(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 119234042 A
(43)申请公布日 2024.12.31

(21)申请号 202380041777.X

(22)申请日 2023.03.24

(30)优先权数据
63/323748 2022.03.25 US
63/405142 2022.09.09 US
63/419924 2022.10.27 US

(85)PCT国际申请进入国家阶段日
2024.11.20

(86)PCT国际申请的申请数据
PCT/US2023/016192 2023.03.24

(87)PCT国际申请的公布数据
WO2023/183550 EN 2023.09.28

(71)申请人 摩登纳特斯有限公司
地址 美国马萨诸塞州

(72)发明人 D·里德 M·A·齐默

(74)专利代理机构 中国专利代理(香港)有限公司 72001
专利代理师 李波 彭昶

(51)Int.Cl.
C12N 15/88(2006.01)
A61K 31/7088(2006.01)

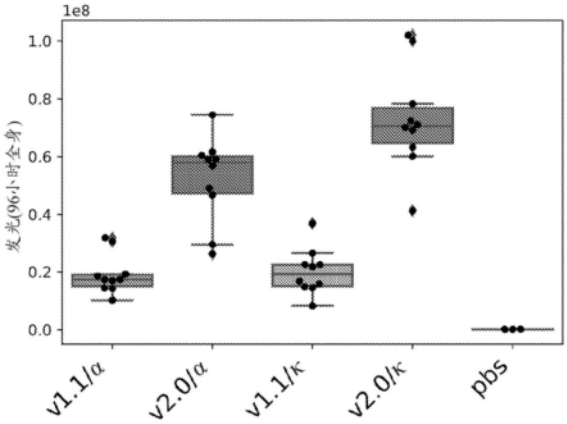
权利要求书6页 说明书226页 附图27页

(54)发明名称

具有延长的半衰期的信使核糖核酸

(57)摘要

本公开涉及一种编码多肽的多核苷酸,所述多核苷酸包含5'UTR、编码多肽的编码区和3'UTR,以及包含所述多核苷酸的脂质纳米颗粒。本公开的多核苷酸和/或脂质纳米颗粒能够通过增加编码多肽的多核苷酸的半衰期和/或表达持续时间来提高所述多肽的水平和/或活性。本文还公开了使用本公开的脂质纳米颗粒治疗受试者的疾病或病症的方法。



CN 119234042 A

1.一种包含5'UTR、编码多肽的开放阅读框和3'UTR的信使RNA(mRNA),其中所述3'UTR包含:

(i)与SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列至少98%同一的核苷酸序列;或

(ii)对应于SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列或其缺失变体的核苷酸序列,在所述缺失变体中1至75个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失,其中所述核酸序列或其缺失变体被修饰为包括:

a)插入所述核酸序列或其缺失变体内的一个或多个miRNA结合位点,和/或

b)插入所述核酸序列或其缺失变体内的TENT募集序列、FUT8募集序列、一个或多个鉴定和比率确定(IDR)序列、一个或多个核糖体接合检测测定(REDA)序列,或一个或多个IDR序列与一个或多个REDA序列的组合。

2.如权利要求1所述的mRNA,其中所述3'UTR包含与SEQ ID NO:139的核酸序列至少99%同一的核苷酸序列。

3.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:139所示的核酸序列。

4.如权利要求1所述的mRNA,其中所述3'UTR包含与SEQ ID NO:140的核酸序列至少99%同一的核苷酸序列。

5.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:140所示的核酸序列。

6.如权利要求1所述的mRNA,其中所述3'UTR包含与SEQ ID NO:141的核酸序列至少99%同一的核苷酸序列。

7.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:141所示的核酸序列。

8.如权利要求1所述的mRNA,其中所述3'UTR包含与SEQ ID NO:142的核酸序列至少99%同一的核苷酸序列。

9.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:142所示的核酸序列。

10.如权利要求1所述的mRNA,其中所述3'UTR包含与SEQ ID NO:143的核酸序列至少99%同一的核苷酸序列。

11.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:143所示的核酸序列。

12.如权利要求1所述的mRNA,其中所述3'UTR包含与SEQ ID NO:144的核酸序列至少99%同一的核苷酸序列。

13.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:144所示的核酸序列。

14.如权利要求1所述的mRNA,其中所述3'UTR包含与SEQ ID NO:145的核酸序列至少99%同一的核苷酸序列。

15.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ IDNO:145所示的核酸序列。

16.如权利要求1所述的mRNA,其中所述3'UTR包含与SEQ ID NO:146的核酸序列至少99%同一的核苷酸序列。

17.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:146所示的核酸序列。

18.如权利要求1所述的mRNA,其中所述3'UTR包含与SEQ ID NO:147的核酸序列至少

99%同一的核苷酸序列。

19.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:147所示的核酸序列。

20.如权利要求1所述的mRNA,其中所述3'UTR包含对应于SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列的核苷酸序列,其中所述核酸序列被修饰为包括插入所述核酸序列内的一个或多个miRNA结合位点。

21.如权利要求20所述的mRNA,其中所述一个或多个miRNA结合位点选自SEQ ID NO:148-157。

22.如权利要求20所述的mRNA,其中所述一个或多个miRNA结合位点包含SEQ ID NO:149的至少一个拷贝和SEQ ID NO:150的至少一个拷贝。

23.如权利要求20所述的mRNA,其中所述一个或多个miRNA结合位点包含SEQ ID NO:150的至少三个拷贝。

24.如权利要求20所述的mRNA,其中所述一个或多个miRNA结合位点包含SEQ ID NO:149的至少两个拷贝。

25.如权利要求20所述的mRNA,其中所述一个或多个miRNA结合位点包含SEQ ID NO:149的至少两个拷贝和SEQ ID NO:150的至少一个拷贝。

26.如权利要求20所述的mRNA,其中所述一个或多个miRNA结合位点包含SEQ ID NO:148的至少三个拷贝。

27.如权利要求1所述的mRNA,其中所述3'UTR包含对应于SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列的核苷酸序列,其中所述核酸序列被修饰为包括插入所述核酸序列内的TENT募集序列。

28.如权利要求1所述的mRNA,其中所述3'UTR包含对应于SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列的核苷酸序列,其中所述核酸序列被修饰为包括插入所述核酸序列内的FUT8募集序列。

29.如权利要求1所述的mRNA,其中所述3'UTR包含对应于SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列的核苷酸序列,其中所述核酸序列被修饰为包括插入所述核酸序列内的一个或多个IDR序列。

30.如权利要求1所述的mRNA,其中所述3'UTR包含对应于SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列的核苷酸序列,其中所述核酸序列被修饰为包括插入所述核酸序列内的一个或多个REDA序列。

31.如权利要求1所述的mRNA,其中在所述缺失变体中1至60个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。

32.如权利要求1所述的mRNA,其中在所述缺失变体中1至50个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、

SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。

33.如权利要求1所述的mRNA,其中在所述缺失变体中1至40个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。

34.如权利要求1所述的mRNA,其中在所述缺失变体中1至30个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。

35.如权利要求1所述的mRNA,其中在所述缺失变体中1至20个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。

36.如权利要求1所述的mRNA,其中在所述缺失变体中1至10个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。

37.如权利要求1所述的mRNA,其中在所述缺失变体中少于10个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。

38.如前述权利要求中任一项所述的mRNA,其中所述5'UTR包含与SEQ ID NO:50至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一的核苷酸序列。

39.如前述权利要求中任一项所述的mRNA,其中所述5'UTR包含SEQ ID NO:50所示的核酸序列。

40.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:139所示的核酸序列,并且其中所述5'UTR包含SEQ ID NO:50所示的核酸序列。

41.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:140所示的核酸序列,并且其中所述5'UTR包含SEQ ID NO:50所示的核酸序列。

42.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:141所示的核酸序列,并且其中所述5'UTR包含SEQ ID NO:50所示的核酸序列。

43.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:142所示的核酸序列,并且其中所述5'UTR包含SEQ ID NO:50所示的核酸序列。

44.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:143所示的核酸序列,并且其中所述5'UTR包含SEQ ID NO:50所示的核酸序列。

45.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:144所示的核酸序列,并且其中所述5'UTR包含SEQ ID NO:50所示的核酸序列。

46.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:145所示的核酸序列,并且其中所述5'UTR包含SEQ ID NO:50所示的核酸序列。

47.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:146所示的核酸序列,并且其中所述5'UTR包含SEQ ID NO:50所示的核酸序列。

48.如权利要求1所述的mRNA,其中所述3'UTR包含SEQ ID NO:147所示的核酸序列,并且其中所述5'UTR包含SEQ ID NO:50所示的核酸序列。

49.如前述权利要求中任一项所述的mRNA,其中所述mRNA包含终止盒。

50.如权利要求49所述的mRNA,其中所述终止盒选自SEQ ID NO:158-174。

51.如权利要求50所述的mRNA,其中所述终止盒是UAAAGC UCCCCGGGG(SEQ ID NO:165)或UAAGCCCCUCCGGGG(SEQ ID NO:164)。

52.如前述权利要求中任一项所述的mRNA,其中所述mRNA包含5'末端帽。

53.如权利要求52所述的mRNA,其中所述5'末端帽包含$m^7GpppG_{2'OMe}$、m7G-ppp-Gm-A、m7G-ppp-Gm-AG、帽0、帽1、ARCA、肌苷、N1-甲基-鸟苷、2'-氟-鸟苷、7-脱氮-鸟苷、8-氧代-鸟苷、2-氨基-鸟苷、LNA-鸟苷、2-叠氮鸟苷、帽2、帽4、5'甲基G帽或其类似物。

54.如前述权利要求中任一项所述的mRNA,其中所述mRNA包含多聚A区。

55.如权利要求54所述的mRNA,其中所述多聚A区的长度为至少约10、至少约20、至少约30、至少约40、至少约50、至少约60、至少约70、至少约80、至少约90个核苷酸,或长度为至少约100个核苷酸。

56.如权利要求54所述的mRNA,其中所述多聚A区的长度为约10至约200、约20至约180、约50至约160、约70至约140或约80至约120个核苷酸。

57.如权利要求54至56中任一项所述的mRNA,其中所述多聚A区包含A100-UCUAG-A20-反向脱氧-胸苷。

58.如前述权利要求中任一项所述的mRNA,其中所述mRNA包含至少一个化学修饰的核碱基、糖、主链或它们的任何组合。

59.如权利要求58所述的mRNA,其中所述至少一个化学修饰的核碱基选自由以下组成的组:假尿嘧啶(ψ)、N1-甲基假尿嘧啶(m1ψ)、1-乙基假尿嘧啶、2-硫尿嘧啶(s2U)、4'-硫尿嘧啶、5-甲基胞嘧啶、5-甲基尿嘧啶、5-甲氧基尿嘧啶和它们的任何组合。

60.如前述权利要求中任一项所述的mRNA,其中所述多肽包含分泌蛋白、膜结合蛋白或细胞间蛋白。

61.如权利要求60所述的mRNA,其中所述多肽是细胞因子、抗体、疫苗、受体、酶、激素、转录因子、配体、膜转运体、结构蛋白、核酸酶或其组分、变体或片段。

62.一种药物组合物,所述药物组合物包含前述权利要求中任一项所述的mRNA和药学上可接受的载体。

63.一种脂质纳米颗粒,所述脂质纳米颗粒包含权利要求1至61中任一项所述的mRNA。

64.如权利要求63所述的脂质纳米颗粒,其中所述脂质纳米颗粒包含:

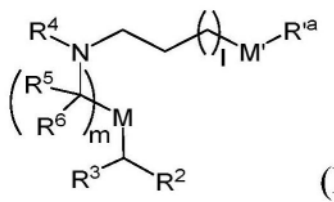
(i)可电离脂质,

(ii)磷脂,

(iii)结构脂质,和

(iv)PEG-脂质。

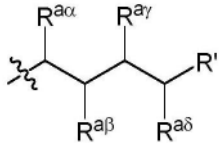
65.如权利要求63或64所述的脂质纳米颗粒,其中所述脂质纳米颗粒包含式(I)化合物:



或其N-氧化物,或其盐或异构体,

(I),

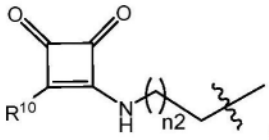
其中R'^{a}是R'^{支链};其中

R'^{支链}是: 其中—≈—表示连接点;

其中$R^{a\alpha}$、$R^{a\beta}$、$R^{a\gamma}$和$R^{a\delta}$各自独立地选自由H、C_{2-12}烷基和C_{2-12}烯基组成的组;

R^2和R^3各自独立地选自由C_{1-14}烷基和C_{2-14}烯基组成的组;

R^4选自由以下组成的组:-$(CH_2)_nOH$,其中n选自由1、2、3、4和5组成的组,和

,

其中—≈—表示连接点;其中

R^{10}是$N(R)_2$;每个R独立地选自由C_{1-6}烷基、C_{2-3}烯基和H组成的组;并且n2选自由1、2、3、4、5、6、7、8、9和10组成的组;

每个R^5独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组;

每个R^6独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组;

M和M'各自独立地选自由-C(O)O-和-OC(O)-组成的组;

R'是C_{1-12}烷基或C_{2-12}烯基;

l选自由1、2、3、4和5组成的组;并且

m选自由5、6、7、8、9、10、11、12和13组成的组。

66.如权利要求63-65中任一项所述的脂质纳米颗粒,其中所述脂质纳米颗粒包含:

(a)(i)化合物II,(ii)胆固醇,和(iii)PEG-DMG或化合物I;

(b)(i)化合物VI,(ii)胆固醇,和(iii)PEG-DMG或化合物I;

(c)(i)化合物II,(ii)DSPC或DOPE,(iii)胆固醇,和(iv)PEG-DMG或化合物I;

(d)(i)化合物VI,(ii)DSPC或DOPE,(iii)胆固醇,和(iv)PEG-DMG或化合物I;

(e)(i)化合物II,(ii)胆固醇,和(iii)化合物I;

(f)(i)化合物II,(ii)DSPC或DOPE,(iii)胆固醇,和(iv)化合物I;

(g)(i)化合物B,(ii)DSPC或DOPE,(iii)胆固醇,和(iv)PEG-DMG或化合物I;

(h)(i)化合物B,(ii)胆固醇,和(iii)化合物I;或

(i)(i)化合物B,(ii)DSPC或DOPE,(iii)胆固醇,和(iv)化合物I。

67.如权利要求63-65中任一项所述的脂质纳米颗粒,其中所述脂质纳米颗粒包含化合物II和化合物I。

68.如权利要求63-65中任一项所述的脂质纳米颗粒,其中所述脂质纳米颗粒包含化合物B和化合物I。

69.如权利要求63-65中任一项所述的脂质纳米颗粒,其中所述脂质纳米颗粒包含化合物II、DSPC、胆固醇和化合物I。

70.如权利要求63-69中任一项所述的脂质纳米颗粒,其中所述脂质纳米颗粒包含摩尔比为约20-60%的可电离脂质:5-25%的磷脂:25-55%的胆固醇:和0.5-15%的PEG脂质。

71.如权利要求63-70中任一项所述的脂质纳米颗粒,其中所述脂质纳米颗粒被配制成

用于静脉内、皮下、肌肉内、鼻内、眼内、经直肠、经肺或经口递送。

72.一种药物组合物,所述药物组合物包含权利要求63至71中任一项所述的脂质纳米颗粒。

73.一种细胞,所述细胞包含权利要求63至71中任一项所述的脂质纳米颗粒。

74.一种增加多肽表达的方法,所述方法包括向细胞施用权利要求63至71中任一项所述的脂质纳米颗粒。

75.一种将权利要求63至71中任一项所述的脂质纳米颗粒递送至细胞的方法,所述方法包括使所述细胞在体外、体内或离体与所述脂质纳米颗粒接触。

76.一种将权利要求63至71中任一项所述的脂质纳米颗粒递送至患有疾病或病症的人类受试者的方法,所述方法包括向有需要的所述人类受试者施用有效量的所述脂质纳米颗粒。

77.一种治疗、预防有需要的人类受试者的疾病或病症或预防所述疾病或病症的症状的方法,所述方法包括向所述人类受试者施用有效量的权利要求63至71中任一项所述的脂质纳米颗粒。

具有延长的半衰期的信使核糖核酸

[0001] 相关申请的交叉引用

[0002] 本申请要求2022年3月25日提交的美国临时申请号63/323,748、2022年9月9日提交的美国临时申请号63/405,142和2022年10月27日提交的美国临时申请号63/419,924的优先权益,所述临时申请的内容特此以引用的方式并入。

背景技术

[0003] 增加信使核糖核酸(mRNA)效力的努力已经专注于具有开放阅读框(ORF)的最佳序列设计的mRNA。然而,需要通过利用RNA生物学来进一步提高mRNA表达的效力和持久性。

发明内容

[0004] 本公开尤其提供了编码多肽的多核苷酸(例如,mRNA),其中所述多核苷酸包含:(a)5'-UTR(例如,如本文所描述);(b)编码区,所述编码区包含终止元件(例如,如本文所描述);以及(c)3'-UTR(例如,如本文所描述),以及包含所述多核苷酸的LNP组合物。在一个实施方案中,编码区包含多核苷酸序列,例如mRNA,例如编码肽或多肽有效载荷,例如治疗性有效载荷或预防性有效载荷的开放阅读框(ORF)。在一个实施方案中,多核苷酸,例如mRNA,或由多核苷酸编码的多肽相比缺乏本文所描述的5'-UTR、3'-UTR或终止元件的型式具有增加的水平和/或活性,例如表达或半衰期。在一个实施方案中,多核苷酸,例如mRNA的水平和/或活性增加。在一个实施方案中,由多核苷酸编码的多肽的水平、活性和/或表达持续时间增加。本文还公开了使用包含本文所公开的多核苷酸的LNP组合物用于治疗疾病或病症或用于促进受试者中的所需生物作用的方法。将了解,任何ORF可以与所公开的元件组合,例如,编码无论是例如细胞内的、跨膜的还是分泌的多肽或肽的ORF。本公开的额外方面在下文进一步详细描述。

[0005] 具体来说,在一些实施方案中,本文提供了包含5′UTR、编码多肽的开放阅读框和3′UTR的信使RNA(mRNA),其中所述3′UTR包含:

[0006] (i)与SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列至少98%同一的核苷酸序列;或

[0007] (ii)对应于SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列或其缺失变体的核苷酸序列,在所述缺失变体中1至75个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失,其中所述核酸序列或其缺失变体被修饰为包括:

[0008] a)插入所述核酸序列或其缺失变体内的一个或多个miRNA结合位点,和/或

[0009] b)插入所述核酸序列或其缺失变体内的TENT募集序列、FUT8募集序列、一个或多个鉴定和比率确定(IDR)序列、一个或多个核糖体接合检测测定(REDA)序列,或一个或多个

IDR序列与一个或多个REDA序列的组合。

[0010] 在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含与SEQ ID NO:139的核酸序列至少99%同一的核苷酸序列。在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含SEQ ID NO:139所示的核酸序列。

[0011] 在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含与SEQ ID NO:140的核酸序列至少99%同一的核苷酸序列。在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含SEQ ID NO:140所示的核酸序列。

[0012] 在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含与SEQ ID NO:141的核酸序列至少99%同一的核苷酸序列。在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含SEQ ID NO:141所示的核酸序列。

[0013] 在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含与SEQ ID NO:142的核酸序列至少99%同一的核苷酸序列。在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含SEQ ID NO:142所示的核酸序列。

[0014] 在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含与SEQ ID NO:143的核酸序列至少99%同一的核苷酸序列。在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含SEQ ID NO:143所示的核酸序列。

[0015] 在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含与SEQ ID NO:144的核酸序列至少99%同一的核苷酸序列。在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含SEQ ID NO:144所示的核酸序列。

[0016] 在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含与SEQ ID NO:145的核酸序列至少99%同一的核苷酸序列。在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含SEQ ID NO:145所示的核酸序列。

[0017] 在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含与SEQ ID NO:146的核酸序列至少99%同一的核苷酸序列。在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含SEQ ID NO:146所示的核酸序列。

[0018] 在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含与SEQ ID NO:147的核酸序列至少99%同一的核苷酸序列。在某些实施方案中,本公开提供了一种3′UTR,所述3′UTR包含SEQ ID NO:147所示的核酸序列。

[0019] 在本公开的一些方面,3′UTR包含对应于SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列的核苷酸序列,其中所述核酸序列被修饰为包括插入所述核酸序列内的一个或多个miRNA结合位点。在一些情况下,一个或多个miRNA结合位点选自SEQ ID NO:148-157。在一些实施方案中,一个或多个miRNA结合位点包含SEQ ID NO:149的至少一个拷贝和SEQ ID NO:150的至少一个拷贝。在一些实施方案中,一个或多个miRNA结合位点包含SEQ ID NO:150的至少三个拷贝。在一些实施方案中,一个或多个miRNA结合位点包含SEQ ID NO:149的至少两个拷贝。在一些实施方案中,一个或多个miRNA结合位点包含SEQ ID NO:149的至少两个拷贝和SEQ ID NO:150的至少一个拷贝。在一些实施方案中,一个或多个miRNA结合位点包含SEQ ID NO:148的至少三个拷贝。

[0020] 在本文所提供的一些方面,3′UTR包含对应于SEQ ID NO:139、SEQ ID NO:140、SEQ

ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列的核苷酸序列,其中所述核酸序列被修饰为包括插入所述核酸序列内的TENT募集序列。

[0021] 在本文所提供的一些实施方案中,3′UTR包含对应于SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列的核苷酸序列,其中所述核酸序列被修饰为包括插入所述核酸序列内的FUT8募集序列。

[0022] 在本文所提供的一些情况下,3′UTR包含对应于SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列的核苷酸序列,其中所述核酸序列被修饰为包括插入所述核酸序列内的一个或多个IDR序列。

[0023] 在本文所提供的一些方面,3′UTR包含对应于SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147的核酸序列的核苷酸序列,其中所述核酸序列被修饰为包括插入所述核酸序列内的一个或多个REDA序列。

[0024] 在本文所提供的一些方面,在缺失变体中1至60个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。在本文所提供的一些方面,在缺失变体中1至50个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。在本文所提供的一些方面,在缺失变体中1至40个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。在本文所提供的一些方面,在缺失变体中1至30个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。在本文所提供的一些方面,在缺失变体中1至20个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。在本文所提供的一些方面,在缺失变体中1至10个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。在本文所提供的一些方面,在缺失变体中少于10个连续核苷酸从SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147中缺失。

[0025] 在上文所描述的mRNA中的任一者的某些实施方案中,5′UTR包含与SEQ ID NO:50至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一的核苷酸序列。在一些实施方案中,5′UTR包含SEQ ID NO:50所示的核酸序列。

[0026] 在本文所提供的某些实施方案中,3′UTR包含SEQ ID NO:139所示的核酸序列,并且其中5′UTR包含SEQ ID NO:50所示的核酸序列。

[0027] 在本文所提供的某些实施方案中,3′UTR包含SEQ ID NO:140所示的核酸序列,并

且其中5′UTR包含SEQ ID NO:50所示的核酸序列。

[0028] 在本文所提供的某些实施方案中,3′UTR包含SEQ ID NO:141所示的核酸序列,并且其中5′UTR包含SEQ ID NO:50所示的核酸序列。

[0029] 在本文所提供的某些实施方案中,3′UTR包含SEQ ID NO:142所示的核酸序列,并且其中5′UTR包含SEQ ID NO:50所示的核酸序列。

[0030] 在本文所提供的某些实施方案中,3′UTR包含SEQ ID NO:143所示的核酸序列,并且其中5′UTR包含SEQ ID NO:50所示的核酸序列。

[0031] 在本文所提供的某些实施方案中,3′UTR包含SEQ ID NO:144所示的核酸序列,并且其中5′UTR包含SEQ ID NO:50所示的核酸序列。

[0032] 在本文所提供的某些实施方案中,3′UTR包含SEQ ID NO:145所示的核酸序列,并且其中5′UTR包含SEQ ID NO:50所示的核酸序列。

[0033] 在本文所提供的某些实施方案中,3′UTR包含SEQ ID NO:146所示的核酸序列,并且其中5′UTR包含SEQ ID NO:50所示的核酸序列。

[0034] 在本文所提供的某些实施方案中,3′UTR包含SEQ ID NO:147所示的核酸序列,并且其中5′UTR包含SEQ ID NO:50所示的核酸序列。

[0035] 在上文所描述的mRNA中的任一者的某些方面,mRNA包含终止盒。在一些实施方案中,终止盒选自SEQ ID NO:158-174。在一些实施方案中,终止盒是UAAAGCUCCCCGGGG(SEQ ID NO:165)或UAAGCCCC UCCGGGG(SEQ ID NO:164)。

[0036] 在上文所描述的mRNA中的任一者的某些方面,mRNA包含5′末端帽。在一些实施方案中,5′末端帽包含$m^7GpppG_{2'OMe}$、m7G-ppp-Gm-A、m7G-ppp-Gm-AG、帽0、帽1、ARCA、肌苷、N1-甲基-鸟苷、2′-氟-鸟苷、7-脱氮-鸟苷、8-氧代-鸟苷、2-氨基-鸟苷、LNA-鸟苷、2-叠氮鸟苷、帽2、帽4、5′甲基G帽或其类似物。

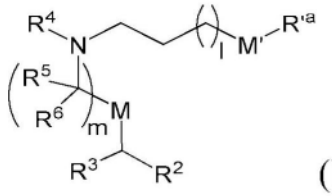
[0037] 在上文所描述的mRNA中的任一者的某些方面,mRNA包含多聚A区。在一些实施方案中,多聚A区的长度为至少约10、至少约20、至少约30、至少约40、至少约50、至少约60、至少约70、至少约80、至少约90个核苷酸,或长度为至少约100个核苷酸。在一些实施方案中,多聚A区的长度为约10至约200、约20至约180、约50至约160、约70至约140或约80至约120个核苷酸。在一些实施方案中,多聚A区包含A100-UCUAG-A20-反向脱氧-胸苷。

[0038] 在上文所描述的mRNA中的任一者的某些方面,mRNA包含至少一个化学修饰的核碱基、糖、主链或它们的任何组合。在一些实施方案中,至少一个化学修饰的核碱基选自由以下组成的组:假尿嘧啶(ψ)、N1-甲基假尿嘧啶(m1ψ)、1-乙基假尿嘧啶、2-硫尿嘧啶(s2U)、4′-硫尿嘧啶、5-甲基胞嘧啶、5-甲基尿嘧啶、5-甲氧基尿嘧啶和它们的任何组合。

[0039] 在上文所描述的mRNA中的任一者的某些方面,多肽包含分泌蛋白、膜结合蛋白或细胞间蛋白。在一些实施方案中,多肽是细胞因子、抗体、疫苗、受体、酶、激素、转录因子、配体、膜转运体、结构蛋白、核酸酶或其组分、变体或片段。

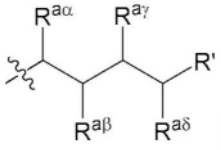
[0040] 本文还提供了药物组合物,所述药物组合物包含上文所描述的mRNA中的任一者和药学上可接受的载体。

[0041] 本文还提供了脂质纳米颗粒,所述脂质纳米颗粒包含上文所描述的mRNA中的任一者。在一些实施方案中,脂质纳米颗粒包含:(i)可电离脂质,(ii)磷脂,(iii)结构脂质,和(iv)PEG-脂质。在一些实施方案中,脂质纳米颗粒包含式(I)化合物:

[0042]  或其N-氧化物,或其盐或异构体,

(I),

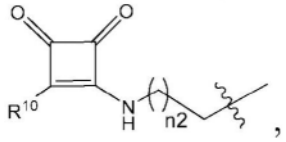
[0043] 其中R′^{a}是R′^{支链};其中

[0044] R′^{支链}是: 其中 表示连接点;

[0045] 其中$R^{a\alpha}$、$R^{a\beta}$、$R^{a\gamma}$和$R^{a\delta}$各自独立地选自由H、C_{2-12}烷基和C_{2-12}烯基组成的组;

[0046] R^{2}和R^{3}各自独立地选自由C_{1-14}烷基和C_{2-14}烯基组成的组;

[0047] R^{4}选自由以下组成的组:-$(CH_2)_n$OH,其中n选自由1、2、3、4和5组成的组,和

,

[0048] 其中 表示连接点;其中

[0049] R^{10}是N(R)$_2$;每个R独立地选自由C_{1-6}烷基、C_{2-3}烯基和H组成的组;并且n2选自由1、2、3、4、5、6、7、8、9和10组成的组;

[0050] 每个R^{5}独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组;

[0051] 每个R^{6}独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组;

[0052] M和M′各自独立地选自由-C(O)O-和-OC(O)-组成的组;

[0053] R′是C_{1-12}烷基或C_{2-12}烯基;

[0054] 1选自由1、2、3、4和5组成的组;并且

[0055] m选自由5、6、7、8、9、10、11、12和13组成的组。

[0056] 在一些实施方案中,脂质纳米颗粒包含:

[0057] (a)(i)化合物II,(ii)胆固醇,和(iii)PEG-DMG或化合物I;

[0058] (b)(i)化合物VI,(ii)胆固醇,和(iii)PEG-DMG或化合物I;

[0059] (c)(i)化合物II,(ii)DSPC或DOPE,(iii)胆固醇,和(iv)PEG-DMG或化合物I;

[0060] (d)(i)化合物VI,(ii)DSPC或DOPE,(iii)胆固醇,和(iv)PEG-DMG或化合物I;

[0061] (e)(i)化合物II,(ii)胆固醇,和(iii)化合物I;

[0062] (f)(i)化合物II,(ii)DSPC或DOPE,(iii)胆固醇,和(iv)化合物I;

[0063] (g)(i)化合物B,(ii)DSPC或DOPE,(iii)胆固醇,和(iv)PEG-DMG或化合物I;

[0064] (h)(i)化合物B,(ii)胆固醇,和(iii)化合物I;或

[0065] (i)(i)化合物B,(ii)DSPC或DOPE,(iii)胆固醇,和(iv)化合物I。

[0066] 在一些实施方案中,脂质纳米颗粒包含化合物II和化合物I。在一些实施方案中,脂质纳米颗粒包含化合物B和化合物I。在一些实施方案中,脂质纳米颗粒包含化合物II、DSPC、胆固醇和化合物I。在一些实施方案中,脂质纳米颗粒包含摩尔比为约20-60%的可电离脂质:5-25%的磷脂:25-55%的胆固醇;和0.5-15%的PEG脂质。在一些实施方案中,脂质

纳米颗粒被配制成用于静脉内、皮下、肌肉内、鼻内、眼内、经直肠、经肺或经口递送。

[0067] 本文还提供了药物组合物,所述药物组合物包含上文所描述的脂质纳米颗粒中的任一者。本文还提供了细胞,所述细胞包含上文所描述的脂质纳米颗粒中的任一者。

[0068] 在一些实施方案中,本文还提供了增加多肽表达的方法,所述方法包括向细胞施用上文所描述的脂质纳米颗粒中的任一者。

[0069] 本文还提供了将上文所描述的脂质纳米颗粒中的任一者递送至细胞的方法,所述方法包括使所述细胞在体外、体内或离体与所述脂质纳米颗粒接触。

[0070] 在一些情况下,本文还提供了将上文所描述的脂质纳米颗粒中的任一者递送至患有疾病或病症的人类受试者的方法,所述方法包括向有需要的所述人类受试者施用有效量的所述脂质纳米颗粒。

[0071] 本文还提供的其它方面是治疗、预防有需要的人类受试者的疾病或病症或预防所述疾病或病症的症状的方法,所述方法包括向所述人类受试者施用有效量的上文所描述的脂质纳米颗粒中的任一者。

[0072] 本文还提供的其它方面是治疗、预防有需要的人类受试者的疾病或病症或预防所述疾病或病症的症状的方法,所述方法包括向所述人类受试者施用有效量的上文所描述的脂质纳米颗粒中的任一者。

[0073] 本发明的其它特征和优点根据以下详细描述和图式以及根据权利要求将显而易见。

附图说明

[0074] 图1A-图1G是描绘由具有v1.1 5′UTR(SEQ ID NO:56)或v2.0 5′UTR(SEQ ID NO:50)与α3'UTR或κ 3′UTR组合的mRNA构建体编码的荧光素酶或靶蛋白表达的图。图1A显示了给药后96小时的全身ffLuc活性。图1B显示了给药后72小时的全身ffLuc活性。图1C显示了给药后0-4天内的全身ffLuc活性。图1D显示了肝中的表达。图1E显示了脾中的表达。图1F显示了给药后96小时血清中的靶蛋白(即,EPO)表达。图1 G显示了给药后0-4天血清中的靶蛋白(即,EPO)表达。

[0075] 图2A-图2I是描绘在各种免疫细胞中给药后1、2和3天由具有v2.0 5′UTR与对照3′UTR(在图例中标注为“三重(triple)”)、κ 3′UTR或ι 3′UTR组合的mRNA构建体编码的总体平均荧光强度或mOX40L+细胞百分比的图。图2A显示了LSK+造血干细胞和祖细胞中的平均荧光强度和mOX40L+细胞%。图2B显示了脾树突状细胞中的平均荧光强度和mOX40L+细胞%。图2C显示了脾巨噬细胞中的平均荧光强度和mOX40L+细胞%。图2D显示了脾嗜中性粒细胞中的平均荧光强度和mOX40L+细胞%。图2E显示了脾单核细胞中的平均荧光强度和mOX40L+细胞%。图2F显示了脾嗜酸性粒细胞中的平均荧光强度和mOX40L+细胞%。图2G显示了脾CD4+T细胞中的平均荧光强度和mOX40L+细胞%。图2H显示了脾CD8+T细胞中的平均荧光强度和mOX40L+细胞%。图2I显示了脾B细胞中的平均荧光强度和mOX40L+细胞%。

[0076] 图3A是显示在用指定构建体转染的细胞的指定时间点的FaDu三联体细胞系中FANCA(每个样品的上行)和核仁素(每个样品的下行)的蛋白质水平的图像。计算FANCA表达的倍数变化,将其针对上样对照和野生型(WT)平均值进行归一化。

[0077] 图3B是显示针对核仁素进行归一化且表示为相对于WT FaDu细胞系中FANCA的平

均表达的倍数变化的FANCA表达水平的图。

[0078] 图4是显示用1ug GFP mRNA构建体转染的FaDu-WT、FaDu-KO或用1ug FANCA构建体转染并在转染后24小时用指定浓度的丝裂霉素C(MMC)处理的FaDu-KO的存活百分比的图。使用cell titer Glo评估处理后5天的存活率。

[0079] 图5A是显示对照(Ctl)或1,3-丁二烯二环氧化物(DEB)和MMC处理的细胞中的细胞周期的G2_M期中的细胞积聚的图。WT-GFP是指用GFP mRNA转染的FaDu WT细胞;KO-GFP是指用GFP mRNA转染的FaDu KO细胞;FANCA_01-FANCA-04是指用FANCA构建体转染的FaDu KO细胞。转染在处理前6小时发生。

[0080] 图5B是显示图5A的数据的G2M频率的图,表示为平均值±SD。使用司徒登t检验(student t-test)来计算统计显著性。*P<0.05,**P＜0.01,***P<0.001,****P<0.0001。

[0081] 图6A是显示对照(Ctl)或1,3-丁二烯二环氧化物(DEB)和MMC处理的细胞中的细胞周期的G2_M期中的细胞积聚的图。WT-GFP是指用GFP mRNA转染的FaDu WT细胞;KO-GFP是指用GFP mRNA转染的FaDu KO细胞;FANCA_01-FANCA-04是指用FANCA构建体转染的FaDu KO细胞。转染在处理前72小时发生。

[0082] 图6B是显示图6A的数据的G2M频率的图,表示为平均值±SD。使用司徒登t检验来计算统计显著性。*P<0.05,**P＜0.01,***P<0.001,****P<0.0001。

[0083] 图6C是显示针对核仁素进行归一化且表示为相对于WT FaDu细胞系中的FANCA的倍数变化的FANCA表达水平的图。

[0084] 图7显示了给药后6小时和48小时后的卵白蛋白浓度。所评价的mRNA具有κ 3′UTR(SEQ ID NO:139)和v1.1 5′UTR(SEQ ID NO:56)或v2.05′UTR(SEQ ID NO:50)。两种mRNA均使用相同的“α”工艺制备。

[0085] 图8是显示当利用包含v2.0 5′UTR(SEQ ID NO:50)和κ 3′UTR(SEQ ID NO:139)的mRNA时的抗体反应的柱状图。

[0086] 图9是显示与不存在FUT8序列时相比,FUT8序列在3′UTR中的作用的柱状图。从左到右显示:PBS对照、含有v1.0UTR的mRNA、具有v2.0 5′UTR(SEQ ID:NO:50)和重新优化的编码序列的mRNA、具有来自(2)的元件加上带有‘δ’终止盒和‘FUT8’序列的3′UTR(3′UTR对应于SEQ ID NO:140)的mRNA以及两个阳性对照。

[0087] 图10是显示编码血凝素的v2.0 5′UTR mRNA的存在增加小鼠中的免疫原性的示意图。

[0088] 图11是显示v2.0 5′UTR和v2.0 3′UTR的存在增强体外编码血凝素的mRNA的表达的柱状图。

具体实施方式

[0089] 可以通过以下方式优化mRNA的效力和持久性:(1)确保递送至细胞质的mRNA与核糖体适当且有效地缔合;以及(2)最大化mRNA主动产生所需蛋白质产物所耗费的时间。mRNA的序列是在这些方面的性能的重要决定因素。

[0090] 本文尤其公开了以下发现:可以优化3′非翻译区(UTR)的序列以增加所述mRNA的效力和/或持久性。在一些实施方案中,可以优化3′UTR的序列与mRNA的5′UTR和/或终止元件的组合以例如通过延长mRNA的半衰期和/或表达持续时间来增加所述mRNA的效力和/或

持久性。在一些实施方案中,本公开提供了包含可以增加mRNA或由mRNA编码的多肽的功效(例如,水平和/或活性)的优化的3′UTR的多核苷酸和脂质纳米颗粒组合物。

[0091] 1.非翻译区(UTR)

[0092] 非翻译区(UTR)是起始密码子之前(5′UTR)和终止密码子之后(3′UTR)不被翻译的多核苷酸的核酸部分。在一些实施方案中,包含编码多肽的开放阅读框(ORF)的本发明的多核苷酸(例如,核糖核酸(RNA)),例如mRNA还包含UTR(例如,5′UTR或其功能片段、3′UTR或其功能片段或它们的组合)。

[0093] UTR可以与多核苷酸中的编码区同源或异源。在一些实施方案中,UTR与编码治疗性有效载荷或预防性有效载荷的ORF同源。在一些实施方案中,UTR与编码治疗性有效载荷或预防性有效载荷的ORF异源。

[0094] 在一些实施方案中,多核苷酸包含两个或更多个5′UTR或其功能片段,各自具有相同或不同的核苷酸序列。在一些实施方案中,多核苷酸包含两个或更多个3′UTR或其功能片段,各自具有相同或不同的核苷酸序列。

[0095] 在一些实施方案中,5′UTR或其功能片段、3′UTR或其功能片段或它们的任何组合是序列优化的。

[0096] 在一些实施方案中,5′UTR或其功能片段、3′UTR或其功能片段或它们的任何组合包含至少一个化学修饰的核碱基,例如,N1-甲基假尿嘧啶或5-甲氧基尿嘧啶。

[0097] UTR可以具有提供调控作用的特征,例如,增加或减少的稳定性、定位和/或翻译效率。可以向细胞、组织或生物体施用包含UTR的多核苷酸,并且可以使用常规方法测量一个或多个调控特征。在一些实施方案中,5′UTR或3′UTR的功能片段分别包含全长5′或3′UTR的一个或多个调控特征。

[0098] 天然5′UTR具有在翻译起始中起作用的特征。它们具有像Kozak序列这样的特征,通常已知这些特征涉及于核糖体启动许多基因翻译的过程中。Kozak序列具有共有CCR(A/G)CCAUGG(SEQ ID NO:125),其中R是起始密码子(AUG)上游三个碱基的嘌呤(腺嘌呤或乌嘌呤),其后是另一个‘G’。还已知5′UTR形成涉及于延伸因子结合中的二级结构。

[0099] 通过对特定靶器官的大量表达的基因中通常发现的特征进行工程改造,可以增强多核苷酸的稳定性和蛋白质产量。举例来说,引入肝表达的mRNA,诸如白蛋白、血清淀粉样蛋白A、载脂蛋白A/B/E、转铁蛋白、α胎蛋白、促红细胞生成素或因子VIII的5′UTR可以增强肝细胞系或肝中多核苷酸的表达。同样,使用来自其它组织特异性mRNA的5′UTR来提高那个组织中的表达对于肌肉(例如,MyoD、肌球蛋白、肌红蛋白、肌细胞生成素、力蛋白(Herculin))、内皮细胞(例如,Tie-1、CD36)、骨髓细胞(例如,C/EBP、AML1、G-CSF、GM-CSF、CD11b、MSR、Fr-1、i-NOS)、白细胞(例如,CD45、CD18、)、脂肪组织(例如,CD36、GLUT4、ACRP30、脂联素)和肺上皮细胞(例如,SP-A/B/C/D)是可能的。

[0100] 在一些实施方案中,UTR选自其蛋白质共享共同功能、结构、特征或特性的转录物家族。举例来说,编码的多肽可以属于在特定细胞、组织中或在发育期间的特定时间表达的蛋白质家族(即,共享至少一种功能、结构、特征、定位、起源或表达模式)。来自基因或mRNA中的任一者的UTR可以与相同或不同蛋白质家族的其它UTR中的任一者交换,以产生新的多核苷酸。

[0101] 在一些实施方案中,5′UTR和3′UTR可以是异源的。在一些实施方案中,5′UTR可以

来源于与3′UTR不同的种类。

[0102] 国际专利申请公布号WO/2014/164253(以全文引用的方式并入本文中)提供了可以在本发明的多核苷酸中用作ORF的侧翼区的示例性UTR的清单。

[0103] 本申请的额外示例性UTR包括但不限于一个或多个来源于以下核酸序列的5′UTR和/或3′UTR:球蛋白,诸如α-球蛋白或β-球蛋白(例如,爪蟾、小鼠、兔或人类球蛋白);强Kozak翻译起始信号;CYBA(例如,人类细胞色素b-245α多肽);白蛋白(例如,人类白蛋白7);HSD17B4(羟基类固醇(17-β)脱氢酶);病毒(例如,烟草蚀纹病毒(TEV)、委内瑞拉马脑炎病毒(VEEV)、登革热病毒、巨细胞病毒(CMV)(例如,CMV立即早期1(IE1))、肝炎病毒(例如,乙型肝炎病毒)、辛德毕斯病毒(sindbis virus)或PAV大麦黄矮病毒);热休克蛋白(例如,hsp70);翻译起始因子(例如,elF4G);葡萄糖转运体(例如,hGLUT1(人类葡萄糖转运体1));肌动蛋白(例如,人类α或β肌动蛋白);GAPDH;微管蛋白;组蛋白;柠檬酸循环酶;拓扑异构酶(例如,缺少5′TOP基序(寡嘧啶序列段)的TOP基因的5′UTR);核糖体蛋白大32(L32);核糖体蛋白(例如人类或小鼠核糖体蛋白,诸如rps9);ATP合酶(例如,ATP5A1或线粒体H+-ATP合酶的β亚基);生长激素e(例如,牛(bGH)或人类(hGH));延伸因子(例如,延伸因子1αα1(EEF1A1));锰超氧化物歧化酶(MnSOD);肌细胞增强因子2A(MEF2A);β-F1-ATPase、肌酸激酶、肌红蛋白、粒细胞集落刺激因子(G-CSF);胶原蛋白(例如,I型胶原蛋白α2(Col1A2)、I型胶原蛋白α1(Col1A1)、VI型胶原蛋白α2(Col6A2)、VI型胶原蛋白α1(Col6A1));核糖体结合蛋白(例如,核糖体结合蛋白I(RPNI));低密度脂蛋白受体相关蛋白(例如,LRP1);心肌营养素样细胞因子(例如,Nnt1);钙网蛋白(Calr);前胶原-赖氨酸2-氧代戊二酸5-双加氧酶1(Plod1);以及核连蛋白(例如,Nucb1)。

[0104] 在一些实施方案中,5′UTR选自由以下组成的组:β-球蛋白5′UTR;含有强Kozak翻译起始信号的5′UTR;细胞色素b-245α多肽(CYBA)5′UTR;羟基类固醇(17-β)脱氢酶(HSD17B4)5′UTR;烟草蚀刻病毒(TEV)5′UTR;委内瑞拉马脑炎病毒(TEEV)5′UTR;编码非结构蛋白的风疹病毒(RV)RNA的5′近端开放阅读框;登革热病毒(DEN)5′UTR;热休克蛋白70(Hsp70)5′UTR;eIF4G 5′UTR;GLUT1 5′UTR;其功能片段以及它们的任何组合。

[0105] 可以将来源于任何基因或mRNA的野生型UTR并入本发明的多核苷酸中。在一些实施方案中,可以相对于野生型或原生UTR改变UTR以产生变体UTR,例如,通过改变UTR相对于ORF的取向或位置;或通过包含额外的核苷酸、核苷酸的缺失、核苷酸的交换或转座。在一些实施方案中,可以利用5′或3′UTR的变体,例如,野生型UTR的突变体,或其中一个或多个核苷酸添加至UTR的末端或从UTR的末端去除的变体。

[0106] 另外,一个或多个合成UTR可以与一个或多个非合成UTR组合使用。参见例如Mandal和Rossi,Nat.Protoc.2013 8(3):568-82,其内容以全文引用的方式并入本文中。

[0107] UTR或其部分可以与从中选择它们的转录物相同的取向放置,或者可改变取向或位置。因此,5′和/或3′UTR可以是反向的、缩短的、加长的或与一个或多个其它5′UTR或3′UTR组合。

[0108] 在一些实施方案中,多核苷酸包含多个UTR,例如,双重、三重或四重5′UTR或3′UTR。举例来说,双重UTR包含同一UTR的两个串联或基本上串联的拷贝。举例来说,可以使用双重β-球蛋白3′UTR(参见US2010/0129877,其内容以全文引用的方式并入本文中)。

[0109] 本发明的多核苷酸可以包含特征的组合。举例来说,ORF可以侧接有包含强Kozak

翻译起始信号的5′UTR和/或包含寡聚(dT)序列用于模板化添加多聚A尾的3′UTR。5′UTR可以包含来自相同和/或不同UTR的第一多核苷酸片段和第二多核苷酸片段(参见例如US2010/0293625,以全文引用的方式并入本文中)。

[0110] 其它非UTR序列可以用作本发明的多核苷酸内的区域或亚区域。举例来说,内含子或内含子序列的部分可以并入本发明的多核苷酸中。内含子序列的并入可以增加蛋白质产量以及多核苷酸表达水平。在一些实施方案中,本发明的多核苷酸包含代替UTR或除UTR之外的内部核糖体进入位点(IRES)(参见例如Yakubov等人,Biochem.Biophys.Res.Commun.2010 394(1):189-193,其内容以全文引用的方式并入本文中)。在一些实施方案中,多核苷酸包含代替5′UTR序列的IRES。在一些实施方案中,多核苷酸包含ORF和病毒衣壳序列。在一些实施方案中,多核苷酸包含合成5′UTR与非合成3′UTR的组合。

[0111] 在一些实施方案中,UTR还可以包括至少一个翻译增强子多核苷酸、翻译增强子元件或翻译性增强子元件(统称为“TEE”,它是指增加从多核苷酸产生的多肽或蛋白质的量的核酸序列)。作为一个非限制性实例,TEE可以位于转录启动子与起始密码子之间。在一些实施方案中,5′UTR包含TEE。

[0112] 在一个方面,TEE是UTR中的保守元件,其可以促进核酸的翻译活性,诸如但不限于帽依赖性或帽非依赖性翻译。

[0113] a.5′UTR序列

[0114] 5′UTR序列对于核糖体募集至mRNA很重要,并且已报道在翻译中起作用(Hinnebusch A等人,(2016)Science,352:6292:1413-6)。

[0115] 本文尤其公开了一种编码多肽的多核苷酸,所述多核苷酸尤其包含5′UTR。在一个实施方案中,本文所公开的多核苷酸包含:(a)5′-UTR(例如,如表1中所提供或其变体或片段);(b)编码区;以及(c)终止元件和3′-UTR(例如,如表2中所提供或其变体或片段),以及包含所述多核苷酸的LNP组合物。在一个实施方案中,多核苷酸包含5′-UTR,所述5′-UTR包含表1中所提供的序列或其变体或片段(例如,其功能变体或片段)。将了解,此类5′UTR并入自然界中未发现的构建体中,例如,此类5′UTR是合成的,在序列中从天然存在的5′UTR改变,是在自然界中发现的那些的截短或加长形式,包含化学修饰的碱基,处于不同于自然界中可能发现的那些的ORF序列的5′,等等。

[0116] 在一个实施方案中,5′UTR包含表1中所提供的序列或与表1中所提供的5′UTR序列具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列或其变体或片段。在一个实施方案中,5′UTR包含与SEQ ID NO:50、SEQ ID NO:51、SEQ ID NO:52、SEQ ID NO:53、SEQ ID NO:54、SEQ ID NO:55、SEQ ID NO:56、SEQ ID NO:57或SEQ ID NO:58具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0117] 在一个实施方案中,5′UTR包含与SEQ ID NO:50具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。在一个实施方案中,5′UTR包含与SEQ ID NO:51具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。在一个实施方案中,5′UTR包含与SEQ ID NO:52具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。在一个实施方案中,5′UTR包含与SEQ ID NO:53具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。在一个实施方案中,5′UTR

包含与SEQ ID NO:54具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。在一个实施方案中,5′UTR包含与SEQ ID NO:55具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。在一个实施方案中,5′UTR包含与SEQ ID NO:56具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。在一个实施方案中,5′UTR包含与SEQ ID NO:57具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。在一个实施方案中,5′UTR包含与SEQ ID NO:58具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0118] 在一个实施方案中,5′UTR包含与SEQ ID NO:64具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0119] 在一个实施方案中,5′UTR包含SEQ ID NO:50的序列。在一个实施方案中,5′UTR由SEQ ID NO:50的序列组成。

[0120] 在一个实施方案中,5′UTR包含SEQ ID NO:64的序列。在一个实施方案中,5′UTR由SEQ ID NO:64的序列组成。

[0121] 在一个实施方案中,表1中所提供的5′UTR序列具有的第一核苷酸为A。在一个实施方案中,表1中所提供的5′UTR序列具有的第一核苷酸为G。在一个实施方案中,表1中所提供的5′UTR序列具有的前两个核苷酸为AG。在一个实施方案中,表1中所提供的5′UTR序列具有的前两个核苷酸为GA。

[0122] 表1:5′UTR序列

[0123]

SEQ ID NO:	序列名称	序列
50	A1	GGAAAUCGCAAAAUUUGCUCUUCGCGUUAGAUUUCUUUUAGUUUUCUCGCAACUAGCAAGCUUUUUGUUCUCGCC
51	A5	GGAAAUCCCCACAACCGCCUCAUAUCCAGGCUCAAGAAUAGAGCUCAGUGUUUUGUUGUUUAAUCAUUCCGACGUGUUUUGCGAUAUUCGCGCAAAGCAGCCAGUCGCGCGCUUGCUUUUAAGUAGAGUUGUUUUUCCACCCGUUUGCCAGGCAUCUUUAAUUUAACAUAUUUUUAUUUUUCAGGCUAACCUACGCCGCCCACC
52	A6	GGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGAUCUCCCUGAGCUUCAGGGAGCCCCGGCGCCGCCACC
53	A7	GGAAACCCCCCACCCCCGUAAGAGAGAAAAGAA

		GAGUAAGAAGAAAUAUAAGAUCUCCCUGAGCUU CAGGGAGCCCCGGCGCCGCCACC
54	A8	GGAGAACUUCCGCUUCCGUUGGCGCAAGCGCUU UCAUUUUUUCUGCUACCGUGACUAAG
55	A9	GGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGAGCCACC
56	A11 (参考)	GGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC
57	A2	GGAAAUCGCAAAAUUUGCUCUUCGCGUUAGAUU UCUUUUAGUUUUCUCGCAACUAGCAAGCUUUUU GUUCUCGCCGCCGCC
58	A3	GGAAAUCGCAAAAUUUUCUUUUCGCGUUAGAUU UCUUUUAGUUUUCUUUCAACUAGCAAGCUUUUU GUUCUCGCCGCCGCC
59	A4	G G A A A U C G C A A A A $(N_2)_X$ $(N_3)_X$ C U $(N_4)_X$ $(N_5)_X$ C G C G U U A G A U U U C U U U U U A G U U U U U C U N_6 N_7 C A A C U A G C A A G C U U U U U U G U U C U C G C C $(N_8$ C C)x $(N_2)_x$ 是尿嘧啶并且 x 是 0 至 5 的整数，例如，其中 x = 3 或 4; $(N_3)_x$ 是乌嘌呤并且 x 是 0 至 1 的整数; $(N_4)_x$ 是胞嘧啶并且 x 是 0 至 1 的整数; $(N_5)_x$ 是尿嘧啶并且 x 是 0 至 5 的整数，例如，其中 x = 2 或 3; N_6 是尿嘧啶或胞嘧啶; N_7 是尿嘧啶或乌嘌呤; N_8 是腺嘌呤或乌嘌呤并且 x 是 0 至 1 的整数。
60	A27	GGAAAAUUUUAGCCUGGAACGUUAGAUAACUGU CCUGUUGUCUUUAUAUACUUGGUCCCCAAGUAG UUUGUCUUCCAAA
61	A12	GGAAACUUUAUUUAGUGUUACUUUAUUUUCUGU UUAUUUGUGUUUCUUCAGUGGGUUUGUUCUAAU UUCCUUGGCCGCC
62	A13	GGAAAAUCUGUAUUAGGUUGGCGUGUUCUUUGG UCGGUUGUUAGUAUUGUUGUUGAUUCGUUUGUG GUCGGUUGCCGCC
63	A14	GGAAAAUUAUUAACAUCUUGGUAUUCUCGAUAA CCAUUCGUUGGAUUUUAUUGUAUUCGUAGUUUG GGUUCCUGCCGCC
64	A15	GGAAAUUAUUAUUAUUUCUAGCUACAAUUUAUC AUUGUAUUAUUUUAGCUAUUCAUCAUUAUUUAC UUGGUGAUCAACA
65	A16	GGAAAUAGGUUGUUAACCAAGUUCAAGCCUAAU AAGCUUGGAUUCUGGUGACUUGCUUCACCGUUG GCGGGCACCGAUC
66	A17	GGAAAUCGUAGAGAGUCGUACUUAGUACAUAUC GACUAUCGGUGGACACCAUCAAGAUUAUAAACC AGGCCAGA
67	A18	GGAAACCCGCCCAAGCGACCCCAACAUAUCAGC AGUUGCCCAAUCCCAACUCCCAACACAAUCCCCA AGCAACGCCGCC
68	A19	GGAAAGCGAUUGAAGGCGUCUUUUCAACUACUC GAUUAAGGUUGGGUAUCGUCGUGGGACUUGGAA AUUUGUUGUUUCC
69	A20	GGAAACUAAUCGAAAUAAAAGAGCCCCGUACUC UUUUAUUUCUAUUAGGUUAGGAGCCUUAGCAUU UGUAUCUUAGGUA

[0124]

[0125]

70	A21	GGAAAUGUGAUUUCCAGCAACUUCUUUUGAAUA UAUUGAAUUCCUAAUUCAAAGCGAACAAAUCUA CAAGCCAUAUACC
71	A22	GGAAAUCGUAGAGAGUCGUACUUACGUGGUCGC CAUUGCAUAGCGCGCGAAAGCAACAGGAACAAG AACGCGCC
72	A23	GGAAAUCGUAGAGAGUCGUACUUAGAAUAAACA GAGUCGGGUCGACUUGUCUCUGAUACUACGACG UCACAAUC
73	A24	GGAAAAUUUGCCUUCGGAGUUGCGUAUCCUGAA CUGCCCAGCCUCCUGAUAUACAACUGUUCCGCU UAUUCGGGCCGCC
74	A25	GGAAAUCUGAGCAGGAAUCCUUUGUGCAUUGAA GACUUUAGAUUCCUCUCUGCGGUAGACGUGCAC UUAUAAGUAUUUG
75	A26	GGAAAGCGAUUGAAGGCGUCUUUUCAACUACUC GAUUAAGGUUGGGUAUCGUCGUGGGACUUGGAA AUUUGUUGCCACC
76	A28	GGAAAUUUUUUUUUGAUAUUAUAAGAGUUUUU UUUUGAUAUUAAGAAAAUUUUUUUUUGAUAUU AGAAGAGUAAGAAGAAAUAUAAGACCCCGGCGC CGCCACC
77	A29	GGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGAGCCAAAAAAAAAAAACC
78	A30	GGAAAUCUCCCUGAGCUUCAGGGAGUAAGAGAG AAAAGAAGAGUAAGAAGAAAUAUAAGACCCCGG CGCCGCCACC
79	A31	GCCRCC，其中 R＝A 或 G
80	A32	GGACUCACUAUUUGUUUUCGCGCCCAGUUGCAA AAA

[0126] 在一个实施方案中,5′UTR包含SEQ ID NO:50的变体。在一个实施方案中,SEQ ID NO:50的变体包含式A的核酸序列:

[0127] $$GGAAAUCGCAAAA(N_2)_X(N_3)_XCU(N_4)_X(N_5)_XCGCGU$$
$$UAGAUUUCUUUUAGUUUUCUN_6N_7CAACUAGCA$$
$$AGCUUUUUGUUCUCGCC(N_8CC)x\ (SEQ\ ID\ NO:59),$$

[0128] 其中:

[0129] $(N_2)_x$是尿嘧啶并且x是0至5的整数,例如,其中x＝3或4;

[0130] $(N_3)_x$是乌嘌呤并且x是0至1的整数;

[0131] $(N_4)_x$是胞嘧啶并且x是0至1的整数;

[0132] $(N_5)_x$是尿嘧啶并且x是0至5的整数,例如,其中x＝2或3;

[0133] N_6是尿嘧啶或胞嘧啶;

[0134] N_7是尿嘧啶或乌嘌呤;

[0135] N_8是腺嘌呤或鸟嘌呤并且x是0至1的整数。

[0136] 在一个实施方案中,$(N_2)_x$是尿嘧啶并且x是0。在一个实施方案中,$(N_2)_x$是尿嘧啶并且x是1。在一个实施方案中,$(N_2)_x$是尿嘧啶并且x是2。在一个实施方案中,$(N_2)_x$是尿嘧啶并且x是3。在一个实施方案中,$(N_2)_x$是尿嘧啶并且x是4。在一个实施方案中,$(N_2)_x$是尿嘧啶并且x是5。

[0137] 在一个实施方案中,$(N_3)_x$是鸟嘌呤并且x是0。在一个实施方案中,$(N_3)_x$是鸟嘌呤并且x是1。

[0138] 在一个实施方案中,$(N_4)_x$是胞嘧啶并且x是0。在一个实施方案中,$(N_4)_x$是胞嘧啶并且x是1。

[0139] 在一个实施方案中,$(N_5)_x$是尿嘧啶并且x是0。在一个实施方案中,$(N_5)_x$是尿嘧啶并且x是1。在一个实施方案中,$(N_5)_x$是尿嘧啶并且x是2。在一个实施方案中,$(N_5)_x$是尿嘧啶并且x是3。在一个实施方案中,$(N_5)_x$是尿嘧啶并且x是4。在一个实施方案中,$(N_5)_x$是尿嘧啶并且x是5。

[0140] 在一个实施方案中,N6是尿嘧啶。在一个实施方案中,N6是胞嘧啶。

[0141] 在一个实施方案中,N7是尿嘧啶。在一个实施方案中,N7是鸟嘌呤。

[0142] 在一个实施方案中,N8是腺嘌呤并且x是0。在一个实施方案中,N8是腺嘌呤并且x是1。

[0143] 在一个实施方案中,N8是乌嘌呤并且x是0。在一个实施方案中,N8是鸟嘌呤并且x是1。

[0144] 在一个实施方案中,5′UTR包含SEQ ID NO:50的变体。在一个实施方案中,SEQ ID NO:50的变体包含与SEQ ID NO:50具有至少50%、60%、70%、80%、85%、90%、95%、96%、97%、98%或99%同一性的序列。在一个实施方案中,SEQ ID NO:50的变体包含与SEQ ID NO:50具有至少50%同一性的序列。在一个实施方案中,SEQ ID NO:50的变体包含与SEQ ID NO:50具有至少60%同一性的序列。在一个实施方案中,SEQ ID NO:50的变体包含与SEQ ID NO:50具有至少70%同一性的序列。在一个实施方案中,SEQ ID NO:50的变体包含与SEQ ID NO:50具有至少80%同一性的序列。在一个实施方案中,SEQ ID NO:50的变体包含与SEQ ID NO:50具有至少90%同一性的序列。在一个实施方案中,SEQ ID NO:50的变体包含与SEQ ID NO:50具有至少95%同一性的序列。在一个实施方案中,SEQ ID NO:50的变体包含与SEQ ID NO:50具有至少96%同一性的序列。在一个实施方案中,SEQ ID NO:50的变体包含与SEQ ID NO:50具有至少97%同一性的序列。在一个实施方案中,SEQ ID NO:50的变体包含与SEQ ID NO:50具有至少98%同一性的序列。在一个实施方案中,SEQ ID NO:50的变体包含与SEQ ID NO:50具有至少99%同一性的序列。

[0145] 在一个实施方案中,5′UTR包含SEQ ID NO:64的变体。在一个实施方案中,SEQ ID NO:64的变体包含与SEQ ID NO:64具有至少64%、60%、70%、80%、85%、90%、95%、96%、97%、98%或99%同一性的序列。在一个实施方案中,SEQ ID NO:64的变体包含与SEQ ID NO:64具有至少64%同一性的序列。在一个实施方案中,SEQ ID NO:64的变体包含与SEQ ID NO:64具有至少60%同一性的序列。在一个实施方案中,SEQ ID NO:64的变体包含与SEQ ID NO:64具有至少70%同一性的序列。在一个实施方案中,SEQ ID NO:64的变体包含与SEQ ID NO:64具有至少80%同一性的序列。在一个实施方案中,SEQ ID NO:64的变体包含与SEQ ID NO:64具有至少90%同一性的序列。在一个实施方案中,SEQ ID NO:64的变体包含与SEQ ID NO:64具有至少95%同一性的序列。在一个实施方案中,SEQ ID NO:64的变体包含与SEQ ID NO:64具有至少96%同一性的序列。在一个实施方案中,SEQ ID NO:64的变体包含与SEQ ID NO:64具有至少97%同一性的序列。在一个实施方案中,SEQ ID NO:64的变体包含与SEQ ID NO:64具有至少98%同一性的序列。在一个实施方案中,SEQ ID NO:64的变体包含与SEQ ID

NO:64具有至少99%同一性的序列。

[0146] 在一个实施方案中,5′UTR包含SEQ ID NO:55的变体。在一个实施方案中,SEQ ID NO:55的变体包含与SEQ ID NO:55具有至少55%、60%、70%、80%、85%、90%、95%、96%、97%、98%或99%同一性的序列。在一个实施方案中,SEQ ID NO:55的变体包含与SEQ ID NO:55具有至少55%同一性的序列。在一个实施方案中,SEQ ID NO:55的变体包含与SEQ ID NO:55具有至少60%同一性的序列。在一个实施方案中,SEQ ID NO:55的变体包含与SEQ ID NO:55具有至少70%同一性的序列。在一个实施方案中,SEQ ID NO:55的变体包含与SEQ ID NO:55具有至少80%同一性的序列。在一个实施方案中,SEQ ID NO:55的变体包含与SEQ ID NO:55具有至少90%同一性的序列。在一个实施方案中,SEQ ID NO:55的变体包含与SEQ ID NO:55具有至少95%同一性的序列。在一个实施方案中,SEQ ID NO:55的变体包含与SEQ ID NO:55具有至少96%同一性的序列。在一个实施方案中,SEQ ID NO:55的变体包含与SEQ ID NO:55具有至少97%同一性的序列。在一个实施方案中,SEQ ID NO:55的变体包含与SEQ ID NO:55具有至少98%同一性的序列。在一个实施方案中,SEQ ID NO:55的变体包含与SEQ ID NO:55具有至少99%同一性的序列。

[0147] 在一个实施方案中,5′UTR包含SEQ ID NO:56的变体。在一个实施方案中,SEQ ID NO:56的变体包含与SEQ ID NO:56具有至少56%、60%、70%、80%、85%、90%、95%、96%、97%、98%或99%同一性的序列。在一个实施方案中,SEQ ID NO:56的变体包含与SEQID NO:56具有至少56%同一性的序列。在一个实施方案中,SEQ IDNO:56的变体包含与SEQ ID NO:56具有至少60%同一性的序列。在一个实施方案中,SEQ ID NO:56的变体包含与SEQ ID NO:56具有至少70%同一性的序列。在一个实施方案中,SEQ ID NO:56的变体包含与SEQ ID NO:56具有至少80%同一性的序列。在一个实施方案中,SEQ ID NO:56的变体包含与SEQ ID NO:56具有至少90%同一性的序列。在一个实施方案中,SEQ ID NO:56的变体包含与SEQ ID NO:56具有至少95%同一性的序列。在一个实施方案中,SEQ ID NO:56的变体包含与SEQ ID NO:56具有至少96%同一性的序列。在一个实施方案中,SEQ ID NO:56的变体包含与SEQ ID NO:56具有至少97%同一性的序列。在一个实施方案中,SEQ ID NO:56的变体包含与SEQ ID NO:56具有至少98%同一性的序列。在一个实施方案中,SEQ IDNO:56的变体包含与SEQ ID No:56具有至少99%同一性的序列。

[0148] 在一个实施方案中,5′UTR包含SEQ ID NO:58的变体。在一个实施方案中,SEQ ID NO:58的变体包含与SEQ ID NO:58具有至少58%、60%、70%、80%、85%、90%、95%、96%、97%、98%或99%同一性的序列。在一个实施方案中,SEQ ID NO:58的变体包含与SEQ ID NO:58具有至少58%同一性的序列。在一个实施方案中,SEQ ID NO:58的变体包含与SEQ ID NO:58具有至少60%同一性的序列。在一个实施方案中,SEQ ID NO:58的变体包含与SEQ ID NO:58具有至少70%同一性的序列。在一个实施方案中,SEQ ID NO:58的变体包含与SEQ ID NO:58具有至少80%同一性的序列。在一个实施方案中,SEQ ID NO:58的变体包含与SEQ ID NO:58具有至少90%同一性的序列。在一个实施方案中,SEQ ID NO:58的变体包含与SEQ ID NO:58具有至少95%同一性的序列。在一个实施方案中,SEQ ID NO:58的变体包含与SEQ ID NO:58具有至少96%同一性的序列。在一个实施方案中,SEQ ID NO:58的变体包含与SEQ ID NO:58具有至少97%同一性的序列。在一个实施方案中,SEQ ID NO:58的变体包含与SEQID NO:58具有至少98%同一性的序列。在一个实施方案中,SEQ ID NO:58的变体包含与SEQ ID

NO:58具有至少99%同一性的序列。

[0149] 在一个实施方案中,5′UTR包含SEQ ID NO:76的变体。在一个实施方案中,SEQ ID NO:76的变体包含与SEQ ID NO:76具有至少76%、60%、70%、80%、85%、90%、95%、96%、97%、98%或99%同一性的序列。在一个实施方案中,SEQ ID NO:76的变体包含与SEQ ID NO:76具有至少76%同一性的序列。在一个实施方案中,SEQ ID NO:76的变体包含与SEQ ID NO:76具有至少60%同一性的序列。在一个实施方案中,SEQ ID NO:76的变体包含与SEQ ID NO:76具有至少70%同一性的序列。在一个实施方案中,SEQ ID NO:76的变体包含与SEQ ID NO:76具有至少80%同一性的序列。在一个实施方案中,SEQ ID NO:76的变体包含与SEQ ID NO:76具有至少90%同一性的序列。在一个实施方案中,SEQ ID NO:76的变体包含与SEQ ID NO:76具有至少95%同一性的序列。在一个实施方案中,SEQ ID NO:76的变体包含与SEQ ID NO:76具有至少96%同一性的序列。在一个实施方案中,SEQ ID NO:76的变体包含与SEQ ID NO:76具有至少97%同一性的序列。在一个实施方案中,SEQ ID NO:76的变体包含与SEQ ID NO:76具有至少98%同一性的序列。在一个实施方案中,SEQ ID NO:76的变体包含与SEQ ID NO:76具有至少99%同一性的序列。

[0150] 在一个实施方案中,5′UTR包含SEQ ID NO:78的变体。在一个实施方案中,SEQ ID NO:78的变体包含与SEQ ID NO:78具有至少78%、60%、70%、80%、85%、90%、95%、96%、97%、98%或99%同一性的序列。在一个实施方案中,SEQ ID NO:78的变体包含与SEQID NO:78具有至少78%同一性的序列。在一个实施方案中,SEQ ID NO:78的变体包含与SEQ ID NO:78具有至少60%同一性的序列。在一个实施方案中,SEQ ID NO:78的变体包含与SEQ ID NO:78具有至少70%同一性的序列。在一个实施方案中,SEQ ID NO:78的变体包含与SEQ ID NO:78具有至少80%同一性的序列。在一个实施方案中,SEQ ID NO:78的变体包含与SEQ ID NO:78具有至少90%同一性的序列。在一个实施方案中,SEQ ID NO:78的变体包含与SEQ ID NO:78具有至少95%同一性的序列。在一个实施方案中,SEQ ID NO:78的变体包含与SEQ ID NO:78具有至少96%同一性的序列。在一个实施方案中,SEQ ID NO:78的变体包含与SEQ ID NO:78具有至少97%同一性的序列。在一个实施方案中,SEQ ID NO:78的变体包含与SEQ ID NO:78具有至少98%同一性的序列。在一个实施方案中,SEQ ID NO:78的变体包含与SEQ ID NO:78具有至少99%同一性的序列。

[0151] 在一个实施方案中,SEQ ID NO:50的变体包含至少5%、10%、20%、30%、40%、50%、60%、70%或80%的尿苷含量。在一个实施方案中,SEQ ID NO:50的变体包含至少5%的尿苷含量。在一个实施方案中,SEQ ID NO:50的变体包含至少10%的尿苷含量。在一个实施方案中,SEQ ID NO:50的变体包含至少20%的尿苷含量。在一个实施方案中,SEQ ID NO:50的变体包含至少30%的尿苷含量。在一个实施方案中,SEQ ID NO:50的变体包含至少40%的尿苷含量。在一个实施方案中,SEQ ID NO:50的变体包含至少50%的尿苷含量。在一个实施方案中,SEQ ID NO:50的变体包含至少60%的尿苷含量。在一个实施方案中,SEQ ID NO:50的变体包含至少70%的尿苷含量。在一个实施方案中,SEQ ID NO:50的变体包含至少80%的尿苷含量。

[0152] 在一个实施方案中,SEQ ID NO:64的变体包含至少5%、10%、20%、30%、40%、64%、60%、70%或80%的尿苷含量。在一个实施方案中,SEQ ID NO:64的变体包含至少5%的尿苷含量。在一个实施方案中,SEQ ID NO:64的变体包含至少10%的尿苷含量。在一个实

施方案中,SEQ ID NO:64的变体包含至少20%的尿苷含量。在一个实施方案中,SEQ ID NO:64的变体包含至少30%的尿苷含量。在一个实施方案中,SEQ ID NO:64的变体包含至少40%的尿苷含量。在一个实施方案中,SEQ ID NO:64的变体包含至少64%的尿苷含量。在一个实施方案中,SEQ ID NO:64的变体包含至少60%的尿苷含量。在一个实施方案中,SEQ ID NO:64的变体包含至少70%的尿苷含量。在一个实施方案中,SEQ ID NO:64的变体包含至少80%的尿苷含量。

[0153] 在一个实施方案中,SEQ ID NO:55的变体包含至少5%、10%、20%、30%、40%、55%、60%、70%或80%的尿苷含量。在一个实施方案中,SEQ ID NO:55的变体包含至少5%的尿苷含量。在一个实施方案中,SEQ ID NO:55的变体包含至少10%的尿苷含量。在一个实施方案中,SEQ ID NO:55的变体包含至少20%的尿苷含量。在一个实施方案中,SEQ ID NO:55的变体包含至少30%的尿苷含量。在一个实施方案中,SEQ ID NO:55的变体包含至少40%的尿苷含量。在一个实施方案中,SEQ ID NO:55的变体包含至少55%的尿苷含量。在一个实施方案中,SEQ ID NO:55的变体包含至少60%的尿苷含量。在一个实施方案中,SEQ ID NO:55的变体包含至少70%的尿苷含量。在一个实施方案中,SEQ ID NO:55的变体包含至少80%的尿苷含量。

[0154] 在一个实施方案中,SEQ ID NO:56的变体包含至少5%、10%、20%、30%、40%、56%、60%、70%或80%的尿苷含量。在一个实施方案中,SEQ ID NO:56的变体包含至少5%的尿苷含量。在一个实施方案中,SEQ ID NO:56的变体包含至少10%的尿苷含量。在一个实施方案中,SEQ ID NO:56的变体包含至少20%的尿苷含量。在一个实施方案中,SEQ ID NO:56的变体包含至少30%的尿苷含量。在一个实施方案中,SEQ ID NO:56的变体包含至少40%的尿苷含量。在一个实施方案中,SEQ ID NO:56的变体包含至少56%的尿苷含量。在一个实施方案中,SEQ ID NO:56的变体包含至少60%的尿苷含量。在一个实施方案中,SEQ ID NO:56的变体包含至少70%的尿苷含量。在一个实施方案中,SEQ ID NO:56的变体包含至少80%的尿苷含量。

[0155] 在一个实施方案中,SEQ ID NO:58的变体包含至少5%、10%、20%、30%、40%、58%、60%、70%或80%的尿苷含量。在一个实施方案中,SEQ ID NO:58的变体包含至少5%的尿苷含量。在一个实施方案中,SEQ ID NO:58的变体包含至少10%的尿苷含量。在一个实施方案中,SEQ ID NO:58的变体包含至少20%的尿苷含量。在一个实施方案中,SEQ ID NO:58的变体包含至少30%的尿苷含量。在一个实施方案中,SEQ ID NO:58的变体包含至少40%的尿苷含量。在一个实施方案中,SEQ ID NO:58的变体包含至少58%的尿苷含量。在一个实施方案中,SEQ ID NO:58的变体包含至少60%的尿苷含量。在一个实施方案中,SEQ ID NO:58的变体包含至少70%的尿苷含量。在一个实施方案中,SEQ ID NO:58的变体包含至少80%的尿苷含量。

[0156] 在一个实施方案中,SEQ ID NO:76的变体包含至少5%、10%、20%、30%、40%、76%、60%、70%或80%的尿苷含量。在一个实施方案中,SEQ ID NO:76的变体包含至少5%的尿苷含量。在一个实施方案中,SEQ ID NO:76的变体包含至少10%的尿苷含量。在一个实施方案中,SEQ ID NO:76的变体包含至少20%的尿苷含量。在一个实施方案中,SEQ ID NO:76的变体包含至少30%的尿苷含量。在一个实施方案中,SEQ ID NO:76的变体包含至少40%的尿苷含量。在一个实施方案中,SEQ ID NO:76的变体包含至少76%的尿苷含量。在一

个实施方案中,SEQ ID NO:76的变体包含至少60%的尿苷含量。在一个实施方案中,SEQ ID NO:76的变体包含至少70%的尿苷含量。在一个实施方案中,SEQ ID NO:76的变体包含至少80%的尿苷含量。

[0157] 在一个实施方案中,SEQ ID NO:78的变体包含至少5%、10%、20%、30%、40%、78%、60%、70%或80%的尿苷含量。在一个实施方案中,SEQ ID NO:78的变体包含至少5%的尿苷含量。在一个实施方案中,SEQ ID NO:78的变体包含至少10%的尿苷含量。在一个实施方案中,SEQ ID NO:78的变体包含至少20%的尿苷含量。在一个实施方案中,SEQ ID NO:78的变体包含至少30%的尿苷含量。在一个实施方案中,SEQ ID NO:78的变体包含至少40%的尿苷含量。在一个实施方案中,SEQ ID NO:78的变体包含至少78%的尿苷含量。在一个实施方案中,SEQ ID NO:78的变体包含至少60%的尿苷含量。在一个实施方案中,SEQ ID NO:78的变体包含至少70%的尿苷含量。在一个实施方案中,SEQ ID NO:78的变体包含至少80%的尿苷含量。

[0158] 在一个实施方案中,SEQ ID NO:50的变体包含至少2、3、4、5、6或7个连续尿苷(例如,聚尿苷序列段)。在一个实施方案中,SEQ ID NO:50的变体中的聚尿苷序列段包含至少1-7、2-7、3-7、4-7、5-7、6-7、1-6、1-5、1-4、1-3、1-2、2-6或3-5个连续尿苷。在一个实施方案中,SEQ ID NO:50的变体中的聚尿苷序列段包含4个连续尿苷。在一个实施方案中,SEQ ID NO:50的变体中的聚尿苷序列段包含5个连续尿苷。

