但频繁仅此和且周转时间短, 也可以抑制传播。频繁测试的实际实施的主要挑战包括测试成本以及每天收集和成百上千个样品的后勤。

[0219] 在诊断测试中使用Illumina测序已引起在周转时间和成本方面的担忧。SwabSeq使用短测序运行, 在5小时内读出靶序列的分子索引和26个碱基对, 然后在5分钟内在台式计算机上进行计算分析。当在一次MiSeq运行中分析1,000个样品时, 测序试剂的成本低于每个样品\$1。在NextSeq550上运行10,000个样品, 每个流动池产生13倍以上的读段, 可以将此成本降低大约10倍。减少反应体积和使用更便宜的RT-PCR试剂的进一步优化可以进一步降低每次测试的总成本。

[0220] 最后,扩大对SARS-CoV-2的测试需要高通量样品收集和处理工作流程。大多数学术临床实验室中常见的手动过程不容易与简单的自动化兼容。当前将鼻咽拭子放入VTM、Amies或NS中的方案是追溯到前分子遗传学时代的收集方法,在那时使用活病毒培养物来鉴定对细胞系的细胞病变影响并且测试以低通量手动进行。对可扩展的收集方法的全新视角将对测试社区有巨大益处。

[0221] 几个小组已经试行了“轻量级”样本收集方法,通过智能电话应用程序将样品登记和患者信息收集直接推送到测试的个体。样品申请的大部分工作是由于电子健康系统之间缺乏互操作性,实验室专业人员手动输入每个样品的信息。通过开发符合HIPAA的注册过程,可以简化劳动密集型的样品申请。为了提高可扩展性,可以使用这样的样品收集方法,其使用与简单自动化(诸如自动化加盖器和96头液体处理器)兼容的小体积管。这些方法减少了实验室处理和进行测试所需的动手工作的量。

[0222] 通过提高满足诊断和广泛监视测试的需求的测试容量,SwabSeq诊断平台补充了传统的临床诊断测试,以及越来越多的针对COVID-19新兴的即时诊断平台。展望未来,SwabSeq可容易扩展以适应其他病原体和病毒靶标。这可以进一步增加测试的有用性,特别是在冬季感冒和流感季节期间,当多种呼吸道病原体在人群中传播并且不能单独基于症状而轻易区分时。随着社会计划安全地重新开放社会的教育、商业和娱乐行业,监视测试可能会成为新常态的一部分。

方法

[0223] 样品收集:研究中使用的患者样品均去鉴定化。所有样品均在UCLA IRB批准的情况下获得。健康护理提供者在疑似COVID19的患者中收集鼻咽样品。

[0224] 人为样本的创建:对于检测实验的临床限制,汇集了UCLA临床微生物学实验室收集的确认的COVID-19阴性残留鼻咽拭子样本。然后将汇集的临床样品加标指定浓度的ATCC灭活病毒(ATCC1986-HK),并如下所述进行提取。对于临床纯化的RNA样品,将它们以鼻咽拭子的形式收集,并使用KingFisherFlex (ThermoFisher Scientific) 仪器,使用MagMax珠子提取进行纯化。所有提取均根据制造商的方案进行。对于无提取样品,首先将样品以指定浓度制备成汇集的、确认为阴性的临床样品,并且在TE缓冲液或水中稀释样品,之后添加至RT-PCR主混合物中。

[0225] 无提取唾液样本的处理:使用小漏斗(来自Amazon的零件编号)将直接唾液收集到基体管中。将唾液样品收集到基体管中并加热至95°C持续30分钟。然后将样品冷冻或用2X TBE和1% Tween-20稀释处理,最终浓度为1x TBE和0.5% Tween-2。测试了具有Qiagen蛋白酶和RNA Secure (ThermoFisher) 的1xTween,这也有效但导致更多的样品间变异性,并且需要额外的孵育步骤。

[0226] 无提取鼻拭子裂解物的处理:使用在56C下持续30分钟的加热灭活将所有无提取裂解物灭活。然后将样品以1:4的比率用水稀释,然后直接添加至主混合物中。稀释量根据所使用的液体培养基而变化。在CDC推荐的培养基中,生理盐水表现得最稳健。病毒运送培养基和Amies缓冲液显示了显著的PCR抑制,即使在水中稀释也难以克服。在一个实施方案中,将拭子直接放入稀释的TE缓冲液中,这对PCR抑制没有影响。

[0227] 条形码引物设计:从1,536个唯一10bp i5条形码的组和1,536个唯一10bp i7条形码的组选择条形码引物。这些10bp条形码满足以下标准,即任何两个索引(在i5和i7组内)

之间的最小莱文斯坦距离为3,并且条形码不含有大于2个核苷酸的均聚物重复。此外,将python API中的辅助函数用于Primer3,对条形码进行选择以使同二聚化和异二聚化最小。

[0228] S2和RPP30合成标的构建:使用表2中的引物对SARS-CoV-2的gRNA (Twist BioSciences,#1) 进行RT-PCR,以构建S2加标DNA模板。对HEK293T裂解物进行RT PCR (FP_1, R) 和第二轮PCR (FP_2,R),以构建RPP30加标DNA模板。产品在凝胶上运行以鉴定约150bp的特定产物。使用Ampure珠子 (Axygen),使用1.8的珠子:样品体积比,纯化DNA。将混合物涡旋并在室温下孵育5分钟。使用磁体结合珠子1分钟,用70% EtOH洗涤两次,将珠子风干5分钟,然后从磁体除去并在100uL IDTE缓冲液中洗脱。将珠子溶液放回磁体上,并且1分钟后除去洗脱液。通过nanodrop (Denovix) 对DNA进行定量。

[0229] 此制备的DNA模板用于标准HiScribe T7体外转录 (NEB)。根据制造商的说明,使用每20uL反应物300ng模板DNA,在37°C下孵育16小时,制备IVT反应物。根据制造商的说明,用DNA酶I处理IVT反应物。根据制造商的说明,用RNA Clean&Concentrator-25试剂盒 (Zymo Research) 纯化RNA,并洗脱到水中。通过nanodrop并且根据制造商的说明 (Agilent),使用针对TapeStation的RNA screen tape试剂盒对RNA加标进行定量,以验证RNA的大小正确 (约133nt)。

表2

| | | |
|---------------|------------------|--|
| SEQ NO: 13 | ID S2_FP | TAATACGACTCACTATAGGGCTGGTG CTGCAGCTTATTATGTGGGTATAGAA CAACCTAGGACTTTTCTATTAA |
| SEQ NO: 14 | ID S2_RP | AACGTACACTTTGTTTCTGAGAGAGG |
| SEQ NO: 15 | ID RPP30_FP_1 | CTGACCTGAAGGCTGACGCCGACTT GTGGAGACAGC |
| SEQ NO: 16 | ID RPP30_FP2 | TAATACGACTCACTATAGGGAGATTT GGACCTGCGAGCGGGTTCTGACCTGA AGGCTGA |
| SEQ NO: 17 | ID RPP30_R | GGTTTTTCAATTCCTGTTTCTTTTCCT TAAAGTCAACG |

[0230] 一步RT-PCR:使用Luna® Universal One-Step RT-qPCR试剂盒 (New England BioSciences E3005) 或TaqPath™1-Step RT-qPCR主混合物 (Thermofisher Scientific, A15300) 进行RT-PCR,反应体积为20μL。两种试剂盒均根据制造商的方案使用。主混合物中引物的最终浓度对于RPP30 F和R引物为50nM,并且对于S2 F和R引物为400nM。以每个反应500个拷贝的拷贝数,将合成S2 RNA直接添加至主混合物中。将样品加载至20μL反应中。所有反应均在96或384孔形式上运行,并且热循环仪条件根据制造商的方案运行。对于纯化的RNA样品,进行了40次循环的PCR。对于未纯化的样品,进行了50次循环的终点PCR。

[0231] 多元文库制备:在RT-PCR反应后,使用多通道移液器或Integra Viaflow Benchtop液体处理器汇集样品。将每孔6μL合并到无菌贮存器中,并将整个体积转移到15mL锥形管中并涡旋。将100uL总体积转移到1.7mL eppendorf管中进行双侧SPRI清洗。简而言之,将50μL AmpureXP珠子 (A63880) 添加至100μL汇集的PCR体积中并涡旋。5分钟后,用磁体

收集珠子持续1分钟,并且将上清液转移至新的ependorf管中。将额外的130uL Ampure XP 珠子添加至150uL上清液中并涡旋。再过5分钟后,使用磁体收集珠子持续1分钟,并用新鲜的70% EtOH洗涤珠子两次。在40uL qiagen EB缓冲液中将DNA从珠子上洗脱下来。用磁体收集珠子持续1分钟,并且将33uL上清液转移至新的管中。

[0232] 测序方案:在Illumina MiSeq (2012) 或Nextseq550上对文库进行测序。在每次运行MiSeq之前,根据Illumina方案使用次氯酸钠溶液(Sigma Aldrich,239305)进行漂洗。将汇集且定量的文库稀释至6nM的浓度(基于Qubit 4荧光计和Illumina的ng/ul与nM之间的转换公式),并且以25pM(MiSeq)或1.5pM(NextSeq)加载至测序仪上。将PhiX Control v3 (Illumina,FC-110-3001)以文库的估计30-40%加标至文库中。PhiX为读段1提供额外的序列多样性,这有助于模板注册并提高运行和碱基质量。

[0233] 对于此应用,MiSeq需要2种定制测序引物混合物,读段1引物混合物和i7引物混合物。两种混合物具有最终浓度为20uM的引物(每种扩增子测序引物10uM)。NextSeq需要额外的测序引物混合物,i5引物混合物,其最终浓度也为20uM。将MiSeq试剂盒v3(150次循环;MS-102-3001)加载在贮存器12中的30uL读段1测序引物混合物和在贮存器13中的30uL i7测序引物混合物。将NextSeq500/550中等产量试剂盒加载在贮存器20中的52uL读段1测序引物混合物、在贮存器22中的85uL i7测序引物混合物和在贮存器22中的85uL i5测序引物混合物。索引1和2各为10bp,并且读段1为26bp。

[0234] 分析:生物信息学分析由使用Illumina的bcl2fastq软件(v2.20.0.422)将BCL文件标准转换为FASTQ测序文件组成。使用定制软件进行每个样品的解复用和读段计数。这里读段1与三个预期的扩增子中的一个匹配,允许扩增子序列中的单核苷酸错误的可能性。汉明距离是对应的序列彼此不同的位置的数目并且是常用的序列间距离的度量。使用两个索引读段对样品进行解复用,以鉴定读段来自哪个样品。观察到的索引读段与预期的索引序列匹配,允许两个索引序列中的一个或两个中单核苷酸错误的可能性。如果索引1和索引2与预期索引序列的汉明距离大于1,则丢弃三个读段的组。计算每个扩增子和每个样品的读段的计数。在此分析中,一些定制脚本是以R写的,它们依赖于ShortRead和stringdist包来处理fastq文件并计算观察到的和预期的扩增子和索引之间的汉明距离。这种方法非常保守并且对测序分析的控制水平非常低。然而,kallisto和bustools SwabSeq分析工具的持续开发可以为其他实施SwabSeq的小组提供更加用户友好且计算效率更高的解决方案。

[0235] 纯化的患者样品的分类标准:对于分析链(pipeline),开发了针对每种类型的样本的QC量度。对于纯化的RNA,每个样品要求针对RPP30检测至少10个读段并且S2和S2合成加标读段的总和超过2,000个读段。如果不满足这些条件,样品将重新运行一次,并且如果第二次未通过,则需要重新取样。为了确定SARS-CoV-2是否存在,计算S2与S2标的比率超过0.003。(在计算此比率之前将1个计数添加到S2和S2标有助于在对数标度上绘制结果。)如果比率大于0.003,则得出结论,所述样品检测到SARS-CoV-2,并且如果比率小于或等于0.003,则得出结论,未检测到SARS-CoV-2(图10,A,B和C图)。

[0236] 同一对引物将同时扩增S和S标扩增子两者。因为运行是终点测定,所以引物将是继续扩增的限制试剂。在开发此测定时,观察到,随着样品的S2计数增加,S2标计数将减少(图8)。在非常高的病毒负载下,S2标读段减少至小于1000个读段。因此,将S2和S2标一起考虑使得即使在极高的病毒滴度下,QC也可以检出SARS-CoV-2。因此,考虑到因为S2扩增子计

数非常高并且样品含有大量SARS-CoV-2RNA而导致S2标计数较低的情况也很重要(图8)。

[0237] 因此,由于S2和S2标来源于同一引物对,所以QC中要求S2和S2标计数的总和超过2000。例如,检测大于2000个S2计数和0个S2标计数,这肯定是SARS-CoV-2阳性样品,并会导致:检测到SARS-CoV-2。

[0238] 索引错误分配的分析:唯一双重指数和扩增子特异性索引用于研究索引错误分配。在此方案中,为每个样品分配了S或标扩增子的两个唯一索引和RPP30扩增子的两个唯一索引,每个样品总计四个唯一索引。具有每个索引对(一个i7和一个i5)的所有可能的成对组合的计数矩阵用于生成匹配矩阵。匹配矩阵对角线上的计数对应于真实样品,而对角线以外的计数对应于索引交换事件。i7和i5索引的索引错误分配的程度是通过分别获取匹配矩阵的非对角元的行和列和来确定的。通过取各孔的病毒S计数与标的比率的平均值来计算在已知零病毒RNA量的情况下那些孔的观察到的索引交换率。

补充结果

[0239] 提高检测限需要最大限度地减少噪声源:运行高灵敏度分子诊断测定中的主要挑战之一是,即使是单一污染物或噪声源也会降低测试的分析灵敏度。在开发SwabSeq的过程中,从不存在SARS-CoV-2RNA的对照样品中观察到S2读段(图4,D图)。这些读段被称为“无模板对照”(NTC)读段。SwabSeq优化的关键部分是了解和最小化NTC读段的来源,以提高测定的检测限(LoD)。鉴定了NTC读段的两种来源:分子污染和错误分配测序读段。

[0240] 为了使分子污染最少,遵循了分子遗传诊断实验室常用的方案和程序。为了限制分子污染,使用了用于稀释合成RNA对照和主混合物的专用罩。在每次新运行开始时,将移液器、稀释溶液和PCR板用10%漂白剂,然后通过UV光处理15分钟来灭菌。

[0241] 为了防止高浓度的PCR后产物污染PCR前过程,将PCR前和PCR后步骤在物理上分为两个单独的房间,其中在PCR前实验室空间内从未打开任何扩增的板。为了进一步防止PCR后污染,比较了具有或不具有尿嘧啶-N-糖基化酶(UNG)的RT-PCR主混合物。TaqPath™1-Step RT-qPCR主混合物(Thermofisher Scientific)中UNG的存在显示与Luna一步RT-PCR混合物(New England Biosciences)相比,减少阴性患者样品中存在的S2读段的PCR后污染的显著改良(图14)。RT-PCR主混合物含有dTTP和dUTP的混合物,使得PCR后扩增子是含有尿嘧啶的DNA。因此,作为先前运行的SwabSeq实验的残余物的这些PCR后可以通过UNG选择性地消除。

[0242] 分子污染的第三种来源是在Illumina MiSeq的测序仪模板线上的滞留污染。在没有维护漂洗的情况下,在不包括前一次测序运行中运行的索引的实验中检测到索引。虽然一些索引的多个读段以多个S2读段存在,但滞留污染的存在影响测定的灵敏度和特异性。在额外的维护和漂洗后,存在的滞留读段的量显著减少至小于10个读段(图15)。

[0243] NTC读段的另一种来源是扩增子的错误分配。当测序(可能以较低速率,寡核苷酸合成)错误得到来源于S2标但被错误分配至给定样品内的S2序列的扩增子序列时,发生扩增子的错误分配。在读段1的开始,仅6bp将S2与S2标区分开。测序错误可能使得S2标读段被错误分类为S2读段,因为在读段开始时错误率似乎更高(图16,A图)。如果扩增子读段的计算错误校正太过宽容,这些读段可能被不注意地计入错误的类别。为了减少这种S2读段错误分配的来源,使用了更保守的编辑距离阈值(图16,B图)。未来对其他病毒扩增子的重新设计或延伸应考虑在此处工程化更长的序列多样性区。

[0244] NTC读段的另一种来源是S2扩增子读段基于索引策略被错误分配至错误的样品的时候。在测定中,单独的样品通过数对索引读数进行鉴定(图4,B图)。由于索引引物序列的污染、索引序列中的合成错误、索引序列中的测序错误或“索引跳跃”,可能发生将样品错误分配至错误的索引。

[0245] 在SwabSeq的开发中利用了多种索引策略,从完全组合索引(其中i5和i7索引中的每个可能的组合用于对测定中的样品加标签)到唯一双重索引(UDI),其中每个样品具有不同且不相关的i7和i5索引(图17)。然而,由于开发许多唯一引物的前期成本很高,因此扩展能力可能受到限制。完全组合索引方法显著扩大了唯一引物集合的数目。还可以使用在完全组合索引与UDI之间的折衷策略,其中各索引组仅在小部分样品之间共享。此类设计减少了样品错误分配的影响,并且有助于扩展到数万个患者样品(图17)。使用完全组合索引(图17),NTC读段深度与跨所有共享同一i7序列的样品求和的S2读段的总数相关(图18,A图)。这与从具有高S2病毒读段的样品到共享相同索引的样品的索引跳跃的效果一致。当阳性样品在索引中随机化时,有可能计算校正这种影响,例如使用线性混合模型。

[0246] 最后,与组合和半组合索引策略相关联的挑战可以通过使用唯一双重索引(UDI)来缓和,这是一种将索引跳跃的读段的数目减少两个数量级的策略。观察到阴性对照样品UDI的S2病毒读段始终较低。它还通过计数不应在测定中出现的索引组合的读段来实现索引错误分配的定量(图21,A和B图)。索引跳跃事件的数目与S2+S2标读段的总数相关(图21,C和D图),指示跳跃读段更有可能来自预期索引具有强病毒信号的孔。在MiSeq上定量了总体条约率为1-2%,已知其在图案化流动池仪器上更高。