[0159] 在一个实施方案中,SEQ ID NO:64的变体包含至少2、3、4、5、6或7个连续尿苷(例如,聚尿苷序列段)。在一个实施方案中,SEQ ID NO:64的变体中的聚尿苷序列段包含至少1-7、2-7、3-7、4-7、5-7、6-7、1-6、1-5、1-4、1-3、1-2、2-6或3-5个连续尿苷。在一个实施方案中,SEQ ID NO:64的变体中的聚尿苷序列段包含4个连续尿苷。在一个实施方案中,SEQ ID NO:64的变体中的聚尿苷序列段包含5个连续尿苷。

[0160] 在一个实施方案中,SEQ ID NO:55的变体包含至少2、3、4、5、6或7个连续尿苷(例如,聚尿苷序列段)。在一个实施方案中,SEQ ID NO:55的变体中的聚尿苷序列段包含至少1-7、2-7、3-7、4-7、5-7、6-7、1-6、1-5、1-4、1-3、1-2、2-6或3-5个连续尿苷。在一个实施方案中,SEQ ID NO:55的变体中的聚尿苷序列段包含4个连续尿苷。在一个实施方案中,SEQ ID NO:55的变体中的聚尿苷序列段包含5个连续尿苷。

[0161] 在一个实施方案中,SEQ ID NO:56的变体包含至少2、3、4、5、6或7个连续尿苷(例如,聚尿苷序列段)。在一个实施方案中,SEQ ID NO:56的变体中的聚尿苷序列段包含至少1-7、2-7、3-7、4-7、5-7、6-7、1-6、1-5、1-4、1-3、1-2、2-6或3-5个连续尿苷。在一个实施方案中,SEQ ID NO:56的变体中的聚尿苷序列段包含4个连续尿苷。在一个实施方案中,SEQ ID NO:56的变体中的聚尿苷序列段包含5个连续尿苷。

[0162] 在一个实施方案中,SEQ ID NO:58的变体包含至少2、3、4、5、6或7个连续尿苷(例如,聚尿苷序列段)。在一个实施方案中,SEQ ID NO:58的变体中的聚尿苷序列段包含至少1-7、2-7、3-7、4-7、5-7、6-7、1-6、1-5、1-4、1-3、1-2、2-6或3-5个连续尿苷。在一个实施方案中,SEQ ID NO:58的变体中的聚尿苷序列段包含4个连续尿苷。在一个实施方案中,SEQ ID NO:58的变体中的聚尿苷序列段包含5个连续尿苷。

[0163] 在一个实施方案中,SEQ ID NO:76的变体包含至少2、3、4、5、6或7个连续尿苷(例如,聚尿苷序列段)。在一个实施方案中,SEQ ID NO:76的变体中的聚尿苷序列段包含至少

1-7、2-7、3-7、4-7、5-7、6-7、1-6、1-5、1-4、1-3、1-2、2-6或3-5个连续尿苷。在一个实施方案中,SEQ ID NO:76的变体中的聚尿苷序列段包含4个连续尿苷。在一个实施方案中,SEQ ID NO:76的变体中的聚尿苷序列段包含5个连续尿苷。

[0164] 在一个实施方案中,SEQ ID NO:78的变体包含至少2、3、4、5、6或7个连续尿苷(例如,聚尿苷序列段)。在一个实施方案中,SEQ ID NO:78的变体中的聚尿苷序列段包含至少1-7、2-7、3-7、4-7、5-7、6-7、1-6、1-5、1-4、1-3、1-2、2-6或3-5个连续尿苷。在一个实施方案中,SEQ ID NO:78的变体中的聚尿苷序列段包含4个连续尿苷。在一个实施方案中,SEQ ID NO:78的变体中的聚尿苷序列段包含5个连续尿苷。

[0165] 在一个实施方案中,SEQ ID NO:50的变体包含1、2、3、4、5、6、7、8、9、10、11、12、13、14或15个聚尿苷序列段。在一个实施方案中,SEQ ID NO:50的变体包含3个聚尿苷序列段。在一个实施方案中,SEQ ID NO:50的变体包含4个聚尿苷序列段。在一个实施方案中,SEQ ID NO:50的变体包含5个聚尿苷序列段。

[0166] 在一个实施方案中,SEQ ID NO:64的变体包含1、2、3、4、5、6、7、8、9、10、11、12、13、14或15个聚尿苷序列段。在一个实施方案中,SEQ ID NO:64的变体包含3个聚尿苷序列段。在一个实施方案中,SEQ ID NO:64的变体包含4个聚尿苷序列段。在一个实施方案中,SEQ ID NO:64的变体包含5个聚尿苷序列段。

[0167] 在一个实施方案中,SEQ ID NO:55的变体包含1、2、3、4、5、6、7、8、9、10、11、12、13、14或15个聚尿苷序列段。在一个实施方案中,SEQ ID NO:55的变体包含3个聚尿苷序列段。在一个实施方案中,SEQ ID NO:55的变体包含4个聚尿苷序列段。在一个实施方案中,SEQ ID NO:55的变体包含5个聚尿苷序列段。

[0168] 在一个实施方案中,SEQ ID NO:56的变体包含1、2、3、4、5、6、7、8、9、10、11、12、13、14或15个聚尿苷序列段。在一个实施方案中,SEQ ID NO:56的变体包含3个聚尿苷序列段。在一个实施方案中,SEQ ID NO:56的变体包含4个聚尿苷序列段。在一个实施方案中,SEQ ID NO:56的变体包含5个聚尿苷序列段。

[0169] 在一个实施方案中,SEQ ID NO:58的变体包含1、2、3、4、5、6、7、8、9、10、11、12、13、14或15个聚尿苷序列段。在一个实施方案中,SEQ ID NO:58的变体包含3个聚尿苷序列段。在一个实施方案中,SEQ ID NO:58的变体包含4个聚尿苷序列段。在一个实施方案中,SEQ ID NO:58的变体包含5个聚尿苷序列段。

[0170] 在一个实施方案中,SEQ ID NO:76的变体包含1、2、3、4、5、6、7、8、9、10、11、12、13、14或15个聚尿苷序列段。在一个实施方案中,SEQ ID NO:76的变体包含3个聚尿苷序列段。在一个实施方案中,SEQ ID NO:76的变体包含4个聚尿苷序列段。在一个实施方案中,SEQ ID NO:76的变体包含5个聚尿苷序列段。

[0171] 在一个实施方案中,SEQ ID NO:78的变体包含1、2、3、4、5、6、7、8、9、10、11、12、13、14或15个聚尿苷序列段。在一个实施方案中,SEQ ID NO:78的变体包含3个聚尿苷序列段。在一个实施方案中,SEQ ID NO:78的变体包含4个聚尿苷序列段。在一个实施方案中,SEQ ID NO:78的变体包含5个聚尿苷序列段。

[0172] 在一个实施方案中,聚尿苷序列段中的一者或多者与不同聚尿苷序列段相邻。在一个实施方案中,聚尿苷序列段中的每一者,例如全部彼此相邻,例如,聚尿苷序列段全部是相连的。

[0173] 在一个实施方案中,聚尿苷序列段中的一者或多者由1、2、3、4、5、6、7、8、9、10、11、2、13、14、15、16、17、18、19、20、30、40、50或60个核苷酸隔开。在一个实施方案中,聚尿苷序列段中的每一者,例如全部由1、2、3、4、5、6、7、8、9、10、11、2、13、14、15、16、17、18、19、20、30、40、50或60个核苷酸隔开。

[0174] 在一个实施方案中,第一聚尿苷序列段和第二聚尿苷序列段彼此相邻。

[0175] 在一个实施方案中,后续的,例如第三、第四、第五、第六或第七、第八、第九或第十聚尿苷序列段由1、2、3、4、5、6、7、8、9、10、11、2、13、14、15、16、17、18、19、20、30、40、50或60个核苷酸与第一聚尿苷序列段、第二聚尿苷序列段或后续聚尿苷序列段中的任一者隔开。

[0176] 在一个实施方案中,第一聚尿苷序列段由1、2、3、4、5、6、7、8、9、10、11、2、13、14、15、16、17、18、19、20、30、40、50或60个核苷酸与后续聚尿苷序列段,例如第二、第三、第四、第五、第六或第七、第八、第九或第十聚尿苷序列段隔开。在一个实施方案中,后续聚尿苷序列段中的一者或多者与不同聚尿苷序列段相邻。

[0177] 在一个实施方案中,5′UTR包含Kozak序列,例如,GCCRCC核苷酸序列(SEQ ID NO:79),其中R是腺嘌呤或鸟嘌呤。在一个实施方案中,Kozak序列位于5′UTR序列的3′端。

[0178] 在一个实施方案中,本文所公开的包含5′UTR序列的多核苷酸包含编码有效载荷,例如治疗性或预防性有效载荷的编码区。

[0179] 在一个方面,本文所公开的包含5′UTR序列的多核苷酸(例如,mRNA)被配制成LNP。在一个实施方案中,LNP组合物包含:(i)可电离脂质,例如,氨基脂质;(ii)固醇或其它结构脂质;(iii)非阳离子辅助脂质或磷脂;以及(iv)PEG-脂质。

[0180] 在另一个方面,本公开的LNP组合物用于治疗疾病或病症的方法中,或用于抑制受试者的免疫反应的方法中。

[0181] 在一个方面,包含本文所公开的编码治疗性有效载荷或预防性有效载荷(例如,如本文所描述)的多核苷酸的LNP组合物可以与额外剂(例如,如本文所描述)一起施用。

[0182] b.终止元件件+3′UTR序列

[0183] 翻译终止密码子UAA、UAG和UGA是遗传密码的重要组分并且是mRNA翻译终止的信号。在蛋白质合成期间,终止密码子与蛋白质释放因子相互作用并且这种相互作用可以调节核糖体活性,由此对翻译具有影响(Tate WP等人,(2018)Biochem Soc Trans,46(6):1615-162)。

[0184] 3′UTR序列已经显示影响mRNA的翻译、半衰期和亚细胞定位(Mayr C.,Cold Spring Harb Persp Biol 2019年10月1日;11(10):a034728)。

[0185] 本文尤其公开了一种编码多肽的多核苷酸,所述多核苷酸具有终止元件与3′UTR的组合,其向由所述多核苷酸编码的多肽或多核苷酸本身赋予增加的半衰期、增加的表达和/或增加的活性。在一个实施方案中,本文所公开的多核苷酸包含:(a)5′-UTR;(b)编码区;以及(c)终止元件和3′-UTR(例如,如本文所描述),以及包含所述多核苷酸的LNP组合物。

[0186] 本文尤其公开了一种编码多肽的多核苷酸,所述多核苷酸尤其包含3′UTR。在一个实施方案中,本文所公开的多核苷酸包含:(a)5′-UTR(例如,如表1中所提供或其变体或片段);(b)编码区;以及(c)终止元件和3′-UTR(例如,如表2中所提供或其变体或片段),以及包含所述多核苷酸的LNP组合物。在一个实施方案中,多核苷酸包含3′-UTR,所述3′-UTR包

含表2中所提供的序列或其变体或片段(例如,其功能变体或片段)。将了解,此类3′UTR并入自然界中未发现的构建体中,例如,此类3′UTR是合成的,在序列中从天然存在的3′UTR改变,是在自然界中发现的那些的截短或加长形式,包含化学修饰的碱基,处于不同于自然界中可能发现的那些的ORF序列的3′,等等。

[0187] 在一个实施方案中,3′UTR包含表2中所提供的序列或与表2中所提供的3′UTR序列具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列或其变体或片段。在一个实施方案中,3′UTR包含与SEQ ID NO:139、SEQ ID NO:140、SEQ ID NO:141、SEQ ID NO:142、SEQ ID NO:143、SEQ ID NO:144、SEQ ID NO:145、SEQ ID NO:146或SEQ ID NO:147具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0188] 表2:3′UTR序列(终止盒为斜体;miR结合位点为粗体)

[0189]

SEQ ID NO:	序列名称	序列
SEQ ID NO: 139	D1	*UAAAGCUCCCCCGGGG*GCCUCGGUGGCCUAGCUUCU UGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGUACCCCCGUGGUCUUUGAAUAAA GUCUGAGUGGGCGGC
SEQ ID NO: 140	D2	*UAAGUCUAA*GCUGGAGCCUCCUGAGAGACCUGUGU GAACUAUUGAGAAGAUCGGAACAGCUCCUUACUC UGAGGAAGUUGGUACCCCCGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGC
SEQ ID NO: 141	D3	*UAAAGCUCCCCCGGGG***CAAACACCAUUGUCACACU CCA**GCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CC**CAAACACCAUUGUCACACUCCA**UCCCCCCAGC CCCUCCUCCCCUUCCUGCACCCGUACCCCC**CAAAC ACCAUUGUCACACUCCA**GUGGUCUUUGAAUAAA GUCUGAGUGGGCGGC (miR122 结合位点为粗体)
SEQ ID NO: 142	D4	*UAAAGCUCCCCCGGGG***UCCAUAAAGUAGGAAACAC UACA**GCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCC**CAAACACCAUUGUCACACUCCA**UCC CCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (miR-142-3p 和 miR122 结合位点为粗体)
SEQ ID NO: 143	D5	*UAAAGCUCCCCCGGGG*GCCUCGGUGGCCUAGCUUCU UGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGUACCCCC**CAAACACCAUUGUCAC ACUCCA**GUGGUCUUUGAAUAAAGUCUGAGUGGGC

[0190]

		GGC (miR122 结合位点为粗体)
SEQ ID NO: 144	D6	*UAAGCCCCUCCGGGG***CAAACACCAUUGUCACACU** **CCA**GCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CC**CAAACACCAUUGUCACACUCCA**UCCCCCCAGC CCCUCCUCCCCUUCCUGCACCCGUACCCCC**CAAAC** **ACCAUUGUCACACUCCA**GUGGUCUUUGAAUAAA GUCUGAGUGGGCGGC (miR122 结合位点为粗体)
SEQ ID NO: 145	D7	*UAAGCCCCUCCGGGG***UCCAUAAAGUAGGAAACAC** **UACA**GCCUCGGUGGCCUAGCUUCUUGCCCCUUGG GCC**UCCAUAAAGUAGGAAACACUACA**UCCCCCCA GCCCCUCCUCCCCUUCCUGCACCCGUACCCCC**CGC** **AUUAUUACUCACGGUACGA**GUGGUCUUUGAAUA AAGUCUGAGUGGGCGGC (miR-142-3p 和 miR-126-3p 结合位点为粗体)
SEQ ID NO: 146	D8	*UAAGCCCCUCCGGGG***UCCAUAAAGUAGGAAACAC** **UACA**GCUGGAGCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCC**CAAACACCAUUGUCACACUCCA**UCC CCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (miR-142-3p 和 miR122 结合位点为粗体)
SEQ ID NO: 147	D9	*UAAGCCCCUCCGGGG*GCCUCGGUGGCCUAGCUUCU UGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGUACCCCC**CAAACACCAUUGUCAC** **ACUCCA**GUGGUCUUUGAAUAAAGUCUGAGUGGGC GGC (miR122 结合位点为粗体)

[0191] 在一个实施方案中,多核苷酸包含终止元件和3′-UTR,其中序列是(终止元件为斜体):

[0192] *UAAAGCUCCCCGGGG*GCCUCGGUGGCCUAGCUUCUUGCC CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGU ACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (**SEQ ID NO:139**)或其变体或片段(例如,缺少SEQ ID NO:139的核苷酸的初始一个、两个、三个、四个、五个、六个或更多个核苷酸的片段)。

[0193] 在一个实施方案中,具有SEQ ID NO:139中所提供的3′UTR序列或其变体或片段的多核苷酸使得多核苷酸的半衰期增加,例如,多核苷酸的半衰期增加约1.5-10倍。在一个实施方案中,半衰期增加约1.5、2、3、4、5、6、7、8、9或10倍或更多。在一个实施方案中,半衰期增加约1.5倍或更多。在一个实施方案中,半衰期增加约2倍或更多。在一个实施方案中,半衰期增加约3倍或更多。在一个实施方案中,半衰期增加约4倍或更多。在一个实施方案中,半衰期增加约5倍或更多。在一个实施方案中,半衰期增加约6倍或更多。在一个实施方案中,半衰期增加约7倍或更多。在一个实施方案中,半衰期增加约8倍。在一个实施方案中,半衰期增加约9倍或更多。在一个实施方案中,半衰期增加约10倍或更多。

[0194] 在一个实施方案中,具有SEQ ID NO:139中所提供的3′UTR序列或其变体或片段的多核苷酸使得由多核苷酸编码的多肽的水平和/或活性(例如,输出)增加。

[0195] 在一个实施方案中,将所述增加与不具有3′UTR、具有不同3′UTR或不具有SEQ ID

NO:139的3′UTR或其变体或片段的其它方面类似的多核苷酸进行比较。

[0196] 在一个实施方案中,多核苷酸包含SEQ ID NO:139中所提供的3′UTR序列或与SEQ ID NO:139中所提供的3′UTR序列具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0197] 在一个实施方案中,多核苷酸包含终止元件和3'-UTR,其中序列是(终止元件为斜体):

[0198] *UAAGUCUAA*GCUGGAGCCUCCUGAGAGACCUGUGUGAAC UAUUGAGAAGAUCGGAACAGCUCCUUACUCUGAGGAAGUUG GUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (**SEQ ID NO:140**)或其变体或片段(例如,缺少SEQ ID NO:140的核苷酸的初始一个、两个、三个、四个、五个、六个或更多个核苷酸的片段)。

[0199] 在一个实施方案中,具有SEQ ID NO:140中所提供的3′UTR序列或其变体或片段的多核苷酸使得多核苷酸的半衰期增加,例如,多核苷酸的半衰期增加约1.5--10倍。在一个实施方案中,半衰期增加约1.5、2、3、4、5、6、7、8、9或10倍或更多。在一个实施方案中,半衰期增加约1.5倍或更多。在一个实施方案中,半衰期增加约2倍或更多。在一个实施方案中,半衰期增加约3倍或更多。在一个实施方案中,半衰期增加约4倍或更多。在一个实施方案中,半衰期增加约5倍或更多。在一个实施方案中,半衰期增加约6倍或更多。在一个实施方案中,半衰期增加约7倍或更多。在一个实施方案中,半衰期增加约8倍。在一个实施方案中,半衰期增加约9倍或更多。在一个实施方案中,半衰期增加约10倍或更多。

[0200] 在一个实施方案中,具有SEQ ID NO:140中所提供的3′UTR序列或其变体或片段的多核苷酸使得由多核苷酸编码的多肽的水平和/或活性(例如,输出)增加。

[0201] 在一个实施方案中,将所述增加与不具有3′UTR、具有不同3′UTR或不具有SEQ ID NO:140的3′UTR或其变体或片段的其它方面类似的多核苷酸进行比较。

[0202] 在一个实施方案中,多核苷酸包含SEQ ID NO:140中所提供的3′UTR序列或与SEQ ID NO:140中所提供的3′UTR序列具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0203] 在一个实施方案中,多核苷酸包含终止元件和3'-UTR,其中序列是(终止元件为斜体):

[0204] *UAAAGCUCCCCGGGGC*AAACACCAUUGUCACACUCCAGCC UCGGUGGCCUAGCUUCUUGCCCCUUGGGCCCAAACACCAUUG UCACACUCCAUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCG UACCCCCCAAACACCAUUGUCACACUCCAGUGGUCUUUGAAU AAAGUCUGAGUGGGCGGC (**SEQ ID NO:141**)或其变体或片段(例如,缺少SEQ ID NO:141的核苷酸的初始一个、两个、三个、四个、五个、六个或更多个核苷酸的片段)。

[0205] 在一个实施方案中,具有SEQ ID NO:141中所提供的3′UTR序列或其变体或片段的多核苷酸使得多核苷酸的半衰期增加,例如,多核苷酸的半衰期增加约1.5-10倍。在一个实

施方案中,半衰期增加约1.5、2、3、4、5、6、7、8、9或10倍或更多。在一个实施方案中,半衰期增加约1.5倍或更多。在一个实施方案中,半衰期增加约2倍或更多。在一个实施方案中,半衰期增加约3倍或更多。在一个实施方案中,半衰期增加约4倍或更多。在一个实施方案中,半衰期增加约5倍或更多。在一个实施方案中,半衰期增加约6倍或更多。在一个实施方案中,半衰期增加约7倍或更多。在一个实施方案中,半衰期增加约8倍。在一个实施方案中,半衰期增加约9倍或更多。在一个实施方案中,半衰期增加约10倍或更多。

[0206] 在一个实施方案中,具有SEQ ID NO:141中所提供的3′UTR序列或其变体或片段的多核苷酸使得由多核苷酸编码的多肽的水平和/或活性(例如,输出)增加。

[0207] 在一个实施方案中,将所述增加与不具有3′UTR、具有不同3′UTR或不具有SEQ ID NO:141的3′UTR或其变体或片段的其它方面类似的多核苷酸进行比较。

[0208] 在一个实施方案中,多核苷酸包含SEQ ID NO:141中所提供的3′UTR序列或与SEQ ID NO:141中所提供的3′UTR序列具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0209] 在一个实施方案中,多核苷酸包含终止元件和3'-UTR,其中序列是(终止元件为斜体):

[0210] *UAAAGCUCCCCGGGG*UCCAUAAAGUAGGAAACACUACAG
CUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCCAAA
CACCAUUGUCACACUCCAUCCCCCCAGCCCCUCCUCCCCUUCC
UGCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGG

GCGGC(SEQ ID NO:142)或其变体或片段(例如,缺少SEQ ID NO:142的核苷酸的初始一个、两个、三个、四个、五个、六个或更多个核苷酸的片段)。

[0211] 在一个实施方案中,具有SEQ ID NO:142中所提供的3′UTR序列或其变体或片段的多核苷酸使得多核苷酸的半衰期增加,例如,多核苷酸的半衰期增加约1.5-10倍。在一个实施方案中,半衰期增加约1.5、2、3、4、5、6、7、8、9或10倍或更多。在一个实施方案中,半衰期增加约1.5倍或更多。在一个实施方案中,半衰期增加约2倍或更多。在一个实施方案中,半衰期增加约3倍或更多。在一个实施方案中,半衰期增加约4倍或更多。在一个实施方案中,半衰期增加约5倍或更多。在一个实施方案中,半衰期增加约6倍或更多。在一个实施方案中,半衰期增加约7倍或更多。在一个实施方案中,半衰期增加约8倍。在一个实施方案中,半衰期增加约9倍或更多。在一个实施方案中,半衰期增加约10倍或更多。

[0212] 在一个实施方案中,具有SEQ ID NO:142中所提供的3′UTR序列或其变体或片段的多核苷酸使得由多核苷酸编码的多肽的水平和/或活性(例如,输出)增加。

[0213] 在一个实施方案中,将所述增加与不具有3′UTR、具有不同3′UTR或不具有SEQ ID NO:142的3′UTR或其变体或片段的其它方面类似的多核苷酸进行比较。

[0214] 在一个实施方案中,多核苷酸包含SEQ ID NO:142中所提供的3′UTR序列或与SEQ ID NO:142中所提供的3′UTR序列具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0215] 在一个实施方案中,多核苷酸包含终止元件和3'-UTR,其中序列是(终止元件为斜体):

*UAAAGCUCCCCGGGG*GCCUCGGUGGCCUAGCUUCUUGCC

[0216] CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGU

ACCCCCCAAACACCAUUGUCACACUCCAGUGGUCUUUGAAUA

AAGUCUGAGUGGGCGGC **(SEQ ID NO:143)**或其变体或片段(例如,缺少SEQ ID NO:143的核苷酸的初始一个、两个、三个、四个、五个、六个或更多个核苷酸的片段)。

[0217] 在一个实施方案中,具有SEQ ID NO:143中所提供的3′UTR序列或其变体或片段的多核苷酸使得多核苷酸的半衰期增加,例如,多核苷酸的半衰期增加约1.5-10倍。在一个实施方案中,半衰期增加约1.5、2、3、4、5、6、7、8、9或10倍或更多。在一个实施方案中,半衰期增加约1.5倍或更多。在一个实施方案中,半衰期增加约2倍或更多。在一个实施方案中,半衰期增加约3倍或更多。在一个实施方案中,半衰期增加约4倍或更多。在一个实施方案中,半衰期增加约5倍或更多。在一个实施方案中,半衰期增加约6倍或更多。在一个实施方案中,半衰期增加约7倍或更多。在一个实施方案中,半衰期增加约8倍。在一个实施方案中,半衰期增加约9倍或更多。在一个实施方案中,半衰期增加约10倍或更多。

[0218] 在一个实施方案中,具有SEQ ID NO:143中所提供的3′UTR序列或其变体或片段的多核苷酸使得由多核苷酸编码的多肽的水平和/或活性(例如,输出)增加。

[0219] 在一个实施方案中,将所述增加与不具有3′UTR、具有不同3′UTR或不具有SEQ ID NO:143的3′UTR或其变体或片段的其它方面类似的多核苷酸进行比较。

[0220] 在一个实施方案中,多核苷酸包含SEQ ID NO:143中所提供的3′UTR序列或与SEQ ID NO:143中所提供的3′UTR序列具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0221] 在一个实施方案中,多核苷酸包含终止元件和3′-UTR,其中序列是(终止元件为斜体):

*UAAGCCCCUCCGGGG*CAAACACCAUUGUCACACUCCAGC

[0222] CUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCCAAACACCAUU

GUCACACUCCAUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCC

GUACCCCCCAAACACCAUUGUCACACUCCAGUGGUCUUUGAA

UAAAGUCUGAGUGGGCGGC **(SEQ ID NO:144)**或其变体或片段(例如,缺少SEQ ID NO:144的核苷酸的初始一个、两个、三个、四个、五个、六个或更多个核苷酸的片段)。

[0223] 在一个实施方案中,具有SEQ ID NO:144中所提供的3′UTR序列或其变体或片段的多核苷酸使得多核苷酸的半衰期增加,例如,多核苷酸的半衰期增加约1.5-10倍。在一个实施方案中,半衰期增加约1.5、2、3、4、5、6、7、8、9或10倍或更多。在一个实施方案中,半衰期增加约1.5倍或更多。在一个实施方案中,半衰期增加约2倍或更多。在一个实施方案中,半衰期增加约3倍或更多。在一个实施方案中,半衰期增加约4倍或更多。在一个实施方案中,半衰期增加约5倍或更多。在一个实施方案中,半衰期增加约6倍或更多。在一个实施方案中,半衰期增加约7倍或更多。在一个实施方案中,半衰期增加约8倍。在一个实施方案中,半衰期增加约9倍或更多。在一个实施方案中,半衰期增加约10倍或更多。

[0224] 在一个实施方案中,具有SEQ ID NO:144中所提供的3′UTR序列或其变体或片段的多核苷酸使得由多核苷酸编码的多肽的水平和/或活性(例如,输出)增加。

[0225] 在一个实施方案中,将所述增加与不具有3′UTR、具有不同3′UTR或不具有SEQ ID NO:144的3′UTR或其变体或片段的其它方面类似的多核苷酸进行比较。

[0226] 在一个实施方案中,多核苷酸包含SEQ ID NO:144中所提供的3′UTR序列或与SEQ ID NO:144中所提供的3′UTR序列具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0227] 在一个实施方案中,多核苷酸包含终止元件和3'-UTR,其中序列是(终止元件为斜体):

[0228] *UAAGCCCCUCCGGGG*UCCAUAAAGUAGGAAACACUACAG CCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCAUAAAGU AGGAAACACUACAUCCCCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCCGCAUUAUUACUCACGGUACGAGUGGUCUUU GAAUAAAGUCUGAGUGGGCGGC (**SEQ ID NO:145**)或其变体或片段(例如,缺少SEQ ID NO:145的核苷酸的初始一个、两个、三个、四个、五个、六个或更多个核苷酸的片段)。

[0229] 在一个实施方案中,具有SEQ ID NO:145中所提供的3′UTR序列或其变体或片段的多核苷酸使得多核苷酸的半衰期增加,例如,多核苷酸的半衰期增加约1.5-10倍。在一个实施方案中,半衰期增加约1.5、2、3、4、5、6、7、8、9或10倍或更多。在一个实施方案中,半衰期增加约1.5倍或更多。在一个实施方案中,半衰期增加约2倍或更多。在一个实施方案中,半衰期增加约3倍或更多。在一个实施方案中,半衰期增加约4倍或更多。在一个实施方案中,半衰期增加约5倍或更多。在一个实施方案中,半衰期增加约6倍或更多。在一个实施方案中,半衰期增加约7倍或更多。在一个实施方案中,半衰期增加约8倍。在一个实施方案中,半衰期增加约9倍或更多。在一个实施方案中,半衰期增加约10倍或更多。

[0230] 在一个实施方案中,具有SEQ ID NO:145中所提供的3′UTR序列或其变体或片段的多核苷酸使得由多核苷酸编码的多肽的水平和/或活性(例如,输出)增加。

[0231] 在一个实施方案中,将所述增加与不具有3′UTR、具有不同3′UTR或不具有SEQ ID NO:145的3′UTR或其变体或片段的其它方面类似的多核苷酸进行比较。

[0232] 在一个实施方案中,多核苷酸包含SEQ ID NO:145中所提供的3′UTR序列或与SEQ ID NO:145中所提供的3′UTR序列具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0233] 在一个实施方案中,多核苷酸包含终止元件和3'-UTR,其中序列是(终止元件为斜体):

[0234] *UAAGCCCCUCCGGGG*UCCAUAAAGUAGGAAACACUACAG CUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCCAAA CACCAUUGUCACACUCCAUCCCCCCAGCCCCUCCUCCCCUUCC UGCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGG

GCGGC (**SEQ ID NO:146**)或其变体或片段(例如,缺少SEQ ID NO:146的核苷酸的初始一个、两个、三个、四个、五个、六个或更多个核苷酸的片段)。

[0235] 在一个实施方案中,具有SEQ ID NO:1 46中所提供的3′UTR序列或其变体或片段的多核苷酸使得多核苷酸的半衰期增加,例如,多核苷酸的半衰期增加约1.5-10倍。在一个实施方案中,半衰期增加约1.5、2、3、4、5、6、7、8、9或10倍或更多。在一个实施方案中,半衰期增加约1.5倍或更多。在一个实施方案中,半衰期增加约2倍或更多。在一个实施方案中,半衰期增加约3倍或更多。在一个实施方案中,半衰期增加约4倍或更多。在一个实施方案中,半衰期增加约5倍或更多。在一个实施方案中,半衰期增加约6倍或更多。在一个实施方案中,半衰期增加约7倍或更多。在一个实施方案中,半衰期增加约8倍。在一个实施方案中,半衰期增加约9倍或更多。在一个实施方案中,半衰期增加约10倍或更多。

[0236] 在一个实施方案中,具有SEQ ID NO:146中所提供的3′UTR序列或其变体或片段的多核苷酸使得由多核苷酸编码的多肽的水平和/或活性(例如,输出)增加。

[0237] 在一个实施方案中,将所述增加与不具有3′UTR、具有不同3′UTR或不具有SEQ ID NO:146的3′UTR或其变体或片段的其它方面类似的多核苷酸进行比较。

[0238] 在一个实施方案中,多核苷酸包含SEQ ID NO:146中所提供的3′UTR序列或与SEQ ID NO:146中所提供的3′UTR序列具有至少80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0239] 在一个实施方案中,多核苷酸包含终止元件和3′-UTR,其中序列是(终止元件为斜体):

*UAAGCCCCUCCGGGG*GCCUCGGUGGCCUAGCUUCUUGCC

[0240] CCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGU

ACCCCCCAAACACCAUUGUCACACUCCAGUGGUCUUUGAAUA

AAGUCUGAGUGGGCGGC (**SEQ ID NO:147**)或其变体或片段(例如,缺少SEQ ID NO:147的核苷酸的初始一个、两个、三个、四个、五个、六个或更多个核苷酸的片段)。

[0241] 在一个实施方案中,具有SEQ ID NO:147中所提供的3′UTR序列或其变体或片段的多核苷酸使得多核苷酸的半衰期增加,例如,多核苷酸的半衰期增加约1.5-10倍。在一个实施方案中,半衰期增加约1.5、2、3、4、5、6、7、8、9或10倍或更多。在一个实施方案中,半衰期增加约1.5倍或更多。在一个实施方案中,半衰期增加约2倍或更多。在一个实施方案中,半衰期增加约3倍或更多。在一个实施方案中,半衰期增加约4倍或更多。在一个实施方案中,半衰期增加约5倍或更多。在一个实施方案中,半衰期增加约6倍或更多。在一个实施方案中,半衰期增加约7倍或更多。在一个实施方案中,半衰期增加约8倍。在一个实施方案中,半衰期增加约9倍或更多。在一个实施方案中,半衰期增加约10倍或更多。

[0242] 在一个实施方案中,具有SEQ ID NO:147中所提供的3′UTR序列或其变体或片段的多核苷酸使得由多核苷酸编码的多肽的水平和/或活性(例如,输出)增加。

[0243] 在一个实施方案中,将所述增加与不具有3′UTR、具有不同3′UTR或不具有SEQ ID NO:147的3′UTR或其变体或片段的其它方面类似的多核苷酸进行比较。

[0244] 在一个实施方案中,多核苷酸包含SEQ ID NO:147中所提供的3′UTR序列或与SEQ ID NO:147中所提供的3′UTR序列具有至少80%、85%、90%、95%、96%、97%、98%、99%或

100％同一性的序列。

[0245] 2.3′稳定区

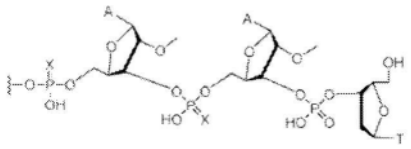
[0246] 本文尤其公开了一种编码多肽的多核苷酸,其中所述多核苷酸包含:(a)5′-UTR(例如,如本文所描述);(b)编码区,所述编码区包含终止元件(例如,如本文所描述);(c)3'-UTR(例如,如本文所描述);以及(d)3′稳定区。本文还公开了包含所述多核苷酸的LNP组合物。

[0247] 在一个实施方案中,多核苷酸包含3′稳定区,例如稳定的尾,例如,如本文所描述。含有3′-稳定区(例如,包括替代性核碱基、糖和/或主链的3′-稳定区)的多核苷酸可能特别有效地用于治疗性组合物中,这是因为它们可能受益于增加的稳定性、高表达水平。

[0248] 在一个实施方案中,3′稳定区包含多聚A尾,例如,包含80-150个,例如120个腺嘌呤的多聚A尾(SEQ ID NO:123)。在一个实施方案中,多聚A尾包含一个或多个非腺苷残基,例如一个或多个鸟苷,例如,如本文所描述。在一个实施方案中,多聚A尾包含UCUAG序列(SEQ ID NO:44)。在一个实施方案中,多聚A尾在SEQ ID NO:44的上游包含约80-120个,例如100个腺嘌呤。在一个实施方案中,多聚A尾在SEQ ID NO:44的下游包含约1-40个,例如20个腺嘌呤。

[0249] 在一个实施方案中,3′稳定区包含至少一个替代性核苷。在一个实施方案中,替代性核苷是反向胸苷(idT)。在一个实施方案中,替代性核苷位于3′稳定区的3′端。

[0250] 在一个实施方案中,3′稳定区包含式VII的结构:

[0251] 

[0252] 或其盐,其中每个X独立地是O或S,并且A代表腺嘌呤并且T代表胸腺嘧啶。

[0253] 在一个方面,本文公开了一种编码多肽的多核苷酸,其中所述多核苷酸包含:(a)5'-UTR,例如,如本文所描述;(b)编码区,所述编码区包含终止元件(例如,如本文所描述);(c)3'-UTR(例如,如本文所描述);以及(d)3′稳定区,例如,如本文所描述。

[0254] 在一个方面,包含本文所公开的包含稳定区的多核苷酸的LNP组合物包含:(i)可电离脂质,例如,氨基脂质;(ii)固醇或其它结构脂质;(iii)非阳离子辅助脂质或磷脂;以及(iv)PEG-脂质。

[0255] 在另一个方面,本公开的LNP组合物用于治疗疾病或病症的方法中,或用于抑制受试者的免疫反应的方法中。

[0256] 在一个方面,包含本文所公开的编码治疗性有效载荷或预防性有效载荷(例如,如本文所描述)的多核苷酸的LNP组合物可以与额外剂(例如,如本文所描述)一起施用。

[0257] 3.微RNA(miRNA)结合位点

[0258] 本公开的核酸分子(例如RNA,例如mRNA)可以包括调控元件,例如,微RNA(miRNA)结合位点、转录因子结合位点、结构化mRNA序列和/或基序、经工程改造以充当内源核酸结合分子的假受体的人工结合位点以及它们的组合。

[0259] 在一些实施方案中,本公开的核酸分子(例如RNA,例如mRNA)包含编码所关注的多肽的开放阅读框(ORF)并且还包含一个或多个miRNA结合位点。基于天然存在的miRNA的组织特异性和/或细胞类型特异性表达,包含或并入一个或多个miRNA结合位点提供了对本公

开的核酸分子(例如RNA,例如mRNA)的调控,进而提供了对由其编码的多肽的调控。

[0260] miRNA,例如天然存在的miRNA是19-25个核苷酸长的非编码RNA,其结合至核酸分子(例如RNA,例如mRNA)并且通过降低多核苷酸的稳定性或通过抑制其翻译来下调基因表达。miRNA序列包含“种子”区,即,成熟miRNA的位置2-8的区域中的序列。miRNA种子可以包含成熟miRNA的位置2-8或2-7。在一些实施方案中,miRNA种子可以包含7个核苷酸(例如,成熟miRNA的核苷酸2-8),其中相应miRNA结合位点中的种子互补位点侧接有与miRNA位置1相对的腺苷(A)。在一些实施方案中,miRNA种子可以包含6个核苷酸(例如,成熟miRNA的核苷酸2-7),其中相应miRNA结合位点中的种子互补位点侧接有与miRNA位置1相对的腺苷(A)。参见例如Grimson A,Farh KK,Johnston WK,Garrett-Engele P,Lim LP,Bartel DP;Mol Cell.2007年7月6日;27(1):91-105。可以对靶细胞或组织的miRNA进行谱分析,以确定细胞或组织中miRNA的存在或不存在。在一些实施方案中,本公开的核酸分子(例如RNA,例如mRNA)包含一个或多个微RNA结合位点、微RNA靶序列、微RNA互补序列或微RNA种子互补序列。此类序列可以对应于任何已知的微RNA,诸如美国公布US2005/0261218和美国公布US2005/0059005中教示的那些,例如与所述微RNA具有互补性,这些公布中的每一者的内容以全文引用的方式并入本文中。

[0261] 如本文所用,术语“微RNA(miRNA或miR)结合位点”是指核酸分子内(例如DNA内或RNA转录物内,包括5′UTR和/或3′UTR中)的序列,其与全部miRNA或其区域具有充足互补性以与miRNA相互作用、缔合或结合。在一些实施方案中,本公开的核酸分子(例如RNA,例如mRNA)包含编码所关注的多肽的ORF并且还包含一个或多个miRNA结合位点。在示例性实施方案中,核酸分子(例如RNA,例如mRNA)的5′UTR和/或3′UTR包含一个或多个miRNA结合位点。

[0262] 与miRNA具有充足互补性的miRNA结合位点是指互补程度足以促进miRNA介导的核酸分子(例如RNA,例如mRNA)调控,例如,miRNA介导的核酸分子(例如RNA,例如mRNA)的翻译抑制或降解。在本公开的示例性方面,与miRNA具有充足互补性的miRNA结合位点是指互补程度足以促进miRNA介导的核酸分子(例如RNA,例如mRNA)降解,例如,miRNA指导的RNA诱导的沉默复合物(RISC)介导的mRNA裂解。miRNA结合位点可以与例如19-25个核苷酸的miRNA序列、19-23个核苷酸的miRNA序列或22个核苷酸的miRNA序列具有互补性。miRNA结合位点可以仅与miRNA的一部分互补,例如,与天然存在的miRNA序列的全长的小于1、2、3或4个核苷酸的部分互补。当所需调控是mRNA降解时,充分或完全互补性(例如,在天然存在的miRNA的长度的全部或大部分上充分互补或完全互补)是优选的。

[0263] 在一些实施方案中,miRNA结合位点包括与miRNA种子序列具有互补性(例如,部分或完全互补性)的序列。在一些实施方案中,miRNA结合位点包括与miRNA种子序列具有完全互补性的序列。在一些实施方案中,miRNA结合位点包括与miRNA序列具有互补性(例如,部分或完全互补性)的序列。在一些实施方案中,miRNA结合位点包括与miRNA序列具有完全互补性的序列。在一些实施方案中,miRNA结合位点与miRNA序列具有完全互补性,但具有1、2或3个核苷酸取代、末端添加和/或截短。

[0264] 在一些实施方案中,miRNA结合位点的长度与相应miRNA相同。在其它实施方案中,在5′末端、3′末端或两个末端处,miRNA结合位点比相应miRNA短一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个或十二个核苷酸。在其它实施方案中,在5′末端、3′

末端或两个末端处,微RNA结合位点比相应微RNA短两个核苷酸。比相应miRNA短的miRNA结合位点仍然能够使并入一个或多个miRNA结合位点的mRNA降解或防止mRNA翻译。

[0265] 在一些实施方案中,miRNA结合位点结合相应成熟miRNA,所述成熟miRNA是含有Dicer的活性RISC的一部分。在另一个实施方案中,miRNA结合位点与RISC中的相应miRNA的结合使含有miRNA结合位点的mRNA降解或防止mRNA翻译。在一些实施方案中,miRNA结合位点与miRNA具有充足互补性,使得包含miRNA的RISC复合物使包含miRNA结合位点的核酸分子(例如RNA,例如mRNA)裂解。在其它实施方案中,miRNA结合位点具有不完美互补性,使得包含miRNA的RISC复合物诱导包含miRNA结合位点的核酸分子(例如RNA,例如mRNA)中的不稳定性。在另一个实施方案中,miRNA结合位点具有不完美互补性,使得包含miRNA的RISC复合物抑制包含miRNA结合位点的核酸分子(例如RNA,例如mRNA)的转录。

[0266] 在一些实施方案中,miRNA结合位点与相应miRNA具有一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个或十二个错配。在一些实施方案中,miRNA结合位点具有分别与相应miRNA的至少约十个、至少约十一个、至少约十二个、至少约十三个、至少约十四个、至少约十五个、至少约十六个、至少约十七个、至少约十八个、至少约十九个、至少约二十个或至少约二十一个连续核苷酸互补的至少约十个、至少约十一个、至少约十二个、至少约十三个、至少约十四个、至少约十五个、至少约十六个、至少约十七个、至少约十八个、至少约十九个、至少约二十个或至少约二十一个连续核苷酸。

[0267] 通过将一个或多个miRNA结合位点工程改造至本公开的核酸分子(例如RNA,例如mRNA)中,所述核酸分子(例如RNA,例如mRNA)可以被靶向用于降解或减少的翻译,前提条件是可获得所讨论的miRNA。这可以减少核酸分子(例如RNA,例如mRNA)递送时的脱靶效应。举例来说,如果本公开的核酸分子((例例如RNA,例如mRNA)不欲递送至组织或细胞,而是终止于所述组织或细胞,则如果将miRNA的一个或多个结合位点工程改造至所述核酸分子(例如RNA,例如mRNA)的5′UTR和/或3′UTR中,则所述组织或细胞中丰富的miRNA可以抑制所关注基因的表达。

[0268] 举例来说,本领域技术人员将了解一个或多个miR结合位点可以包括于核酸分子(例如RNA,例如mRNA)中以使除淋巴样细胞以外的细胞类型中的表达最小化。在一个实施方案中,可以使用miR122结合位点。在另一个实施方案中,可以使用miR126结合位点。在又一个实施方案中,可以使用这些miR结合位点或组合的多个拷贝。

[0269] 相反地,miRNA结合位点可从其中天然存在这些结合位点的核酸分子(例如RNA,例如mRNA)序列中去除,以增加特定组织中的蛋白质表达。举例来说,针对特定miRNA的结合位点可以从核酸分子(例如RNA,例如mRNA)中去除以改善含有miRNA的组织或细胞中的蛋白质表达。

[0270] 对多种组织中的表达的调控可以通过引入或去除一个或多个miRNA结合位点,例如一个或多个不同miRNA结合位点来实现。可以基于发育和/或疾病中的组织和/或细胞中的miRNA表达模式和/或其谱分析来决定是否去除或插入miRNA结合位点。已经报告了对miRNA、miRNA结合位点及其表达模式和在生物学中的作用的鉴定(例如,Bonauer等人,Curr Drug Targets 2010 11:943-949;Anand和Cheresh Curr Opin Hematol 2011 18:171-176;Contreras和Rao Leuk emia 2012 26:404-413(2011年12月20日.doi:10.1038/leu.2011.356);Bartel Cell 2009 136:215-233;Landgraf等人,Cell,2007 129:1401-

1414;Gentner和Naldini,Tissue Antigens.2012 80:393-403以及其中所有参考文献;所述文献中的每一者以全文引用的方式并入本文中)。

[0271] miRNA和miRNA结合位点可对应于任何已知序列,包括描述于美国公布第2014/0200261号、第2005/0261218号和第2005/0059005号中的非限制性实例,所述公布中的每一者通过引用整体并入本文。

[0272] 其中已知miRNA调控mRNA并且由此调控蛋白质表达的组织的实例包括但不限于肝(miR-122)、肌肉(miR-133、miR-206、miR-208)、内皮细胞(miR-17-92、miR-126)、骨髓细胞(miR-142-3p、miR-142-5p、miR-16、miR-21、miR-223、miR-24、miR-27)、脂肪组织(let-7、miR-30c)、心脏(miR-1d、miR-149)、肾(miR-192、miR-194、miR-204)和肺上皮细胞(let-7、miR-133、miR-126)。

[0273] 具体来说,已知miRNA在免疫细胞(也称为造血细胞),诸如抗原呈递细胞(APC)(例如,树突状细胞和单核细胞)、单核细胞、单核细胞、B淋巴细胞、T淋巴细胞、粒细胞、自然杀伤细胞等中差异性地表达。免疫细胞特异性miRNA涉及于免疫原性、自身免疫性、对感染的免疫反应、炎症以及在基因疗法和组织/器官移植之后的非所需免疫反应中。免疫细胞特异性miRNA还调控造血细胞(免疫细胞)的发育、增殖、分化和凋亡的许多方面。举例来说,miR-142和miR-146仅在免疫细胞中表达,尤其在骨髓树突状细胞中很丰富。已经证明了对核酸分子(例如RNA,例如mRNA)的免疫反应可以通过向多核苷酸的3′-UTR中添加miR-142结合位点来关闭,使得能够在组织和细胞中进行更稳定的基因转移。miR-142有效地使抗原呈递细胞中的外源核酸分子(例如RNA,例如mRNA)降解并且抑制转导的细胞的细胞毒性消除(例如,Annoni A等人,blood,2009,114,5152-5161;Brown BD等人,Nat med.2006,12(5),585-591;Brown BD等人,blood,2007,110(13):4144-4152,所述文献中的每一者以全文引用的方式并入本文中)。

[0274] 抗原介导的免疫反应可以指由外来抗原触发的免疫反应,这些外来抗原在进入生物体时由抗原呈递细胞处理并且呈现于抗原呈递细胞的表面上。T细胞可以识别呈递的抗原并且诱导表达所述抗原的细胞的细胞毒性消除。

[0275] 将miR-142结合位点引入本公开的核酸分子的5′UTR和/或3′UTR中可以通过miR-142介导的降解选择性地抑制抗原呈递细胞中的基因表达,从而限制抗原呈递细胞(例如,树突状细胞)中的抗原呈递并且由此防止在递送核酸分子(例如RNA,例如mRNA)之后出现抗原介导的免疫反应。核酸分子(例如RNA,例如mRNA)然后在靶组织或细胞中稳定表达而不会触发细胞毒性消除。

[0276] 在一个实施方案中,可以将已知在免疫细胞、尤其抗原呈递细胞中表达的miRNA的结合位点工程改造至本公开的核酸分子(例如RNA,例如mRNA)中,以通过miRNA介导的RNA降解来抑制抗原呈递细胞中核酸分子(例如RNA,例如mRNA)的表达,从而抑制抗原介导的免疫反应。核酸分子(例如RNA,例如mRNA)在其中未表达免疫细胞特异性miRNA的非免疫细胞中维持表达。举例来说,在一些实施方案中,为了防止针对肝特异性蛋白的免疫原性反应,可以去除任何miR-122结合位点并且可以将miR-142(和/或mirR-146)结合位点工程改造至本公开的核酸分子的5'UTR和/或3'UTR中。

[0277] 为了进一步驱动APC和巨噬细胞中的选择性降解和抑制,本公开的核酸分子(例如RNA,例如mRNA)可以在5'UTR和/或3'UTR中包括另一负调控元件,单独或与miR-142和/或

miR-146结合位点组合。作为一个非限制性实例,另一负调控元件是组成性衰减元件(CDE)。

[0278] 免疫细胞特异性miRNA包括但不限于hsa-let-7a-2-3p、hsa-let-7a-3p、hsa-7a-5p、hsa-let-7c、hsa-let-7e-3p、hsa-let-7e-5p、hsa-1et-7g-3p、hsa-let-7g-5p、hsa-let-7i-3p、hsa-let-7i-5p、miR-10a-3p、miR-10a-5p、miR-1184、hsa-let-7f-1--3p、hsa-let-7f-2--5p、hsa-let-7f-5p、miR-125b-1-3p、miR-125b-2-3p、miR-125b-5p、miR-1279、miR-130a-3p、miR-130a-5p、miR-132-3p、miR-132-5p、miR-142-3p、miR-142-5p、miR-143-3p、miR-143-5p、miR-146a-3p、miR-146a-5p、miR-146b-3p、miR-146b-5p、miR-147a、miR-147b、miR-148a-5p、miR-148a-3p、miR-150-3p、miR-150-5p、miR-151b、miR-155-3p、miR-155-5p、miR-15a-3p、miR-15a-5p、miR-15b-5p、miR-15b-3p、miR-16-1-3p、miR-16-2-3p、miR-16-5p、miR-17-5p、miR-181a-3p、miR-181a-5p、miR-181a-2-3p、miR-182-3p、miR-182-5p、miR-197-3p、miR-197-5p、miR-21-5p、miR-21-3p、miR-214-3p、miR-214-5p、miR-223-3p、miR-223-5p、miR-221-3p、miR-221-5p、miR-23b-3p、miR-23b-5p、miR-24-1-5p、miR-24-2-5p、miR-24-3p、miR-26a-1-3p、miR-26a-2-3p、miR-26a-5p、miR-26b-3p、miR-26b-5p、miR-27a-3p、miR-27a-5p、miR-27b-3p、miR-27b-5p、miR-28-3p、miR-28-5p、miR-2909、miR-29a-3p、miR-29a-5p、miR-29b-1-5p、miR-29b-2-5p、miR-29c-3p、miR-29c-5p、miR-30e-3p、miR-30e-5p、miR-331-5p、miR-339-3p、miR-339-5p、miR-345-3p、miR-345-5p、miR-346、miR-34a-3p、miR-34a-5p、miR-363-3p、miR-363-5p、miR-372、miR-377-3p、miR-377-5p、miR-493-3p、miR-493-5p、miR-542、miR-548b-5p、miR548c-5p、miR-548i、miR-548j、miR-548n、miR-574-3p、miR-598、miR-718、miR-935、miR-99a-3p、miR-99a-5p、miR-99b-3p和miR-99b-5p。此外,可以通过微阵列杂交和切片机分析在免疫细胞中鉴定新颖miRNA(例如,Jima DD等人,Blood,2010,116:e118-e127;Vaz C等人,BMC Genomics,2010,11,288,所述文献中的每一者的内容以全文引用的方式并入本文中)。

[0279] 在一些实施方案中,将miRNA结合位点在核酸分子(例如RNA,例如mRNA)的任何位置(例如,5′UTR和/或3′UTR)处插入本公开的核酸分子(例如RNA,例如mRNA)中。在一些实施方案中,5′UTR包含miRNA结合位点。在一些实施方案中,3′UTR包含miRNA结合位点。在一些实施方案中,5′UTR和3′UTR包含miRNA结合位点。核酸分子(例如RNA,例如mRNA)中的插入位点可以在核酸分子(例如RNA,例如mRNA)中的任何位置,只要在不存在相应miRNA的情况下,在核酸分子(例如RNA,例如mRNA)中插入miRNA结合位点不干扰功能性多肽的翻译;并且在存在miRNA的情况下,在核酸分子(例如RNA,例如mRNA)中插入miRNA结合位点以及miRNA结合位点与相应miRNA的结合能够使多核苷酸降解或防止核酸分子(例如RNA,例如mRNA)翻译。

[0280] 在一些实施方案中,将miRNA结合位点插入本公开的包含ORF的核酸分子(例如RNA,例如mRNA)中ORF的终止密码子下游的至少约30个核苷酸中。在一些实施方案中,将miRNA结合位点插入本公开的多核苷酸中ORF的终止密码子下游的至少约10个核苷酸、至少约15个核苷酸、至少约20个核苷酸、至少约25个核苷酸、至少约30个核苷酸、至少约35个核苷酸、至少约40个核苷酸、至少约45个核苷酸、至少约50个核苷酸、至少约55个核苷酸、至少约60个核苷酸、至少约65个核苷酸、至少约70个核苷酸、至少约75个核苷酸、至少约80个核苷酸、至少约85个核苷酸、至少约90个核苷酸、至少约95个核苷酸或至少约100个核苷酸中。在一些实施方案中,将miRNA结合位点插入本公开的核酸分子(例如RNA,例如mRNA)中ORF的

终止密码子下游的约10个核苷酸至约100个核苷酸、约20个核苷酸至约90个核苷酸、约30个核苷酸至约80个核苷酸、约40个核苷酸至约70个核苷酸、约50个核苷酸至约60个核苷酸、约45个核苷酸至约65个核苷酸中。

[0281] miRNA基因调控可能受miRNA周围的序列影响,诸如但不限于周围序列的种类、序列的类型(例如异源、同源、外源、内源或人工)、周围序列中的调控元件和/或周围序列中的结构元件。miRNA可能受5′UTR和/或3′UTR影响。作为一个非限制性实例,与相同序列类型的人类3′UTR相比,非人类3′UTR可以增加miRNA序列对所关注多肽的表达的调控作用。

[0282] 在一个实施方案中,5′UTR的其它调控元件和/或结构元件可能影响miRNA介导的基因调控。调控元件和/或结构元件的一个实例是5′UTR中的结构化IRES(内部核糖体进入位点),它是结合翻译延伸因子以启动蛋白质翻译所必需的。EIF4A2与5′-UTR中的这种二级结构化元件的结合是miRNA介导的基因表达所必需的(Meijer HA等人,Science,2013,340,82-85,以全文引用的方式并入本文中)。本公开的核酸分子(例如RNA,例如mRNA)还可以包括这种结构化5′UTR以增强微RNA介导的基因调控。

[0283] 可以将至少一个miRNA结合位点工程改造至本公开的多核苷酸的3′UTR中。在这种情形中,可以将至少两个、至少三个、至少四个、至少五个、至少六个、至少七个、至少八个、至少九个、至少十个或更多个miRNA结合位点工程改造至本公开的核酸分子(例如RNA,例如mRNA)的3′UTR中。举例来说,可以将1至10个、1至9个、1至8个、1至7个、1至6个、1至5个、1至4个、1至3个、2个或1个miRNA结合位点工程改造至本公开的核酸分子(例如RNA,例如mRNA)的3′UTR中。在一个实施方案中,并入本公开的核酸分子(例如RNA,例如mRNA)中的miRNA结合位点可以是相同的或可以是不同的miRNA位点。并入本公开的核酸分子(例如RNA,例如mRNA)中的不同miRNA结合位点的组合可以包括其中并入不同miRNA位点中的任一者的多于一个拷贝的组合。在另一个实施方案中,并入本公开的核酸分子(例如RNA,例如mRNA)中的miRNA结合位点可以靶向体内相同或不同的组织。作为一个非限制性实例,通过在本公开的核酸分子(例如RNA,例如mRNA)的3’-UTR中引入组织、细胞类型或疾病特异性miRNA结合位点,可以降低特定细胞类型(例如,肝细胞、骨髓细胞、内皮细胞、癌细胞等)中的表达程度。

[0284] 在一个实施方案中,可以将miRNA结合位点工程改造至本公开的核酸分子(例如RNA,例如mRNA)中的3′UTR的5′末端附近、3′UTR的5′末端与3′末端之间约一半处和/或3′UTR的3′末端附近。作为一个非限制性实例,可以将miRNA结合位点工程改造至3′UTR的5′末端附近和3′UTR的5′末端与3′末端之间约一半处。作为另一个非限制性实例,可以将miRNA结合位点工程改造至3′UTR的3′末端附近和3′UTR的5′末端与3′末端之间约一半处。作为又一个非限制性实例,可以miRNA结合位点工程改造至3′UTR的5′末端附近和3′UTR的3′末端附近。

[0285] 在另一个实施方案中,3′UTR可以包含1、2、3、4、5、6、7、8、9或10个miRNA结合位点。miRNA结合位点可以与miRNA、miRNA种子序列和/或侧接种子序列的miRNA序列互补。

[0286] 可以对本公开的核酸分子(例如RNA,例如mRNA)进行工程改造以基于不同组织、细胞类型或生物条件中miRNA的表达模式在特定组织、细胞类型或生物条件中进行更具靶向性的表达。通过引入组织特异性miRNA结合位点,可以设计本公开的核酸分子(例如RNA,例如mRNA)以在组织或细胞中或在生物条件的情形中进行最佳蛋白质表达。

[0287] 在一些实施方案中,本公开的核酸分子(例如RNA,例如mRNA)可以在3′UTR中包含

至少一个miRNA结合位点,以选择性地使免疫细胞中的mRNA治疗剂降解以抑制由治疗性递送引起的非所需免疫原性反应。作为一个非限制性实例,miRNA结合位点可以使本公开的核酸分子(例如RNA,例如mRNA)在抗原呈递细胞中更不稳定。这些miRNA的非限制性实例示于下表3中。

[0288] 表3:miRNA结合位点序列

[0289]

序列名称	SEQ ID NO:	序列
miR122 bs	148	CAAACACCAUUGUCACACUCCA
miR-142-3p bs	149	UCCAUAAAGUAGGAAACACUACA
miR-126-3p bs	150	CGCAUUAUUACUCACGGUACGA
miR-142	151	GACAGUGCAGUCACCCAUAAAGUAGAAA GCACUACUAACAGCACUGGAGGGUGUAG UGUUUCCUACUUUAUGGAUGAGUGUACU

[0290]

		GUG
miR-142-3p	152	UGUAGUGUUUCCUACUUUAUGGA
miR-142-5p	153	CAUAAAGUAGAAAGCACUACU
miR-142-5p 结合位点	154	AGUAGUGCUUUCUACUUUAUG
miR-126	155	CGCUGGCGACGGGACAUUAUUACUUUUG GUACGCGCUGUGACACUUCAAACUCGUA CCGUGAGUAAUAAUGCGCCGUCCACGGC A
miR-126-5p	156	CAUUAUUACUUUUGGUACGCG
miR-126-5p 结合位点	157	CGCGUACCAAAAGUAAUAAUG

[0291] 在一些实施方案中,本文所描述的核酸分子的3'UTR包含miR122bs(即,如上表3中所示的SEQ ID NO:148)。在一些实施方案中,本文所描述的核酸分子的3'UTR包含miR-142-3p bs(即,如上表3中所示的SEQ ID NO:149)。在一些实施方案中,本文所描述的核酸分子的3'UTR包含miR-126-3p bs(即,如上表3中所示的SEQ ID NO:150)。

[0292] 在一些实施方案中,本文所描述的核酸分子的3′UTR包含多于一个miRNA结合位点。在一些实施方案中,本文所描述的核酸分子的3'UTR包含1、2、3、4、5、6、7、8、9或10个miRNA结合位点。在一些实施方案中,在存在多于一个miRNA结合位点的情况下,miRNA结合位点是相同的。在存在多于一个miRNA结合位点的一些情况下,miRNA结合位点(例如,上表3中所列出的任何miRNA结合位点的任何组合)。在一些实施方案中,在存在多于一个miRNA结合位点的情况下,约1-25个核苷酸可以存在于每个miRNA结合位点之间。举例来说,1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个核苷酸可以存在于每个miRNA结合位点之间。

[0293] 在一些实施方案中,本文所描述的核酸分子的3'UTR包含miR142-3p bs和miR-126-3p bs两者。在一些实施方案中,本文所描述的核酸分子的3'UTR包含miR-142-3p bs的

三个拷贝。在一些实施方案中,本文所描述的核酸分子的3′UTR包含miR-142-3p bs的两个拷贝。在一些实施方案中,本文所描述的核酸分子的3′UTR包含miR-142-3p bs的两个拷贝和miR-126-3p bs的一个拷贝。在一些实施方案中,本文所描述的核酸分子的3′UTR包含miR122bs的三个拷贝。

[0294] 4.核苷酸帽

[0295] 本公开还包括一种包含5′帽和本发明的多核苷酸(例如,包含编码待表达的多肽的核苷酸序列的多核苷酸)两者的多核苷酸。

[0296] 天然mRNA的5′帽结构涉及于核输出中,增加mRNA稳定性并结合mRNA帽结合蛋白(CBP),后者通过CBP与多聚(A)结合蛋白的缔合而负责细胞中的mRNA稳定性和翻译能力以形成成熟环状mRNA种类。所述帽进一步有助于在mRNA剪接期间去除5′近端内含子。

[0297] 内源mRNA分子可以在5′端加帽,从而在mRNA分子的末端鸟苷帽残基与5′-末端转录的有义核苷酸之间产生5’-ppp-5’-三磷酸键。然后可以将这种5′-鸟苷酸帽甲基化以产生N7-甲基-鸟苷酸残基。mRNA的5′端的末端和/或末端前转录核苷酸的核糖也可以任选地被2′-O-甲基化。通过鸟苷酸帽结构的水解和裂解使5′-脱帽可以靶向核酸分子,诸如mRNA分子,以进行降解。

[0298] 在一些实施方案中,本发明的多核苷酸(例如,包含编码多肽的核苷酸序列的多核苷酸)并入帽部分。

[0299] 在一些实施方案中,本发明的多核苷酸包含防止脱帽并因此增加mRNA半衰期的不可水解帽结构。因为帽结构水解需要5’-ppp-5′磷酸二酯键裂解,所以可以在加帽反应期间使用修饰的核苷酸。举例来说,来自New England Biolabs(Ipswich,MA)的牛痘加帽酶可以根据制造商的说明书与αα-硫-鸟苷核苷酸一起使用,以在5’-ppp-5’帽中产生硫代磷酸酯键。可以使用额外修饰的鸟苷核苷酸,诸如$_{\alpha}$α-甲基-膦酸和硒代-磷酸核苷酸。

[0300] 额外的修饰包括但不限于多核苷酸(如上文所提及)的5′-末端和/或5′-末端前核苷酸的核糖在糖环的2′-羟基上的2’-O-甲基化。多个不同的5′-帽结构可以用于产生核酸分子的5′-帽,诸如用作mRNA分子的多核苷酸。帽类似物,在本文中也称为合成帽类似物、化学帽、化学帽类似物或结构或功能帽类似物,其化学结构不同于天然(即,内源、野生型或生理)5′-帽,同时保留帽功能。帽类似物可以是化学(即,非酶促)或酶促合成的和/或连接至本发明的多核苷酸。

[0301] 举例来说,抗反向帽类似物(ARCA)帽含有两个由5′-5′-三磷酸基团连接的鸟嘌呤,其中一个鸟嘌呤含有N7甲基以及3′-O-甲基(即,N7,3’-O-二甲基-鸟苷-5′-三磷酸-5′-鸟苷(m^7G-3’mppp-G;可以等效地命名为3′O-Me-m^7G(5′)ppp(5′)G)。另一个未修饰的鸟嘌呤的3′-O原子变成连接至加帽多核苷酸的5′-末端核苷酸。N7-和3′-O-甲基化鸟嘌呤提供了加帽多核苷酸的末端部分。

[0302] 另一个示例性帽是mCAP,它类似于ARCA,但在鸟苷上具有2′-O-甲基(即,N7,2′-O-二甲基-鸟苷-5′-三磷酸-5′-鸟苷,m^7Gm-ppp-G)。

[0303] 另一个示例性帽是$m^7GpppG_{2',OMc}$或m^7G-ppp-Gm-A(即,N7,鸟苷-5′-三磷酸-2’-O--二甲基-鸟苷-腺苷)。

[0304] 在一些实施方案中,帽是二核苷酸帽类似物。作为一个非限制性实例,二核苷酸帽类似物可以在不同的磷酸位置用硼代磷酸基团或硒代磷酸基团进行修饰,诸如美国专利号

US 8,519,110中所描述的二核苷酸帽类似物,所述专利的内容以全文引用的方式并入本文中。

[0305] 在另一个实施方案中,帽是帽类似物,是本领域中已知和/或本文所描述的帽类似物的N7-(4-氯苯氧基乙基)取代的二核苷酸形式。帽类似物的N7-(4-氯苯氧基乙基)取代的二核苷酸形式的非限制性实例包括N7-(4-氯苯氧基乙基)-G(5′)ppp(5′)G和N7-(4-氯苯氧基乙基)-$m^{3'-^{0}}$G(5′)ppp(5′)G帽类似物(参见例如Kore等人,Bioorganic&Medicinal Chemistry 2013 21:4570-4574中所描述的各种帽类似物和合成帽类似物的方法;所述文献的内容以全文引用的方式并入本文中)。在另一个实施方案中,本发明的帽类似物是4-氯/溴苯氧基乙基类似物。

[0306] 本发明的多核苷酸也可以在制造后(无论是IVT还是化学合成)使用酶加帽,以产生更真实的5′-帽结构。如本文所用,短语“更真实”是指在结构上或功能上密切反映或模仿内源或野生型特征的特征。也就是说,“更真实”的特征与现有技术的合成特征或类似物等相比更能代表内源、野生型、天然或生理细胞功能和/或结构,或在一个或多个方面优于相应的内源、野生型、天然或生理特征。本发明的更真实的5′帽结构的非限制性实例是那些帽结构,它们与本领域中已知的合成5′帽结构(或野生型、天然或生理性5′帽结构)相比具有增强的帽结合蛋白结合、增加的半衰期、降低的对5′核酸内切酶的敏感性和/或减少的5′脱帽等。举例来说,重组牛痘病毒加帽酶和重组2’-O-甲基转移酶可以在多核苷酸的5′-末端核苷酸与鸟嘌呤帽核苷酸之间产生规范的5′-5′-三磷酸键,其中帽鸟嘌呤含有N7甲基化并且mRNA的5′-末端核苷酸含有2′-O-甲基。此种结构称为帽1结构。与例如本领域中已知的其它5′帽类似物结构相比,这种帽使得翻译能力和细胞稳定性更高以及细胞促炎细胞因子的活化减少。帽结构包括但不限于7mG(5′)ppp(5′)N1pN2p(帽0)、7mG(5′)ppp(5′)N1mpNp(帽1)和7mG(5′)-ppp(5′)N1mpN2mp(帽2)。

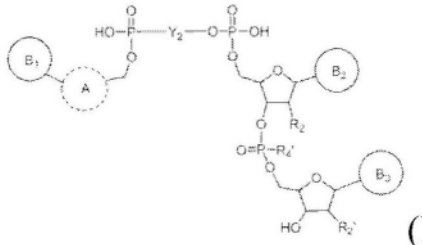
[0307] 作为一个非限制性实例,制造后对嵌合多核苷酸加帽可能更有效,因为几乎100%的嵌合多核苷酸可以被加帽。这与当在体外转录反应过程中帽类似物连接至嵌合多核苷酸时约80%成对比。

[0308] 根据本发明,5′末端帽可以包括内源帽或帽类似物。根据本发明,5′末端帽可以包含鸟嘌呤类似物。可用的鸟嘌呤类似物包括但不限于肌苷、N1-甲基-鸟苷、2′氟-鸟苷、7-脱氮-鸟苷、8-氧代-鸟苷、2-氨基-鸟苷、LNA-鸟苷和2-叠氮基-鸟苷。

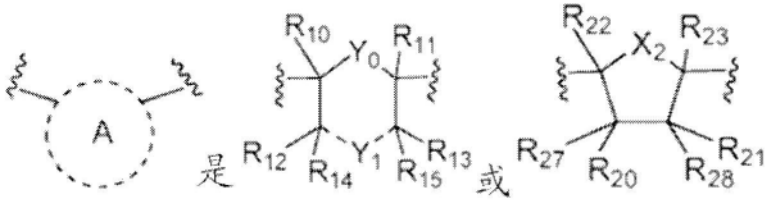
[0309] 本文还提供了示例性帽,所述帽包括可以用于核糖核酸(RNA)合成的共转录加帽方法中的那些帽,所述合成使用RNA聚合酶,例如野生型RNA聚合酶或其变体,诸如本文所描述的那些变体。在一个实施方案中,当在“一锅”反应中产生RNA时可以添加帽,而不需要单独的加帽反应。因此,在一些实施方案中,所述方法包括使多核苷酸模板与RNA聚合酶变体、三磷酸核苷和帽类似物在体外转录反应条件下反应以产生RNA转录物。

[0310] 如本文所用,术语“帽”包括反向G核苷酸并且可以包含反向G核苷酸3′的一个或多个额外核苷酸,例如,反向G核苷酸3′和5′UTR(例如,本文所描述的5′UTR)5′的1、2、3个或更多个核苷酸。

[0311] 示例性帽包含GG、GA或GGA的序列,其中加下划线的斜体G是反向G核苷酸,其后是5’-5’-三磷酸基团。

[0312] 在一个实施方案中,帽包含式(I)化合物  (I),

或其立体异构体、互变异构体或盐,其中

[0313] ;

[0314] 环B_1是修饰或未修饰的鸟嘌呤;

[0315] 环B_2和环B_3各自独立地是核碱基或修饰的核碱基;

[0316] X_2是O、$S(O)_p$、NR_{24}或$CR_{25}R_{26}$,其中p是0、1或2;

[0317] Y_0是O或CR_6R_7;

[0318] Y1是O、$S(O)_n$、CR_6R_7或NR_8,其中n是0、1或2;

[0319] 每个---是单键或不存在,其中当每个---是单键时,Yi是O、$S(O)_n$、CR_6R_7或NR_8;并且当每个---不存在时,Y_1是缺失的;

[0320] Y_2是$(OP(O)R_4)_m$,其中m是0、1或2,或-O-$(CR_{40}R_{41})u-Q_0-(CR_{42}R_{43})v-$,其中$Q_0$是键、O、$S(O)_r$、$NR_{44}$或$CR_{45}R_{46}$,r是0、1或2,并且u和v中的每一者独立地是1、2、3或4;

[0321] 每个R_2和R_2'独立地是卤基、LNA或OR_3;

[0322] 每个R_3独立地是H、C_1-C_6烷基、C_2-C_6烯基或C_2-C_6炔基,并且当R_3是C_1-C_6烷基、C_2-C_6烯基或C_2-C_6炔基时任选地被以下一者或多者取代:卤基、OH和任选地被一个或多个OH或OC(O)-C_1-C_6烷基取代的C_1-C_6烷氧基;

[0323] 每个R_4和R_4′独立地是H、卤基、C_1-C_6烷基、OH、SH、SeH或BH_3^-;

[0324] R_6、R_7和R_8中的每一者独立地是-Q_1-T_1,其中Q_1是键或任选地被卤基、氰基、OH和C_1-C_6烷氧基中的一者或多者取代的C_1-C_3烷基连接子,并且T_1是H、卤基、OH、COOH、氰基或R_{s1},其中R_{s1}是C_1-C_3烷基、C_2-C_6烯基、C_2-C_6炔基、C_1-C_6烷氧基、C(O)O-C_1-C_6烷基、C_3-C_8环烷基、C_6-C_{10}芳基、$NR_{31}R_{32}$、$(NR_{31}R_{32}R_{33})^+$、4元至12元杂环烷基或5元或6元杂芳基,并且R_{s1}任选地被一个或多个选自由以下组成的组的取代基取代:卤基、OH、氧代、C_1-C_6烷基、COOH、C(O)O-C_1-C_6烷基、氰基、C_1-C_6烷氧基、$NR_{31}R_{32}$、$(NR_{31}R_{32}R_{33})^+$、C_3-C_8环烷基、C_6-C_{10}芳基、4元至12元杂环烷基和5元或6元杂芳基;

[0325] R_{10}、R_{11}、R_{12}、R_{13}、R_{14}和R_{15}中的每一者独立地是-Q_2-T_2,其中Q_2是键或视情况被卤基、氰基、OH和C_1-C_6烷氧基中的一者或多者取代的C_1-C_3烷基连接子,并且T_2是H、卤基、OH、NH_2、氰基、NO_2、N_3、R_{s2}或OR_{s2},其中R_{s2}是C_1-C_6烷基、C_2-C_6烯基、C_2-C_6炔基、C_3-C_8环烷基、C_6-C_{10}芳基、NHC(O)-C_1-C_6烷基、$NR_{31}R_{32}$、$(NR_{31}R_{32}R_{33})^+$、4元至12元杂环烷基或5元或6元杂芳基,并且R_{s2}任选地被一个或多个选自由以下组成的组的取代基取代:卤基、OH、氧代、C_1-C_6烷基、COOH、C(O)O-C_1-C_6烷基、氰基、C_1-C_6烷氧基、$NR_{31}R_{32}$、$(NR_{31}R_{32}R_{33})^+$、C_3-C_8环烷基、C_6-C_{10}芳基、4

元至12元杂环烷基和5元或6元杂芳基;或替代地,R_{12}与R_{14}一起是氧代,或R_{13}与R_{15}一起是氧代,

[0326] R_{20}、R_{21}、R_{22}和R_{23}中的每一者独立地是-Q_3-T_3,其中Q_3是键或任选地被卤基、氰基、OH和C_1-C_6烷氧基中的一者或多者取代的C_1-C_3烷基连接子,并且T_3是H、卤基、OH、NH_2、氰基、NO_2、N_3、R_{S3}或OR_{S3},其中R_{S3}是C_1-C_6烷基、C_2-C_6烯基、C_2-C_6炔基、C_3-C_8环烷基、C_6-C_{10}芳基、NHC(O)-C_1-C_6烷基、单C_1-C_6烷基氨基、二C_1-C_6烷基氨基、4元至12元杂环烷基或5元或6元杂芳基,并且Rs_3任选地被一个或多个选自由以下组成的组的取代基取代:卤基、OH、氧代、C_1-C_6烷基、COOH、C(O)O-C_1-C_6烷基、氰基、C_1-C_6烷氧基、氨基、单C_1-C_6烷基氨基、二C_1-C_6烷基氨基、C_3-C_8环烷基、C_6-C_{10}芳基、4元至12元杂环烷基和5元或6元杂芳基;

[0327] R_{24}、R_{25}和R_{26}中的每一者独立地是H或C_1-C_6烷基;

[0328] R_{27}和R_{28}中的每一者独立地是H或OR_{29};或R_{27}与R_{28}一起形成OR_{30}-O;每个R_{29}独立地是H、C_1-C_6烷基、C_2-C_6烯基或C_2-C_6炔基,并且当R_{29}是C_1-C_6烷基、C_2-C_6烯基或C_2-C_6炔基时任选地被以下一者或多者取代:卤基、OH和任选地被一个或多个OH或OC(O)-C_1-C_6烷基取代的C_1-C_6烷氧基;

[0329] R_{30}是任选地被卤基、OH和C_1-C_6烷氧基中的一者或多者取代的C_1-C_6亚烷基;

[0330] R_{31}、R_{32}和R_{33}中的每一者独立地是H、C_1-C_6烷基、C_3-C_8环烷基、C_6-C_{10}芳基、4元至12元杂环烷基或5元或6元杂芳基;

[0331] R_{40}、R_{41}、R_{42}和R_{43}中的每一者独立地是H、卤基、OH、氰基、N_3、$OP(O)R_{47}R_{48}$或任选地被一个或多个$OP(O)R_{47}R_{48}$取代的C_1-C_6烷基,或一个R_{41}和一个R_{43}与它们所连接的碳原子和Q_0一起形成C_4-C_{10}环烷基、4元至14元杂环烷基、C_6-C_{10}芳基或5元至14元杂芳基,并且所述环烷基、杂环烷基、苯基或5元至6元杂芳基中的每一者任选地被以下一者或多者取代:OH、卤基、氰基、N_3、氧代、$OP(O)R_{47}R_{48}$、C_1-C_6烷基、C_1-C_6卤烷基、COOH、C(O)O-C_1-C_6烷基、C_1-C_6烷氧基、C_1-C_6卤烷氧基、氨基、单C_1-C_6烷基氨基和二C_1-C_6烷基氨基;

[0332] R_{44}是H、C_1-C_6烷基或胺保护基;

[0333] R_{45}和R_{46}中的每一者独立地是H、$OP(O)R_{47}R_{48}$或任选地被一个或多个$OP(O)R_{47}R_{48}$取代的C_1-C_6烷基,并且

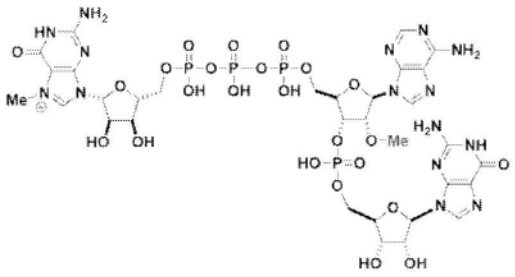
[0334] R_{47}和R_{48}中的每一者独立地是H、卤基、C_1-C_6烷基、OH、SH、SeH或BH_3。

[0335] 应了解,如本文所提供的帽类似物可以包括2017年4月20日公布的国际公布WO 2017/066797中所描述的帽类似物中的任一者,所述公布以全文引用的方式并入本文中。

[0336] 在一些实施方案中,B_2中间位置可以是非核糖分子,诸如阿拉伯糖。

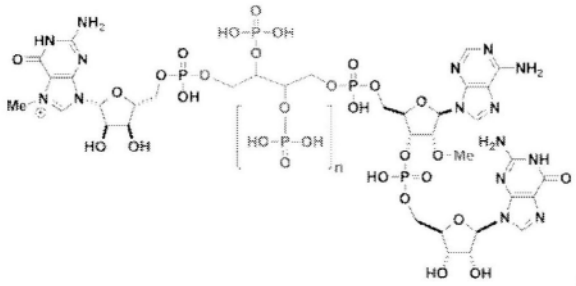
[0337] 在一些实施方案中,R_2是基于乙基的。

[0338] 因此,在一些实施方案中,帽包括以下结构:

[0339] 

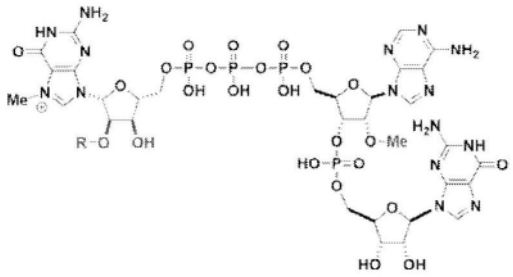
(II)

[0340] 在其它实施方案中,帽包括以下结构:

[0341] 

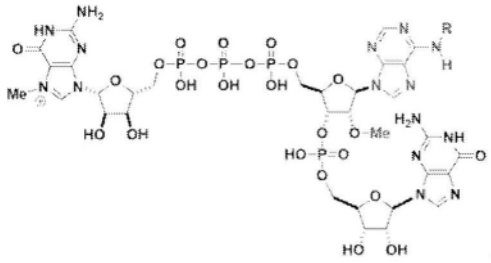
(III)

[0342] 在其它实施方案中,帽包括以下结构:

[0343] 

(IV)

[0344] 在其它实施方案中,帽包括以下结构:

[0345] 

(V)

[0346] 在一些实施方案中,R是烷基(例如,C_1-C_6烷基)。在一些实施方案中,R是甲基(例如,C_1烷基)。在一些实施方案中,R是乙基(例如,C_2烷基)。

[0347] 在一些实施方案中,帽包含选自以下序列的序列:GAA、GAC、GAG、GAU、GCA、GCC、GCG、GCU、GGA、GGC、GGG、GGU、GUA、GUC、GUG和GUU。在一些实施方案中,帽包含GAA。在一些实施方案中,帽包含GAC。在一些实施方案中,帽包含GAG。在一些实施方案中,帽包含GAU。在一些实施方案中,帽包含GCA。在一些实施方案中,帽包含GCC。在一些实施方案中,帽包含GCG。在一些实施方案中,帽包含GCU。在一些实施方案中,帽包含GGA。在一些实施方案中,帽包含GGC。在一些实施方案中,帽包含GGG。在一些实施方案中,帽包含GGU。在一些实施方案中,帽包含GUA。在一些实施方案中,帽包含GUC。在一些实施方案中,帽包含GUG。在一些实施方案中,帽包含GUU。

[0348] 在一些实施方案中,帽包含选自以下序列的序列:m^7GpppG、m^7GpppApA、m^7GpppApC、m^7GpppApG、m^7GpppApU、m^7GpppCpA、m^7GpppCpC、m^7GpppCpG、m^7GpppCpU、m^7GpppGpA、m^7GpppGpC、m^7GpppGpG、m^7GpppGpU、m^7GpppUpA、m^7GpppUpC、m^7GpppUpG和m^7GpppUpU。

[0349] 在一些实施方案中,帽包含m^7GpppApA。在一些实施方案中,帽包含m^7GpppApC。在一些实施方案中,帽包含m^7GpppApG。在一些实施方案中,帽包含m^7GpppApU。在一些实施方案中,帽包含m^7GpppCpA。在一些实施方案中,帽包含m^7GpppCpC。在一些实施方案中,帽包含

$m^7GpppCpG$。在一些实施方案中,帽包含$m^7GpppCpU$。在一些实施方案中,帽包含$m^7GpppGpA$。在一些实施方案中,帽包含$m^7GpppGpC$。在一些实施方案中,帽包含$m^7GpppGpG$。在一些实施方案中,帽包含$m^7GpppGpU$。在一些实施方案中,帽包含$m^7GpppUpA$。在一些实施方案中,帽包含$m^7GpppUpC$。在一些实施方案中,帽包含$m^7GpppUpG$。在一些实施方案中,帽包含$m^7GpppUpU$。

[0350] 在一些实施方案中,帽包含选自以下序列的序列:$m^7G_{3'OMc}pppApA$、$m^7G_{3'OMc}pppApC$、$m^7G_{3'OMc}pppApG$、$m^7G_{3'OMc}pppApU$、$m^7G_{3'OMe}pppCpA$、$m^7G_{3'OMe}pppCpC$、$m^7G_{3'OMe}pppCpG$、$m^7G_{3'OMe}pppCpU$、$m^7G_{3'OMe}pppGpA$、$m^7G_{3'OMe}pppGpC$、$m^7G_{3'OMe}pppGpG$、$m^7G_{3'OMe}pppGpU$、$m^7G_{3'OMc}pppUpA$、$m^7G_{3'OMc}pppUpC$、$m^7G_{3'OMc}pppUpG$和$m^7G_{3'OMc}pppUpU$。

[0351] 在一些实施方案中,帽包含$m^7G_{3'OMc}pppApA$。在一些实施方案中,帽包含$m^7G_{3'OMc}pppApC$。在一些实施方案中,帽包含$m^7G_{3'OMc}pppApG$。在一些实施方案中,帽包含$m^7G_{3'OMc}pppApU$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppCpA$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppCpC$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppCpG$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppCpU$。在一些实施方案中,帽包含$m^7G_{3'OMc}pppGpA$。在一些实施方案中,帽包含$m^7G_{3'OMc}pppGpC$。在一些实施方案中,帽包含$m^7G_{3'OMc}pppGpG$。在一些实施方案中,帽包含$m^7G_{3'OMc}pppGpU$。在一些实施方案中,帽包含$m^7G_{3'OMc}pppUpA$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppUpC$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppUpG$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppUpU$。

[0352] 在其它实施方案中,帽包含选自以下序列的序列:$m^7G_{3'OMe}pppA_{2'OMe}pA$、$m^7G_{3'OMe}pppA_{2'OMe}pC$、$m^7G_{3'OMe}pppA_{2'OMe}pG$、$m^7G_{3'OMe}pppA_{2'OMe}pU$、$m^7G_{3'OMe}pppC_{2'OMe}pA$、$m^7G_{3'OMe}pppC_{2'OMe}pC$、$m^7G_{3'OMe}pppC_{2'OMc}pG$、$m^7G_{3'OMc}pppC_{2'OMc}pU$、$m^7G_{3'OMc}pppG_{2'OMc}pA$、$m^7G_{3'OMc}pppG_{2'OMc}pC$、$m^7G_{3'OMc}pppG_{2'OMc}pG$、$m^7G_{3'OMc}pppG_{2'OMc}pU$、$m^7G_{3'OMc}pppU_{2'OMc}pA$、$m^7G_{3'OMc}pppU_{2'OMc}pC$、$m^7G_{3'OMc}pppU_{2'OMc}pG$和$m^7G_{3'OMc}pppU_{2'OMe}pU$。

[0353] 在一些实施方案中,帽包含$m^7G_{3'OMc}pppA_{2'OMc}pA$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppA_{2'OMe}pC$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppA_{2'OMe}pG$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppA_{2'OMe}pU$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppC_{2'OMe}pA$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppC_{2'OMc}pC$。在一些实施方案中,帽包含$m^7G_{3'OMc}pppC_{2'OMc}pG$。在一些实施方案中,帽包含$m^7G_{3'OMc}pppC_{2'OMc}pU$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppG_{2'OMe}pA$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppG_{2'OMe}pC$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppG_{2'OMc}pG$。在一些实施方案中,帽包含$m^7G_{3'OMc}pppG_{2'OMc}pU$。在一些实施方案中,帽包含$m^7G_{3'OMc}pppU_{2'OMc}pA$。在一些实施方案中,帽包含$m^TG_{3'OMe}pppU_{2'OMe}pC$。在一些实施方案中,帽包含$m^TG_{3'OMc}pppU_{2'OMc}pG$。在一些实施方案中,帽包含$m^7G_{3'OMe}pppU_{2'OMe}pU$。

[0354] 在其它实施方案中,帽包含选自以下序列的序列:$rn^7GpppG_{2'OMe}$、$m^7GpppA_{2'OMe}pA$、$m^7GpppA_{2'OMe}pC$、$m^7GpppA_{2'OMe}pG$、$m^7GpppA_{2'OMe}pU$、$m^7GpppC_{2'OMe}pA$、$m^7GpppC_{2'OMe}pC$、$m^7GpppC_{2'OMe}pG$、$m^7GpppC_{2'OMe}pU$、$m^7GpppG_{2'OMep}A$、$m^7GpppG_{2'OMe}pC$、$m^7GpppG_{2'OMe}pG$、$m^TGpppG_{2'OMo}pU$、$m^TGpppU_{2'OMo}pA$、$m^TGpppU_{2'OMo}pC$、$m^7GpppU_{2'OMc}pG$和$m^7GpppU_{2'OMc}pU$。

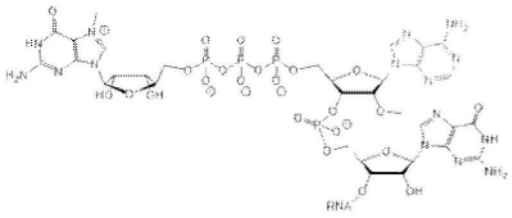
[0355] 在一些实施方案中,帽包含$m^7GpppA_{2'OMc}pA$。在一些实施方案中,帽包含$m^TGpppA_{2,OMc}pC$。在一些实施方案中,帽包含$m^7GpppA_{2,OMc}pG$。在一些实施方案中,帽包含$m^7GpppA_{2,OMc}pU$。在一些实施方案中,帽包含$m^TGpppC_{2,OMe}pA$。在一些实施方案中,帽包含$m^TGpppC_{2,OMe}pC$。在一些实施方案中,帽包含$m^7GpppC_{2,OMe}pG$。在一些实施方案中,帽包含$m^7GpppC_{2'OMe}pU$。在一些实

施方案中,帽包含$m^{T}GpppG_{2,OMo}pA$。在一些实施方案中,帽包含$m^{T}GpppG_{2'OMc}pC$。在一些实施方案中,帽包含$m^{T}GpppG_{2,OMc}pG$。在一些实施方案中,帽包含$m^{7}GpppG_{2'OMop}U$。在一些实施方案中,帽包含$m^{7}GpppU_{2'OMc}pA$。在一些实施方案中,帽包含$m^{7}GpppU_{2,OM}$。pC。在一些实施方案中,帽包含$m^{T}GpppU_{2,OMe}pG$。在一些实施方案中,帽包含$m^{T}GpppU_{2,OMe}pU$。

[0356] 在一些实施方案中,帽包含$m^{T}Gpppm^{6}A_{2,ome}pG$。在一些实施方案中,帽包含$m^{T}Gpppe^{6}A_{2,ome}pG$。

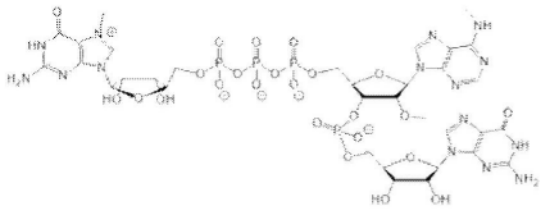
[0357] 在一些实施方案中,帽包含GAG。在一些实施方案中,帽包含GCG。在一些实施方案中,帽包含GUG。在一些实施方案中,帽包含GGG。

[0358] 在一些实施方案中,帽包括以下结构中的任一者:

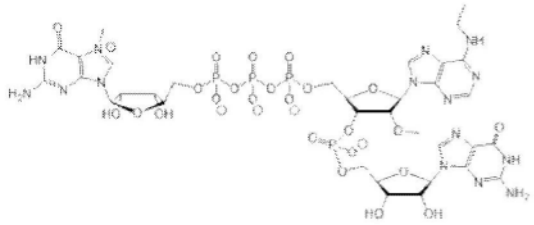


(VI);

[0359]



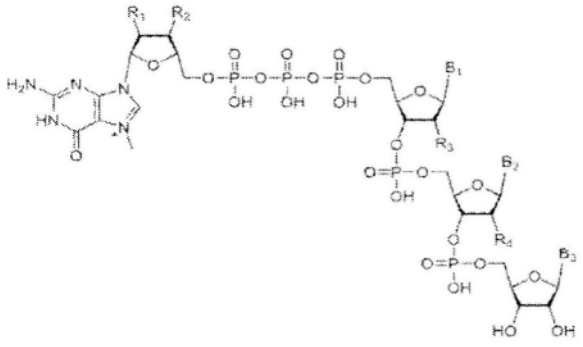
(VII); 或



(VIII)。

[0360] 在一些实施方案中,帽包含$^{m7}GpppN_1N_2N_3$,其中N_1、N_2和N_3是任选的(即,可以不存在或一者或多者可以存在)并且独立地是天然的、修饰的或非天然的核苷碱基。在一些实施方案中,^{m7}G被进一步甲基化,例如,在3′位置处。在一些实施方案中,^{m7}G在3′位置处包含O-甲基。在一些实施方案中,N_1、N_2和N_3,如果存在的话,任选地独立地是腺嘌呤、尿嘧啶、胍、胸腺嘧啶或胞嘧啶。在一些实施方案中,N_1、N_2和N_3中的一者或多者(或全部),如果存在的话,例如在2′位置处被甲基化。在一些实施方案中,N_1、N_2和N_3中的一者或多者(或全部),如果存在的话,在2′位置处具有O-甲基。

[0361] 在一些实施方案中,帽包括以下结构:

[0362] 

(IX)

[0363] 其中B_1、B_2和B_3独立地是天然的、修饰的或非天然的核苷碱基;并且R_1、R_2、R_3和R_4独立地是OH或O-甲基。在一些实施方案中,R_3是O-甲基并且R_4是OH。在一些实施方案中,R_3和R_4是O-甲基。在一些实施方案中,R_4是O-甲基。在一些实施方案中,R_1是OH,R_2是OH,R_3是O-甲基,并且R_4是OH。在一些实施方案中,R_1是OH,R_2是OH,R_3是O-甲基,并且R_4是O-甲基。在一些实施方案中,R_1和R_2中的至少一者是O-甲基,R_3是O-甲基,并且R_4是OH。在一些实施方案中,R_1和R_2中的至少一者是O-甲基,R_3是O-甲基,并且R_4是O-甲基。

[0364] 在一些实施方案中,B_1、B_3和B_3是天然核苷碱基。在一些实施方案中,B_1、B_2和B_3中的至少一者是修饰的或非天然的碱基。在一些实施方案中,B_1、B_2和B_3中的至少一者是N6-甲基腺嘌呤。在一些实施方案中,B_1是腺嘌呤、胞嘧啶、胸腺嘧啶或尿嘧啶。在一些实施方案中,B_1是腺嘌呤,B_2是尿嘧啶,并且B_3是腺嘌呤。在一些实施方案中,R_1和R_2是OH,R_3和R_4是O-甲基,B_1是腺嘌呤,B_2是尿嘧啶,并且B_3是腺嘌呤。

[0365] 在一些实施方案中,帽包含选自以下序列的序列:GAAA、GACA、GAGA、GAUA、GCAA、GCCA、GCGA、GCUA、GGAA、GGCA、GGGA、GGUA、GUCA和GUUA。在一些实施方案中,帽包含选自以下序列的序列:GAAG、GACG、GAGG、GAUG、GCAG、GCCG、GCGG、GCUG、GGAG、GGCG、GGGG、GGUG、GUCG、GUGG和GUUG。在一些实施方案中,帽包含选自以下序列的序列:GAAU、GACU、GAGU、GAUU、GCAU、GCCU、GCGU、GCUU、GGAU、GGCU、GGGU、GGUU、GUAU、GUCU、GUGU和GUUU。在一些实施方案中,帽包含选自以下序列的序列:GAAC、GACC、GAGC、GAUC、GCAC、GCCC、GCGC、GCUC、GGAC、GGCC、GGGC、GGUC、GUAC、GUCC、GUGC和GUUC。

[0366] 在一些实施方案中,帽包含选自以下序列的序列:$m^7G_{3'OMe}pppApApN$、$m^7G_{3'OMe}pppApCpN$、$m^7G_{3'OMe}pppApGpN$、$m^7G_{3'OMe}pppApUpN$、$m^7G_{3'OMe}pppCpApN$、$m^7G_{3'OMe}pppCpCpN$、$m^7G_{3'OMe}pppCpGpN$、$m^7G_{3'OMe}pppCpUpN$、$m^7G_{3'OMe}pppGpApN$、$m^7G_{3'OMe}pppGpCpN$、$m^7G_{3'OMe}pppGpGpN$、$m^7G_{3'OMe}pppGpUpN$、$m^7G_{3'OMe}pppUpApN$、$m^7G_{3'OMe}pppUpCpN$、$m^7G_{3'OMe}pppUpGpN$和$m^7G_{3'OMe}pppUpUpN$,其中N是天然的、修饰的或非天然的核苷碱基。

[0367] 在其它实施方案中,帽包含选自以下序列的序列:$m^7G_{3'OMc}pppA_{2'OMe}pApN$、$m^7G_{3'OMc}pppA_{2'OMe}pCpN$、$m^7G_{3'OMe}pppA_{2'OMe}pGpN$、$m^7G_{3'OMe}pppA_{2'OMe}pUpN$、$m^7G_{3'OMe}pppC_{2'OMe}pApN$、$m^7G_{3'OMe}pppC_{2'OMe}pCpN$、$m^7G_{3'OMe}pppC_{2'OMe}pGpN$、$m^7G_{3'OMe}pppC_{2'OMe}pUpN$、$m^7G_{3'OMe}pppG_{2'OMe}pApN$、$m^7G_{3'OMe}pppG_{2'OMe}pCpN$、$m^7G_{3'OMe}pppG_{2'OMe}pGpN$、$m^7G_{3'OMe}pppG_{2'OMe}pUpN$、$m^7G_{3'OMe}pppU_{2'OMe}pApN$、$m^7G_{3'OMe}pppU_{2'OMe}pCpN$、$m^7G_{3'OMc}pppU_{2'OMc}pGpN$和$m^7G_{3'OMe}pppU_{2'OMc}pUpN$,其中N是天然的、修饰的或非天然的核苷碱基。

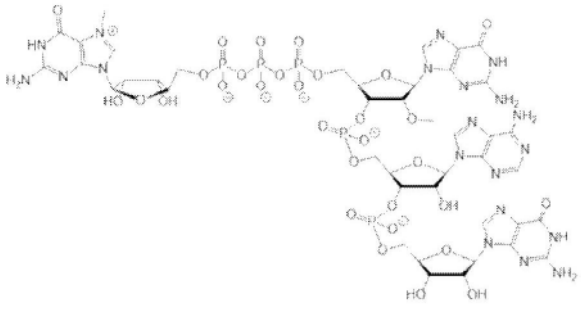
[0368] 在其它实施方案中,帽包含选自以下序列的序列:$m^7GpppA_{2'OMe}pApN$、$m^7GpppA_{2'OMe}pCpN$、$m^7GpppA_{2'OMe}pGpN$、$m^7GpppA_{2'OMe}pUpN$、$m^7GpppC_{2'OMe}pApN$、$m^7GpppC_{2'OMe}pCpN$、$m^7GpppC_{2'}$

$_{OMe}$pGpN、m^7GpppC$_{2'OMe}$pUpN、m^7GpppG$_{2'OMe}$pApN、m^7GpppG$_{2'OMe}$pCpN、m^7GpppG$_{2'OMe}$pGpN、m^7GpppG$_{2'OMe}$pUpN、m^7GpppU$_{2'OMe}$pApN、m^7GpppU$_{2'OMe}$pCpN、m^7GpppU$_{2'OMe}$pGpN和m^7GpppU$_{2'OMe}$pUpN，其中N是天然的、修饰的或非天然的核苷碱基。

[0369] 在其它实施方案中，帽包含选自以下序列的序列：m^7G$_{3'OMe}$pppA$_{2'OMe}$pA$_{2'OMe}$pN、m^7G$_{3'OMe}$pppA$_{2'OMe}$pC$_{2'OMe}$pN、m^7G$_{3'OMe}$pppA$_{2'OMe}$pG$_{2'OMc}$pN、m^7G$_{3'OMe}$pppA$_{2'OMc}$pU$_{2'OMc}$pN、m^7G$_{3'OMc}$pppC$_{2'OMc}$pA$_{2'OMc}$pN、m^7G$_{3'OMc}$pppC$_{2'OMc}$pC$_{2'OMc}$pN、m^7G$_{3'OMc}$pppC$_{2'OMc}$pG$_{2'OMc}$pN、m^7G$_{3'OMc}$pppC$_{2'OMc}$pU$_{2'OMc}$pN、m^7G$_{3'OMc}$pppG$_{2'OMc}$pA$_{2'OMc}$pN、m^7G$_{3'OMc}$pppG$_{2'OMc}$pC$_{2'OMe}$pN、m^7G$_{3'OMe}$pppG$_{2'OMe}$pG$_{2'OMe}$pN、m^7G$_{3'OMe}$pppG$_{2'OMe}$pU$_{2'OMe}$pN、m^7G$_{3'OMe}$pppU$_{2'OMe}$pA$_{2'OMe}$pN、m^7G$_{3'OMe}$pppU$_{2'OMe}$pC$_{2'OMe}$pN、m^7G$_{3'OMe}$pppU$_{2'OMe}$pG$_{2'OMe}$pN和m^7G$_{3'OMe}$pppU$_{2'OMe}$pU$_{2'OMe}$pN，其中N是天然的、修饰的或非天然的核苷碱基。

[0370] 在其它实施方案中，帽包含选自以下序列的序列：m^7GpppA$_{2'OMe}$pA$_{2'OMc}$pN、m^7GpppA$_{2'OMc}$pC$_{2'OMc}$pN、m^7GpppA$_{2'OMc}$pG$_{2'OMc}$pN、m^7GpppA$_{2'OMc}$pU$_{2'OMc}$pN、m^7GpppC$_{2'OMc}$pA$_{2'OMc}$pN、m^7GpppC$_{2'OMc}$pC$_{2'OMc}$pN、m^7GpppC$_{2'OMc}$pG$_{2'OMc}$pN、m^7GpppC$_{2'OMc}$pU$_{2'OMc}$pN、m^7GpppG$_{2'OMc}$pA$_{2'OMe}$pN、m^7GpppG$_{2'OMe}$pC$_{2'OMe}$pN、m^7GpppG$_{2'OMe}$pG$_{2'OMe}$pN、m^7GpppG$_{2'OMe}$pU$_{2'OMe}$pN、m^7GpppU$_{2'OMe}$pA$_{2'OMe}$pN、m^7GpppU$_{2'OMe}$pC$_{2'OMe}$pN、m^7GpppU$_{2'OMe}$pG$_{2'OMe}$pN和m^7GpppU$_{2'OMe}$pU$_{2'OMe}$pN，其中N是天然的、修饰的或非天然的核苷碱基。

[0371] 在一些实施方案中，帽包含GGAG。在一些实施方案中，帽包括以下结构：

[0372]  (X)。

[0373] 5.终止元件

[0374] 翻译终止密码子UAA、UAG和UGA是遗传密码的重要组分并且是mRNA翻译终止的信号。在蛋白质合成期间，终止密码子与蛋白质释放因子相互作用并且这种相互作用可以调节核糖体活性，由此对翻译具有影响(Tate WP等人，(2018)Biochem Soc Trans，46(6)：1615-162)。

[0375] 本文尤其公开了一种编码多肽的多核苷酸，所述多核苷酸具有包含终止元件的编码区，所述终止元件向由所述多核苷酸编码的多肽或多核苷酸本身赋予增加的半衰期、增加的表达和/或增加的活性。在一个实施方案中，多核苷酸包含：(a)5′-UTR(例如，如本文所描述)；(b)编码区，所述编码区包含终止元件(例如，如本文所描述)；以及(c)3′-UTR(例如，如本文所描述)，以及包含所述多核苷酸的LNP组合物。在一个实施方案中，多核苷酸包含编码区，所述编码区包含表4中所提供的终止元件。

[0376] 如本文所用的终止元件是指包含终止密码子的核酸序列。在DNA的情况下，终止密码子可以选自TGA、TAA和TAG；或在RNA的情况下，选自UGA、UAA和UAG。在一个实施方案中，终止元件包含两个连续的终止密码子。在一个实施方案中，终止元件包含三个连续的终止密码子。在一个实施方案中，终止元件包含四个连续的终止密码子。在一个实施方案中，终止

元件包含五个连续的终止密码子。

[0377] 在一个实施方案中,终止元件包含多个相同的终止密码子。在一个实施方案中,终止元件包含多个不同的终止密码子。

[0378] 在一个实施方案中,终止元件在一个或多个终止密码子的上游和/或下游还包含至少1、2、3、4、5或10个核苷酸。在一个实施方案中,终止元件在一个或多个终止密码子的上游还包含至少1、2、3、4、5或10个核苷酸。在一个实施方案中,终止元件在一个或多个终止密码子的下游还包含至少1、2、3、4、5或10个核苷酸。

[0379] 本发明还包括一种包含终止密码子元件和本文所描述的多核苷酸两者的多核苷酸。在一些实施方案中,终止密码子元件包含终止密码子区域。在一些实施方案中,多核苷酸的编码区包含终止元件。在一些实施方案中,终止元件处于多核苷酸中的3′UTR序列的上游,例如,在其之前。

[0380] 在一些实施方案中,本发明的多核苷酸可以在3′非翻译区(UTR)之前包括至少两个终止密码子。在DNA的情况下,终止密码子可以选自TGA、TAA和TAG;或在RNA的情况下,选自UGA、UAA和UAG。在一些实施方案中,本发明的多核苷酸在DNA的情况下包括终止密码子TGA,或在RNA的情况下包括终止密码子UGA,以及一个额外的终止密码子。在又一个实施方案中,额外的终止密码子可以是TAA或UAA。在另一个实施方案中,本发明的多核苷酸包括三个连续的终止密码子、四个终止密码子或更多。

[0381] 已经观察到包含表4中所提供的序列的终止元件可以使得多核苷酸的半衰期增加和/或由多核苷酸编码的多肽的水平或活性增加。

[0382] 在一个实施方案中,具有表4中所提供的终止元件的多核苷酸使得多核苷酸的半衰期增加或由多核苷酸编码的多肽的水平和/或活性(例如,输出)增加。在一个实施方案中,半衰期增加约1.5-20倍。在一个实施方案中,半衰期增加约1.5、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20倍或更多。在一个实施方案中,半衰期增加约1.5倍或更多。在一个实施方案中,半衰期增加约2倍或更多。在一个实施方案中,半衰期增加约3倍或更多。在一个实施方案中,半衰期增加约4倍。在一个实施方案中,半衰期增加约5倍或更多。

[0383] 在一个实施方案中,具有表4中所提供的终止元件的多核苷酸使得由多核苷酸编码的多肽的水平和/或活性(例如,输出或表达持续时间)增加。在一个实施方案中,终止元件使得由多核苷酸编码的多肽的水平和/或活性,例如可检测水平或活性增加约1.5-20倍,持续约1、2、3、4、5、6、7、8、9、10或14天。在一个实施方案中,终止元件使得由多核苷酸编码的多肽的可检测水平或活性持续约1、2、3、4、5、6、7、8、9、10或14天。

[0384] 在一个实施方案中,活性增加约1.5、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20倍或更多。在一个实施方案中,活性增加约1.5倍或更多。在一个实施方案中,活性增加约2倍或更多。在一个实施方案中,活性增加约3倍或更多。在一个实施方案中,活性增加约4倍或更多。在一个实施方案中,活性增加约5倍或更多。

[0385] 在一个实施方案中,将所述增加与不具有终止元件、具有不同终止元件或不具有表4中所提供的终止元件的其它方面类似的多核苷酸进行比较。

[0386] 在一个实施方案中,终止元件包含表4中所提供的序列。在一个实施方案中,终止元件包含SEQ ID NO:158、SEQ ID NO:159、SEQ ID NO:160、SEQ ID NO:161、SEQ ID NO:162、SEQ ID NO:163、SEQ ID NO:164、SEQ ID NO:165、SEQ ID NO:166、SEQ ID NO:167或

SEQ ID NO:168、SEQ ID NO;169、SEQ ID NO:173或SEQ ID NO:174的序列。在一个实施方案中,终止元件包含SEQ ID NO:158的序列。在一个实施方案中,终止元件包含SEQ ID NO:159的序列。在一个实施方案中,终止元件包含SEQ ID NO:160的序列。在一个实施方案中,终止元件包含SEQ ID NO:161的序列。在一个实施方案中,终止元件包含SEQ ID NO:162的序列。在一个实施方案中,终止元件包含SEQ ID NO:163的序列。在一个实施方案中,终止元件包含SEQ ID NO:164的序列。在一个实施方案中,终止元件包含SEQ ID NO:165的序列。在一个实施方案中,终止元件包含SEQ ID NO:166的序列。在一个实施方案中,终止元件包含SEQ ID NO:167的序列。在一个实施方案中,终止元件包含SEQ ID NO:168的序列。在一个实施方案中,终止元件包含SEQ ID NO:169的序列。在一个实施方案中,终止元件包含SEQ ID NO:173的序列。在一个实施方案中,终止元件包含SEQ ID NO:174的序列。

[0387] 在一些实施方案中,多核苷酸包括κ终止盒(即,UAAAGCUCCCCGGGG(SEQ ID NO:165))或ι终止盒(即,UAAGCCCCUCCGGGG(SEQ ID NO:164))。

[0388] 在一个实施方案中,(b)的编码区包含终止元件,所述终止元件包含式B的共有序列:

[0389] X_{-3}-X_{-2}-X_{-1}-U-A-A-X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}-X_{11}-X_{12}(SEQ ID NO:170)

[0390] 其中:

[0391] X_1是G或A;

[0392] X_2、X_4、X_5、X_6或X_7各自独立地是C或U;

[0393] X_3是C或A;

[0394] X_8、X_{10}、X_{11}、X_{12}、X_{-1}或X_{-3}各自独立地是C或G;

[0395] X_9是G或U;和/或

[0396] X_{-2}是A或U。

[0397] 在一个实施方案中,X_1是G。在一个实施方案中,X_1是A。

[0398] 在一个实施方案中,X_2是C。在一个实施方案中,X_2是U。

[0399] 在一个实施方案中,X_4是C。在一个实施方案中,X_4是U。

[0400] 在一个实施方案中,X_5是C。在一个实施方案中,X_5是U。

[0401] 在一个实施方案中,X_6是C。在一个实施方案中,X_6是U。

[0402] 在一个实施方案中,X_7是C。在一个实施方案中,X_7是U。

[0403] 在一个实施方案中,X_3是C。在一个实施方案中,X_3是A。

[0404] 在一个实施方案中,X_8是C。在一个实施方案中,X_8是G。

[0405] 在一个实施方案中,X_{10}是C。在一个实施方案中,X_{10}是G。

[0406] 在一个实施方案中,X_{11}是C。在一个实施方案中,X_{11}是G。

[0407] 在一个实施方案中,X_{12}是C。在一个实施方案中,X_{12}是G。

[0408] 在一个实施方案中,X_{-1}是C。在一个实施方案中,X_{-1}是G。

[0409] 在一个实施方案中,X_{-3}是C。在一个实施方案中,X_{-3}是G。

[0410] 在一个实施方案中,X_9是G。在一个实施方案中,X_9是U。

[0411] 在一个实施方案中,X_{-2}是A。在一个实施方案中,X_{-2}是U。

[0412] 在一个实施方案中,式B(SEQ ID NO:170)的共有序列具有高GC含量,例如,约50%、60%、70%、80%、90%或99%的GC含量。在一个实施方案中,GC含量为约50%。在一个

实施方案中,GC含量为约60%。在一个实施方案中,GC含量为约70%。在一个实施方案中,GC含量为约80%。在一个实施方案中,GC含量为约90%。在一个实施方案中,GC含量为约99%。

[0413] 在一个实施方案中,(b)的编码区包含终止元件,所述终止元件包含式C的共有序列:

[0414] X_{-3}-X_{-2}-X_{-1}-U-G-A-X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}-X_{11}-X_{12}(SEQ ID NO:171)

[0415] 其中:

[0416] X_{-3}、X_{-1}、X_2、X_5、X_6、X_7、X_8、X_9或X_{12}各自独立地是G或C;

[0417] X_{-2}、X_3或X_4各自独立地是A或C;

[0418] X_1是A或G;和/或

[0419] X_{10}或X_{11}各自独立地是C或U。

[0420] 在一个实施方案中,X_{-3}是G。在一个实施方案中,X_{-3}是C。

[0421] 在一个实施方案中,X_{-1}是G。在一个实施方案中,X_{-1}是C。

[0422] 在一个实施方案中,X_2是G。在一个实施方案中,X_2是C。

[0423] 在一个实施方案中,X_5是G。在一个实施方案中,X_5是C。

[0424] 在一个实施方案中,X_6是G。在一个实施方案中,X_6是C。

[0425] 在一个实施方案中,X_7是G。在一个实施方案中,X_7是C。

[0426] 在一个实施方案中,X_8是G。在一个实施方案中,X_8是C。

[0427] 在一个实施方案中,X_9是G。在一个实施方案中,X_9是C。

[0428] 在一个实施方案中,X_{12}是G。在一个实施方案中,X_{12}是C。

[0429] 在一个实施方案中,X_{-2}是A。在一个实施方案中,X_{-2}是C。

[0430] 在一个实施方案中,X_3是A。在一个实施方案中,X_3是C。

[0431] 在一个实施方案中,X_4是A。在一个实施方案中,X_4是C。

[0432] 在一个实施方案中,X_1是A。在一个实施方案中,X_1是G。

[0433] 在一个实施方案中,X_{10}是C。在一个实施方案中,X_{10}是U。

[0434] 在一个实施方案中,X_{11}是C。在一个实施方案中,X_{11}是U。

[0435] 在一个实施方案中,式C(SEQ ID NO:171)的共有序列具有高GC含量,例如,约50%、60%、70%、80%、90%或99%的GC含量。在一个实施方案中,GC含量为约50%。在一个实施方案中,GC含量为约60%。在一个实施方案中,GC含量为约70%。在一个实施方案中,GC含量为约80%。在一个实施方案中,GC含量为约90%。在一个实施方案中,GC含量为约99%。

[0436] 在一个实施方案中,(b)的编码区包含终止元件,所述终止元件包含式D的共有序列:

[0437] X_{-3}-X_{-2}-X_{-1}-U-A-G-X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}-X_{11}-X_{12}(SEQ ID NO:172)

[0438] 其中:

[0439] X_{-3}、X_{-1}、X_2、X_3、X_{10}各自独立地是G或C;

[0440] X_{-2}或X_9各自独立地是A或U;

[0441] X_1或X_4各自独立地是A或G;

[0442] X_5或X_8各自独立地是A或C;和/或

[0443] X_6、X_7、X_{11}或X_{12}各自独立地是C或U。

[0444] 在一个实施方案中,X_{-3}是G。在一个实施方案中,X_{-3}是C。

[0445] 在一个实施方案中,X_{-1}是G。在一个实施方案中,X_{-1}是C。

[0446] 在一个实施方案中,X_2是G。在一个实施方案中,X_2是C。

[0447] 在一个实施方案中,X_3是G。在一个实施方案中,X_3是C。

[0448] 在一个实施方案中,X_{10}是G。在一个实施方案中,X_{10}是C。

[0449] 在一个实施方案中,X_{-2}是A。在一个实施方案中,X_{-2}是U。

[0450] 在一个实施方案中,X_9是A。在一个实施方案中,X_9是U。

[0451] 在一个实施方案中,X_1是A。在一个实施方案中,X_1是G。

[0452] 在一个实施方案中,X_4是A。在一个实施方案中,X_4是G。

[0453] 在一个实施方案中,X_5是A。在一个实施方案中,X_5是C。

[0454] 在一个实施方案中,X_8是A。在一个实施方案中,X_8是C。

[0455] 在一个实施方案中,X_6是C。在一个实施方案中,X_6是U。

[0456] 在一个实施方案中,X_7是C。在一个实施方案中,X_7是U。

[0457] 在一个实施方案中,X_{11}是C。在一个实施方案中,X_{11}是U。

[0458] 在一个实施方案中,X_{12}是C。在一个实施方案中,X_{12}是U。

[0459] 在一个实施方案中,式D(SEQ ID NO:172)的共有序列具有高GC含量,例如,约50%、60%、70%、80%、90%或99%的GC含量。在一个实施方案中,GC含量为约50%。在一个实施方案中,GC含量为约60%。在一个实施方案中,GC含量为约70%。在一个实施方案中,GC含量为约80%。在一个实施方案中,GC含量为约90%。在一个实施方案中,GC含量为约99%。

[0460] 表4:终止元件

[0461]

SEQ ID NO	序列信息	序列
158	C1 (参考)	UGAUAAUAG
159	C2	UAAUAGUAA
160	C3	UAAGUCUAA
161	C4	UAAAGCUAA
162	C5	UAAGUCUCC
163	C6	UAAGGCUAA
164	C7	UAAGCCCCUCCGGGG

165	C8	UAAAGCUCCCCGGGG
166	C9	UAAGCCCCU
167	C10	UAAAGCUCC
168	C11	UAGGGUUAA
169	C15	UAAGCACCC
170	C12 (UAA共有)	X_{-3}-X_{-2}-X_{-1}-U-A-A-X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}-X_{11}-X_{12} 其中: X_1是G或A; X_2、X_4、X_5、X_6或X_7各自独立地是C或U; X_3是C或A; X_8、X_{10}、X_{11}、X_{12}、X_{-1}或X_{-3}各自独立地是C或G; X_9是G或U;和/或 X_{-2}是A或U。
171	C13 (UGA共有)	X_{-3}-X_{-2}-X_{-1}-U-G-A-X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}-X_{11}-X_{12} 其中: X_{-3}、X_{-1}、X_2、X_5、X_6、X_7、X_8、X_9或X_{12}各自独立地是G或C; X_{-2}、X_3或X_4各自独立地是A或C; X_1是A或G;和/或 X_{10}或X_{11}各自独立地是C或U。
172	C14 (UAG共有)	X_{-3}-X_{-2}-X_{-1}-U-A-G-X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}-X_{11}-X_{12} 其中: X_{-3}、X_{-1}、X_2、X_3、X_{10}各自独立地是G或C; X_{-2}或X_9各自独立地是A或U; X_1或X_4各自独立地是A或G; X_5或X_8各自独立地是A或C;和/或 X_6、X_7、X_{11}或X_{12}各自独立地是C或U。
173	C16	UGAUAGUAA
174	C17	UAAAGCGCU

[0462]

[0463] 在一个方面,本文公开了一种编码多肽的多核苷酸,其中所述多核苷酸包含:(a)5′-UTR,例如,如本文所描述;(b)编码区,所述编码区包含终止元件(例如,如表4中所提供);以及(c)3′-UTR(例如,如本文所描述)。

[0464] 6.多聚A尾

[0465] 在一些实施方案中,本公开的多核苷酸还包含多聚A尾。在其它实施方案中,可以在多聚A尾上并入末端基团以用于稳定化。在其它实施方案中,多聚A尾包含脱3′羟基尾。

[0466] 在RNA加工期间,可以将长链腺嘌呤核苷酸(多聚A尾)添加至多核苷酸(诸如mRNA分子)中以增加稳定性。转录后随即,可以使转录物的3′端裂解以释放3′羟基。然后多聚A聚合酶将腺嘌呤核苷酸链添加至RNA中。所述过程,称为多聚腺苷酸化,添加多聚A尾,其长度可以介于例如约80个至约250个残基之间,包括约80、90、100、110、120、130、140、150、160、

170、180、190、200、210、220、230、240或250个残基的长度。在一个实施方案中,多聚A尾的长度为100个核苷酸(SEQ ID NO:121)。

[0467] aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa (SEQ ID NO:121)

[0468] 也可以在构建体从细胞核中导出后添加多聚A尾。

[0469] 根据本发明,可以在多聚A尾上并入末端基团以用于稳定化。本发明的多核苷酸可以包括脱3′羟基尾。它们还可以包括Junjie Li等人(Current Biology,第15卷,1501-1507,2005年8月23日,其内容以全文引用的方式并入本文中)教示的结构部分或2′-O甲基修饰。

[0470] 本发明的多核苷酸可以被设计成编码具有包括组蛋白mRNA的替代性多聚A尾结构的转录物。根据Norbury,在人类复制依赖的组蛋白mRNA上也检测到末端尿苷酸化。这些mRNA的周转被认为对于在完成或抑制染色体DNA复制后防止潜在的毒性组蛋白积累很重要。这些mRNA的特点是缺少3′多聚(A)尾,其功能替代地由稳定的茎环结构及其同源茎环结合蛋白(SLBP)承担;后者在多聚腺苷酸化mRNA上执行与PABP相同的功能(Norbury,"Cytoplasmic RNA:a case of the tail wagging the dog,"Nature Reviews Molecular Cell Biology;AOP,2013年8月29日在线公布;doi:10.1038/nrm3645),其内容以全文引用的方式并入本文中。

[0471] 独特的多聚A尾长度为本发明的多核苷酸提供了某些优势。一般来说,当存在时,多聚A尾的长度大于30个核苷酸长度。在另一个实施方案中,多聚A尾的长度大于35个核苷酸(例如,至少或大于约35、40、45、50、55、60、70、80、90、100、120、140、160、180、200、250、300、350、400、450、500、600、700、800、900、1,000、1,100、1,200、1,300、1,400、1,500、1,600、1,700、1,800、1,900、2,000、2,500和3,000个核苷酸)。

[0472] 在一些实施方案中,多核苷酸或其区域包括约30个至约3,000个核苷酸(例如,30至50、30至100、30至250、30至500、30至750、30至1,000、30至1,500、30至2,000、30至2,500、50至100、50至250、50至500、50至750、50至1,000、50至1,500、50至2,000、50至2,500、50至3,000、100至500、100至750、100至1,000、100至1,500、100至2,000、100至2,500、100至3,000、500至750、500至1,000、500至1,500、500至2,000、500至2,500、500至3,000、1,000至1,500、1,000至2,000、1,000至2,500、1,000至3,000、1,500至2,000、1,500至2,500、1,500至3,000、2,000至3,000、2,000至2,500以及2,500至3,000)。

[0473] 在一些实施方案中,多聚A尾是相对于整个多核苷酸的长度或多核苷酸的特定区域的长度而设计的。这种设计可以基于编码区的长度、特定特征或区域的长度或基于从多核苷酸表达的最终产物的长度。

[0474] 在这种情形中,多聚A尾的长度可以比多核苷酸或其特征长10、20、30、40、50、60、70、80、90或100%。多聚A尾也可以被设计成它所属的多核苷酸的一部分。在这种情形中,多聚A尾可以是构建体总长度、构建体区域或构建体总长度减去多聚A尾的10、20、30、40、50、60、70、80或90%或更多。此外,多聚A结合蛋白的工程化结合位点和多核苷酸缀合可以增强表达。

[0475] 另外,多个不同的多核苷酸可以经由PABP(多聚A结合蛋白)通过3′端在多聚A尾的3′末端使用修饰的核苷酸连接在一起。转染实验可以在相关细胞系中进行,并且可以在转

染后12小时、24小时、48小时、72小时和第7天通过ELISA测定蛋白质产量。

[0476] 在一些实施方案中,本发明的多核苷酸被设计成包括多聚A-G四联体区。G-四联体是四个鸟嘌呤核苷酸的环状氢键阵列,所述阵列可以由DNA和RNA中富含G的序列形成。在这个实施方案中,G-四联体被并入多聚A尾的末端。测定所得多核苷酸的稳定性、蛋白质产量和其它参数,包括各个时间点的半衰期。已经发现,多聚A-G四联体使得来自mRNA的蛋白质产量相当于单独使用120个核苷酸的多聚A尾(SEQ ID NO:51)所见的至少75%。

[0477] aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa aaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa aaaa (SEQ ID NO: 51)

[0478] 在一些实施方案中,多聚A尾是具有间断非腺苷残基(例如,鸟苷)的混合多聚A尾。在一些实施方案中,多聚A尾是鸟苷酸化的。不希望受理论束缚,据信在一些实施方案中,混合多聚A尾可以保护mRNA免于快速脱腺苷酸化。

[0479] 在一些实施方案中,多聚A尾包含一个或多个非腺苷残基。在一些实施方案中,非腺苷残基是鸟苷。在一些实施方案中,多聚A尾包含1-20,例如1-15、1-10、1-5、15-20、10-20、5-20、2-15、5-10、1-5、2-10或5-15个非腺苷残基(例如,鸟苷)。举例来说,多聚A尾可以包含1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或更多个非腺苷残基(例如,鸟苷)。在一些实施方案中,多聚A尾中至少1%,例如至少2%、5%、10%、15%、20%或25%的残基是非腺苷残基(例如,鸟苷)。在一些实施方案中,多聚A尾是鸟苷酸化的,例如,包含一个或多个鸟苷残基。

[0480] 在一个实施方案中,包含一个或多个非腺苷残基的多聚A尾是化学合成的。

[0481] 在一个实施方案中,3′UTR包含TENT募集序列,例如,如本文所描述,其将一种或多种末端核苷酸转移酶(TENT)募集至包含3′UTR的多核苷酸。在一个实施方案中,TENT是TENT4,例如,TENT4A和/或TENT4B。不希望受理论束缚,据信在一些实施方案中,一个或多个TENT(例如,TENT4A和/或TENT4B)产生具有间断非腺苷残基(例如,鸟苷)的混合多聚A尾,其保护rnRNA免于快速脱腺苷酸化。

[0482] 示例性TENT募集序列包括但不限于

[0483] CACCGCGUUAUCCGUUCCUCGUAGGCUGGUCCUGGGGAA

[0484] CGGGUCGGCGG (SEQ ID NO: 91)和

CCACCCCCAGCGCCACCACCGCUGCCGUCGCCACCGCGUU

[0485] AUCCGUUCCUCGUAGGCUGGUCCUGGGGAACGGGUCGGCGGC CGGUCGGCUUCUGUUUUA (SEQ ID NO: 92)

[0486] 在一个实施方案中,TENT募集序列包含SEQ ID NO:91的核苷酸序列,或与其具有至少80%、85%、90%、95%、96%、97%、98%或99%同一性,或与其相差不超过1、2、3、4、5、6、7、8、9或10个核苷酸的核苷酸序列。在一个实施方案中,TENT募集序列包含SEQ ID NO:91的核苷酸序列。

[0487] 在一个实施方案中,TENT募集序列包含SEQ ID NO:92的核苷酸序列,或与其具有至少80%、85%、90%、95%、96%、97%、98%或99%同一性,或与其相差不超过1、2、3、4、5、

6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个核苷酸的核苷酸序列。在一个实施方案中,TENT募集序列包含SEQ ID NO:92的核苷酸序列。

[0488] 在一个实施方案中,3′UTR包含一个或多个(例如,2、3、4、5个或更多个)TENT募集序列,例如,一个或多个本文所描述的TENT募集序列。在一个实施方案中,3′UTR包含一个TENT募集序列。在一个实施方案中,3′UTR包含两个TENT募集序列。在一个实施方案中,3′UTR包含三个TENT募集序列。在一个实施方案中,3′UTR包含四个TENT募集序列。在一个实施方案中,3′UTR包含五个TENT募集序列。举例来说,3′UTR中的多个TENT募集序列可以相同或不同。

[0489] 在一个实施方案中,3′UTR包含TENT募集序列,所述TENT募集序列包含SEQ ID NO:91的核苷酸序列,或与其具有至少80%、85%、90%、95%、96%、97%、98%或99%同一性,或与其相差不超过1、2、3、4、5、6、7、8、9或10个核苷酸的核苷酸序列。在一个实施方案中,3′UTR包含TENT募集序列,所述TENT募集序列包含SEQ ID NO:91的核苷酸序列。

[0490] 在一个实施方案中,3′UTR包含一个或多个(例如,2、3、4、5个或更多个)TENT募集序列,所述TENT募集序列包含SEQ ID NO:91的核苷酸序列,或与其具有至少80%、85%、90%、95%、96%、97%、98%或99%同一性,或与其相差不超过1、2、3、4、5、6、7、8、9或10个核苷酸的核苷酸序列。在一个实施方案中,3′UTR包含一个TENT募集序列,所述TENT募集序列包含SEQ ID NO:91的核苷酸序列,或与其具有至少80%、85%、90%、95%、96%、97%、98%或99%同一性,或与其相差不超过1、2、3、4、5、6、7、8、9或10个核苷酸的核苷酸序列。在一个实施方案中,3′UTR包含两个TENT募集序列,所述TENT募集序列各自包含SEQ ID NO:91的核苷酸序列,或与其具有至少80%、85%、90%、95%、96%、97%、98%或99%同一性,或与其相差不超过1、2、3、4、5、6、7、8、9或10个核苷酸的核苷酸序列。在一个实施方案中,3′UTR包含三个TENT募集序列,所述TENT募集序列各自包含SEQ ID NO:91的核苷酸序列,或与其具有至少80%、85%、90%、95%、96%、97%、98%或99%同一性,或与其相差不超过1、2、3、4、5、6、7、8、9或10个核苷酸的核苷酸序列。在一个实施方案中,3′UTR包含四个TENT募集序列,所述TENT募集序列各自包含SEQ ID NO:91的核苷酸序列,或与其具有至少80%、85%、90%、95%、96%、97%、98%或99%同一性,或与其相差不超过1、2、3、4、5、6、7、8、9或10个核苷酸的核苷酸序列。在一个实施方案中,3′UTR包含五个TENT募集序列,所述TENT募集序列各自包含SEQ ID NO:91的核苷酸序列,或与其具有至少80%、85%、90%、95%、96%、97%、98%或99%同一性,或与其相差不超过1、2、3、4、5、6、7、8、9或10个核苷酸的核苷酸序列。

[0491] 在一个实施方案中,3′UTR包含一个或多个(例如,2、3、4、5个或更多个)TENT募集序列,所述TENT募集序列包含SEQ ID NO:91的核苷酸序列。在一个实施方案中,3′UTR包含两个TENT募集序列,所述TENT募集序列各自包含SEQ ID NO:91的核苷酸序列。在一个实施方案中,3′UTR包含三个TENT募集序列,所述TENT募集序列各自包含SEQ ID NO:91的核苷酸序列。在一个实施方案中,3′UTR包含四个TENT募集序列,所述TENT募集序列各自包含SEQ ID NO:91的核苷酸序列。在一个实施方案中,3′UTR包含五个TENT募集序列,所述TENT募集序列各自包含SEQ ID NO:91的核苷酸序列。

[0492] 7.额外的3′UTR元件

[0493] A)鉴定和比率确定(IDR)

[0494] 鉴定和比率确定(IDR)序列是生物分子(例如,核酸或蛋白质)的序列,当与靶生物分子的序列组合时,其用于鉴定靶生物分子。通常,IDR序列是结合在靶生物分子的序列内或附加到靶生物分子的序列上的异源序列,并且可以用作鉴定靶分子的参考。因此,在一些实施方案中,核酸(例如,mRNA)包含(i)目标靶序列(例如,编码治疗性和/或抗原性肽或蛋白质的编码序列);以及(ii)独特的IDR序列。

[0495] RNA种类(例如,具有给定编码序列的RNA)可以包含与其它RNA种类(例如,具有不同编码序列的RNA)的IDR序列不同的IDR序列。因此,每个IDR序列鉴定特定的RNA种类,并且因此可以测量IDR序列的丰度以确定组合物中每个RNA种类的丰度。使用不同的IDR序列来鉴定RNA种类允许分析多价RNA组合物(例如,含有多个RNA种类),所述多价RNA组合物含有编码序列和/或长度类似的RNA种类,使用基于PCR或色谱法的全长RNA分析可能难以区分这些RNA种类。

[0496] 多价RNA组合物中的每个RNA种类可以包含IDR序列,所述IDR序列并非是多价RNA组合物中的另一个RNA种类的IDR序列的序列异构体(例如,IDR序列不具有与组合物中的另一种IDR序列相同数量的腺苷核苷酸、相同数量的胞嘧啶核苷酸、相同数量的鸟嘌呤核苷酸和相同数量的尿嘧啶核苷酸,即使那些序列具有不同的序列)。具有相同的核苷酸组成导致序列异构体具有相同的质量,这对使用基于质量的鉴定方法(例如,质谱法)区分序列异构体提出了挑战。

[0497] 多价RNA组合物中的每个RNA种类可以包含IDR序列,其质量不同于多价RNA组合物中每个其它RNA种类的IDR序列的质量。举例来说,每个IDR序列的质量可以与其它IDR序列的质量相差至少9Da、至少25Da、至少25Da或至少50Da。使用具有不同质量的IDR序列允许使用基于质量的分析方法(例如,质谱法)来区分包含不同IDR序列的RNA片段,所述分析方法不需要对RNA进行逆转录、扩增或测序。

[0498] RNA组合物中的每个RNA种类可以包含具有不同长度的IDR序列。举例来说,每个IDR序列可以具有独立地选自0至25个核苷酸的长度。核酸的长度影响核酸穿过色谱柱的速率,因此在不同的RNA种类上使用不同长度的IDR序列允许使用基于色谱的方法(例如,LC-UV)来区分具有不同IDR序列的RNA片段。

[0499] 可以选择IDR序列,使得没有IDR序列包含起始密码子‘AUG’。在IDR序列中缺少起始密码子防止IDR序列内和/或下游的核苷酸序列发生不期望的翻译。

[0500] 可以选择IDR序列,使得没有IDR序列包含限制性酶的识别位点。在一个实例中,没有IDR序列包含XbaI的识别位点‘UCUAG’。缺少限制性酶的识别位点(例如,XbaI识别位点‘UCUAG’)允许限制性酶用于生成和修饰用于体外转录的DNA模板,而不影响IDR序列或转录后的RNA的序列。

[0501] B)FUT8

[0502] 在一些实施方案中,3′UTR包含FUT8序列。举例来说,FUT8序列包含以下序列:

[0503] CUGAGAGACCUGUGUGAACUAUUGAGAAGAUCGGAACAGCUCCUUACUCUGAGGAAGUUG SEQ ID NO:93。在一个实施方案中,3′UTR包含FUT8序列,所述FUT8序列包含SEQ ID NO:93的核苷酸序列,或与其具有至少80%、85%、90%、95%、96%、97%、98%或99%同一性,或与其相差不超过1、2、3、4、5、6、7、8、9或10个核苷酸的核苷酸序列。在一些实施方案中,FUT8序列可以与存在于3′UTR中并如本文所描述的任何miRNA结合位点组合。

[0504] C)核糖体接合检测测定(REDA)

[0505] REDA可以用于评估细胞脂质纳米颗粒-核酸摄取和所制造核酸的mRNA翻译的效力和有效性。所述测定并入核糖体接合检测测定(REDA)的一些方面,以便测量在细胞中的翻译步骤期间结合至核糖体的mRNA。所述测定不需要涉及实际的蛋白质表达,而是通过证明有效的mRNA摄取和与核糖体缔合,并且因此有效地细胞内翻译来代表诸如mRNA的核酸在细胞中产生蛋白质的有效性。

[0506] 因此,在一些实施方案中,如本文所描述的3′UTR序列中的任一者包含可以在REDA中通过qPCR检测的序列。

[0507] RNA种类(例如,具有给定编码序列的RNA)可以包含与其它RNA种类(例如,具有不同编码序列的RNA)的REDA序列不同的REDA序列。因此,每个REDA序列鉴定特定的RNA种类,并且因此可以测量REDA序列的丰度以确定组合物中每个RNA种类的丰度。使用不同的REDA序列来鉴定RNA种类允许分析多价RNA组合物(例如,含有多个RNA种类),所述多价RNA组合物含有编码序列和/或长度类似的RNA种类,使用基于PCR或色谱法的全长RNA分析可能难以区分这些RNA种类。

[0508] 多价RNA组合物中的每个RNA种类可以包含REDA序列,所述REDA序列并非是多价RNA组合物中的另一个RNA种类的REDA序列的序列异构体(例如,IDR序列不具有与组合物中的另一种REDA序列相同数量的腺苷核苷酸、相同数量的胞嘧啶核苷酸、相同数量的鸟嘌呤核苷酸和相同数量的尿嘧啶核苷酸,即使那些序列具有不同的序列)。具有相同的核苷酸组成导致序列异构体具有相同的质量,这对使用基于质量的鉴定方法(例如,质谱法)区分序列异构体提出了挑战。

[0509] 多价RNA组合物中的每个RNA种类可以包含REDA序列,其质量不同于多价RNA组合物中每个其它RNA种类的REDA序列的质量。举例来说,每个REDA序列的质量可以与其它REDA序列的质量相差至少9Da、至少25Da、至少25Da或至少50Da。使用具有不同质量的REDA序列允许使用基于质量的分析方法(例如,质谱法)来区分包含不同IDR序列的RNA片段,所述分析方法不需要对RNA进行逆转录、扩增或测序。

[0510] RNA组合物中的每个RNA种类可以包含具有不同长度的REDA序列。举例来说,每个IDR序列可以具有独立地选自0至25个核苷酸的长度。核酸的长度影响核酸穿过色谱柱的速率,因此在不同的RNA种类上使用不同长度的REDA序列允许使用基于色谱的方法(例如,LC-UV)来区分具有不同REDA序列的RNA片段。

[0511] 可以选择REDA序列,使得没有REDA序列包含起始密码子‘AUG’。在IDR序列中缺少起始密码子防止REDA序列内和/或下游的核苷酸序列发生不期望的翻译。

[0512] 可以选择REDA序列,使得没有REDA序列包含限制性酶的识别位点。在一个实例中,没有REDA序列包含XbaI的识别位点‘UCUAG’。缺少限制性酶的识别位点(例如,XbaI识别位点‘UCUAG’)允许限制性酶用于生成和修饰用于体外转录的DNA模板,而不影响IDR序列或转录后的RNA的序列。

[0513] 8.起始密码子区域

[0514] 本发明还包括一种包含起始密码子区域和本文所描述的多核苷酸两者的多核苷酸。在一些实施方案中,本发明的多核苷酸可以具有与起始密码子区域类似或功能类似的区域。

[0515] 在一些实施方案中,多核苷酸的翻译可以在不是起始密码子AUG的密码子上开始。多核苷酸的翻译可以在替代性起始密码子上开始,诸如但不限于ACG、AGG、AAG、CTG/CUG、GTG/GUG、ATA/AUA、ATT/AUU、TTG/UUG(参见Touriol等人,Biology of the Cell 95(2003)169-178以及Matsuda和Mauro PLoS ONE,2010 5:11;所述文献中每一者的内容以全文引用的方式并入本文中)。

[0516] 作为一个非限制性实例,多核苷酸的翻译始于替代性起始密码子ACG。作为另一个非限制性实例,多核苷酸的翻译始于替代性起始密码子CTG或CUG。作为又一个非限制性实例,多核苷酸的翻译始于替代性起始密码子GTG或GUG。

[0517] 已知位于启动翻译的密码子(诸如但不限于起始密码子或替代性起始密码子)侧翼的核苷酸会影响多核苷酸的翻译效率、长度和/或结构。(参见例如Matsuda和Mauro PLoS ONE,2010 5:11;其内容以全文引用的方式并入本文中)。掩蔽位于启动翻译的密码子侧翼的核苷酸中的任一者可以用于改变翻译起始的位置、多核苷酸的翻译效率、长度和/或结构。

[0518] 在一些实施方案中,可以在起始密码子或替代性起始密码子附近使用掩蔽剂来掩蔽或隐藏密码子以降低在掩蔽的起始密码子或替代性起始密码子处翻译起始的概率。掩蔽剂的非限制性实例包括反义锁核酸(LNA)多核苷酸和外显子连接复合物(EJC)(参见例如Matsuda和Mauro描述了掩蔽剂LNA多核苷酸和EJC(PLoS ONE,2010 5:11);其内容以全文引用的方式并入本文中)。

[0519] 在另一个实施方案中,可以使用掩蔽剂来掩蔽多核苷酸的起始密码子以增加翻译将在替代性起始密码子上开始的可能性。在一些实施方案中,可以使用掩蔽剂来掩蔽第一起始密码子或替代性起始密码子,以增加翻译在掩蔽的起始密码子或替代性起始密码子下游的起始密码子或替代性起始密码子上开始的机会。

[0520] 在一些实施方案中,起始密码子或替代性起始密码子可以位于miRNA结合位点的完美补体内。miRNA结合位点的完美补体可以类似于掩蔽剂来帮助控制多核苷酸的翻译、长度和/或结构。作为一个非限制性实例,起始密码子或替代性起始密码子可以位于miRNA结合位点的完美补体的中间。起始密码子或替代性起始密码子可以位于第一个核苷酸、第二个核苷酸、第三个核苷酸、第四个核苷酸、第五个核苷酸、第六个核苷酸、第七个核苷酸、第八个核苷酸、第九个核苷酸、第十个核苷酸、第十一个核苷酸、第十二个核苷酸、第十三个核苷酸、第十四个核苷酸、第十五个核苷酸、第十六个核苷酸、第十七个核苷酸、第十八个核苷酸、第十九个核苷酸、第二十个核苷酸或第二十一个核苷酸之后。

[0521] 在另一个实施方案中,可以从多核苷酸序列中去除多核苷酸的起始密码子,以使多核苷酸的翻译在不是起始密码子的密码子上开始。多核苷酸的翻译可以在去除的起始密码子之后的密码子上或在下游起始密码子或替代性起始密码子上开始。在一个非限制性实例中,起始密码子ATG或AUG作为多核苷酸序列的初始3个核苷酸被去除以在下游起始密码子或替代性起始密码子上开始翻译。去除起始密码子的多核苷酸序列还可以包含至少一种用于下游起始密码子和/或替代性起始密码子的掩蔽剂,以控制或试图控制翻译的起始、多核苷酸的长度和/或多核苷酸的结构。

[0522] 9.mRNA元件的组合

[0523] 本文所公开的多核苷酸中的任一者可以包含以下元件中的一者、两者、三者或全

部:5'-UTR,例如,如本文所描述;编码区;终止元件+3'-UTR(例如,如本文所描述);以及任选地,3′稳定区,例如,如本文所描述。本文还公开了包含所述多核苷酸的LNP组合物。

[0524] 在一个实施方案中,本公开的多核苷酸包含表1中所描述的5′UTR或其变体或片段,以及终止元件+包含SEQ ID NO:139的序列的3′UTR或其变体或片段。在一个实施方案中,多核苷酸还包含帽结构,例如,如本文所描述,或多聚A尾,例如,如本文所描述。在一个实施方案中,多核苷酸还包含3′稳定区,例如,如本文所描述。

[0525] 在一个实施方案中,本公开的多核苷酸包括包含SEQ ID NO:50的序列的5′UTR或其变体或片段,以及终止元件+包含SEQ ID NO:139的序列的3′UTR或其变体或片段。在一个实施方案中,多核苷酸还包含帽结构,例如,如本文所描述,或多聚A尾,例如,如本文所描述。在一个实施方案中,多核苷酸还包含3′稳定区,例如,如本文所描述。

[0526] 在一个实施方案中,本公开的多核苷酸包括包含SEQ ID NO:56的序列的5′UTR或其变体或片段,以及终止元件+包含SEQ ID NO:139的序列的3′UTR或其变体或片段。在一个实施方案中,多核苷酸还包含帽结构,例如,如本文所描述,或多聚A尾,例如,如本文所描述。在一个实施方案中,多核苷酸还包含3′稳定区,例如,如本文所描述。

[0527] 在一些实施方案中,5'UTR是SEQ ID NO:50并且3'UTR是SEQ ID NO:139或SEQ ID NO:144。举例来说,5'UTR是SEQ ID NO:50并且3'UTR是SEQ ID NO:139,或者5'UTR是SEQ ID NO:50并且3′UTR是SEQ ID NO:144。在一个实施方案中,多核苷酸还包含帽结构,例如,如本文所描述,或多聚A尾,例如,如本文所描述。在一个实施方案中,多核苷酸还包含3′稳定区,例如,如本文所描述。

[0528] 在一些实施方案中,选自表3的miRNA结合位点序列中的任一者或多者可以与如表4中所示的终止盒中的任一者组合。另外,在一些实施方案中,Tent募集序列可以与选自表3的miRNA结合位点序列中的任一者或多者和如表4中所示的终止盒中的任一者组合。此外,在一些实施方案中,FUT8序列可以与选自表3的miRNA结合位点序列中的任一者或多者和如表4中所示的终止盒中的任一者组合。

[0529] 10.治疗性有效载荷或预防性有效载荷

[0530] 本文尤其公开了一种具有本文所描述的5′UTR、本文所描述的3′UTR和/或包含终止元件的编码区的多核苷酸,所述编码区还包含编码有效载荷,例如治疗性有效载荷或预防性有效载荷的序列。在一个实施方案中,编码区编码一个有效载荷。在一个实施方案中,编码区编码多于一个有效载荷,例如2、3、4、5、6个或更多个有效载荷,例如相同或不同的有效载荷。在一个实施方案中,编码每个有效载荷的序列在多核苷酸中是连续的。在一个实施方案中,编码每个有效载荷的序列由至少1-1000个核苷酸隔开。在一些实施方案中,治疗性有效载荷或预防性有效载荷包含编码以下的mRNA:分泌蛋白;膜结合蛋白;或细胞间蛋白,或肽、多肽或其生物活性片段。

[0531] 本文还公开了一种包含多核苷酸的LNP,所述多核苷酸包含编码有效载荷,例如治疗性有效载荷或预防性有效载荷的编码区。在一些实施方案中,治疗性有效载荷或预防性有效载荷包含编码以下的mRNA:分泌蛋白;膜结合蛋白;或细胞间蛋白,或肽、多肽或其生物活性片段。

[0532] 在一些实施方案中,治疗性有效载荷或预防性有效载荷包含编码分泌蛋白或肽、多肽或其生物活性片段的mRNA。在一些实施方案中,分泌蛋白包含细胞因子或其变体或片

段(例如,生物活性片段)。在一些实施方案中,分泌蛋白包含抗体或其变体或片段(例如,生物活性片段)。在一些实施方案中,分泌蛋白包含酶或其变体或片段(例如,生物活性片段)。在一些实施方案中,分泌蛋白包含激素或其变体或片段(例如,生物活性片段)。在一些实施方案中,分泌蛋白包含配体或其变体或片段(例如,生物活性片段)。在一些实施方案中,分泌蛋白包含疫苗(例如,抗原、免疫原性表位)或其组分、变体或片段(例如,生物活性片段)。在一些实施方案中,疫苗是预防性疫苗。在一些实施方案中,疫苗是治疗性疫苗,例如,癌症疫苗。在一些实施方案中,分泌蛋白包含生长因子或其组分、变体或片段(例如,生物活性片段)。在一些实施方案中,分泌蛋白包含免疫调节剂,例如,免疫检查点激动剂或拮抗剂。

[0533] 在一些实施方案中,治疗性有效载荷或预防性有效载荷包含编码膜结合蛋白或肽、多肽或其生物活性片段的mRNA。在一些实施方案中,膜结合蛋白包含疫苗(例如,抗原、免疫原性表位)或其组分、变体或片段(例如,生物活性片段)。在一些实施方案中,疫苗是预防性疫苗。在一些实施方案中,疫苗是治疗性疫苗,例如,癌症疫苗。在一些实施方案中,膜结合蛋白包含配体、其变体或片段(例例如,生物活性片段)。在一些实施方案中,膜结合蛋白包含膜转运体、其变体或片段(例如,生物活性片段)。在一些实施方案中,膜结合蛋白包含结构蛋白、其变体或片段(例如,生物活性片段)。在一些实施方案中,膜结合蛋白包含免疫调节剂,例如,免疫检查点激动剂或拮抗剂。

[0534] 在一些实施方案中,治疗性有效载荷或预防性有效载荷包含编码细胞内蛋白或肽、多肽或其生物活性片段的mRNA。在一些实施方案中,细胞内蛋白包含酶或其变体或片段(例如,生物活性片段)。在一些实施方案中,细胞内蛋白包含转录因子或其变体或片段(例如,生物活性片段)。在一些实施方案中,细胞内蛋白包含核酸酶或其变体或片段(例如,生物活性片段)。在一些实施方案中,细胞内蛋白包含结构蛋白或其变体或片段(例如,生物活性片段)。

[0535] 在一些实施方案中,治疗性有效载荷或预防性有效载荷选自细胞因子、抗体、疫苗(例如,抗原、免疫原性表位)、受体、酶、激素、转录因子、配体、膜转运体、结构蛋白、核酸酶、生长因子、免疫调节剂或其组分、变体或片段(例如,生物活性片段)。

[0536] 在一些实施方案中,治疗性有效载荷或预防性有效载荷包含蛋白质或肽。

[0537] 将了解,本文所公开的调控元件(例如,5'UTR、终止元件、3'UTR、稳定区(例如,idT或修饰的多聚A尾)可以与编码本文所描述的有效载荷的ORF一起使用。将进一步了解,本文所公开的调控元件可以按模块化方式使用,即,可以在与来自本领域的其它调控元件组合的mRNA构建体中使用(例如,本发明的5'UTR与ORF和来自本领域的其它调控区组合),或可以与本文所公开的其它调控元件(例如,本发明的5'UTR和本发明的3'UTR等)组合使用。将进一步了解,本发明的终止元件可以与缺少终止密码子的所需ORF组合使用。还将了解,当所需ORF包含终止密码子时,额外的终止密码子或终止元件将不包括于最终构建体中。在一些实施方案中,所需ORF中的终止密码子可以用本文所描述的终止元件置换。

[0538] 11.制备多核苷酸的方法

[0539] 本公开还提供了用于制备本文所公开的多核苷酸或其补体的方法。在一些方面,本文所公开的多核苷酸(例如,mRNA)可以使用体外转录来构建。

[0540] 在其它方面,本文所公开的多核苷酸(例如,mRNA)可以通过使用寡核苷酸合成仪进行化学合成来构建。在其它方面,本文所公开的多核苷酸(例如,mRNA)通过使用宿主细胞

来制备。在某些方面,本文所公开的多核苷酸(例如,mRNA)通过IVT、化学合成、宿主细胞表达或本领域中已知的任何其它方法的一种或多种组合来制备。

[0541] 天然存在的核苷、非天然存在的核苷或它们的组合可以完全地或部分地置换存在于候选核苷酸序列中的天然存在的核苷,并且可以并入编码治疗性有效载荷或预防性有效载荷的序列优化的核苷酸序列(例如,mRNA)中。然后可以检查所得mRNA产生蛋白质和/或产生治疗结果的能力。

[0542] 虽然可以使用本领域中熟知的方法以合成方式制备RNA,但在一个实施方案中,通过在会产生RNA转录物的条件下使DNA模板与RNA聚合酶(例如,T7 RNA聚合酶或T7 RNA聚合酶变体)接触来合成RNA转录物(例如,mRNA转录物)。

[0543] 在一些方面,本公开提供了进行IVT(体外转录)反应的方法,所述方法包括在三磷酸核苷和缓冲液存在下在会产生RNA转录物的条件下使DNA模板与RNA聚合酶(例如T7 RNA聚合酶,诸如T7 RNA聚合酶变体)接触。

[0544] 本公开的其它方面提供了加帽方法,例如,共转录加帽方法或本领域中已知的其它方法。在一个实施方案中,加帽方法包括在体外转录反应条件下使多核苷酸模板与T7 RNA聚合酶变体、三磷酸核苷和帽类似物反应以产生RNA转录物。

[0545] IVT条件通常需要纯化的线性DNA模板,所述模板含有启动子、三磷酸核苷、包含二硫苏糖醇(DTT)和镁离子的缓冲系统以及RNA聚合酶。转录反应中使用的确切条件取决于特定应用所需的RNA量。典型的IVT反应通过在转录缓冲液中孵育DNA模板与RNA聚合酶和三磷酸核苷来进行,所述三磷酸核苷包括GTP、ATP、CTP和UTP(或核苷酸类似物)。由这种反应产生具有5′末端三磷酸鸟苷的RNA转录物。

[0546] 脱氧核糖核酸(DNA)仅仅是RNA聚合酶的核酸模板。DNA模板可以包括编码所关注的多肽(例如,抗原性多肽)的多核苷酸。在一些实施方案中,DNA模板包括RNA聚合酶启动子(例如,T7 RNA聚合酶启动子),所述启动子位于编码所关注多肽的多核苷酸的5′并且可操作地连接至所述多核苷酸。DNA模板还可以包括编码位于所关注基因的3′端的多聚腺苷酸化(多聚A)尾的核苷酸序列。

[0547] 所关注的多肽包括但不限于生物剂、抗体、抗原(疫苗)和治疗性蛋白质。术语“蛋白质”涵盖肽。

[0548] 在一些实施方案中,RNA转录物是IVT反应的产物,并且如本领域普通技术人员将了解,用于制备RNA分子的DNA模板是基于碱基互补性而已知的。在一些实施方案中,RNA转录物是信使RNA(mRNA),其包括编码连接至多聚A尾的所关注多肽的核苷酸序列。在一些实施方案中,mRNA是修饰的rnRNA(mmRNA),其包括至少一个修饰的核苷酸。

[0549] 核苷酸包括含氮碱基、五碳糖(核糖或脱氧核糖)和至少一个磷酸基团。核苷酸包括单磷酸核苷、二磷酸核苷和三磷酸核苷。单磷酸核苷(NMP)包括连接至核糖和单个磷酸的核碱基;二磷酸核苷(NDP)包括连接至核糖和两个磷酸的核碱基;并且三磷酸核苷(NTP)包括连接至核糖和三个磷酸的核碱基。核苷酸类似物是具有核苷酸的一般结构或在结构上类似于核苷酸的化合物。举例来说,核苷酸类似物包括核碱基的类似物、糖的类似物和/或核苷酸的一个或多个磷酸基团的类似物。

[0550] 核苷包括含氮碱基和5碳糖。因此,核苷加上磷酸基团产生核苷酸。核苷类似物是具有核苷的一般结构或在结构上类似于核苷的化合物。举例来说,核苷类似物包括核碱基

的类似物和/或核苷糖的类似物。

[0551] 应了解,除非另有指示,否则术语“核苷酸”包括天然存在的核苷酸、合成核苷酸和修饰的核苷酸。如本文所提供,用于产生RNA(例如,在IVT反应中)的天然存在的核苷酸的实例包括三磷酸腺苷(ATP)、三磷酸鸟苷(GTP)、三磷酸胞苷(CTP)、三磷酸尿苷(UTP)和三磷酸5-甲基尿苷(m^5UTP)。在一些实施方案中,使用二磷酸腺苷(ADP)、二磷酸鸟苷(GDP)、二磷酸胞苷(CDP)和/或二磷酸尿苷(UDP)。

[0552] 核苷酸类似物的实例包括但不限于抗病毒核苷酸类似物、磷酸类似物(可溶的或固定的、可水解的或不可水解的)、二核苷酸、三核苷酸、四核苷酸,例如,帽类似物,或用于酶加帽(牛痘或连接酶)的前体/底物,用官能团标记以促进帽或5’部分(IRES)的连接/缀合的核苷酸,用5′PO_4标记以促进帽或5′部分的连接的核苷酸,或用可以化学或酶促裂解的官能团/保护基标记的核苷酸。抗病毒核苷酸/核苷类似物的实例包括但不限于更昔洛韦(Ganciclovir)、恩替卡韦(Entecavir)、替比夫定(Telbivudine)、阿糖腺苷(Vidarabine)和西多福韦(Cidofovir)。

[0553] 修饰的核苷酸可以包括修饰的核碱基。举例来说,本公开的RNA转录物(例如,mRNA转录物)可以包括选自以下的修饰的核碱基:假尿苷(ψ)、1-甲基假尿苷(m1ψ)、1-乙基假尿苷、2-硫尿苷、4′-硫尿苷、2-硫-1-甲基-1-脱氮-假尿苷、2-硫-1-甲基-假尿苷、2-硫-5-氮杂-尿苷、2-硫-二氢假尿苷、2-硫-二氢尿苷、2-硫-假尿苷、4-甲氧基-2-硫-假尿苷、4-甲氧基-假尿苷、4-硫-1-甲基-假尿苷、4-硫-假尿苷、5-氮杂-尿苷、二氢假尿苷、5-甲基尿苷、5-甲氧基尿苷(mo5U)和2′-O-甲基尿苷。在一些实施方案中,RNA转录物(例如,mRNA转录物)包括至少两种(例如,2种、3种、4种或更多种)前述修饰的核碱基的组合。