[0247] 在基于扩增子的测序中有许多噪声源,从RT-PCR和测序步骤中的环境污染到基于Illumina流动池上的计算校正和“索引跳跃”的读段的错误分配。防止和校正这些错误来源可大大提高SwabSeq测定的检测限。

实施例5-引物浓度对SARS-CoV-2测试的影响

[0248] 对S2和N1扩增子测试了降低的引物浓度的影响。使用唾液(1:1稀释)或鼻拭子,在20 μ L反应中将引物浓度从400nM降低至100nM。RPP30扩增子引物浓度相同,为50nM。降低引物浓度允许保留定量能力。较低的引物浓度还导致较少的引物二聚体和占用测序读段的非特异性扩增产物。在图22中所示的凝胶中,此引物二聚体条带在无模板对照(NTC)泳道中约120bp处。每个凝胶都加载相同量的DNA。在显示用400nM引物进行的RT-PCR的凝胶中,可以看到,在NTC泳道中的120bp处的此引物二聚体在显示以100nM引物进行的RT-PCR的凝胶中亮得多。在扩增反应中使用200nM浓度的引物还显示引物二聚体和非特异性产物的显著改善和减少。

实施例6-多样化的合成序列

[0249] 本文所述的方法因合成核酸序列的存在而得到改进,所述序列充当与样品共同提取、转录和扩增的内部对照序列。

S2合成核酸序列

[0250] 设计了四种合成RNA S2寡核苷酸以增加测序期间的核苷酸碱基密度,从而实现足够的碱基多样性,以确保测序期间无需PhiX即可进行稳健的碱基检出。合成RNA寡核苷酸以等摩尔方式汇集在一起,并且以每个反应250个RNA拷贝的总浓度加标至RT-PCR主混合物中。有几种组合,但发现第1组最好。这些序列来源于野生型S2序列,具有特定碱基对变化,

以等同地表示每个位置处的四个碱基中的每一个(靶向25% A、25% T、25% G、25% C)。实现核苷酸的均匀分布需要序列随机化,导致最终多样化的S2标序列与26bp读段区中的原始模板不同。这些序列经过优化以减少二级结构,并且全部都具有相似的解链温度。侧翼序列与S2扩增子共享同源性。使用此标的LoD一样好,序列如图23所示。用于此分析的引物是正向引物GCTGGTGCTGCAGCTTATTATGTGGGT (SEQ ID NO:18)和反向引物AGGGTCAAGTGCACAGTCTA (SEQ ID NO:19)。

N1合成核酸序列

[0251] 这在使用可以使用N1特异性引物进行扩增的多样化合成序列时也很明显。这些N1序列用2019-nCoV_N1-F,GACCCCAAATCAGCGAAAT (SEQ ID NO:20)和2019-nCoV_N1-R,TCTGGTACTGCCAGTTGAATCTG (SEQ ID NO:21)扩增。

[0252] 初步检测线实验用加标至阴性唾液中并连续稀释的人为样品(使用ATCC热灭活的SARS-CoV-2标准物)进行。这使用三种引物条件进行:N1和S2、仅N1和仅S2。RPP30以最终反应浓度50nM存在,S2为100nM并且N1为100nM。我们仅使用S2的测定的灵敏度被确定为8000GCE/mL。此实验的N1的初步LoD是6000GCE/mL,并且N1+S2的初步LoD是4000GCE/mL,指示添加N1提高了SwabSEQ测定的灵敏度,如图25所示。

[0253] 包含多样化的合成序列使得提高SwabSEQ测试的灵敏度和准确性,同时减少测序过程期间对PhiX的依赖。参见以下表3中的示例性合成序列。

| 表 3.COVID 19 S2 和 N1 合成标序列的序列 | |
|--------------------------------------|--|
| S2_01 | GCTGGTGCTGCAGCTTATTATGTGGGTGTGTATCTCACGA AGCGACCCTTTGGAAAATATAATGAAAATGGAACCATTAC AGATGCTGTAGACTGTGCACTTGACCCT |
| S2_02 | GCTGGTGCTGCAGCTTATTATGTGGGTCCTCGCTAGGACG TCGCTATgacgccAAAATATAATGAAAATGGAACCATTACAG ATGCTGTAGACTGTGCACTTGACCCT |
| S2_03 | GCTGGTGCTGCAGCTTATTATGTGGGTAGCACGACTTGAT CTAACTgacactaAAAATATAATGAAAATGGAACCATTACAG ATGCTGTAGACTGTGCACTTGACCCT |
| S2_04 | GCTGGTGCTGCAGCTTATTATGTGGGTTAAGTAGGACTTC GATTggaTggaatAAAATATAATGAAAATGGAACCATTACAG ATGCTGTAGACTGTGCACTTGACCCT |
| N1_001 | TCTGGTACTGCCAGTTGAATCTGAGGGTCCGGACGGATA TCGCACTAAGTGTACCTGGTGCATTTGCTGATTTGGGGT C |
| N | TCTGGTACTGCCAGTTGAATCTGAGGGTCCCATGACCAT |

| | |
|--------|--|
| 1_002 | GTCACTGGCTACACTGAGGTGCATTTTCGCTGATTTTGGGGTC |
| N1_003 | TCTGGTACTGCCAGTTGAATCTGAGGGTCCTTGACATGGCATGTGACTCCACTGTCGGTGCATTTTCGCTGATTTTGGGGTC |
| N1_004 | TCTGGTACTGCCAGTTGAATCTGAGGGTCCACCTTTGCCAGATGACTGAGTGGAAGGGTGCATTTTCGCTGATTTTGGGGTC |
| N1_005 | TCTGGTACTGCCAGTTGAATCTGAGGGTCCAGCTTGAAGCGTTCGCGACAAGTGTCGGTGCATTTTCGCTGATTTTGGGGTC |
| N1_006 | TCTGGTACTGCCAGTTGAATCTGAGGGTCCTCACGTCCTGAGATCAACTGCTACATGGTGCATTTTCGCTGATTTTGGGGTC |
| N1_007 | TCTGGTACTGCCAGTTGAATCTGAGGGTCCGTTGACTGATCACATGCTGCTCCACGGGTGCATTTTCGCTGATTTTGGGGTC |
| N1_008 | TCTGGTACTGCCAGTTGAATCTGAGGGTCCCAGACAGTCATCGGATTGATGAGTGAGGTGCATTTTCGCTGATTTTGGGGTC |

实施例7-SwabSeq可以检测多个不同的SARS-CoV-2扩增子

[0254] SwabSeq能够检测两种SARS-CoV-2基因：N1和S2以及合成N1和S2对照，提高了检出阳性样品的能力。每个扩增子的数据示出于图24中。

N1 F:CAAGCAGAAGACGGCATAACGAGAT XXXXXXXXXXXX ACCCCAAAATCAGCGAAAT (SEQ ID NO:22); N1 R:AATGATACGGCGACCACCGAGATCTACAC XXXXXXXXXXXX TCTGGTACTGCCAGTTGAATCTG (SEQ ID NO:23); X表示每个引物的唯一双重索引 (UDI)。

实施例8-使用不同样品的SwabSeq性能

[0255] 从最少或少量样品开始工作的能力对于测试很重要，尤其是在可以使用侵入性较小的唾液样本的情况下。图26显示了在文库制备期间可以使用多达10uL唾液样品而不必抑制RT或PCR，并且图27显示在文库制备期间可以使用多达6uL鼻拭子样品（鼻拭子接种到1mL 1X PBS中）而不必抑制RT或PCR。

实施例9-SwabSeq可用于检测流感病毒和SARS-CoV-2病毒

[0256] 可以同时检测多种病毒的测序和检测系统和平台具有优势。进行了共同检测SARS-CoV-2以及甲型流感病毒和乙型流感病毒的实验。

[0257] 具体而言，SwabSeq能够鉴定FluA MP、FluA NS和FluB NS的靶标扩增子。基因组序

列是从NCBI流感数据库下载的。然后将初步引物与这些序列比对,并提取这些基因组中每一个的靶标扩增区。然后使用此信息进一步优化引物以尽可能多地靶向这些扩增子。然后导出所有唯一扩增区并添加到靶标列表中。图28中示出的结果指示SwabSeq可以检测人为样品中的甲型流感病毒或乙型流感病毒。

[0258] 为了对其他呼吸道病原体实施Swabseq,设计了通用的甲型流感病毒和乙型流感病毒Swabseq引物。对于甲型流感病毒,测试了2个靶向甲型流感病毒的引物对,M1 (FluA_5_f1GACCAATCYTGTCACCTCTGAC (SEQ ID NO:24)、FluA_5_f2GACCAATYCTGTCACCTYTGAC (SEQ ID NO:25)、FluA_5_rAGGGCATTGAYAAAGCGTCTA (SEQ ID NO:26)) 0和NS1 (FluA_12_fTTGGGTCTCATCGGAG (SEQ ID NO:27)、FluA_12_rTTCTCCAAGCGAATCTCTGTA (SEQ ID NO:28)) 基因。

对于乙型流感病毒,对NS1 (FluB_11_f AAGATGGCCATCGGATCC (SEQ ID NO:29) FluB_11_r GTCTCCCTCTTCTGGTGATAATC (SEQ ID NO:30)) 基因具有特异性的引物对。

合成核酸标序列如下所示:

FluA_5_标_004

GACCGATCCTGTCACCTCTGACTAAGGGTGGCACAACGCAGTGTGTTGAGCTCCCTTAGTAGGGTGAC TAGCACCGGCAGCGTAGACGCTTTGTCCAAAATGCCCT (SEQ ID NO:31)

FluA_12_标_002

TTGGGTCTCATCGGAGGACGCTTAAATAATAGTAGGAACGTTTCGAGTCTCTAAAAATATACAGAGA TTCGCTTGAGAA (SEQ ID NO:32)

FluB_11_标_004

AAGATGGCCATCGGATCCTCAACTCACTCCTGTAATGAGTCGTAGACAAGGATAAGAGGCCCGATCGG CAGTATTTGCGCCCTTTCTAAAAATAATGTGACCTGGGACGCACTGCACCGATTATCACCAGAAGAGGGAGAC (SEQ ID NO:33)

实施例10-S2引物的包容性和交叉反应性

[0259] 鉴于SARS-CoV-2的突变率高,重要的是用于测序测试的引物能够扩增许多不同的变体。对S2引物集和S2扩增子序列进行计算机分析,以评估SwabSeq测试的包容性。对于引物分析,评估了在GISAID上可用的SARS-CoV-2序列。我们首先过滤掉低质量基因组,我们将它们定义为具有大于或等于1%的未鉴定核苷酸(N)。对剩余的324,355个高质量基因组进行了BLASTn (NCBI) 分析,以通过针对下载的SARS-CoV-2序列查询两个SARS-CoV-2引物序列中的每一个来定量这些序列中引物同源性的水平。分析表明,所有分析毒株的91.24%与引物序列具有100%的同源性,并且324,355个完整基因组中的28,398个毒株(或8.76%)与引物具有低于100%的同源性。对于正向S2引物,303,840个(93.67%) GISAID基因组具有100%同源性。对于反向S2引物,305,019个(94.04%) GISAID基因组具有100%同源性。S2扩增子是26个碱基对,并且303,934(93.70%) GISAID基因组具有100%同源性。

[0260] 进行了计算机分析以评估SwabSeq引物与代表性常见呼吸道病原体的交叉反应性。对于引物分析,从NCBI下载了38个非SARS-CoV-2共有基因组作为阴性样品队列。然后进行BLASTn (NCBI) 分析以定量与队列中的每个基因组具有多于80%同源性的引物对的数目。没有病原体表现出与任何引物大于80%的同源性。

实施例11-测序数据的分析

[0261] 本文呈现了确定病原体感染的分析平台的实施例,所述平台可以与SwabSeq的提取、扩增和测序步骤组合。

[0262] 首先,生成了所有靶向扩增子和条形码索引的所有单核苷酸取代、缺失和插入。然后将它们用作嵌套数据结构中的键,所述嵌套数据结构包含原始序列、靶标名称和匹配类型(精确匹配、错配、缺失、插入)。由于此数据结构中的键是索引的,因此可以对测序数据进行精确的字符串匹配,并在预定的匹配容差(单碱基取代、缺失和插入)内快速检索靶标信息。所有重复的序列都被删除,并且将剩余的唯一序列的匹配类型重新分类为“未确定”。

[0263] 然后从FASTQ文件中提取测序数据,并与其对应的词典(索引1、索引2或扩增子)进行匹配。在匹配之前,读入仪器和试剂盒信息以确定哪些序列需要在匹配之前进行反向互补。

[0264] 匹配函数返回原始靶标序列(用于索引)或靶标名称(用于扩增子)。这是关键的一步,因为许多靶标,包括多样化的标和流感病毒靶标都有几个需要聚集在一起的序列。还可以返回其他数据,诸如匹配类型。

[0265] 然后将匹配的序列按其索引和扩增子靶标聚集,并返回每个索引对的总读段的表。在下游使用此信息以实现质量控制(可用于确定残留污染、索引跳跃、扩增失败等)以及是否存在COVID-19和甲型流感病毒/乙型流感病毒。

[0266] 此信息以自动生成的PDF的形式返回,其含有质量控制曲线图以及具有关于每个样品的信息的表格,所述信息诸如QC通过/未通过、COVID-19阳性/阴性、甲型流感病毒/乙型流感病毒阳性/阴性。

[0267] 基于词典的扩增子和条形码的精确匹配在不依赖现有生物信息学分析库的情况下实现快速分析,同时考虑所有单碱基缺失、插入和错配。

[0268] 用于在使用S2引物、S2合成核酸和RPP30引物时检出阳性样品的算法的实例显示在以下表4和图29中并且需要RPP30扩增子的读段的绝对数目、S2+S2标的绝对数目以及大于0.1的S2/S2标的比率。

| 孔对照 | | 结果 | | | |
|-----------|-------------|-----------|-----------------|--------------------------|-------------------------------|
| 总 S2+S2 标 | RPP30 读段设计数 | S2/S2 标比率 | 结果 | 判读 | 动作 |
| >100 | >10 | >0.1 | SARS-CoV-2 检测到 | 所述样品 ID 对 SARS-CoV-2 呈阳性 | 向医生、患者和适当的公共健康管理机构报告结果。 |
| >100 | >10 | <0.1 | SARS-CoV-2 未检测到 | 所述样品 ID 对 SARS-CoV-2 呈阴性 | 向医生、患者和适当的公共健康管理机构报告结果。 |
| <100 | >10 | - | 不确定 | 所述样品 ID 无效 | 所述样品 ID 的质量控制为未通过。重复取样或重新收集样品 |
| <100 | <10 | - | 不确定 | 所述样品 ID 无效 | 所述样品 ID 的质量控制为未通过。重复取样或重新收集样品 |

实施例12-检测限

[0269] 通过使用稀释系列在阴性唾液样本中加标热灭活的SARS-CoV-2病毒(ATCC, VR-1986)来进行检测限(LoD)研究。每个浓度进行十一次提取重复。初步LoD定义为11次重复中的11次测试为阳性的最低浓度。LoD确定针对具有UDG的Luna[®]探针一步RT-qPCR 4X混合物以及TaqPath[™]1步RT-qPCR主混合物独立进行。对于两种RT-qPCR主混合物,初步LoD确定为8,000GCE/ml。结果在表5中示出。

表5. 初步LoD确定结果

| 病毒浓度 (GCE/ml) | 检测率 (%) Luna | 检测率 (%) Taqpath |
|---------------|--------------|-----------------|
| 24,000 | 11/11 (100%) | 11/11 (100%) |
| 20,000 | 11/11 (100%) | 11/11 (100%) |
| 16,000 | 11/11 (100%) | 11/11 (100%) |
| 12,000 | 11/11 (100%) | 11/11 (100%) |
| 8,000 | 11/11 (100%) | 11/11 (100%) |
| 6,000 | 11/11 (100%) | 10/11 (90.9%) |
| 4,000 | 6/11 (54.5%) | 6/11 (54.5%) |
| 2,000 | 1/11 (9.1%) | 1/11 (9.1%) |

[0270] 具有UDG的 **Luna**[®]探针一步RT-qPCR 4X混合物和TaqPath[™]1步RT-qPCR主混合物的检测限确认研究通过对应于1.5X和1X确定的检测限以12,000GCE/mL和8,000GCE/mL的最终浓度在阴性唾液样本中加标热灭活的SARS-CoV-2病毒 (ATCC, VR-1986) 来进行。每个浓度进行十 (10) 次提取重复。在这两个浓度下的LoD确认针对具有UDG的 **Luna**[®]探针一步RT-qPCR 4X混合物以及TaqPath[™]1步RT-qPCR主混合物独立进行。对于具有UDG的 **Luna**[®]探针一步RT-qPCR 4X混合物和TaqPath[™]1步RT-qPCR主混合物,两种RT-qPCR主混合物的最小LoD确认为8,000GCE/mL。

实施例13-临床评估

[0271] 进行了一项研究,以评估SwabSeq的性能,比较了先前使用实验室开发的测试 (LDT) 分析并确认为阳性或阴性的唾液临床样品。

[0272] 使用MiniSeq Illumina测序仪收集并运行唾液样品。表6汇总了使用具有UDG的Luna探针一步RT-qPCR 4x混合物的结果。使用具有UDG的Luna探针一步RT-qPCR 4x混合物,阳性符合率 (25/25) 为100%,并且阴性符合率 (93/93) 为100%。使用Taqpath 1步RT-qPCR主混合物的结果汇总在表7中。阳性符合率 (25/25) 为100%,并且阴性符合率 (93/93) 为100%。