[0554] 如本文所提供的三磷酸核苷(NTP)可以包含未修饰或修饰的ATP、修饰或未修饰的UTP、修饰或未修饰的GTP和/或修饰或未修饰的CTP。在一些实施方案中,IVT反应的NTP包含未修饰的ATP。在一些实施方案中,IVT反应的NTP包含修饰的ATP。在一些实施方案中,IVT反应的NTP包含未修饰的UTP。在一些实施方案中,IVT反应的NTP包含修饰的UTP。在一些实施方案中,IVT反应的NTP包含未修饰的GTP。在一些实施方案中,IVT反应的NTP包含修饰的GTP。在一些实施方案中,IVT反应的NTP包含未修饰的CTP。在一些实施方案中,IVT反应的NTP包含修饰的CTP。

[0555] 存在于IVT反应中的三磷酸核苷和帽类似物的浓度可能有变化。在一些实施方案中,NTP和帽类似物以等摩尔浓度存在于反应中。在一些实施方案中,反应中帽类似物(例如,三核苷酸帽)与三磷酸核苷的摩尔比大于1:1。举例来说,反应中帽类似物与三磷酸核苷的摩尔比可以为2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、15:1、20:1、25:1、50:1或100:1。在一些实施方案中,反应中帽类似物(例如,三核苷酸帽)与三磷酸核苷的摩尔比小于1:1。举例来说,反应中帽类似物(例如,三核苷酸帽)与三磷酸核苷的摩尔比可以为1:2、1:3、1:4、1:5、1:6、1:7、1:8、1:9、1:10、1:1 5、1:20、1:25、1:50或1:100。

[0556] IVT反应中NTP的组成也可能有变化。举例来说,ATP可以在过量GTP、CTP和UTP下使用。作为一个非限制性实例,IVT反应可以包括7.5毫摩尔GTP、7.5毫摩尔CTP、7.5毫摩尔UTP和3.75毫摩尔ATP。相同的IVT反应可以包括3.75毫摩尔帽类似物(例如,三核苷酸帽)。在一些实施方案中,G:C:U:A:帽的摩尔比为1:1:1:0.5:0.5。在一些实施方案中,G:C:U:A:帽的摩尔比为1:1:0.5:1:0.5。在一些实施方案中,G:C:U:A:帽的摩尔比为1:0.5:1:1:0.5。在一

些实施方案中,G:C:U:A:帽的摩尔比为0.5:1:1:1:0.5。

[0557] 在一些实施方案中,RNA转录物(例如,mRNA转录物)包括选自以下的修饰的核碱基:假尿苷(ψ)、1-甲基假尿苷($m^1\psi$)、5-甲氧基尿苷(mo^5U)、5-甲基胞苷(m^5C)、α-硫-鸟苷或α-硫-腺苷。在一些实施方案中,RNA转录物(例如,mRNA转录物)包括至少两种(例如,2种、3种、4种或更多种)前述修饰的核碱基的组合。

[0558] 在一些实施方案中,RNA转录物(例如,mRNA转录物)包括假尿苷(ψ)。在一些实施方案中,RNA转录物(例如,mRNA转录物)包括1-甲基假尿苷($m^1\psi$)。在一些实施方案中,RNA转录物(例如,mRNA转录物)包括5-甲氧基尿苷(mo^5U)。在一些实施方案中,RNA转录物(例如,mRNA转录物)包括5-甲基胞苷(m^5C)。在一些实施方案中,RNA转录物(例如,mRNA转录物)包括α-硫-鸟苷。在一些实施方案中,RNA转录物(例如,mRNA转录物)包括αα-硫-腺苷。

[0559] 在一些实施方案中,多核苷酸(例如RNA多核苷酸,诸如mRNA多核苷酸)被统一修饰(例如,完全修饰,遍及整个序列进行修饰)以用于特定修饰。举例来说,多核苷酸可以统一被1-甲基假尿苷($m^1\psi$)修饰,意指mRNA序列中的所有尿苷残基均被1-甲基假尿苷($m^1\psi$)置换。类似地,多核苷酸可以通过用修饰的残基(诸如上文所阐述的那些残基中的任一者)置换而针对存在于序列中的任何类型的核苷残基进行统一修饰。或者,多核苷酸(例如RNA多核苷酸,诸如mRNA多核苷酸)可以不被统一修饰(例如,部分修饰,序列的一部分被修饰)。每种可能性代表本发明的单独实施方案。

[0560] 在一些实施方案中,缓冲系统含有tris。举例来说,IVT反应中使用的tris的浓度可以为至少10mM、至少20mM、至少30mM、至少40mM、至少50mM、至少60mM、至少70mM、至少80mM、至少90mM、至少100mM或至少110 mM磷酸盐。在一些实施方案中,磷酸盐的浓度为20-60mM或10-100mM。

[0561] 在一些实施方案中,缓冲系统含有二硫苏糖醇(DTT)。举例来说,IVT反应中使用的DTT的浓度可以为至少1mM、至少5mM或至少50mM。在一些实施方案中,IVT反应中使用的DTT的浓度为1-50mM或5-50mM。在一些实施方案中,IVT反应中使用的DTT的浓度为5mM。

[0562] 在一些实施方案中,缓冲系统含有镁。在一些实施方案中,存在于IVT反应中的NTP与镁离子(Mg^{2+};例如,$MgCl_2$)的摩尔比为1:1至1:5。举例来说,NTP与镁离子的摩尔比可以为1:1、1:2、1:3、1:4或1:5。

[0563] 在一些实施方案中,存在于IVT反应中的NTP加上帽类似物(例如三核苷酸帽,诸如GAG)与镁离子(Mg^{2+};例如,$MgCl_2$)的摩尔比为1:1至1:5。举例来说,NTP+三核苷酸帽(例如,GAG)与镁离子的摩尔比可以为1:1、1:2、1:3、1:4或1:5。

[0564] 在一些实施方案中,缓冲系统含有Tris-HCl、亚精胺(例如,浓度为1-30mM)、TRITON®X-100(聚乙二醇对(1,1,3,3-四甲基丁基)-苯基醚)和/或聚乙二醇(PEG)。

[0565] 将三磷酸核苷(NTP)添加至生长的RNA链的3′端是由聚合酶催化的,诸如T7 RNA聚合酶,例如,本公开的T7 RNA聚合酶变体(例如,G47A)中的任何一者或多者。在一些实施方案中,RNA聚合酶(例如,T7 RNA聚合酶变体)以0.01mg/ml至1mg/ml的浓度存在于反应(例如,IVT反应)中。举例来说,RNA聚合酶可能以0.01mg/mL、0.05mg/ml、0.1mg/ml、0.5mg/ml或1.0mg/ml的浓度存在于反应中。

[0566] 在一些实施方案中,本公开的多核苷酸是IVT多核苷酸。传统上,mRNA分子的基本组分至少包括编码区、5′UTR、3′UTR、5′帽和多聚A尾。本公开的IVT多核苷酸可以作为mRNA

起作用,但在它们的功能和/或结构设计特征上不同于野生型mRNA,这些特征用于例如克服使用基于核酸的治疗剂有效产生多肽的现有问题。

[0567] IVT多核苷酸的初级构建体包含连接的核苷酸的第一区,其侧接有第一侧翼区和第二侧翼区。这个第一区可以包括但不限于编码的治疗性有效载荷或预防性有效载荷。第一侧翼区可以包括作为5′非翻译区(UTR)起作用的连接核苷的序列,诸如编码多肽的原生5′UTR或非原生5'UTR的核酸中的任一者的5'UTR,诸如但不限于异源5′UTR或合成5′UTR。编码治疗性有效载荷或预防性有效载荷的IVT可以在其5末端包含编码一个或多个信号序列的信号序列区。侧翼区可以包含连接的核苷酸的区域,所述连接的核苷酸包含一个或多个完整或不完整的5′UTR序列。侧翼区还可以包括5′末端帽。第二侧翼区可以包含连接的核苷酸的区域,所述连接的核苷酸包含一个或多个完整或不完整的3′UTR,其可以编码治疗性有效载荷或预防性有效载荷的原生3′UTR,或非原生3′UTR,诸如但不限于异源3′UTR或合成3′UTR。侧翼区还可以包含3′加尾序列。3′加尾序列可以是但不限于多聚A尾、多聚A-G四联体和/或茎环序列。

[0568] IVT多核苷酸架构的额外和示例性特征以及制备多核苷酸的方法在2017年5月18日提交的国际PCT申请WO 2017/201325中公开,所述申请的全部内容特此以引用的方式并入。

[0569] 12.纯化

[0570] 在其它方面,可以纯化本文所公开的多核苷酸(例如,mRNA)。本文所描述的多核苷酸(例如,mRNA)的纯化可以包括但不限于多核苷酸清理、质量保证和质量控制。可以通过本领域中已知的方法进行清理,诸如但不限于AGENCOURT®珠粒(Beckman Coulter Genomics,Danvers,MA)、多聚T珠粒、LNATM寡聚T捕获探针(EXIQON®Inc,Vedbaek,Denmark)或基于HPLC的纯化方法,诸如但不限于强阴离子交换HPLC、弱阴离子交换HPLC、反相HPLC(RP-HPLC)和疏水相互作用HPLC(HIC-HPLC)。术语“纯化的”当关于多核苷酸,诸如“纯化的多核苷酸”使用时,是指与至少一种污染物分离的多核苷酸。如本文所用,“污染物”是使另一种物质不适合、不纯或劣等的任何物质。因此,纯化的多核苷酸(例如,DNA和RNA)以不同于其在自然界中发现的形式或配置存在,或以不同于对其进行处理或纯化方法之前存在的形式或配置存在。

[0571] 在一些实施方案中,本公开的多核苷酸(例如,m_{mm}RNA)的纯化去除了杂质,这可以减少或去除非所需的免疫反应,例如,降低细胞因子活性。

[0572] 在一些实施方案中,本公开的多核苷酸(例如,mRNA)在施用之前使用柱色谱法(例如,强阴离子交换HPLC、弱阴离子交换HPLC、反相HPLC(RP-HPLC)和疏水相互作用HPLC(HIC-HPLC)或(LCMS))进行纯化。在一些实施方案中,柱色谱法(例如,强阴离子交换HPLC、弱阴离子交换HPLC、反相HPLC(RP-HPLC)和疏水相互作用HPLC(HIC-HPLC)或(LCMS))纯化的多核苷酸,其编码本文所公开的治疗性有效载荷或预防性有效载荷,与通过不同纯化方法纯化的编码治疗性有效载荷或预防性有效载荷的多核苷酸相比,增加了治疗性有效载荷或预防性有效载荷的表达。

[0573] 在一些实施方案中,柱色谱法(例如,强阴离子交换HPLC、弱阴离子交换HPLC、反相HPLC(RP-HPLC)和疏水相互作用HPLC(HIC-HPLC)或(LCMS))纯化的多核苷酸编码治疗性有效载荷或预防性有效载荷。在一些实施方案中,纯化的多核苷酸编码治疗性有效载荷或预

防性有效载荷。

[0574] 在一些实施方案中,纯化的多核苷酸为至少约80%纯的、至少约85%纯的、至少约90%纯的、至少约95%纯的、至少约96%纯的、至少约97%纯的、至少约98%纯的、至少约99%纯的或约100%纯的。

[0575] 质量保证和/或质量控制检查可以使用诸如但不限于凝胶电泳、UV吸光度或分析型HPLC的方法进行。

[0576] 在另一个实施方案中,可以通过包括但不限于逆转录酶-PCR的方法对多核苷酸进行测序。

[0577] 13.多核苷酸的化学修饰

[0578] 如上文所描述,核酸(例如RNA核酸,诸如mRNA核酸)的修饰的核苷和核苷酸可以包括于本发明的多核苷酸中。"核苷"是指含有与有机碱(例如,嘌呤或嘧啶)或其衍生物(本文中也称为"核碱基")组合的糖分子(例例如,戊糖或核糖)或其衍生物的化合物。"核苷酸"是指包括磷酸基团的核苷。修饰的核苷酸可以通过诸如化学、酶促或重组的任何有用方法合成,以包括一个或多个修饰的或非天然的核苷。核酸可以包含连接的核苷的一个或多个区域。此类区域可能具有可变的主链键联。键联可以是标准的磷酸二酯键联,在这种情况下,核酸将包含核苷酸区域。

[0579] 修饰的核苷酸碱基配对不仅涵盖标准的腺苷-胸腺嘧啶、腺苷-尿嘧啶或鸟苷-胞嘧啶碱基对,而且涵盖核苷酸和/或包含非标准或修饰碱基的修饰核苷酸之间形成的碱基对,其中氢键供体和氢键接受体的排列允许非标准碱基与标准碱基之间或两个互补的非标准碱基结构之间的氢键结,例如,在具有至少一个化学修饰的那些核酸中。此种非标准碱基配对的一个实例是修饰的核苷酸肌苷与腺嘌呤、胞嘧啶或尿嘧啶之间的碱基配对。碱基/糖或连接子的任何组合可以并入本公开的核酸中。

[0580] 在一些实施方案中,核酸(例如RNA核酸,诸如mRNA核酸)中的修饰核碱基包括N1-甲基-假尿苷(m1ψ)、1-乙基-假尿苷(e1ψ)、5-甲氧基-尿苷(mo5U)、5-甲基-胞苷(m5C)和/或假尿苷(ψ)。在一些实施方案中,核酸(例如RNA核酸,诸如mRNA核酸)中的修饰核碱基包括5-甲氧基甲基尿苷、5-甲硫基尿苷、1-甲氧基甲基假尿苷、5-甲基胞苷和/或5-甲氧基胞苷。在一些实施方案中,多核糖核苷酸包括至少两种(例如,2种、3种、4种或更多种)任何上述修饰的核碱基的组合,包括但不限于化学修饰。

[0581] 在一些实施方案中,本公开的RNA核酸在核酸的一个或多个或所有尿苷位置处包含N1-甲基-假尿苷(m1ψ)取代。

[0582] 在一些实施方案中,本公开的RNA核酸在核酸的一个或多个或所有尿苷位置处包含N1-甲基-假尿苷(m1ψ)取代并且在核酸的一个或多个或所有胞苷位置处包含5-甲基胞苷取代。

[0583] 在一些实施方案中,本公开的RNA核酸在核酸的一个或多个或所有尿苷位置处包含假尿苷(ψ)取代。

[0584] 在一些实施方案中,本公开的RNA核酸在核酸的一个或多个或所有尿苷位置处包含假尿苷(ψ)取代并且在核酸的一个或多个或所有胞苷位置处包含5-甲基胞苷取代。

[0585] 在一些实施方案中,本公开的RNA核酸在核酸的一个或多个或所有尿苷位置处包含尿苷。

[0586] 在一些实施方案中,核酸(例如RNA核酸,诸如mRNA核酸)被统一修饰(例如,完全修饰,遍及整个序列进行修饰)以用于特定修饰。举例来说,核酸可以统一被N1-甲基-假尿苷修饰,意指mRNA序列中的所有尿苷残基均被N1-甲基-假尿苷置换。类似地,核酸可以通过用修饰的残基(诸如上文所阐述的那些残基)置换而针对存在于序列中的任何类型的核苷残基进行统一修饰。

[0587] 本公开的核酸可以沿着分子的整个长度被部分或完全修饰。举例来说,一种或多种或所有或给定类型的核苷酸(例如,嘌呤或嘧啶,或A、G、U、C中的任何一者或多者或全部)可以在本公开的核酸中或在其预定序列区域中(例如,在包括或不包括多聚A尾的mRNA中)被统一修饰。在一些实施方案中,本公开的核酸中(或其序列区域中)的所有核苷酸X都是修饰的核苷酸,其中X可以是核苷酸A、G、U、C中的任一者,或组合A+G、A+U、A+C、G+U、G+C、U+C、A+G+U、A+G+C、G+U+C或A+G+C中的任一者。

[0588] 核酸可以含有约1%至约100%的修饰核苷酸(相对于总核苷酸含量,或相对于一种或多种类型的核苷酸,即,A、G、U或C中的任何一者或多者)或任何中间百分比(例如,1%至20%、1%至25%、1%至50%、1%至60%、1%至70%、1%至80%、1%至90%、1%至95%、10%至20%、10%至25%、10%至50%、10%至60%、10%至70%、10%至80%、10%至90%、10%至95%、10%至100%、20%至25%、20%至50%、20%至60%、20%至70%、20%至80%、20%至90%、20%至95%、20%至100%、50%至60%、50%至70%、50%至80%、50%至90%、50%至95%、50%至100%、70%至80%、70%至90%、70%至95%、70%至100%、80%至90%、80%至95%、80%至100%、90%至95%、90%至100%及95%至100%)。将了解,任何剩余百分比均由未修饰的A、G、U或C的存在来解释。

[0589] 核酸可以含有最少1%和最多100%的修饰核苷酸,或任何中间百分比,诸如至少5%的修饰核苷酸、至少10%的修饰核苷酸、至少25%的修饰核苷酸、至少50%的修饰核苷酸、至少80%的修饰核苷酸或至少90%的修饰核苷酸。举例来说,核酸可以含有修饰的嘧啶,例如修饰的尿嘧啶或胞嘧啶。在一些实施方案中,核酸中至少5%、至少10%、至少25%、至少50%、至少80%、至少90%或100%的尿嘧啶被修饰的尿嘧啶(例如,5-取代的尿嘧啶)置换。修饰的尿嘧啶可以被具有单一独特结构的化合物置换,或者可以被具有不同结构(例如,2个、3个、4个或更多个独特结构)的多种化合物置换。在一些实施方案中,核酸中至少5%、至少10%、至少25%、至少50%、至少80%、至少90%或100%的胞嘧啶被修饰的胞嘧啶(例如,5-取代的胞嘧啶)置换。修饰的胞嘧啶可以被具有单一独特结构的化合物置换,或者可以被具有不同结构(例如,2个、3个、4个或更多个独特结构)的多种化合物置换。

[0590] 14.序列优化及其方法

[0591] 在一些实施方案中,本公开的多核苷酸包含编码本文所公开的多肽的序列优化的核苷酸序列,例如,编码治疗性有效载荷或预防性有效载荷的多核苷酸。在一些实施方案中,本公开的多核苷酸包含编码治疗性有效载荷或预防性有效载荷的开放阅读框(ORF),其中ORF已经进行序列优化。

[0592] 本文所公开的序列优化的核苷酸序列不同于相应的野生型核苷酸序列和其它已知的序列优化的核苷酸序列,例如,这些序列优化的核酸具有独特的组成特征。

[0593] 在一些实施方案中,相对于参考野生型核苷酸序列中的尿嘧啶或胸腺嘧啶核碱基的百分比,序列优化的核苷酸序列中的尿嘧啶或胸腺嘧啶核碱基的百分比被修饰(例如,降

低)。此种序列被称为尿嘧啶修饰的或胸腺嘧啶修饰的序列。核苷酸序列中尿嘧啶或胸腺嘧啶含量的百分比可以通过将序列中尿嘧啶或胸腺嘧啶的数目除以核苷酸总数并乘以100来确定。在一些实施方案中,序列优化的核苷酸序列具有比参考野生型序列中的尿嘧啶或胸腺嘧啶含量更低的尿嘧啶或胸腺嘧啶含量。在一些实施方案中,本公开的序列优化的核苷酸序列中的尿嘧啶或胸腺嘧啶含量大于参考野生型序列中的尿嘧啶或胸腺嘧啶含量并且仍维持有益作用,例如,与参考野生型序列相比时增加的表达和/或信号响应。

[0594] 在一些实施方案中,本公开的优化序列在序列中含有独特范围的尿嘧啶或胸腺嘧啶(如果是DNA的话)。优化序列的尿嘧啶或胸腺嘧啶含量可以按多种方式表示,例如,相对于理论最小值(UTM%或TTM%)、相对于野生型(UWT%或TWT%)以及相对于总核苷酸含量(UTL%或TTL%)的优化序列的尿嘧啶或胸腺嘧啶含量。对于DNA,认识到胸腺嘧啶(T)代替尿嘧啶(U)存在,并且在U出现之处用T取代。对于RNA,认识到尿嘧啶(U)代替胸腺嘧啶(T)存在。当通过将DNA序列中的胸腺嘧啶替换为尿嘧啶来提供DNA序列时,本领域技术人员可以容易地获得RNA序列。因此,与例如UTM%、UWT%或UTL%(关于RNA)有关的所有公开内容同样适用于TTM%、TWT%或TTL%(关于DNA)。

[0595] 相对于尿嘧啶或胸腺嘧啶理论最小值的尿嘧啶或胸腺嘧啶含量是指通过将序列优化的核苷酸序列中的尿嘧啶或胸腺嘧啶的数目除以假设核苷酸序列中的尿嘧啶或胸腺嘧啶的总数并乘以100而确定的参数,其中假设序列中的所有密码子都被具有尽可能低的尿嘧啶或胸腺嘧啶含量的同义密码子置换。这个参数在本文中缩写为UTM%或TTM%。

[0596] 在一些实施方案中,本公开的尿嘧啶修饰的序列相对于相应的野生型核酸序列具有数目减少的连续尿嘧啶。举例来说,两个连续亮氨酸可以由序列CUUUUG编码,所述序列包括四尿嘧啶簇。此种子序列可以例如用CUGCUC替代,所述序列去除尿嘧啶簇。苯丙氨酸可以由UUC或UUU编码。因此,即使由UUU编码的苯丙氨酸被UUC替代,同义密码子仍然含有尿嘧啶对(UU)。因此,序列中苯丙氨酸的数目确立了在不改变编码的多肽中苯丙氨酸数目的情况下不能消除的尿嘧啶对(UU)的最小数目。

[0597] 在一些实施方案中,本公开的尿嘧啶修饰的序列相对于野生型核酸序列具有数目减少的尿嘧啶三联体(UUU)。在一些实施方案中,相对于野生型核酸序列中尿嘧啶对(UU)的数目,尿嘧啶修饰的序列具有数目减少的尿嘧啶对(UU)。在一些实施方案中,本公开的尿嘧啶修饰的序列具有对应于野生型核酸序列中的最小可能数目的尿嘧啶对(UU)的尿嘧啶对(UU)的数目。

[0598] 短语“相对于野生型核酸序列中的尿嘧啶对(UU)的尿嘧啶对(UU)”是指通过将序列优化的核苷酸序列中尿嘧啶对(UU)的数目除以相应野生型核苷酸序列中尿嘧啶对(UU)的总数并乘以100而确定的参数。这个参数在本文中缩写为UUwt%。在一些实施方案中,尿嘧啶修饰的序列具有低于100%的UUwt%。

[0599] 在一些实施方案中,本公开的多核苷酸包含尿嘧啶修饰的序列。在一些实施方案中,尿嘧啶修饰的序列包含至少一个化学修饰的核碱基,例如,5-甲氧基尿嘧啶。在一些实施方案中,本公开的尿嘧啶修饰的序列中至少95%的核碱基(例如,尿嘧啶)是修饰的核碱基。在一些实施方案中,尿嘧啶修饰的序列中至少95%的尿嘧啶是5-甲氧基尿嘧啶。

[0600] 在一些实施方案中,本公开的多核苷酸是序列优化的。

[0601] 序列优化的核苷酸序列(核苷酸序列在本文中也称为“核酸”)包含相对于参考序

列(例如,编码治疗性有效载荷或预防性有效载荷的野生型序列)的至少一个密码子修饰。因此,在序列优化的核酸中,至少一个密码子不同于参考序列(例如,野生型序列)中的相应密码子。

[0602] 一般来说,序列优化的核酸通过至少一个步骤产生,所述步骤包括用同义密码子(即,编码相同氨基酸的密码子)取代参考序列中的密码子。举例来说,可以通过应用密码子取代图(即,提供将编码密码子优化序列中的每个氨基酸的密码子的表)或通过应用一组规则(例如,如果甘氨酸紧接在中性氨基酸之后,则甘氨酸将由某个密码子编码,但如果紧接在极性氨基酸之后,则它将由另一个密码子编码)来实现此类取代。除了密码子取代(即,“密码子优化”)之外,本文所公开的序列优化方法包括不严格针对密码子优化的额外优化步骤,诸如去除有害基序(去稳定基序取代)。可以向有需要的受试者施用包含这些序列优化的核酸(例如RNA,例如mRNA)的组合物和制剂,以促进编码治疗性有效载荷或预防性有效载荷的功能活性物质的体内表达。

[0603] 额外和示例性序列优化方法在2017年5月18日提交的国际PCT申请WO 2017/201325中公开,所述申请的全部内容特此以引用的方式并入。

[0604] 15.脂质纳米颗粒的脂质含量

[0605] 如上文所阐述,关于脂质,用作本文所公开的递送运载体的LNP包含(i)可电离脂质;(ii)固醇或其它结构脂质;(iii)非阳离子辅助脂质或磷脂;以及任选地(iv)PEG脂质。这些类别的脂质在下文更详细地阐述。

[0606] 在一些实施方案中,本发明的核酸被配制成脂质纳米颗粒(LNP)组合物。脂质纳米颗粒通常包含氨基脂质、磷脂、结构脂质和PEG脂质组分以及所关注的核酸货物。本发明的脂质纳米颗粒可以使用本领域中一般已知的组分、组合物和方法产生,参见例如PCT/US2016/052352;PCT/US2016/068300;PCT/US2017/037551;PCT/US2015/027400;PCT/US2016/047406;PCT/US2016000129;PCT/US2016/014280;PCT/US2016/014280;PCT/US2017/038426;PCT/US2014/027077;PCT/US2014/055394;PCT/US2016/52117;PCT/US2012/069610;PCT/US2017/027492;PCT/US2016/059575;PCT/US2016/069491;PCT/US2016/069493;和PCT/US2014/66242,全部以全文引用的方式并入本文中。

[0607] 在一些实施方案中,脂质纳米颗粒包含相对于其它脂质组分,摩尔比为20-60%的氨基脂质。举例来说,脂质纳米颗粒可以包含摩尔比为20-50%、20-40%、20-30%、30-60%、30-50%、30-40%、40-60%、40-50%或50-60%的氨基脂质。在一些实施方案中,脂质纳米颗粒包含摩尔比为20%、30%、40%、50或60%的氨基脂质。

[0608] 在一些实施方案中,脂质纳米颗粒包含相对于其它脂质组分,摩尔比为5-25%的磷脂。举例来说,脂质纳米颗粒可以包含摩尔比为5-30%、5-15%、5-10%、10-25%、10-20%、10-25%、15-25%、15-20%、20-25%或25-30%的磷脂。在一些实施方案中,脂质纳米颗粒包含摩尔比为5%、10%、15%、20%、25%或30%的非阳离子脂质。

[0609] 在一些实施方案中,脂质纳米颗粒包含相对于其它脂质组分,摩尔比为25-55%的结构脂质。举例来说,脂质纳米颗粒可以包含摩尔比为10-55%、25-50%、25-45%、25-40%、25-35%、25-30%、30-55%、30-50%、30-45%、30-40%、30-35%、35-55%、35-50%、35-45%、35-40%、40-55%、40-50%、40-45%、45-55%、45-50%或50-55%的结构脂质。在一些实施方案中,脂质纳米颗粒包含摩尔比为10%、15%、20%、25%、30%、35%、40%、

45%、50%或55%的结构脂质。

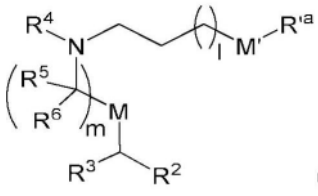
[0610] 在一些实施方案中,脂质纳米颗粒包含相对于其它脂质组分,摩尔比为0.5-15%的PEG脂质。举例来说,脂质纳米颗粒可以包含摩尔比为0.5-10%、0.5-5%、1-15%、1-10%、1-5%、2-15%、2-10%、2-5%、5-15%、5-10%或10-15%的PEG脂质。在一些实施方案中,脂质纳米颗粒包含摩尔比为0.5%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%或15%的PEG-脂质。

[0611] 在一些实施方案中,脂质纳米颗粒包含摩尔比为20-60%的氨基脂质、5-25%的磷脂、25-55%的结构脂质和0.5-15%的PEG脂质。

[0612] 在一些实施方案中,脂质纳米颗粒包含摩尔比为20-60%的氨基脂质、5-30%的磷脂、10-55%的结构脂质和0.5-15%的PEG脂质。

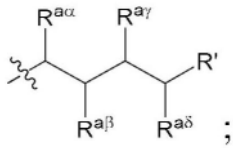
[0613] 可电离氨基脂质

[0614] 在一些方面,本公开涉及式(I)化合物:

[0615]  或其N-氧化物,或其盐或异构体,

(I),

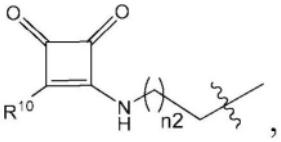
[0616] 其中R'a是R′支链;其中

[0617] R′支链是: 其中~表示连接点;

[0618] 其中Raα、Raβ、Ray和Raδ各自独立地选自由H、C2-12烷基和C2-12烯基组成的组;

[0619] R2和R3各自独立地选自由C1-14烷基和C2-14烯基组成的组;

[0620] R4选自由以下组成的组:-(CH2)nOH,其中n选自由1、2、3、4和5组成的组,和

,

[0621] 其中~表示连接点;其中

[0622] R10是N(R)2;每个R独立地选自由C1-6烷基、C2-3烯基和H组成的组;并且n2选自由1、2、3、4、5、6、7、8、9和10组成的组;

[0623] 每个R5独立地选自由C1-3烷基、C2-3烯基和H组成的组;

[0624] 每个R6独立地选自由C1-3烷基、C2-3烯基和H组成的组;

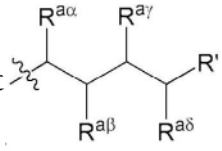
[0625] M和M′各自独立地选自由-C(O)O-和-OC(O)-组成的组;

[0626] R′是C1-12烷基或C2-12烯基;

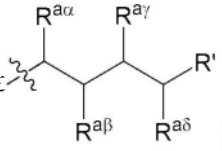
[0627] 1选自由1、2、3、4和5组成的组;并且

[0628] m选自由5、6、7、8、9、10、11、12和13组成的组。

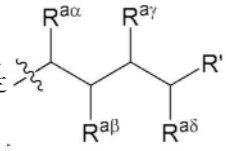
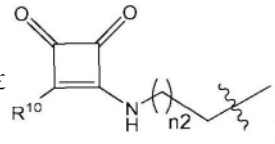
[0629] 在式(I)化合物的一些实施方案中,R′a是R′支链;

[0630] R′支链是表示连接点;Raα、Raβ、Ra γ 和Raδ各自是H;R2和R3各自是C1-14烷基;R4是-(CH2)nOH;n是2;每个R5是H;每个R6是H;M和M′各自是-C(O)O-;R′是C1-12烷基;1是5;并且m是7。

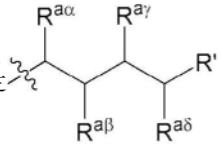
[0631] 在式(I)化合物的一些实施方案中,R′a是R′支链;

[0632] R′支链是表示连接点;Raα、Raβ、Ray和Raδ各自是H;R2和R3各自是C1-14烷基;R4是-(CH2)nOH;n是2;每个R5是H;每个R6是H;M和M′各自是-C(O)O-;R′是C1-12烷基;1是3;并且m是7。

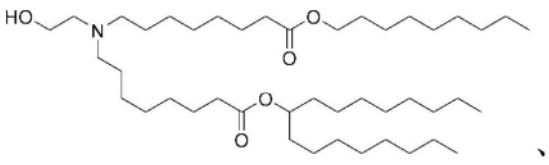
[0633] 在式(I)化合物的一些实施方案中,R′a是R′支链;

[0634] R′支链是表示连接点;Raα是C2-12烷基;Raβ、Ra γ 和Raδ各自是H;R2和R3各自是C1-14烷基;R4是R10是NH(C1-6烷基);n2是2;R5是H;每个R6是H;M和M′各自是-C(O)O-;R′是C1-12烷基;1是5;并且m是7。

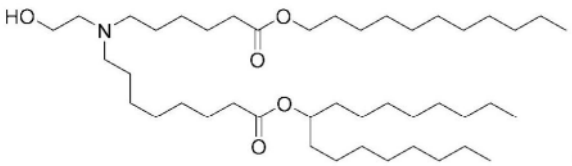
[0635] 在式(I)化合物的一些实施方案中,R′a是R′支链;

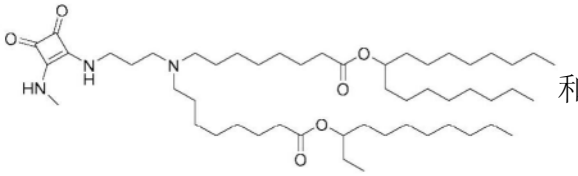
[0636] R′支链是表示连接点;Raα、Raβ和Raδ各自是H;Ra γ 是C2-12烷基;R2和R3各自是C1-14烷基;R4是-(CH2)nOH;n是2;每个R5是H;每个R6是H;M和M′各自是-C(O)O-;R′是C1-12烷基;1是5;并且m是7。

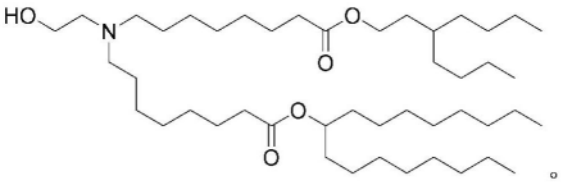
[0637] 在一些实施方案中,式(I)化合物选自:

、

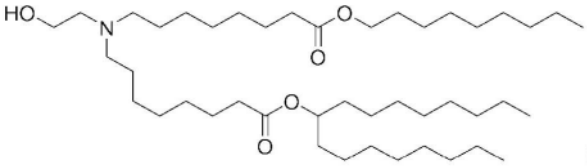
[0638]

、

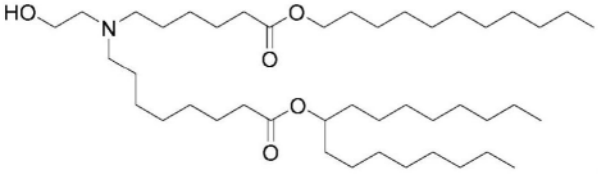
[0639]  和

[0640] 。

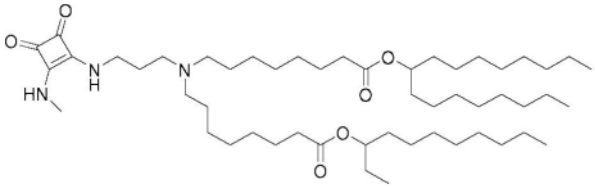
[0641] 在一些实施方案中,式(I)化合物是:

[0642] (化合物 II)。

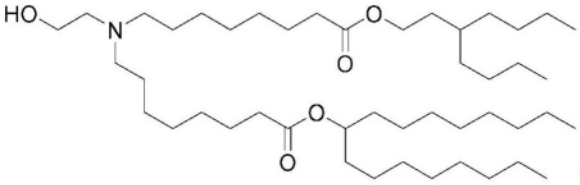
[0643] 在一些实施方案中,式(I)化合物是:

[0644] 。

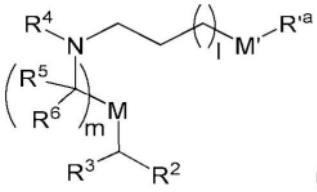
[0645] 在一些实施方案中,式(I)化合物是:

[0646] 。

[0647] 在一些实施方案中,式(I)化合物是:

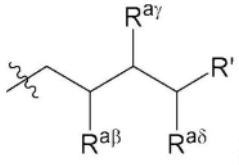
[0648] (化合物 B)。

[0649] 在一些方面,本公开涉及式(Ia)化合物:

[0650]  或其N-氧化物,或其盐或异构体,其中R'a是R′支链;

(Ia),

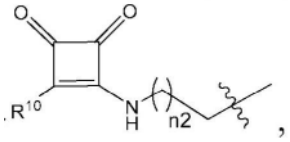
其中

[0651] R′支链是： 其中—表示连接点；

[0652] 其中Raβ、Ray和Raδ各自独立地选自由C2-12烷基和C2-12烯基组成的组；

[0653] R2和R3各自独立地选自由C1-14烷基和C2-14烯基组成的组；

[0654] R4选自由以下组成的组：-(CH2)nOH，其中n选自由1、2、3、4和5组成的组，和



[0655] 其中—表示连接点；其中

[0656] R10是N(R)2；每个R独立地选自由C1-6烷基、C2-3烯基和H组成的组；并且n2选自由1、2、3、4、5、6、7、8、9和10组成的组；

[0657] 每个R5独立地选自由C1-3烷基、C2-3烯基和II组成的组；

[0658] 每个R6独立地选自由C1-3烷基、C2-3烯基和H组成的组；

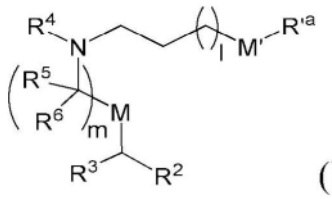
[0659] M和M′各自独立地选自由-C(O)O-和-OC(O)-组成的组；

[0660] R′是C1-12烷基或C2-12烯基；

[0661] l选自由1、2、3、4和5组成的组；并且

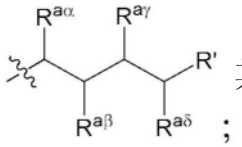
[0662] m选自由5、6、7、8、9、10、11、12和13组成的组。

[0663] 在一些方面，本公开涉及式(Ib)化合物：

[0664]  或其N-氧化物，或其盐或异构体，

(Ib)，

[0665] 其中R'a是R′支链；其中

[0666] R′支链是： 其中—表示连接点；

[0667] 其中Raα、Raβ、Ray和Raδ各自独立地选自由H、C2-12烷基和C2-12烯基组成的组；

[0668] R2和R3各自独立地选自由C1-14烷基和C2-14烯基组成的组；

[0669] R4是-(CH2)nOH，其中n选自由1、2、3、4和5组成的组；

[0670] 每个R5独立地选自由C1-3烷基、C2-3烯基和H组成的组；

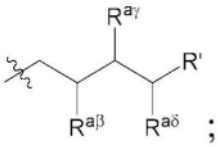
[0671] 每个R6独立地选自由C1-3烷基、C2-3烯基和H组成的组；

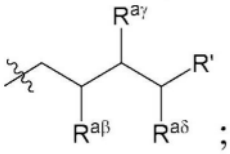
[0672] M和M′各自独立地选自由-C(O)O-和-OC(O)-组成的组；

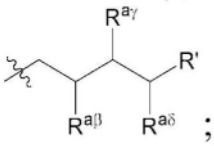
[0673] R′是C1-12烷基或C2-12烯基；

[0674] l选自由1、2、3、4和5组成的组；并且

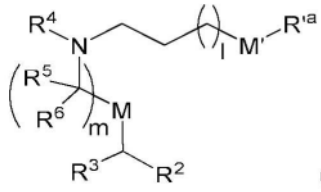
[0675] m选自由5、6、7、8、9、10、11、12和13组成的组。

[0676] 在式(I)或式(Ib)的一些实施方案中,R′a是R′支链;R′支链是表示连接点;Raβ、Raγ和Raδ各自是H;R2和R3各自是C1-14烷基;R4是-(CH2)nOH;n是2;每个R5是H;每个R6是H;M和M′各自是-C(O)O-;R′是C1-12烷基;l是5;并且m是7。

[0677] 在式(I)或式(Ib)的一些实施方案中,R′a是R′支链;R′支链是表示连接点;Raβ、Ray和Raδ各自是H;R2和R3各自是C1-14烷基;R4是-(CH2)nOH;n是2;每个R5是H;每个R6是H;M和M′各自是-C(O)O-;R′是C1-12烷基;l是3;并且m是7。

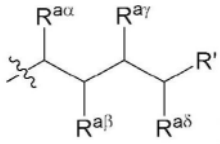
[0678] 在式(I)或式(Ib)的一些实施方案中,R′a是R′支链;R′支链是表示连接点;Raβ和Raδ各自是H;Ray是C2-12烷基;R2和R3各自是C1-14烷基;R4是-(CH2)nOH;n是2;每个R5是H;每个R6是H;M和M′各自是-C(O)O-;R′是C1-12烷基;l是5;并且m是7。

[0679] 在一些方面,本公开涉及式(Ic)化合物:

[0680] 或其N-氧化物,或其盐或异构体,

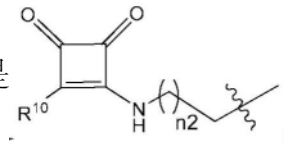
(Ic),

[0681] 其中R'a是R′支链;其中

[0682] R′支链是:其中表示连接点;

[0683] 其中Raα、Raβ、Raγ和Raδ各自独立地选自由H、C2-12烷基和C2-12烯基组成的组;

[0684] R2和R3各自独立地选自由C1-14烷基和C2-14烯基组成的组;

[0685] R4是,

[0686] 其中表示连接点;其中

[0687] R10是N(R)2;每个R独立地选自由C1-6烷基、C2-3烯基和H组成的组;n2选自由1、2、3、4、5、6、7、8、9和10组成的组;

[0688] 每个R5独立地选自由C1-3烷基、C2-3烯基和H组成的组;

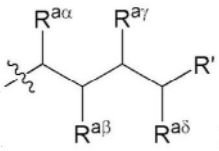
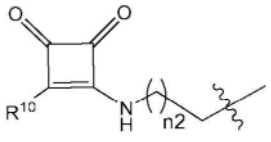
[0689] 每个R6独立地选自由C1-3烷基、C2-3烯基和H组成的组；

[0690] M和M′各自独立地选自由-C(O)O-和-OC(O)-组成的组；

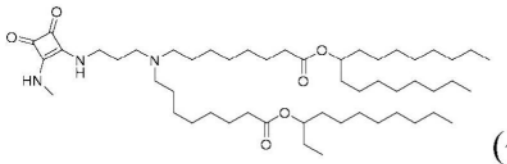
[0691] R′是C1-12烷基或C2-12烯基；

[0692] 1选自由1、2、3、4和5组成的组；并且

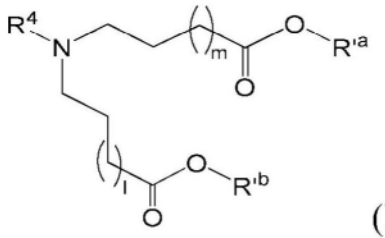
[0693] m选自由5、6、7、8、9、10、11、12和13组成的组。

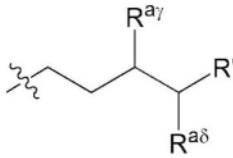
[0694] 在一些实施方案中，R'a是R′支链；R′支链是；表示连接点；Raβ、Ray和Raδ各自是H；Raα是C2-12烷基；R2和R3各自是C1-14烷基；R4是；表示连接点；R10是NH(C1-6烷基)；n2是2；每个R5是H；每个R6是H；M和M′各自是-C(O)O-；R′是C1-12烷基；1是5；并且m是7。

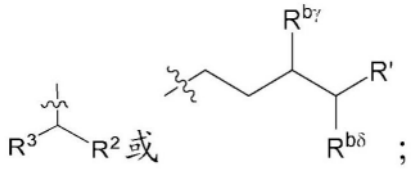
[0695] 在一些实施方案中，式(Ic)化合物是：

[0696] (化合物 A)。

[0697] 在一些方面，本公开涉及式(II)化合物：

[0698] 或其N-氧化物，或其盐或异构体，其中R′a是R′支链或R′环状；其中

[0699] R′支链是：并且R′环状是：；并且

[0700] R′b是：；

[0701] 其中表示连接点；

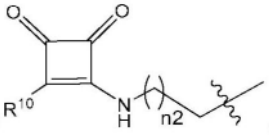
[0702] Raγ和Raδ各自独立地选自由H、C1-12烷基和C2-12烯基组成的组，其中Raγ和Raδ中的至少一者选自由C1-12烷基和C2-12烯基组成的组；

[0703] Rbγ和Rbδ各自独立地选自由H、C1-12烷基和C2-12烯基组成的组，其中Rbγ和Rbδ

中的至少一者选自由C1-12烷基和C2-12烯基组成的组;

[0704] R2和R3各自独立地选自由C1-14烷基和C2-14烯基组成的组;

[0705] R4选自由以下组成的组:-(CH2)nOH,其中n选自由1、2、3、4和5组成的组,和



,

[0706] 其中表示连接点;其中

[0707] R10是N(R)2;每个R独立地选自由C1-6烷基、C2-3烯基和H组成的组;并且n2选自由1、2、3、4、5、6、7、8、9和10组成的组;

[0708] 每个R′独立地是C1-12烷基或C2-12烯基;

[0709] Ya是C3-6碳环;

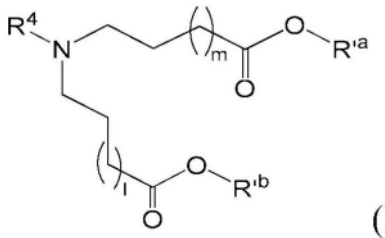
[0710] R*″a选自由C1-15烷基和C2-15烯基组成的组;并且

[0711] s是2或3;

[0712] m选自1、2、3、4、5、6、7、8和9;

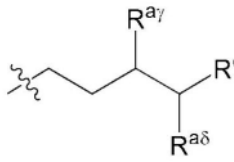
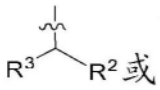
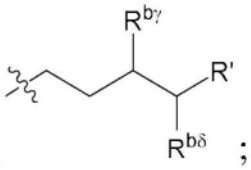
[0713] 1选自1、2、3、4、5、6、7、8和9。

[0714] 在一些方面,本公开涉及式(II-a)化合物:

[0715] 

(II-a),

或其N-氧化物,或其盐或异构体,其中R'a是R′支链或R′环状;其中

[0716] R′支链是:并且R'b是:或;

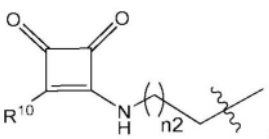
[0717] 其中表示连接点;

[0718] Ra γ和Raδ各自独立地选自由H、C1-12烷基和C2-12烯基组成的组,其中Ray和Raδ中的至少一者选自由C1-12烷基和C2-12烯基组成的组;

[0719] Rb γ和Rbδ各自独立地选自由H、C1-12烷基和C2-12烯基组成的组,其中Rb γ和Rbδ中的至少一者选自由C1-12烷基和C2-12烯基组成的组;

[0720] R2和R3各自独立地选自由C1-14烷基和C2-14烯基组成的组;

[0721] R4选自由以下组成的组:-(CH2)nOH,其中n选自由1、2、3、4和5组成的组,和



,

[0722] 其中~表示连接点;其中

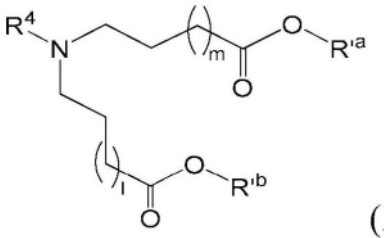
[0723] R70是N(R)2;每个R独立地选自由C1-6烷基、C2-3烯基和H组成的组;并且n2选自由1、2、3、4、5、6、7、8、9和10组成的组;

[0724] 每个R′独立地是C1-12烷基或C2-12烯基;

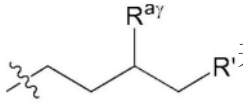
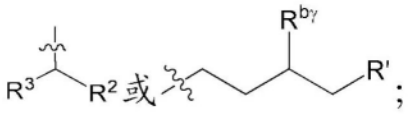
[0725] m选自1、2、3、4、5、6、7、8和9;

[0726] 1选自1、2、3、4、5、6、7、8和9。

[0727] 在一些方面,本公开涉及式(II-b)化合物:

[0728]  (II-b),

或其N-氧化物,或其盐或异构体,其中R'a是R′支链或R′环状;其中

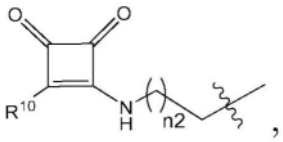
[0729] R′支链是:并且R′b是:;

[0730] 其中~表示连接点;

[0731] Ra γ和Rb γ各自独立地选自由C1-12烷基和C2-12烯基组成的组;

[0732] R2和R3各自独立地选自由C1-14烷基和C2-14烯基组成的组;

[0733] R4选自由以下组成的组:-(CH2)nOH,其中n选自由1、2、3、4和5组成的组,和

,

[0734] 其中~表示连接点;其中

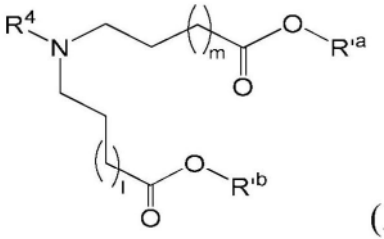
[0735] R10是N(R)2;每个R独立地选自由C1-6烷基、C2-3烯基和H组成的组;并且n2选自由1、2、3、4、5、6、7、8、9和10组成的组;

[0736] 每个R′独立地是C1-12烷基或C2-12烯基;

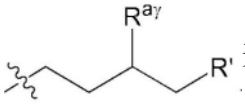
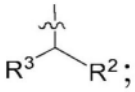
[0737] m选自1、2、3、4、5、6、7、8和9;

[0738] 1选自1、2、3、4、5、6、7、8和9。

[0739] 在一些方面,本公开涉及式(II-c)化合物:

[0740]  (II-c),

或其N-氧化物,或其盐或异构体,其中R'a是R′支链或R′环状;其中

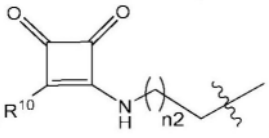
[0741] R′支链是：并且R′b是：；

[0742] 其中表示连接点；

[0743] 其中RaT选自由C1-12烷基和C2-12烯基组成的组；

[0744] R2和R3各自独立地选自由C1-14烷基和C2-14烯基组成的组；

[0745] R4选自由以下组成的组：-(CH2)nOH，其中n选自由1、2、3、4和5组成的组，和

，

[0746] 其中表示连接点；其中

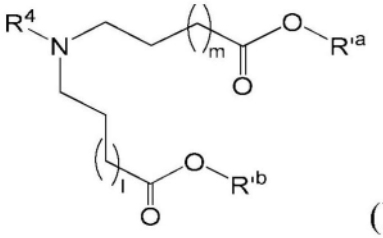
[0747] R10是N(R)2；每个R独立地选自由C1-6烷基、C2-3烯基和H组成的组；并且n2选自由1、2、3、4、5、6、7、8、9和10组成的组；

[0748] R′是C1-12烷基或C2-12烯基；

[0749] m选自1、2、3、4、5、6、7、8和9；

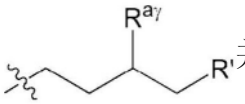
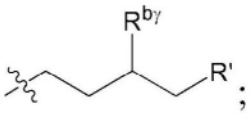
[0750] l选自1、2、3、4、5、6、7、8和9。

[0751] 在一些方面，本公开涉及式(II-d)化合物：

[0752] 或其N-氧化物，或其盐或异构体，其中R′a是R′

(II-d)，

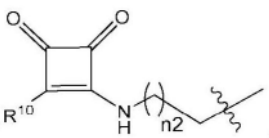
支链或R′环状；其中

[0753] R′支链是：并且R′b是：；

[0754] 其中表示连接点；

[0755] 其中Ra γ和Rb γ各自独立地选自由C1-12烷基和C2-12烯基组成的组；

[0756] R4选自由以下组成的组：-(CH2)nOH，其中n选自由1、2、3、4和5组成的组，和

，

[0757] 其中表示连接点；其中

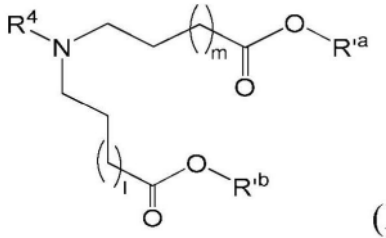
[0758] R10是N(R)2；每个R独立地选自由C1-6烷基、C2-3烯基和H组成的组；并且n2选自由1、2、3、4、5、6、7、8、9和10组成的组；

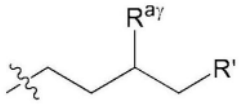
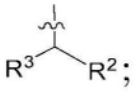
[0759] 每个R′独立地是C1-12烷基或C2-12烯基；

[0760] m选自1、2、3、4、5、6、7、8和9；

[0761] 1选自1、2、3、4、5、6、7、8和9。

[0762] 在一些方面，本公开涉及式(II-e)化合物：

[0763]  (II-e)，或其N-氧化物，或其盐或异构体，其中R′a是R′支链或R′环状；其中

[0764] R′支链是：并且R′b是：；

[0765] 其中表示连接点；

[0766] 其中RaT选自由C1-12烷基和C2-12烯基组成的组；

[0767] R2和R3各自独立地选自由C1-14烷基和C2-14烯基组成的组；

[0768] R4是-(CH2)nOH，其中n选自由1、2、3、4和5组成的组；

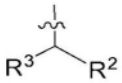
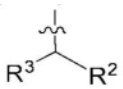
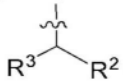
[0769] R′是C1-12烷基或C2-12烯基；

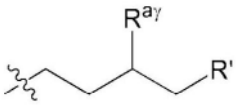
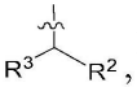
[0770] m选自1、2、3、4、5、6、7、8和9；

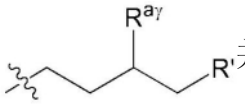
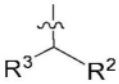
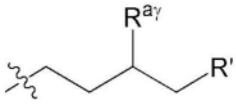
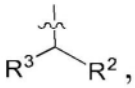
[0771] 1选自1、2、3、4、5、6、7、8和9。

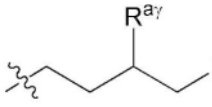
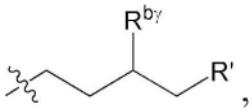
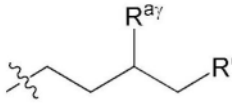
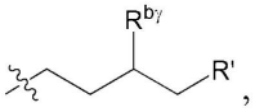
[0772] 在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中，m和1各自独立地选自4、5和6。在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中，m和1各自是5。

[0773] 在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中，每个R′独立地是C1-12烷基。在式(II)、式(II-a)、式(II-b)、式(TT-c)、式(II-d)或式(II-e)化合物的一些实施方案中，每个R′独立地是C2-5烷基。

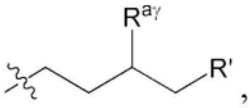
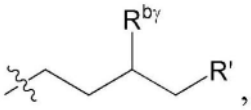
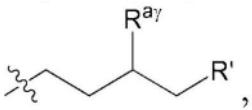
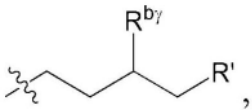
[0774] 在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中，R′b是：并且R2和R3各自独立地是C1-14烷基。在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中，R′b是：并且R2和R3各自独立地是C6-10烷基。在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中，R′b是：并且R2和R3各自是C8烷基。

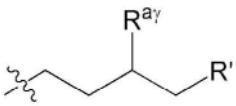
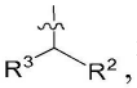
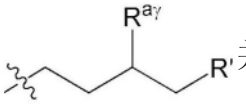
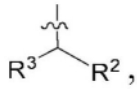
[0775] 在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中，R′支链是：并且R′b是：，Ra γ是C1-12烷基并且R2和R3各自独立地是C6-10烷基。在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)

化合物的一些实施方案中,R′支链是:并且R'b是:,Ra γ是C2-6烷基并且R2和R3各自独立地是C6-10烷基。在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R′支链是:并且R'b是:,Ra γ是C2-6烷基,并且R2和R3各自是C8烷基。

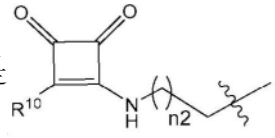
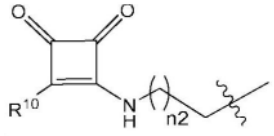
[0776] 在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R′支链是:,R′b是:,并且Ra γ和Rb γ各自是C1-12烷基。在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R′支链是:,R'b是:,并且Ray和Rb γ各自是C2-6烷基。

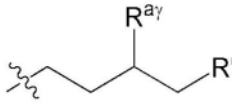
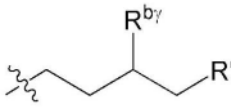
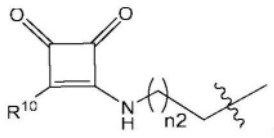
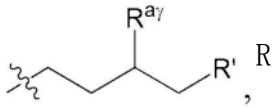
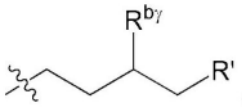
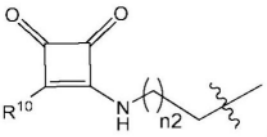
[0777] 在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,m和l各自独立地选自4、5和6,并且每个R′独立地是C1-12烷基。在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,m和l各自是5并且每个R′独立地是C2-5烷基。

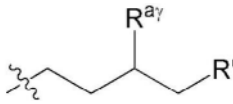
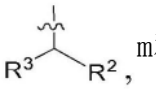
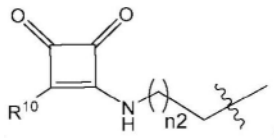
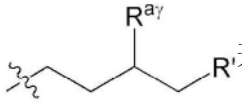
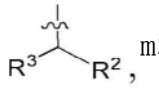
[0778] 在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R′支链是:,R′b是:,m和l各自独立地选自4、5和6,每个R′独立地是C1-12烷基,并且Ra γ和Rb γ各自是C1-12烷基。在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R′支链是:,R'b是:,m和l各自是5,每个R′独立地是C2-5烷基,并且Ray和Rb γ各自是C2-6烷基。

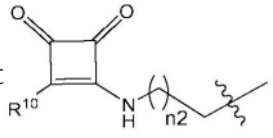
[0779] 在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R′支链是:并且R'b是:,m和l各自独立地选自4、5和6,R′是C1-12烷基,Ra γ是C1-12烷基并且R2和R3各自独立地是C6-10烷基。在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R′支链是:并且R'b是:,m和l各自是5,R′是C2-5烷基,Ra γ是C2-6烷基,并且R2和R3各自是C8烷基。

[0780] 在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实

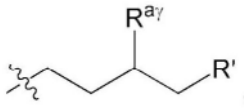
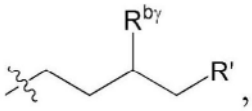
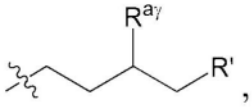
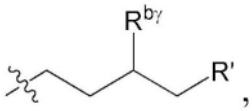
施方案中,R4是其中R10是NH(C1-6烷基)并且n2是2。在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R4是其中R10是NH(CH3)并且n2是2。

[0781] 在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R′支链是:R′b是:m和l各自独立地选自4、5和6,每个R′独立地是C1-12烷基,Raγ和Rbγ各自是C1-12烷基,并且R4是其中R10是NH(C1-6烷基),并且n2是2。在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R′支链是:R′b是:m和l各自是5,每个R′独立地是C2-5烷基,Ray和Rbγ各自是C2-6烷基,并且R4是其中R10是NH(CH3)并且n2是2。

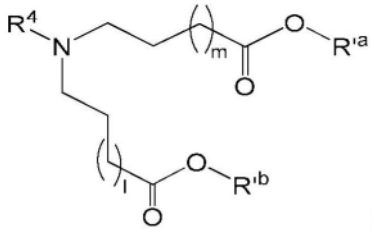
[0782] 在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R′支链是:并且R'b是:m和l各自独立地选自4、5和6,R′是C1-12烷基,R2和R3各自独立地是C6-10烷基,Ray是C1-12烷基,并且R4是其中R10是NH(C1-6烷基)并且n2是2。在式(II)、式(Il-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R′支链是:并且R'b是:m和l各自是5,R′是C2-5烷基,Ray是C2-6烷基,R2和R3各自是C8烷基,并

且R4是其中R10是NH(CH3)并且n2是2。

[0783] 在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R4是-(CH2)nOH并且n是2、3或4。在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R4是-(CH2)nOH并且n是2。

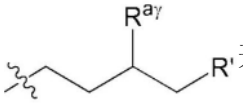
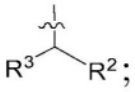
[0784] 在式(II)、式(TI-a)、式(TT-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R′支链是:,R′b是:,m和1各自独立地选自4、5和6,每个R′独立地是C1-12烷基,Ra γ和Rb γ各自是C1-12烷基,R4是-(CH2)nOH,并且n是2、3或4。在式(II)、式(II-a)、式(II-b)、式(II-c)、式(II-d)或式(II-e)化合物的一些实施方案中,R′支链是:,R′b是:,m和1各自是5,每个R′独立地是C2-5烷基,Ray和Rb γ各自是C2-6烷基,R4是-(CH2)nOH,并且n是2。

[0785] 在一些方面,本公开涉及式(II-f)化合物:

[0786] 

(II-f),

或其N-氧化物,或其盐或异构体,其中R'a是R′支链或R′环状;其中

[0787] R′支链是:并且R'b是:;

[0788] 其中表示连接点;

[0789] Ra γ是C1-12烷基;

[0790] R2和R3各自独立地是C1-14烷基;

[0791] R4是-(CH2)nOH,其中n选自由1、2、3、4和5组成的组;

[0792] R′是C1-12烷基;

[0793] m选自4、5和6;并且

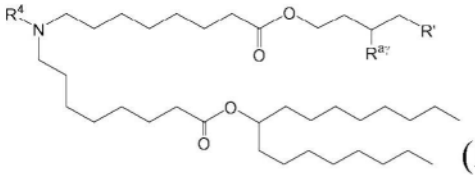
[0794] 1选自4、5和6。

[0795] 在式(II-f)化合物的一些实施方案中,m和1各自是5,并且n是2、3或4。

[0796] 在式(II-f)化合物的一些实施方案中,R′是C2-5烷基,Ray是C2-6烷基,并且R2和R3各自是C6-10烷基。

[0797] 在式(II-f)化合物的一些实施方案中,m和1各自是5,并且n是2、3或4,R′是C2-5烷基,Ray是C2-6烷基,并且R2和R3各自是C6-10烷基。

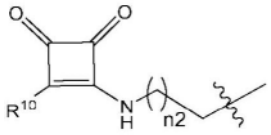
[0798] 在一些方面,本公开涉及式(II-g)化合物:

[0799]  其中 (II-g),

[0800] Ray是C2-6烷基;

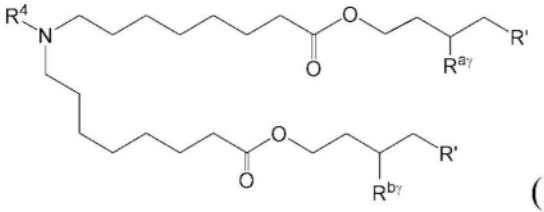
[0801] R′是C2-5烷基;并且

[0802] R4选自由以下组成的组:-(CH2)nOH,其中n选自由3、4和5组成的组,和

,

[0803] 其中~表示连接点,R10是NH(C1-6烷基),并且n2选自由1、2和3组成的组。

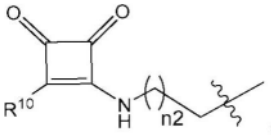
[0804] 在一些方面,本公开涉及式(II-h)化合物:

[0805]  其中 (II-h),

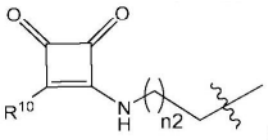
[0806] Ray和Rby独立地是C2-6烷基;

[0807] 每个R′独立地是C2-5烷基;并且

[0808] R4选自由以下组成的组:-(CH2)nOH,其中n选自由3、4和5组成的组,和

,

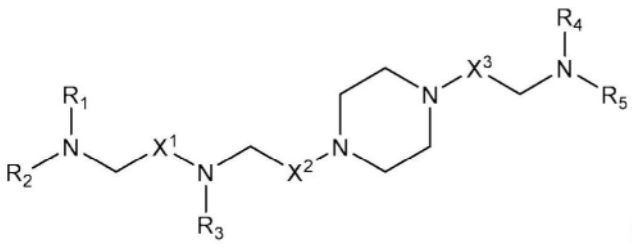
[0809] 其中~表示连接点,R10是NH(C1-6烷基),并且n2选自由1、2和3组成的组。

[0810] 在式(II-g)或式(II-h)化合物的一些实施方案中,R4是其中,

[0811] R10是NH(CH3)并且n2是2。

[0812] 在式(II-g)或式(II-h)化合物的一些实施方案中,R4是-(CH2)2OH。

[0813] 在一些方面,本公开涉及具有式(III)的化合物:

[0814]  (III),

[0815] 或其盐或异构体,其中

[0816] R1、R2、R3、R4和R5独立地选自由C5-20烷基、C5-20烯基、-R″MR′、-R*YR″、-YR″和-

R*OR″组成的组；

[0817] 每个M独立地选自由-C(O)O-、-OC(O)-、-OC(O)O-、-C(O)N(R′)-、-N(R′)C(O)-、-C(O)-、-C(S)-、-C(S)S-、-SC(S)-、-CH(OH)-、-P(O)(OR′)O-、-S(O)2-、芳基和杂芳基组成的组；

[0818] X1、X2和X3独立地选自由键、-CH2-、-(CH2)2-、-CHR-、-CHY-、-C(O)-、-C(O)O-、-OC(O)-、-C(O)-CH2-、-CH2-C(O)-、-C(O)O-CH2-、-OC(O)-CH2-、-CH2-C(O)O-、-CH2-OC(O)—、-CH(OH)-、-C(S)-和-CH(SH)-组成的组；

[0819] 每个Y独立地是C3-6碳环；

[0820] 每个R*独立地选自由C1-12烷基和C2-12烯基组成的组；

[0821] 每个R独立地选自由C1-3烷基和C3-6碳环组成的组；

[0822] 每个R′独立地选自由C1-12烷基、C2-12烯基和H组成的组；并且

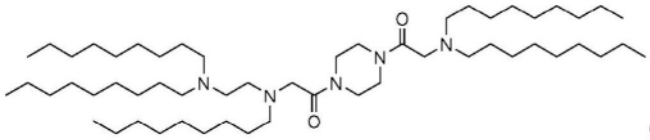
[0823] 每个R″独立地选自由C3-12烷基和C3-12烯基组成的组，并且其中：

[0824] i)X1、X2和X3中的至少一者不是-CH2-；并且/或者

[0825] ii)R1、R2、R3、R4和R5中的至少一者是-R″MR′。

[0826] 在一些实施方案中，R1、R2、R3、R4和R5各自是C5-20烷基；X1是-CH2-；并且X2和X3各自是-C(O)-。

[0827] 在一些实施方案中，式(III)化合物是：

[0828] 

(化合物 VI)，或其盐或异构体。

[0829] 磷脂

[0830] 本文所公开的脂质纳米颗粒组合物的脂质组合物可以包含一种或多种磷脂，例如，一种或多种饱和或(多)不饱和磷脂或它们的组合。一般来说，磷脂包含磷脂部分和一个或多个脂肪酸部分。

[0831] 磷脂部分可以选自例如由磷脂酰胆碱、磷脂酰乙醇胺、磷脂酰甘油、磷脂酰丝氨酸、磷脂酸、2-溶血磷脂酰胆碱和鞘磷脂组成的非限制性组。

[0832] 脂肪酸部分可以选自例如由月桂酸、肉豆蔻酸、肉豆蔻油酸、棕榈酸、棕榈油酸、硬脂酸、油酸、亚麻油酸、α-次亚麻油酸、芥子酸、植烷酸、花生酸、花生油酸、二十碳五烯酸、山嵛酸、二十二碳五烯酸和二十二碳六烯酸组成的非限制性组。

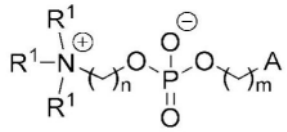
[0833] 特定磷脂可以促进与膜的融合。举例来说，阳离子磷脂可以与膜(例如，细胞膜或细胞内膜)的一种或多种带负电荷的磷脂相互作用。磷脂与膜的融合可以允许含脂质的组合物(例如，LNP)的一种或多种要素(例如，治疗剂)传递通过膜，从而允许例如将所述一种或多种要素递送至靶组织。

[0834] 还预期非天然磷脂物质，包括具有包括支链、氧化、环化和炔烃在内的修饰和取代的天然物质。举例来说，磷脂可以用一种或多种炔烃(例如，其中一个或多个双键被三键替换的烯基)官能化或与一种或多种炔烃交联。在适当反应条件下，炔烃基团可以在暴露于叠氮化物后经历铜催化的环加成。此类反应可用于官能化纳米颗粒组合物的脂质双层以促进膜渗透或细胞识别，或者可用于使纳米颗粒组合物缀合于可用组分，诸如靶向或成像部分(例如染料)。

[0835] 磷脂包括但不限于甘油磷脂,诸如磷脂酰胆碱、磷脂酰乙醇胺、磷脂酰丝氨酸、磷脂酰肌醇、磷脂酰甘油和磷脂酸。磷脂还包括磷酸鞘脂,诸如鞘磷脂。

[0836] 在一些实施方案中,本发明的磷脂包括1,2-二硬脂酰基-sn-甘油-3-磷酸胆碱(DSPC)、1,2-二油酰基-sn-甘油-3-磷酸乙醇胺(DOPE)、1,2-二亚油酰基-sn-甘油-3-磷酸胆碱(DLPC)、1,2--二肉豆蔻酰基-sn-甘油-磷酸胆碱(DMPC)、1,2-二油酰基-sn-甘油-3-磷酸胆碱(DOPC)、1,2-二棕榈酰基-sn-甘油-3-磷酸胆碱(DPPC)、1,2-二-十-烷酰基-sn-甘油-磷酸胆碱(DUPC)、1-棕榈酰基-2-油酰基-sn-甘油-3-磷酸胆碱(POPC)、1,2-二-O-十八烯基-sn-甘油-3-磷酸胆碱(18:0二醚PC)、1-油酰基-2胆固醇半琥珀酰基-sn-甘油-3-磷酸胆碱(OChemsPC)、1-十六烷基-sn-甘油-3-磷酸胆碱(C16 Lyso PC)、1,2--二亚麻酰基-sn-甘油-3-磷酸胆碱、1,2-二花生四烯酰基-sn-甘油-3-磷酸胆碱、1,2-二-二十二碳六烯酰基-sn-甘油-3-磷酸胆碱、1,2-二植酰基-sn-甘油-3-磷酸乙醇胺(ME 16.0PE)、1,2--二硬脂酰基-sn-甘油-3-磷酸乙醇胺、1,2-二亚油酰基-sn-甘油-3-磷酸乙醇胺、1,2-二亚麻酰基-sn-甘油-3-磷酸乙醇胺、1,2-二花生四烯酰基-sn-甘油-3-磷酸乙醇胺、1,2-二-二十二碳六烯酰基-sn-甘油-3-磷酸乙醇胺、1,2-二油酰基-sn-甘油-3-磷酸-外消旋-(1-甘油)钠盐(DOPG)、鞘磷脂以及它们的混合物。

[0837] 在某些实施方案中,在本发明中有用或可能有用的磷脂是DSPC的类似物或变体。在某些实施方案中,可用于或潜在可用于本发明的磷脂是式(IV)化合物:

[0838] 

(IV),

[0839] 或其盐,其中:

[0840] 每个R1独立地是任选取代的烷基;或者任选地两个R1与居间原子接合在一起以形成任选取代的单环碳环基或任选取代的单环杂环基;或者任选地三个R1与居间原子接合在一起以形成任选取代的双环碳环基或任选取代的双环杂环基;

[0841] n是1、2、3、4、5、6、7、8、9或10;

[0842] m是0、1、2、3、4、5、6、7、8、9或10;

[0843] A具有下式:  ;

[0844] L2的每个实例独立地是键或任选取代的C1-6亚烷基,其中任选取代的C1-6亚烷基的一个亚甲基单元任选地被O、N(RN)、S、C(O)、C(O)N(RN)、NRNC(O)、C(O)O、OC(O)、OC(O)O、OC(O)N(RN)、NRNC(O)O或NRNC(O)N(RN)置换;

[0845] R2的每个实例独立地是任选取代的C1-30烷基、任选取代的C1-30烯基或任选取代的C1-30炔基;任选地其中R2的一个或多个亚甲基单元独立地被任选取代的亚碳环基、任选取代的亚杂环基、任选取代的亚芳基、任选取代的亚杂芳基、N(RN)、O、S、C(O)、C(O)N(RN)、NRNC(O)、NRNC(O)N(RN)、C(O)O、OC(O)、OC(O)O、OC(O)N(RN)、NRNC(O)O、C(O)S、SC(O)、C(=NRN)、C(=NRN)N(RN)、NRNC(=NRN)、NRNC(=NRN)N(RN)、C(S)、C(S)N(RN)、NRNC(S)、NRNC

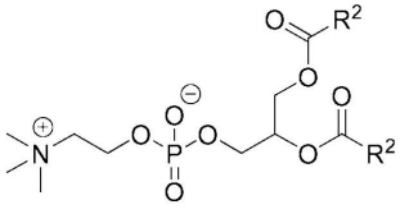
(S)N(RN)、S(O)、OS(O)、S(O)O、OS(O)O、OS(O)2、S(O)2O、OS(O)2O、N(RN)S(O)、S(O)N(RN)、N(RN)S(O)N(RN)、oS(O)N(RN)、N(RN)S(O)O、S(O)2、N(RN)S(O)2、S(O)2N(RN)、N(RN)S(O)2N(RN)、OS(O)2N(RN)或N(RN)S(O)2O置换;

[0846] RN的每个实例独立地是氢、任选取代的烷基或氮保护基;

[0847] 环B是任选取代的碳环基、任选取代的杂环基、任选取代的芳基或任选取代的杂芳基;并且

[0848] p是1或2;

[0849] 前提条件是化合物不具有下式:

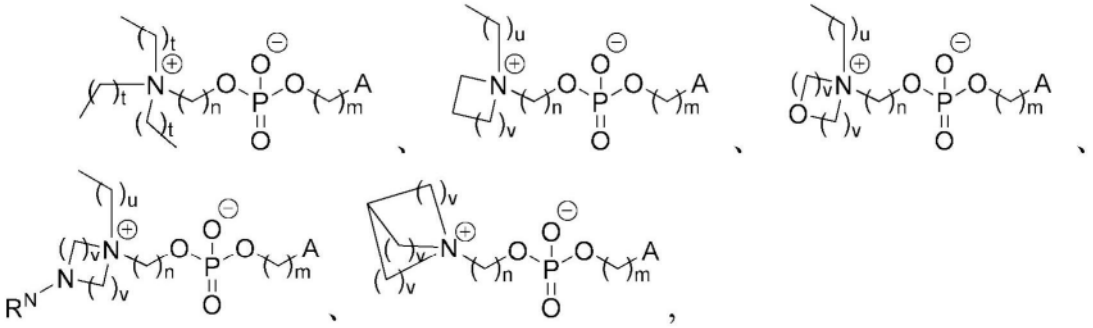
[0850] ,

[0851] 其中R2的每个实例独立地是未取代的烷基、未取代的烯基或未取代的炔基。

[0852] 在一些实施方案中,磷脂可以是描述于美国申请号62/520,530中的一种或多种磷脂。

[0853] i)磷脂头修饰

[0854] 在某些实施方案中,在本发明中有用或可能有用的磷脂包含修饰的磷脂头部(例如,修饰的胆碱基团)。在某些实施方案中,具有修饰的头部的磷脂是具有修饰的季胺的DSPC或其类似物。举例来说,在式(IV)的实施方案中,至少一个R1不是甲基。在某些实施方案中,至少一个R1不是氢或甲基。在某些实施方案中,式(IV)化合物具有下式之一:

[0855] 

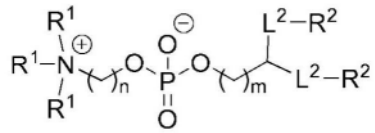
[0856] 或其盐,其中:

[0857] 每个t独立地是1、2、3、4、5、6、7、8、9或10;

[0858] 每个u独立地是0、1、2、3、4、5、6、7、8、9或10;并且

[0859] 每个v独立地是1、2或3。

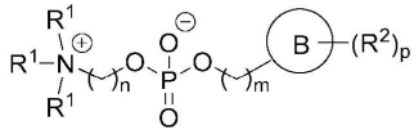
[0860] 在某些实施方案中,式(IV)化合物具有式(IV-a):

[0861] 

(IV-a),

[0862] 或其盐。

[0863] 在某些实施方案中,在本发明中有用或可能有用的磷脂包含环状部分代替甘油酯部分。在某些实施方案中,在本发明中有用的磷脂是DSPC或其类似物,其具有环状部分代替甘油酯部分。在某些实施方案中,式(IV)化合物具有式(IV-b):

[0864] ,

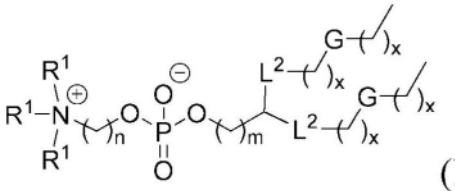
(IV-b),

[0865] 或其盐。

[0866] (ii)磷脂尾修饰

[0867] 在某些实施方案中,在本发明中有用或可能有用的磷脂包含修饰的尾部。在某些实施方案中,在本发明中有用或可能有用的磷脂是具有修饰的尾部的DSPC或其类似物。如本文所述,"修饰的尾部"可以是具有较短或较长脂族链、引入支链的脂族链、引入取代基的脂族链、其中一个或多个亚甲基由环状或杂原子基团替换的脂族链或它们的任何组合的尾部。举例来说,在某些实施方案中,(IV)的化合物具有式(IV-a),或其盐,其中R2的至少一个实例是R2的每个实例是任选取代的C1-30烷基,其中R2的一个或多个亚甲基单元独立地被任选取代的亚碳环基、任选取代的亚杂环基、任选取代的亚芳基、任选取代的亚杂芳基、N(RN)、O、S、C(O)、C(O)N(RN)、NRNC(O)、NRNC(O)N(RN)、C(O)O、OC(O)、OC(O)O、OC(O)N(RN)、NRNC(O)O、C(O)S、SC(O)、C(=NRN)、C(=NRN)N(RN)、NRNC(=NRN)、NRNC(=NRN)N(RN)、C(S)、C(S)N(RN)、NRNC(S)、NRNC(S)N(RN)、S(O)、OS(O)、S(O)O、OS(O)O、OS(O)2、S(O)2O、OS(O)2O、N(RN)S(O)、S(O)N(RN)、N(RN)S(O)N(RN)、OS(O)N(RN)、N(RN)S(O)O、S(O)2、N(RN)S(O)2、S(O)2N(RN)、N(RN)S(O)2N(RN)、OS(O)2N(RN)或N(RN)S(O)2O置换。

[0868] 在某些实施方案中,式(IV)化合物具有式(IV-c):

[0869] 

(IV-c),

[0870] 或其盐,其中:

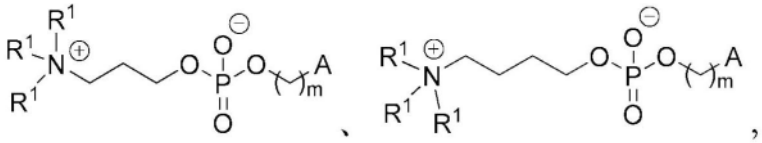
[0871] 每个x独立地是0-30之间的整数,包括端值;并且

[0872] G的每个实例独立地选自由以下组成的组:任选取代的亚碳环基、任选取代的亚杂环基、任选取代的亚芳基、任选取代的亚杂芳基、N(RN)、O、S、C(O)、C(O)N(RN)、NRNC(O)、NRNC(O)N(RN)、C(O)O、OC(O)、OC(O)O、OC(O)N(RN)、NRNC(O)O、C(O)S、SC(O)、C(=NRN)、C(=NRN)N(RN)、NRNC(=NRN)、NRNC(=NRN)N(RN)、C(S)、C(S)N(RN)、NRNC(S)、NRNC(S)N(RN)、S(O)、OS(O)、S(O)O、OS(O)O、OS(O)2、S(O)2O、OS(O)2O、N(RN)S(O)、S(O)N(RN)、N(RN)S(O)N(RN)、OS(O)N(RN)、N(RN)S(O)O、S(O)2、N(RN)S(O)2、S(O)2N(RN)、N(RN)S(O)2N(RN)、OS(O)2N(RN)或N(RN)S(O)2O。每种可能性代表本发明的单独实施方案。

[0873] 在某些实施方案中,在本发明中有用或可能有用的磷脂包含修饰的磷酸胆碱部分,其中将季胺连接至磷酰基的烷基链不是亚乙基(例如,n不是2)。因此,在某些实施方案

中,在本发明中有用或可能有用的磷脂是式(IV)化合物,其中n是1、3、4、5、6、7、8、9或10。

[0874] 举例来说,在某些实施方案中,式(IV)化合物具有下式之一:

[0875] 

[0876] 或其盐。

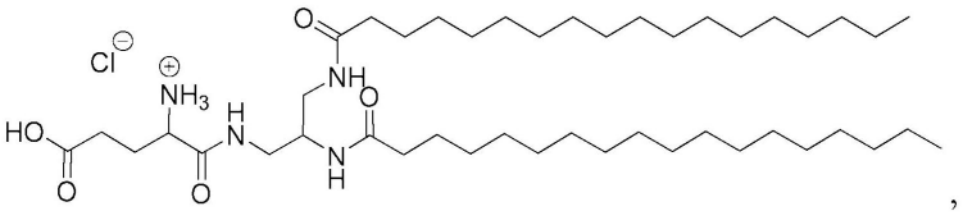
[0877] 替代性脂质

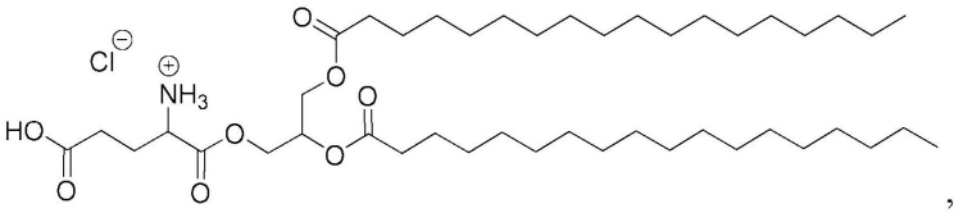
[0878] 在某些实施方案中,在本发明中有用或可能有用的磷脂包含修饰的磷酸胆碱部分,其中将季胺连接至磷酰基的烷基链不是亚乙基(例如,n不是2)。因此,在某些实施方案中,磷脂是有用的。

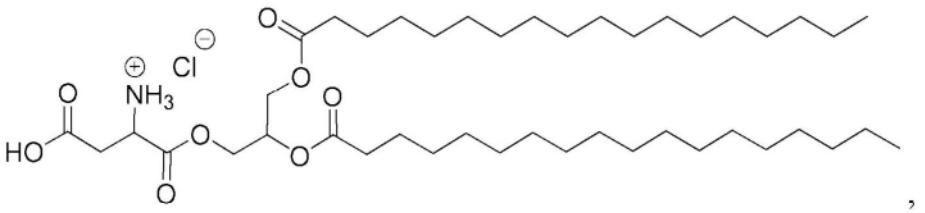
[0879] 在某些实施方案中,使用替代性脂质代替本公开的磷脂。

[0880] 在某些实施方案中,本发明的替代性脂质是油酸。

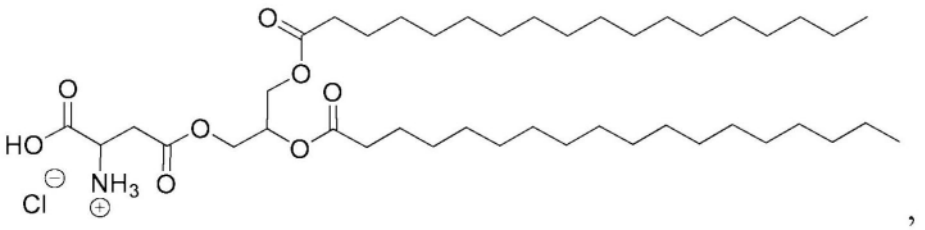
[0881] 在某些实施方案中,替代性脂质是以下之一:

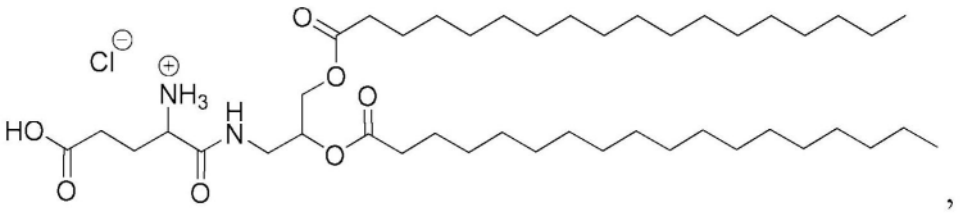


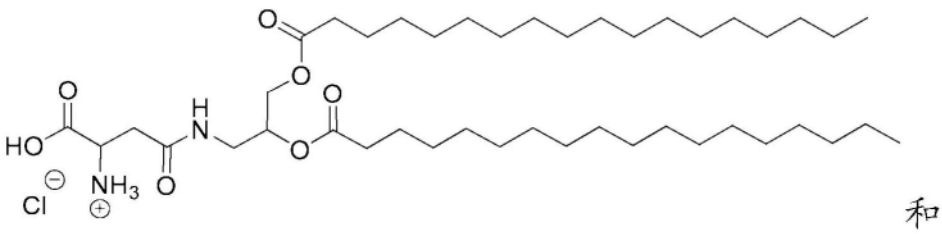




[0882]

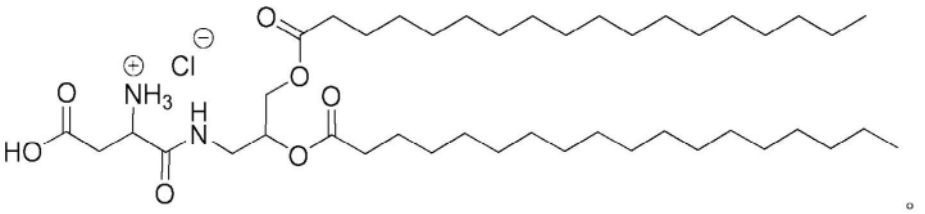






和

[0883]



[0884] 结构脂质

[0885] 本文所公开的药物组合物的脂质组合物可以包含一种或多种结构脂质。如本文所

用,术语“结构脂质”是指固醇并且还指含有固醇部分的脂质。

[0886] 在脂质纳米颗粒中并入结构脂质可以帮助减轻颗粒中其它脂质的聚集。结构脂质可以选自包括但不限于胆固醇、粪固醇、谷固醇、麦角固醇、菜油固醇、豆固醇、菜子固醇、番茄碱(tomatidine)、番茄素(tomatine)、熊果酸、α-生育酚、藿烷、植物固醇、类固醇以及它们的混合物的组。在一些实施方案中,结构脂质是固醇。如本文所定义,“固醇”是由类固醇类醇组成的类固醇子组。在某些实施方案中,结构脂质是类固醇。在某些实施方案中,结构脂质是胆固醇。在某些实施方案中,结构脂质是胆固醇的类似物。在某些实施方案中,结构脂质是$_{\alpha}$α-生育酚。

[0887] 在一些实施方案中,结构脂质可以是描述于美国申请号62/520,530中的一种或多种结构脂质。

[0888] 聚乙二醇(PEG)-脂质

[0889] 本文所公开的药物组合物的脂质组合物可以包含一种或多种聚乙二醇(PEG)脂质。

[0890] 如本文所用,术语“PEG-脂质”是指聚乙二醇(PEG)修饰的脂质。PEG-脂质的非限制性实例包括PEG修饰的磷脂酰乙醇胺和磷脂酸、PEG-神经酰胺缀合物(例如,PEG-CerC14或PEG-CerC20)、PEG修饰的二烷基胺和PEG修饰的1,2--二酰氧基丙-3-胺。此类脂质也被称为聚乙二醇化脂质。举例来说,PEG脂质可以是PEG-c-DOMG、PEG-DMG、PEG-DLPE、PEG-DMPE、PEG-DPPC或PEG-DSPE脂质。

[0891] 在一些实施方案中,PEG-脂质包括但不限于1,2-二肉豆蔻酰基-snn-甘油甲氧基聚乙二醇(PEG-DMG)、1,2-二硬脂酰基-sn-甘油-3-磷酸乙醇胺-N-[氨基(聚乙二醇)](PEG-DSPE)、PEG-二硬脂基甘油(PEG-DSG)、PEG-二棕榈油烯基、PEG-二油烯基、PEG-二硬脂基、PEG-二酰基甘酰胺(PEG-DAG)、PEG-二棕榈酰基磷脂酰乙醇胺(PEG-DPPE)或PEG-1,2--二肉豆蔻基氧基丙基-3-胺(PEG-c-DMA)。

[0892] 在一个实施方案中,PEG-脂质选自由PEG修饰的磷脂酰乙醇胺、PEG修饰的磷脂酸、PEG修饰的神经酰胺、PEG修饰的二烷基胺、PEG修饰的二酰基甘油、PEG修饰的二烷基甘油和它们的混合物组成的组。

[0893] 在一些实施方案中,PEG-脂质的脂质部分包括长度为约C14至约C22、优选约C14至约C16的那些。在一些实施方案中,PEG部分(例如,mPEG-NH2)的大小为约1000、2000、5000、10,000、15,000或20,000道尔顿。在一个实施方案中,PEG-脂质是PEG2k-DMG。

[0894] 在一个实施方案中,本文所描述的脂质纳米颗粒可以包含PEG脂质,其为不可扩散PEG。不可扩散PEG的非限制性实例包括PEG-DSG和PEG-DSPE。

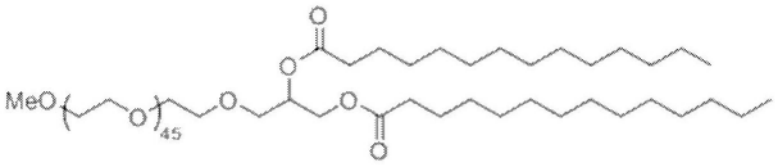
[0895] PEG-脂质是本领域中已知的,诸如美国专利号8158601和国际公布号WO 2015/130584 A2中所描述的那些,所述文献以全文引用的方式并入本文中。

[0896] 一般来说,本文所描述的各式的一些其它脂质组分(例如,PEG脂质)可以如2016年12月10日提交的标题为“Compositions and Methods for Delivery of Therapeutic Agents”的国际专利申请号PCT/US2016/000129中所描述来合成,所述专利申请以全文引用的方式并入。

[0897] 脂质纳米颗粒组合物的脂质组分可以包括一种或多种包含聚乙二醇的分子,诸如PEG或PEG修饰的脂质。此类物质可以替代地称为PEG化脂质。PEG脂质是用聚乙二醇修饰的

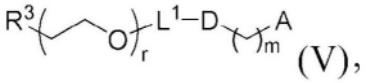
脂质。PEG脂质可以选自包括PEG修饰的磷脂酰乙醇胺、PEG修饰的磷脂酸、PEG修饰的神经酰胺、PEG修饰的二烷基胺、PEG修饰的二酰基甘油、PEG修饰的二烷基甘油以及它们的混合物的非限制性组。举例来说,PEG脂质可以是PEG-c-DOMG、PEG-DMG、PEG-DLPE、PEG-DMPE、PEG-DPPC或PEG-DSPE脂质。

[0898] 在一些实施方案中,PEG修饰的脂质是PEG DMG的修饰形式。PEG-DMG具有以下结构:

[0899] 

[0900] 在一个实施方案中,可用于本发明的PEG脂质可以是描述于国际公布第WO2012099755号中的PEG化脂质,所述公布的内容通过引用整体并入本文。本文所描述的这些示例性PEG脂质中的任一者可以被修饰为在PEG链上包含羟基。在某些实施方案中,PEG脂质是PEG-OH脂质。如本文一般所定义,"PEG-OH脂质"(本文中也称为"羟基-PEG化脂质")是在脂质上具有一个或多个羟基(-OH)基团的PEG化脂质。在某些实施方案中,PEG-OH脂质包含在PEG链上的一个或多个羟基。在某些实施方案中,PEG-OH或羟基-PEG化脂质包含在PEG链的末端的-OH基团。每种可能性代表本发明的单独实施方案。

[0901] 在某些实施方案中,可用于本发明的PEG脂质是式(V)化合物。本文提供了式(V)化合物:

[0902] 

[0903] 或其盐,其中:

[0904] R3是-ORO;

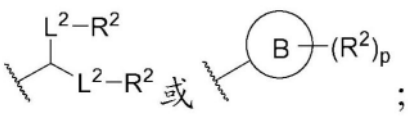
[0905] RO为氢、任选取代的烷基或氧保护基;

[0906] r是介于1与100之间的整数,首尾包括在内;

[0907] L1是任选取代的C1-10亚烷基,其中任选取代的C1-10亚烷基的至少一个亚甲基独立地被任选取代的亚碳环基、任选取代的亚杂环基、任选取代的亚芳基、任选取代的亚杂芳基、O、N(RN)、S、C(O)、C(O)N(RN)、NRNC(O)、C(O)O、OC(O)、OC(O)O、OC(O)N(RN)、NRNC(O)O或NRNC(O)N(RN)置换;

[0908] D是通过点击化学获得的部分或在生理条件下可裂解的部分;

[0909] m是0、1、2、3、4、5、6、7、8、9或10;

[0910] A具有下式:  ;

[0911] L2的每个实例独立地是键或任选取代的C1-6亚烷基,其中任选取代的C1-6亚烷基的一个亚甲基单元任选地被O、N(RN)、S、C(O)、C(O)N(RN)、NRNC(O)、C(O)O、OC(O)、OC(O)O、OC(O)N(RN)、NRNC(O)O或NRNC(O)N(RN)置换;

[0912] R2的每个实例独立地是任选取代的C1-30烷基、任选取代的C1-30烯基或任选取代的C1-30炔基;任选地其中R2的一个或多个亚甲基单元独立地被任选取代的亚碳环基、任选

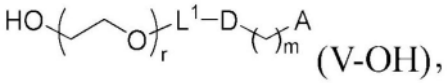
取代的亚杂环基、任选取代的亚芳基、任选取代的亚杂芳基、N(RN)、O、S、C(O)、C(O)N(RN)、NRNC(O)、NRNC(O)N(RN)、C(O)O、OC(O)、OC(O)O、OC(O)N(RN)、NRNC(O)O、C(O)S、SC(O)、C(=NRN)、C(=NRN)N(RN)、NRNC(=NRN)、NRNC(=NRN)N(RN)、C(S)、C(S)N(RN)、NRNC(S)、NRNC(S)N(RN)、S(O)、OS(O)、S(O)O、OS(O)O、OS(O)2、S(O)2O、OS(O)2O、N(RN)S(O)、S(O)N(RN)、N(RN)S(O)N(RN)、OS(O)N(RN)、N(RN)S(O)O、S(O)2、N(RN)S(O)2、S(O)2N(RN)、N(RN)S(O)2N(RN)、OS(O)2N(RN)或N(RN)S(O)2O置换;

[0913] RN的每个实例独立地是氢、任选取代的烷基或氮保护基;

[0914] 环B是任选取代的碳环基、任选取代的杂环基、任选取代的芳基或任选取代的杂芳基;并且

[0915] $_p$是1或2。

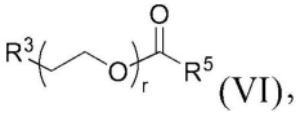
[0916] 在某些实施方案中,式(V)化合物是PEG-OH脂质(即,R3是-ORO,并且RO是氢)。在某些实施方案中,式(V)化合物具有式(V-OH):

[0917] 
(V-OH),

[0918] 或其盐。

[0919] 在某些实施方案中,可用于本发明的PEG脂质是PEG化的脂肪酸。在某些实施方案中,可用于本发明的PEG脂质是式(VI)化合物。

[0920] 本文提供了式(VI)化合物:

[0921] 
(VI),

[0922] 或其盐,其中:

[0923] R3是-ORO;

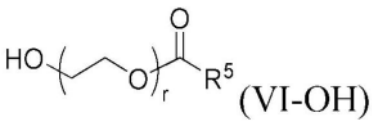
[0924] RO为氢、任选取代的烷基或氧保护基;

[0925] r是介于1与100之间的整数,首尾包括在内;

[0926] R5是任选取代的C10-40烷基、任选取代的C10-40烯基或任选取代的C10-40炔基;并且任选地R5的一个或多个亚甲基被任选取代的亚碳环基、任选取代的亚杂环基、任选取代的亚芳基、任选取代的亚杂芳基、N(RN)、O、S、C(O)、C(O)N(RN)、NRNC(O)、NRNC(O)N(RN)、C(O)O、OC(O)、OC(O)O、OC(O)N(RN)、NRNC(O)O、C(O)S、SC(O)、C(=NRN)、C(=NRN)N(RN)、NRNC(=NRN)、NRNC(=NRN)N(RN)、C(S)、C(S)N(RN)、NRNC(S)、NRNC(S)N(RN)、S(O)、OS(O)、S(O)O、OS(O)O、OS(O)2、S(O)2O、OS(O)2O、N(RN)S(O)、S(O)N(RN)、N(RN)S(O)N(RN)、OS(O)N(RN)、N(RN)S(O)O、S(O)2、N(RN)S(O)2、S(O)2N(RN)、N(RN)S(O)2N(RN)、OS(O)2N(RN)或N(RN)S(O)2O置换;并且

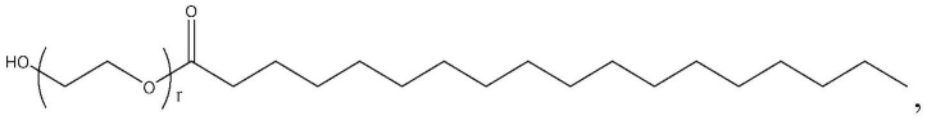
[0927] RN的每个实例独立地是氢、任选取代的烷基或氮保护基。

[0928] 在某些实施方案中,式(VI)化合物具有式(VI-OH):

[0929] 
(VI-OH),

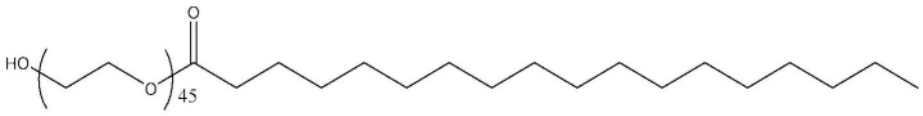
[0930] 或其盐。在一些实施方案中,r是45。

[0931] 在其它实施方案中,式(VI)化合物是:

[0932] 

[0933] 或其盐。

[0934] 在一个实施方案中,式(VI)化合物是

[0935] 

[0936] (化合物I)。

[0937] 在一些方面,本文所公开的药物组合物的脂质组合物不包含PEG-脂质。

[0938] 在一些实施方案中,PEG-脂质可以是描述于美国申请号62/520,530中的一种或多种PEG脂质。

[0939] 在一些实施方案中,本发明的PEG脂质包括PEG修饰的磷脂酰乙醇胺、PEG修饰的磷脂酸、PEG修饰的神经酰胺、PEG修饰的二烷基胺、PEG修饰的二酰基甘油、PEG修饰的二烷基甘油以及它们的混合物。在一些实施方案中,PEG修饰的脂质是PEG-DMG、PEG-c-DOMG(也称为PEG-DOMG)、PEG-DSG和/或PEG-DPG。

[0940] 在一些实施方案中,本发明的LNP包含式I、式II或式III中的任一者的可电离阳离子脂质、包含DSPC的磷脂、结构脂质和包含PEG-DMG的PEG脂质。

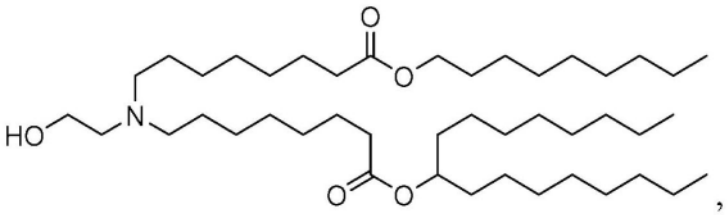
[0941] 在一些实施方案中,本发明的LNP包含式I、式II或式III中的任一者的可电离阳离子脂质、包含DSPC的磷脂、结构脂质和包含具有式VI的化合物的PEG脂质。

[0942] 在一些实施方案中,本发明的LNP包含式I、式II或式III的可电离阳离子脂质、包含具有式IV的化合物的磷脂、结构脂质和包含具有式V或式VI的化合物的PEG脂质。

[0943] 在一些实施方案中,本发明的LNP包含式I、式11或式III的可电离阳离子脂质、包含具有式IV的化合物的磷脂、结构脂质和包含具有式V或式VI的化合物的PEG脂质。

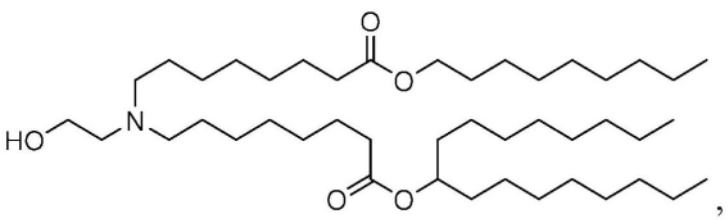
[0944] 在一些实施方案中,本发明的LNP包含式I、式II或式III的可电离阳离子脂质、具有式IV的磷脂、结构脂质和包含具有式VI的化合物的PEG脂质。

[0945] 在一些实施方案中,本发明的LNP包含下式的可电离阳离子脂质

[0946] 

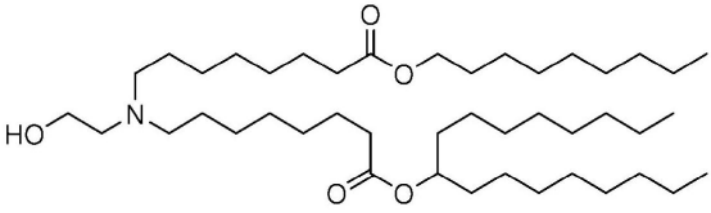
[0947] 和包含式VI的PEG脂质。

[0948] 在一些实施方案中,本发明的LNP包含下式的可电离阳离子脂质

[0949] 

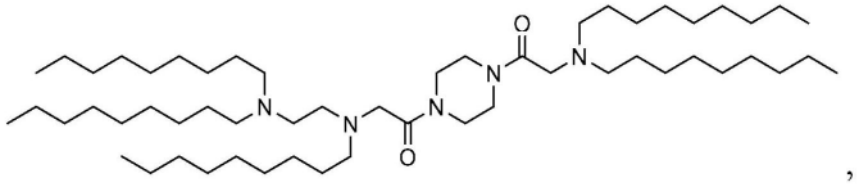
[0950] 和包含油酸的替代性脂质。

[0951] 在一些实施方案中,本发明的LNP包含下式的可电离阳离子脂质

[0952] 

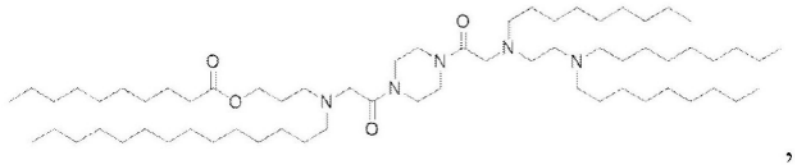
[0953] 包含油酸的替代性脂质、包含胆固醇的结构脂质和包含具有式VI的化合物的PEG脂质。

[0954] 在一些实施方案中,本发明的LNP包含下式的可电离阳离子脂质

[0955] 

[0956] 包含DOPE的磷脂、包含胆固醇的结构脂质和包含具有式VI的化合物的PEG脂质。

[0957] 在一些实施方案中,本发明的LNP包含下式的可电离阳离子脂质

[0958] 

[0959] 包含DOPE的磷脂、包含胆固醇的结构脂质和包含具有式VI的化合物的PEG脂质。

[0960] 在一些实施方案中,本发明的LNP包含约2:1至约30:1的N:P比率。

[0961] 在一些实施方案中,本发明的LNP包含约6:1的N:P比率。

[0962] 在一些实施方案中,本发明的LNP包含约3:1的N:P比率。

[0963] 在一些实施方案中,本发明的LNP包含约10:1至约100:1的可电离阳离子脂质组分与RNA的重量/重量比。

[0964] 在一些实施方案中,本发明的LNP包含约20:1的可电离阳离子脂质组分与RNA的重量/重量比。

[0965] 在一些实施方案中,本发明的LNP包含约10:1的可电离阳离子脂质组分与RNA的重量/重量比。

[0966] 在一些实施方案中,本发明的LNP具有约50nm至约150nm的平均直径。

[0967] 在一些实施方案中,本发明的LNP具有约70nm至约120nm的平均直径。

[0968] 如本文所用,术语“烷基”、“烷基基团”或“亚烷基”意指包括一个或多个碳原子(例如,一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个、十二个、十三个、十四个、十五个、十六个、十七个、十八个、十九个、二十个或更多个碳原子)的直链或支链饱和烃,其任选被取代。注记“C1-14烷基”意指任选被取代的包括1-14个碳原子的直链或支链饱和烃。除非另有指定,否则本文所描述的烷基基团是指未被取代和被取代的烷基基团。

[0969] 如本文所用,术语“烯基”、“烯基基团”或“亚烯基”意指包括两个或更多个碳原子(例如,两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个、十二个、十三个、十四

个、十五个、十六个、十七个、十八个、十九个、二十个或更多个碳原子)和至少一个双键的直链或支链烃,其任选被取代。注记"C2-14烯基"意指任选被取代的包括2-14个碳原子和至少一个碳碳双键的直链或支链烃。烯基基团可以包括一个、两个、三个、四个或更多个碳碳双键。举例来说,C18烯基可以包括一个或多个双键。包括两个双键的C18烯基基团可以是亚油烯基。除非另有指定,否则本文所描述的烯基基团是指未被取代和被取代的烯基基团。

[0970] 如本文所用,术语"炔基"、"炔基基团"或"亚炔基"意指包括两个或更多个碳原子(例如,两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个、十二个、十三个、十四个、十五个、十六个、十七个、十八个、十九个、二十个或更多个碳原子)和至少一个碳碳三键的直链或支链烃,其任选被取代。注记"C2-14炔基"意指任选被取代的包括2-14个碳原子和至少一个碳碳三键的直链或支链烃。炔基基团可以包括一个、两个、三个、四个或更多个碳碳三键。举例来说,C18炔基可以包括一个或多个碳碳三键。除非另有指定,否则本文所描述的炔基基团是指未被取代和被取代的炔基基团。

[0971] 如本文所用,术语"碳环"或"碳环基团"意指任选被取代的包括一个或多个碳原子环的单环或多环体系。环可以是三元、四元、五元、六元、七元、八元、九元、十元、十一元、十二元、十三元、十四元、十五元、十六元、十七元、十八元、十九元或二十元环。标注"C3-6碳环"意指包含具有3-6个碳原子的单环的碳环。碳环可以包括一个或多个碳碳双键或三键并且可以是非芳族的或芳族的(例如,环烷基或芳基基团)。碳环的实例包括环丙基、环戊基、环已基、苯基、萘基和1,2二氢萘基基团。如本文所用,术语"环烷基"意指非芳族碳环并且可以包括或可不包括任何双键或三键。除非另有指定,否则本文所描述的碳环是指未被取代和被取代的碳环基团,即,任选被取代的碳环。

[0972] 如本文所用,术语"杂环"或"杂环基团"意指任选被取代的包括一个或多个环的单环或多环体系,其中至少一个环包括至少一个杂原子。杂原子可以是例如氮、氧或硫原子。环可以是三元、四元、五元、六元、七元、八元、九元、十元、十一元、十二元、十三元或十四元环。杂环可以包括一个或多个双键或三键并且可以是非芳族的或芳族的(例如,杂环烷基或杂芳基基团)。杂环的实例包括咪唑基、咪唑烷基、噁唑基、噁唑烷基、噻唑基、噻唑烷基、吡唑烷基、吡唑基、异噁唑烷基、异噁唑基、异噻唑烷基、异噻唑基、吗啉基、吡咯基、吡咯烷基、呋喃基、四氢呋喃基、噻吩基、吡啶基、哌啶基、喹啉基和异喹啉基基团。如本文所用,术语"杂环烷基"意指非芳族杂环并且可以包括或可不包括任何双键或三键。除非另有指定,否则本文所描述的杂环是指未被取代和被取代的杂环基团,即,任选被取代的杂环。

[0973] 如本文所用,术语"杂烷基"、"杂烯基"或"杂炔基"分别是指如本文所定义的烷基、烯基、炔基,其还包含一个或多个(例如,1、2、3或4个)杂原子(例如,氧、硫、氮、硼、硅、磷),其中一个或多个杂原子插入母体碳链内的相邻碳原子之间,并且/或者一个或多个杂原子插入碳原子与母体分子之间,即,连接点之间。除非另有指定,否则本文所描述的杂烷基、杂烯基或杂炔基是指未被取代和被取代的杂烷基、杂烯基或杂炔基,即,任选被取代的杂烷基、杂烯基或杂炔基。

[0974] 如本文所用,"生物可降解基团"是可以促进哺乳动物实体中脂质的更快速代谢的基团。生物可降解基团可以选自由但不限于-C(O)O-、-OC(O)-、-C(O)N(R′)-、-N(R′)C(O)-、-C(O)-、-C(S)-、-C(S)S-、-SC(S)-、-CH(OH)-、-P(O)(OR')O-、-S(O)2-、芳基基团和杂芳基基团组成的组。如本文所用,"芳基基团"是任选被取代的包括一个或多个芳环的碳环

基团。芳基基团的实例包括苯基和萘基基团。如本文所用,“杂芳基基团”是任选被取代的包括一个或多个芳环的杂环基团。杂芳基基团的实例包括吡咯基、呋喃基、噻吩基、咪唑基、噁唑基和噻唑基。芳基和杂芳基基团均可以任选被取代。举例来说,M和M′可以选自由任选被取代的苯基、噁唑和噻唑组成的非限制性组。在本文中的各式中,M和M′可独立地选自上文生物可降解基团的列表。除非另有指定,否则本文所描述的芳基或杂芳基基团是指未被取代和被取代的基团,即,任选被取代的芳基或杂芳基基团。

[0975] 除非另有指定,否则烷基、烯基和环基(例如,碳环基和杂环基)基团可以任选被取代。任选的取代基可以选自由以下组成但不限于以下的组:卤素原子(例如,氯、溴、氟或碘基团)、羧酸(例如,C(O)OH)、醇(例如,羟基、OH)、酯(例如,C(O)OR、OC(O)R)、醛(例如,C(O)H)、羰基(例如,C(O)R,替代地用C=O表示)、酰基卤(例如,C(O)X,其中X是选自溴化物、氟化物、氯化物和碘化物的卤化物)、碳酸酯(例如,OC(O)OR)、烷氧基(例如,OR)、缩醛(例如,C(OR)2R″″,其中每个OR是可以相同或不同的烷氧基并且R″′″是烷基或烯基基团)、磷酸根(例如,P(O)43-)、硫醇(例如,SH)、亚砜(例如,S(O)R)、亚磺酸(例如,S(O)OH)、磺酸(例如,S(O)2OH)、硫醛(例如,C(S)H)、硫酸根(例如,S(O)42-)、磺酰基(例如,S(O)2)、酰胺(例如,C(O)NR2或N(R)C(O)R)、叠氮基(例如,N3)、硝基(例如,NO2)、氰基(例如,CN)、异氰基(例如,NC)、酰氧基(例如,OC(O)R)、氨基(例如,NR2、NRH或NH2)、氨基甲酰基(例如,OC(O)NR2、OC(O)NRH或OC(O)NH2)、磺酰胺(例如,S(O)2NR2、S(O)2NRH、S(O)2NH2、N(R)S(O)2R、N(H)S(O)2R、N(R)S(O)2H或N(H)S(O)2H)、烷基基团、烯基基团和环基(例如,碳环基或杂环基)基团。在前述任一者中,R是如本文所定义的烷基或烯基基团。在一些实施方案中,所述取代基基团自身可以进一步被例如一个、两个、三个、四个、五个或六个如本文所定义的取代基取代。举例来说,C1-6烷基基团可以进一步被一个、两个、三个、四个、五个或六个如本文所描述的取代基取代。

[0976] 含氮的本公开化合物可通过用氧化剂(例如3-氯过氧基苯甲酸(mCPBA)和/或过氧化氢)处理而转化为N-氧化物以提供其它本公开化合物。因此,当价态和结构允许时,所有所显示和要求保护的含氮化合物均被视为包括如所示的化合物及其N-氧化物衍生物(其可以被指定为N→O或N+-O-)。此外,在其它情况下,本公开化合物中的氮可转化为N-羟基或N-烷氧基化合物。举例来说,N-羟基化合物可通过利用诸如m CPBA的氧化剂氧化母体胺来制备。当价态和结构允许时,所有所显示和要求保护的含氮化合物也被视为涵盖如所示的化合物及其N-羟基(即,N-OH)和N-烷氧基(即,N-OR,其中R是被取代或未被取代的C1-C6烷基、C1-C6烯基、C1-C6炔基、3-14元碳环或3-14元杂环)衍生物。

[0977] mRNA-脂质加合物

[0978] 已经确定某些可电离脂质容易形成脂质-多核苷酸加合物。具体地,包含叔胺基团的可电离脂质可分解成仲胺和反应性醛物质中的一种或两种,所述仲胺和反应性醛物质能够与多核苷酸(诸如mRNA)相互作用以形成可电离脂质-多核苷酸加合物杂质,其可以通过反相离子对色谱法(RP-IP HPLC)检测。例如,叔胺的氧化可导致N-氧化物形成,其可在胺处经历酸/碱催化的水解以生成醛和仲胺,所述醛和仲胺可以与mRNA形成加合物。因此,在一些方面,可电离脂质-多核苷酸加合物杂质是醛-mRNA加合物杂质。

[0979] 还已经确定此类加合物可破坏mRNA翻译并影响脂质纳米颗粒(LNP)配制的mRNA产品的活性。因此,制备和使用可电离脂质-多核苷酸加合物杂质含量降低的LNP组合物可能

是有利的,诸如其中小于约20%、小于约10%、小于约5%、或小于约1%的mRNA是可电离脂质-多核苷酸加合物杂质的形式,如可以通过RP-IP HPLC测量的。因此,根据一些方面,提供了一种LNP组合物,其中小于约10%、小于约5%、或小于约1%的mRNA是可电离脂质-多核苷酸加合物杂质的形式,包括小于10%、小于5%、或小于1%,如可通过RP-IP HPLC测量的。

[0980] 在一些方面,组合物中脂质醛的量小于约50ppm,包括小于50ppm。此外或另选地,在一些方面,组合物中的N-氧化物化合物的量小于约50ppm,包括小于50ppm。此外或另选地,在一些方面,组合物中的过渡金属(诸如Fe)的量小于约50ppm,包括小于50ppm。此外或另选地,在一些方面,组合物中的烷基卤化物化合物的量小于约50ppm,包括小于50ppm。此外或另选地,在一些方面,组合物中的酸酐化合物的量小于约50ppm,包括小于50ppm。此外或另选地,在一些方面,组合物中的酮化合物的量小于约50ppm,包括小于50ppm。此外或另选地,在一些方面,组合物中的共轭二烯化合物的量小于约50ppm,包括小于50ppm。

[0981] 在一些方面,组合物对于可电离脂质-多核苷酸加合物杂质的形成是稳定的。在一些方面,当在约25℃或更低的温度下储存时,组合物中的可电离脂质-多核苷酸加合物杂质的量以小于每天约2%的平均速率增加,包括以小于每天2%的平均速率增加。在一些方面,当在约5℃或更低的温度下储存时,组合物中的可电离脂质-多核苷酸加合物杂质的量以小于每天约0.5%的平均速率增加,包括以小于每天0.5%的平均速率增加。在一些方面,当在冷藏温度下储存时,组合物中的可电离脂质-多核苷酸加合物杂质的量以小于每天约0.5%的平均速率增加,任选地其中所述冷藏温度为约5℃。

[0982] 具有降低的可电离脂质-多核苷酸加合物杂质含量的脂质媒介物(例如,LNP)组合物可以通过抑制N-氧化物和醛中的一者或两者形成的方法来制备。此类方法可以包括处理包含可电离脂质的组合物,所述可电离脂质包含叔胺基团,以抑制N-氧化物和醛中的一者或两者的形成,诸如通过用还原剂处理组合物;用螯合剂处理组合物;调节组合物的pH;调节组合物的温度;并调节组合物中的缓冲液。此类方法可以包括以下中的一项或多项:在将可电离脂质与多核苷酸组合之前,用清除剂处理可电离脂质;用还原处理剂处理可电离脂质;用还原剂处理可电离脂质;用螯合剂处理可电离脂质;用还原剂处理多核苷酸;以及用螯合剂处理多核苷酸。

[0983] 根据前述项中的任一项,清除剂、还原处理剂和/或还原剂可以是与醛、酮、酸酐和/或二烯化合物反应的剂。清除剂可以包括选自以下的一种或多种:(O-(2,3,4,5,6-五氟苄基)羟胺盐酸盐)(PFBHA)、甲氧胺(例如,甲氧胺盐酸盐)、苄氧胺(例如,苄氧胺盐酸盐)、乙氧胺(例如,乙氧胺盐酸盐)、4-[2-(氨基氧基)乙基]吗啉二盐酸盐、丁氧胺(例如,叔丁氧胺盐酸盐)、4-二甲基氨基吡啶(DMAP)、1,4-二氮杂双环[2.2.2]辛烷(DABCO)、三乙胺(TEA)、哌啶4-羧酸酯(BPPC)及其组合。还原处理剂可以包括硼化合物(例如,硼氢化钠和/或联硼酸频那醇酯)。还原处理剂可以包括硼化合物,诸如硼氢化钠和或联硼酸频那醇酯中的一种或两种。螯合剂可以包括固定化亚氨基二乙酸。还原剂可以包括固定化还原剂,诸如二氧化硅上的固定化二苯膦(Si-DPP)、琼脂糖上的固定化硫醇(Ag-Thiol)、二氧化硅上的固定化半胱氨酸(Si-半胱氨酸)、二氧化硅上的固定化硫醇(Si-Thiol)或其组合。还原剂可以包括游离还原剂,诸如焦亚硫酸钾、硫代乙醇酸钠、三(2-羧乙基)膦(TCEP)、硫代硫酸钠、N-乙酰基半胱氨酸、谷胱甘肽、二硫苏糖醇(DTT)、胱胺、二硫赤藓糖醇(DTE)、二氯二苯基三氯乙烷(DDT)、高半胱氨酸、硫辛酸或其组合。

[0984] 根据前述项中的任一项,pH可以为或者可以调节为约7至约9的pH。

[0985] 根据前述项中的任一项,缓冲液可以选自磷酸钠、柠檬酸钠、琥珀酸钠、组氨酸、组氨酸-HCl、苹果酸钠、碳酸钠和TRIS(三(羟甲基)氨基甲烷)。根据前述项中的任一项,缓冲液可以是TRIS,并且可以是或调节为约20mM至约150mM TRIS。

[0986] 根据前述项中的任一项,组合物的温度可以是或调节至25℃或更低。

[0987] 组合物还可以包含游离还原剂或抗氧化剂。

[0988] 16.示例性额外LNP组分

[0989] 表面活性剂

[0990] 在某些实施方案中,本公开的脂质纳米颗粒任选地包括一种或多种表面活性剂。

[0991] 在某些实施方案中,表面活性剂是两亲性聚合物。如本文所用,两亲性"聚合物"是包含寡聚物或聚合物的两亲性化合物。

[0992] 举例来说,两亲性聚合物可以包含寡聚物片段,诸如两个或更多个PEG单体单元。举例来说,本文所描述的两亲性聚合物可以是PS20。

[0993] 举例来说,两亲性聚合物是嵌段共聚物。

[0994] 举例来说,两亲性聚合物是冻干保护剂。

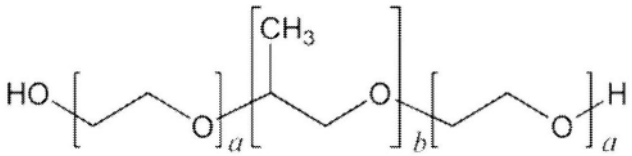
[0995] 举例来说,两亲性聚合物在约30℃和大气压下于水中具有小于2×10-4M的临界胶束浓度(CMC)。

[0996] 举例来说,两亲性聚合物在约30℃和大气压下于水中具有范围介于约0.1×10^{-4}M与约1.3×10^{-4}M之间的临界胶束浓度(CMC)。

[0997] 举例来说,制剂中两亲性聚合物的浓度范围介于约其CMC与CMC的约30倍之间(例如,高达其CMC的约25倍、约20倍、约15倍、约10倍、约5倍或约3倍),例如,在冷冻或冻干之前。

[0998] 举例来说,两亲性聚合物选自泊洛沙姆(poloxamer,Pluronic®)、泊洛沙胺(poloxamine,Tetronic®)、聚氧乙二醇脱水山梨糖醇烷基酯(聚山梨醇酯)和聚乙烯吡咯烷酮(PVP)。

[0999] 举例来说,两亲性聚合物是泊洛沙姆。举例来说,两亲性聚合物具有以下结构:

[1000] 

,

[1001] 其中a是介于10与150之间的整数并且b是介于20与60之间的整数。举例来说,a为约12并且b为约20,或a为约80并且b为约27,或a为约64并且b为约37,或a为约141并且b为约44,或a为约101并且b为约56。

[1002] 举例来说,两亲性聚合物是P124、P188、P237、P338或P407。

[1003] 举例来说,两亲性聚合物是P188(例如,泊洛沙姆188,CAS编号9003-11-6,也称为Kolliphor P188)。

[1004] 举例来说,两亲性聚合物是泊洛沙胺,例如,tetronic 304或tetronic 904。

[1005] 举例来说,两亲性聚合物是聚乙烯吡咯烷酮(PVP),诸如分子量为3kDa、10kDa或29kDa的PVP。

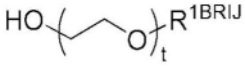
[1006] 举例来说,两亲性聚合物是聚山梨醇酯,诸如PS20。

[1007] 在某些实施方案中,表面活性剂是非离子表面活性剂。

[1008] 在一些实施方案中,脂质纳米颗粒包含表面活性剂。在一些实施方案中,表面活性剂是两亲性聚合物。在一些实施方案中,表面活性剂是非离子表面活性剂。

[1009] 举例来说,非离子表面活性剂选自由聚乙二醇醚(Brij)、泊洛沙姆、聚山梨醇酯、脱水山梨糖醇及其衍生物组成的组。

[1010] 举例来说,聚乙二醇醚是式(VIII)化合物:

[1011] 
(VIII),

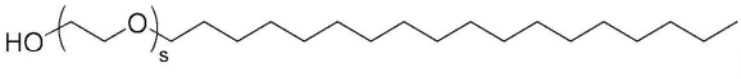
[1012] 或其盐或异构体,其中:

[1013] t是介于1与100之间的整数;

[1014] R^{1BRIJ}独立地是C_{10-40}烷基、C_{10-40}烯基或C_{10-40}炔基;并且任选地R^{5PEG}的一个或多个亚甲基独立地被C_{3-10}亚碳环基、4至10元亚杂环基、C_{6-10}亚芳基、4至10元亚杂芳基、-N(R^N)-、-O-、-S-、-C(O)-、-C(O)N(R^N)-、-NR^NC(O)-、-NR^NC(O)N(R^N)-、-C(O)O—、-OC(O)-、-OC(O)O-、-OC(O)N(R^N)-、-NR^NC(O)O-、-C(O)S-、-SC(O)-、-C(=NR^N)-、-C(=NR^N)N(R^N)-、-NR^NC(=NR^N)-、-NR^NC(=NR^N)N(R^N)-、-C(S)-、-C(S)N(R^N)-、-NR^NC(S)-、-NR^NC(S)N(R^N)-、-S(O)-、-OS(O)-、-S(O)O-、-OS(O)O-、-OS(O)$_2$-、-S(O)$_2$O-、-oS(o)$_2$O-、-N(R^N)S(O)-、-S(O)N(R^N)-、-N(R^N)S(O)N(R^N)-、-OS(O)N(R^N)-、-N(R^N)S(O)O-、-S(O)$_2$-、-N(R^N)S(O)$_2$-、-S(O)$_2$N(R^N)-、-N(R^N)S(O)$_2$N(R^N)-、-OS(O)$_2$N(R^N)-或-N(R^N)S(O)$_2$O-置换;并且

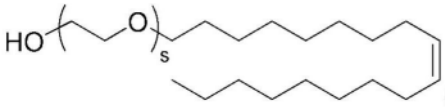
[1015] R^N的每个实例独立地是氢、C_{1-6}烷基或氮保护基。

[1016] 在某个实施方案中,R^{1BRIJ}是C_{18}烷基。举例来说,聚乙二醇醚是式(VIII-a)化合物:

[1017] 
(VIII-a),

[1018] 或其盐或异构体。

[1019] 在一些实施方案中,R^{1BRIJ}是C_{18}烯基。举例来说,聚乙二醇醚是式(VIII-b)化合物:

[1020] 
(VIII-b),

[1021] 或其盐或异构体。

[1022] 在一些实施方案中,泊洛沙姆选自由泊洛沙姆101、泊洛沙姆105、泊洛沙姆108、泊洛沙姆122、泊洛沙姆123、泊洛沙姆124、泊洛沙姆181、泊洛沙姆182、泊洛沙姆183、泊洛沙姆184、泊洛沙姆185、泊洛沙姆188、泊洛沙姆212、泊洛沙姆215、泊洛沙姆217、泊洛沙姆231、泊洛沙姆234、泊洛沙姆235、泊洛沙姆237、泊洛沙姆238、泊洛沙姆282、泊洛沙姆284、泊洛沙姆288、泊洛沙姆331、泊洛沙姆333、泊洛沙姆334、泊洛沙姆335、泊洛沙姆338、泊洛沙姆401、泊洛沙姆402、泊洛沙姆403和泊洛沙姆407组成的组。

[1023] 在一些实施方案中,聚山梨醇酯是Tween®20、Tween®40、Tween®60或Tween®80。

[1024] 在一些实施方案中,脱水山梨糖醇的衍生物是Span®20、Span®60、Span®65、Span®80或Span®85。

[1025] 在一些实施方案中,脂质纳米颗粒中非离子表面活性剂的浓度在约0.00001%w/v至约1%w/v的范围内,例如,约0.00005%w/v至约0.5%w/v,或约0.0001%w/v至约0.1%w/v。

[1026] 在一些实施方案中,脂质纳米颗粒中非离子表面活性剂的浓度在约0.000001重量%至约1重量%的范围内,例如,约0.000002重量%至约0.8重量%,或约0.000005重量%至约0.5重量%。

[1027] 在一些实施方案中,脂质纳米颗粒中PEG脂质的浓度在约0.01摩尔%至约50摩尔%的范围内,例如,约0.05摩尔%至约20摩尔%,约0.07摩尔%至约10摩尔%,约0.1摩尔%至约8摩尔%,约0.2摩尔%至约5摩尔%,或约0.25摩尔%至约3摩尔%。

[1028] 佐剂

[1029] 在一些实施方案中,本发明的LNP任选地包括一种或多种佐剂,例如,吡喃葡萄糖基脂质佐剂(GLA)、CpG寡脱氧核苷酸(例如,A类或B类)、多聚(I:C)、氢氧化铝和Pam3CSK4。

[1030] 其它组分

[1031] 除了前述章节中所描述的那些以外,本发明的LNP可以任选地包括一种或多种组分。举例来说,脂质纳米颗粒可以包括一种或多种小疏水性分子,诸如维生素(例如,维生素A或维生素E)或固醇。

[1032] 脂质纳米颗粒还可以包括一种或多种渗透性增强剂分子、碳水化合物、聚合物、表面改变剂或其它组分。渗透性增强剂分子可以是例如美国专利申请公布号2005/0222064所描述的分子。碳水化合物可以包括单糖(例如,葡萄糖)和多糖(例如,糖原及其衍生物和类似物)。

[1033] 聚合物可以包括于脂质纳米颗粒中和/或用于囊封或部分地囊封脂质纳米颗粒。聚合物可以是生物可降解的和/或生物相容性的。聚合物可以选自但不限于聚胺、聚醚、聚酰胺、聚酯、聚氨基甲酸酯、聚脲、聚碳酸酯、聚苯乙烯、聚酰亚胺、聚砜、聚氨酯、聚乙炔、聚乙烯、聚乙烯亚胺、聚异氰酸酯、聚丙烯酸酯、聚甲基丙烯酸酯、聚丙烯腈和聚丙烯酸酯。举例来说,聚合物可以包括聚(己内酯)(PCL)、乙烯乙酸乙烯酯聚合物(EVA)、聚(乳酸)(PLA)、聚(L-乳酸)(PLLA)、聚(乙醇酸)(PGA)、聚(乳酸-共-乙醇酸)(PLGA)、聚(L-乳酸-共-乙醇酸)(PLLGA)、聚(D,L-丙交酯)(PDLA)、聚(L-丙交酯)(PLLA)、聚(D,L-丙交酯-共-己内酯)、聚(D,L-丙交酯-共-己内酯-共-乙交酯)、聚(D,L-丙交酯-共-PEO-共-D,L-丙交酯)、聚(D,L-丙交酯-共-PPO-共-D,L-丙交酯)、聚氰基丙烯酸烷酯、聚氨酯、聚-L-赖氨酸(PLL)、甲基丙烯酸羟丙酯(HPMA)、聚乙二醇、聚-L-谷氨酸、聚(羟基酸)、聚酸酐、聚原酸酯、聚(酯酰胺)、聚酰胺、聚(酯醚)、聚碳酸酯、聚烯(诸如聚乙烯和聚丙烯)、聚烷二醇(诸如聚(乙二醇)(PEG))、聚氧化烯(PEO)、聚对苯二甲酸烷二酯(诸如聚(对苯二甲酸乙二酯))、聚乙烯醇(PVA)、聚乙烯醚、聚乙烯酯(诸如聚(乙酸乙烯酯))、聚卤化乙烯(诸如聚(氯乙烯)(PVC))、聚乙烯吡咯烷酮(PVP)、聚硅氧烷、聚苯乙烯、聚氨酯、衍生化纤维素(诸如烷基纤维素、羟烷基纤维素、纤维素醚、纤维素酯、硝基纤维素、羟丙基纤维素、羧甲基纤维素)、丙烯酸聚合物(诸如聚((甲基)丙烯酸甲酯)(PMMA)、聚((甲基)丙烯酸乙酯)、聚((甲基)丙烯酸丁酯)、聚((甲基)丙烯酸异丁酯)、聚((甲基)丙烯酸己酯)、聚((甲基)丙烯酸异癸酯)、聚((甲基)丙烯酸月桂酯)、聚((甲基)丙烯酸苯酯)、聚(丙烯酸甲酯)、聚(丙烯酸异丙酯)、聚(丙烯酸异丁酯)、聚(丙烯酸十八烷酯)及其共聚物和混合物)、聚二噁烷酮及其共聚物、聚羟基烷酸

酯、聚富马酸丙二醇酯、聚甲醛、泊洛沙姆、聚氧胺、聚(原酸)酯、聚(丁酸)、聚(戊酸)、聚(丙交酯-共-己内酯)、三亚甲基碳酸酯、聚(N-丙烯酰基吗啉)(PAcM)、聚(2-甲基-2-噁唑啉)(PMOX)、聚(2-乙基-2-噁唑啉)(PEOZ)和聚甘油。

[1034] 表面改变剂可以包括但不限于阴离子蛋白(例如,牛血清白蛋白)、表面活性剂(例如,阳离子表面活性剂,诸如二甲基二(十八烷基)-溴化铵)、糖或糖衍生物(例如,环糊精)、核酸、聚合物(例如,肝素、聚乙二醇和泊洛沙姆)、粘液溶解剂(例如,乙酰基半胱氨酸、艾蒿、菠罗蛋白酶、木瓜蛋白酶、大青属(clerodendrum)、溴己新(bromhexine)、羧甲司坦(carbocisteine)、依普拉酮(eprazinone)、美司钠(mesna)、胺溴索(ambroxol)、索布瑞醇(sobrerol)、多米奥醇(domiodol)、来托司坦(letosteine)、司替罗宁(stepronin)、硫普罗宁(tiopronin)、凝溶胶蛋白、胸腺素β4、阿法链道酶(dornase alfa)、奈替克新(neltenexine)和厄多司坦(erdosteine))和DNase(例如,rhDNase)。表面改变剂可以安置于纳米颗粒内和/或LNP的表面上(例如,通过涂布、吸附、共价键联或其它工艺)。

[1035] 脂质纳米颗粒还可以包含一种或多种官能化脂质。举例来说,脂质可以经炔烃基团官能化,所述炔烃基团当在适当反应条件下暴露于叠氮化物时可以经历环加成反应。特别地,脂质双层可以按这种方式经可用于促进膜渗透、细胞识别或成像的一种或多种基团官能化。LNP的表面还可以与一种或多种可用的抗体缀合。可用于靶向细胞递送、成像和膜渗透的官能团和缀合物在本领域中是众所周知的。

[1036] 除了这些组分以外,脂质纳米颗粒还可以包括可用于药物组合物的任何物质。举例来说,脂质纳米颗粒可以包括一种或多种药学上可接受的赋形剂或辅助成分,诸如但不限于一种或多种溶剂、分散介质、稀释剂、分散助剂、悬浮助剂、造粒助剂、崩解剂、填充剂、助流剂、液体媒剂、粘合剂、表面活性剂、等渗剂、增稠剂或乳化剂、缓冲剂、润滑剂、油、防腐剂和其它物质。还可以包括诸如蜡、黄油、着色剂、涂布剂、调味剂和芳香剂的赋形剂。药学上可接受的赋形剂在本领域中是众所周知的(参见例如Remington's The Science and Practice of Pharmacy,第21版,A.R.Gennaro;Lippincott,Williams&Wilkins,Baltimore,MD,2006)。

[1037] 稀释剂的实例可以包括但不限于碳酸钙、碳酸钠、磷酸钙、磷酸二钙、硫酸钙、磷酸氢钙、磷酸钠、乳糖、蔗糖、纤维素、微晶纤维素、高岭土、甘露糖醇、山梨糖醇、肌醇、氯化钠、干淀粉、玉米淀粉、糖粉和/或它们的组合。造粒剂和分散剂可以选自由马铃薯淀粉、玉米淀粉、木薯淀粉、乙醇酸淀粉钠、粘土、海藻酸、瓜尔胶、柑橘渣、琼脂、膨润土、纤维素和木制品、天然海绵、阳离子交换树脂、碳酸钙、硅酸盐、碳酸钠、交联聚(乙烯-吡咯烷酮)(交联聚维酮)、羧甲基淀粉钠(乙醇酸淀粉钠)、羧甲基纤维素、交联羧甲基纤维素钠(交联羧甲纤维素)、甲基纤维素、预胶凝淀粉(淀粉1500)、微晶淀粉、水不溶性淀粉、羧甲基纤维素钙、硅酸镁铝(VEEGUM®)、月桂基硫酸钠、季铵化合物和/或它们的组合组成的非限制性清单。

[1038] 表面活性剂和/或乳化剂可以包括但不限于天然乳化剂(例如,阿拉伯胶、琼脂、海藻酸、海藻酸钠、黄蓍胶、角叉菜(chondrux)、胆固醇、黄原胶、果胶、明胶、蛋黄、酪蛋白、羊毛脂、胆固醇、蜡和卵磷脂)、胶质粘土(例如,膨润土[硅酸铝]和VEEGUM®[硅酸镁铝])、长链氨基酸衍生物、高分子量醇(例如,硬脂醇、鲸蜡醇、油醇、三乙酸甘油酯单硬脂酸酯、乙二醇二硬脂酸酯、甘油单硬脂酸酯和丙二醇单硬脂酸酯、聚乙烯醇)、卡波姆(carbomer)(例如,羧聚乙烯、聚丙烯酸、丙烯酸聚合物和羧乙烯基聚合物)、角叉菜胶、纤维

素衍生物(例如,羧甲基纤维素钠、粉状纤维素、羟甲基纤维素、羟丙基纤维素、羟丙基甲基纤维素、甲基纤维素)、脱水山梨糖醇脂肪酸酯(例如,聚氧乙烯脱水山梨糖醇单月桂酸酯[TWEEN®20]、聚氧乙烯脱水山梨糖醇[TWEEN® 60]、聚氧乙烯脱水山梨糖醇单油酸酯[TWEEN®80]、脱水山梨糖醇单棕榈酸酯[SPAN®40]、脱水山梨糖醇单硬脂酸酯[SPAN®60]、脱水山梨糖醇三硬脂酸酯[SPAN®65]、甘油单油酸酯、脱水山梨糖醇单油酸酯[SPAN®80])、聚氧乙烯酯(例如,聚氧乙烯单硬脂酸酯[MYRJ® 45]、聚氧乙烯氢化蓖麻油、聚乙氧基化蓖麻油、聚氧亚甲基硬脂酸酯和SOLUTOL®)、蔗糖脂肪酸酯、聚乙二醇脂肪酸酯(例如,CREMOPHOR®)、聚氧乙烯醚(例如,聚氧乙烯月桂基醚[BRIJ® 30])、聚(乙烯-吡咯烷酮)、二乙二醇单月桂酸酯、三乙醇胺油酸酯、油酸钠、油酸钾、油酸乙酯、油酸、月桂酸乙酯、月桂基硫酸钠、PLURONIC®F 68、POLOXAMER®188、西曲溴铵(cetrimonium bromide)、西吡氯铵(cetylpyridinium chloride)、苯扎氯铵(benzalkonium chloride)、多库酯钠(docusate sodium)和/或它们的组合。

[1039] 粘合剂可以是淀粉(例如,玉米淀粉和淀粉糊);明胶;糖(例如,蔗糖、葡萄糖、右旋糖、糊精、糖蜜、乳糖、乳糖醇、甘露糖醇);天然和合成胶(例如,阿拉伯胶、海藻酸钠、爱尔兰苔提取物、潘瓦尔胶(panwar gum)、茄替胶(ghatti gum)、伊莎珀尔果壳(isapol husk)的粘液、羧甲基纤维素、甲基纤维素、乙基纤维素、羟乙基纤维素、羟丙基纤维素、羟丙基甲基纤维素、微晶纤维素、乙酸纤维素、聚(乙烯-吡咯烷酮)、硅酸镁铝(VEEGUM®)和落叶松阿拉伯半乳聚糖);海藻酸盐;聚氧化乙烯;聚乙二醇;无机钙盐;硅酸;聚甲基丙烯酸酯;蜡;水;醇;和它们的组合,或任何其它适合的粘合剂。

[1040] 防腐剂的实例可以包括但不限于抗氧化剂、螯合剂、抗微生物防腐剂、抗真菌防腐剂、醇防腐剂、酸性防腐剂和/或其它防腐剂。抗氧化剂的实例包括但不限于α生育酚、抗坏血酸、抗坏血酸棕榈酸酯、丁基化羟基茴香醚、丁基化羟基甲苯、单硫代甘油、偏亚硫酸氢钾、丙酸、没食子酸丙酯、抗坏血酸钠、亚硫酸氢钠、偏亚硫酸氢钠和/或亚硫酸钠。螯合剂的实例包括乙二胺四乙酸(EDTA)、单水合柠檬酸、依地酸二钠、依地酸二钾、依地酸、富马酸、苹果酸、磷酸、依地酸钠、酒石酸和/或依地酸三钠。抗微生物防腐剂的实例包括但不限于苯扎氯铵、苄索氯铵(benzethonium chloride)、苯甲醇、布罗波尔(bronopol)、西曲溴铵(cetrimide)、西吡氯铵、氯己定(chlorhexidine)、氯丁醇、氯甲酚、氯二甲苯酚、甲酚、乙醇、甘油、海克替啶(hexetidine)、咪脲、苯酚、苯氧乙醇、苯乙醇、硝酸苯汞、丙二醇和/或硫柳汞(thimerosal)。抗真菌防腐剂的实例包括但不限于对羟基苯甲酸丁酯、对羟基苯甲酸甲酯、对羟基苯甲酸乙酯、对羟基苯甲酸丙酯、苯甲酸、羟基苯甲酸、苯甲酸钾、山梨酸钾、苯甲酸钠、丙酸钠和/或山梨酸。醇防腐剂的实例包括但不限于乙醇、聚乙二醇、苯甲醇、苯酚、酚类化合物、双酚、氯丁醇、羟基苯甲酸酯和/或苯乙醇。酸性防腐剂的实例包括但不限于维生素A、维生素C、维生素E、β-胡萝素、柠檬酸、乙酸、去氢抗坏血酸、抗坏血酸、山梨酸和/或植酸。其它防腐剂包括但不限于生育酚、乙酸生育酚、甲磺酸去铁胺、西曲溴铵、丁基化羟基茴香醚(BHA)、丁基化羟基甲苯(BHT)、乙二胺、月桂基硫酸钠(SLS)、月桂基醚硫酸钠(SLES)、亚硫酸氢钠、偏亚硫酸氢钠、亚硫酸钾、偏亚硫酸氢钾、GLYDANTPLUS®、PHENONIP®、对羟基苯甲酸甲酯、GERMALL®115、GERMABEN®II、NEOLONE™、

KATHON™和/或EUXYL®。

[1041] 缓冲剂的实例包括但不限于柠檬酸盐缓冲溶液、乙酸盐缓冲溶液、磷酸盐缓冲溶液、氯化铵、碳酸钙、氯化钙、柠檬酸钙、葡乳醛酸钙、葡庚糖酸钙、葡萄糖酸钙、d-葡萄糖酸、甘油磷酸钙、乳酸钙、乳糖酸钙、丙酸、乙酰丙酸钙、戊酸、磷酸氢钙、磷酸、磷酸三钙、磷酸氢氧化钙、乙酸钾、氯化钾、葡萄糖酸钾、钾混合物、磷酸氢二钾、磷酸二氢钾、磷酸钾混合物、乙酸钠、碳酸氢钠、氯化钠、柠檬酸钠、乳酸钠、磷酸氢二钠、磷酸二氢钠、磷酸钠混合物、缓血酸胺(tromethamine)、氨基-磺酸盐缓冲液(例如,HEPES)、氢氧化镁、氢氧化铝、海藻酸、无热原质水、等渗盐水、林格氏溶液(Ringer's solution)、乙醇和/或它们的组合。润滑剂可以选自由硬脂酸镁、硬脂酸钙、硬脂酸、二氧化硅、滑石、麦芽、甘油山嵛酸酯、氢化植物油、聚乙二醇、苯甲酸钠、乙酸钠、氯化钠、亮氨酸、月桂基硫酸镁、月桂基硫酸钠和它们的组合组成的非限制性组。

[1042] 油的实例包括但不限于扁桃仁、杏仁、鳄梨、巴西棕榈、佛手柑、黑加仑籽、琉璃苣、刺桧、甘菊、芥花、香菜、巴西棕榈、蓖麻、肉桂、可可脂、椰子、鱼肝、咖啡、玉米、棉籽、鸸鹋、桉树、月见草、鱼、亚麻仁、香草醇、葫芦、葡萄籽、榛子、海索草、肉豆蔻酸异丙酯、荷荷巴、夏威夷核果、熏衣草花、熏衣草、柠檬、山苍子、澳洲坚果、锦葵、芒果核、池花籽、貂、肉豆蔻、橄榄、橙、橙鲷、棕榈、棕榈仁、桃仁、花生、罂粟籽、南瓜籽、油菜籽、米糠、迷迭香、红花、白檀、山茶花、欧洲薄荷、沙棘、芝麻、乳木果油、硅酮、大豆、向日葵、茶树、蓟、椿花、香根草、胡桃和小麦胚芽油以及硬脂酸丁酯、辛酸三酸甘油酯、癸酸三酸甘油酯、环甲基硅酮、癸二酸二乙酯、二甲硅油360、西甲硅油、肉豆蔻酸异丙酯、矿物油、辛基十二醇、油醇、硅酮油和/或它们的组合。

[1043] 额外的和示例性脂质纳米颗粒和化合物公开于2020年9月18日提交的国际申请号PCT/US2020/051609中,其全部内容特此以引用的方式并入。

[1044] 17.使用LNP组合物的方法

[1045] 本公开提供了LNP组合物,所述组合物可以例如体外或体内递送至细胞,例如靶细胞。对于体外蛋白质表达,通过离体孵育LNP和细胞使细胞与LNP接触。随后可将此类细胞引入体内。对于体内蛋白质表达,通过向受试者施用LNP使细胞与LNP接触,从而增加或诱导受试者内的细胞中或细胞上的蛋白质表达。举例来说,在一个实施方案中,静脉内施用LNP。在另一个实施方案中,肌肉内施用LNP。在又一个实施方案中,通过选自由皮下、结内和瘤内组成的组的途径施用LNP。

[1046] 对于体外递送,在一个实施方案中,通过离体孵育LNP和靶细胞使细胞与LNP接触。在一个实施方案中,细胞是人类细胞。已经证明各种类型的细胞可以被LNP转染。

[1047] 在另一个实施方案中,细胞与LNP接触例如至少30分钟、至少1小时、至少2小时、至少3小时、至少4小时、至少5小时、至少6小时、至少12小时或至少24小时。

[1048] 在一个实施方案中,细胞与LNP接触持续单次治疗/转染。在另一个实施方案中,细胞与LNP接触持续多次治疗/转染(例如,相同细胞之两次、三次、四次或更多次治疗/转染)。

[1049] 在另一个实施方案中,对于体内递送,通过向受试者施用LNP使细胞与LNP接触,从而将核酸递送至受试者内的细胞。举例来说,在一个实施方案中,静脉内施用LNP。在另一个实施方案中,肌肉内施用LNP。在又一个实施方案中,通过选自由皮下、结内和瘤内组成的组的途径施用LNP。

[1050] 在一个方面,本文提供了一种增加细胞中的治疗性有效载荷或预防性有效载荷的表达的方法,所述方法包括向所述细胞施用本文所公开的LNP组合物。

[1051] 在一个相关方面,本文提供了一种LNP组合物,所述组合物用于增加细胞中的治疗性有效载荷或预防性有效载荷的表达的方法中。

[1052] 在另一个方面,本公开提供了一种增加受试者中的治疗性有效载荷或预防性有效载荷的表达的方法,所述方法包括向所述受试者施用有效量的本文所公开的LNP组合物。

[1053] 在一个相关方面,本文提供了一种LNP组合物,所述组合物用于增加受试者中的治疗性有效载荷或预防性有效载荷的表达的方法中。

[1054] 在又一个方面,本文提供了一种递送本文所公开的LNP组合物的方法。

[1055] 在一个相关方面,本文提供了一种LNP组合物,所述组合物用于将LNP组合物递送至细胞的方法中。

[1056] 在一个实施方案中,方法或用途包括体外、体内或离体使细胞与LNP组合物接触。

[1057] 在一个实施方案中,本公开的LNP组合物与细胞例如离体或体内接触,并且可以用于将分泌多肽、细胞内多肽或跨膜多肽递送至受试者。

[1058] 在一个方面,本公开提供了一种将本文所公开的LNP组合物递送至患有例如本文所描述的疾病或病症的受试者的方法。

[1059] 在一个相关方面,本文提供了一种LNP组合物,所述组合物用于将LNP组合物递送至患有例如本文所描述的疾病或病症的受试者的方法中。

[1060] 在另一个方面,本文提供了一种调节受试者的免疫反应的方法,所述方法包括向有需要的受试者施用有效量的本文所公开的LNP组合物。

[1061] 在一个相关方面,本文提供了一种LNP组合物,所述组合物用于调节受试者中的免疫反应的方法中,所述方法包括向所述受试者施用有效量的LNP组合物。

[1062] 在另一个方面,本文提供了一种将分泌多肽、细胞内多肽或跨膜多肽递送至受试者的方法。

[1063] 在一个方面,本文提供了一种治疗、预防疾病或病症或预防疾病或病症的症状的方法,所述方法包括向有需要的受试者施用有效量的本文所公开的LNP组合物。

[1064] 在一个相关方面,本文提供了一种LNP组合物,所述组合物用于治疗、预防受试者的疾病或病症或预防疾病或病症的症状的方法中,所述方法包括向有需要的受试者施用有效量的LNP组合物。

[1065] 在一些实施方案中,供使用的方法或组合物使得编码治疗性有效载荷或预防性有效载荷的mRNA的表达和/或水平增加。

[1066] 在一些实施方案中,供使用的方法或组合物使得编码治疗性有效载荷或预防性有效载荷的mRNA的表达和/或水平持续。

[1067] 在一些实施方案中,供使用的方法或组合物使得治疗性有效载荷或预防性有效载荷的表达和/或水平增加。

[1068] 在一些实施方案中,供使用的方法或组合物使得治疗性有效载荷或预防性有效载荷的表达和/或水平持续。

[1069] 在一些实施方案中,将本文所描述的功能效应中的任一者与以下细胞进行比较:

[1070] (a)尚未与本文所公开的LNP组合物接触;或者

[1071] (b)尚未与包含多核苷酸的LNP接触,所述多核苷酸包含本文所描述的5′UTR、本文所描述的3′UTR和/或包含本文所描述的终止元件的编码区。

[1072] 18.组合疗法

[1073] 在一些实施方案中,本文所公开的供使用的治疗方法或组合物包括施用本文所公开的LNP与额外剂的组合。在一个实施方案中,额外剂是疾病或病症,例如自身免疫疾病的标准护理。在一个实施方案中,额外剂是mRNA。

[1074] 在一些方面,本发明方法或组合物的受试者已经用一种或多种标准护理疗法治疗。在其它方面,本发明方法或组合物的受试者对一种或多种标准护理疗法或抗癌疗法没有反应。

[1075] 19.药物组合物

[1076] 本公开提供了包含本文所公开的LNP组合物中的任一者的药物制剂。

[1077] 在本公开的一些实施方案中,多核苷酸与一种或多种药学上可接受的赋形剂组合配制于组合物和复合物中。药物组合物可以任选地包含一种或多种额外的活性物质,例如,治疗和/或预防活性物质。本公开的药物组合物可以是无菌的和/或无热原的。药剂的配制和/或制造的一般准则可以见于例如Remington:The Science and Practice of Pharmacy,第21版,Lippincott Williams&Wilkins,2005中。

[1078] 在一些实施方案中,向人类、人类患者或受试者施用组合物。出于本公开的目的,短语“活性成分”一般是指如本文所描述而递送的多核苷酸。

[1079] 尽管本文所提供的药物组合物的描述主要针对适合于向人类施用的药物组合物,但熟练的技术人员将了解,此类组合物一般适合于向任何其它动物,例如非人类动物,例如非人类哺乳动物施用。充分了解到对适合于向人类施用的药物组合物进行修饰以使所述组合物适合于向各种动物施用,并且普通兽医药理学家可以设计和/或仅通过普通实验(如果有的话)进行此种修饰。药物组合物预期施用的受试者包括但不限于人类和/或其它灵长类动物;哺乳动物。

[1080] 在一些实施方案中,本公开的多核苷酸被配制成用于皮下、静脉内、腹膜内、肌肉内、关节内、滑膜内、胸骨内、鞘内、肝内、病灶内、颅内、心室内、经口、吸入喷雾、经肺、局部、经直肠、经鼻、经颊、经阴道或植入式储库肌肉内、皮下或皮内递送。在其它实施方案中,多核苷酸被配制成用于皮下或静脉内递送。

[1081] 本文所描述的药物组合物的制剂可以通过药理学领域中已知的或以后开发的任何方法来制备。一般来说,此类制备方法包括以下步骤:使活性成分与赋形剂和/或一种或多种其它辅助成分缔合,然后必要或需要时,将产物分割、成形和/或包装成所需的单剂量或多剂量单位。

[1082] 根据本公开的药物组合物中的活性成分、药学上可接受的赋形剂和/或任何附加成分的相对量将取决于所治疗的受试者的身份、体型和/或状况并且进一步取决于组合物的施用途径而变化。举例来说,组合物可以包含介于0.1%与100%之间,例如介于0.5%与50%之间、介于1%与30%之间、介于5%与80%之间或至少80%(w/w)的活性成分。

[1083] 20.制剂和递送

[1084] 可以使用一种或多种赋形剂来配制包含本公开的rnRNA的多核苷酸。

[1085] 一种或多种赋形剂的功能是例如:(1)增加稳定性;(2)增加细胞转染;(3)允许持

续或延迟释放(例如,从多核苷酸的贮库制剂中);(4)改变生物分布(例如,将多核苷酸靶向特定组织或细胞类型);(5)增加体内编码蛋白的翻译;和/或(6)改变体内编码蛋白的释放型态。除了诸如任何和所有溶剂、分散介质、稀释剂或其它液体媒剂、分散或悬浮助剂、表面活性剂、等渗剂、增稠剂或乳化剂、防腐剂的传统赋形剂以外,本公开的赋形剂还可以包括但不限于类脂质、脂质体、脂质纳米颗粒、聚合物、脂质复合物、核-壳纳米颗粒、肽、蛋白质、用多核苷酸转染的细胞(例如,用于移植至受试者中)、透明质酸酶、纳米颗粒模拟物和它们的组合。因此,本公开的制剂可以包括一种或多种赋形剂,每种赋形剂的量共同增加多核苷酸的稳定性,增加多核苷酸的细胞转染,增加多核苷酸编码的蛋白质的表达,和/或改变多核苷酸编码的蛋白质的释放型态。此外,本公开的多核苷酸可以使用自组装的核酸纳米颗粒来配制。

[1086] 本文所描述的药物组合物的制剂可以通过药理学领域中已知的或以后开发的任何方法来制备。一般来说,此类制备方法包括将活性成分与赋形剂和/或一种或多种其它辅助成分缔合的步骤。

[1087] 根据本公开的药物组合物可以大批、以单一单位剂量和/或以多个单一单位剂量来制备、包装和/或销售。如本文所用,“单位剂量”是指包含预定量的活性成分的药物组合物的个别量。活性成分的量一般等于将向受试者使用的活性成分的剂量,和/或此种剂量的适宜分率,诸如此种剂量的一半或三分之一。

[1088] 根据本公开的药物组合物中活性成分、药学上可接受的赋形剂和/或任何附加成分的相对量可根据所治受试者的特性、大小和/或状况,并进一步根据施用组合物的途径而变化。举例来说,组合物可以包含介于0.1%与99%(w/w)之间的活性成分。举例来说,组合物可以包含介于0.1%与100%之间,例如介于0.5%与50%之间、介于1-30%之间、介于5-80%之间、至少80%(w/w)的活性成分。

[1089] 在一些实施方案中,本文所描述的制剂含有至少一种多核苷酸。作为一个非限制性实例,制剂含有1、2、3、4或5种多核苷酸。

[1090] 药物制剂另外可以包含药学上可接受的赋形剂,如本文所用的赋形剂包括但不限于对所需特定剂型适合的任何和所有的溶剂、分散介质、稀释剂或其它液体媒剂、分散或悬浮助剂、表面活性剂、等渗剂、增稠剂或乳化剂、防腐剂等。用于配制药物组合物的各种赋形剂和用于制备组合物的技术在本领域中是已知的(参见Remington:The Science and Practice of Pharmacy,第21版,A.R.Gennaro,Lippincott,Williams&Wilkins,Baltimore,MD,2006)。除非任何常规赋形剂介质可能与物质或其衍生物不相容,诸如通过产生任何不合需要的生物作用或在其它方面以有害方式与药物组合物的任何其它组分相互作用,否则可以预期常规赋形剂介质的使用处于本公开的范围内。