表6. 使用具有UDG的 **Luna**[®]探针一步RT-qPCR 4X混合物以临床样本的评估

| | | LDT 比较测定 | | |
|--|-----|---|----|-----|
| | | 阳性 | 阴性 | 总计 |
| Ginkgo SARS-CoV-2 NGS 测试 (Octant v1) – Luna | 阳性 | 25 | 0 | 25 |
| | 阴性 | 0 | 93 | 93 |
| | 不确定 | 0 | 0 | 0 |
| | 总计 | 25 | 93 | 118 |
| 阳性符合率 | | 100% (25/25) ; | | |
| | | 82.4%-100% ¹ | | |
| 阴性符合率 | | 100% (93/93) ; 96.1%-100% ¹ | | |
| 总体符合率 | | 100% (118/118) ; 96.9%-100% ¹ | | |

¹双边95%分数置信区间

表7. 使用TaqPath™1步RT-qPCR主混合物以临床样本的评估

| | | LDT 比较测定 | | |
|---|-----|---|----|-----|
| | | 阳性 | 阴性 | 总计 |
| Ginkgo SARS-CoV-2 NGS 测试 (Octant v1) – Taqpath | 阳性 | 25 | 0 | 25 |
| | 阴性 | 0 | 93 | 93 |
| | 不确定 | 0 | 0 | 0 |
| | 总计 | 25 | 93 | 118 |
| 阳性符合率 | | 100% (25/25) ; 82.4%-100% ¹ | | |
| 阴性符合率 | | 100% (93/93) ; 96.1%-100% ¹ | | |
| 总体符合率 | | 100% (118/118) ; 96.9%-100% ¹ | | |

¹双边95%分数置信区间

[0273] 第二天,使用相同的临床样品进行了中间精密度研究,以评估SwabSeq的可靠性。第二天使用的进行测试的人员和测试中使用的仪器(Viaflo-96、手动移液器组、MiniSeq测序仪、热循环仪)全部都与第一天不同。然后比较先前使用实验室开发的测试(LDT)分析并确认为阳性或阴性的唾液临床样品。

[0274] 使用SwabSeq的结果汇总在表8和表9中。表8汇总了使用具有UDG的Luna探针一步RT-qPCR 4x混合物的临床样品。阳性符合率(25/25)为100%,并且阴性符合率(93/93)为100%,并且两个中间精密度运行的一致性为100%。使用Taqpath 1步RT-qPCR主混合物的临床样品汇总在表9中。阳性符合率(25/25)为100%,并且阴性符合率(93/93)为100%,并且两个中间精密度运行的一致性为100%。

表8. 使用具有UDG的 **Luna**[®] 探针一步RT-qPCR 4X混合物以临床样本的评估

| | | LDT 比较测定 | | |
|--|-----|---|----|-----|
| | | 阳性 | 阴性 | 总计 |
| Ginkgo SARS-CoV-2 NGS 测试 (Octant v1) – Luna | 阳性 | 25 | 0 | 25 |
| | 阴性 | 0 | 93 | 93 |
| | 不确定 | 0 | 0 | 0 |
| | 总计 | 25 | 93 | 118 |
| 阳性符合率 | | 100% (25/25) ; 82.4%-100% ¹ | | |
| 阴性符合率 | | 100% (93/93) ; 96.1%-100% ¹ | | |
| 总体符合率 | | 100% (118/118) ; 96.9%-100% ¹ | | |

¹双边95%分数置信区间

表9. 使用TaqPath™1步RT-qPCR主混合物以临床样本的评估

| | | LDT 比较测定 | | |
|---|-----|---|----|-----|
| | | 阳性 | 阴性 | 总计 |
| Ginkgo SARS-CoV-2 NGS 测试 (Octant v1) – Taqpath | 阳性 | 25 | 0 | 25 |
| | 阴性 | 0 | 93 | 93 |
| | 不确定 | 0 | 0 | 0 |
| | 总计 | 25 | 93 | 118 |
| 阳性符合率 | | 100% (25/25) ; 82.4%-100% ¹ | | |
| 阴性符合率 | | 100% (93/93) ; 96.1%-100% ¹ | | |
| 总体符合率 | | 100% (118/118) ; 96.9%-100% ¹ | | |

¹双边95%分数置信区间

[0275] 使用SwabSeq进行的两次中间精密度运行比较了先前分析并确认为阳性或阴性的唾液临床样本,显示100%一致性。

实施例14-精密度

[0276] 为了评估运行内和运行之间的精密度,在MiniSeq测序仪上的3次运行中以两个人为浓度病毒/mL进行复制:12,000个拷贝/mL和24,000个拷贝/mL。重复此过程,使得同时使用具有UDG的Luna[®]探针一步RT-qPCR 4X混合物(表10)和TaqPath[™]1步RT-qPCR主混合物(表11)。这些浓度表示我们的测定的检测限的1.5X和3X。据观察,我们使用Taqpath主混合物进行的第三次精密度运行的符合率不为100%(参见表11)。有时观察到,如果不快速处理,由加标至阴性对照唾液的病毒RNA SARS-CoV-2标准物组成的人为样品可能降解。Luna和Taqpath的第四次精密度运行,其中使用人为样品的运行没有在室温下静置这么长时间,并且发现在第四次运行中有100%符合率。

表10. 使用具有UDG的Luna[®]探针一步RT-qPCR 4X混合物的人为样品的精密度

| | 检测率% Luna | | | | 运行之间的一致性 |
|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 拷贝/mL | 运行 1 | 运行 2 | 运行 3 | 运行 4 | |
| 12,000 | 6/6 (100%) | 6/6 (100%) | 6/6 (100%) | 6/6 (100%) | 24/24 (100%) |
| 24,000 | 6/6 (100%) | 6/6 (100%) | 6/6 (100%) | 6/6 (100%) | 24/24 (100%) |
| 运行中的一致性 | 12/12 (100%) | 12/12 (100%) | 12/12 (100%) | 12/12 (100%) | |

表11. 使用TaqPath™1步RT-qPCR主混合物的人为样品的精密度

| | 检测率% Taqpath | | | | 运行之间的一致性 |
|----------------|-----------------|-----------------|----------------|-----------------|------------------|
| 拷贝/mL | 运行 1 | 运行 2 | 运行 3 | 运行 4 | |
| 12,000 | 6/6 (100%) | 6/6 (100%) | 5/6 (83.3%) | 6/6 (100%) | 23/24 (95.8%) |
| 24,000 | 6/6 (100%) | 6/6 (100%) | 4/6 (66.6%) | 6/6 (100%) | 22/24 (91.6%) |
| 运行中的一致性 | 12/12 (100%) | 12/12 (100%) | 9/12 (75%) | 12/12 (100%) | |

[0277] 尽管已经在本文中示出和描述了本发明的优选实施方案,但是对本领域技术人员显而易见的是,此类实施方案仅通过举例方式提供。在不脱离本发明的情况下,本领域技术人员现在将想到许多变化、改变和替换。应理解,本文所述的本发明的实施方案的各种替代方案可以用于实践本发明。

[0278] 本说明书中提及的所有出版物、专利申请、授权专利和其他文件都以引用的方式并入本文,如同具体且单独地指示每个单独的出版物、专利申请、授权专利或其他文件以引用的方式整体并入一样。若以引用的方式并入的文本中所包含的定义与本公开中的定义相抵触,则将其排除。