[1091] 在一些实施方案中,脂质纳米颗粒的粒度增加和/或减小。粒度的变化能够帮助对抗生物反应,诸如但不限于炎症,或者可以增加递送至哺乳动物的修饰mRNA的生物作用。

[1092] 用于制造药物组合物的药学上可接受的赋形剂包括但不限于惰性稀释剂、表面活性剂和/或乳化剂、防腐剂、缓冲剂、润滑剂和/或油。此类赋形剂可以任选地包含于本公开的药物制剂中。

[1093] 在一些实施方案中,多核苷酸在纳米结构中施用或与纳米结构一起施用,配制于纳米结构中或与纳米结构一起递送,所述纳米结构可以螯合诸如胆固醇的分子。这些纳米

结构的非限制性实例和制造这些纳米结构的方法描述于美国专利公布号US20130195759中。这些纳米结构的示例性结构示于美国专利公布号US20130195759中,并且可以包括核和包围核的壳。

[1094] 可以使用本领域中已知的任何方法将包含本公开的mRNA的多核苷酸递送至细胞。举例来说,包含本公开的mRNA的多核苷酸可以通过基于脂质的递送(例如,转染)或通过电穿孔而递送至细胞。

[1095] 21.定义

[1096] 为了更容易地理解本公开,首先对某些术语进行了定义。如在本申请中所用,除非本文另外明确地提供,否则以下术语中的每一个应该具有下面阐述的含义。额外定义阐述于整个申请中。

[1097] 本发明包括如下实施方案,其中确切地组的一个成员存在于、用于给定产物或工艺中或以其它方式与给定产物或工艺相关。本发明包括如下实施方案,其中组成员中的超过一者或全部存在于、用于给定产物或工艺中或以其它方式与给定产物或工艺相关。

[1098] 在本说明书和所附权利要求书中,除非上下文另外明确指出,否则单数形式“一个”、“一种”和“所述”包括复数指示物。术语“一个”(或“一种”)以及术语“一个或多个”和“至少一个”在本文可以互换使用。在某些方面,术语“一个”或“一种”意指“单个”。在其它方面,术语“一个”或“一种”包括“两个或更多个”或“多个”。

[1099] 此外,当在本文中使用时,“和/或”应当视为具有或不具有另一者的两种特定特征或组分的每一者的具体公开。因此,本文在短语中使用的术语“和/或”诸如“A和/或B”意为包括:“A和B”;“A或B”;“A”(单独);和“B”(单独)。同样地,在诸如“A、B和/或C”的短语中使用的术语“和/或”旨在涵盖以下方面中的每一个:A、B和C;A、B或C;A或C;A或B;B或C;A和C;A和B;B和C;A(单独);B(单独);和C(单独)。

[1100] 除非另有定义,否则本文使用的所有技术和科技术语具有与本公开相关领域的普通技术人员通常所理解的相同含义。举例来说,Concise Dictionary of Biomedicine and Molecular Biology,Juo,Pei-Show,第2版,2002,CRC Press;Dictionary of Cell and Molecular Biology,第3版,1999,Academic Press;和Oxford Dictionary Of Biochemistry And Molecular Biology,修订板,2000,Oxford University Press,向普通技术人员提供本公开中使用的许多术语的通用字典。

[1101] 当本文中用语言“包含”描述方面时,还提供了根据“由.....组成”和/或“基本上由.....组成”所描述的其它类似方面。

[1102] 单位、前缀和符号以其国际单位制(Système International de Unites,SI)可接受的形式表示。数值范围将限定范围的数字包括在内。在列举值的范围的情况下,应当理解,还具体公开了在该范围的所列举的上限与下限之间的每个居间整数值及其每个分数,以及在此类值之间的每个子范围。任何范围的上限和下限可以独立地包括在该范围内或从该范围中排除,并且其中包括任一个限值、不包括两个限值或包括两个限值的每个范围也涵盖在本发明内。在明确列举值的情况下,应理解作为与所列举的值大约相同的数量或量的值也在本发明的范围内。在公开组合的情况下,该组合的要素的每个子组合也被具体公开并且在本发明的范围内。相反,在单独公开不同要素或要素的组的情况下,也公开了它们的组合。在将本发明的任何元素公开为具有多个替代的情况下,由此还公开了本发明的实

例,其中单独地或以与其它替代方案的任何组合形式排除每个替代方案;本发明的多于一个元素可以具有此类排除,并且由此公开了具有此类排除的元素的所有组合。

[1103] 核苷酸由其普遍接受的单字母码表示。除非另有指示,否则核酸是以5′至3′取向自左至右书写。核碱基在本文中以其由IUPAC-IUB生物化学命名委员会推荐的众所周知的单字母符号提及。因此,A代表腺嘌呤,C代表胞嘧啶,G代表鸟嘌呤,T代表胸腺嘧啶,U代表尿嘧啶。

[1104] 氨基酸可以在本文中由其通常已知的三字母符号或由IUPAC-IUB生物化学命名委员会推荐的单字母符号来表示。除非另有指示,否则氨基酸序列以氨基到羧基的取向从左到右书写。

[1105] 约:在整个说明书和权利要求书中连同数值使用的术语“约”表示本领域技术人员熟悉和可接受的精确度区间,此种精确度区间为±10%。

[1106] 在给出范围的情况下,包括端点。此外,除非另有指示或另外从上下文和本领域普通技术人员的理解中显而易见,除非上下文另外清楚地指示,否则表示为范围的值可以假定在本发明的不同实施方案中规定范围内的任何具体值或子范围,至该范围下限的单位的十分之一。

[1107] 尿苷含量:术语“尿苷含量”或“尿嘧啶含量”是可互换的并且是指存在于某个核酸序列中的尿嘧啶或尿苷的量。尿苷含量或尿嘧啶含量可以表示为绝对值(序列中尿苷或尿嘧啶的总数)或相对值(相对于核酸序列中核碱基总数的尿苷或尿嘧啶百分比)。

[1108] 终止元件.如本文所用的术语“终止元件”是指包含终止密码子的核酸序列。在DNA的情况下,终止密码子可以选自TGA、TAA和TAG;或在RNA的情况下,选自UGA、UAA和UAG。在一个实施方案中,终止元件包含两个连续的终止密码子。在一个实施方案中,终止元件包含三个连续的终止密码子。在一个实施方案中,终止元件包含四个连续的终止密码子。在一个实施方案中,终止元件包含五个连续的终止密码子。在一个实施方案中,终止元件在一个或多个终止密码子的上游和/或下游还包含至少1、2、3、4、5、6、7、8、9、10、11或12个核苷酸。

[1109] 3′稳定区.如本文所用的术语“3′稳定区”是指形成或变得稳定的区域。3′稳定区可以存在于核酸序列的3′末端。在一个实施方案中,3′稳定区包含多聚A尾,例如,如本文所描述。在一个实施方案中,3′稳定区包含替代性核苷,例如,反向胸苷。

[1110] 序列同一性:序列之间的序列同一性的计算可以如下进行。为了确定两个核苷酸序列的同一性百分比,出于最佳比较目的将序列进行比对(例如,可以在第一核苷酸序列和第二核苷酸序列的一者或两者中引入空位以供最佳比对)。在一些实施方案中,为比较目的而比对的参考序列的长度为参考序列长度的至少50%,例如,至少60%、70%、80%、90%或100%。比较相应核苷酸位置处的核苷酸。当第一个序列中的一个位置被与第二个序列中相应位置处相同的核苷酸占据时,则分子在那个位置处是相同的。两个序列之间的同一性百分比是序列共享的相同位置数的函数,这考虑到空位数和每个空位的长度,需要引入空位以供两个序列的最佳比对。同一性百分比通常是指匹配残基数与比对总长度的比率。可以使用数学算法实现序列的比较和两个序列之间的同一性百分比的确定。在一些实施方案中,使用成对序列比对程序或多序列比对程序确定两个核苷酸序列之间的同一性百分比。示例性序列比对程序包括但不限于lalign程序(embnet.vital-it.ch;Huang和Miller,(1991)Adv.Appl.Math.12:337-357);Clustal Omega程序(www.ebi.ac.uk;Sievers等人

(2011)Mol.Syst.Biol.7:539)。在一些实施方案中,使用程序的默认参数。本文所描述的核苷酸序列可以用作“查询序列”以针对公共数据库执行搜索以例如鉴定其它家族成员或相关序列。可以使用BLAST®程序(blast.ncbi.nlm.nhi.gov;Altschul等人,(1990)J.Mol.Biol.215:403-10)执行此类搜索。举例来说,可以用blastn程序执行BLAST核苷酸搜索以获得与本文所描述的核苷酸序列相同或相似的核苷酸序列。在一些实施方案中,使用程序的默认参数。

[1111] 替代性核苷.如本文所用的术语“替代性核苷”在提及核苷酸、核苷或多核苷酸(诸如本发明的多核苷酸,例如mRNA分子)时,是指关于A、G、U或C核糖核苷酸的改变。一般来说,在本文中,这些术语并非旨在指代天然存在的5′末端mRNA帽部分中的核糖核苷酸改变。改变可以是各种不同的改变。在一些实施方案中,在多核苷酸是mRNA的情况下,编码区、侧翼区和/或末端区(例如,3′-稳定区)可以含有一个、两个或更多个(任选不同的)核苷或核苷酸改变。在一些实施方案中,与未改变的多核苷酸相比,引入细胞中的替代性多核苷酸可以展现出细胞中减少的降解。

[1112] 施用:如本文所用,“施用”是指将组合物递送至受试者或患者的方法。可以选择施用方法以靶向递送(例如,特异性地递送)至身体的特定区域或系统。举例来说,施用可以是肠道外(例如,皮下、皮内、静脉内、腹膜内、肌肉内、关节内、动脉内、滑膜内、胸骨内、鞘内、病灶内或颅内注射,以及任何适合的输注技术)、经口、经皮或皮内、皮间、经直肠、阴道内、局部(例如,通过粉末、软膏、乳膏、凝胶、洗剂和/或滴剂)、经粘膜、经鼻、经颊、经肠、玻璃体、肿瘤内、舌下、鼻内;通过气管内滴注、支气管滴注和/或吸入;作为口腔喷雾和/或粉末、鼻喷雾和/或气雾剂,和/或经由门静脉导管。优选的施用方式是静脉内或皮下。

[1113] 抗体分子:在一个实施方案中,抗体分子可以用于靶向所需的细胞类型。如本文所用,“抗体分子”是指天然存在的抗体、工程化抗体或其片段,例如,天然存在的抗体或工程化抗体的抗原结合部分。抗体分子可以包括例如抗体或其抗原结合片段(例如,Fab、Fab’、F(ab′)2、Fv片段、scFv抗体片段、二硫键连接的Fv(sdFv)、由VH和CH1结构域组成的Fd片段、线性抗体、单结构域抗体(诸如sdAb(VL或VH))、纳米抗体或骆驼科VHH结构域)、抗原结合纤连蛋白III型(Fn3)支架(诸如纤连蛋白多肽微抗体)、配体、细胞因子、趋化因子或T细胞受体(TCR)。示例性抗体分子包括但不限于人源化抗体分子;完整IgA、IgG、IgE或IgM抗体;双特异性或多特异性抗体(例如,Zybodies®等);抗体片段,诸如Fab片段、Fab’片段、F(ab′)2片段、Fd’片段、Fd片段和分离的CDR或它们的集合;单链Fv;多肽-Fc融合体;单结构域抗体(例如鲨鱼单结构域抗体,诸如IgNAR或其片段);骆驼抗体;掩蔽抗体(例如,Probodies®);小型模块化免疫药物(“SMIPsTM”);单链或串联双抗体(TandAb®);VHH;Anticalins®;Nanobodies®;微抗体;BiTE®;锚蛋白重复蛋白或DARPINs®;Avimers®;DART;类TCR抗体;Adnectins®;Affilins®;Trans-bodies®;Affibodies®;TrimerX®;MicroProteins;Fynomers®;Centyrins®;和KALBITOR®。

[1114] 接触:如本文所用,术语“接触”意指在两个或更多个实体之间建立物理连接。举例来说,使细胞与mRNA或脂质纳米颗粒组合物接触意味着使细胞和mRNA或脂质纳米颗粒共享

物理连接。体内、体外和离体使细胞与外部实体接触的方法是生物学领域中众所周知的。在本公开的示例性实施方案中,使哺乳动物细胞与组合物(例如,本公开的纳米颗粒或药物组合物)接触的步骤在体内进行。举例来说,可以通过任何适合的施用途径(例如,肠道外施用于生物体,包括静脉内、肌肉内、皮内和皮下施用)使脂质纳米颗粒组合物和可以位于生物体(例如,哺乳动物)内的细胞(例如,哺乳动物细胞)接触。对于存在于体外的细胞,可以例如通过将组合物添加至细胞的培养基中使组合物(例如,脂质纳米颗粒)和细胞接触,并且可以涉及或引起转染。此外,纳米颗粒组合物可以接触多于一个细胞。

[1115] 递送:如本文所用,术语"递送"意指向标的物提供实体。举例来说,将治疗剂和/或预防剂递送至受试者可以涉及向所述受试者施用包括所述治疗剂和/或预防剂的LNP(例如,通过静脉内、肌肉内、皮内、经肺或皮下途径)。向哺乳动物或哺乳动物细胞施用LNP可以涉及使一个或多个细胞与脂质纳米颗粒接触。

[1116] 有效量:如本文所用,剂的术语"有效量"是足以实现有益或期望的结果,例如临床结果的量,并且因此"有效量"取决于所应用的情形。举例来说,在本公开的脂质组合物(例如,LNP)中靶细胞递送增强脂质的量的情形中,靶细胞递送增强脂质的有效量是足以实现与缺乏靶细胞递送增强脂质的脂质组合物(例如,LNP)相比有益或期望结果的量。由脂质组合物(例如,LNP)实现的有益或期望结果的非限制性实例包括增加转染细胞的百分比和/或增加由与脂质组合物(例如,LNP)相关/由脂质组合物(例如,LNP)囊封的核酸编码的蛋白质的表达水平。在施用含靶细胞递送增强脂质的脂质纳米颗粒以使有效量的脂质纳米颗粒被受试者中的靶细胞摄取的情形中,含靶细胞递送增强脂质的LNP的有效量是足以实现与缺乏靶细胞递送增强脂质的LNP相比有益或期望结果的量。受试者中的有益或期望结果的非限制性实例包括与缺乏靶细胞递送增强脂质的LNP相比,增加转染细胞的百分比,增加由与含靶细胞递送增强脂质的LNP相关/由所述LNP囊封的核酸编码的蛋白质的表达水平,和/或增加与含靶细胞递送增强脂质的LNP相关和/或由所述LNP囊封的核酸或其编码的蛋白质在体内的预防或治疗作用。在一些实施方案中,当向罹患或易患感染、疾病、病症和/或疾患的受试者施用时,含靶细胞递送增强脂质的LNP的治疗有效量足以治疗所述感染、疾病、病症和/或疾患,改善其症状,诊断、预防所述感染、疾病、病症和/或疾患,和/或延迟其发作。在另一个实施方案中,脂质纳米颗粒的有效量足以使所需蛋白质在至少约5%、10%、15%、20%、25%或更多的靶细胞中表达。举例来说,含靶细胞递送增强脂质的LNP的有效量可以是在单次静脉注射后使至少5%、10%、15%、20%、25%、30%或35%的靶细胞转染的量。

[1117] 表达:如本文所用,核酸序列的"表达"是指以下事件中的一者或多者:(1)从DNA序列产生RNA模板(例如,通过转录);(2)RNA转录物的加工(例如,通过剪接、编辑、5′帽形成和/或3′端加工);(3)将RNA翻译成多肽或蛋白质;以及(4)多肽或蛋白质的翻译后修饰。

[1118] 离体:如本文所用,术语"离体"是指事件在生物体(例如,动物、植物或微生物或其细胞或组织)外部发生。离体事件可以在从天然(例如,体内)环境最低程度地改变的环境中发生。

[1119] 片段:如本文所用,"片段"是指一部分。举例来说,蛋白质的片段可以包括通过消化从培养细胞中分离的全长蛋白质获得或通过重组DNA技术获得的多肽。蛋白质的片段可以是例如蛋白质的一部分,包括一个或多个功能结构域,使得蛋白质的片段保留蛋白质的功能活性。

[1120] 异源:如本文所用,“异源”指示序列(例如,氨基酸序列或编码氨基酸序列的多核苷酸)通常不存在于给定的多肽或多核苷酸中。举例来说,对应于一种蛋白质的结构域或基序的氨基酸序列可能与第二种蛋白质异源。

[1121] 分离的:如本文所用,术语“分离的”是指已经与至少一些与其相关的组分(无论是在自然界中还是在实验环境中)分离的物质或实体。相对于与其相关的物质,分离的物质可能具有不同水平的纯度。分离的物质和/或实体可以与至少约10%、约20%、约30%、约40%、约50%、约60%、约70%、约80%、约90%或更多的最初与其相关的其它组分分离。在一些实施方案中,分离的剂是超过约80%、约85%、约90%、约91%、约92%、约93%、约94%、约95%、约96%、约97%、约98%、约99%或超过约99%纯的。如本文所用,如果物质基本上不含其它组分,则这种物质是“纯的”。

[1122] Kozak序列:术语“Kozak序列”(也称为“Kozak共有序列”)是指用于增强基因或开放阅读框的表达并且在真核生物中位于5′UTR中的翻译起始增强子元件。在分析起始密码子(AUG)周围的单个突变对前胰岛素原基因翻译的影响(Kozak(1986)Cell 44:283-292)之后,将Kozak共有序列最初定义为序列GCCRCC(SEQ ID NO:43),其中R=嘌呤。本文所公开的多核苷酸包含Kozak共有序列或其衍生物或修饰。(翻译增强子组合物及其使用方法的实例,参见Andrews等人的美国专利号5,807,707,以全文引用的方式并入本文中;Chernajovsky的美国专利号5,723,332,以全文引用的方式并入本文中;Wilson的美国专利号5,891,665,以全文引用的方式并入本文中。)

[1123] 脂质体:如本文所用,“脂质体”意指包括包封水性内部的含脂质膜的结构。脂质体可以具有一个或多个脂质膜。脂质体包括单层脂质体(在本领域中也称为单室脂质体)和多层脂质体(在本领域中也称为多室脂质体)。

[1124] 修饰的:如本文所用,“修饰的”是指本公开的分子的改变的状态或结构,例如,多核苷酸(例如,mRNA)的组成或结构的改变。分子,例如多核苷酸,可以按各种方式进行修饰,包括化学上、结构上和/或功能上。举例来说,分子,例如多核苷酸,可以通过并入一个或多个RNA元件在结构上进行修饰,其中RNA元件包含提供一种或多种功能(例如,翻译调控活性)的序列和/或一个或多个RNA二级结构。因此,本公开的分子,例如多核苷酸,可以由一个或多个修饰组成(例如,可以包括一个或多个化学、结构或功能修饰,包括它们的任何组合)。在一个实施方案中,通过引入非天然核苷和/或核苷酸来修饰本公开的多核苷酸,例如mRNA分子,例如,因为它与天然核糖核苷酸A、U、G和C相关。非规范核苷酸(诸如帽结构)不应视为“修饰的”,尽管它们与A、C、G、U核糖核苷酸的化学结构不同。

[1125] mRNA:如本文所用,“mRNA”是指信使核糖核酸。mRNA可以是天然或非天然存在的。举例来说,mRNA可以包括修饰的和/或非天然存在的组分,诸如一种或多种核碱基、核苷、核苷酸或连接子。mRNA可以包括帽结构、链终止核苷、茎环、多聚A序列和/或多聚腺苷酸化信号。mRNA可以具有编码多肽的核苷酸序列。mRNA的翻译,例如哺乳动物细胞内的mRNA的体内翻译,可以产生多肽。传统上,mRNA分子的基本组分至少包括编码区、5′-非翻译区(5′-UTR)、3'UTR、5′帽和多聚A序列。在一个实施方案中,mRNA是环状mRNA。

[1126] 纳米颗粒:如本文所用,“纳米颗粒”是指在小于约1000nm的标度上具有任何一种结构特征的颗粒,所述颗粒与相同材料的大块样品相比展现出新颖特性。常规地,纳米颗粒在小于约500nm、小于约200nm或约100nm的标度上具有任何一种结构特征。同样常规地,纳

米颗粒在约50nm至约500nm、约50nm至约200nm或约70至约120nm的标度上具有任何一种结构特征。在示例性实施方案中,纳米颗粒是具有约1-1000nm量级的一个或多个尺寸的颗粒。在其它示例性实施方案中,纳米颗粒是具有约10-500nm量级的一个或多个尺寸的颗粒。在其它示例性实施方案中,纳米颗粒是具有约50-200nm量级的一个或多个尺寸的颗粒。球形纳米颗粒将具有例如介于约50-100或70-120纳米之间的直径。就其转运和特性来说,纳米颗粒最常表现为一个单元。值得注意的是,将纳米颗粒与相应的大块材料相区分的新颖特性通常以1000nm以下的尺寸标度或以约100nm的尺寸发展,但纳米颗粒可以具有更大的尺寸,例如,对于长椭圆形、管状等颗粒。尽管大多数分子的尺寸将符合上述概况,但个别分子通常不被称为纳米颗粒。

[1127] 核酸:如本文所用,术语“核酸”以其最广泛的含义使用并且涵盖包括核苷酸聚合物的任何化合物和/或物质。这些聚合物通常被称为多核苷酸。本公开的示例性核酸或多核苷酸包括但不限于核糖核酸(RNA)、脱氧核糖核酸(DNA)、DNA-RNA杂合体、RNAi诱导剂、RNAi剂、siRNA、shRNA、miRNA、反义RNA、核酶、催化DNA、诱导三螺旋形成的RNA、苏糖核酸(TNA)、乙二醇核酸(GNA)、肽核酸(PNA)、锁核酸(LNA,包括具有β-D-核糖构型的LNA、具有α-L-核糖构型的α-LNA(LNA的非对映异构体)、具有2′-氨基官能化的2′-氨基-LNA和具有2′-氨基官能化的2′-氨基-α-LNA)或它们的杂合体。

[1128] 核碱基:如本文所用,术语“核碱基”(或者“核苷酸碱基”或“含氮碱基”)是指在核酸中发现的嘌呤或嘧啶杂环化合物,包括向核酸或其部分或区段赋予改善的特性(例如,结合亲和力、核酸酶抗性、化学稳定性)的天然存在的嘌呤和嘧啶的任何衍生物或类似物。腺嘌呤、胞嘧啶、鸟嘌呤、胸腺嘧啶和尿嘧啶是主要在天然核酸中发现的核碱基。如本领域中已知和/或本文所描述的其它天然、非天然和/或合成核碱基可以并入核酸中。除非另有指定,否则本文所描述的SEQ ID NO的核碱基序列涵盖天然核碱基和化学修饰的核碱基(例如,SEQ ID NO中的“U”名称涵盖尿嘧啶和化学修饰的尿嘧啶)。

[1129] 核苷/核苷酸:如本文所用,术语“核苷”是指含有共价连接至核碱基(例如,嘌呤或嘧啶)或其衍生物或类似物(本文中也称为“核碱基”),但缺少核苷间连接基团(例如,磷酸基团)的糖分子(例如,RNA中的核糖或DNA中的脱氧核糖)或其衍生物或类似物的化合物。如本文所用,术语“核苷酸”是指共价键结至核苷间连接基团(例如,磷酸基团)的核苷或其任何衍生物、类似物或修饰,它向核酸或其部分或区段赋予改善的化学和/或功能特性(例如,结合亲和力、核酸酶抗性、化学稳定性)。

[1130] 开放阅读框:如本文所用,术语“开放阅读框”,缩写为“ORF”,是指编码多肽的mRNA分子的区段或区域。ORF包含不重叠的框内密码子的连续段,从起始密码子开始并且以终止密码子结束,并且由核糖体翻译。

[1131] 患者:如本文所用,“患者”是指可以寻求或需要治疗、要求治疗、正在接受治疗、将接受治疗的受试者,或处于针对特定疾病或疾患受过训练的专业人员的护理下的受试者。在特定实施方案中,患者是人类患者。在一些实施方案中,患者是罹患自身免疫疾病(例如,如本文所描述)的患者。

[1132] 药学上可接受:短语“药学上可接受”在本文中用于指在合理医学判断的范围内适合与人类和动物组织接触使用而无过度毒性、刺激、过敏反应或其它问题或并发症,与合理的效益/风险比相称的那些化合物、材料、组合物和/或剂型。

[1133] 药学上可接受的赋形剂:如本文所用,短语“药学上可接受的赋形剂”是指除本文所描述的化合物以外(例如,能够悬浮或溶解活性化合物的媒剂)并且具有在患者中基本上无毒且非炎性的特性的任何成分。赋形剂可以包括例如:抗粘剂、抗氧化剂、粘合剂、包衣、压缩助剂、崩解剂、染料(着色剂)、软化剂、乳化剂、填充剂(稀释剂)、成膜剂或包衣、调味剂、芳香剂、助流剂(流动增强剂)、润滑剂、防腐剂、印刷油墨、吸附剂、悬浮或分散剂、甜味剂和水合水。示例性赋形剂包括但不限于:丁基化羟基甲苯(BHT)、碳酸钙、磷酸钙(二碱式)、硬脂酸钙、交联羧甲纤维素、交联聚乙烯吡咯烷酮、柠檬酸、交联聚维酮、半胱氨酸、乙基纤维素、明胶、羟丙基纤维素、羟丙基甲基纤维素、乳糖、硬脂酸镁、麦芽糖醇、甘露糖醇、甲硫氨酸、甲基纤维素、对羟基苯甲酸甲酯、微晶纤维素、聚乙二醇、聚乙烯吡咯烷酮、聚维酮、预胶凝淀粉、对羟基苯甲酸丙酯、棕榈酸视黄酯、虫胶、二氧化硅、羧甲基纤维素钠、柠檬酸钠、乙醇酸淀粉钠、山梨糖醇、淀粉(玉米)、硬脂酸、蔗糖、滑石、二氧化钛、维生素A、维生素E、维生素C和木糖醇。

[1134] 药学上可接受的盐:如本文所用,“药学上可接受的盐”是指所公开化合物的衍生物,其中母体化合物通过将现有酸或碱部分转化为其盐形式(例如,通过使游离碱基团与适合的有机酸反应)而进行修饰。药学上可接受的盐的实例包括但不限于诸如胺的碱性残基的无机或有机酸盐;诸如羧酸的酸性残基的碱金属或有机盐;等等。代表性酸加成盐包括乙酸盐、乙酸、己二酸盐、海藻酸盐、抗坏血酸盐、天冬氨酸盐、苯磺酸盐、苯磺酸、苯甲酸盐、硫酸氢盐、硼酸盐、丁酸盐、樟脑酸盐、樟脑磺酸盐、柠檬酸盐、环戊烷丙酸盐、二葡糖酸盐、十二烷基硫酸盐、乙磺酸盐、富马酸盐、葡庚糖酸盐、甘油磷酸盐、半硫酸盐、庚酸盐、己酸盐、氢溴酸盐、盐酸盐、氢碘酸盐、2-羟基-乙磺酸盐、乳糖酸盐、乳酸盐、月桂酸盐、月桂基硫酸盐、苹果酸盐、马来酸盐、丙二酸盐、甲磺酸盐、2-萘磺酸盐、烟酸盐、硝酸盐、油酸盐、草酸盐、棕榈酸盐、双羟萘酸盐、果胶酸盐、过硫酸盐、3-苯基丙酸盐、磷酸盐、苦味酸盐、特戊酸盐、丙酸盐、硬脂酸盐、琥珀酸盐、硫酸盐、酒石酸盐、硫氰酸盐、甲苯磺酸盐、十一烷酸盐、戊酸盐等。代表性碱金属或碱土金属盐包括钠、锂、钾、钙、镁等;以及无毒铵、季铵和胺阳离子,包括但不限于铵、四甲基铵、四乙基铵、甲胺、二甲胺、三甲胺、三乙胺、乙胺等。本公开的药学上可接受的盐包括例如由无毒无机或有机酸形成的母体化合物的常规无毒盐。本公开的药学上可接受的盐可以通过常规化学方法由含有碱性或酸性部分的母体化合物合成。一般来说,此类盐可以通过使游离酸或碱形式的这些化合物与化学计算量的适当碱或酸在水中或有机溶剂中或两者的混合物中反应来制备;一般优选非水性介质,如乙醚、乙酸乙酯、乙醇、异丙醇或乙腈。适合盐的清单见于Remington's Pharmaceutical Sciences,第17版,Mack Publishing Company,Easton,Pa.,1985,第1418页,Pharmaceutical Salts:Properties,Selection,and Use,P.H.Stahl和C.G.Wermuth(编),Wiley-VCH,2008,以及Berge等人,Journal of Pharmaceutical Science,66,1-19(1977)中,所述文献中的每一者以全文引用的方式并入本文中。

[1135] 多肽:如本文所用,术语“多肽”或“所关注的多肽”是指通常由肽键接合的氨基酸残基的聚合物,所述聚合物可以天然地(例如,分离或纯化)或以合成方式产生。

[1136] RNA:如本文所用,“RNA”是指可以天然或非天然存在的核糖核酸。举例来说,RNA可以包括修饰的和/或非天然存在的组分,诸如一种或多种核碱基、核苷、核苷酸或连接子。RNA可以包括帽结构、链终止核苷、茎环、多聚A序列和/或多聚腺苷酸化信号。RNA可以具有

编码所关注多肽的核苷酸序列。举例来说,RNA可以是信使RNA(mRNA)。编码特定多肽的mRNA的翻译(例如,哺乳动物细胞内的mRNA的体内翻译)可以产生编码的多肽。RNA可以选自由小干扰RNA(siRNA)、不对称干扰RNA(aiRNA)、微RNA(miRNA)、Dicer-底物RNA(dsRNA)、小发夹RNA(shRNA)、mRNA、长链非编码RNA(lncRNA)和它们的混合物组成的非限制性组。

[1137] RNA元件:如本文所用,术语“RNA元件”是指提供生物功能和/或具有生物活性(例如,翻译调控活性)的RNA分子的部分、片段或区段。通过并入一个或多个RNA元件(诸如本文所描述的那些)对多核苷酸进行修饰,为修饰的多核苷酸提供了一种或多种所需的功能特性。如本文所描述,RNA元件可以是天然存在的、非天然存在的、合成的、工程改造的或它们的任何组合。举例来说,提供调控活性的天然存在的RNA元件包括在病毒、原核和真核生物体(例如,人类)的整个转录组中发现的元件。RNA元件,特别是真核mRNA和翻译的病毒RNA,已经显示参与介导细胞中的许多功能。示例性天然RNA元件包括但不限于翻译起始元件(例如,内部核糖体进入位点(IRES),参见Kieft等人,(2001)RNA 7(2):194-206)、翻译增强子元件(例如,APP mRNA翻译增强子元件,参见Rogers等人,(1999)J Biol Che m 274(10):6421-6431)、mRNA稳定性元件(例如,富含AU的元件(ARE),参见Gameau等人,(2007)Nat Rev Mol Cell Biol 8(2):113-126)、翻译抑制元件(参见例如Blumer等人,(2002)Mech Dev 110(1-2):97-112)、蛋白质结合RNA元件(例如,铁反应元件,参见Selezne va等人,(2013)J Mol Biol 425(18):3301-3310)、细胞质多聚腺苷酸化元件(Villalba等人,(2011)Curr Opin Genet Dev 21(4):452-457)和催化RNA元件(例如,核酶,参见Scott等人,(2009)Biochim Bioph ys Acta 1789(9-10):634-641)。

[1138] 基本上:如本文所用,术语“基本上”是指展现出所关注的特征或特性的全部或接近全部范围或程度的定性条件。生物学领域的普通技术人员将了解,生物和化学现象很少(如果有的话)完成和/或继续完成或者达到或避免绝对结果。因此,术语“基本上”在本文中用于捕获许多生物和化学现象中固有的潜在完整性缺乏。

[1139] 罹患:“罹患”疾病、病症和/或疾患的个体已经被诊断出患有或表现出疾病、病症和/或疾患的一种或多种症状。

[1140] 治疗剂:术语“治疗剂”是指当向受试者施用时,具有治疗、诊断和/或预防作用和/或引发所需的生物和/或药理作用的任何剂。在一些实施方案中,治疗剂包含或是治疗性有效载荷。在一些实施方案中,治疗剂包含或是小分子或生物剂(例如,抗体分子)。

[1141] 转染:如本文所用,术语“转染”是指将物质(例如多核苷酸,诸如mRNA)引入细胞中的方法。

[1142] 翻译调控居性:如本文所用,术语“翻译调控活性”(与“翻译调控功能”可互换使用)是指调节(例如,调控、影响、控制、改变)翻译装置的活性,包括PIC和/或核糖体的活性的生物功能、机制或过程。在一些方面,所需的翻译调控活性促进和/或提高mRNA翻译的翻译保真度。在一些方面,所需的翻译调控活性减少和/或抑制遗漏扫描。

[1143] 受试者:如本文所用,术语“受试者”是指可以例如出于实验、诊断、预防和/或治疗目的施用根据本公开的组合物的任何生物体。典型的受试者包括动物(例如哺乳动物,诸如小鼠、大鼠、兔、非人类灵长类动物和人类)和/或植物。在一些实施方案中,受试者可以是患者。

[1144] 治疗:如本文所用,术语“治疗”是指部分地或完全地缓解、改善、改进、减轻特定感

染、疾病、病症和/或疾患，延迟其发作，抑制其进展，降低其严重度，和/或降低其一种或多种症状或特征的发生率。举例来说，“治疗”癌症可以指抑制肿瘤的存活、生长和/或扩散。出于降低发展与疾病、病症和/或疾患相关的病理的风险的目的，治疗可以向未展现疾病、病症和/或疾患的征象的受试者施用，和/或向仅展现疾病、病症和/或疾患的早期征象的受试者施用。

[1145] 预防：如本文所用，术语“预防”是指部分地或完全地抑制特定感染、疾病、病症和/或疾患的一种或多种症状或特征的发作。

[1146] 未修饰：如本文所用，“未修饰”是指在以任何方式改变之前的任何物质、化合物或分子。未修饰可以但不总是指生物分子的野生型或原生形式。分子可以经历一系列修饰，借此每个修饰的分子可以充当“未修饰的”起始分子以供后续修饰。

[1147] 变体：如本文所用，术语“变体”是指具有野生型分子的至少50％、60％、70％、80％、85％、90％、95％、96％、97％、98％、99％或100％活性或结构相似性的分子，例如，如通过本领域公认的测定所测量。

[1148] 22.等效方案和范围

[1149] 本领域技术人员将认识到或能够仅使用常规实验确定根据本文所描述的公开内容的特定实施方案的许多等效方案。本公开的范围不旨在局限于以下描述，而是如所附权利要求书中所阐述。

[1150] 除非相反指示或从上下文另外显而易见，否则在权利要求书中，诸如“一个”、“一种”和“所述”的冠词可以意指一个(种)或超过一个(种)。除非相反指示或从上下文另外显而易见，否则在组的一个或多个成员之间包括“或”的权利要求或描述在组成员中的一者、超过一者或全部存在于、用于给定产物或工艺中或以其它方式与给定产物或工艺相关时被视为满足条件的。本公开包括如下实施方案，其中确切地组的一个成员存在于、用于给定产物或工艺中或以其它方式与给定产物或工艺相关。本公开包括如下实施方案，其中组成员中的超过一者或全部存在于、用于给定产物或工艺中或以其它方式与给定产物或工艺相关。

[1151] 还应注意，术语“包含”旨在是开放性的并且允许但不需要包括额外的要素或步骤。当术语“包含”在本文中使用时，因此还预期并公开术语“由.....组成”。

[1152] 在给出范围的情况下，包括端点。此外，应了解，除非另有指示或从上下文和本领域普通技术人员的理解另外显而易见，否则以范围表述的值可以在本公开的不同实施方案中假设所陈述的范围内的任何特定值或子范围，除非上下文另外清楚地指示，否则达到所述范围的下限的单位的十分之一。

[1153] 所有引用的来源(例如，参考文献、公布、数据库、数据库条目和本文所引用的技术)以引用的方式并入本申请中，即使在引用中未明确陈述。在引用的来源与本申请的陈述有冲突的情况下，应以本申请中的陈述为准。

[1154] 实施例

[1155] 本发明在以下实施例中进一步描述，这些实施例不限制权利要求中所描述的本发明的范围。

[1156] 下表5和表6中提供了与实施例1和2的5′UTR和3′UTR相关的序列信息。

[1157]

表 5: 5'UTR 序列

名称:	序列:
v1.1 5' UTR	GGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAA UAUAAGACCCCCGGCGCCGCCACC (SEQ ID NO:56)
v2.0 5' UTR	GGAAAUCGCAAAAUUUGCUCUUCGCGUUAGAUUU CUUUUAGUUUUCUCGCAACUAGCAAGCUUUUUGU UCUCGCC (SEQ ID NO:50)

表 6: 带有加下划线的终止元件序列的 3' UTR

名称:	序列:
α 3'UTR/对照 3'UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCC UUGGGCCUCCCCCCCCAGCCCCCUCCCCCUUCCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (SEQ ID NO:180)
κ 3'UTR	UAAAGCUCCCCCGGGGGCCUCGGUGGCCUAGCUUCUUGCCCC UUGGGCCUCCCCCCCCAGCCCCCUCCCCCUUCCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (SEQ ID NO:139)
ι 3'UTR	UAAGCCCCCUCCCGGGGGCCUCGGUGGCCUAGCUUCUUGCCCC UUGGGCCUCCCCCCCCAGCCCCCUCCCCCUUCCCUGCACCCGUA CCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (SEQ ID NO:181)

[1158] 实施例1.具有v1.1 5′UTR或v2.0 5′UTR与α3′UTR或κ 3′UTR的组合的mRNA的体内作用

[1159] 本实施例描述了由具有v1.1 5′UTR或v2.0 5′UTR与α3′UTR(本文中也称为“对照3′UTR”)或κ 3′UTR的组合的mRNA编码的萤火虫荧光素酶发光(ffluc)和/或靶蛋白的体内评估。

[1160] 为了评价两种3′UTR(“α”或“κ”)在小鼠体内的作用,通过大剂量静脉内尾静脉注射向CD-1小鼠静脉内给与0.25mg/kg配制的ffLuc mRNA或hEPO mRNA化合物II/DMG(50％化

合物II、10％DSPC、38.5胆固醇、1.5％PEG-DMG 2500MW)。mRNA构建体具有v1.1 5′UTR或v2.0 5′UTR。在给药后0-4天(或0-96小时)对动物进行成像。每组有10只动物。还收集血清,并且通过ELISA(12小时、48小时和96小时)分析hEPO蛋白水平。第4天(即,96小时)处死小鼠;解剖肝和脾并且分析发光。

[1161] 本实验的结果示于图1A-图1G中。具有v2.0 5′UTR和κ 3′UTR两者的mRNA构建体产生最高的全身发光(图1A-图1C)、肝发光(图1D)和脾发光(图1E)。对于血清中的靶蛋白(即,EPO)表达观察到类似作用(图1F-图1G)。

[1162] 实施例2.具有v2.0 5′UTR与κ 3′UTR或ι 3′UTR的组合的mRNA的体内作用

[1163] 本实施例描述了免疫细胞中由具有v2.0 5′UTR与以下的组合的mRNA编码的靶蛋白表达(密码子优化的mOX40L)的体内评估:对照3′UTR(v1.1)、κ 3′UTR或ι 3′UTR。

[1164] 为了评价与对照3'UTR相比两种3'UTR(“κ”或“ι”)在小鼠体内的作用,向C57BL/6小鼠静脉内给与0.50mg/kg在含有化合物II和化合物I的LNP中配制的mOX40L_D99KmRNA。mRNA构建体具有v2.0 5′UTR、密码子优化的mOX40L_D99K ORF和κ 3UTR或l 3′UTR。在给药后1天、2天和3天针对LSK+造血干细胞和祖细胞(图2A)、脾树突状细胞(图2B)、脾巨噬细胞(图2C)、脾嗜中性粒细胞(图2D)、脾单核细胞(图2E)、脾嗜酸性粒细胞(图2F)、脾CD4+T细胞(图2G)、脾CD8+T细胞(图2H)和脾B细胞(图2I)评估OX40L的平均荧光强度和mOX40L+细胞百分比。

[1165] 含有ι 3′UTR或κ 3′UTR的mRNA与大多数免疫细胞中mOX40L的表达增加相关。

[1166] 实施例3.编码FANCA的mmRNA的合成

[1167] 可以例如通过使用SEQ ID NO:1中提供的ORF序列(氨基酸)来构建编码人类范可尼贫血互补组A(Fanconi anemia complementation group A,FANCA)多肽的mRNA。参见下表7。

[1168] SEQ ID NO:2中提供了编码SEQ ID NO:1的氨基酸序列的示例性序列优化的核苷酸序列。SEQ ID NO:3中提供了编码SEQ ID NO:1的氨基酸序列的另一个示例性序列优化的核苷酸序列。SEQ ID NO:14-17中提供了编码SEQ ID NO:1的氨基酸序列的额外示例性序列优化的核苷酸序列。参见下表7。

[1169] mRNA序列包括侧接ORF序列(核苷酸)的5′UTR和3′UTR区。在示例性构建体中,5′UTR和3′UTR序列分别是SEQ ID NO:64和139。

[1170] 5′UTR:

[1171] GGAAAUUAUUAUUAUUUCUAGCUACAAUUUAUCAUUGUAUUAUUUUAGCUAUUCAUCAUUAUUUACUUGGUGAUCAACA (SEQ ID NO:64)

[1172] 3′UTR:

[1173] UAAAGCUCCCCGGGGGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (SEQ ID NO:139)

[1174] 在另一个示例性构建体中,5′UTR和3′UTR序列分别是SEQ ID NO:50和139。

[1175] 5′UTR:

[1176] GGAAAUCGCAAAAUUUGCUCUUCGCGUUAGAUUUCUUUUAGUUUUCUCGCAACUAGCAAGCUUUUUGUUCUCGCC (SEQ ID NO:50)

[1177] 3′UTR:

[1178] UAAAGCUCCCCGGGGGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (SEQ ID NO:139)

[1179] 将FANCAmRNA序列制备成修饰的mRNA。具体地,在体外转录期间,可以使用N1-甲基假尿苷-5′-三磷酸产生修饰的mRNA,以确保mRNA含有100%的N1-甲基假尿苷而不是尿苷。或者,在体外转录期间,可以使用N1-甲氧基尿苷-5′-三磷酸产生修饰的mRNA,以确保mRNA含有100%的5-甲氧基尿苷而不是尿苷。此外,可以用引入多聚A尾的引物合成FANCA-mRNA,并且使用经由m^7G-ppp-Gm-AG四核苷酸的共转录加帽以并入m^7G-ppp-Gm-AG 5′帽1而在两种mRNA上产生帽结构。或者,可以合成FANCA-mRNA并且在DNA模板的Gibson组装期间引入多聚A尾。

[1180] 以下实施例中使用的FANCA构建体序列描述如下。“G5”意指mRNA中的所有尿嘧啶(U)都被N1-甲基假尿嘧啶置换。

[1181]

表 7：FANCA 序列

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
SEQ ID NO:		1	2	64	139	4
	FANCA_01 化学：G5 帽：C1 多聚 A 尾： 100nt	MSDSWVPNSASG QDPGGRRRAWAE LLAGRVKREKYNP ERAQKLKESAVRL LRSHQDLNALLLE VEGPLCKKLSLSK VIDCDSSEAYANH SSSFIGSALQDQAS RLGVPVGILSAGM VASSVGQICTAPAE TSHPVLLTVEQRK KLSSLLEFAQYLL AHSMFSRLSFCQE LWKIQSSLLLEAV WHLHVQGIVSLQE LLESHPDMHAVGS WLFRNLCCLCEQ MEASCQHADVAR AMLSDFVQMFVL RGFQKNSDLRRTV EPEKMPQVTVDVL QRMLIFALDALAA GVQEESSTHKIVR CWFGVFSGHTLGS VISTDPLKRFFSHT LTQILTHSPVLKAS DAVQMQREWSFA RTHPLLTSLYRRLF VMLSAEELVGHLQ	AUGUCCGACUCCU GGGUGCCCAACUC CGCUAGCGGACAA GAUCCAGGUGGCC GCAGGAGAGCGUG GGCUGAACUGCUU GCCGGGCGAGUGA AGCGCGAGAAGUA CAACCCCGAGCGC GCCCAGAAGCUGA AGGAGAGCGCCGU ACGGCUGCUGCGC AGCCACCAGGACC UGAACGCCCUGCU GCUGGAGGUGGAG GGGCCCCUGUGCA AGAAGCUGAGCCU GUCCAAGGUGAUC GACUGCGACAGCA GCGAGGCCUACGC CAACCACAGCUCC UCCUUCAUCGGCU CCGCCCUGCAGGA UCAAGCUAGCCGG CUGGGCGUUCCAG UGGGCAUCCUGAG CGCCGGCAUGGUG GCCAGCAGCGUGG GCCAGAUCUGCAC	GGAAAUUA UUAUUAUU UCUAGCUA CAAUUUAU CAUUGUAU UAUUUUAG CUAUUCAU CAUUAUUU ACUUGGUG AUCAACA	UAAAGCUCCCC GGGGGCCUCGG UGGCCUAGCUU CUUGCCCCUUG GGCCUCCCCCC AGCCCCUCCUC CCCUUCCUGCA CCCGUACCCCC GUGGUCUUUGA AUAAAGUCUGA GUGGGCGGC	SEQ ID NO:4 从 5' 至 3'端由以下组成：SEQ ID NO:64 的 5' UTR、SEQ ID NO:2 的 ORF 序列和 SEQ ID NO:139 的 3' UTR

[1182]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		EVLETQEVHWQR	CGCUCCCGCCGAG			
		VLSFVSALVVCFPE	ACCAGCCACCCCG			
		AQQLLEDWVARL	UGCUGCUGACCGU			
		MAQAFESCQLDS	GGAGCAGCGCAAG			
		MVTAFLVVRQAAL	AAGCUGUCAAGCC			
		EGPSAFLSYADWF	UGCUGGAGUUCGC			
		KASFGSTRGYHGC	CCAGUACCUGCUG			
		SKKALVFLFTFLSE	GCCCACAGCAUGU			
		LVPFESPRYLQVHI	UCUCCCGCCUGAG			
		LHPPLVPGKYRSL	CUUCUGCCAGGAG			
		LTDYISLAKTRLA	CUGUGGAAGAUCC			
		DLKVSIENMGLYE	AGUCCAGCCUGCU			
		DLSSAGDITEPHSQ	ACUGGAGGCCGUG			
		ALQDVEKAIMVFE	UGGCACUUACACG			
		HTGNIPVTVMEASI	UGCAGGGCAUCGU			
		FRRPYYVSHFLPA	GAGCCUGCAGGAG			
		LLTPRVLPKVPDSR	CUACUGGAGAGCC			
		VAFIESLKRADKIP	ACCCCGACAUGCA			
		PSLYSTYCQACSA	CGCCGUGGGGUCC			
		AEEKPEDAALGVR	UGGCUGUUCCGCA			
		AEPNSAEEPLGQL	ACCUGUGCUGCCU			
		TAALGELRASMTD	GUGCGAGCAGAUG			
		PSQRDVISAQVAVI	GAAGCCUCCUGCC			
		SERLRAVLGHNED	AACACGCCGACGU			
		DSSVEISKIQLSINT	UGCACGGGCAAUG			
		PRLEPREHMAVDL	CUGAGCGACUUCG			
		LLTSFCQNLMAAS	UGCAGAUGUUCGU			
		SVAPPERQGPWAA	GCUGCGCGGGUUC			
		LFVRTMCGRVLPA	CAGAAGAACUCAG			
		VLTRLCQLLRHQG	ACCUCCGGAGGAC			
		PSLSAPHVLGLAA	CGUGGAGCCUGAG			
		LAVHLGESRSALP	AAGAUGCCCCAGG			
		EVDVGPPAPGAGL	UGACCGUGGACGU			
		PVPALFDSLLTCRT	GCUGCAGCGGAUG			

[1183]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		RDSLFFCLKFCTA	CUGAUCUUUGCCC			
		AISYSLCKFSSQSR	UGGACGCUCUCGC			
		DTLCSCLSPGLIKK	UGCCGGAGUGCAG			
		FQFLMFRLFSEAR	GAGGAGAGCUCCA			
		QPLSEEDVASLSW	CCCACAAGAUCGU			
		RPLHLPSADWQRA	GCGGUGCUGGUUC			
		ALSLWTHRTFREV	GGCGUGUUCUCCG			
		LKEEDVHLTYQD	GCCACACCCUGGG			
		WLHLELEIQPEAD	CUCCGUGAUCUCC			
		ALSDTERQDFHQ	ACCGAUCCCCUGA			
		WAIHEHFLPESSAS	AGCGGUUCUUCUC			
		GGCDGDLQAACTI	CCACACCCUGACC			
		LVNALMDFHQSSR	CAGAUCCCUGACCC			
		SYDHSENSDLVFG	ACUCCCCUGUGCU			
		GRTGNEDIISRLQE	GAAGGCCAGCGAC			
		MVADLELQQDLIV	GCCGUGCAGAUGC			
		PLGHTPSQEHFLFE	AGCGCGAGUGGUC			
		IFRRRLQALTSGW	CUUCGCCCCGCACC			
		SVAASLQRQRELL	CAUCCCCCUGCUGA			
		MYKRILLRLPSSVL	CCAGCCUGUACCG			
		CGSSFQAEQPITAR	GCGGCUGUUCGUG			
		CEQFFHLVNSEMR	ALGCUGUCCGCCG			
		NFCSHGGALTQDI	AGGAGCUGGUGGG			
		TAHFFRGLLNACL	CCACCUGCAGGAG			
		RSRDPSLMVDFIL	GUGCUGGAGACCC			
		AKCQTKCPLILTSA	AGGAGGUGCACUG			
		LVWWPSLEPVLLC	GCAGCGCGUGCUG			
		RWRRHCQSPLPRE	AGCUUCGUGAGCG			
		LQKLQEGRQFASD	CCCUGGUGGUGUG			
		FLSPEAASPAPNPD	CUUCCCCGAGGCC			
		WLSAAALHFAIQQ	CAGCAGCUCCCUGG			
		VREENIRKQLKKL	AGGACUGGGUGGC			
		DCEREELLVFLFFF	CCGGCUGAUGGCC			
		SLMGLLSSHLTSNS	CAGGCCUUCGAGA			

[1184]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		TTDLPKAFHVCAA	GCUGCCAGCUGGA			
		ILECLEKRKISWLA	CAGCAUGGUGACC			
		LFQLTESDLRLGRL	GCCUUCCUGGUGG			
		LLRVAPDQHTRLL	UGCGUCAAGCAGC			
		PFAFYSLLSYFHED	CCUGGAGGGCCCA			
		AAIREEAFLHVAV	AGCGCCUUCCUGA			
		DMYLKLVQLFVAG	GCUACGCCGACUG			
		DTSTVSPPAGRSLE	GUUCAAGGCCAGC			
		LKGQGNPVELITK	UUCGGGUCCACCC			
		ARLFLLQLIPRCPK	GCGGGUACCACGG			
		KSFSHVAELLADR	CUGCUCCAAGAAG			
		GDCDPEVSAALQS	GCCCUGGUGUUCC			
		RQQAAPDADLSQE	UGUUUACCUUCCU			
		PHLF	GAGCGAGCUGGUG			
			CCCUUCGAGUCGC			
			CCCGCUACCUGCA			
			GGUGCACAUCCUG			
			CACCCACCCCUGG			
			UCCCUGGCAAGUA			
			CCGGAGCCUGCUG			
			ACCGACUACAUCU			
			CCCUGGCCAAGAC			
			CCGGCUGGCCGAC			
			CUGAAGGUGAGCA			
			UCGAGAACAUGGG			
			CCUGUACGAGGAC			
			CUGUCCAGCGCCG			
			GCGACAUCACCGA			
			GCCCCACAGCCAG			
			GCCCUGCAGGACG			
			UGGAGAAGGCCAU			
			CAUGGUGUUCGAG			
			CACACCGGCAACA			
			UCCCCGUGACCGU			

[1185]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GAUGGAGGCUAGU			
			AUCUUCCGGCGGC			
			CCUACUACGUGUC			
			CCACUUCCUGCCC			
			GCCCUGCUGACAC			
			CCCGGGUGCUGCC			
			CAAGGUGCCCGAC			
			AGUCGCGUCGCCU			
			UCAUCGAGAGCCU			
			GAAGCGGGCCGAC			
			AAGAUCCCGCCCU			
			CCCUGUACUCCAC			
			CUACUGCCAGGCU			
			UGCUCCGCCGCCG			
			AGGAGAAGCCCGA			
			GGACGCAGCCUUA			
			GGAGUACGCGCCG			
			AGCCCAACAGCGC			
			CGAGGAGCCACUG			
			GGCCAGUUGACCG			
			CUGCCCUGGGAGA			
			GCUGCGGGCCAGC			
			AUGACCGACCCCA			
			GCCAACGCGACGU			
			GAUCAGCGCCCAG			
			GUGGCCGUGAUCA			
			GCGAGAGGCUGCG			
			GGCAGUGCUGGGC			
			CACAACGAGGACG			
			ACAGCAGCGUGGA			
			GAUCAGCAAGAUC			
			CAGCUGAGCAUCA			
			ACACACCCAGGCU			
			GGAGCCCAGAGAG			

[1186]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CACAUGGCCGUGG			
			ACCUGCUGCUGAC			
			CUCCUUCUGCCAG			
			AACCUGAUGGCCG			
			CAAGCAGCGUGGC			
			UCCUCCAGAACGC			
			CAAGGCCCUUGGG			
			CUGCCCUGUUCGU			
			GCGCACCAUGUGU			
			GGCCGCGUCCUGC			
			CAGCAGUGCUGAC			
			CCGGCUGUGCCAG			
			CUUCUGCGGCACC			
			AAGGGCCCUCCCU			
			GAGCGCUCCCCAC			
			GUGCUUGGGUUGG			
			CCGCCCUCGCUGU			
			GCACCUGGGCGAG			
			AGCAGGAGCGCCC			
			UUCCCGAGGUGGA			
			CGUGGGUCCUCCA			
			GCCCCAGGUGCCG			
			GACUCCCCGUGCC			
			AGCCCUGUUCGAC			
			AGCCUGCUGACUU			
			GUCGCACCCGGGA			
			CUCCCUGUUCUUC			
			UGCCUGAAGUUCU			
			GCACCGCCGCCAU			
			CUCCUACUCCCUG			
			UGCAAGUUCUCCA			
			GCCAGUCCCGGGA			
			CACCCUGUGCUCC			
			UGCCUGUCUCCCG			

[1187]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GCCUGAUCAAGAA			
			GUUCCAGUUCCUG			
			AUGUUCCGCCUGU			
			UCAGCGAAGCACG			
			CCAGCCCCUGUCC			
			GAGGAGGACGUGG			
			CCUCCCUGUCCUG			
			GCGACCCCUGCAC			
			CUGCCCUCCGCCG			
			AUUGGCAGCGCGC			
			CGCCCUUAGCCUG			
			UGGACCCACCGGA			
			CCUUCCGCGAGGU			
			GCUGAAGGAGGAA			
			GACGUGCACCUGA			
			CCUACCAGGACUG			
			GCUGCACCUGGAG			
			CUGGAGAUCCAGC			
			CCGAAGCCGACGC			
			CCUGAGCGACACC			
			GAGCGCCAGGACU			
			UUCACCAGUGGGC			
			CAUCCACGAGCAC			
			UUUCUGCCCGAGA			
			GCAGCGCCAGUGG			
			AGGCUGCGACGGC			
			GACCUGCAAGCCG			
			CCUGCACCAUCCU			
			GGUGAACGCCCUG			
			ALGGACUUCCACC			
			AGAGCUCCCGCAG			
			CUACGACCACUCC			
			GAGAACAGCGACC			
			UGGUGUUUGGAGG			

[1188]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCGCACCGGGAAC			
			GAGGACAUCAUCA			
			GCCGCCUGCAGGA			
			GAUGGUGGCCGAC			
			CUGGAGCUCCAGC			
			AGGACCUGAUCGU			
			GCCACUGGGGCAC			
			ACCCCAAGCCAGG			
			AGCACUUCCUGUU			
			CGAGAUCUUCCGC			
			AGACGCCUGCAGG			
			CACUGACCUCCGG			
			CUGGAGUGUGGCC			
			GCAAGCCUGCAGC			
			GCCAACGCGAGCU			
			GCUGAUGUACAAG			
			CGGAUCCUGCUGC			
			GCCUGCCCUCCUC			
			CGUGCUGUGCGGC			
			UCCUCCUUCCAGG			
			CCGAGCAGCCCAU			
			CACCGCCCGGUGC			
			GAGCAGUUCUUCC			
			ACCUGGUGAACAG			
			CGAGAUGCGGAAC			
			UUCUGCUCUCACG			
			GCGGCGCUCUGAC			
			CCAGGACAUCACC			
			GCCCACUUCUUCC			
			GCGGCCUGCUGAA			
			CGCCUGCCUGCGG			
			AGCAGGGACCCCA			
			GCCUGAUGGUCGA			
			CUUCAUCCUGGCC			

[1189]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			AAGUGCCAGACCA			
			AGUGUCCCCUGAU			
			CCUGACCAGCGCC			
			CUGGUUUGGUGGC			
			CCUCACUGGAGCC			
			CGUGCUCCUGUGU			
			CGCUGGAGGCGGC			
			ACUGCCAGUCACC			
			CCUGCCCCGAGAG			
			CUGCAGAAGCUGC			
			AGGAGGGCCGGCA			
			GUUCGCCAGCGAC			
			UUCCUGUCACCAG			
			AGGCAGCCUCUCC			
			CGCUCCUAACCCC			
			GAUUGGCUAUCCG			
			CGGCCGCUCUGCA			
			CUUCGCCAUCCAG			
			CAGGUGCGGGAGG			
			AGAACAUCCGGAA			
			GCAGCUGAAGAAG			
			CUGGACUGCGAGA			
			GGGAGGAGCUGCU			
			UGUGUUCCUGUUC			
			UUCUUCAGCCUGA			
			UGGGCCUGCUGUC			
			CUCCCACCUGACC			
			AGCAACUCCACCA			
			CCGACCUGCCCAA			
			GGCCUUCCACGUG			
			UGCGCCGCCAUCC			
			UGGAGUGCCUGGA			
			GAAGCGGAAGAUC			
			AGCUGGCUGGCCC			

[1190]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			UGUUCCAGCUGAC			
			CGAGAGCGACCUG			
			CGCCUGGGGAGAC			
			UGCUGCUGCGCGU			
			GGCUCCCGACCAG			
			CACACCAGGCUGC			
			UGCCCUUCGCCUU			
			CUACUCCCUGCUG			
			UCCUACUUCCACG			
			AGGACGCCGCCAU			
			CCGGGAGGAGGCC			
			UUCCUGCACGUGG			
			CCGUGGACAUGUA			
			CCUGAAGCUGGUG			
			CAGCUGUUCGUGG			
			CCGGCGACACCUC			
			CACCGUGAGCCCU			
			CCUGCUGGCCGAA			
			GCCUGGAGCUGAA			
			GGGCCAGGGCAAC			
			CCCGUGGAGCUGA			
			UCACCAAGGCCCG			
			CCUGUUCCUGCUG			
			CAGCUGAUCCCUC			
			GCUGCCCCAAGAA			
			GAGCUUCAGCCAC			
			GUGGCCGAGCUGC			
			UGGCUGACAGGGG			
			AGACUGCGACCCC			
			GAGGUGAGCGCUG			
			CCCUGCAGAGCCG			
			UCAACAAGCGGCC			
			CCUGACGCCGAUC			
			UGAGCCAGGAGCC			

[1191]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCACCUGUUC			
SEQ ID NO		1	3	50	139	5
	FANCA_02 化学：G5 帽：m7Gp-ppGm-A 多聚 A 尾：100nt	MSDSWVPNSASG QDPGGRRRAWAE LLAGRVKREKYNP ERAQKLKESAVRL LRSHQDLNALLLE VEGPLCKKLSLSK VIDCDSSEAYANH SSSFIGSALQDQAS RLGVPVGILSAGM VASSVGQICTAPAE TSHPVLLTVEQRK KLSSLLEFAQYLL AHSMFSRLSFCQE LWKIQSSLLLEAV WHLHVQGIVSLQE LLESHPDMHAVGS WLFRNLCCLCEQ MEASCQHADVAR AMLSDFVQMFVL RGFQKNSDLRRTV EPEKMPQVTVDVL QRMLIFALDALAA GVQEESSTHKIVR CWFGVFSGHTLGS VISTDPLKRFFSHT LTQILTHSPVLKAS DAVQMQREWSFA RTHPLLTSLYRRLF VMLSAEELVGHLQ EVLETQEVHWQR VLSFVSALVVCFPE	AUGAGCGACAGCU GGGUGCCUAAUAG CGCCAGCGGACAG GAUCCUGGUGGAA GGAGAAGGGCGUG GGCAGAGCUACUG GCUGGACGGGUGA AGCGAGAGAAGUA CAACCCCGAGCGA GCACAGAAGCUGA AGGAGUCAGCCGU GAGACUGCUGCGG AGCCACCAAGACC UGAACGCCUUACU GCUGGAAGUGGAG GGACCCCUGUGCA AGAAGCUGAGCCU GAGCAAGGUGAUC GACUGCGACAGCA GCGAGGCCUACGC CAACCACAGCAGC AGCUUCAUCGGCA GCGCAUUGCAGGA CCAAGCUUCUAGG CUAGGCGUCCCCG UUGGUAUCCUGAG CGCCGGUAUGGUA GCAAGCAGCGUGG GCCAGAUUUGUAC CGCACCUGCCGAG ACUAGCCACCCCG	GGAAAUCG CAAAAUUU GCUCUUCG CGUUAGAU UUCUUUUA GUUUUCUC GCAACUAG CAAGCUUU UUGUUCUC GCC	UAAAGCUCCCC GGGGGCCUCGG UGGCCUAGCUU CUUGCCCCUUG GGCCUCCCCCC AGCCCCUCCUC CCCUUCCUGCA CCCGUACCCCC GUGGUCUUUGA AUAAAGUCUGA GUGGGCGGC	SEQ ID NO:5 从 5' 至 3'端由以下组成：SEQ ID NO:50 的 5' UTR、SEQ ID NO:3 的 ORF 序列和 SEQ ID NO:139 的 3' UTR

[1192]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		AQQLLEDWVARL	UGCUUCUGACCGU			
		MAQAFESCQLDS	GGAGCAGCGGAAG			
		MVTAFLVVRQAAL	AAGCUGAGCAGCC			
		EGPSAFLSYADWF	UGCUGGAGUUCGC			
		KASFGSTRGYHGC	CCAGUACCUGCUU			
		SKKALVFLFTFLSE	GCGCACAGCAUGU			
		LVPFESPRYLQVHI	UCAGCCGGCUGAG			
		LHPPLVPGKYRSL	CUUCUGCCAGGAG			
		LTDYISLAKTRLA	CUGUGGAAGAUCC			
		DLKVSIENMGLYE	AGAGCAGCCUGCU			
		DLSSAGDITEPHSQ	GCUGGAAGCCGUG			
		ALQDVEKAIMVFE	UGGCAUCUGCACG			
		HTGNIPVTVMEASI	UGCAGGGCAUCGU			
		FRRPYYVSHFLPA	GAGCCUGCAGGAG			
		LLTPRVLPKVPDSR	CUGCUGGAAAGCC			
		VAFIESLKRADKIP	ACCCAGACAUGCA			
		PSLYSTYCQACSA	CGCCGUAGGGAGU			
		AEEKPEDAALGVR	UGGCUGUUCCGGA			
		AEPNSAEEPLGQL	ACCUGUGCUGCCU			
		TAALGELRASMTD	GUGCGAGCAGAUG			
		PSQRDVISAQVAVI	GAGGCCAGCUGUC			
		SERLRAVLGHNED	AACACGCUGACGU			
		DSSVEISKIQLSINT	AGCUCCGCGCAAUG			
		PRLEPREHMAVDL	CUGAGCGACUUCG			
		LLTSFCQNLMAAS	UGCAGAUGUUCGU			
		SVAPPERQGPWAA	GCUGCGGGGCUUC			
		LFVRTMCGRVLPA	CAGAAGAACAGCG			
		VLTRLCQLLRHQG	ACCUCAGACGGAC			
		PSLSAPHVLGLAA	CGUGGAACCCGAG			
		LAVHLGESRSALP	AAGAUGCCACAGG			
		EVDVGPPAPGAGL	UGACAGUGGACGU			
		PVPALFDSLLTCRT	GCUGCAGCGGAUG			
		RDSLFFCLKFCTA	CUGAUUUUCGCGU			
		AISYSLCKFSSQSR	UAGACGCACACUAGC			

[1193]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		DTLCSCLSPGLIKK	CGCAGGCGUUCAG			
		FQFLMFRLFSEAR	GAGGAGAGCAGCA			
		QPLSEEDVASLSW	CCCACAAGAUAGU			
		RPLHLPSADWQRA	GCGGUGCUGGUUC			
		ALSLWTHRTFREV	GGCGUGUUCUCUG			
		LKEEDVHLTYQD	GCCACACACUGGG			
		WLHLELEIQPEAD	GAGCGUGAUCAGC			
		ALSDTERQDFHQ	ACCGAUCCCCUGA			
		WAIHEHFLPESSAS	AGCGGUUCUUCAG			
		GGCDGDLQAACTI	CCACACCCUGACC			
		LVNALMDFHQSSR	CAGAUCCUGACCC			
		SYDHSENSDLVFG	AUAGCCCCGUGCU			
		GRTGNEDIISRLQE	GAAGGCAAGCGAC			
		MVADLELQQDLIV	GCCGUGCAGAUGC			
		PLGHTPSQEHFLFE	AGCGUGAGUGGAG			
		IFRRRLQALTSGW	CUUCGCACGGACC			
		SVAASLQRQRELL	CAUCCGCUACUGA			
		MYKRILLRLPSSVL	CCAGCCUAUACCG			
		CGSSFQAEQPITAR	GCGCCUGUUCGUG			
		CEQFFHLVNSEMR	AUGCUAUCCGCCG			
		NFCSHGGALTQDI	AGGAGUUGGUCGG			
		TAHFFRGLLNACL	ACACCUGCAGGAG			
		RSRDPSLMVDFIL	GUGCUGGAGACCC			
		AKCQTKCPLILTSA	AGGAGGUGCACUG			
		LVWWPSLEPVLLC	GCAGAGAGUGCUG			
		RWRRHCQSPLPRE	AGCUUCGUGAGCG			
		LQKLQEGRQFASD	CCCUGGUGGUGUG			
		FLSPEAASPAPNPD	CUUUCCCGAGGCA			
		WLSAAALHFAIQQ	CAGCAGCUGCUGG			
		VREENIRKQLKKL	AAGACUGGGUAGC			
		DCEREELLVFLFFF	UCGGCUUAUGGCC			
		SLMGLLSSHLTSNS	CAGGCCUUCGAGA			
		TTDLPKAFHVCAA	GCUGCCAGCUGGA			
		ILECLEKRKISWLA	CAGCAUGGUGACC			

[1194]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		LFQLTESDLRLGRL	GCCUUUCUCGUGG			
		LLRVAPDQHTRLL	UGAGGCAAGCAGC			
		PFAFYSLLSYFHED	CCUAGAAGGCCCU			
		AAIREEAFLHVAV	AGCGCCUUUCUGA			
		DMYLKLVQLFVAG	GCUACGCCGACUG			
		DTSTVSPPAGRSLE	GUUCAAAGCCAGC			
		LKGQGNPVELITK	UUCGGCUCCACUC			
		ARLFLLQLIPRCPK	GGGGUUACCACGG			
		KSFSHVAELLADR	CUGCAGCAAGAAG			
		GDCDPEVSAALQS	GCCCUGGUGUUCC			
		RQQAAPDADLSQE	UGUUCACCUUCCU			
		PHLF	GAGCGAGCUGGUG			
			CCCUUCGAGUCAC			
			CCCGGUACCUGCA			
			GGUGCACAUCCUG			
			CAUCCACCACUGG			
			UCCCCGGGAAAUA			
			CCGGAGCCUGCUG			
			ACCGACUACAUCA			
			GCCUGGCCAAGAC			
			UCGGCUGGCAGAC			
			CUGAAGGUGAGCA			
			UCGAGAACAUGGG			
			CCUGUACGAGGAU			
			CUGAGCAGCGCUG			
			GCGACAUCACAGA			
			ACCGCACAGCCAG			
			GCCUUGCAGGACG			
			UGGAGAAGGCCAU			
			CAUGGUGUUCGAG			
			CACACCGGCAACA			
			UCCCCGUGACCGU			
			GAUGGAGGCCAGC			
			AUUUUCCGCCGGC			

[1195]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CAUACUACGUUAG			
			CCACUUCCUGCCC			
			GCACUGCUGACUC			
			CACGCGUGCUGCC			
			AAAGGUCCCCGAU			
			AGCAGAGUGGCUU			
			UCAUCGAGAGCCU			
			GAAACGGGCCGAC			
			AAGAUCCCACCUA			
			GCCUGUACAGCAC			
			CUACUGCCAAGCU			
			UGUAGCGCUGCCG			
			AGGAGAAACCGGA			
			AGACGCUGCACUG			
			GGUGUUAGGGCCG			
			AGCCCAAUAGCGC			
			AGAGGAACCCUUG			
			GGCCAACUGACUG			
			CUGCACUUGGUGA			
			GCUUCGGGCAAGC			
			ALGACCGACCCCU			
			CUCAACGGGACGU			
			GAUCAGCGCCCAA			
			GUGGCCGUGAUCA			
			GCGAGCGGUUACG			
			UGCCGUGCUGGGC			
			CACAACGAGGACG			
			ACAGCAGCGUGGA			
			GAUCAGCAAGAUC			
			CAGCUGAGCAUCA			
			ALACGCCUCGACU			
			GGAACCUCGGGAG			
			CACAUGGCCGUGG			
			ACCUGCUGCUGAC			

[1196]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CAGCUUCUGCCAG			
			AACUUGAUGGCCG			
			CUAGCAGUGUGGC			
			UCCACCCGAACGG			
			CAAGGACCUUGGG			
			CUGCGCUGUUCGU			
			CAGGACCAUGUGU			
			GGCCGGGUUCUUC			
			CCGCUGUGUUGAC			
			CCGGCUUUGCCAG			
			CUGCUGCGGCAUC			
			AAGGGCCAUCUUU			
			GUCCGCACCUCAC			
			GUUCUGGGUCUUG			
			CUGCUCUGGCCGU			
			GCAUCUGGGCGAG			
			AGCAGAUCGGCAC			
			UGCCGGAAGUGGA			
			CGUGGGACCUCCU			
			GCACCUGGUGCAG			
			GUCUUCCCGUGCC			
			AGCCUUAUUCGAC			
			AGCCUGCUUACCU			
			GCCGUACUCGGGA			
			CAGCCUGUUCUUC			
			UGCCUGAAGUUCU			
			GCACCGCCGCCAU			
			CAGCUACAGCCUG			
			UGCAAGUUCAGCA			
			GCCAGAGCCGGGA			
			CACUCUGUGUAGC			
			UGUCUGAGCCCUG			
			GCCUCAUCAAGAA			
			GUUCCAGUUCCUG			

[1197]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			AUGUUCAGAUUAU			
			UUAGCGAGGCUCG			
			GCAACCUCUGAGC			
			GAGGAGGACGUUG			
			CCAGCCUGAGUUG			
			GCGUCCCCUACAC			
			CUUCCCAGCGCCG			
			AUUGGCAGAGGGC			
			UGCUCUGUCCCUG			
			UGGACUCACCGCA			
			CUUUCCGGGAGGU			
			GCUGAAGGAGGAG			
			GACGUGCACCUGA			
			CCUACCAGGACUG			
			GCUGCACCUGGAG			
			CUCGAGAUACAGC			
			CCGAAGCUGACGC			
			CCUGUCAGACACC			
			GAGCGGCAGGACU			
			UCCACCAGUGGGC			
			CAUCCACGAGCAC			
			UUCCUGCCUGAGA			
			GCAGCGCCUCAGG			
			AGGGUGUGACGGG			
			GACCUGCAAGCAG			
			CCUGCACCAUCCU			
			GGUGAACGCCCUG			
			AUGGACUUCCACC			
			AGAGCAGCCGGAG			
			CUACGACCACAGC			
			GAGAACAGCGAUC			
			UAGUCUUCGGCGG			
			CAGAACAGGUAAC			
			GAGGACAUCAUCA			

[1198]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GCCGGCUGCAGGA			
			GAUGGUGGCCGAC			
			CUGGAACUGCAGC			
			AGGACCUGAUUGU			
			CCCCUUGGGUCAC			
			ACCCCUAGCCAGG			
			AGCACUUCCUGUU			
			CGAGAUUUUCAGG			
			AGGCGGCUGCAAG			
			CACUGACCAGCGG			
			CUGGUCUGUCGCU			
			GCUUCCCUUCAGC			
			GUCAGCGGGAGCU			
			GCUGAUGUACAAG			
			CGGAUUCUGCUGC			
			GGCUACCCAGCAG			
			UGUGCUGUGCGGC			
			UCUAGCUUUCAAG			
			CCGAGCAGCCAAU			
			CACCGCUCGCUGC			
			GAGCAGUUCUUCC			
			ACCUGGUGAACAG			
			CGAGAUGCGGAAU			
			UUCUGUUCCCACG			
			GAGGUGCUCUGAC			
			CCAGGACAUCACC			
			GCCCAUUUCUUCC			
			GGGGCCUGCUGAA			
			CGCUUGUCUCCGA			
			UCGAGAGACCCCA			
			GCCUGAUGGUGGA			
			CUUCAUCCUGGCC			
			AAGUGCCAGACCA			
			AGUGUCCCCUGAU			

[1199]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCUGACCUCUGCC			
			CUGGUGUGGUGGC			
			CUUCACUGGAGCC			
			CGUGUUGCUGUGU			
			CGCUGGCGGAGAC			
			AUUGUCAAUCACC			
			ACUGCCACGGGAG			
			CUGCAGAAGCUGC			
			AAGAGGGUCGGCA			
			AUUCGCCUCGGAC			
			UUUCUGUCUCCUG			
			AAGCUGCUAGCCC			
			UGCCCCCUAAUCCU			
			GACUGGUUGUCUG			
			CGGCCGCACUACA			
			CUUUGCCAUCCAG			
			CAGGUGCGGGAGG			
			AGAACAUCCGGAA			
			GCAGCUGAAGAAA			
			CUGGACUGCGAGC			
			GGGAAGAGCUGCU			
			GGUGUUCCUGUUC			
			UUCUUCAGCCUGA			
			UGGGCCUGCUGAG			
			CUCACACCUGACC			
			AGCAACAGCACUA			
			CCGACCUGCCCAA			
			AGCCUUCCACGUG			
			UGCGCUGCCAUCC			
			UGGAGUGCCUGGA			
			GAAGCGGAAGAUC			
			AGCUGGCUGGCCC			
			UGUUCCAGCUGAC			
			AGAGUCUGAUCUU			

[1200]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
SEQ ID		1	14	50	139	23
			CGGCUGGGGCGGU UACUCCUGAGAGU AGCACCCGAUCAG CAUACCAGGCUGC UCCCCUUCGCAUU CUACAGCCUGCUG AGCUACUUCCACG AAGACGCCGCCAU ACGCGAGGAAGCU UUCCUGCACGUGG CCGUGGACAUGUA CCUGAAGCUGGUG CAGCUGUUCGUUG CUGGCGACACCAG CACCGUGUCACCU CCUGCAGGCAGGA GCCUGGAGCUGAA AGGUCAGGGCAAC CCCGUGGAGCUGA UCACCAAGGCCCG GCUGUUCCUGCUU CAGCUGAUACCCC GGUGCCCCAAGAA GAGCUUUAGCCAC GUGGCCGAACUGC UGGCAGACCGGGG UGAUUGCGACCCA GAAGUGUCAGCCG CCCUUCAAAGCCG GCAACAAGCUGCC CCUGACGCCGACU UGAGCCAGGAGCC GCACCUGUUC			

[1201]

NO:	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
	FANCA_03 化学：G5 帽：C1 多聚 A 尾： 100nt	MSDSWVPNSASG QDPGGRRRAWAE LLAGRVKREKYNP ERAQKLKESAVRL LRSHQDLNALLLE VEGPLCKKLSLSK VIDCDSSEAYANH SSSFIGSALQDQAS RLGVPVGILSAGM VASSVGQICTAPAE TSHPVLLTVEQRK KLSSLLEFAQYLL AHSMFSRLSFCQE LWKIQSSLLLEAV WHLHVQGIVSLQE LLESHPDMHAVGS WLFRNLCCLCEQ MEASCQHADVAR AMLSDFVQMFVL RGFQKNSDLRRTV EPEKMPQVTVDVL QRMLIFALDALAA GVQEESSTHKIVR CWFGVFSGHTLGS VISTDPLKRFFSHT LTQILTHSPVLKAS DAVQMQREWSFA RTHPLLTSLYRRLF VMLSAEELVGHLQ EVLETQEVHWQR VLSFVSALVVCFPE AQQLLEDWVARL MAQAFESCQLDS	ALGAGCGACAGCU GGGUGCCCAACAG UGCUUCCGGGCAG GAUCCCGGCGGUA GAAGACGGGCUUG GGCUGAGCUGCUG GCCGGCAGGGUGA AGCGGGAGAAGUA CAACCCCGAGCGG GCCCAGAAGCUGA AGGAGAGUGCCGU GCGGCUGCUGCGG AGCCACCAGGACC UGAACGCCCUCCU GCUGGAGGUCGAG GGCCCACUGUGCA AGAAGCUCAGCCU GAGCAAGGUGAUC GACUGUGACAGCA GCGAGGCCUACGC CAACCACAGCUCC AGCUUCAUCGGCA GCGCCCUGCAGGA CCAGGCCAGCAGA CUGGGCGUGCCCG UGGGCAUCCUGAG CGCCGGCAUGGUG GCCAGCUCGGUGG GCCAGAUCUGCAC CGCUCCCGCCGAG ACCAGCCACCCCG UGCUGCUCACCGU GGAGCAGCGGAAG	GGAAAUCG CAAAAUUU GCUCUUCG CGUUAGAU UUCUUUUA GUUUUCUC GCAACUAG CAAGCUUU UUGUUCUC GCC	UAAAGCUCCCC GGGGGCCUCGG UGGCCUAGCUU CUUGCCCCUUG GGCCUCCCCCC AGCCCCUCCUC CCCUUCCUGCA CCCGUACCCCC GUGGUCUUUGA AUAAAGUCUGA GUGGGCGGC	SEQ ID NO:23 从 5'至 3'端由以下组成：SEQ ID NO:50 的 5' UTR、SEQ ID NO:14 的 ORF 序列和 SEQ ID NO:139 的 3' UTR

[1202]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		MVTAFLVVRQAAL	AAGCUGAGCAGCC			
		EGPSAFLSYADWF	UGCUGGAGUUCGC			
		KASFGSTRGYHGC	CCAGUACCUGCUG			
		SKKALVFLFTFLSE	GCCCACAGCAUGU			
		LVPFESPRYLQVHI	UCUCCCGGCUGAG			
		LHPPLVPGKYRSL	CUUUUGCCAAGAG			
		LTDYISLAKTRLA	CUGUGGAAGAUCC			
		DLKVSIENMGLYE	AGAGCUCCCUGCU			
		DLSSAGDITEPHSQ	GCUGGAGGCCGUG			
		ALQDVEKAIMVFE	UGGCACCUUCACG			
		HTGNIPVTVMEASI	UGCAGGGCAUCGU			
		FRRPYYVSHFLPA	GAGCCUGCAGGAA			
		LLTPRVLPKVPDSR	CUGCUGGAGAGCC			
		VAFIESLKRADKIP	ACCCCGACAUGCA			
		PSLYSTYCQACSA	CGCCGUGGGCAGC			
		AEEKPEDAALGVR	UGGCUGUUCCGGA			
		AEPNSAEEPLGQL	ACCUGUGCUGCCU			
		TAALGELRASMTD	GUGCGAGCAGAUG			
		PSQRDVISAQVAVI	GAGGCUAGCUGCC			
		SERLRAVLGHNED	AGCACGCCGACGU			
		DSSVEISKIQLSINT	GGCACGGGCCAUG			
		PRLEPREHMAVDL	CUGAGCGACUUCG			
		LLTSFCQNLMAAS	UGCAGAUGUUCGU			
		SVAPPERQGPWAA	GCUGCGGGGCUUC			
		LFVRTMCGRVLPA	CAGAAGAAUAGCG			
		VLTRLCQLLRHQG	ACCUACGGCGGAC			
		PSLSAPHVLGLAA	CGUGGAGCCCGAG			
		LAVHLGESRSALP	AAGAUGCCCCAGG			
		EVDVGPPAPGAGL	UGACCGUGGACGU			
		PVPALFDSLLTCRT	GCUGCAGCGGAUG			
		RDSLFFCLKFCTA	CUGAUCUUCGCCC			
		AISYSLCKFSSQSR	UGGACGCUCUGGC			
		DTLCSCLSPGLIKK	CGCUGGCGUGCAG			
		FQFLMFRLFSEAR	GAGGAGAGCAGCA			

[1203]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		QPLSEEDVASLSW	CCCACAAGAUCGU			
		RPLHLPSADWQRA	GCGGUGCUGGUUC			
		ALSLWTHRTFREV	GGCGUGUUCAGCG			
		LKEEDVHLTYQD	GCCACACCCUGGG			
		WLHLELEIQPEAD	CAGCGUGAUCAGC			
		ALSDTERQDFHQ	ACCGACCCACUGA			
		WAIHEHFLPESSAS	AGCGGUUCUUCAG			
		GGCDGDLQAACTI	CCACACCCUCACC			
		LVNALMDFHQSSR	CAGAUCCUGACCC			
		SYDHSENSDLVFG	ACAGCCCCGUGCU			
		GRTGNEDIISRLQE	GAAGGCCAGCGAC			
		MVADLELQQDLIV	GCCGUGCAGAUGC			
		PLGHTPSQEHFLFE	AGCGGGAGUGGAG			
		IFRRRLQALTSGW	CUUCGCCCGGACC			
		SVAASLQRQRELL	CACCCACUGCUGA			
		MYKRILLRLPSSVL	CUAGCCUGUACCG			
		CGSSFQAEQPITAR	GCGGCUGUUCGUG			
		CEQFFHLVNSEMR	AUGCUGAGCGCAG			
		NFCSHGGALTQDI	AGGAGCUGUGGGG			
		TAHFFRGLLNACL	CCACCUGCAGGAG			
		RSRDPSLMVDFIL	GUGCUGGAGACCC			
		AKCQTKCPLILTSA	AGGAGGUGCACUG			
		LVWWPSLEPVLLC	GCAGCGGGUGCUG			
		RWRRHCQSPLPRE	AGCUUCGUGAGCG			
		LQKLQEGRQFASD	CCCUGGUGGUGUG			
		FLSPEAASPAPNPD	CUUCCCCGAGGCC			
		WLSAAALHFAIQQ	CAGCAGCUCCUGG			
		VREENIRKQLKKL	AGGACUGGGUGGC			
		DCEREELLVFLFFF	CCGGCUGAUGGCC			
		SLMGLLSSHLTSNS	CAGGCCUUCGAGA			
		TTDLPKAFHVCAA	GCUGCCAGCUGGA			
		ILECLEKRKISWLA	CAGCAUGGUGACC			
		LFQLTESDLRLGRL	GCCUUCCUGGUGG			
		LLRVAPDQHTRLL	UGCGGCAGGCUGC			

[1204]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		PFAFYSLLSYFHED	CCUGGAAGGCCCC			
		AAIREEAFLHVAV	AGCGCUUUCCUGA			
		DMYLKLVQLFVAG	GCUACGCCGACUG			
		DTSTVSPPAGRSLE	GUUCAAGGCCAGC			
		LKGQGNPVELITK	UUCGGCUCCCACCC			
		ARLFLLQLIPRCPK	GGGGCUACCACGG			
		KSFSHVAELLADR	CUGCAGCAAGAAG			
		GDCDPEVSAALQS	GCCCUGGUGUUCC			
		RQQAAPDADLSQE	UGUUCACCUUCCU			
		PHLF	GAGCGAGCUGGGUG			
			CCCUUCGAGUCAC			
			CCCGGUACCUGCA			
			GGUGCACAUCCUG			
			CAUCCUCCGCUGG			
			UGCCCGGCAAGUA			
			CCGGAGCCUGCUG			
			ACCGACUACAUCA			
			GCCUGGCCAAGAC			
			CCGGCUGGCCGAC			
			CUGAAGGUGAGCA			
			UCGAGAACAUGGG			
			CCUGUACGAGGAC			
			CUGAGCAGCGCCG			
			GCGACAUCACCGA			
			GCCCCACAGCCAG			
			GCCCUGCAAGACG			
			UGGAGAAGGCCAU			
			CAUGGUGUUCGAG			
			CACACCGGCAACA			
			UCCCCGUGACCGU			
			CAUGGAGGCCAGC			
			ACUUCCGGCGGC			
			CCUACUACGUGAG			
			CCACUUCCUGCCC			

[1205]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GCCCUGCUGACCC			
			CACGGGUGCUGCC			
			CAAGGUGCCCCGAC			
			AGCCGGGUGGCCU			
			UCAUCGAGAGCCU			
			GAAGCGGGCCCGAC			
			AAGAUCCCACCCA			
			GCCUGUACAGCAC			
			CUACUGCCAGGCC			
			UGCUCUGCCGCCG			
			AGGAGAAGCCCGA			
			GGACGCCGCAUUA			
			GGCGUGCGGGCCG			
			AGCCCAACAGCGC			
			CGAGGAACCCCUG			
			GGGCAGCUGACCG			
			CAGCCCUGGGAGA			
			GCUGCGGGCCAGC			
			AUGACCGACCCCA			
			GCCAGCGGGACGU			
			GAUCAGCGCCCAG			
			GUGGCCGUCAUCA			
			GCGAGCGGCUGCG			
			GGCAGUGCUGGGC			
			CACAACGAGGACG			
			ACAGCAGCGUGGA			
			GAUCAGCAAGAUC			
			CAGCUGAGCAUCA			
			ACACCCCACGGCU			
			GGAACCUCGGGAG			
			CACAUGGCCGUGG			
			ACCUGCUGCUGAC			
			CAGCUUCUGCCAG			
			AACCUGAUGGCCG			

[1206]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCUCUAGCGUGGC			
			CCCACCAGAACGG			
			CAGGGUCCCCUGGG			
			CUGCCCUGUUCGU			
			GCGGACCAUGUGC			
			GGCCGGGUUCUGC			
			CCGCUGUGCUGAC			
			CCGGCUGUGCCAG			
			CUGCUGCGGCACC			
			AGGGCCCUAGCCU			
			GAGCGCUCCCCAC			
			GUGCUGGGCCUAG			
			CUGCCCUGGCCGU			
			GCACCUUGGGGAG			
			UCCCGAAGCGCCC			
			UGCCCGAGGUGGA			
			CGUGGGCCCUCCA			
			GCCCCAGGUGCUG			
			GCCUUCCCGUGCC			
			CGCUCUGUUCGAC			
			AGUCUGCUGACCU			
			GCCGGACCCGGGA			
			CAGCCUGUUCUUC			
			UGCCUGAAGUUCU			
			GCACCGCCGCCAU			
			CAGCUACAGCCUG			
			UGCAAGUUCAGCA			
			GCCAGAGCCGGGA			
			CACCCUGUGCAGC			
			UGCCUGAGCCCCG			
			GCCUGAUCAAGAA			
			GUUCCAGUUCCUG			
			ALGUUCCGGCUGU			
			UCAGCGAAGCCCG			

[1207]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GCAGCCCCUGAGC			
			GAGGAGGACGUCG			
			CCAGCCUGUCCUG			
			GAGACCCCUGCAC			
			CUGCCCAGCGCCG			
			ACUGGCAGAGAGC			
			CGCCCUGAGCCUG			
			UGGACCCACCGGA			
			CCUUCCGGGAGGU			
			GCUGAAGGAAGAG			
			GACGUGCACCUGA			
			CCUACCAGGACUG			
			GCUGCACCUGGAG			
			CUGGAGAUUCAGC			
			CCGAGGCCGACGC			
			CCUGAGCGACACC			
			GAGCGGCAGGAUU			
			UCCACCAGUGGGC			
			CAUCCACGAGCAU			
			UUCCUGCCCGAGA			
			GCAGCGCCAGCGG			
			CGGCUGUGACGGC			
			GACCUGCAGGCCG			
			CCUGCACCAUCCU			
			GGUGAACGCCCUG			
			AUGGACUUCCACC			
			AGAGCAGCCGGAG			
			CUACGACCACAGC			
			GAGAACAGCGACC			
			UGGUGUUCGGCGG			
			ACGGACCGGCAAC			
			GAGGACAUCAUCA			
			GCCGGCUGCAGGA			
			GAUGGUCGCCGAC			

[1208]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CUGGAGUUGCAGC			
			AGGACCUGAUCGU			
			GCCCCUGGGCCAC			
			ACUCCCAGCCAGG			
			AGCACUUCCUGUU			
			CGAGAUCUUCCGC			
			CGGAGACUGCAGG			
			CCCUGACCUCAGG			
			CUGGAGCGUGGCC			
			GCCAGCCUGCAAC			
			GGCAGCGGGAGCU			
			GCUGAUGUACAAG			
			CGGAUCCUGCUGC			
			GGCUGCCCAGCAG			
			CGUGCUGUGCGGC			
			AGCAGCUUCCAGG			
			CCGAGCAGCCCAU			
			CACGGCCCGGUGC			
			GAGCAGUUCUUCC			
			ACCUGGUGAACAG			
			CGAGAUGCGGAAC			
			UUCUGCAGCCACG			
			GAGGCGCCCUGAC			
			CCAGGACAUCACC			
			GCCCACUUCUUCC			
			GGGGCCUGCUGAA			
			CGCCUGCCUGCGG			
			AGCAGAGACCCCA			
			GCCUGAUGGUGGA			
			CUUCAUCCCUGGCC			
			AAGUGCCAGACCA			
			AGUGUCCCCUGAU			
			CCUGACCUCCCGCC			
			CUGGUGUGGUGGC			

[1209]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCAGCCUGGAGCC			
			CGUGCUUCUGUGU			
			CGGUGGCGCGGC			
			ACUGCCAGAGCCC			
			UCUGCCCCGGGAG			
			CUGCAGAAGCUGC			
			AGGAGGGCCGGCA			
			GUUCGCCAGCGAC			
			UUCCUGUCUCCCG			
			AGGCCGCUAGCCC			
			AGCCCCAAACCCC			
			GACUGGCUGAGCG			
			CUGCCGCCCUGCA			
			CUUCGCCAUCCAG			
			CAGGUGCGGGAGG			
			AGAACAUCCGGAA			
			GCAGCUGAAGAAG			
			CUGGACUGCGAGC			
			GGGAGGAGCUGCU			
			CGUGUUCCUCUUC			
			UUCUUCAGCCUCA			
			UGGGCCUUCUGAG			
			CAGCCACCUGACC			
			AGCAACAGCACCA			
			CCGACCUGCCCAA			
			GGCCUUCCACGUG			
			UGCGCCGCCAUCC			
			UGGAGUGCCUGGA			
			GAAGCGGAAGAUC			
			AGCUGGCUGGCCC			
			UGUUCCAGCUGAC			
			CGAGAGCGACCUG			
			CGGCUGGGCCGAC			
			UGCUGCUGCGGGU			

[1210]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
SEQ ID NO:		1	15	50	139	24
	FANCA_04	MSDSWVPNSASG	AUGAGCGACAGCU	GGAAAUCG	UAAAGCUCCCC	SEQ ID NO:24 从

[1211]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
	化学：G5 帽：C1 多聚 A 尾： 100nt	QDPGGRRRAWAE LLAGRVKREKYNP ERAQKLKESAVRL LRSHQDLNALLLE VEGPLCKKLSLSK VIDCDSSEAYANH SSSFIGSALQDQAS RLGVPVGILSAGM VASSVGQICTAPAE TSHPVLLTVEQRK KLSSLLEFAQYLL AHSMFSRLSFCQE LWKIQSSLLLEAV WHLHVQGIVSLQE LLESHPDMHAVGS WLFRNLCCLCEQ MEASCQHADVAR AMLSDFVQMFVL RGFQKNSDLRRTV EPEKMPQVTVDVL QRMLIFALDALAA GVQEESSTHKIVR CWFGVFSGHTLGS VISTDPLKRFFSHT LTQILTHSPVLKAS DAVQMQREWSFA RTHPLLTSLYRRLF VMLSAEELVGHLQ EVLETQEVHWQR VLSFVSALVVCFPE AQQLLEDWVARL MAQAFESCQLDS MVTAFLVVRQAAL EGPSAFLSYADWF	GGGUGCCCAACAG CGCCAGUGGCCAA GACCCCGGCGGUC GGAGACGAGCGUG GGCUGAGCUGCUG GCCGGACGGGUGA AGCGGGAGAAGUA CAACCCCGAGCGG GCCCAGAAGCUGA AGGAGAGCGCUGU GCGGCUGCUGCGG AGCCACCAGGACC UGAACGCCCUGCU GCUGGAGGUGGAG GGCCCACUGUGCA AGAAGCUGAGCCU GAGCAAGGUGAUC GACUGCGACAGCA GCGAGGCCUACGC CAACCACAGCAGC AGCUUCAUCGGCA GCGCCCUGCAGGA CCAGGCCAGCAGA CUGGGCGUGCCCG UGGGCAUCCUGAG CGCCGGCAUGGUG GCCAGCAGCGUGG GCCAGAUCUGCAC CGCCCCGGCCGAG ACAAGCCACCCCG UGCUGCUGACCGU GGAGCAGCGGAAG UGCUGGAGUUCGC	CAAAAUUU GCUCUUCG CGUUAGAU UUCUUUUA GUUUUCUC GCAACUAG CAAGCUUU UUGUUCUC GCC	GGGGGCCUCGG UGGCCUAGCUU CUUGCCCCUUG GGCCUCCCCCC AGCCCCUCCUC CCCUUCCUGCA CCCGUACCCCC GUGGUCUUUGA AUAAAGUCUGA GUGGGCGGC	5'至 3'端由以下组成：SEQ ID NO:50 的 5' UTR、SEQ ID NO:15 的 ORF 序列和 SEQ ID NO:139 的 3' UTR

[1212]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		KASFGSTRGYHGC	CCAGUACCUGCUG			
		SKKALVFLFTFLSE	GCCCACAGCAUGU			
		LVPFESPRYLQVHI	UCAGCCGGCUGAG			
		LHPPLVPGKYRSL	CUUCUGCCAGGAG			
		LTDYISLAKTRLA	CUGUGGAAGAUCC			
		DLKVSIENMGLYE	AGAGCAGCCUGCU			
		DLSSAGDITEPHSQ	GCUGGAGGCCGUG			
		ALQDVEKAIMVFE	UGGCACCUGCACG			
		HTGNIPVTVMEASI	UGCAGGGCAUCGU			
		FRRPYYVSHFLPA	GAGCCUGCAGGAG			
		LLTPRVLPKVPDSR	CUGCUGGAGAGCC			
		VAFIESLKRADKIP	ACCCCGACAUGCA			
		PSLYSTYCQACSA	CGCCGUGGGGCAGC			
		AEEKPEDAALGVR	UGGCUGUUCCGGA			
		AEPNSAEEPLGQL	ACCUGUGCUGCCU			
		TAALGELRASMTD	GUGCGAGCAGAUG			
		PSQRDVISAQVAVI	GAGGCCAGCUGCC			
		SERLRAVLGHNED	AGCACGCCGACGU			
		DSSVEISKIQLSINT	GGCUCGGGCCAUG			
		PRLEPREHMAVDL	CUGAGCGACUUCG			
		LLTSFCQNLMAAS	UGCAGAUGUUCGU			
		SVAPPERQGPWAA	GCUGCGGGGCUUC			
		LFVRTMCGRVLPA	CAGAAGAACAGCG			
		VLTRLCQLLRHQG	ACCUGCGGCGGAC			
		PSLSAPHVLGLAA	CGUAGAGCCCGAG			
		LAVHLGESRSALP	AAGAUGCCCCAGG			
		EVDVGPPAPGAGL	UGACCGUGGACGU			
		PVPALFDSLLTCRT	GCUGCAGCGGAUG			
		RDSLFFCLKFCTA	CUGAUCUUCGCAC			
		AISYSLCKFSSQSR	UCGACGCCCUGGC			
		DTLCSCLSPGLIKK	CGCUGGUGUGCAG			
		FQFLMFRLFSEAR	GAGGAGAGCAGCA			
		QPLSEEDVASLSW	CCCACAAGAUCGU			
		RPLHLPSADWQRA	GCGGUGCUGGUUC			

[1213]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		ALSLWTHRTFREV	GGCGUGUUCAGCG			
		LKEEDVHLTYQD	GCCACACCCUGGG			
		WLHLELEIQPEAD	CAGCGUGAUCAGC			
		ALSDTERQDFHQ	ACCGAUCCCCUGA			
		WAIHEHFLPESSAS	AGCGGUUCUUCAG			
		GGCDGDLQAACTI	CCACACCCUGACC			
		LVNALMDFHQSSR	CAGAUCCUGACCC			
		SYDHSENSDLVFG	ACAGCCCCGUGCU			
		GRTGNEDIISRLQE	GAAGGCCAGCGAC			
		MVADLELQQDLIV	GCCGUGCAGAUGC			
		PLGHTPSQEHFLFE	AGCGGGAGUGGAG			
		IFRRRLQALTSGW	CUUCGCCCGGACC			
		SVAASLQRQRELL	CAUCCCCUGCUGA			
		MYKRILLRLPSSVL	CCAGCCUGUACCG			
		CGSSFQAEQPITAR	GCGGCUGUUCGUG			
		CEQFFHLVNSEMR	AUGCUGAGCGCCG			
		NFCSHGGALTQDI	AGGAGCUGGUGGG			
		TAHFFRGLLNACL	CCACCUGCAGGAG			
		RSRDPSLMVDFIL	GUGCUGGAGACCC			
		AKCQTKCPLILTSA	AGGAGGUGCACUG			
		LVWWPSLEPVLLC	GCAGCGGGUGCUG			
		RWRRHCQSPLPRE	AGCUUCGUGAGCG			
		LQKLQEGRQFASD	CCCUGGUGGUGUG			
		FLSPEAASPAPNPD	CUUCCCCGAGGCC			
		WLSAAALHFAIQQ	CAGCAGCUGCUGG			
		VREENIRKQLKKL	AGGACUGGGUGGC			
		DCEREELLVFLFFF	CCGGCUGAUGGCC			
		SLMGLLSSHLTSNS	CAGGCCUUCGAGA			
		TTDLPKAFHVCAA	GCUGCCAGCUGGA			
		ILECLEKRKISWLA	CAGCAUGGUGACC			
		LFQLTESDLRLGRL	GCCUUCCUGGUGG			
		LLRVAPDQHTRLL	UAAGGCAAGCCGC			
		PFAFYSLLSYFHED	CCUGGAGGGUCCC			
		AAIREEAFLHVAV	AGCGCCUUCCCUGA			

[1214]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		DMYLKLVQLFVAG DTSTVSPPAGRSLE LKGQGNPVELITK ARLFLLQLIPRCPK KSFSHVAELLADR GDCDPEVSAALQS RQQAAPDADLSQE PHLF	GCUACGCCGACUG GUUCAAGGCCAGC UUCGGCAGCACCC GGGGCUACCACGG CUGCAGCAAGAAG GCCCUGGUGUUCC UGUUCACCUUCCU GAGCGAGCUGGGUG CCCUUCGAGUCAC CCCGGUACCUGCA GGUGCACAUCCUG CACCCACCCCUGG UGCCCGGCAAGUA CCGGAGCCUGCUG ACCGACUACAUCA GCCUGGCCAAGAC CCGGCUGGCCGAC CUGAAGGUGAGCA UCGAGAACAUGGG CCUGUACGAGGAC CUGAGCAGCGCCG GCGACAUCACCGA GCCCCACAGCCAG GCCCUGCAGGACG UGGAGAAGGCCAU CAUGGUGUUCGAG CACACCGGCAACA UCCCCGUGACCGU GAUGGAGGCCAGC ACUUCCGGCGGC CCUACUACGUGAG CCACUUCCUGCCC GCCCUGCUGACUC CCCGGGUGCUGCC			

[1215]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CAAGGUGCCCGAC			
			AGCCGGGUGGCCU			
			UCAUCGAGAGCCU			
			GAAGCGGGCCGAC			
			AAGAUUCCUCCCA			
			GCCUGUACAGCAC			
			CUACUGCCAGGCC			
			UGCAGUGCCGCCG			
			AGGAGAAGCCCGA			
			AGACGCCGCUCUG			
			GGCGUGAGAGCCG			
			AGCCCAACAGCGC			
			CGAGGAGCCUUUG			
			GGCCAGCUGACAG			
			CAGCCCUUGGCGA			
			GCUGCGGGCCAGC			
			AUGACCGACCCCA			
			GCCAGCGGGACGU			
			GAUCAGCGCCCAG			
			GUGGCCGUGAUCU			
			CCGAGAGACUGCG			
			GGCCGUGCUGGGC			
			CACAACGAGGACG			
			ACAGCAGCGUGGA			
			GAUCAGCAAGAUC			
			CAGCUGAGCAUCA			
			ACACUCCCCGGCU			
			GGAGCCCAGGGAG			
			CACAUGGCCGUGG			
			ACCUGCUGCUGAC			
			CAGCUUCUGCCAG			
			AACCUGAUGGCCG			
			CCUCAAGCGUGGC			
			CCCUCCAGAGCGG			

[1216]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CAAGGCCCUUGGG			
			CGGCCCUGUUCGU			
			GCGGACCAUGUGC			
			GGCCGGGUGUUGC			
			CUGCCGUGCUGAC			
			CCGGCUGUGCCAG			
			CUGCUGCGGCACC			
			AGGGCCCUAGCCU			
			GAGCGCUCCCCAC			
			GUGCUGGGGCUGG			
			CCGCCUUAGCAGU			
			GCACCUGGGCGAG			
			UCACGGAGCGCCC			
			UGCCCGAGGUGGA			
			CGUGGGCCCUCCU			
			GCGCCAGGAGCAG			
			GGCUUCCCGUGCC			
			CGCCCUGUUCGAC			
			AGCCUGCUGACCU			
			GCCGGACCCGGGA			
			CAGCCUGUUCUUC			
			UGCCUGAAGUUCU			
			GCACCGCCGCCAU			
			CAGCUACAGCCUG			
			UGCAAGUUCAGCA			
			GCCAGAGCCGGGA			
			CACCCUGUGCAGC			
			UGCCUGAGCCCCG			
			GCCUGAUCAAGAA			
			GUUCCAGUUCCUG			
			ALGUUCCGGCUGU			
			UCAGCGAGGCCCG			
			GCAACCCCUGAGC			
			GAGGAGGACGUGG			

[1217]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCAGCCUGAGCUG			
			GCGGCCCCUGCAC			
			CUUCCCAGCGCCG			
			ACUGGCAACGCGC			
			CGCCCUGAGCCUG			
			UGGACCCACCGGA			
			CCUUCCGGGAGGU			
			GCUGAAGGAGGAG			
			GACGUGCACCUGA			
			CCUACCAGGACUG			
			GCUGCACCUGGAG			
			CUGGAGAUUCAGC			
			CCGAGGCCGACGC			
			CCUGAGCGACACC			
			GAGCGGCAGGACU			
			UCCACCAGUGGGC			
			CAUCCACGAGCAC			
			UUCCUGCCCGAGA			
			GCAGUGCCAGCGG			
			CGGCUGUGACGGC			
			GACCUGCAGGCCG			
			CCUGCACCAUCCU			
			GGUGAACGCCCUG			
			AUGGACUUCCACC			
			AGAGCAGCCGGAG			
			CUACGACCACAGC			
			GAGAACAGCGACC			
			UGGUGUUCGGAGG			
			CCGGACCGGCAAC			
			GAGGACAUCAUCA			
			GCCGGCUGCAGGA			
			GAUGGUGGCCGAC			
			CUGGAGCUGCAGC			
			AGGACCUGAUCGU			

[1218]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GCCCCUGGGCCAC			
			ACCCCUAGCCAGG			
			AGCACUUCCUGUU			
			CGAGAUCUUUCGG			
			CGGCGGCUGCAAG			
			CGCUGACCAGCGG			
			UUGGAGCGUGGCC			
			GCCUCCCUGCAAC			
			GGCAGCGGGAGCU			
			GCUGAUGUACAAG			
			CGGAUCCUGCUGC			
			GGCUGCCCAGCAG			
			CGUGCUGUGCGGC			
			AGCAGCUUCCAGG			
			CCGAGCAGCCCAU			
			CACCGCCCGGUGC			
			GAGCAGUUCUUCC			
			ACCUGGUGAACAG			
			CGAGAUGCGGAAC			
			UUCUGCAGCCACG			
			GUGGCGCCCUGAC			
			CCAGGACAUCACC			
			GCCCACUUCUUCC			
			GGGGCCUGCUGAA			
			CGCCUGCCUGCGU			
			AGCCGGGACCCCA			
			GCCUGAUGGUGGA			
			CUUCAUCCUGGCC			
			AAGUGCCAGACCA			
			AGUGCCCACUGAU			
			CCUGACCAGCGCC			
			CUGGUGUGGUGGC			
			CCAGCCUGGAGCC			
			CGUGCUGCUGUGC			

[1219]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CGGUGGCGCACGGC			
			ACUGCCAGUCACC			
			GCUGCCCCCGGGAG			
			CUGCAGAAGCUGC			
			AGGAGGGCCGGCA			
			GUUCGCCAGCGAC			
			UUCCUGAGCCCCG			
			AGGCUGCCAGCCC			
			UGCCCCCGAACCCC			
			GACUGGCUGAGUG			
			CCGCCGCCCUGCA			
			CUUCGCCAUCCAG			
			CAGGUGCGGGAGG			
			AGAACAUCCGGAA			
			GCAGCUGAAGAAG			
			CUGGACUGCGAGC			
			GGGAGGAGCUGCU			
			GGUGUUCCUGUUC			
			UUCUUCAGCCUGA			
			UGGGCCUGCUGAG			
			CAGCCACCUGACC			
			AGCAACAGCACCA			
			CCGACCUGCCCAA			
			GGCCUUCCACGUG			
			UGCGCCGCCAUCC			
			UGGAGUGCCUGGA			
			GAAGCGGAAGAUC			
			AGCUGGCUGGCCC			
			UGUUCCAGCUGAC			
			CGAGUCUGACCUG			
			CGCCUGGGCAGGC			
			UGCUGCUGCGGGU			
			GGCUCCCCGACCAG			
			CACACCCGGCUGC			

[1220]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			UGCCCUUCGCCUU CUACAGCCUGCUG AGCUACUUCCACG AGGACGCCGCCAU CCGGGAGGAGGCC UUCCUGCACGUGG CCGUGGACAUGUA CCUGAAGCUGGUG CAGCUGUUCGUGG CCGGCGACACCAG CACUGUAUCUCCA CCGGCUGGGCGGU CCCUGGAGCUGAA GGGCCAGGGCAAC CCCGUGGAGCUGA UCACCAAGGCCCG GCUGUUCCUGCUG CAGCUGAUCCCAC GGUGCCCCAAGAA GAGCUUCAGCCAC GUGGCCGAGCUGC UGGCCGACCGGGG AGACUGCGACCCC GAGGUUAGUGCCG CCCUCCAGAGCCG GCAGCAAGCCGCC CCUGACGCCGACC UGAGCCAGGAGCC CCACCUGUUC			
SEQ ID NO:		1	16	50	139	25
	FANCA_05 化学：G5 帽：C1	MSDSWVPNSASG QDPGGRRRAWAE LLAGRVKREKYNP	ALGUCCGACUCUU GGGUCCCCAACUC CGCCUCUGGUCAA	GGAAAUCG CAAAAUUU GCUCUUCG	UAAAGCUCCCC GGGGGCCUCGG UGGCCUAGCUU	SEQ ID NO:25 从5'至 3'端由以下组成：SEQ ID

[1221]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
	多聚 A 尾; 100nt	ERAQKLKESAVRL LRSHQDLNALLLLE VEGPLCKKLSLSLSK VIDCDSSEAYANH SSSFIGSALQDQAS RLGVPVGILSAGM VASSVGQICTAPAE TSHPVLLTVEQRK KLSSLLEFAQYLL AHSMFSRLSFCQE LWKIQSSLLLEAV WHLHVQGIVSLQE LLESHPDMHAVGS WLFRNLCCLCEQ MEASCQHADVAR AMLSDFVQMFVL RGFQKNSDLRRTV EPEKMPQVTVDVL QRMLIFALDALAA GVQEESSTHKIVR CWFGVFSGHTLGS VISTDPLKRFFSHT LTQILTHSPVLKAS DAVQMQREWSFA RTHPLLTSLYRRLF VMLSAEELVGHLQ EVLETQEVHWQR VLSFVSALVVCFPE AQQLLEDWVARL MAQAFESCQLDS MVTAFLVVRQAAL EGPSAFLSYADWF KASFGSTRGYHGC SKKALVFLFTFLSE	GAUCCGGGAGGAA GGAGACGUGCCUG GGCGGAACUGUUA GCCGGACGUGUGA AGAGGGAGAAGUA CAAUCCCGAGCGC GCACAGAAGCUCA AGGAAUCCGCAGU UAGGCUCCUCCGC UCUCACCAGGACC UCAACGCACUGCU CCUCGAGGUCGAG GGACCACUCUGCA AGAAGCUCUCCCU CUCCAAGGUCAUC GACUGCGACUCCU CCGAGGCCUACGC CAACCACUCCUCC UCCUUCAUCGGCU CCGCCCUCCAAGA UCAAGCUUCCCGC CUCGGAGUUCCUG UGGGCAUCCUAAG CGCCGGGAUGGUC GCAUCUUCCGUCG GCCAGAUUUGCAC AGCACCCGCCGAG ACAUCCCAUCCUG UCCUUCUCACCGU CGAGCAGCGCAAG AAGCUCUCCUCCC UCCUCGAGUUUGC CCAGUACCUCCUC GCCCACUCCAUGU	CGUUAGAU UUCUUUUA GUUUUCUC GCAACUAG CAAGCUUU UUGUUCUC GCC	CUUGCCCCUUG GGCCUCCCCCCC AGCCCCUCCUC CCCUUCCUGCA CCCGUACCCCC GUGGUCUUUGA AUAAAGUCUGA GUGGGCGGC	NO:50 的 5' UTR 、SEQ ID NO:16 的 ORF 序列和 SEQ ID NO:139 的 3' UTR

[1222]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		LVPFESPRYLQVHI	UCUCCCGCCUCUC			
		LHPPLVPGKYRSL	CUUCUGCCAGGAG			
		LTDYISLAKTRLA	CUCUGGAAGAUCC			
		DLKVSIENMGLYE	AGUCCUCCCUCCU			
		DLSSAGDITEPHSQ	GCUCGAAGCCGUC			
		ALQDVEKAIMVFE	UGGCAUCUCCACG			
		HTGNIPVTVMEASI	UCCAGGGCAUCGU			
		FRRPYYVSHFLPA	CUCCCUCCAGGAG			
		LLTPRVLPKVPDSR	CUCCUCGAAUCCC			
		VAFIESLKRADKIP	AUCCAGACAUGCA			
		PSLYSTYCQACSA	CGCCGUCGGUUCU			
		AEEKPEDAALGVR	UGGCUCUUCCGCA			
		AEPNSAEEPLGQL	ACCUCUGCUGCCU			
		TAALGELRASMTD	CUGCGAGCAGAUG			
		PSQRDVISAQVAVI	GAAGCCUCCUGCC			
		SERLRAVLGHNED	AACACGCUGACGU			
		DSSVEISKIQLSINT	CGCUCCGAGCCAUG			
		PRLEPREHMAVDL	CUCUCCGACUUCG			
		LLTSFCQNLMAAS	UCCAGAUGUUCGU			
		SVAPPERQGPWAA	CCUCCGCGGCUUC			
		LFVRTMCGRVLPA	CAGAAGAAUUCCG			
		VLTRLCQLLRHQG	ACCUCCGUCGCAC			
		PSLSAPHVLGLAA	AGUCGAGCCUGAG			
		LAVHLGESRSALP	AAGAUGCCUCAGG			
		EVDVGPPAPGAGL	UCACCGUCGACGU			
		PVPALFDSLLTCRT	CCUCCAGAGAAUG			
		RDSLFFCLKFCTA	CUCAUCUUCGCAC			
		AISYSLCKFSSQSR	UGGACGCACUAGC			
		DTLCSCLSPGLIKK	UGCAGGCGUCCAG			
		FQFLMFRLFSEAR	GAGGAGUCCUCCA			
		QPLSEEDVASLSW	CCCACAAGAUCGU			
		RPLHLPSADWQRA	CCGCUGUUGGUUC			
		ALSLWTHRTFREV	GGCGUCUUUAGCG			
		LKEEDVHLTYQD	GGCACACCUUGGG			

[1223]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		WLHLELEIQPEAD	GUCAGUCAUCUCC			
		ALSDTERQDFHQ	ACCGAUCCCCUCA			
		WAIHEHFLPESSAS	AGCGCUUCUUCUC			
		GGCDGDLQAACTI	CCACACCCUCACC			
		LVNALMDFHQSSR	CAGAUCCUCACCC			
		SYDHSENSDLVFG	ACUCCCCAGUCCU			
		GRTGNEDIISRLQE	CAAAGCCUCCCGAC			
		MVADLELQQDLIV	GCCGUCCAGAUGC			
		PLGHTPSQEHFLFE	AACGCGAGUGGUC			
		IFRRRLQALTSGW	CUUUGCCCGAACC			
		SVAASLQRQRELL	CACCCACUCCCUCA			
		MYKRILLRLPSSVL	CCUCCCUCUACCG			
		CGSSFQAEQPITAR	CAGGCUCUUCGUC			
		CEQFFHLVNSEMR	AUGCUCUCCGCCG			
		NFCSHGGALTQDI	AAGAACUCGUCGG			
		TAHFFRGLLNACL	CCACUUGCAGGAG			
		RSRDPSLMVDFIL	GUCCUCGAGACCC			
		AKCQTKCPLILTSA	AAGAAGUCCACUG			
		LVWWPSLEPVLLC	GCAGCGCGUCCUC			
		RWRRHCQSPLPRE	UCCUUCGUCUCCG			
		LQKLQEGRQFASD	CCCUCGUCGUCUG			
		FLSPEAASPAPNPD	UUUUCCAGAGGCC			
		WLSAAALHFAIQQ	CAGCAGCUCCUUG			
		VREENIRKQLKKL	AGGAUUGGGUGGC			
		DCEREELLVFLFFF	ACGGCUCAUGGCU			
		SLMGLLSSHLTSNS	CAGGCCUUCGAGU			
		TTDLPKAFHVCAA	CCUGCCAGCUCGA			
		ILECLEKRKISWLA	CUCCAUGGUCACC			
		LFQLTESDLRLGRL	GCCUUUCUUGUUG			
		LLRVAPDQHTRLL	UCCGCCAAGCUGC			
		PFAFYSLLSYFHED	UCUCGAGGGUCCU			
		AAIREEAFLHVAV	UCGGCAUUCCUGA			
		DMYLKLVQLFVAG	GCUACGCCGACUG			
		DTSTVSPPAGRSLE	GUUCAAGGCCUCC			

[1224]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		LKGQGNPVELITK ARLFLLQLIPRCPK KSFSHVAELLADR GDCDPEVSAALQS RQQAAPDADLSQE PHLF	UUCGGUAGCACAA GGGGCUACCACGG UUGCUCCAAGAAG GCCCUCGUCUUCC UCUUCACCUUCCU CUCCGAGCUCGUC CCCUUCGAGAGUC CCCGCUACCUCCA GGUCCACAUCCUC CAUCCUCCUCUGG UCCCCGGAAAGUA CCGCUCCCUCCUC ACCGACUACAUCU CCCUCGCUAAGAC CCGACUCGCCGAC UUGAAGGUCUCCA UCGAGAACAUGGG CCUCUACGAAGAU CUCUCCAGUGCAG GCGACAUCACCGA GCCACACUCCCAA GCCCUCCAGGACG UCGAGAAGGCCAU CAUGGUCUUCGAG CACACCGGCAACA UCCCCGUCACCGU CAUGGAGGCCUCC AUCUUCAGACGCC CCUACUACGUCUC CCACUUUCUCCCU GCCCUCCUAACCC CUCGGGUACUCCC CAAGGUCCCCCGAC UCCAGAGUCGCCU			

[1225]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			UCAUCGAGUCCCU			
			CAAGCGCGCAGAC			
			AAGAUCCCACCAU			
			CCCUCUACUCCAC			
			CUAUUGUCAAGCC			
			UGUUCUGCCGCCG			
			AAGAGAAACCCGA			
			GGACGCCGCAUUA			
			GGCGUCCGUGCAG			
			AGCCCAACUCAGC			
			UGAGGAGCCACUG			
			GGCCAGCUAACAG			
			CUGCACUAGGCGA			
			GCUGAGGGCGAGU			
			AUGACCGACCCUU			
			CCCAGCGAGACGU			
			CAUCUCCGCCCAA			
			GUCGCCGUCAUUU			
			CCGAACGCUUGAG			
			GGCAGUGCUGGGC			
			CACAACGAGGACG			
			ACUCCUCCGUCGA			
			GAUCUCCAAGAUC			
			CAGCUCUCCAUCA			
			AUACACCUCGCCU			
			GGAACCCCGAGAG			
			CACAUGGCCGUCG			
			ACCUCCUCCUCAC			
			CUCCUUCUGCCAG			
			AACUUGAUGGCUG			
			CCUCCUCCGUUGC			
			UCCUCCGGAAAGA			
			CAAGGUCCCUGGG			
			CCGCUCUCUUUGU			

[1226]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCGCACCAUGUGU			
			GGACGAGUGCUCC			
			CAGCUGUCCUCAC			
			UCGCCUCUGCCAG			
			UUGCUCCGCCACC			
			AAGGACCCUCACU			
			CUCUGCACCGCAC			
			GUACUGGGACUCG			
			CAGCACUCGCAGU			
			CCACCUCGGAGAA			
			UCCCGCUCUGCUC			
			UCCCAGAAGUGGA			
			CGUGGGACCUCCU			
			GCACCUGGUGCUG			
			GCUUGCCCGUACC			
			AGCUCUCUUCGAC			
			UCCCUCCUCACGU			
			GCAGAACCCGCGA			
			CUCCCUCUUCUUC			
			UGCCUCAAGUUCU			
			GCACCGCCGCCAU			
			CUCCUACUCCCUC			
			UGCAAGUUCAGCU			
			CCCAGUCACGCGA			
			CACCCUCUGCUCC			
			UGCUUGUCCCCUG			
			GCCUCAUCAAGAA			
			GUUCCAGUUCCUC			
			AUGUUCCGCCUCU			
			UCUCCGAGGCCAG			
			ACAACCCCUCUCC			
			GAGGAGGACGUGG			
			CUUCCCUCAGUUG			
			GAGACCCCUCCAC			

[1227]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			UUGCCUUCCGCCG			
			AUUGGCAAAGAGC			
			CGCGCUCUCUCUC			
			UGGACCCAUCGCA			
			CCUUCCGCGAGGU			
			CCUCAAGGAGGAG			
			GACGUCCACCUCA			
			CCUACCAGGACUG			
			GCUCCACCUCGAG			
			CUCGAGAUCCAAC			
			CCGAAGCAGACGC			
			CCUCUCCGACACC			
			GAACGCCAGGACU			
			UCCACCAGUGGGC			
			CAUCCACGAGCAC			
			UUCCUCCCCGAAA			
			GCAGUGCUUCCGG			
			UGGUUGCGACGGG			
			GAUCUCCAAGCCG			
			CUUGCACCAUCCU			
			CGUCAACGCCCUC			
			ALGGACUUCCACC			
			AGUCCUCCCGCUC			
			CUACGACCACUCC			
			GAGAACUCCGACC			
			UAGUGUUUGGCGG			
			CCGCACUGGAAAC			
			GAGGACAUCAUCU			
			CCCGCCUCCAGGA			
			GAUGGUGGCCGAU			
			CUCGAGCUCCAGC			
			AGGAUCUCAUCGU			
			CCCUCUCGGCCAC			
			ACACCUUCGCAGG			

[1228]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			AGCACUUCCUCUU			
			CGAAAUCUUCAGA			
			AGACGCCUCCAAG			
			CGCUCACUUCCGG			
			UUGGUCUGUGGCC			
			GCAAGUCUCCAAC			
			GCCAGAGGGAACU			
			CCUCAUGUACAAG			
			CGCAUCCUCCUAC			
			GCCUCCCUUCCUC			
			CGUCCUCUGUGGG			
			UCCUCCUUCCAGG			
			CAGAACAGCCCAU			
			CACUGCCCGCUGU			
			GAGCAGUUCUUCC			
			ACCUCGUCAACUC			
			CGAGAUGCGCAAC			
			UUCUGUUCACACG			
			GCGGUGCACUCAC			
			CCAAGACAUCACC			
			GCGCACUUCUUCC			
			GCGGACUCCUGAA			
			CGCCUGCUUGAGA			
			UCCCGCGACCCAU			
			CCCUCAUGGUCGA			
			CUUCAUCCUCGCC			
			AAGUGCCAGACCA			
			AGUGUCCCCUCAU			
			CCUCACCAGUGCC			
			CUCGUUUGGUGGC			
			CCUCCCUUGAACC			
			CGUCCUCCUCUGC			
			AGGUGGAGACGCC			
			ACUGUCAGUCUCC			

[1229]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GUUACCCCGCGAG			
			CUCCAGAAACUCC			
			AGGAGGGACGCCA			
			GUUCGCUUCCGAC			
			UUCCUCUCACCCG			
			AAGCUGCUAGCCC			
			AGCACCCAACCCU			
			GACUGGUUAUCCG			
			CCGCCGCUUUGCA			
			CUUCGCCAUCCAG			
			CAAGUGCGCGAGG			
			AGAACAUCCGCAA			
			GCAGCUCAAGAAG			
			CUCGACUGCGAGC			
			GAGAGGAGCUCCU			
			CGUUUUCCUCUUC			
			UUCUUCUCCCUCA			
			UGGGCCUCCUCUC			
			ACCCCACCUCACC			
			UCCAACUCCCACCA			
			CCGACUUGCCCAA			
			GGCCUUUCACGUC			
			UGUGCCGCCAUCC			
			UCGAGUGCCUCGA			
			GAAGCGCAAGAUC			
			UCCUGGCUCGCCC			
			UCUUCCAGCUAAC			
			AGAGUCCGACCUG			
			AGACUCGGUCGCC			
			UCCUACUCAGAGU			
			CGCACCCGACCAG			
			CACACUCGCCUCC			
			UCCCCUUCGCAUU			
			CUACUCCCUCCUC			

[1230]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			UCCUACUUCCACG AAGACGCAGCAAU CCGCGAGGAGGCC UUCCUUCACGUCG CCGUCGACAUGUA CCUCAAGCUCGUC CAGCUCUUUGUCG CGGGCGACACGAG UACAGUUAGCCCU CCAGCAGGUAGGU CCCUCGAGCUUAA AGGCCAGGGCAAC CCAGUCGAGCUCA UCACCAAGGCCCG CCUGUUCCUGCUA CAGCUCAUCCCAC GCUGCCCCAAGAA GUCCUUCUCCCAC GUAGCCGAACUCC UCGCCGAUCGCGG AGAUUGCGACCCC GAAGUUUCCGCUG CCUUGCAGUCACG GCAGCAAGCAGCC CCAGACGCAGACC UCUCACAGGAGCC CCACCUCUUC			
SEQ ID NO:		1	17	50	139	26
	FANCA_06 化学：G5 帽：C1 多聚 A 尾：100nt	MSDSWVPNSASG QDPGGRRRAWAE LLAGRVKREKYNP ERAQKLKESAVRL LRSHQDLNALLLE	AUGUCAGAUUCCU GGGUGCCCAACUC CGCUAGCCGGACAA GAUCCAGGUGGCC GCAGGAGAGCCUG	GGAAAUCG CAAAAUUU GCUCUUCG CGUUAGAU UUCUUUUA	UAAAGCUCCCC GGGGGCCUCGG UGGCCUAGCUU CUUGCCCCUUG GGCCUCCCCCC	SEQ ID NO:26 从 5'至 3'端由以下组成：SEQ ID NO:50 的 5' UTR

[1231]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		VEGPLCKKLSLSK	GGCUGAACUGCUU	GUUUUCUC	AGCCCCUCCUC	、SEQ ID NO:17
		VIDCDSSEAYANH	GCCGGGAGAGUGA	GCAACUAG	CCCUUCCUGCA	的 ORF 序列和
		SSSFIGSALQDQAS	AGAGAGAGAAGUA	CAAGCUUU	CCCGUACCCCC	SEQ ID NO:139 的
		RLGVPVGILSAGM	CAACCCAGAGCGC	UUGUUCUC	GUGGUCUUUGA	3' UTR
		VASSVGQICTAPAE	GCCCAGAAGUUAA	GCC	AUAAAGUCUGA	
		TSHPVLLTVEQRK	AGGAGAGCGCCGU		GUGGGCGGC	
		KLSSLLEFAQYLL	AAGACUGCUGCGC			
		AHSMFSRLSFCQE	AGCCACCAAGACU			
		LWKIQSSLLLEAV	UAAACGCUCUGCU			
		WHLHVQGIVSLQE	GCUGGAGGUGGAG			
		LLESHPDMHAVGS	GGGCCCCUGUGCA			
		WLFRNLCCLCEQ	AGAAGCUGAGCCU			
		MEASCQHADVAR	GUCCAAGGUGAUC			
		AMLSDFVQMFVL	GAUUGUGACAGCA			
		RGFQKNSDLRRTV	GCGAGGCUUACGC			
		EPEKMPQVTVDVL	CAAUCAUUCAUCU			
		QRMLIFALDALAA	UCCUUCAUCGGCU			
		GVQEESSTHKIVR	CAGCACUGCAGGA			
		CWFGVFSGHTLGS	UCAAGCAAGCCGG			
		VISTDPLKRFFSHT	CUGGGCGUUCCAG			
		LTQILTHSPVLKAS	UUGGCAUACUGAG			
		DAVQMQREWSFA	CGCCGGCAUGGUG			
		RTHPLLTSLYRRLF	GCCAGCAGCGUGG			
		VMLSAEELVGHLQ	GACAGAUCUGCAC			
		EVLETQEVHWQR	CGCUCCCGCCGAG			
		VLSFVSALVVCFPE	ACCUCACAUCCCG			
		AQQLLEDWVARL	UUCUGGCUGACAGU			
		MAQAFESCQLDS	UGAGCAACGCAAG			
		MVTAFLVVRQAAL	AAGCUGUCAAGCC			
		EGPSAFLSYADWF	UGCUGGAGUUCGC			
		KASFGSTRGYHGC	ACAGUACCUGUUA			
		SKKALVFLFTFLSE	GCCCACAGCAUGU			
		LVPFESPRYLQVHI	UUUCCCGCUUAAG			
		LHPPLVPGKYRSL	CUUCUGCCAGGAG			

[1232]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		LTDYISLAKTRLA	CUGUGGAAGAUCC			
		DLKVSIENMGLYE	AGUCAAGUCUGCU			
		DLSSAGDITEPHSQ	ACUGGAGGCCGUG			
		ALQDVEKAIMVFE	UGGCACUUACACG			
		HTGNIPVTVMEASI	UGCAGGGUAUCGU			
		FRRPYYVSHFLPA	GUCACUGCAAGAG			
		LLTPRVLPKVPDSR	CUACUGGAGAGCC			
		VAFIESLKRADKIP	AUCCCGAUAUGCA			
		PSLYSTYCQACSA	CGCCGUGGGGUCU			
		AEEKPEDAALGVR	UGGCUGUUCCGCA			
		AEPNSAEEPLGQL	AUCUGUGCUGCCU			
		TAALGELRASMTD	GUGCGAACAGAUG			
		PSQRDVISAQVAVI	GAAGCCUCCUGCC			
		SERLRAVLGHNED	AACACGCCGACGU			
		DSSVEISKIQLSINT	UGCUAGAGCAAUG			
		PRLEPREHMAVDL	CUGAGUGACUUCG			
		LLTSFCQNLMAAS	UUCAGAUGUUCGU			
		SVAPPERQGPWAA	GCUGCGCGGGUUC			
		LFVRTMCGRVLPA	CAGAAGAACUCAG			
		VLTRLCQLLRHQG	AUCUCCGGAGGAC			
		PSLSAPHVLGLAA	CGUGGAGCCUGAG			
		LAVHLGESRSALP	AAGAUGCCCCAGG			
		EVDVGPPAPGAGL	UAACCGUGGACGU			
		PVPALFDSLLTCRT	GCUGCAGCGGAUG			
		RDSLFFCLKFCTA	UUAAUCUUUGCCC			
		AISYSLCKFSSQSR	UGGACGCUCUCGC			
		DTLCSCLSPGLIKK	UGCCGGAGUGCAG			
		FQFLMFRLFSEAR	GAGGAGAGCUCCA			
		QPLSEEDVASLSW	CCCACAAGAUCGU			
		RPLHLPSADWQRA	GCGGUGCUGGUUC			
		ALSLWTHRTFREV	GGCGUGUUUUCCG			
		LKEEDVHLTYQD	GCCAUACCUUAGG			
		WLHLELEIQPEAD	CUCCGUGAUCUCC			
		ALSDTERQDFHQ	ACCGAUCCCCCUGA			

[1233]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		WAIHEHFLPESSAS	AACGGUUCUUCUC			
		GGCDGDLQAACTI	CCACACACUGACC			
		LVNALMDFHQSSR	CAGAUACUGACCC			
		SYDHSENSDLVFG	ACUCCCCUGUGUU			
		GRTGNEDIISRLQE	AAAGGCUAGCGAC			
		MVADLELQQDLIV	GCCGUGCAAAUGC			
		PLGHTPSQEHFLFE	AACGCGAGUGGUC			
		IFRRRLQALTSGW	CUUUGCCCGCACC			
		SVAASLQRQRELL	CAUCCACUGCUGA			
		MYKRILLRLPSSVL	CCAGUCUGUACCG			
		CGSSFQAEQPITAR	GCGGCUGUUUGUU			
		CEQFFHLVNSEMR	AUGCUGUCCGCCG			
		NFCSHGGALTQDI	AGGAGUUAGUGGG			
		TAHFFRGLLNACL	CCACCUGCAGGAG			
		RSRDPSLMVDFIL	GUUCUGGAAACCC			
		AKCQTKCPLILTSA	AGGAGGUGCAUUG			
		LVWWPSLEPVLLC	GCAGCGCGUGUUA			
		RWRRHCQSPLPRE	AGCUUCGUGAGCG			
		LQKLQEGRQFASD	CCCUGGUGGUGUG			
		FLSPEAASPAPNPD	CUUUCCCGAGGCC			
		WLSAAALHFAIQQ	CAGCAGUUACUGG			
		VREENIRKQLKKL	AGGACUGGGUGGC			
		DCEREELLVFLFFF	CCGGCUGAUGGCC			
		SLMGLLSSHLTSNS	CAGGCCUUCGAGA			
		TTDLPKAFHVCAA	GCUGCCAGCUGGA			
		ILECLEKRKISWLA	CAGCAUGGUGACC			
		LFQLTESDLRLGRL	GCCUUCCUGGUGG			
		LLRVAPDQHTRLL	UACGUCAAGCUGC			
		PFAFYSLLSYFHED	CCUGGAGGGCCCA			
		AAIREEAFLHVAV	AGCGCCUUCCUGA			
		DMYLKLVQLFVAG	GUUACGCUGACUG			
		DTSTVSPPAGRSLE	GUUCAAGGCCAGC			
		LKGQGNPVELITK	UUCGGGUCCACCC			
		ARLFLLQLIPRCPK	GCGGGUACCACGG			

[1234]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		KSFSHVAELLADR GDCDPEVSAALQS RQQAAPDADLSQE PHLF	CUGCUCCAAGAAG GCCCUGGUGUUCC UGUUUACCUUCCU GAGCGAGCUGGUG CCAUUUGAGUCUC CCCGCUACCUGCA GGUGCACAUCCUG CACCCACCCCUGG UCCCUGGAAAGUA CAGAUCACUGUUA ACCGACUACAUCU CUCUGGCCAAGAC CCGGUUUAGCCGAC CUGAAGGUGUCUA UCGAGAAUAUGGG CCUGUACGAGGAC CUGUCCAGCGCCG GCGAUAUCACCGA GCCCCACAGCCAA GCCUUACAGGACG UGGAGAAGGCCAU UAUGGUGUUCGAG CACACCGGCAAUA UCCCCGUGACCGU AAUGGAGGCUUCU AUCUUCCGGCGGC CCUACUACGUGUC CCACUUCCUGCCC GCCCUGUUAACAC CCAGAGUGUUACC CAAGGUGCCCGAC AGUCGCGUUGCCU UCAUAGAGAGCUU AAAGAGAGCCGAC			

[1235]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			AAGAUCCCUCCUA			
			GUUUAUACUCCAC			
			CUACUGCCAGGCU			
			UGCUCCCGCCGCCG			
			AGGAGAAGCCCGA			
			GGACGCAGCCUUA			
			GGAGUAAGAGCUG			
			AGCCCAACUCAGC			
			CGAGGAGCCACUG			
			GGCCAGUUAACCG			
			CUGCCCUGGGAGA			
			GUUACGGGCCUCU			
			AUGACCGACCCCA			
			GCCAACGCGACGU			
			UAUCUCUGCCCAA			
			GUGGCAGUGAUAA			
			GCGAGAGGCUGCG			
			GGCAGUAUUAGGA			
			CACAACGAGGACG			
			ACAGCAGCGUGGA			
			GAUCAGUAAGAUC			
			CAAUUAAGCAUCA			
			ACACACCCAGGCU			
			GGAGCCCAGAGAG			
			CACAUGGCCGUGG			
			ACCUGCUGUUAAC			
			CUCCUUCUGUCAG			
			AACUUAAUGGCCG			
			CAAGUAGCGUGGC			
			UCCUCCCAGAACGC			
			CAAGGUCCUUGGG			
			CUGCCUUAUUCGU			
			UCGCACCAUGUGU			
			GGCCGCGUCCUGC			

[1236]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CAGCAGUGCUGAC			
			CCGGCUGUGCCAG			
			UUACUGCGGCACC			
			AAGGGCCCUCCCU			
			GAGCGCUCCUCAC			
			GUGCUUGGGUUGG			
			CCGCCCUCGCUGU			
			ACACUUAGGUGAA			
			AGCAGAAGCGCCC			
			UUCCUGAGGUGGA			
			CGUGGGUCCUCCA			
			GCCCCAGGUGCCG			
			GACUCCCCGUGCC			
			AGCCCUGUUCGAC			
			AGCUUACUGACUU			
			GUCGCACCCGGGA			
			CUCCCUGUUCUUU			
			UGUCUGAAGUUCU			
			GCACCGCCGCCAU			
			CUCCUACUCACUG			
			UGCAAGUUCUCUA			
			GCCAGUCCCGGGA			
			UACCCCUGUGCUCC			
			UGCCUGUCUCCAG			
			GCCUGAUCAAGAA			
			GUUCCAGUUCCUG			
			ALGUUCAGACUGU			
			UCUCAGAAGCAAG			
			ACAGCCUCUGUCC			
			GAGGAGGACGUGG			
			CCUCCUUAUCCUG			
			GCGACCCCUGCAC			
			CUGCCCUCCGCUG			
			AUUGGCAAAGAGC			

[1237]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			UGCACUUAGCCUG			
			UGGACCCACCGGA			
			CAUUUCGCGAGGU			
			GCUGAAAGAAGAA			
			GACGUACACCUGA			
			CCUACCAAGACUG			
			GCUGCACCUGGAG			
			CUGGAGAUCCAGC			
			CUGAAGCCGACGC			
			CCUGAGCGACACC			
			GAGCGCCAAGACU			
			UUCACCAGUGGGC			
			CAUACACGAGCAC			
			UUUCUGCCCGAGU			
			CAAGCGCCAGUGG			
			AGGCUGCGACGGU			
			GAUCUGCAAGCAG			
			CCUGCACCAUCUU			
			AGUGAACGCCCUG			
			ALGGACUUCCACC			
			AAUCUUCCCGCAG			
			CUACGACCACUCC			
			GAGAACAGCGACC			
			UGGUGUUUGGAGG			
			CCGCACUGGGAAC			
			GAGGACAUCAUUA			
			GCCGCCUGCAAGA			
			AAUGGUGGCUGAC			
			CUGGAACUCCAAC			
			AGGACUUAAUCGU			
			ACCUCUGGGGCAC			
			ACACCAAGCCAGG			
			AGCACUUCUUAUU			
			UGAGAUAUUUAGA			

[1238]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			AGACGCUUACAAG			
			CACUGACUUCAGG			
			CUGGAGUGUAGCU			
			GCAAGCCUGCAAC			
			GCCAACGCGAGCU			
			GCUGAUGUAUAAG			
			CGGAUACUGCUGC			
			GCCUGCCCUCCUC			
			UGUACUGUGCGGC			
			UCCUCCUUCCAGG			
			CCGAGCAGCCCAU			
			CACCGCCCGGUGU			
			GAGCAGUUCUUCC			
			ACCUGGUGAACAG			
			CGAGAUGCGGAAC			
			UUCUGUUCUCACG			
			GAGGCGCUCUGAC			
			CCAAGACAUCACA			
			GCCCACUUCUUCC			
			GCGGCCUGCUGAA			
			CGCCUGUCUGCGG			
			UCAAGGGACCCCA			
			GCCUGAUGGUCGA			
			CUUCAUACUGGCA			
			AAGUGUCAGACCA			
			AGUGUCCCCUGAU			
			CCUGACCAGCGCU			
			CUGGUUUGGUGGC			
			CAUCACUGGAGCC			
			CGUACUCCUGUGU			
			CGCUGGAGGCGGC			
			AUUGCCAGUCACC			
			UUUACCCCGAGAA			
			CUGCAGAAGCUGC			

[1239]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			AGGAGGGCCGGCA			
			GUUUGCCAGCGAC			
			UUCCUGUCACCAG			
			AGGCAGCCUCUCC			
			CGCUCCUAACCCC			
			GAUUGGCUAUCCG			
			CGGCCGCUCUGCA			
			CUUCGCCAUCCAG			
			CAGGUGCGGGAGG			
			AGAACAUACGGAA			
			ACAACUGAAGAAA			
			UUAGACUGCGAGA			
			GAGAGGAGUUACU			
			UGUGUUCCUGUUC			
			UUCUUCAGUCUGA			
			UGGGCCUGCUGUC			
			CUCCCACCUGACC			
			AGCAAUUCAACCA			
			CCGACCUGCCCAA			
			GGCCUUCCACGUU			
			UGUGCCGCCAUCC			
			UGGAGUGCCUGGA			
			GAAGCGGAAGAUU			
			AGCUGGUUAGCCC			
			UGUUUCAGCUGAC			
			CGAGAGUGAUUUA			
			CGCCUGGGGAGAC			
			UGCUGUUACGCGU			
			GGCUCCCGACCAG			
			CACACCAGGCUGC			
			UGCCUUUCGCCUU			
			CUACAGUCUGUUA			
			UCCUACUUCCACG			
			AGGACGCCGCCAU			

[1240]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CAGAGAGGAGGCC UUCCUGCACGUGG CCGUGGACAUGUA CUUAAAGCUGGUG CAGCUGUUUGUGG CCGGCGAUACCUC CACAGUAAGCCCU CCUGCUGGCAGAA GCCUGGAGCUGAA GGGCCAGGGUAAC CCCGUAGAGCUGA UUACUAAGGCCCG CCUGUUUCUGCUG CAACUGAUCCCUC GCUGCCCCAAGAA GAGCUUUAGCCAC GUGGCCGAAUUAC UGGCUGACAGGGG AGACUGUGAUCCU GAGGUGAGCGCUG CCUUACAGAGCCG UCAACAAGCGGCC CCUGACGCCGAUU UAAGCCAAGAGCC UCACCUCUUC			
SEQ ID NO:		1	2	50	117	27
	FANCA_07 化学：G5 帽：C1 多聚 A 尾： 100nt	MSDSWVPNSASG QDPGGRRRAWAE LLAGRVKREKYNP ERAQKLKESAVRL LRSHQDLNALLLE VEGPLCKKLSLSK VIDCDSSEAYANH	AUGUCCGACUCCU GGGUGCCCCAACUC CGCUAGCGGACAA GAUCCAGGUGGCC GCAGGAGAGCGUG GGCUGAACUGCUU GCCGGGCGAGUGA	GGAAAUCG CAAAAUUU GCUCUUCG CGUUAGAU UUCUUUUA GUUUUCUC GCAACUAG	UGAUAAUAGGC UGGAGCCUCCA CCGCGUUAUCC GUUCCUCCGUAG GCUGGUCCUGG GGAACGGGUCG GCGGGUACCCC	SEQ ID NO:27 从 5'至 3'端由以下组成：SEQ ID NO:50 的 5' UTR、SEQ ID NO:2 的 ORF 序列和 SEQ

[1241]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		SSSFIGSALQDQAS	AGCGCCGAGAAGUA	CAAGCUUU	CGUGGUCUUUG	ID NO:117 的 3'
		RLGVPVGILSAGM	CAACCCCCGAGCGC	UUGUUCUC	AAUAAAGUCUG	UTR
		VASSVGQICTAPAE	GCCCAGAAGCUGA	GCC	AGUGGGCGGC	
		TSHPVLLTVEQRK	AGGAGAGCGCCGU			
		KLSSLLEFAQYLL	ACGGCUGCUGCGC			
		AHSMFSRLSFCQE	AGCCACCAGGACC			
		LWKIQSSLLLEAV	UGAACGCCCUGCU			
		WHLHVQGIVSLQE	GCUGGAGGUGGAG			
		LLESHPDMHAVGS	GGGCCCCUGUGCA			
		WLFRNLCCLCEQ	AGAAGCUGAGCCU			
		MEASCQHADVAR	GUCCAAGGUGAUC			
		AMLSDFVQMFVL	GACUGCGACAGCA			
		RGFQKNSDLRRTV	GCGAGGCCUACGC			
		EPEKMPQVTVDVL	CAACCACAGCUCC			
		QRMLIFALDALAA	UCCUUCAUCGGCU			
		GVQEESSTHKIVR	CCGCCCUGCAGGA			
		CWFGVFSGHTLGS	UCAAGCUAGCCGG			
		VISTDPLKRFFSHT	CUGGGCGUUCCAG			
		LTQILTHSPVLKAS	UGGGCAUCCUGAG			
		DAVQMQREWSFA	CGCCGGCAUGGUG			
		RTHPLLTSLYRRLF	GCCAGCAGCGUGG			
		VMLSAEELVGHLQ	GCCAGAUCUGCAC			
		EVLETQEVHWQR	CGCUCCCGCCGAG			
		VLSFVSALVVCFPE	ACCAGCCACCCCG			
		AQQLLEDWVARL	UGCUGCUGACCGU			
		MAQAFESCQLDS	GGAGCAGCGCAAG			
		MVTAFLVVRQAAL	AAGCUGUCAAGCC			
		EGPSAFLSYADWF	UGCUGGAGUUCGC			
		KASFGSTRGYHGC	CCAGUACCUGCUG			
		SKKALVFLFTFLSE	GCCCACAGCAUGU			
		LVPFESPRYLQVHI	UCUCCCGCCUGAG			
		LHPPLVPGKYRSL	CUUCUGCCAGGAG			
		LTDYISLAKTRLA	CUGUGGAAGAUCC			
		DLKVSIENMGLYE	AGUCCAGCCUGCU			

[1242]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		DLSSAGDITEPHSQ	ACUGGAGGCCGUG			
		ALQDVEKAIMVFE	UGGCACUUACACG			
		HTGNIPVTVMEASI	UGCAGGGCAUCGU			
		FRRPYYVSHFLPA	GAGCCUGCAGGAG			
		LLTPRVLPKVPDSR	CUACUGGAGAGCC			
		VAFIESLKRADKIP	ACCCCGACAUGCA			
		PSLYSTYCQACSA	CGCCGUGGGGUCC			
		AEEKPEDAALGVR	UGGCUGUUCCGCA			
		AEPNSAEEPLGQL	ACCUGUGCUGCCU			
		TAALGELRASMTD	GUGCGAGCAGAUG			
		PSQRDVISAQVAVI	GAAGCCUCCUGCC			
		SERLRAVLGHNED	AACACGCCGACGU			
		DSSVEISKIQLSINT	UGCACGGGCAAUG			
		PRLEPREHMAVDL	CUGAGCGACUUCG			
		LLTSFCQNLMAAS	UGCAGAUGUUCGU			
		SVAPPERQGPWAA	GCUGCGCGGGUUC			
		LFVRTMCGRVLPA	CAGAAGAACUCAG			
		VLTRLCQLLRHQG	ACCUCCGGAGGAC			
		PSLSAPHVLGLAA	CGUGGAGCCUGAG			
		LAVHLGESRSALP	AAGAUGCCCCAGG			
		EVDVGPPAPGAGL	UGACCGUGGGACGU			
		PVPALFDSLLTCRT	GCUGCAGCGGAUG			
		RDSLFFCLKFCTA	CUGAUCUUUGCCC			
		AISYSLCKFSSQSR	UGGACGCUCUCGC			
		DTLCSCLSPGLIKK	UGCCGGAGUGCAG			
		FQFLMFRLFSEAR	GAGGAGAGCUCCA			
		QPLSEEDVASLSW	CCCACAAGAUCGU			
		RPLHLPSADWQRA	GCGGUGCUGGUUC			
		ALSLWTHRTFREV	GGCGUGUUCUCCG			
		LKEEDVHLTYQD	GCCACACCCUGGG			
		WLHLELEIQPEAD	CUCCGUGAUCUCC			
		ALSDTERQDFHQ	ACCGAUCCCCUGA			
		WAIHEHFLPESSAS	AGCGGUUCUUCUC			
		GGCDGDLQAACTI	CCACACCCUGACC			

[1243]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		LVNALMDFHQSSR	CAGAUCCUGACCC			
		SYDHSENSDLVFG	ACUCCCCCUGUGCU			
		GRTGNEDIISRLQE	GAAGGCCAGCGAC			
		MVADLELQQDLIV	GCCGUGCAGAUGC			
		PLGHTPSQEHFLFE	AGCGCGAGUGGUC			
		IFRRRLQALTSGW	CUUCGCCCCGCACC			
		SVAASLQRQRELL	CAUCCCCCUGCUGA			
		MYKRILLRLPSSVL	CCAGCCUGUACCG			
		CGSSFQAEQPITAR	GCGGCUGUUCGUG			
		CEQFFHLVNSEMR	AUGCUGUCCGCCG			
		NFCSHGGALTQDI	AGGAGCUGGUGGG			
		TAHFFRGLLNACL	CCACCUGCAGGAG			
		RSRDPSLMVDFIL	GUGCUGGAGACCC			
		AKCQTKCPLILTSA	AGGAGGUGCACUG			
		LVWWPSLEPVLLC	GCAGCGCGUGCUG			
		RWRRHCQSPLPRE	AGCUUCGUGAGCG			
		LQKLQEGRQFASD	CCCUGGUGGUGUG			
		FLSPEAASPAPNPD	CUUCCCCGAGGCC			
		WLSAAALHFAIQQ	CAGCAGCUCCUGG			
		VREENIRKQLKKL	AGGACUGGGUGGC			
		DCEREELLVFLFFF	CCGGCUGAUGGCC			
		SLMGLLSSHLTSNS	CAGGCCUUCGAGA			
		TTDLPKAFHVCAA	GCUGCCAGCUGGA			
		ILECLEKRKISWLA	CAGCAUGGUGACC			
		LFQLTESDLRLGRL	GCCUUCCUGGUGG			
		LLRVAPDQHTRLL	UGCGUCAAGCAGC			
		PFAFYSLLSYFHED	CCUGGAGGGCCCA			
		AAIREEAFLHVAV	AGCGCCUUCCUGA			
		DMYLKLVQLFVAG	GCUACGCCGACUG			
		DTSTVSPPAGRSLE	GUUCAAGGCCAGC			
		LKGQGNPVELITK	UUCGGGUCCACCC			
		ARLFLLQLIPRCPK	GCGGGUACCACGG			
		KSFSHVAELLADR	CUGCUCCAAGAAG			
		GDCDPEVSAALQS	GCCCUGGUGUUCC			

[1244]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		RQQAAPDADLSQE PHLF	UGUUUACCUUCCU GAGCGAGCUGGUG CCCUUCGAGUCGC CCCGCUACCUGCA GGUGCACAUCCUG CACCCACCCCUGG UCCCUGGCAAGUA CCGGAGCCUGCUG ACCGACUACAUCU CCCUGGCCAAGAC CCGGCUGGCCGAC CUGAAGGUGAGCA UCGAGAACAUGGG CCUGUACGAGGAC CUGUCCAGCGCCG GCGACAUCACCGA GCCCCACAGCCAG GCCCUGCAGGACG UGGAGAAGGCCAU CAUGGUGUUCGAG CACACCGGCAACA UCCCCGUGACCGU GAUGGAGGCUAGU AUCUUCCGGCGGC CCUACUACGUGUC CCACUUCCUGCCC GCCCUGCUGACAC CCCGGGUGCUGCC CAAGGUGCCCGAC AGUCGCGUCGCCU UCAUCGAGAGCCU GAAGCGGGCCGAC AAGAUCCCGCCCU CCCUGUACUCCAC			

[1245]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CUACUGCCAGGCU			
			UGCUCCCGCCGCCG			
			AGGAGAAGCCCGA			
			GGACGCAGCCUUA			
			GGAGUACGCGCCG			
			AGCCCAACAGCGC			
			CGAGGAGCCACUG			
			GGCCAGUUGACCG			
			CUGCCCUGGGAGA			
			GCUGCGGGCCAGC			
			AUGACCGACCCCA			
			GCCAACGCGACGU			
			GAUCAGCGCCCAG			
			GUGGCCGUGAUCA			
			GCGAGAGGCUGCG			
			GGCAGUGCUGGGC			
			CACAACGAGGACG			
			ACAGCAGCGUGGA			
			GAUCAGCAAGAUC			
			CAGCUGAGCAUCA			
			ACACACCCAGGCU			
			GGAGCCCAGAGAG			
			CACAUGGCCGUGG			
			ACCUGCUGCUGAC			
			CUCCUUCUGCCAG			
			AACCUGAUGGCCG			
			CAAGCAGCGUGGC			
			UCCUCCAGAACGC			
			CAAGGCCCUUGGG			
			CUGCCCUGUUCGU			
			GCGCACCAUGUGU			
			GGCCGCGUCCUGC			
			CAGCAGUGCUGAC			
			CCGGCUGUGCCAG			

[1246]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CUUCUGCGGCACC			
			AAGGGCCCUCCCU			
			GAGCGCUCCCCAC			
			GUGCUUGGGUUGG			
			CCGCCCUCGGCUGU			
			GCACCUGGGCGAG			
			AGCAGGAGCGCCC			
			UUCCCGAGGUGGA			
			CGUGGGUCCUCCA			
			GCCCCAGGUGCCG			
			GACUCCCCGUGCC			
			AGCCCUGUUCGAC			
			AGCCUGCUGACUU			
			GUCGCACCCGGGA			
			CUCCCUGUUCUUC			
			UGCCUGAAGUUCU			
			GCACCGCCGCCAU			
			CUCCUACUCCCUG			
			UGCAAGUUCUCCA			
			GCCAGUCCCGGGA			
			CACCCUGUGCUCC			
			UGCCUGUCUCCCG			
			GCCUGAUCAAGAA			
			GUUCCAGUUCCUG			
			AUGUUCCGCCUGU			
			UCAGCGAAGCACG			
			CCAGCCCCUGUCC			
			GAGGAGGACGUGG			
			CCUCCCUGUCCUG			
			GCGACCCCUGCAC			
			CUGCCCUCCGCCG			
			AUUGGCAGCGCGC			
			CGCCCUUAGCCUG			
			UGGACCCACCGGA			

[1247]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCUUCCGCGAGGU			
			GCUGAAGGAGGAA			
			GACGUGCACCUGA			
			CCUACCAGGACUG			
			GCUGCACCUGGAG			
			CUGGAGAUCCAGC			
			CCGAAGCCGACGC			
			CCUGAGCGACACC			
			GAGCGCCAGGACU			
			UUCACCAGUGGGC			
			CAUCCACGAGCAC			
			UUUCUGCCCGAGA			
			GCAGCGCCAGUGG			
			AGGCUGCGACGGC			
			GACCUGCAAGCCG			
			CCUGCACCAUCCU			
			GGUGAACGCCCUG			
			AUGGACUUCCACC			
			AGAGCUCCCGCAG			
			CUACGACCACUCC			
			GAGAACAGCGACC			
			UGGUGUUUGGAGG			
			CCGCACCGGGAAC			
			GAGGACAUCAUCA			
			GCCGCCUGCAGGA			
			GAUGGUGGCCGAC			
			CUGGAGCUCCAGC			
			AGGACCUGAUCGU			
			GCCACUGGGGCAC			
			ACCCCAAGCCAGG			
			AGCACUUCCUGUU			
			CGAGAUCUUCCGC			
			AGACGCCUGCAGG			
			CACUGACCUCCGG			