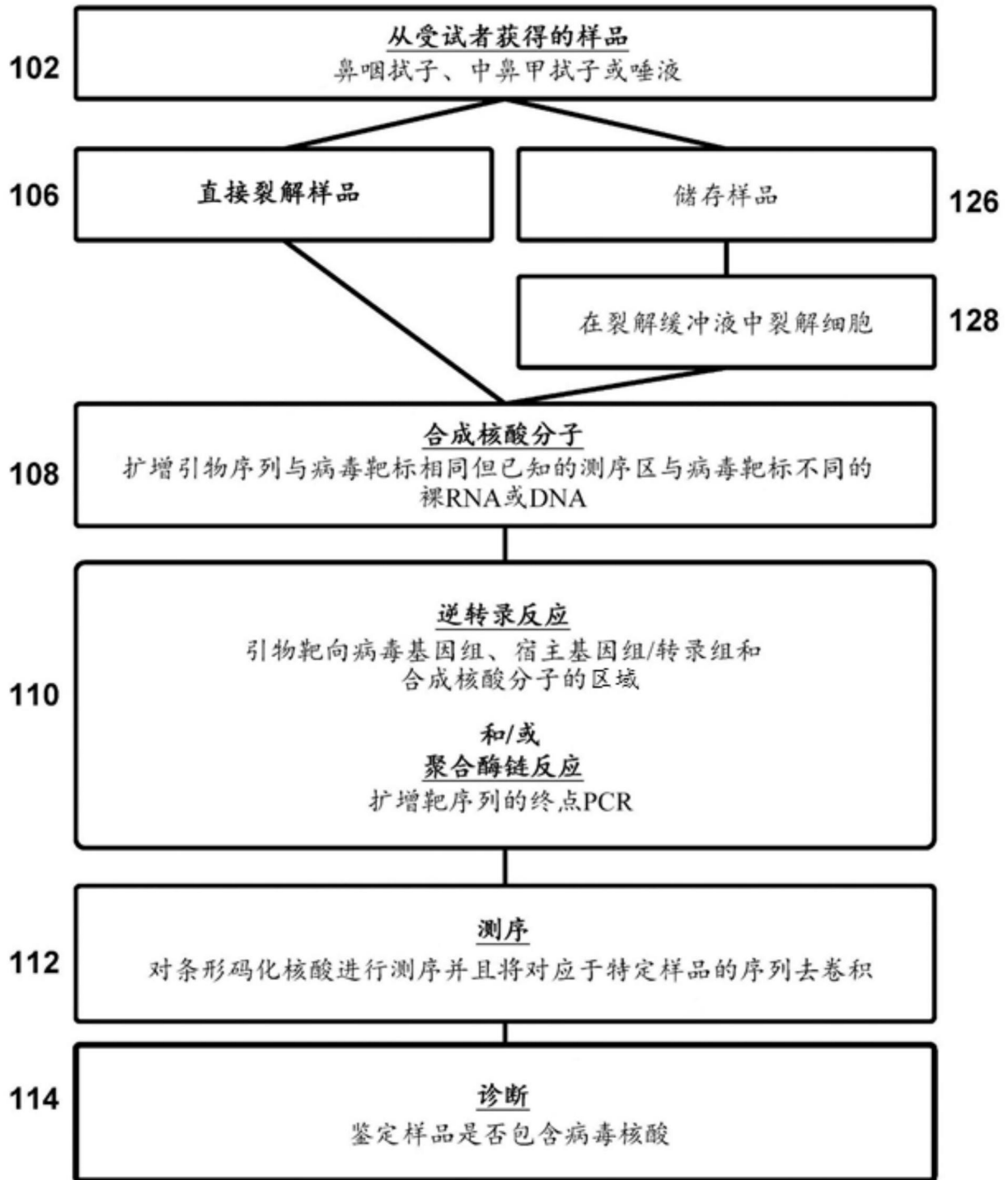


图1

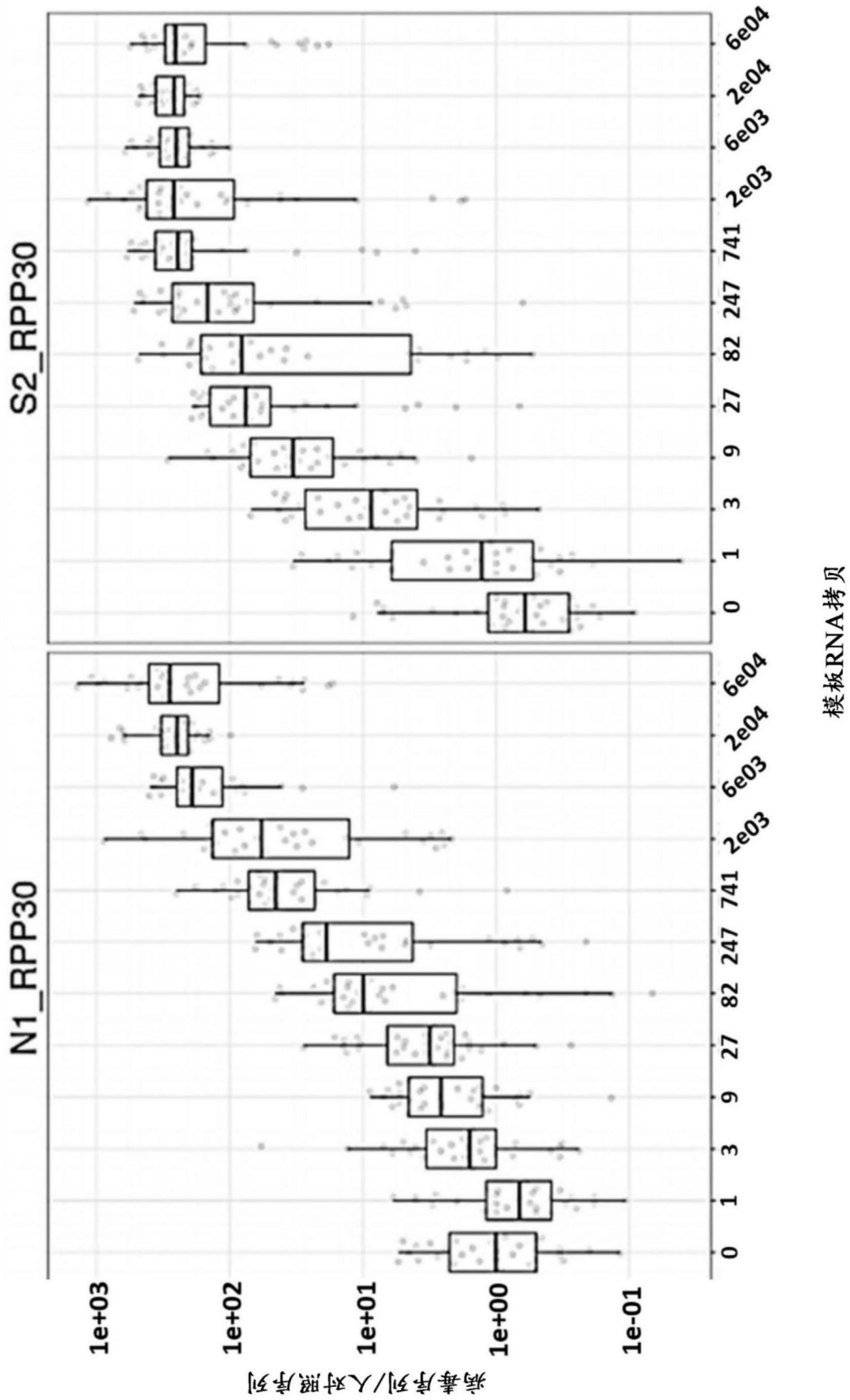


图2A

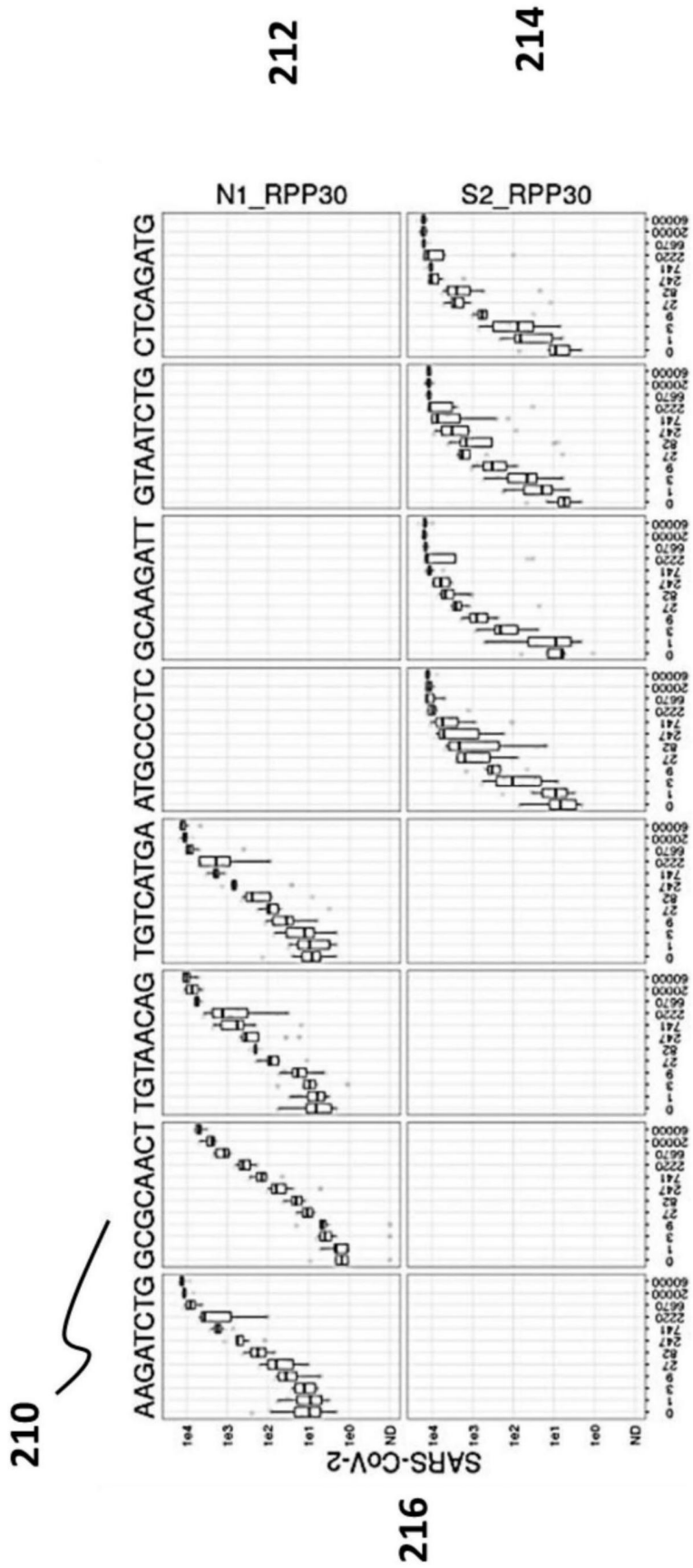


图2B

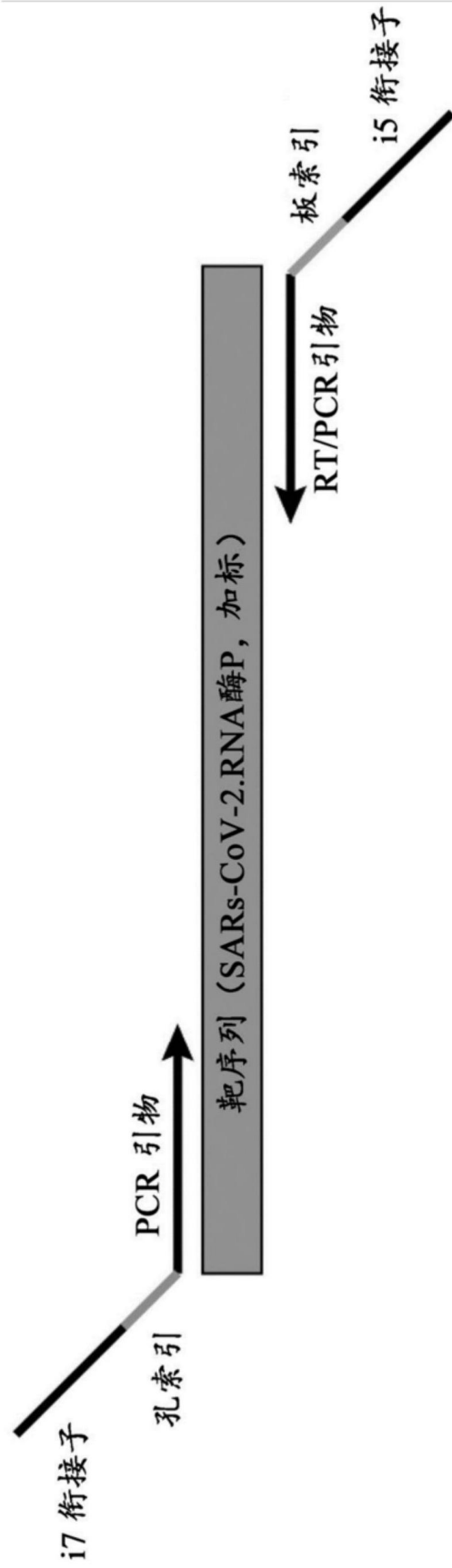


图3

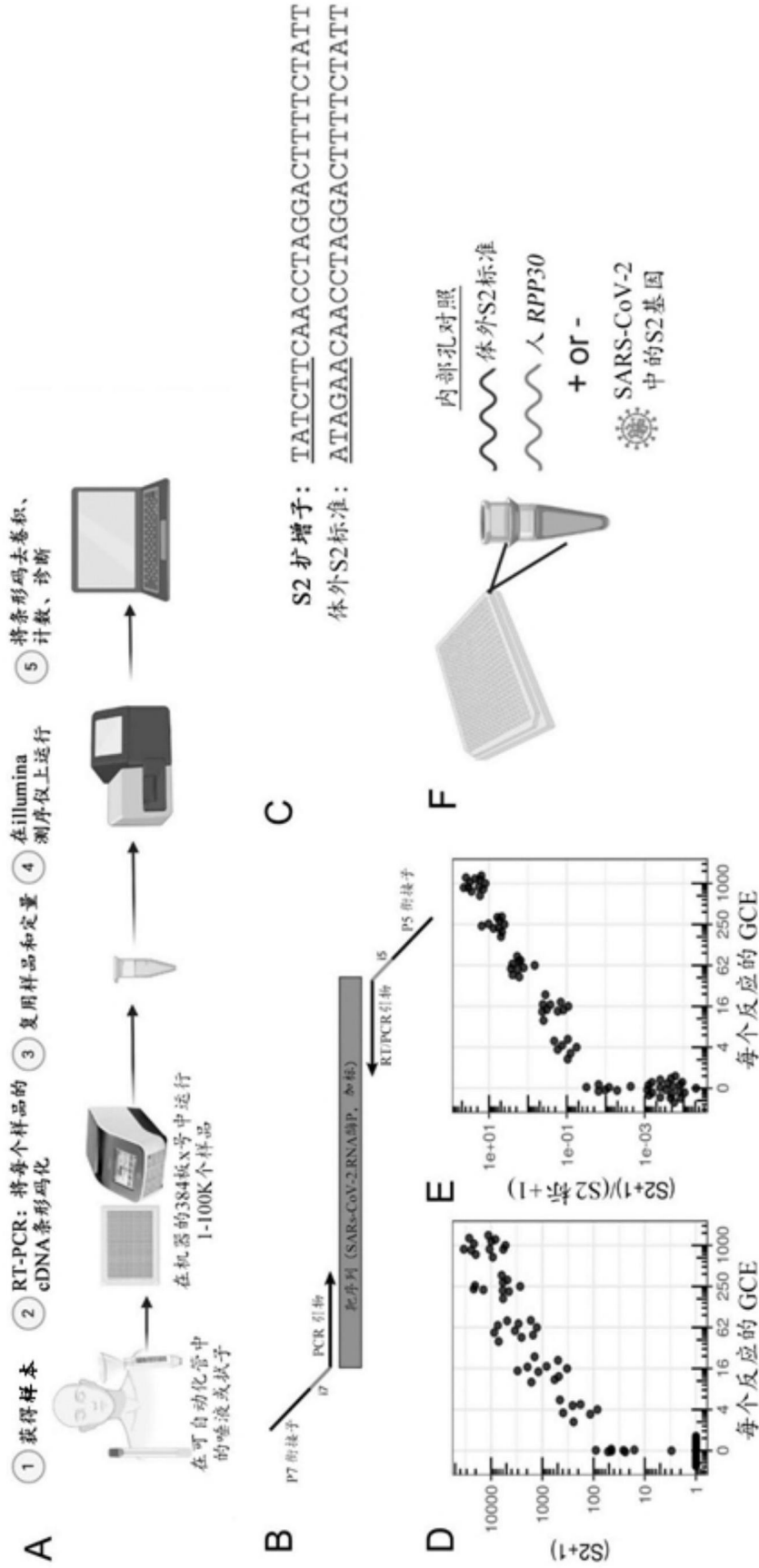


图4

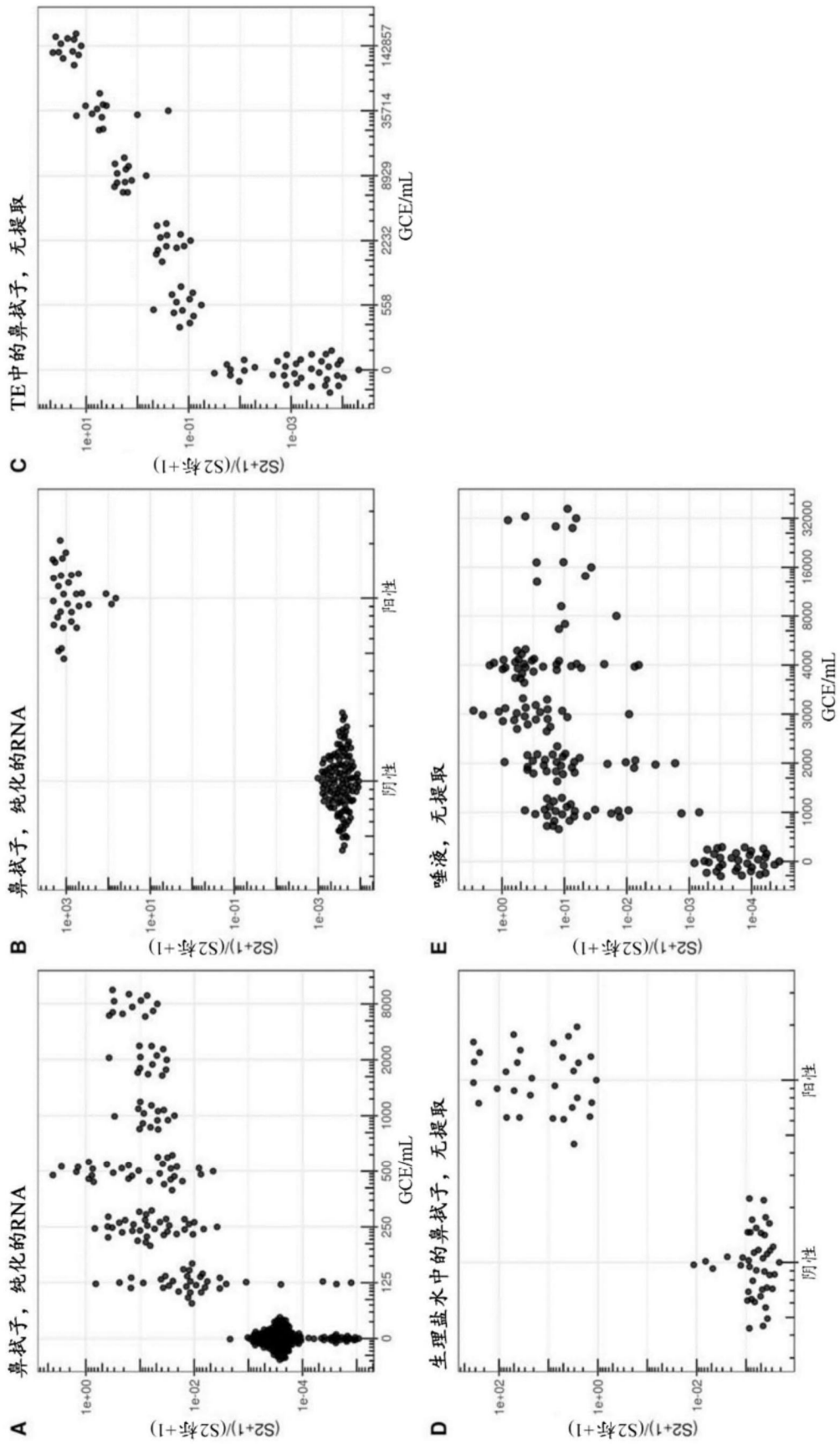


图5

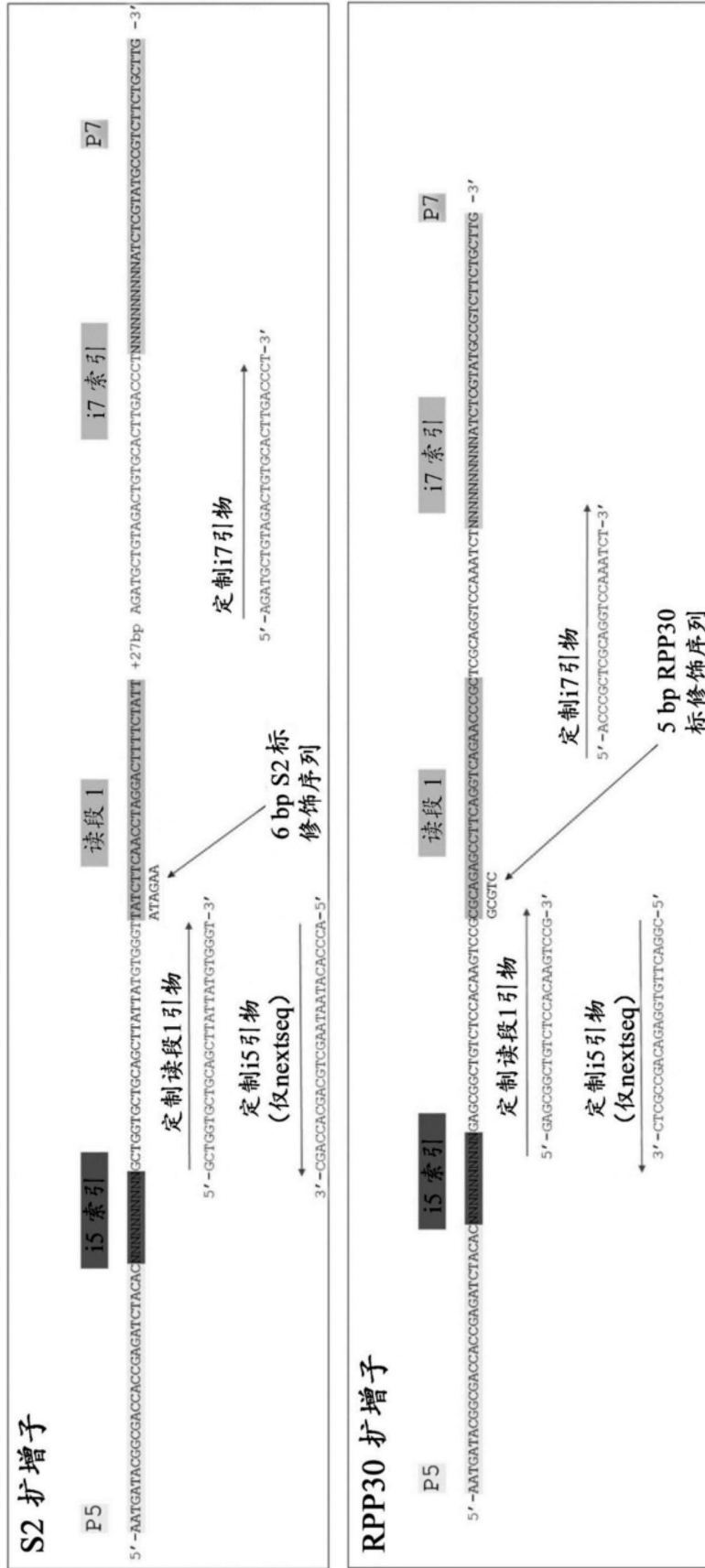


图6

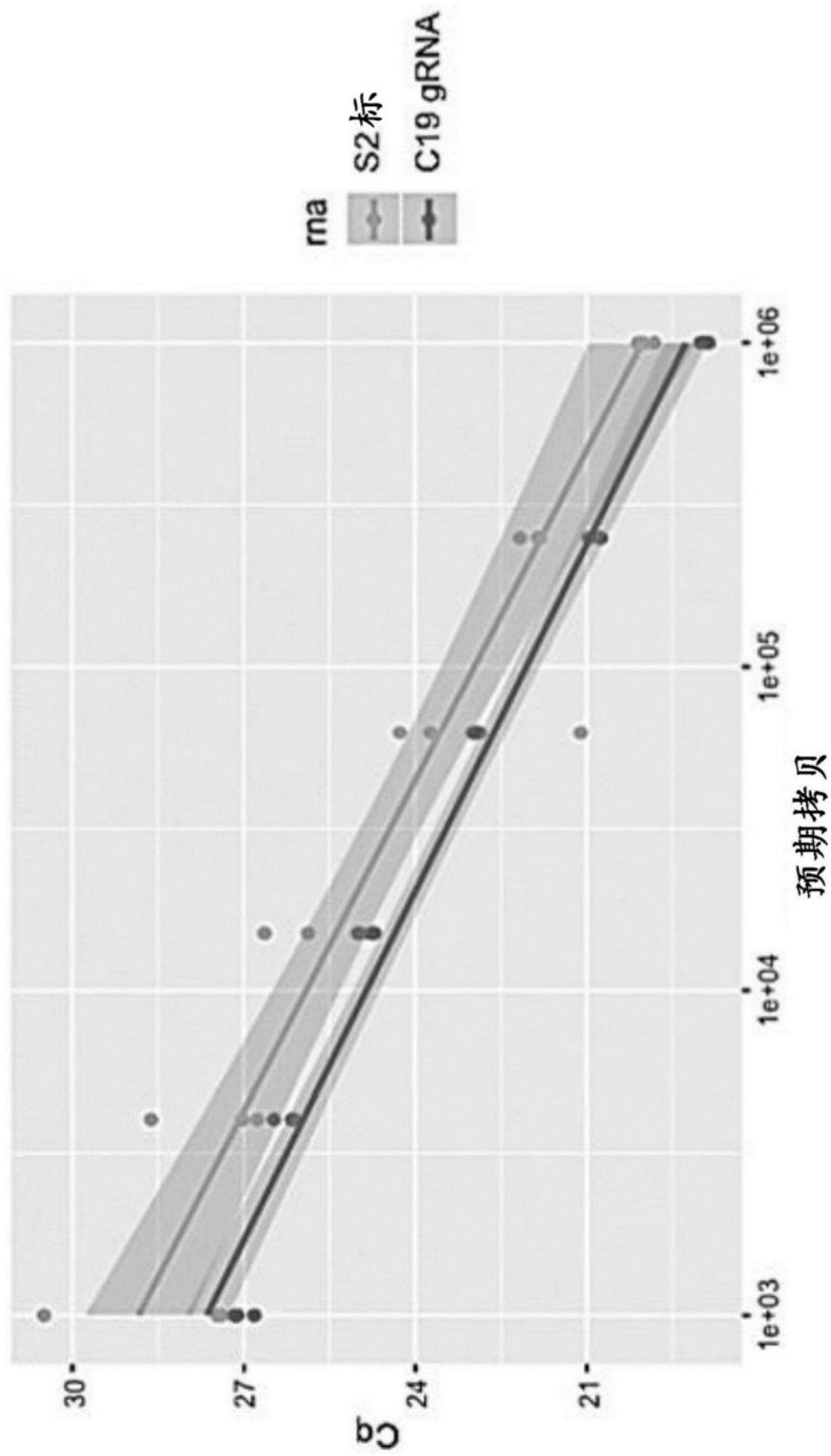


图7

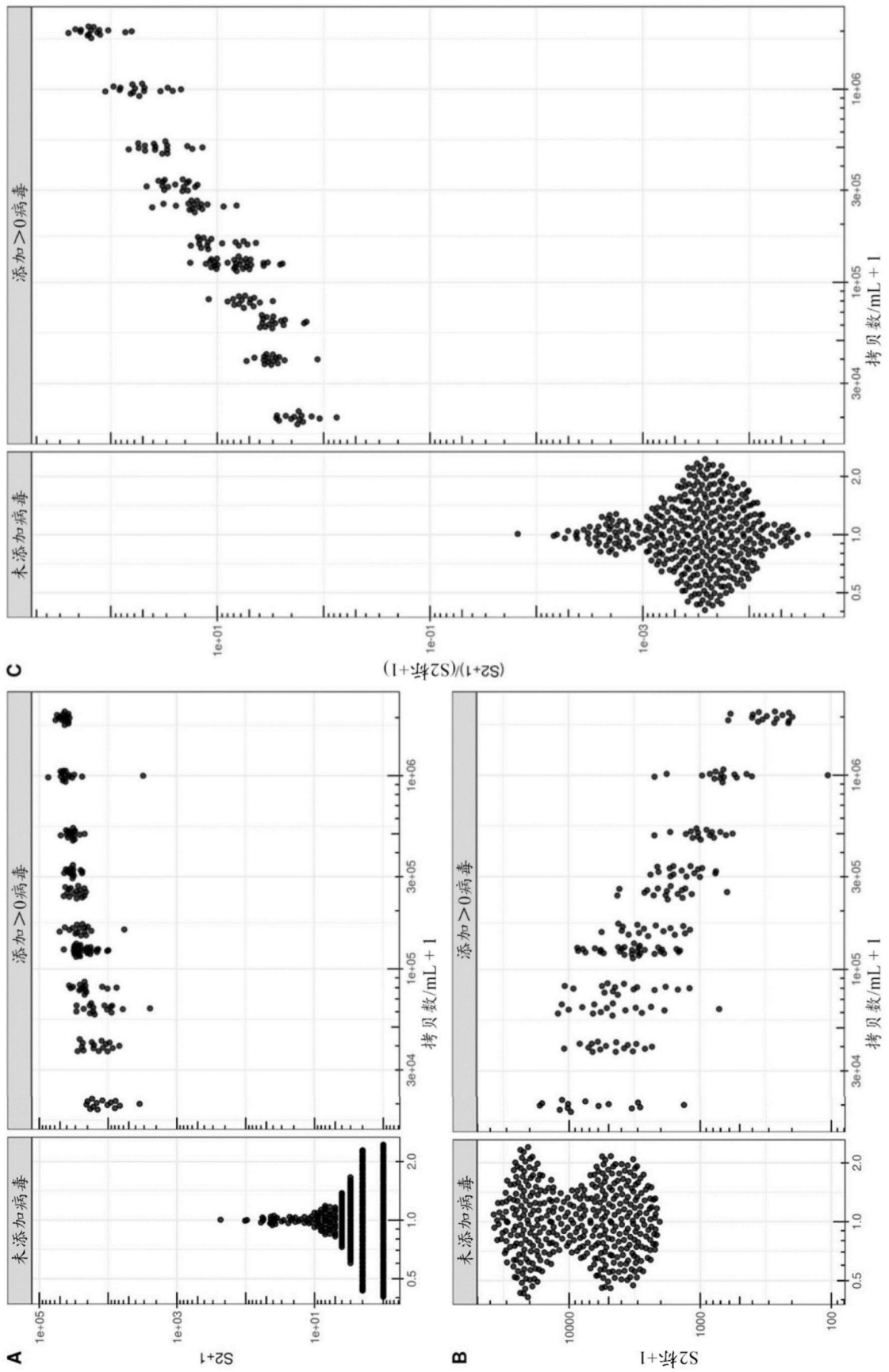


图8

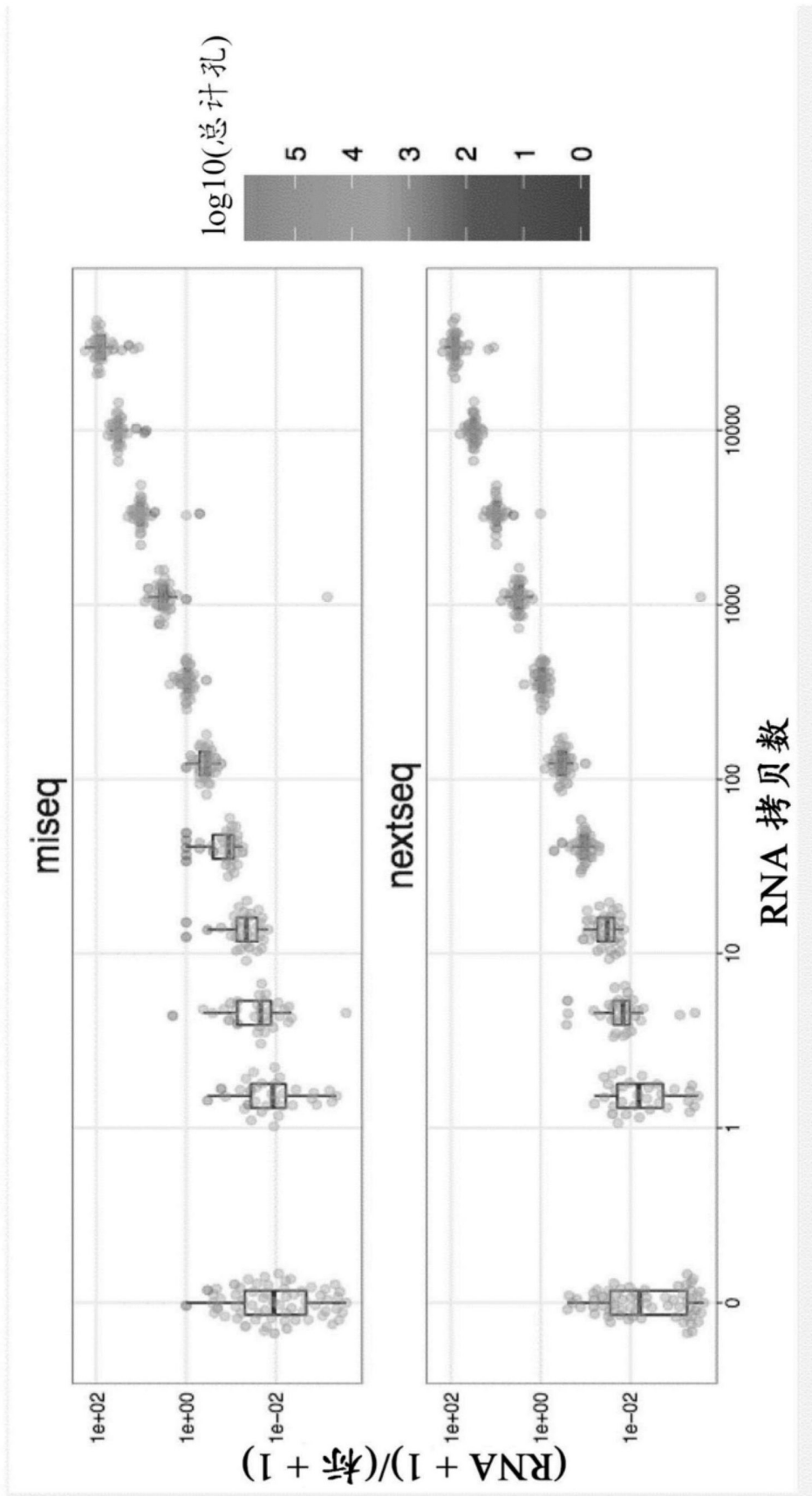


图9

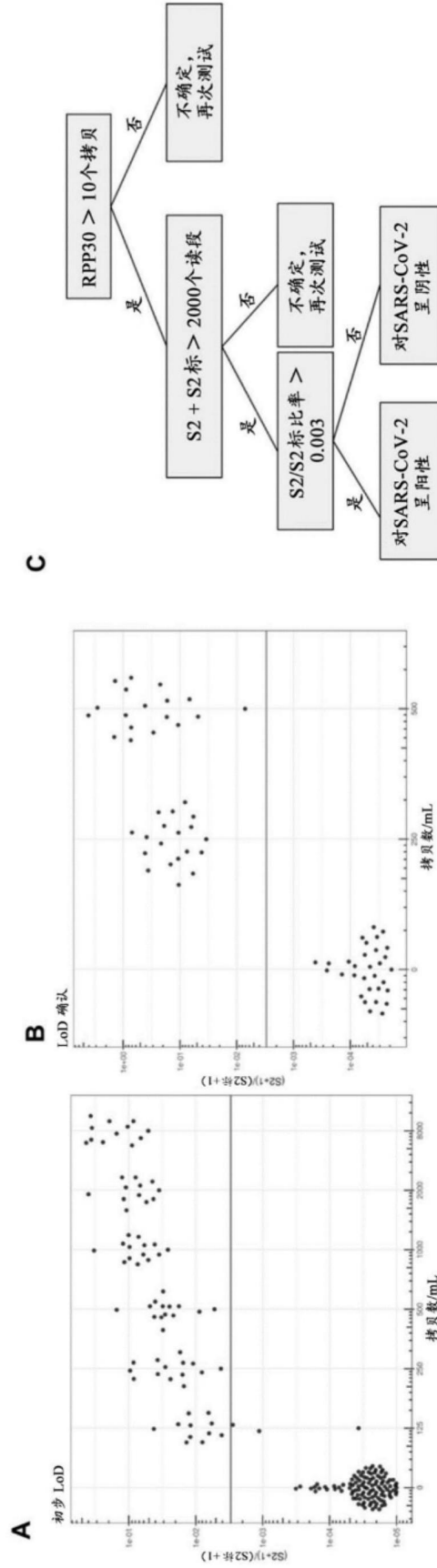


图10

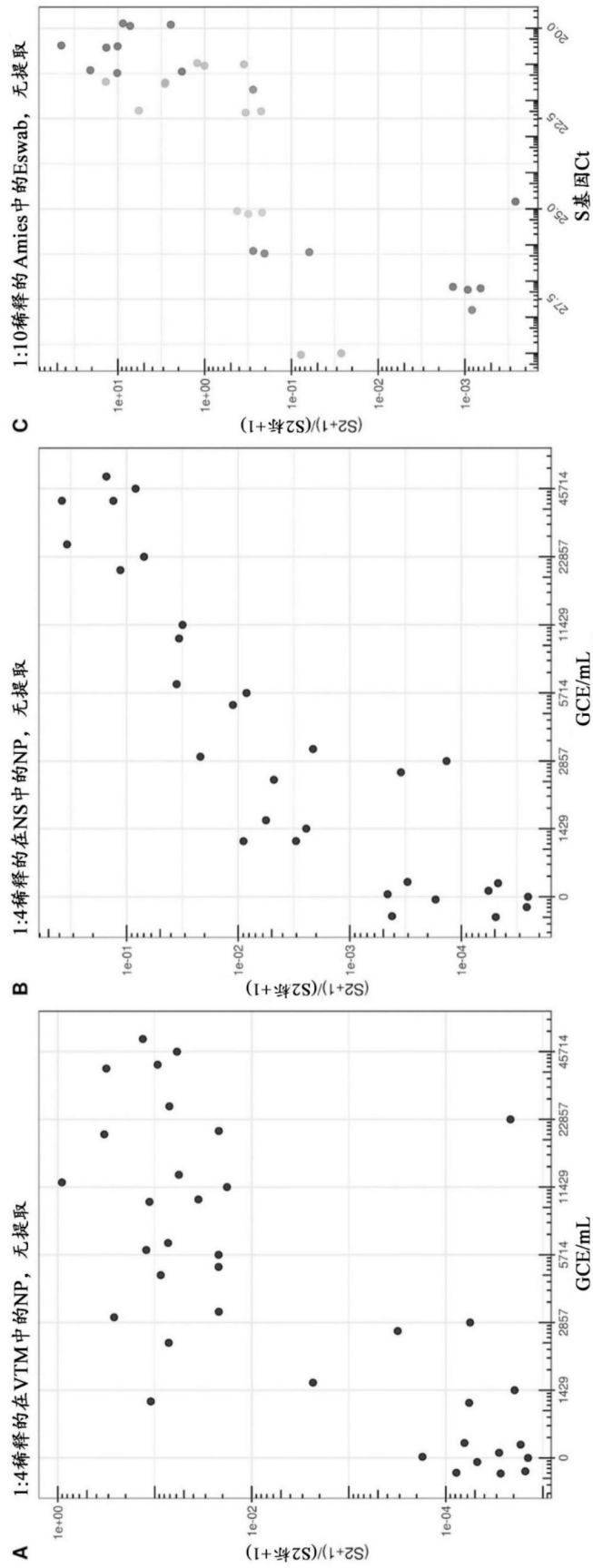


图11

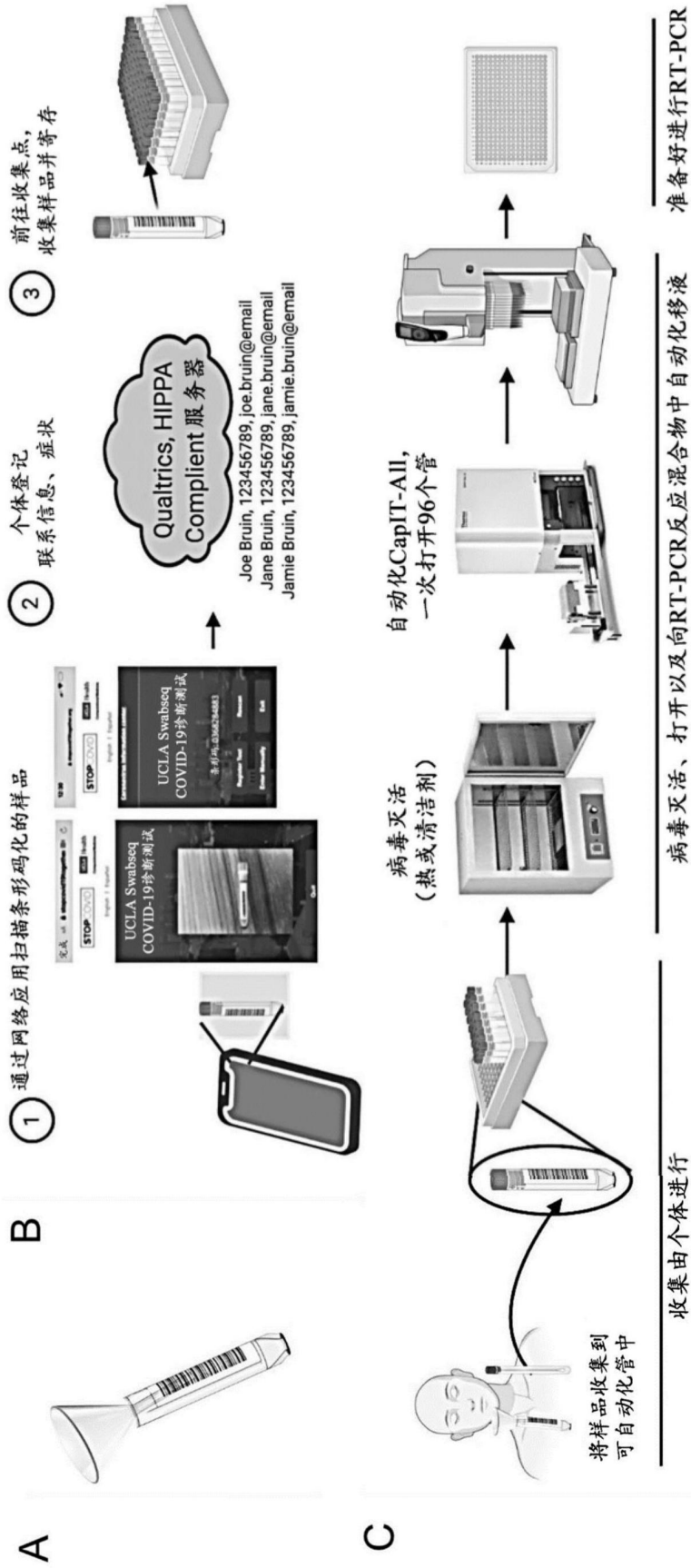


图12

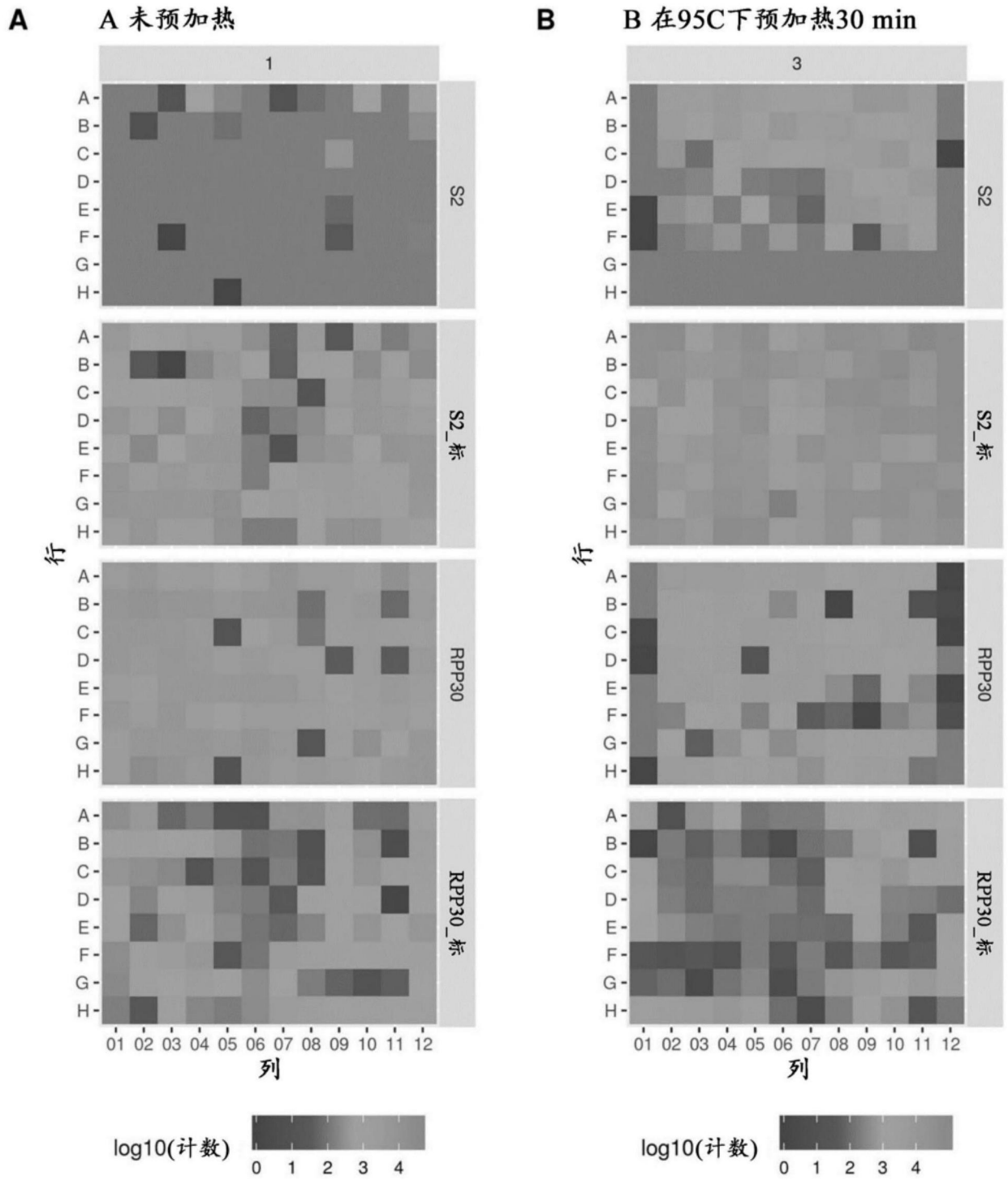


图13

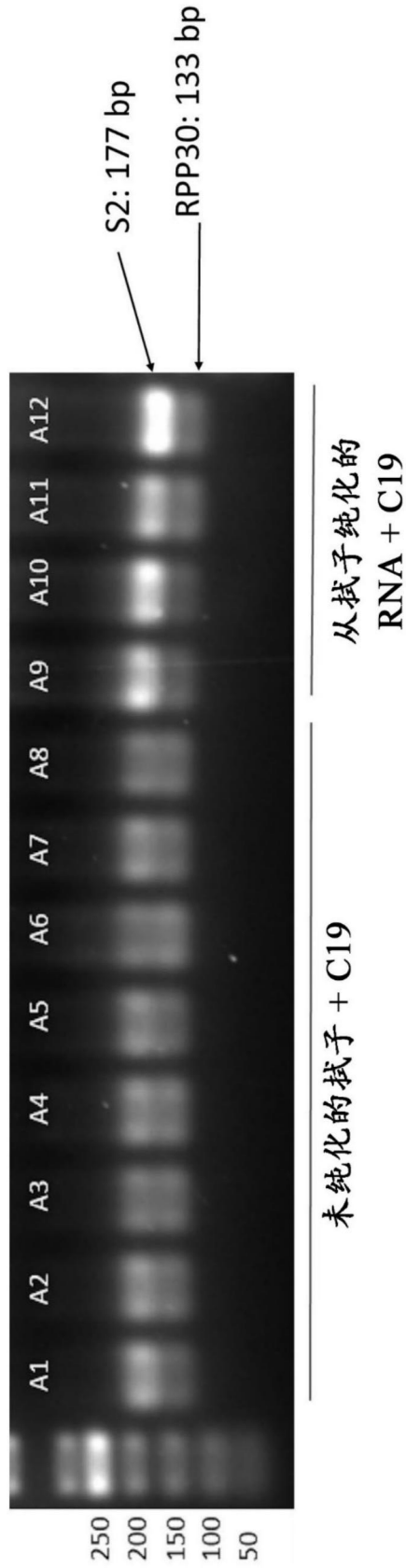


图14

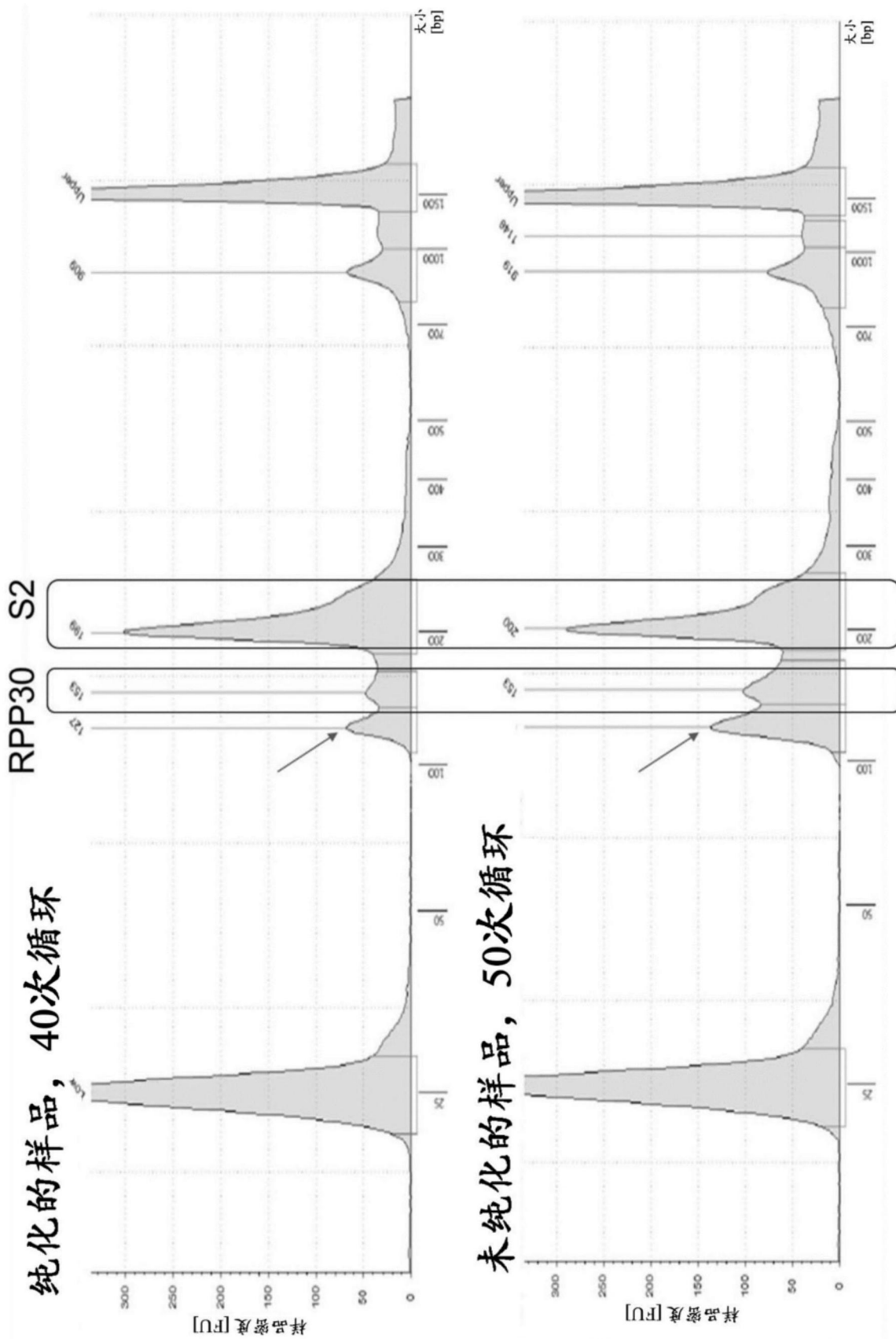


图15

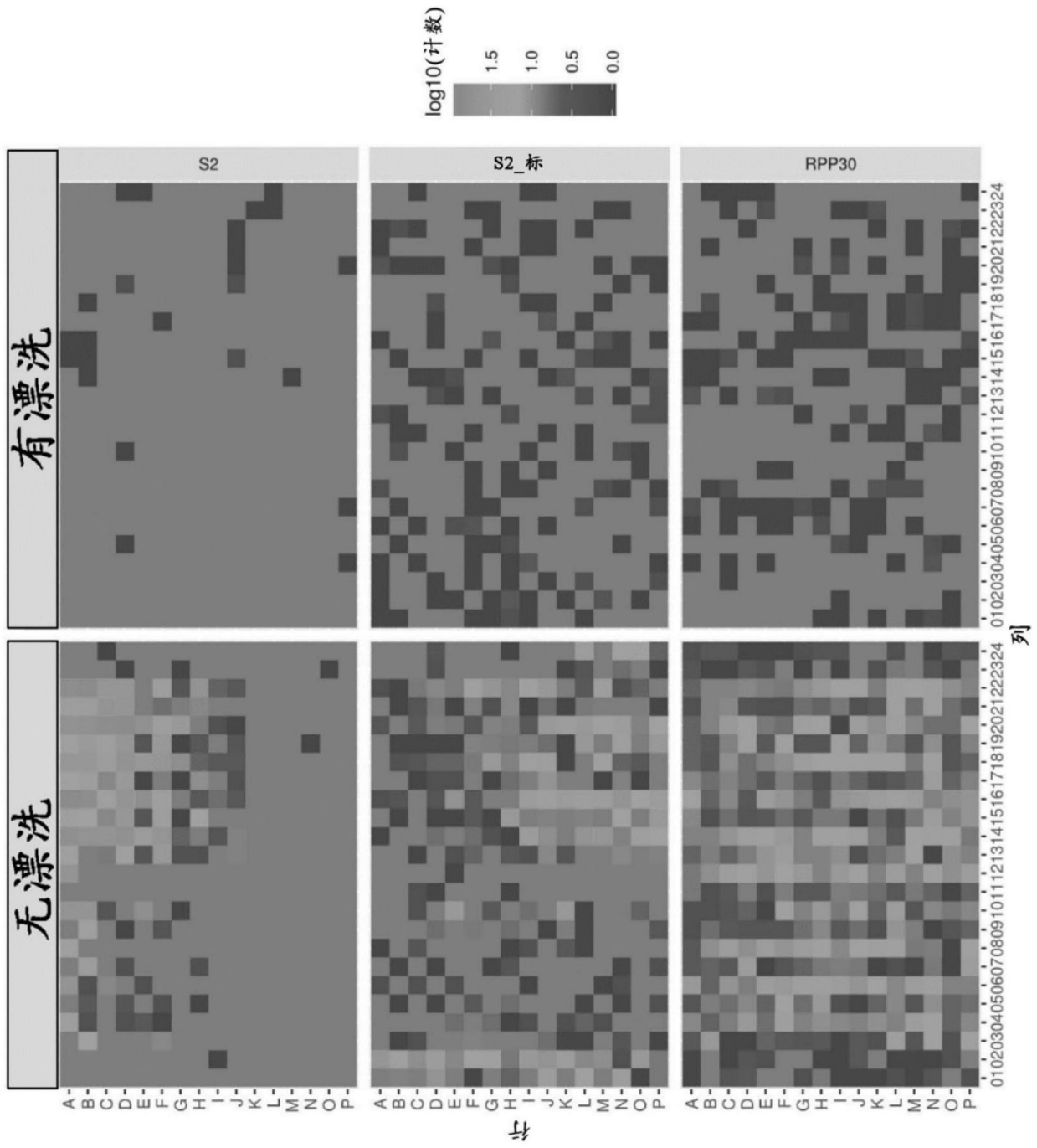
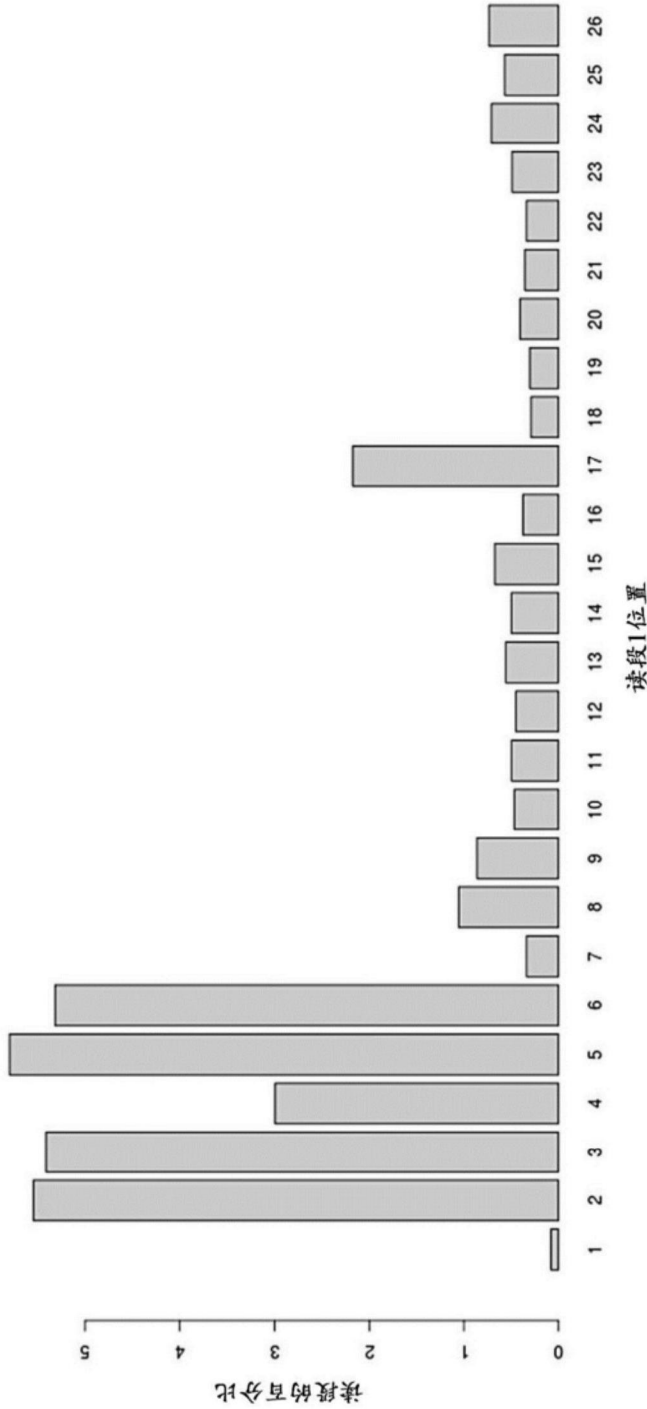


图17

Q<12的碱基的百分比 (碱基检出准确度<92%)



A

离S2标的汉明距离

| 离S2的汉明距离 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|----------|---------|--------|-------|------|-------|-------|--------|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 981607 |
| 1 | 0 | 0 | 0 | 0 | 0 | 49010 | 235315 |
| 2 | 0 | 0 | 0 | 0 | 22578 | 19024 | 111529 |
| 3 | 0 | 0 | 25516 | 7089 | 15226 | 59012 | |
| 4 | 0 | 74519 | 8036 | 7576 | 10021 | 36688 | |
| 5 | 0 | 522655 | 21671 | 9469 | 3292 | 5824 | 13217 |
| 6 | 9062990 | 222210 | 48991 | 2270 | 1337 | 1669 | 2140 |
| 7 | 0 | 928836 | 19700 | 2439 | 277 | 253 | 348 |

B

图18

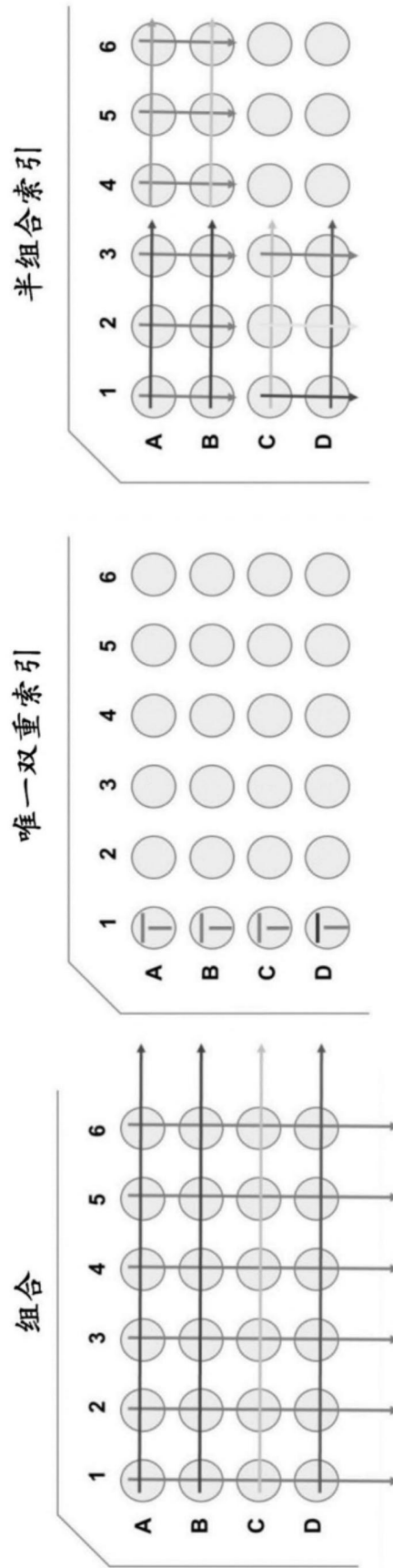


图19

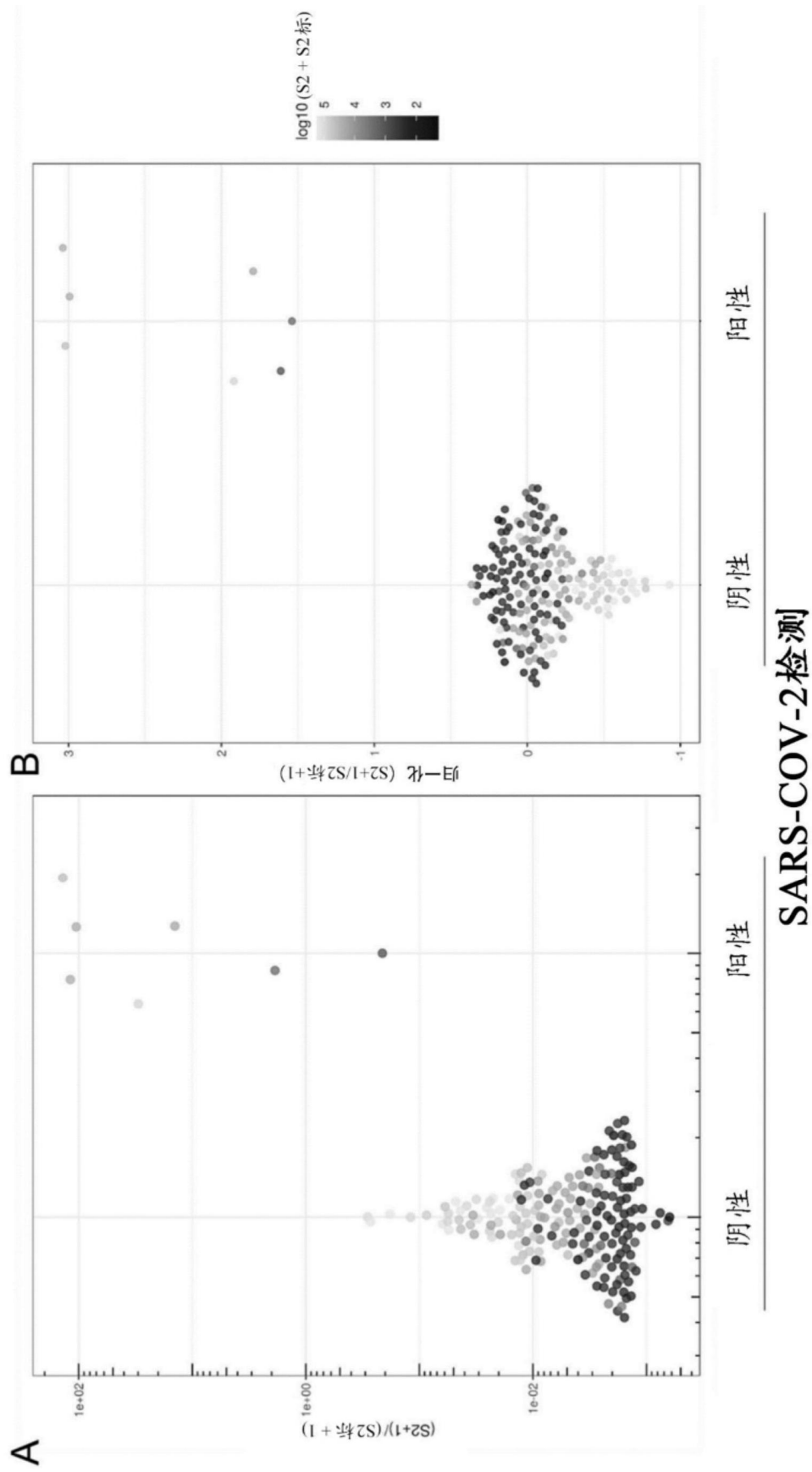


图20

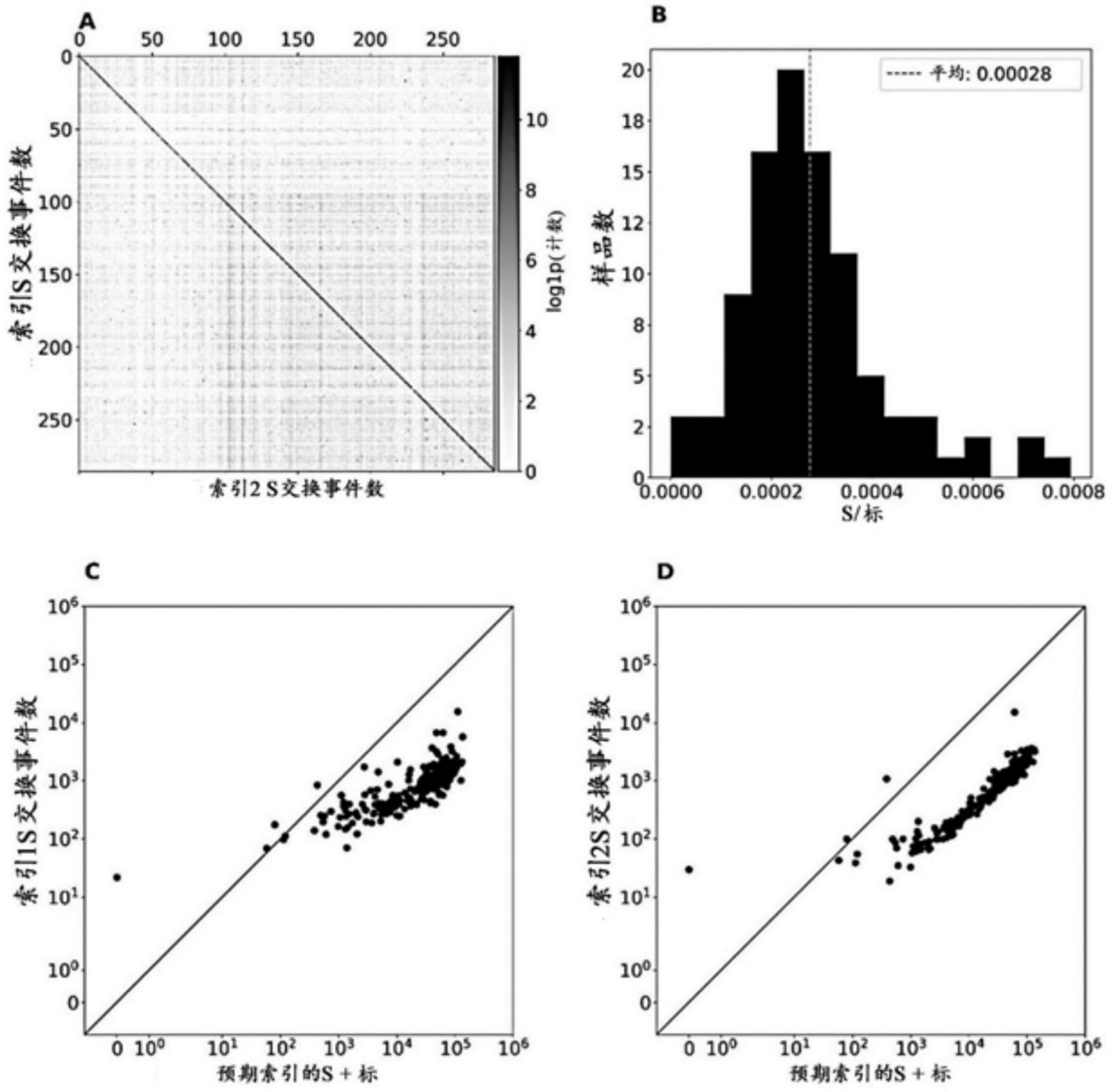
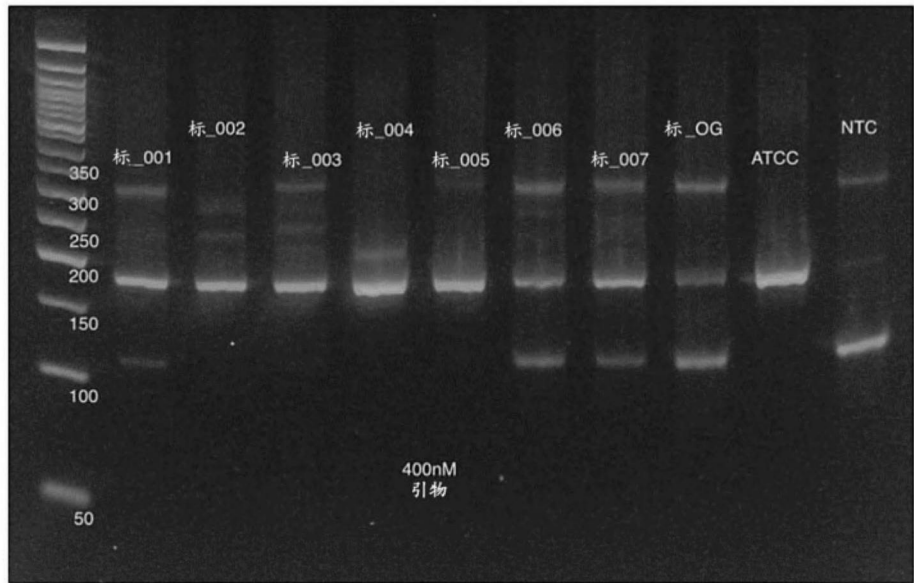


图21

以400 nM引物的RT-PCR



以100 nM引物的RT-PCR

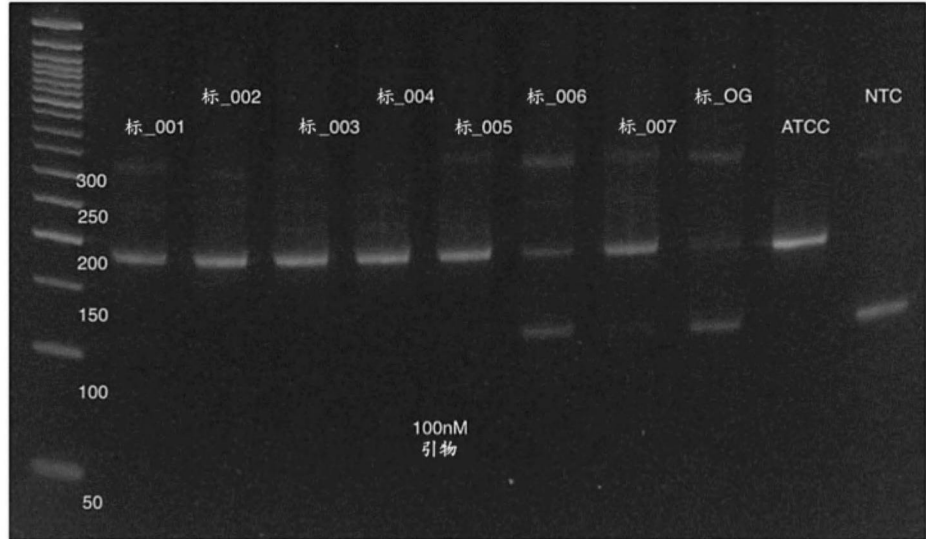


图22

```

> SARS-COV2_S2
-----S2_F引发位点-----S2_R引发位点-----
GCTGGTCTGCAGCTTATTATGTGGGT   TATCTTCAACCTAGGACTTTTCTATT   AAAATATAATGAAAATGGAACCATTACAGATGCTG   TAGACTGTGCACCTTGACCCCT

> 初始S2加标
-----S2_F引发位点-----S2_R引发位点-----
GCTGGTCTGCAGCTTATTATGTGGGT   ATAGAACAACCTAGGACTTTTCTATT   AAAATATAATGAAAATGGAACCATTACAGATGCTG   TAGACTGTGCACCTTGACCCCT

> 多样化加标组1
-----S2_F引发位点-----S2_R引发位点-----
GCTGGTCTGCAGCTTATTATGTGGGT   GTGTATCTCAGGAGCGACCCCTTGG   AAAATATAATGAAAATGGAACCATTACAGATGCTG   TAGACTGTGCACCTTGACCCCT
GCTGGTCTGCAGCTTATTATGTGGGT   CCTCGCTAGGACGTCGCTATGACGCC   AAAATATAATGAAAATGGAACCATTACAGATGCTG   TAGACTGTGCACCTTGACCCCT
GCTGGTCTGCAGCTTATTATGTGGGT   AGCAGACTTGATCTAACTGACACTA   AAAATATAATGAAAATGGAACCATTACAGATGCTG   TAGACTGTGCACCTTGACCCCT
GCTGGTCTGCAGCTTATTATGTGGGT   TAAGTAGGACTTCGATTGGATGGAAT   AAAATATAATGAAAATGGAACCATTACAGATGCTG   TAGACTGTGCACCTTGACCCCT

#多样化加标组2
-----S2_F引发位点-----S2_R引发位点-----
GCTGGTCTGCAGCTTATTATGTGGGT   atagaaCAACCTAGGACTTTTCTATT   AAAATATAATGAAAATGGAACCATTACAGATGCTG   TAGACTGTGCACCTTGACCCCT
GCTGGTCTGCAGCTTATTATGTGGGT   tggactACCTAGGACTGCGCGACTAA   AAAATATAATGAAAATGGAACCATTACAGATGCTG   TAGACTGTGCACCTTGACCCCT
GCTGGTCTGCAGCTTATTATGTGGGT   caCCTCGTAGGACTTCTGAAATAGCC   AAAATATAATGAAAATGGAACCATTACAGATGCTG   TAGACTGTGCACCTTGACCCCT
GCTGGTCTGCAGCTTATTATGTGGGT   GCTTGGTGGATCTCAGAACCGCGCGG   AAAATATAATGAAAATGGAACCATTACAGATGCTG   TAGACTGTGCACCTTGACCCCT

```

图23

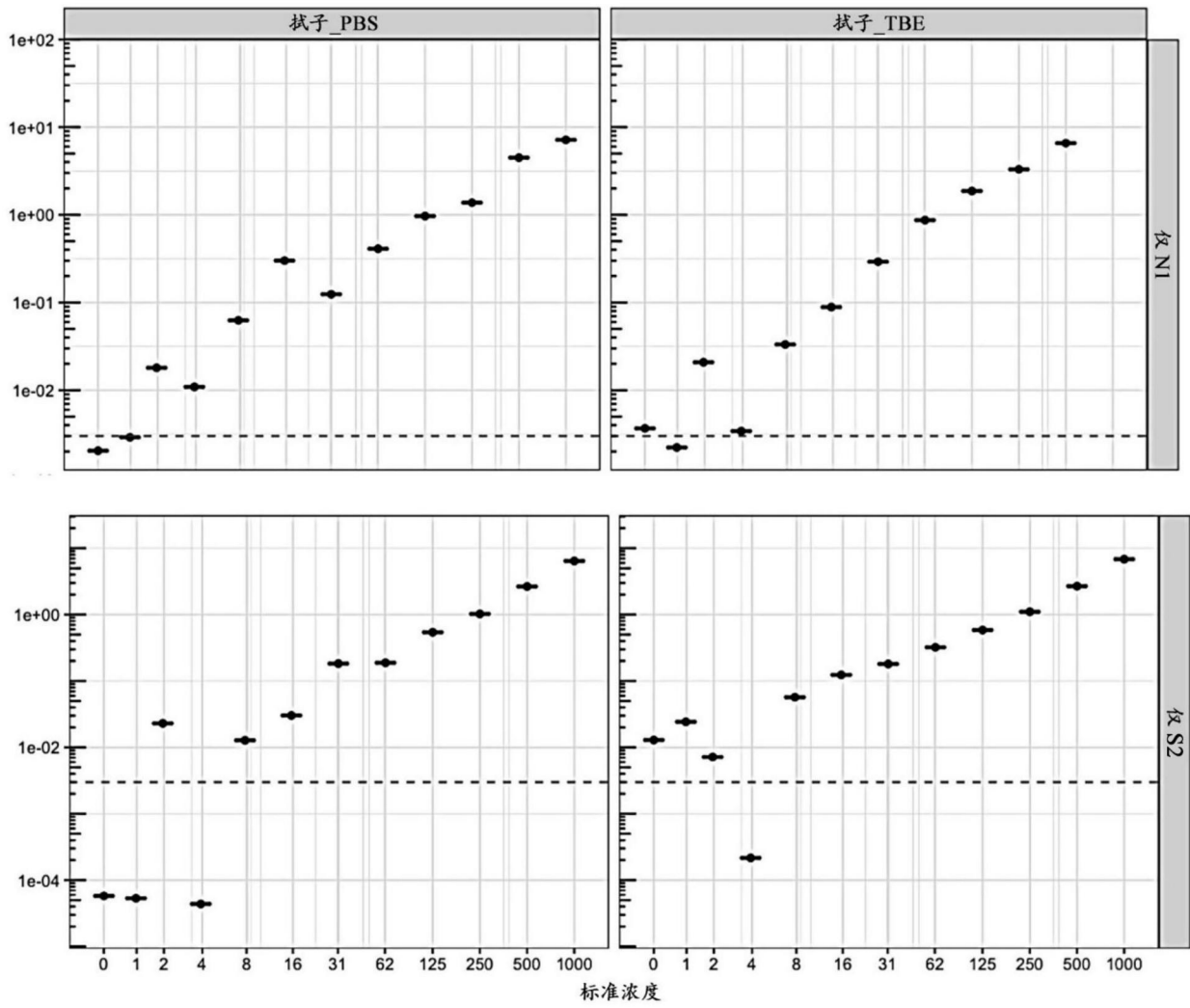


图24

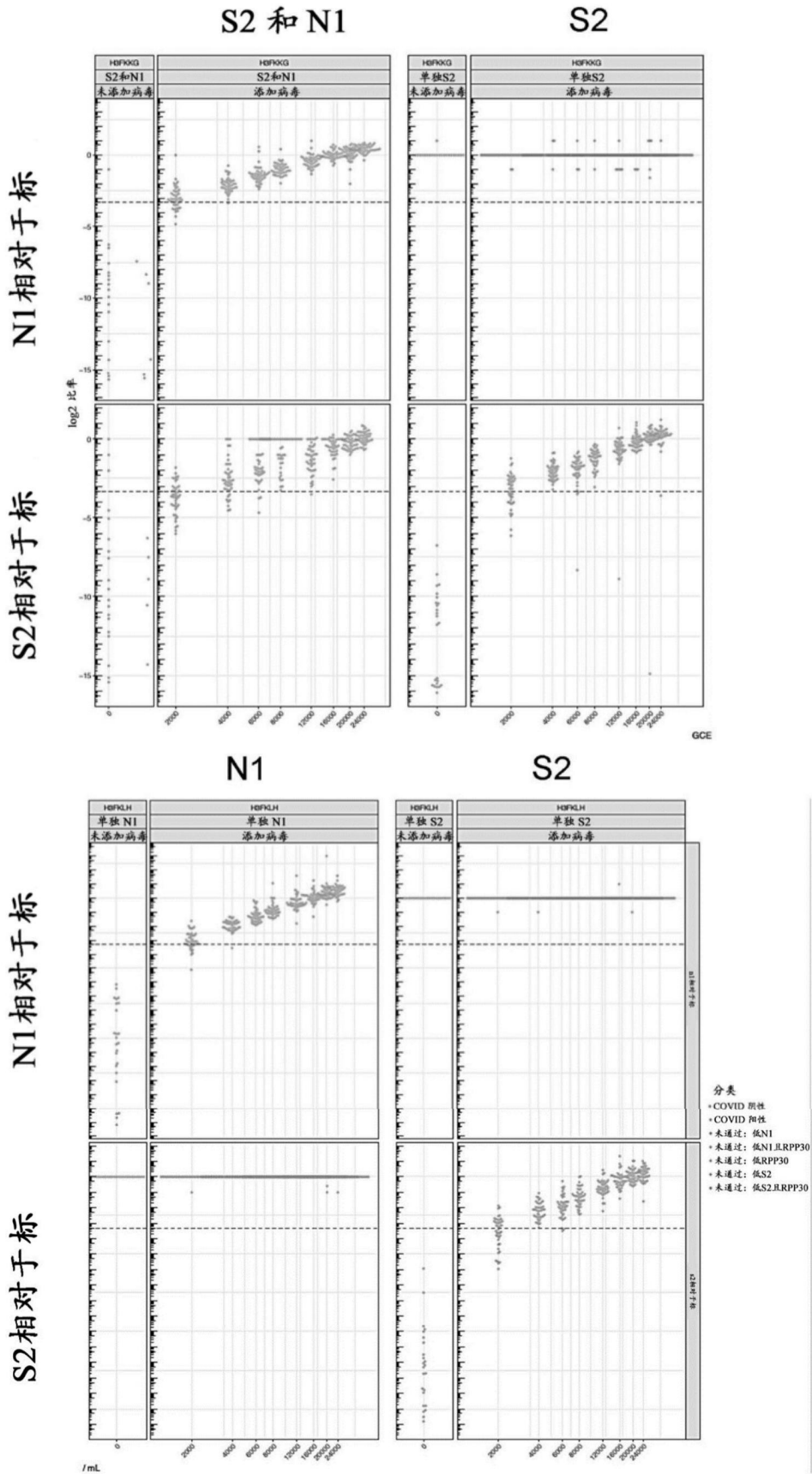


图25

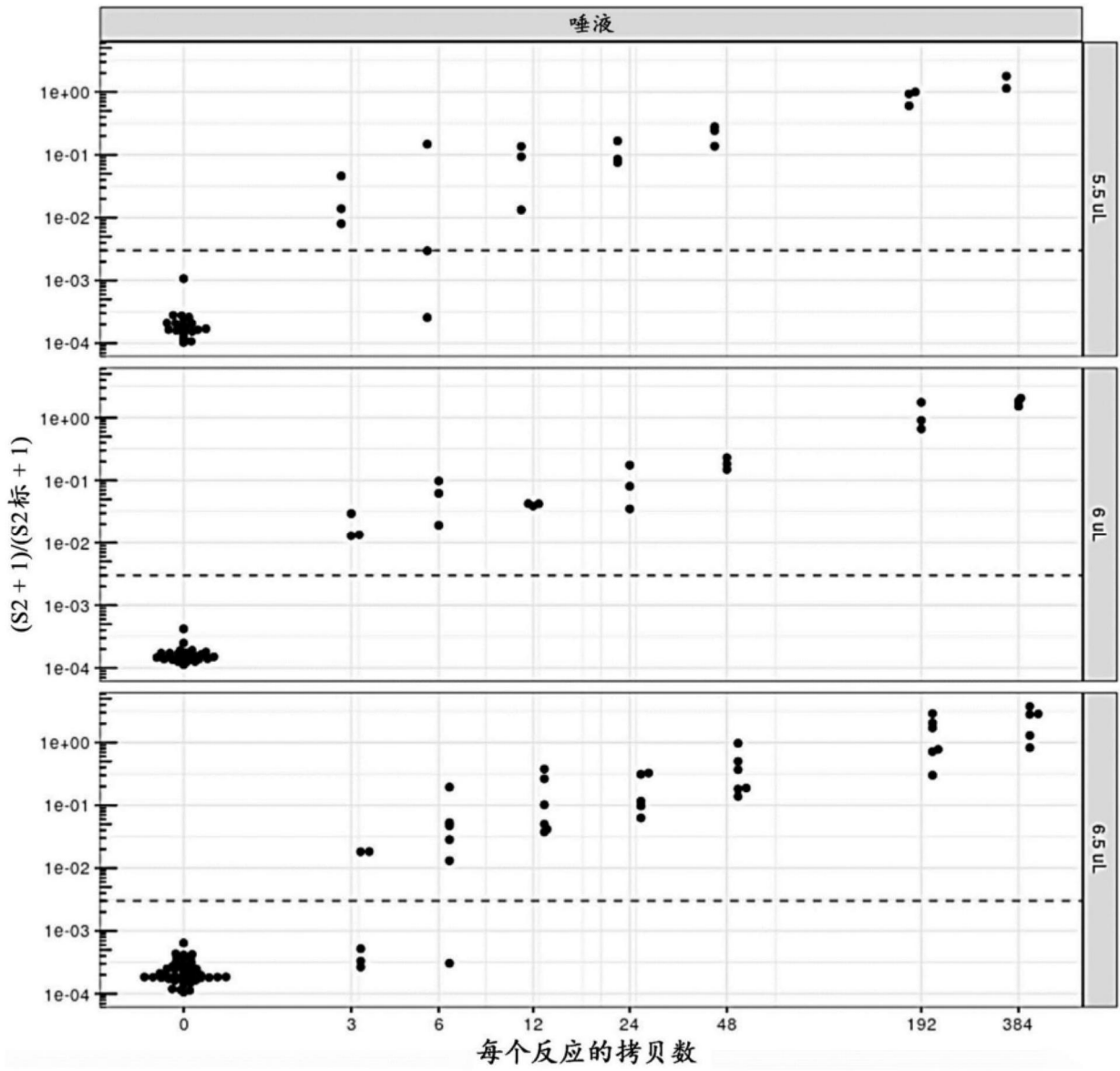


图26

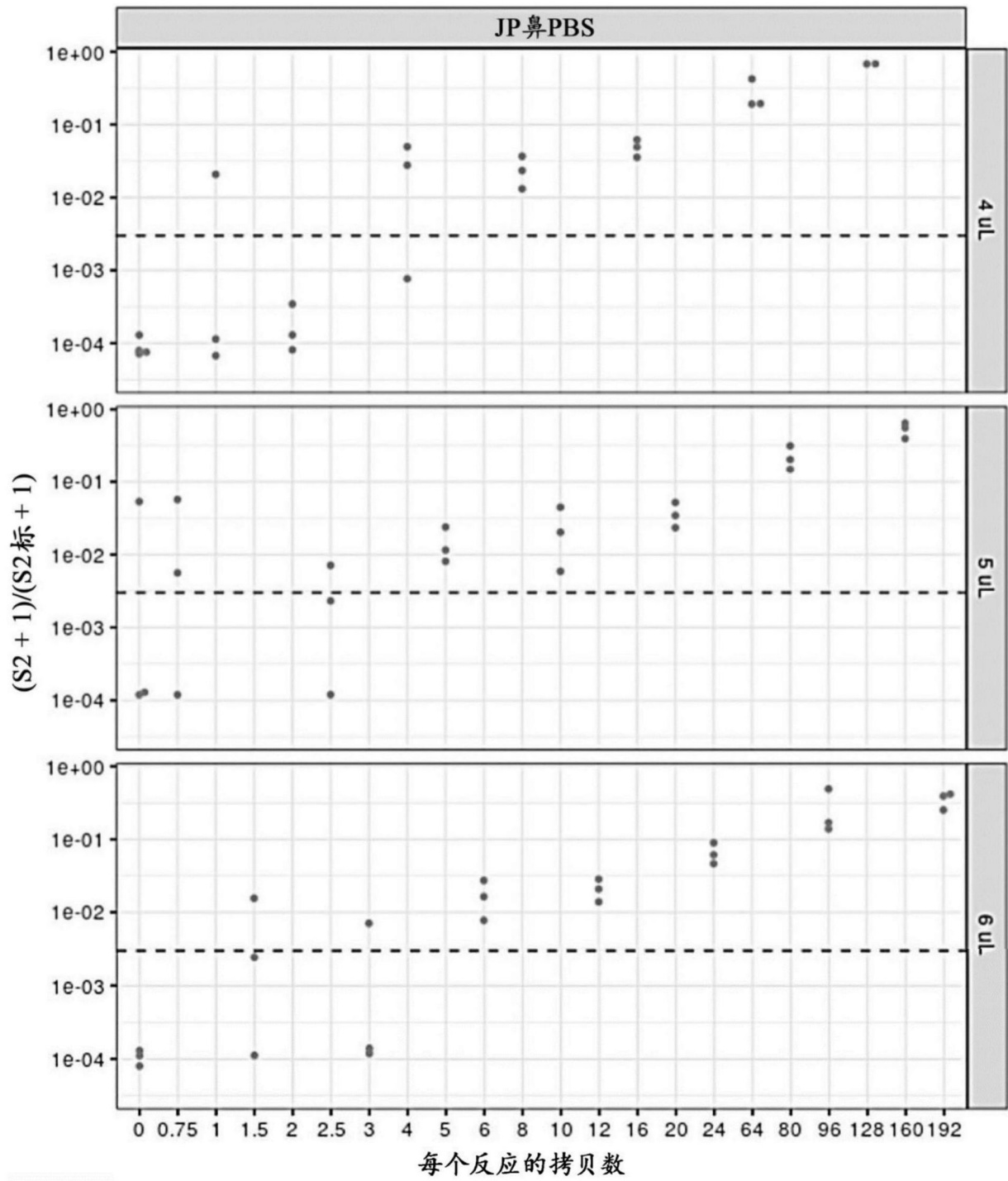


图27

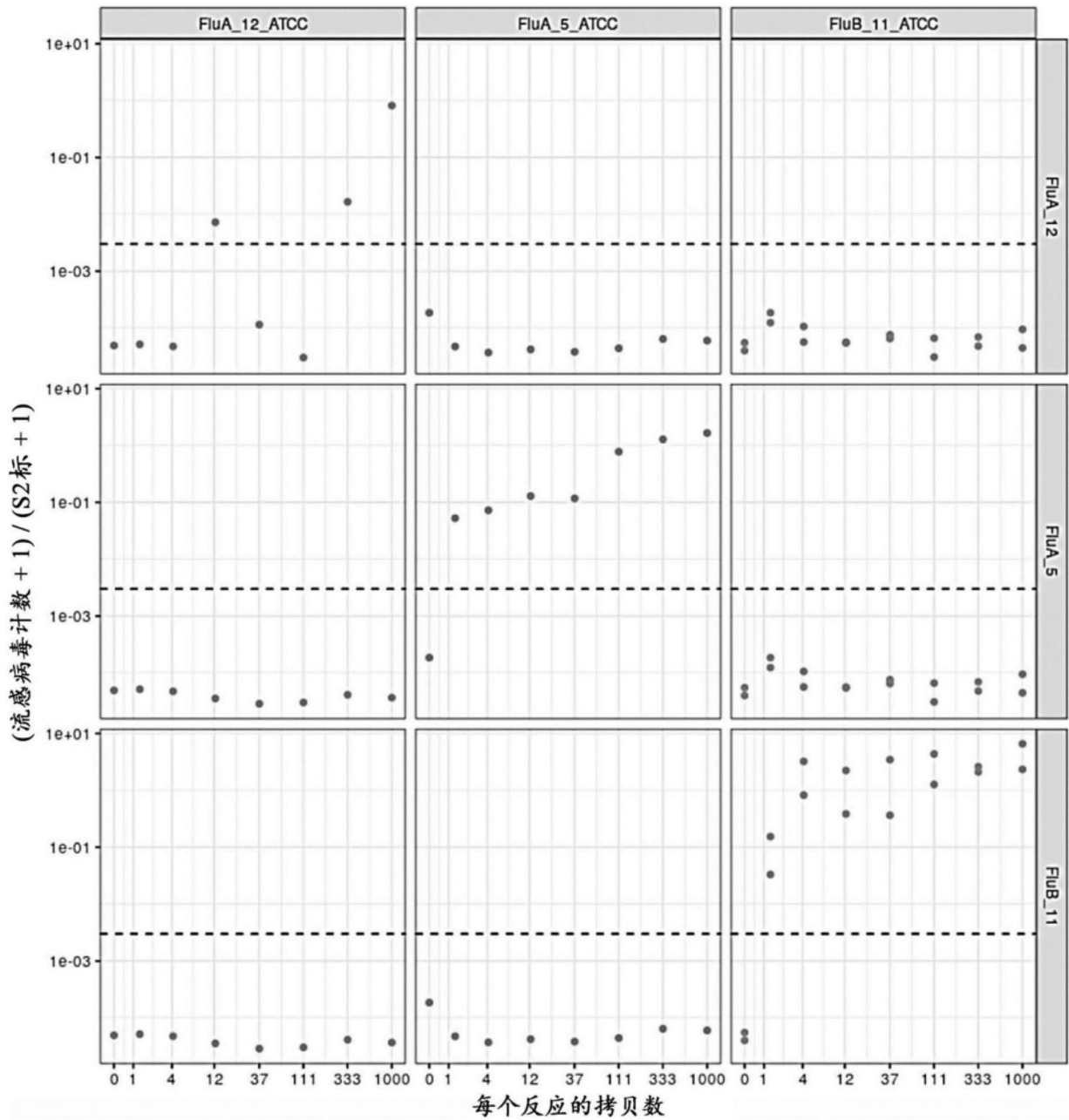


图28

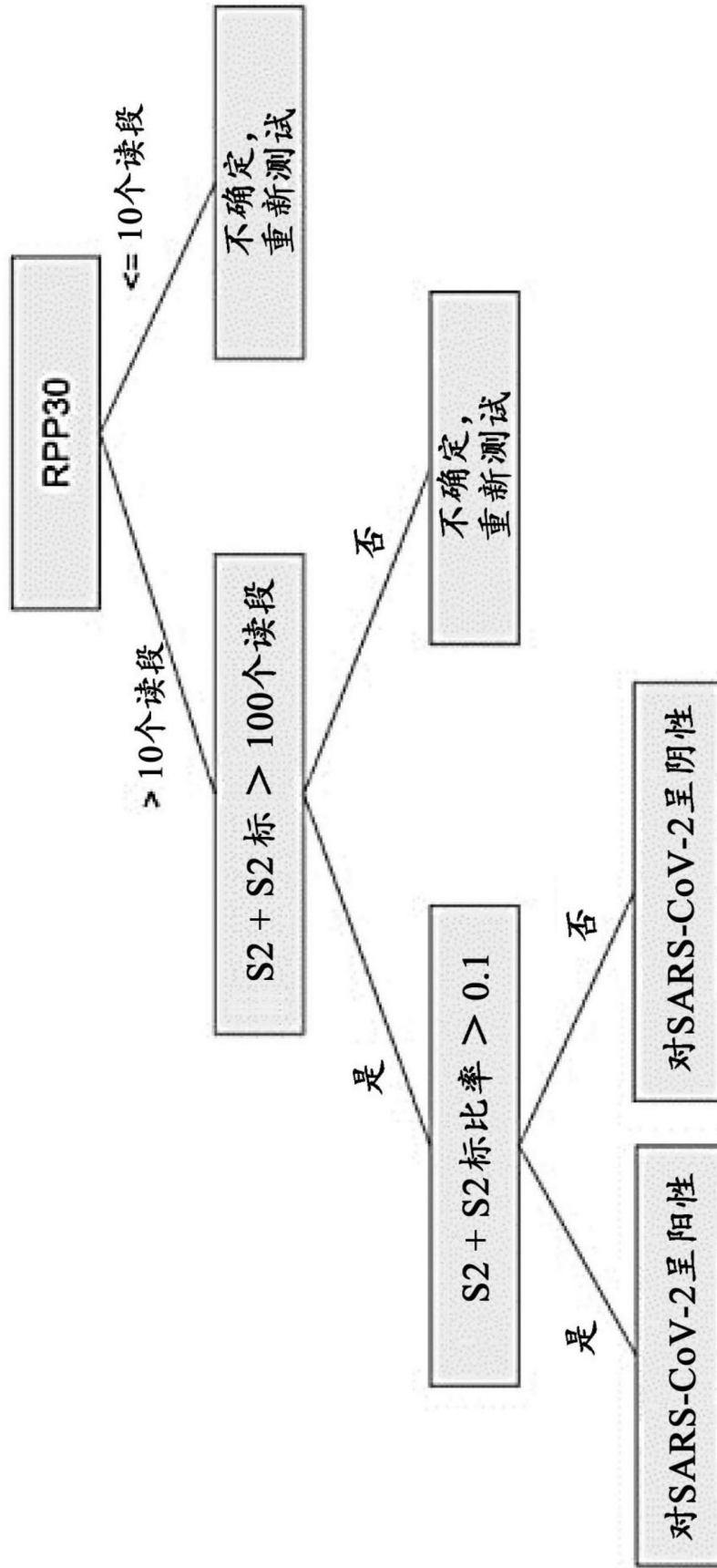


图29