[1248]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CUGGAGUGUGGCC			
			GCAAGCCUGCAGC			
			GCCAACGCGAGCU			
			GCUGAUGUACAAG			
			CGGAUCCUGCUGC			
			GCCUGCCCUCCUC			
			CGUGCUGUGCGGC			
			UCCUCCUUCCAGG			
			CCGAGCAGCCCAU			
			CACCGCCCGGUGC			
			GAGCAGUUCUUCC			
			ACCUGGUGAACAG			
			CGAGAUGCGGAAC			
			UUCUGCUCUCACG			
			GCGGCGCUCUGAC			
			CCAGGACAUCACC			
			GCCCACUUCUUCC			
			GCGGCCUGCUGAA			
			CGCCUGCCUGCGG			
			AGCAGGGACCCCA			
			GCCUGAUGGUCGA			
			CUUCAUCCUGGCC			
			AAGUGCCAGACCA			
			AGUGUCCCCUGAU			
			CCUGACCAGCGCC			
			CUGGUUUGGUGGC			
			CCUCACUGGAGCC			
			CGUGCUCCUGUGU			
			CGCUGGAGGCGGC			
			ACUGCCAGUCACC			
			CCUGCCCCGAGAG			
			CUGCAGAAGCUGC			
			AGGAGGGCCGGCA			
			GUUCGCCAGCGAC			

[1249]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			UUCCUGUCACCAG			
			AGGCAGCCUCUCC			
			CGCUCCCUAACCCCC			
			GAUUGGCUAUCCG			
			CGGCCGCUCUGCA			
			CUUCGCCAUCCAG			
			CAGGUGCGGGAGG			
			AGAACAUCCGGAA			
			GCAGCUGAAGAAG			
			CUGGACUGCGAGA			
			GGGAGGAGCUGCU			
			UGUGUUCCUGUUC			
			UUCUUCAGCCUGA			
			UGGGCCUGCUGUC			
			CUCCCACCUGACC			
			AGCAACUCCCACCA			
			CCGACCUGCCCCAA			
			GGCCUUCCACGUG			
			UGCGCCGCCAUCC			
			UGGAGUGCCUGGA			
			GAAGCGGAAGAUC			
			AGCUGGCUGGCCC			
			UGUUCCAGCUGAC			
			CGAGAGCGACCUG			
			CGCCUGGGGAGAC			
			UGCUGCUGCGCGU			
			GGCUCCCGACCAG			
			CACACCAGGCUGC			
			UGCCCUUCGCCUU			
			CUACUCCCCUGCUG			
			UCCUACUUCCCACG			
			AGGACGCCGCCAU			
			CCGGGAGGAGGCC			
			UUCCUGCACGUGG			

[1250]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCGUGGACAUGUA CCUGAAGCUGGGUG CAGCUGUUCGUGG CCGGCGACACCUC CACCGUGAGCCCU CCUGCUGGCCGAA GCCUGGAGCUGAA GGGCCAGGGCAAC CCCGUGGAGCUGA UCACCAAGGCCCG CCUGUUCCUGCUG CAGCUGAUCCCUC GCUGCCCCAAGAA GAGCUUCAGCCAC GUGGCCGAGCUGC UGGCUGACAGGGG AGACUGCGACCCC GAGGUGAGCGCUG CCCUGCAGAGCCG UCAACAAGCGGCC CCUGACGCCGAUC UGAGCCAGGAGCC CCACCUGUUC			
SEQ ID NO:		1	2	80	117	28
	FANCA_08 化学：G5 帽：C1 多聚 A 尾： 100nt	MSDSWVPNSASG QDPGGRRRAWAE LLAGRVKREKYNP ERAQKLKESAVRL LRSHQDLNALLLE VEGPLCKKLSLSK VIDCDSSEAYANH SSSFIGSALQDQAS RLGVPVGILSAGM	AUGUCCGACUCCU GGGUGGCCCAACUC CGCUAGCGGACAA GAUCCAGGUGGCC GCAGGAGAGCGUG GGCUGAACUGCUU GCCGGGCGAGUGA AGCGCGAGAAGUA CAACCCCGAGCGC	GGACUCAC UAUUUGUU UUCGCGCC CAGUUGCA AAAA	UGAUAAUAGGC UGGAGCCUCCA CCGCGUUAUCC GUUCCUCGUAG GCUGGUCCUGG GGAACGGGUCG GCGGGUACCCC CGUGGUCUUUG AAUAAAGUCUG	SEQ ID NO:28 从 5'至 3'端由以下组成：SEQ ID NO:80 的 5' UTR 、SEQ ID NO:2 的 ORF 序列和 SEQ ID NO:117 的 3' UTR

[1251]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		VASSVGQICTAPAE	GCCCAGAAGCUGA		AGUGGGCGGC	
		TSHPVLLTVEQRK	AGGAGAGCGCCGU			
		KLSSLLEFAQYLL	ACGGCUGCUGCGC			
		AHSMFSRLSFCQE	AGCCACCAGGACC			
		LWKIQSSLLLEAV	UGAACGCCCUGCU			
		WHLHVQGIVSLQE	GCUGGAGGUGGAG			
		LLESHPDMHAVGS	GGGCCCCUGUGCA			
		WLFRNLCCLCEQ	AGAAGCUGAGCCU			
		MEASCQHADVAR	GUCCAAGGUGAUC			
		AMLSDFVQMFVL	GACUGCGACAGCA			
		RGFQKNSDLRRTV	GCGAGGCCUACGC			
		EPEKMPQVTVDVL	CAACCACAGCUCC			
		QRMLIFALDALAA	UCCUUCAUCGGCU			
		GVQEESSTHKIVR	CCGCCCUGCAGGA			
		CWFGVFSGHTLGS	UCAAGCUAGCCGG			
		VISTDPLKRFFSHT	CUGGGCGUUCCAG			
		LTQILTHSPVLKAS	UGGGCAUCCUGAG			
		DAVQMQREWSFA	CGCCGGCAUGGUG			
		RTHPLLTSLYRRLF	GCCAGCAGCGUGG			
		VMLSAEELVGHLQ	GCCAGAUCUGCAC			
		EVLETQEVHWQR	CGCUCCCGCCGAG			
		VLSFVSALVVCFPE	ACCAGCCACCCCG			
		AQQLLEDWVARL	UGCUGCUGACCGU			
		MAQAFESCQLDS	GGAGCAGCGCAAG			
		MVTAFLVVRQAAL	AAGCUGUCAAGCC			
		EGPSAFLSYADWF	UGCUGGAGUUCGC			
		KASFGSTRGYHGC	CCAGUACCUGCUG			
		SKKALVFLFTFLSE	GCCCACAGCAUGU			
		LVPFESPRYLQVHI	UCUCCCGCCUGAG			
		LHPPLVPGKYRSL	CUUCUGCCAGGAG			
		LTDYISLAKTRLA	CUGUGGAAGAUCC			
		DLKVSIENMGLYE	AGUCCAGCCUGCU			
		DLSSAGDITEPHSQ	ACUGGAGGCCGUG			
		ALQDVEKAIMVFE	UGGCACUUACACG			

[1252]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		HTGNIPVTVMEASI	UGCAGGGCAUCGU			
		FRRPYYVSHFLPA	GAGCCUGCAGGAG			
		LLTPRVLPKVPDSR	CUACUGGAGAGCC			
		VAFIESLKRADKIP	ACCCCGACAUGCA			
		PSLYSTYCQACSA	CGCCGUGGGGUCC			
		AEEKPEDAALGVR	UGGCUGUUCCGCA			
		AEPNSAEEPLGQL	ACCUGUGCUGCCU			
		TAALGELRASMTD	GUGCGAGCAGAUG			
		PSQRDVISAQVAVI	GAAGCCUCCUGCC			
		SERLRAVLGHNED	AACACGCCGACGU			
		DSSVEISKIQLSINT	UGCACGGGCAAUG			
		PRLEPREHMAVDL	CUGAGCGACUUCG			
		LLTSFCQNLMAAS	UGCAGAUGUUCGU			
		SVAPPERQGPWAA	GCUGCGCGGGUUC			
		LFVRTMCGRVLPA	CAGAAGAACUCAG			
		VLTRLCQLLRHQG	ACCUCCGGAGGAC			
		PSLSAPHVLGLAA	CGUGGAGCCUGAG			
		LAVHLGESRSALP	AAGAUGCCCCAGG			
		EVDVGPPAPGAGL	UGACCGUGGACGU			
		PVPALFDSLLTCRT	GCUGCAGCGGAUG			
		RDSLFFCLKFCTA	CUGAUCUUUGCCC			
		AISYSLCKFSSQSR	UGGACGCUCUCGC			
		DTLCSCLSPGLIKK	UGCCGGAGUGCAG			
		FQFLMFRLFSEAR	GAGGAGAGCUCCA			
		QPLSEEDVASLSW	CCCACAAGAUCGU			
		RPLHLPSADWQRA	GCGGUGCUGGUUC			
		ALSLWTHRTFREV	GGCGUGUUCUCCG			
		LKEEDVHLTYQD	GCCACACCCUGGG			
		WLHLELEIQPEAD	CUCCGUGAUCUCC			
		ALSDTERQDFHQ	ACCGAUCCCCCUGA			
		WAIHEHFLPESSAS	AGCGGUUCUUCUC			
		GGCDGDLQAACTI	CCACACCCUGACC			
		LVNALMDFHQSSR	CAGAUCCUGACCC			
		SYDHSENSDLVFG	ACUCCCCUGUGCU			

[1253]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		GRTGNEDIISRLQE	GAAGGCCAGCGGAC			
		MVADLELQQDLIV	GCCGUGCAGAUGC			
		PLGHTPSQEHFLFE	AGCGCGAGUGGUC			
		IFRRRLQALTSGW	CUUCGCCCGCACC			
		SVAASLQRQRELL	CAUCCCCUGCUGA			
		MYKRILLRLPSSVL	CCAGCCUGUACCG			
		CGSSFQAEQPITAR	GCGGCUGUUCGUG			
		CEQFFHLVNSEMR	AUGCUGUCCGCCG			
		NFCSHGGALTQDI	AGGAGCUGGUGGG			
		TAHFFRGLLNACL	CCACCUGCAGGAG			
		RSRDPSLMVDFIL	GUGCUGGAGACCC			
		AKCQTKCPLILTSA	AGGAGGUGCACUG			
		LVWWPSLEPVLLC	GCAGCGCGUGCUG			
		RWRRHCQSPLPRE	AGCUUCGUGAGCG			
		LQKLQEGRQFASD	CCCUGGUGGUGUG			
		FLSPEAASPAPNPD	CUUCCCCGAGGCC			
		WLSAAALHFAIQQ	CAGCAGCUCCUGG			
		VREENIRKQLKKL	AGGACUGGGUGGC			
		DCEREELLVFLFFF	CCGGCUGAUGGCC			
		SLMGLLSSHLTSNS	CAGGCCUUCGAGA			
		TTDLPKAFHVCAA	GCUGCCAGCUGGA			
		ILECLEKRKISWLA	CAGCAUGGUGACC			
		LFQLTESDLRLGRL	GCCUUCCUGGUGG			
		LLRVAPDQHTRLL	UGCGUCAAGCAGC			
		PFAFYSLLSYFHED	CCUGGAGGGCCCA			
		AAIREEAFLHVAV	AGCGCCUUCCUGA			
		DMYLKLVQLFVAG	GCUACGCCGACUG			
		DTSTVSPPAGRSLE	GUUCAAGGCCAGC			
		LKGQGNPVELITK	UUCGGGUCCACCC			
		ARLFLLQLIPRCPK	GCGGGUACCACGG			
		KSFSHVAELLADR	CUGCUCCAAGAAG			
		GDCDPEVSAALQS	GCCCUGGUGUUCC			
		RQQAAPDADLSQE	UGUUUACCUUCCU			
		PHLF	GAGCGAGCUGGUG			

[1254]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCCUUCGAGUCGC			
			CCCGCUACCUGCA			
			GGUGCACAUCCUG			
			CACCCACCCCUGG			
			UCCCUGGCAAGUA			
			CCGGAGCCUGCUG			
			ACCGACUACAUCU			
			CCCUGGCCAAGAC			
			CCGGCUGGCCGAC			
			CUGAAGGUGAGCA			
			UCGAGAACAUGGG			
			CCUGUACGAGGAC			
			CUGUCCAGCGCCG			
			GCGACAUCACCGA			
			GCCCCACAGCCAG			
			GCCCUGCAGGACG			
			UGGAGAAGGCCAU			
			CAUGGUGUUCGAG			
			CACACCGGCAACA			
			UCCCCGUGACCGU			
			GAUGGAGGCUAGU			
			ACUUCCGGCGGC			
			CCUACUACGUGUC			
			CCACUUCCUGCCC			
			GCCCUGCUGACAC			
			CCCGGGUGCUGCC			
			CAAGGUGCCCCGAC			
			AGUCGCGUCGCCU			
			UCAUCGAGAGCCU			
			GAAGCGGGCCGAC			
			AAGAUCCCGCCCU			
			CCCUGUACUCCAC			
			CUACUGCCAGGCU			
			UGCUCCCGCCGCCG			

[1255]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			AGGAGAAGCCCGA			
			GGACGCAGCCUUA			
			GGAGUACGCGCCG			
			AGCCCAACAGCGC			
			CGAGGAGCCACUG			
			GGCCAGUUGACCG			
			CUGCCCUGGGAGA			
			GCUGCGGGCCAGC			
			ALGACCGACCCCA			
			GCCAACGCGACGU			
			GAUCAGCGCCCAG			
			GUGGCCGUGAUCA			
			GCGAGAGGCUGCG			
			GGCAGUGCUGGGC			
			CACAACGAGGACG			
			ACAGCAGCGUGGA			
			GAUCAGCAAGAUC			
			CAGCUGAGCAUCA			
			ACACACCCAGGCU			
			GGAGCCCAGAGAG			
			CACAUGGCCGUGG			
			ACCUGCUGCUGAC			
			CUCCUUCUGCCAG			
			AACCUGAUGGCCG			
			CAAGCAGCGUGGC			
			UCCUCCAGAACGC			
			CAAGGCCCUUGGG			
			CUGCCCUGUUCGU			
			GCGCACCAUGUGU			
			GGCCGCGUCCUGC			
			CAGCAGUGCUGAC			
			CCGGCUGUGCCAG			
			CUUCUGCGGCACC			
			AAGGGCCCUCCCU			

[1256]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GAGCGCUCCCCCAC			
			GUGCUUGGGUUGG			
			CCGCCCUCGCUGU			
			GCACCUGGGCGAG			
			AGCAGGAGCGCCC			
			UUCCCGAGGUGGA			
			CGUGGGUCCUCCA			
			GCCCCAGGUGCCG			
			GACUCCCCGUGCC			
			AGCCCUGUUCGAC			
			AGCCUGCUGACUU			
			GUCGCACCCGGGA			
			CUCCCUGUUCUUC			
			UGCCUGAAGUUCU			
			GCACCGCCGCCAU			
			CUCCUACUCCCUG			
			UGCAAGUUCUCCA			
			GCCAGUCCCGGGA			
			CACCCUGUGCUCC			
			UGCCUGUCUCCCG			
			GCCUGAUCAAGAA			
			GUUCCAGUUCCUG			
			ALGUUCCGCCUGU			
			UCAGCGAAGCACG			
			CCAGCCCCUGUCC			
			GAGGAGGACGUGG			
			CCUCCCUGUCCUG			
			GCGACCCCUGCAC			
			CUGCCCUCCGCCG			
			AUUGGCAGCGCGC			
			CGCCCUUAGCCUG			
			UGGACCCACCGGA			
			CCUUCCGCGAGGU			
			GCUGAAGGAGGAA			

[1257]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GACGUGCACCUGA			
			CCUACCAGGACUG			
			GCUGCACCUGGAG			
			CUGGAGAUCCAGC			
			CCGAAGCCGACGC			
			CCUGAGCGACACC			
			GAGCGCCAGGACU			
			UUCACCAGUGGGC			
			CAUCCACGAGCAC			
			UUUCUGCCCGAGA			
			GCAGCGCCAGUGG			
			AGGCUGCGACGGC			
			GACCUGCAAGCCG			
			CCUGCACCAUCCU			
			GGUGAACGCCCUG			
			AUGGACUUCCACC			
			AGAGCUCCCGCAG			
			CUACGACCACUCC			
			GAGAACAGCGACC			
			UGGUGUUUGGAGG			
			CCGCACCGGGAAC			
			GAGGACAUCAUCA			
			GCCGCCUGCAGGA			
			GAUGGUGGCCGAC			
			CUGGAGCUCCAGC			
			AGGACCUGAUCGU			
			GCCACUGGGGCAC			
			ACCCCAAGCCAGG			
			AGCACUUCCUGUU			
			CGAGAUCUUCCGC			
			AGACGCCUGCAGG			
			CACUGACCUCCGG			
			CUGGAGUGUGGCC			
			GCAAGCCUGCAGC			

[1258]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GCCAACGCGAGCU			
			GCUGAUGUACAAG			
			CGGAUCCUGCUGC			
			GCCUGCCCUCCUC			
			CGUGCUGUGCGGC			
			UCCUCCUUCCAGG			
			CCGAGCAGCCCAU			
			CACCGCCCGGUGC			
			GAGCAGUUCUUCC			
			ACCUGGUGAACAG			
			CGAGAUGCGGAAC			
			UUCUGCUCUCACG			
			GCGGCGCUCUGAC			
			CCAGGACAUCACC			
			GCCCACUUCUUCC			
			GCGGCCUGCUGAA			
			CGCCUGCCUGCGG			
			AGCAGGGACCCCA			
			GCCUGAUGGUCGA			
			CUUCAUCCUGGCC			
			AAGUGCCAGACCA			
			AGUGUCCCCUGAU			
			CCUGACCAGCGCC			
			CUGGUUUGGUGGC			
			CCUCACUGGAGCC			
			CGUGCUCCUGUGU			
			CGCUGGAGGCGGC			
			ACUGCCAGUCACC			
			CCUGCCCCGAGAG			
			CUGCAGAAGCUGC			
			AGGAGGGCCGGCA			
			GUUCGCCAGCGAC			
			UUCCUGUCACCAG			
			AGGCAGCCUCUCC			

[1259]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CGCUCCUAACCCC			
			GAUUGGCUAUCCG			
			CGGCCGCUCUGCA			
			CUUCGCCAUCCAG			
			CAGGUGCGGGAGG			
			AGAACAUCCGGAA			
			GCAGCUGAAGAAG			
			CUGGACUGCGAGA			
			GGGAGGAGCUGCU			
			UGUGUUCCUGUUC			
			UUCUUCAGCCUGA			
			UGGGCCUGCUGUC			
			CUCCCACCUGACC			
			AGCAACUCCCACCA			
			CCGACCUGCCCCAA			
			GGCCUUCCACGUG			
			UGCGCCGCCAUCC			
			UGGAGUGCCUGGA			
			GAAGCGGAAGAUC			
			AGCUGGCUGGCCC			
			UGUUCCAGCUGAC			
			CGAGAGCGACCUG			
			CGCCUGGGGAGAC			
			UGCUGCUGCGCGU			
			GGCUCCCGACCAG			
			CACACCAGGCUGC			
			UGCCCUUCGCCUU			
			CUACUCCCUGCUG			
			UCCUACUUCCACG			
			AGGACGCCGCCAU			
			CCGGGAGGAGGCC			
			UUCCUGCACGUGG			
			CCGUGGACAUGUA			
			CCUGAAGCUGGUG			

[1260]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
SEQ ID NO:		1	2	80	140	29
	FANCA_09 化学：G5 帽：C1 多聚 A 尾：100nt	MSDSWVPNSASG QDPGGRRRAWAE LLAGRVKREKYNP ERAQKLKESAVRL LRSHQDLNALLLE VEGPLCKKLSLSK VIDCDSSEAYANH SSSFIGSALQDQAS RLGVPVGILSAGM VASSVGQICTAPAE TSHPVLLTVEQRK	ALGUCCGACUCCU GGGUGCCCAACUC CGCUAGCGGACAA GAUCCAGGUGGCC GCAGGAGAGCGUG GGCUGAACUGCUU GCCGGGCGAGUGA AGCGCGAGAAGUA CAACCCCGAGCGC GCCCAGAAGCUGA AGGAGAGCGCCGU	GGACUCAC UAUUUGUU UUCGCGCC CAGUUGCA AAAA	UAAGCCCCUCC GGGGGCCUCCA CCGCGUUAUCC GUUCCUCGUAG GCUGGUCCUGG GGAACGGGUCG GCGGGUACCCC CGUGGUCUUUG AAUAAAGUCUG AGUGGGCGGC	SEQ ID NO:29 从 5'至 3'端由以下组成：SEQ ID NO:80 的 5' UTR、SEQ ID NO:2 的 ORF 序列和 SEQ ID NO:140 的 3' UTR

[1261]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		KLSSLLEFAQYLL	ACGGCUGCUGCGC			
		AHSMFSRLSFCQE	AGCCACCAGGACC			
		LWKIQSSLLLEAV	UGAACGCCCUGCU			
		WHLHVQGIVSLQE	GCUGGAGGUGGAG			
		LLESHPDMHAVGS	GGGCCCCUGUGCA			
		WLFRNLCCLCEQ	AGAAGCUGAGCCU			
		MEASCQHADVAR	GUCCAAGGUGAUC			
		AMLSDFVQMFVL	GACUGCGACAGCA			
		RGFQKNSDLRRTV	GCGAGGCCUACGC			
		EPEKMPQVTVDVL	CAACCACAGCUCC			
		QRMLIFALDALAA	UCCUUCAUCGGCU			
		GVQEESSTHKIVR	CCGCCCUGCAGGA			
		CWFGVFSGHTLGS	UCAAGCUAGCCGG			
		VISTDPLKRFFSHT	CUGGGCGUUCCAG			
		LTQILTHSPVLKAS	UGGGCAUCCUGAG			
		DAVQMQREWSFA	CGCCGGCAUGGUG			
		RTHPLLTSLYRRLF	GCCAGCAGCGUGG			
		VMLSAEELVGHLQ	GCCAGAUCUGCAC			
		EVLETQEVHWQR	CGCUCCCGCCGAG			
		VLSFVSALVVCFPE	ACCAGCCACCCCG			
		AQQLLEDWVARL	UGCUGCUGACCGU			
		MAQAFESCQLDS	GGAGCAGCGCAAG			
		MVTAFLVVRQAAL	AAGCUGUCAAGCC			
		EGPSAFLSYADWF	UGCUGGAGUUCGC			
		KASFGSTRGYHGC	CCAGUACCUGCUG			
		SKKALVFLFTFLSE	GCCCACAGCAUGU			
		LVPFESPRYLQVHI	UCUCCCGCCUGAG			
		LHPPLVPGKYRSL	CUUCUGCCAGGAG			
		LTDYISLAKTRLA	CUGUGGAAGAUCC			
		DLKVSIENMGLYE	AGUCCAGCCUGCU			
		DLSSAGDITEPHSQ	ACUGGAGGCCGUG			
		ALQDVEKAIMVFE	UGGCACUUACACG			
		HTGNIPVTVMEASI	UGCAGGGCAUCGU			
		FRRPYYVSHFLPA	GAGCCUGCAGGAG			

[1262]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		LLTPRVLPKVPDSR	CUACUGGAGAGCC			
		VAFIESLKRADKIP	ACCCCGACAUGCA			
		PSLYSTYCQACSA	CGCCGUGGGGUCC			
		AEEKPEDAALGVR	UGGCUGUUCCGCA			
		AEPNSAEEPLGQL	ACCUGUGCUGCCU			
		TAALGELRASMTD	GUGCGAGCAGAUG			
		PSQRDVISAQVAVI	GAAGCCUCCUGCC			
		SERLRAVLGHNED	AACACGCCGACGU			
		DSSVEISKIQLSINT	UGCACGGGCAAUG			
		PRLEPREHMAVDL	CUGAGCGACUUCG			
		LLTSFCQNLMAAS	UGCAGAUGUUCGU			
		SVAPPERQGPWAA	GCUGCGCGGGUUC			
		LFVRTMCGRVLPA	CAGAAGAACUCAG			
		VLTRLCQLLRHQG	ACCUCCGGAGGAC			
		PSLSAPHVLGLAA	CGUGGAGCCUGAG			
		LAVHLGESRSALP	AAGAUGCCCCAGG			
		EVDVGPPAPGAGL	UGACCGUGGACGU			
		PVPALFDSLLTCRT	GCUGCAGCGGAUG			
		RDSLFFCLKFCTA	CUGAUCUUUGCCC			
		AISYSLCKFSSQSR	UGGACGCUCUCGC			
		DTLCSCLSPGLIKK	UGCCGGAGUGCAG			
		FQFLMFRLFSEAR	GAGGAGAGCUCCA			
		QPLSEEDVASLSW	CCCACAAGAUCGU			
		RPLHLPSADWQRA	GCGGUGCUGGUUC			
		ALSLWTHRTFREV	GGCGU/GUUCUCCG			
		LKEEDVHLTYQD	GCCACACCCUGGG			
		WLHLELEIQPEAD	CUCCGUGAUCUCC			
		ALSDTERQDFHQ	ACCGAUCCCCUGA			
		WAIHEHFLPESSAS	AGCGGUUCUUCUC			
		GGCDGDLQAACTI	CCACACCCUGACC			
		LVNALMDFHQSSR	CAGAUCCUGACCC			
		SYDHSENSDLVFG	ACUCCCCUGUGCU			
		GRTGNEDIISRLQE	GAAGGCCAGCGAC			
		MVADLELQQDLIV	GCCGUGCAGAUGC			

[1263]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		PLGHTPSQEHFLFE	AGCGCGAGUGGUC			
		IFRRRLQALTSGW	CUUCGCCCCGCACC			
		SVAASLQRQRELL	CAUCCCCUGCUGA			
		MYKRILLRLPSSVL	CCAGCCUGUACCG			
		CGSSFQAEQPITAR	GCGGCUGUUCGUG			
		CEQFFHLVNSEMR	AUGCUGUCCGCCG			
		NFCSHGGALTQDI	AGGAGCUGGUGGG			
		TAHFFRGLLNACL	CCACCUGCAGGAG			
		RSRDPSLMVDFIL	GUGCUGGAGACCC			
		AKCQTKCPLILTSA	AGGAGGUGCACUG			
		LVWWPSLEPVLLC	GCAGCGCGUGCUG			
		RWRRHCQSPLPRE	AGCUUCGUGAGCG			
		LQKLQEGRQFASD	CCCUGGUGGUGUG			
		FLSPEAASPAPNPD	CUUCCCCGAGGCC			
		WLSAAALHFAIQQ	CAGCAGCUCCUGG			
		VREENIRKQLKKL	AGGACUGGGUGGC			
		DCEREELLVFLFFF	CCGGCUGAUGGCC			
		SLMGLLSSHLTSNS	CAGGCCUUCGAGA			
		TTDLPKAFHVCAA	GCUGCCAGCUGGA			
		ILECLEKRKISWLA	CAGCAUGGUGACC			
		LFQLTESDLRLGRL	GCCUUCCUGGUGG			
		LLRVAPDQHTRLL	UGCGUCAAGCAGC			
		PFAFYSLLSYFHED	CCUGGAGGGCCCA			
		AAIREEAFLHVAV	AGCGCCUUCCUGA			
		DMYLKLVQLFVAG	GCUACGCCGACUG			
		DTSTVSPPAGRSLE	GUUCAAGGCCAGC			
		LKGQGNPVELITK	UUCGGGUCCACCC			
		ARLFLLQLIPRCPK	GCGGGUACCACGG			
		KSFSHVAELLADR	CUGCUCCAAGAAG			
		GDCDPEVSAALQS	GCCCUGGUGUUCC			
		RQQAAPDADLSQE	UGUUUACCUUCCU			
		PHLF	GAGCGAGCUGGUG			
			CCCUUCGAGUCGC			
			CCCGCUACCUGCA			

[1264]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GGUGCACAUCCUG			
			CACCCACCCCUGG			
			UCCCUGGCAAGUA			
			CCGGAGCCUGCUG			
			ACCGACUACAUCU			
			CCCUGGCCAAGAC			
			CCGGCUGGCCGAC			
			CUGAAGGUGAGCA			
			UCGAGAACAUGGG			
			CCUGUACGAGGAC			
			CUGUCCAGCGCCG			
			GCGACAUCACCGA			
			GCCCCACAGCCAG			
			GCCCUGCAGGACG			
			UGGAGAAGGCCAU			
			CAUGGUGUUCGAG			
			CACACCGGCAACA			
			UCCCCGUGACCGU			
			GAUGGAGGCUAGU			
			ACUUCCGGCGGC			
			CCUACUACGUGUC			
			CCACUUCCUGCCC			
			GCCCUGCUGACAC			
			CCCGGGUGCUGCC			
			CAAGGUGCCCGAC			
			AGUCGCGUCGCCU			
			UCAUCGAGAGCCU			
			GAAGCGGGCCGAC			
			AAGAUCCCGCCCU			
			CCCUGUACUCCAC			
			CUACUGCCAGGCU			
			UGCUCCGCCGCCG			
			AGGAGAAGCCCGA			
			GGACGCAGCCUUA			

[1265]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GGAGUACGCGCCG			
			AGCCCAACAGCGC			
			CGAGGAGCCACUG			
			GGCCAGUUGACCG			
			CUGCCCUGGGAGA			
			GCUGCGGGCCAGC			
			ALGACCGACCCCA			
			GCCAACGCGACGU			
			GAUCAGCGCCCAG			
			GUGGCCGUGAUCA			
			GCGAGAGGCUGCG			
			GGCAGUGCUGGGC			
			CACAACGAGGACG			
			ACAGCAGCGUGGA			
			GAUCAGCAAGAUC			
			CAGCUGAGCAUCA			
			ACACACCCAGGCU			
			GGAGCCCAGAGAG			
			CACAUGGCCGUGG			
			ACCUGCUGCUGAC			
			CUCCUUCUGCCAG			
			AACCUGAUGGCCG			
			CAAGCAGCGUGGC			
			UCCUCCAGAACGC			
			CAAGGCCCUUGGG			
			CUGCCCUGUUCGU			
			GCGCACCAUGUGU			
			GGCCGCGUCCUGC			
			CAGCAGUGCUGAC			
			CCGGCUGUGCCAG			
			CUUCUGCGGCACC			
			AAGGGCCCUCCCCU			
			GAGCGCUCCCCCAC			
			GUGCUUGGGUUGG			

[1266]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCGCCCUCGGCUGU			
			GCACCUGGGCGAG			
			AGCAGGAGCGCCC			
			UUCCCGAGGUGGA			
			CGUGGGUCCUCCA			
			GCCCCAGGUGCCG			
			GACUCCCCGUGCC			
			AGCCCUGUUCGAC			
			AGCCUGCUGACUU			
			GUCGCACCCGGGA			
			CUCCCUGUUCUUC			
			UGCCUGAAGUUCU			
			GCACCGCCGCCAU			
			CUCCUACUCCCUG			
			UGCAAGUUCUCCA			
			GCCAGUCCCGGGA			
			CACCCUGUGCUCC			
			UGCCUGUCUCCCG			
			GCCUGAUCAAGAA			
			GUUCCAGUUCCUG			
			ALGUUCCGCCUGU			
			UCAGCGAAGCACG			
			CCAGCCCCUGUCC			
			GAGGAGGACGUGG			
			CCUCCCUGUCCUG			
			GCGACCCCUGCAC			
			CUGCCCUCCGCCG			
			AUUGGCAGCGCGC			
			CGCCCUUAGCCUG			
			UGGACCCACCGGA			
			CCUUCCGCGAGGU			
			GCUGAAGGAGGAA			
			GACGUGCACCUGA			
			CCUACCAGGACUG			

[1267]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GCUGCACCUGGAG			
			CUGGAGAUCCAGC			
			CCGAAGCCGACGC			
			CCUGAGCGACACC			
			GAGCGCCAGGACU			
			UUCACCAGUGGGC			
			CAUCCACGAGCAC			
			UUUCUGCCCGAGA			
			GCAGCGCCAGUGG			
			AGGCUGCGACGGC			
			GACCUGCAAGCCG			
			CCUGCACCAUCCU			
			GGUGAACGCCCUG			
			AUGGACUUCCACC			
			AGAGCUCCCGCAG			
			CUACGACCACUCC			
			GAGAACAGCGACC			
			UGGUGUUUGGAGG			
			CCGCACCGGGAAC			
			GAGGACAUCAUCA			
			GCCGCCUGCAGGA			
			GAUGGUGGCCGAC			
			CUGGAGCUCCAGC			
			AGGACCUGAUCGU			
			GCCACUGGGGCAC			
			ACCCCAAGCCAGG			
			AGCACUUCCUGUU			
			CGAGAUCUUCCGC			
			AGACGCCUGCAGG			
			CACUGACCUCCGG			
			CUGGAGUGUGGCC			
			GCAAGCCUGCAGC			
			GCCAACGCGAGCU			
			GCUGAUGUACAAG			

[1268]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CGGAUCCUGCUGC			
			GCCUGCCCUCCUC			
			CGUGCUGUGCGGC			
			UCCUCCUUCCAGG			
			CCGAGCAGCCCAU			
			CACCGCCCGGUGC			
			GAGCAGUUCUUCC			
			ACCUGGUGAACAG			
			CGAGAUGCGGAAC			
			UUCUGCUCUCACG			
			GCGGCGCUCUGAC			
			CCAGGACAUCACC			
			GCCCACUUCUUCC			
			GCGGCCUGCUGAA			
			CGCCUGCCUGCGG			
			AGCAGGGACCCCA			
			GCCUGAUGGUCGA			
			CUUCAUCCUGGCC			
			AAGUGCCAGACCA			
			AGUGUCCCCUGAU			
			CCUGACCAGCGCC			
			CUGGUUUGGUGGC			
			CCUCACUGGAGCC			
			CGUGCUCCUGUGU			
			CGCUGGAGGCGGC			
			ACUGCCAGUCACC			
			CCUGCCCCGAGAG			
			CUGCAGAAGCUGC			
			AGGAGGGCCGGCA			
			GUUCGCCAGCGAC			
			UUCCUGUCACCAG			
			AGGCAGCCUCUCC			
			CGCUCCUAACCCC			
			GAUUGGCUAUCCG			

[1269]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CGGCCGCUCUGCA			
			CUUCGCCAUCCAG			
			CAGGUGCGGGAGG			
			AGAACAUCCGGAA			
			GCAGCUGAAGAAG			
			CUGGACUGCGAGA			
			GGGAGGAGCUGCU			
			UGUGUUCCUGUUC			
			UUCUUCAGCCUGA			
			UGGGCCUGCUGUC			
			CUCCCACCUGACC			
			AGCAACUCCACCA			
			CCGACCUGCCCAA			
			GGCCUUCCACGUG			
			UGCGCCGCCAUCC			
			UGGAGUGCCUGGA			
			GAAGCGGAAGAUC			
			AGCUGGCUGGCCC			
			UGUUCCAGCUGAC			
			CGAGAGCGACCUG			
			CGCCUGGGGAGAC			
			UGCUGCUGCGCGU			
			GGCUCCCCGACCAG			
			CACACCAGGCUGC			
			UGCCCUUCGCCUU			
			CUACUCCCUGCUG			
			UCCUACUUCCACG			
			AGGACGCCGCCAU			
			CCGGGAGGAGGCC			
			UUCCUGCACGUGG			
			CCGUGGACAUGUA			
			CCUGAAGCUGGUG			
			CAGCUGUUCGUGG			
			CCGGCGACACCUC			

[1270]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CACCGUGAGCCCU CCUGCUGGCCGAA GCCUGGAGCUGAA GGGCCAGGGCAAC CCCGUGGAGCUGA UCACCAAGGCCCG CCUGUUCCUGCUG CAGCUGAUCCCUC GCUGCCCCAAGAA GAGCUUCAGCCAC GUGGCCGAGCUGC UGGCUGACAGGGG AGACUGCGACCCC GAGGUGAGCGCUG CCCUGCAGAGCCG UCAACAAGCGGCC CCUGACCGCCGAUC UGAGCCAGGAGCC CCACCUGUUC			
SEQ ID NO:		1	14	50	117	30
	FANCA_010 化学：G5 帽：C1 多聚 A 尾： 100nt	MSDSWVPNSASG QDPGGRRRAWAE LLAGRVKREKYNP ERAQKLKESAVRL LRSHQDLNALLLE VEGPLCKKLSLSK VIDCDSSEAYANH SSSFIGSALQDQAS RLGVPVGILSAGM VASSVGQICTAPAE TSHPVLLTVEQRK KLSSLLEFAQYLL AHSMFSRLSFCQE	AUGAGCGACAGCU GGGUGCCCAACAG UGCUUCCGGGCAG GAUCCCGGCGGUA GAAGACGGGCUUG GGCUGAGCUGCUG GCCGGCAGGGUGA AGCGGGAGAAGUA CAACCCCGAGCGG GCCCAGAAGCUGA AGGAGAGUGCCGU GCGGCUGCUGCGG AGCCACCAGGACC	GGAAAUCG CAAAAUUU GCUCUUCG CGUUAGAU UUCUUUUA GUUUUCUC GCAACUAG CAAGCUUU UUGUUCUC GCC	UGAUAAUAGGC UGGAGCCUCCA CCGCGUUAUCC GUUCCUCCGUAG GCUGGUCCUGG GGAACGGGUCG GCGGGUACCCC CGUGGUCUUUG AAUAAAGUCUG AGUGGGCGGC	SEQ ID NO:30 从 5'至 3'端由以下组成：SEQ ID NO:50 的 5' UTR、SEQ ID NO:14 的 ORF 序列和 SEQ ID NO:117 的 3' UTR

[1271]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		LWKIQSSLLLEAV	UGAACGCCCUCCU			
		WHLHVQGIVSLQE	GCUGGAGGUCGAG			
		LLESHPDMHAVGS	GGCCCACUGUGCA			
		WLFRNLCCLCEQ	AGAAGCUCAGCCU			
		MEASCQHADVAR	GAGCAAGGUGAUC			
		AMLSDFVQMFVL	GACUGUGACAGCA			
		RGFQKNSDLRRTV	GCGAGGCCUACGC			
		EPEKMPQVTVDVL	CAACCACAGCUCC			
		QRMLIFALDALAA	AGCUUCAUCGGCA			
		GVQEESSTHKIVR	GCGCCCUGCAGGA			
		CWFGVFSGHTLGS	CCAGGCCAGCAGA			
		VISTDPLKRFFSHT	CUGGGCGUGCCCG			
		LTQILTHSPVLKAS	UGGGCAUCCUGAG			
		DAVQMQREWSFA	CGCCGGCAUGGUG			
		RTHPLLTSLYRRLF	GCCAGCUCGGUGG			
		VMLSAEELVGHLQ	GCCAGAUCUGCAC			
		EVLETQEVHWQR	CGCUCCCGCCGAG			
		VLSFVSALVVCFPE	ACCAGCCACCCCG			
		AQQLLEDWVARL	UGCUGCUCACCGU			
		MAQAFESCQLDS	GGAGCAGCGGAAG			
		MVTAFLVVRQAAL	AAGCUGAGCAGCC			
		EGPSAFLSYADWF	UGCUGGAGUUCGC			
		KASFGSTRGYHGC	CCAGUACCUGCUG			
		SKKALVFLFTFLSE	GCCCACAGCAUGU			
		LVPFESPRYLQVHI	UCUCCCGGCUGAG			
		LHPPLVPGKYRSL	CUUUUGCCAAGAG			
		LTDYISLAKTRLA	CUGUGGAAGAUCC			
		DLKVSIENMGLYE	AGAGCUCCCCUGCU			
		DLSSAGDITEPHSQ	GCUGGAGGCCGUG			
		ALQDVEKAIMVFE	UGGCACCUUCACG			
		HTGNIPVTVMEASI	UGCAGGGCAUCGU			
		FRRPYYVSHFLPA	GAGCCUGCAGGAA			
		LLTPRVLPKVPDSR	CUGCUGGAGAGCC			
		VAFIESLKRADKIP	ACCCCGACAUGCA			

[1272]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		PSLYSTYCQACSA	CGCCGUGGGCAGC			
		AEEKPEDAALGVR	UGGCUGUUCCGGA			
		AEPNSAEEPLGQL	ACCUGUGCUGCCU			
		TAALGELRASMTD	GUGCGAGCAGAUG			
		PSQRDVISAQVAVI	GAGGCUAGCUGCC			
		SERLRAVLGHNED	AGCACGCCGACGU			
		DSSVEISKIQLSINT	GGCACGGGCCAUG			
		PRLEPREHMAVDL	CUGAGCGACUUCG			
		LLTSFCQNLMAAS	UGCAGAUGUUCGU			
		SVAPPERQGPWAA	GCUGCGGGGCUUC			
		LFVRTMCGRVLPA	CAGAAGAAUAGCG			
		VLTRLCQLLRHQG	ACCUACGGCGGAC			
		PSLSAPHVLGLAA	CGUGGAGCCCGAG			
		LAVHLGESRSALP	AAGAUGCCCCAGG			
		EVDVGPPAPGAGL	UGACCGUGGACGU			
		PVPALFDSLLTCRT	GCUGCAGCGGAUG			
		RDSLFFCLKFCTA	CUGAUCUUCGCCC			
		AISYSLCKFSSQSR	UGGACGCUCUGGC			
		DTLCSCLSPGLIKK	CGCUGGCGUGCAG			
		FQFLMFRLFSEAR	GAGGAGAGCAGCA			
		QPLSEEDVASLSW	CCCACAAGAUCGU			
		RPLHLPSADWQRA	GCGGUGCUGGUUC			
		ALSLWTHRTFREV	GGCGUGUUCAGCG			
		LKEEDVHLTYQD	GCCACACCCUGGG			
		WLHLELEIQPEAD	CAGCGUGAUCAGC			
		ALSDTERQDFHQ	ACCGACCCACUGA			
		WAIHEHFLPESSAS	AGCGGUUCUUCAG			
		GGCDGDLQAACTI	CCACACCCUCACC			
		LVNALMDFHQSSR	CAGAUCCUGACCC			
		SYDHSENSDLVFG	ACAGCCCCGUGCU			
		GRTGNEDIISRLQE	GAAGGCCAGCGAC			
		MVADLELQQDLIV	GCCGUGCAGAUGC			
		PLGHTPSQEHFLFE	AGCGGGAGUGGAG			
		IFRRRLQALTSGW	CUUCGCCCGGACC			

[1273]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		SVAASLQRQRELL	CACCCACUGCUGA			
		MYKRILLRLPSSVL	CUAGCCUGUACCG			
		CGSSFQAEQPITAR	GCGGCUGUUCGUG			
		CEQFFHLVNSEMR	AUGCUGAGCGCAG			
		NFCSHGGALTQDI	AGGAGCUGGUGGG			
		TAHFFRGLLNACL	CCACCUGCAGGAG			
		RSRDPSLMVDFIL	GUGCUGGAGACCC			
		AKCQTKCPLILTSA	AGGAGGUGCACUG			
		LVWWPSLEPVLLC	GCAGCGGGUGCUG			
		RWRRHCQSPLPRE	AGCUUCGUGAGCG			
		LQKLQEGRQFASD	CCCUGGUGGUGUG			
		FLSPEAASPAPNPD	CUUCCCCGAGGCC			
		WLSAAALHFAIQQ	CAGCAGCUCCUGG			
		VREENIRKQLKKL	AGGACUGGGUGGC			
		DCEREELLVFLFFF	CCGGCUGAUGGCC			
		SLMGLLSSHLTSNS	CAGGCCUUCGAGA			
		TTDLPKAFHVCAA	GCUGCCAGCUGGA			
		ILECLEKRKISWLA	CAGCAUGGUGACC			
		LFQLTESDLRLGRL	GCCUUCCUGGUGG			
		LLRVAPDQHTRLL	UGCGGCAGGCUGC			
		PFAFYSLLSYFHED	CCUGGAAGGCCCC			
		AAIREEAFLHVAV	AGCGCUUUCCUGA			
		DMYLKLVQLFVAG	GCUACGCCGACUG			
		DTSTVSPPAGRSLE	GUUCAAGGCCAGC			
		LKGQGNPVELITK	UUCGGCUCCCACCC			
		ARLFLLQLIPRCPK	GGGGCUACCACGG			
		KSFSHVAELLADR	CUGCAGCAAGAAG			
		GDCDPEVSAALQS	GCCCUGGUGUUCC			
		RQQAAPDADLSQE	UGUUCACCUUCCU			
		PHLF	GAGCGAGCUGGUG			
			CCCUUCGAGUCAC			
			CCCGGUACCUGCA			
			GGUGCACAUCCUG			
			CAUCCUCCGCUGG			

[1274]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			UGCCCGGCAAGUA			
			CCGGAGCCUGCUG			
			ACCGACUACAUCA			
			GCCUGGCCAAGAC			
			CCGGCUGGCCGAC			
			CUGAAGGUGAGCA			
			UCGAGAACAUGGG			
			CCUGUACGAGGAC			
			CUGAGCAGCGCCG			
			GCGACAUCACCGA			
			GCCCCACAGCCAG			
			GCCCUGCAAGACG			
			UGGAGAAGGCCAU			
			CAUGGUGUUCGAG			
			CACACCGGCAACA			
			UCCCCGUGACCGU			
			CAUGGAGGCCAGC			
			AUCUUCCGGCGGC			
			CCUACUACGUGAG			
			CCACUUCCUGCCC			
			GCCCUGCUGACCC			
			CACGGGUGCUGCC			
			CAAGGUGCCCGAC			
			AGCCGGGUGGCCU			
			UCAUCGAGAGCCU			
			GAAGCGGGCCGAC			
			AAGAUCCCACCCA			
			GCCUGUACAGCAC			
			CUACUGCCAGGCC			
			UGCUCUGCCGCCG			
			AGGAGAAGCCCGA			
			GGACGCCGCAUUA			
			GGCGUGCGGGCCG			
			AGCCCAACAGCGC			

[1275]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CGAGGAACCCCUG			
			GGGCAGCUGACCG			
			CAGCCCUGGGAGA			
			GCUGCCGGGCCAGC			
			AUGACCGACCCCA			
			GCCAGCGGGACGU			
			GAUCAGCGCCCAG			
			GUGGCCGUCAUCA			
			GCGAGCGGCUGCG			
			GGCAGUGCUGGGC			
			CACAACGAGGACG			
			ACAGCAGCGUGGA			
			GAUCAGCAAGAUC			
			CAGCUGAGCAUCA			
			ACACCCCACGGCU			
			GGAACCUCGGGAG			
			CACAUGGCCGUGG			
			ACCUGCUGCUGAC			
			CAGCUUCUGCCAG			
			AACCUGAUGGCCG			
			CCUCUAGCGUGGC			
			CCCACCAGAACGG			
			CAGGGUCCCCUGGG			
			CUGCCCUGUUCGU			
			GCGGACCAUGUGC			
			GGCCGGGUUCUGC			
			CCGCUGUGCUGAC			
			CCGGCUGUGCCAG			
			CUGCUGCGGCACC			
			AGGGCCCUAGCCU			
			GAGCGCUCCCCCAC			
			GUGCUGGGCCUAG			
			CUGCCCUGGCCGU			
			GCACCUUGGGGAG			

[1276]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			UCCCGAAGCGCCC			
			UGCCCGAGGUGGA			
			CGUGGGCCCUCCA			
			GCCCCAGGUGCUG			
			GCCUUCCCGUGCC			
			CGCUCUGUUCGAC			
			AGUCUGCUGACCU			
			GCCGGACCCGGGA			
			CAGCCUGUUCUUC			
			UGCCUGAAGUUCU			
			GCACCGCCGCCAU			
			CAGCUACAGCCUG			
			UGCAAGUUCAGCA			
			GCCAGAGCCGGGA			
			CACCCUGUGCAGC			
			UGCCUGAGCCCCG			
			GCCUGAUCAAGAA			
			GUUCCAGUUCCUG			
			AUGUUCCGGCUGU			
			UCAGCGAAGCCCG			
			GCAGCCCCUGAGC			
			GAGGAGGACGUCG			
			CCAGCCUGUCCUG			
			GAGACCCCUGCAC			
			CUGCCCAGCGCCG			
			ACUGGCAGAGAGC			
			CGCCCUGAGCCUG			
			UGGACCCACCGGA			
			CCUUCCGGGAGGU			
			GCUGAAGGAAGAG			
			GACGUGCACCUGA			
			CCUACCAGGACUG			
			GCUGCACCUGGAG			
			CUGGAGAUUCAGC			

[1277]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCGAGGCCGACGC			
			CCUGAGCGACACC			
			GAGCGGCAGGAUU			
			UCCACCAGUGGGC			
			CAUCCACGAGCAU			
			UUCCUGCCCGAGA			
			GCAGCGCCAGCGG			
			CGGCUGUGACGGC			
			GACCUGCAGGCCG			
			CCUGCACCAUCCU			
			GGUGAACGCCCUG			
			ALGGACUUCCACC			
			AGAGCAGCCGGAG			
			CUACGACCACAGC			
			GAGAACAGCGACC			
			UGGUGUUCGGCGG			
			ACGGACCGGCAAC			
			GAGGACAUCAUCA			
			GCCGGCUGCAGGA			
			GAUGGUCGCCGAC			
			CUGGAGUUGCAGC			
			AGGACCUGAUCGU			
			GCCCCUGGGCCAC			
			ACUCCCAGCCAGG			
			AGCACUUCCUGUU			
			CGAGAUCUUCCGC			
			CGGAGACUGCAGG			
			CCCUGACCUCAGG			
			CUGGAGCGUGGCC			
			GCCAGCCUGCAAC			
			GGCAGCGGGAGCU			
			GCUGAUGUACAAG			
			CGGAUCCUGCUGC			
			GGCUGCCCAGCAG			

[1278]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CGUGCUGUGCGGC			
			AGCAGCUUCCAGG			
			CCGAGCAGCCCAU			
			CACGGCCCGGUGC			
			GAGCAGUUCUUCC			
			ACCUGGUGAACAG			
			CGAGAUGCGGAAC			
			UUCUGCAGCCACG			
			GAGGCGCCCUGAC			
			CCAGGACAUCACC			
			GCCCACUUCUUCC			
			GGGGCCUGCUGAA			
			CGCCUGCCUGCGG			
			AGCAGAGACCCCA			
			GCCUGAUGGUGGA			
			CUUCAUCCUGGCC			
			AAGUGCCAGACCA			
			AGUGUCCCCUGAU			
			CCUGACCUCCGCC			
			CUGGUGUGGUGGC			
			CCAGCCUGGAGCC			
			CGUGCUUCUGUGU			
			CGGUGGCGGCGGC			
			ACUGCCAGAGCCC			
			UCUGCCCCGGGAG			
			CUGCAGAAGCUGC			
			AGGAGGGCCGGCA			
			GUUCGCCAGCGAC			
			UUCCUGUCUCCCG			
			AGGCCGCUAGCCC			
			AGCCCCAAACCCC			
			GACUGGCUGAGCG			
			CUGCCGCCCUGCA			
			CUUCGCCAUCCAG			

[1279]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CAGGUGCGGGAGG			
			AGAACAUCCGGAA			
			GCAGCUGAAGAAG			
			CUGGACUGCGAGC			
			GGGAGGAGCUGCU			
			CGUGUUCCUCUUC			
			UUCUUCAGCCUCA			
			UGGGCCUUCUGAG			
			CAGCCACCUGACC			
			AGCAACAGCACCA			
			CCGACCUGCCCAA			
			GGCCUUCCACGUG			
			UGCGCCGCCAUCC			
			UGGAGUGCCUGGA			
			GAAGCGGAAGAUC			
			AGCUGGCUGGCCC			
			UGUUCCAGCUGAC			
			CGAGAGCGACCUG			
			CGGCUGGGCCGAC			
			UGCUGCUGCGGGU			
			GGCCCCAGACCAG			
			CACACCCGGCUGC			
			UGCCCUUCGCCUU			
			CUACAGCCUGCUG			
			AGCUACUUCCACG			
			AGGACGCUGCCAU			
			CCGGGAGGAGGCC			
			UUUCUGCACGUGG			
			CAGUGGACAUGUA			
			CCUGAAGCUGGUG			
			CAGCUGUUCGUGG			
			CCGGCGACACCAG			
			CACCGUGUCACCA			
			CCAGCCGGCCGCU			

[1280]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCCUGGAGCUGAA GGGCCAGGGCAAC CCCGUGGAGCUGA UCACCAAGGCCCG GCUGUUCCUGCUG CAGCUGAUCCCAC GGUGCCCCAAGAA GAGCUUCAGCCAC GUGGCCGAGCUGC UUGCCGACCGGGG AGACUGCGACCCC GAGGUGAGCGCCG CCUUGCAGAGCCG GCAGCAGGCCGCA CCUGACGCCGACC UGAGCCAGGAGCC CCACCUGUUC			
SEQ ID NO:		1	2	61	117	31
	FANCA_011 化学：G5 帽：C1 多聚 A 尾： 100nt	MSDSWVPNSASG QDPGGRRRAWAE LLAGRVKREKYNP ERAQKLKESAVRL LRSHQDLNALLLE VEGPLCKKLSLSK VIDCDSSEAYANH SSSFIGSALQDQAS RLGVPVGILSAGM VASSVGQICTAPAE TSHPVLLTVEQRK KLSSLLEFAQYLL AHSMFSRLSFCQE LWKIQSSLLLEAV WHLHVQGIVSLQE	AUGUCCGACUCCU GGGUGCCCAACUC CGCUAGCGGACAA GAUCCAGGUGGCC GCAGGAGAGCGUG GGCUGAACUGCUU GCCGGGCGAGUGA AGCGCGAGAAGUA CAACCCCGAGCGC GCCCAGAAGCUGA AGGAGAGCGCCGU ACGGCUGCUGCGC AGCCACCAGGACC UGAACGCCCUGCU GCUGGAGGGUGGAG	GGAAACUU UAUUUAGU GUUACUUU AUUUUCUG UUUAUUUG UGUUUCUU CAGUGGGU UUGUUCUA AUUUCCUU GGCCGCC	UGAUAAUAGGC UGGAGCCUCCA CCGCGUUAUCC GUUCCUCGUAG GCUGGUCCUGG GGAACGGGUCG GCGGGUACCCC CGUGGUCUUUG AAUAAAGUCUG AGUGGGCGGC	SEQ ID NO:31 从 5'至 3'端由以下组成：SEQ ID NO:61 的 5' UTR 、SEQ ID NO:2 的 ORF 序列和 SEQ ID NO:117 的 3' UTR

[1281]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		LLESHPDMHAVGS	GGGCCCCUGUGCA			
		WLFRNLCCLCEQ	AGAAGCUGAGCCU			
		MEASCQHADVAR	GUCCAAGGUGAUC			
		AMLSDFVQMFVL	GACUGCGACAGCA			
		RGFQKNSDLRRTV	GCGAGGCCUACGC			
		EPEKMPQVTVDVL	CAACCACAGCUCC			
		QRMLIFALDALAA	UCCUUCAUCGGCU			
		GVQEESSTHKIVR	CCGCCCUGCAGGA			
		CWFGVFSGHTLGS	UCAAGCUAGCCGG			
		VISTDPLKRFFSHT	CUGGGCGUUCCAG			
		LTQILTHSPVLKAS	UGGGCAUCCUGAG			
		DAVQMQREWSFA	CGCCGGCAUGGUG			
		RTHPLLTSLYRRLF	GCCAGCAGCGUGG			
		VMLSAEELVGHLQ	GCCAGAUCUGCAC			
		EVLETQEVHWQR	CGCUCCCGCCGAG			
		VLSFVSALVVCFPE	ACCAGCCACCCCG			
		AQQLLEDWVARL	UGCUGCUGACCGU			
		MAQAFESCQLDS	GGAGCAGCGCAAG			
		MVTAFLVVRQAAL	AAGCUGUCAAGCC			
		EGPSAFLSYADWF	UGCUGGAGUUCGC			
		KASFGSTRGYHGC	CCAGUACCUGCUG			
		SKKALVFLFTFLSE	GCCCACAGCAUGU			
		LVPFESPRYLQVHI	UCUCCCGCCUGAG			
		LHPPLVPGKYRSL	CUUCUGCCAGGAG			
		LTDYISLAKTRLA	CUGUGGAAGAUCC			
		DLKVSIENMGLYE	AGUCCAGCCUGCU			
		DLSSAGDITEPHSQ	ACUGGAGGCCGUG			
		ALQDVEKAIMVFE	UGGCACUUACACG			
		HTGNIPVTVMEASI	UGCAGGGCAUCGU			
		FRRPYYVSHFLPA	GAGCCUGCAGGAG			
		LLTPRVLPKVPDSR	CUACUGGAGAGCC			
		VAFIESLKRADKIP	ACCCCGACAUGCA			
		PSLYSTYCQACSA	CGCCGUGGGGUCC			
		AEEKPEDAALGVR	UGGCUGUUCCGCA			

[1282]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		AEPNSAEEPLGQL	ACCUGUGCUGCCU			
		TAALGELRASMTD	GUGCGAGCAGAUG			
		PSQRDVISAQVAVI	GAAGCCUCCCUGCC			
		SERLRAVLGHNED	AACACGCCGACGU			
		DSSVEISKIQLSINT	UGCACGGGCAAUG			
		PRLEPREHMAVDL	CUGAGCGACUUCG			
		LLTSFCQNLMAAS	UGCAGAUGUUCGU			
		SVAPPERQGPWAA	GCUGCGCGGGUUC			
		LFVRTMCGRVLPA	CAGAAGAACUCAG			
		VLTRLCQLLRHQG	ACCUCCGGAGGAC			
		PSLSAPHVLGLAA	CGUGGAGCCUGAG			
		LAVHLGESRSALP	AAGAUGCCCCAGG			
		EVDVGPPAPGAGL	UGACCGUGGACGU			
		PVPALFDSLLTCRT	GCUGCAGCGGAUG			
		RDSLFFCLKFCTA	CUGAUCUUUGCCC			
		AISYSLCKFSSQSR	UGGACGCUCUCGC			
		DTLCSCLSPGLIKK	UGCCGGAGUGCAG			
		FQFLMFRLFSEAR	GAGGAGAGCUCCA			
		QPLSEEDVASLSW	CCCACAAGAUCGU			
		RPLHLPSADWQRA	GCGGUGCUGGUUC			
		ALSLWTHRTFREV	GGCGUGUUCUCCG			
		LKEEDVHLTYQD	GCCACACCCUGGG			
		WLHLELEIQPEAD	CUCCGUGAUCUCC			
		ALSDTERQDFHQ	ACCGAUCCCCUGA			
		WAIHEHFLPESSAS	AGCGGUUCUUCUC			
		GGCDGDLQAACTI	CCACACCCUGACC			
		LVNALMDFHQSSR	CAGAUCCUGACCC			
		SYDHSENSDLVFG	ACUCCCCUGUGCU			
		GRTGNEDIISRLQE	GAAGGCCAGCGAC			
		MVADLELQQDLIV	GCCGUGCAGAUGC			
		PLGHTPSQEHFLFE	AGCGCGAGUGGUC			
		IFRRRLQALTSGW	CUUCGCCCCGCACC			
		SVAASLQRQRELL	CAUCCCCUGCUGA			
		MYKRILLRLPSSVL	CCAGCCUGUACCG			

[1283]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
		CGSSFQAEQPITAR	GCGGCUGUUCGUG			
		CEQFFHLVNSEMR	AUGCUGUCCGCCG			
		NFCSHGGALTQDI	AGGAGCUGGUGGG			
		TAHFFRGLLNACL	CCACCUGCAGGAG			
		RSRDPSLMVDFIL	GUGCUGGAGACCC			
		AKCQTKCPLILTSA	AGGAGGUGCACUG			
		LVWWPSLEPVLLC	GCAGCGCGUGCUG			
		RWRRHCQSPLPRE	AGCUUCGUGAGCG			
		LQKLQEGRQFASD	CCCUGGUGGUGUG			
		FLSPEAASPAPNPD	CUUCCCCGAGGCC			
		WLSAAALHFAIQQ	CAGCAGCUCCUGG			
		VREENIRKQLKKL	AGGACUGGGUGGC			
		DCEREELLVFLFFF	CCGGCUGAUGGCC			
		SLMGLLSSHLTSNS	CAGGCCUUCGAGA			
		TTDLPKAFHVCAA	GCUGCCAGCUGGA			
		ILECLEKRKISWLA	CAGCAUGGUGACC			
		LFQLTESDLRLGRL	GCCUUCCUGGUGG			
		LLRVAPDQHTRLL	UGCGUCAAGCAGC			
		PFAFYSLLSYFHED	CCUGGAGGGCCCA			
		AAIREEAFLHVAV	AGCGCCUUCCUGA			
		DMYLKLVQLFVAG	GCUACGCCGACUG			
		DTSTVSPPAGRSLE	GUUCAAGGCCAGC			
		LKGQGNPVELITK	UUCGGGUCCACCC			
		ARLFLLQLIPRCPK	GCGGGUACCACGG			
		KSFSHVAELLADR	CUGCUCCAAGAAG			
		GDCDPEVSAALQS	GCCCUGGUGUUCC			
		RQQAAPDADLSQE	UGUUUACCUUCCU			
		PHLF	GAGCGAGCUGGUG			
			CCCUUCGAGUCGC			
			CCCGCUACCUGCA			
			GGUGCACAUCCUG			
			CACCCACCCCUGG			
			UCCCUGGCAAGUA			
			CCGGAGCCUGCUG			

[1284]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			ACCGACUACAUCU			
			CCCUGGCCAAGAC			
			CCGGCUGGCCGAC			
			CUGAAGGUGAGCA			
			UCGAGAACAUGGG			
			CCUGUACGAGGAC			
			CUGUCCAGCGCCG			
			GCGACAUCACCGA			
			GCCCCACAGCCAG			
			GCCCUGCAGGACG			
			UGGAGAAGGCCAU			
			CAUGGUGUUCGAG			
			CACACCGGCAACA			
			UCCCCGUGACCGU			
			GAUGGAGGCUAGU			
			ACUUCCGGCGGC			
			CCUACUACGUGUC			
			CCACUUCCUGCCC			
			GCCCUGCUGACAC			
			CCCGGGUGCUGCC			
			CAAGGUGCCCGAC			
			AGUCGCGUCGCCU			
			UCAUCGAGAGCCU			
			GAAGCGGGGCCGAC			
			AAGAUCCCGCCCU			
			CCCUGUACUCCAC			
			CUACUGCCAGGCU			
			UGCUCCGGCCGCCG			
			AGGAGAAGCCCGA			
			GGACGCAGCCUUA			
			GGAGUACGCGCCG			
			AGCCCAACAGCGC			
			CGAGGAGCCACUG			
			GGCCAGUUGACCG			

[1285]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CUGCCCUGGGAGA			
			GCUGCGGGCCAGC			
			AUGACCGACCCCA			
			GCCAACGCGACGU			
			GAUCAGCGCCCAG			
			GUGGCCGUGAUCA			
			GCGAGAGGCUGCG			
			GGCAGUGCUGGGC			
			CACAACGAGGACG			
			ACAGCAGCGUGGA			
			GAUCAGCAAGAUC			
			CAGCUGAGCAUCA			
			ACACACCCAGGCU			
			GGAGCCCAGAGAG			
			CACAUGGCCGUGG			
			ACCUGCUGCUGAC			
			CUCCUUCUGCCAG			
			AACCUGAUGGCCG			
			CAAGCAGCGUGGC			
			UCCUCCCAGAACGC			
			CAAGGCCCUUGGG			
			CUGCCCUGUUCGU			
			GCGCACCAUGUGU			
			GGCCGCGUCCUGC			
			CAGCAGUGCUGAC			
			CCGGCUGUGCCAG			
			CUUCUGCGGCACC			
			AAGGGCCCUCCCU			
			GAGCGCUCCCCAC			
			GUGCUUGGGUUGG			
			CCGCCCUCGCUGU			
			GCACCUGGGCGAG			
			AGCAGGAGCGCCC			
			UUCCCGAGGUGGA			

[1286]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CGUGGGUCCUCCA			
			GCCCCAGGUGCCG			
			GACUCCCCGUGCC			
			AGCCCUGUUCGAC			
			AGCCUGCUGACUU			
			GUCGCACCCGGGA			
			CUCCCUGUUCUUC			
			UGCCUGAAGUUCU			
			GCACCGCCGCCAU			
			CUCCUACUCCCUG			
			UGCAAGUUCUCCA			
			GCCAGUCCCGGGA			
			CACCCUGUGCUCC			
			UGCCUGUCUCCCG			
			GCCUGAUCAAGAA			
			GUUCCAGUUCCUG			
			AUGUUCCGCCUGU			
			UCAGCGAAGCACG			
			CCAGCCCCUGUCC			
			GAGGAGGACGUGG			
			CCUCCCUGUCCUG			
			GCGACCCCUGCAC			
			CUGCCCUCCGCCG			
			AUUGGCAGCGCGC			
			CGCCCUUAGCCUG			
			UGGACCCACCGGA			
			CCUUCCGCGAGGU			
			GCUGAAGGAGGAA			
			GACGUGCACCUGA			
			CCUACCAGGACUG			
			GCUGCACCUGGAG			
			CUGGAGAUCCAGC			
			CCGAAGCCGACGC			
			CCUGAGCGACACC			

[1287]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GAGCGCCAGGACU			
			UUCACCAGUGGGC			
			CAUCCACGAGCAC			
			UUUCUGCCCGAGA			
			GCAGCGCCAGUGG			
			AGGCUGCGACGGC			
			GACCUGCAAGCCG			
			CCUGCACCAUCCU			
			GGUGAACGCCCUG			
			AUGGACUUCCACC			
			AGAGCUCCCGCAG			
			CUACGACCACUCC			
			GAGAACAGCGACC			
			UGGUGUUUGGAGG			
			CCGCACCGGGAAC			
			GAGGACAUCAUCA			
			GCCGCCUGCAGGA			
			GAUGGUGGCCGAC			
			CUGGAGCUCCAGC			
			AGGACCUGAUCGU			
			GCCACUGGGGCAC			
			ACCCCAAGCCAGG			
			AGCACUUCCUGUU			
			CGAGAUCUUCCGC			
			AGACGCCUGCAGG			
			CACUGACCUCCGG			
			CUGGAGUGUGGCC			
			GCAAGCCUGCAGC			
			GCCAACGCGAGCU			
			GCUGAUGUACAAG			
			CGGAUCCUGCUGC			
			GCCUGCCCUCCUC			
			CGUGCUGUGCGGC			
			UCCUCCUUCCAGG			

[1288]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCGAGCAGCCCAU			
			CACCGCCCCGGUGC			
			GAGCAGUUCUUCC			
			ACCUGGUGAACAG			
			CGAGAUGCGGAAC			
			UUCUGCUCUCACG			
			GCGGCGCUCUGAC			
			CCAGGACAUCACC			
			GCCCACUUCUUCC			
			GCGGCCUGCUGAA			
			CGCCUGCCUGCGG			
			AGCAGGGACCCCA			
			GCCUGAUGGUCGA			
			CUUCAUCCUGGCC			
			AAGUGCCAGACCA			
			AGUGUCCCCUGAU			
			CCUGACCAGCGCC			
			CUGGUUUGGUGGC			
			CCUCACUGGAGCC			
			CGUGCUCCUGUGU			
			CGCUGGAGGCGGC			
			ACUGCCAGUCACC			
			CCUGCCCCGAGAG			
			CUGCAGAAGCUGC			
			AGGAGGGCCGGCA			
			GUUCGCCAGCGAC			
			UUCCUGUCACCAG			
			AGGCAGCCUCUCC			
			CGCUCCUAACCCC			
			GAUUGGCUAUCCG			
			CGGCCGCUCUGCA			
			CUUCGCCAUCCAG			
			CAGGUGCGGGAGG			
			AGAACAUCCGGAA			

[1289]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			GCAGCUGAAGAAG			
			CUGGACUGCGAGA			
			GGGAGGAGCUGCU			
			UGUGUUCCUGUUC			
			UUCUUCAGCCUGA			
			UGGGCCUGCUGUC			
			CUCCCACCUGACC			
			AGCAACUCCACCA			
			CCGACCUGCCCAA			
			GGCCUUCCACGUG			
			UGCGCCGCCAUCC			
			UGGAGUGCCUGGA			
			GAAGCGGAAGAUC			
			AGCUGGCUGGCCC			
			UGUUCCAGCUGAC			
			CGAGAGCGACCUG			
			CGCCUGGGGAGAC			
			UGCUGCUGCGCGU			
			GGCUCCCGACCAG			
			CACACCAGGCUGC			
			UGCCCUUCGCCUU			
			CUACUCCCUGCUG			
			UCCUACUUCCACG			
			AGGACGCCGCCAU			
			CCGGGAGGAGGCC			
			UUCCUGCACGUGG			
			CCGUGGACAUGUA			
			CCUGAAGCUGGUG			
			CAGCUGUUCGUGG			
			CCGGCGACACCUC			
			CACCGUGAGCCCU			
			CCUGCUGGCCGAA			
			GCCUGGAGCUGAA			
			GGGCCAGGGCAAC			

[1290]

	mRNA 名称	ORF 序列(氨基酸)	ORF 序列(核苷酸)	5' UTR 序列	3' UTR 序列	构建体序列
			CCCGUGGAGCUGA UCACCAAGGCCCG CCUGUUCCUGCUG CAGCUGAUCCCUC GCUGCCCCAAGAA GAGCUUCAGCCAC GUGGCCGAGCUGC UGGCUGACAGGGG AGACUGCGACCCC GAGGUGAGCGCUG CCCUGCAGAGCCG UCAACAAGCCGGCC CCUGACGCCGAUC UGAGCCAGGAGCC CCACCUGUUC			

[1291] 实施例4.FANCA mRNA的表达

[1292] 使用编码FANCA的密码子优化的ORF(SEQ ID NO:1)和mRNA控制元件(包括5′UTR和3′UTR)各自产生mRNA FANCA_01(SEQ ID NO:4)和FANCA_02(SEQ ID NO:5)。

[1293] FaDu细胞从范可尼贫血癌细胞系资源获得并且来源于散发性头颈部鳞状细胞癌(ATCC#HTB43)。FANCA敲除(KO)细胞系在CAS9介导的FANCA基因的N末端缺失靶向之后以克

隆方式来源于FaDu细胞系,导致FANCA蛋白表达和功能丧失。

[1294] FaDu、Tg(由用GFP mRNA转染的源自克隆的FANCA-/-亚系补充的转基因)和FaDu-KO细胞在转染前一天下午在6孔板(250,000-300,000个细胞/孔)中接种。次日早晨,根据制造商的说明书,使用Lipofectamine MessengerMAX(Invitrogen,参考编号LMRNA003)在1.5:1ul/ug比率的lipofectamine:mRNA下用指定量的FANCA_01-FANCA_11mRNA或用GFP mRNA(作为对照)转染细胞。在转染后6小时更换培养基,并且用适当的指定处理来替代。转染后6、24、48或72小时在1XNuPAGE LDS样品缓冲液中裂解细胞。使用Cell Signaling 14657抗体通过蛋白质印迹来确定FANCA蛋白表达。使用针对核仁素的Cell Signaling 87792抗体对蛋白质上样进行归一化。FANCA/核仁素表达相对于模拟处理的FaDu细胞中的平均表达的倍数变化。

[1295] 图3A是显示在指定时间点的FaDu三联体细胞系中FANCA和核仁素的蛋白质水平的蛋白质印迹图像。计算FANCA表达的倍数变化,将其针对上样对照和野生型(WT)平均值进行归一化。FANCA mRNA在体外表达蛋白质,长达三天。

[1296] 图3B是显示FANCA/核仁素表达相对于野生型中的平均表达的倍数变化的图。直至转染后72小时,所有FANCA构建体都维持与野生型相当或更高的蛋白质表达水平。

[1297] 实施例5.FANCA mRNA挽救经MMC处理后的细胞存活

[1298] FA患者通常对链间DNA交联剂(诸如DEB和MMC)具有增加的敏感性。为了确定FANCA mRNA的表达是否可以使FANCA缺陷细胞免于发生G2_M期中的细胞周期停滞,在PBS(模拟)、DEB或MMC处理后24小时,用1ug GFP mRNA或FANCA mRNA转染FaDu、Tg和FaDu-KO细胞。在MMC的情况下,在处理1小时后将其去除。48小时后,收集细胞,加以固定,并且在RNAse A存在下用DAPI染色以进行细胞周期分析。为了确定FANCA mRNA的表达是否可以挽救FANCA缺陷细胞免于发生对MMC的敏感性,在PBS(模拟)或不同浓度的MMC处理后6小时,用1ug GFP mRNA或FANCA mRNA转染FaDu、Tg和FaDu-KO细胞。在MMC存在下培养后5天,随后使用Cell Titer Glo试剂(Promega#G7570)裂解经处理的细胞以释放总ATP,并且在Bio-Tek Synergy H1上读取发光。将每个数据点的五个重复试样针对未处理或转染对照进行归一化。

[1299] 图4是显示用1ug GFP mRNA构建体转染的FaDu-WT、FaDu-KO或用1ug FANCA构建体转染并在转染后24小时用指定浓度的MMC处理的FaDu-KO的存活百分比的图。使用cell titer Glo评估处理后5天的存活率。

[1300] FANCA在FANCA敲除细胞中的再表达挽救存活达到类似于野生型的水平。

[1301] 实施例6.FANCA mRNA部分地挽救细胞周期

[1302] 为了确定FANCA mRNA的表达是否可以挽救在用DEB或MMC处理后FANCA缺陷细胞免于发生G2_M期中的细胞周期停滞,将GFP或FANCA(1ug)mRNA转染的FaDu、Tg和FaDu-KO细胞用PBS模拟处理或用0.05ug/ml MMC(Sigma,参考编号M0503)处理1小时或用0.1ug/ml DEB(Sigma,参考编号90474)处理48小时。在用mRNA构建体转染后6小时(图3A-图3B)或72小时(图4A-图4B),模拟处理或者用MMC或DEB处理细胞。在添加处理后48小时使用0.25%胰蛋白酶收获细胞,并且用PBS洗涤。

[1303] 根据制造商的说明书,使用BD Cytofix/Cytoperm(BD,参考编号554714)固定之前,使用LIVE/DEAD可固定远红(Invitrogen,参考编号L10120)染色剂对死细胞进行标记。固定后,将细胞在含有1ug/ml4′,6-二脒基-2-苯基吲哚(DAPI)(Invitrogen,参考编号

D21490)和100ug/ml RNaseA(Thermo Scientific,参考编号EN0531)的PBS中重构15分钟,随后使用SONY-SH800Z进行分析。

[1304] 图5A是显示用1ug FANCA mRNA或GFP mRNA转染,继而6小时后用对照(PBS)、DEB或MMC处理后在细胞周期的G2_M期中的细胞积聚的图。与GFP转染的FaDu-WT和FaDu KO细胞相比,细胞周期分析证明,1ug FANCA mRNA构建体转染使得在用DEB和MMC处理后48小时部分挽救细胞周期的G2_M期中的FaDu KO细胞积聚。

[1305] 图5B是显示图5A的数据的G2M频率的图,表示为平均值±SD。

[1306] 图6A是显示用1ug FANCA mRNA或GFP mRNA转染,继而72小时后用对照(PBS)、DEB或MMC处理在细胞周期的G2_M期中的细胞积聚的图。与GFP转染的FaDu-WT和FaDu KO细胞相比,细胞周期分析证明,1ug FANCA mRNA构建体转染使得在用DEB和MMC处理后48小时部分挽救细胞周期的G2_M期中的FaDu KO细胞积聚。

[1307] 图6B是显示图6A的数据的G2M频率的图,表示为平均值±SD。

[1308] 图6C是显示针对核仁素进行归一化且表示为相对于WT FaDu细胞系中的FANCA的倍数变化的FANCA表达水平的图。

[1309] 实施例7.具有v2.0 5′ UTR与κ3′UTR的组合的mRNA的体内作用

[1310] 本实施例描述了疫苗诱导的免疫反应的体内评估。

[1311] 为了评价κ 3′UTR(SEQ ID NO:139)当与v1.1 5′UTR(SEQ ID NO:56)或v2.0 5′UTR(SEQ ID NO:50)组合时在小鼠体内所具有的作用,向CL57BL/6小鼠(N=8)肌肉内注射在8-((2-羟乙基)(6-氧代-6-(十一烷氧基)己基)氨基)辛酸十七烷-9-基酯中配制的2μg mRNA。mRNA编码分泌蛋白卵白蛋白。在给药后6小时和48小时,获取50μl血液并产生血清。通过ELISA评估血清中的卵白蛋白浓度(图7)。

[1312] 两个剂量相隔28天施用。在两个时间点评价血清中的抗体水平:第28天(即,初免后4周)和第43天(即,加强后2周)。抗体水平以终点滴度表示,终点滴度是可以检测到抗体与卵白蛋白抗原包被的板的结合的血清的最大稀释度。

[1313] 结果显示,包含v2.0 5′UTR(SEQ ID NO:50)和κ 3′UTR(SEQ ID NO:139)的mRNA使抗体反应在初免后增加约25倍并且在加强后增加约5倍(图8)。

[1314] 实施例8.在3′UTR中具有FUT8序列的体内作用

[1315] 为了评价3′UTR中的FUT8序列对mRNA所具有的影响,将编码所关注的蛋白质的mRNA配制为LNP(化合物II+GL67),然后分层至原代人类支气管上皮细胞模型上。评估了若干序列变体:

[1316] 1)历史参考序列(v1.0UTR)

[1317] 2)具有v2.0 5'UTR(SEQ ID:NO:50)和重新优化的编码序列的mRNA

[1318] 3)具有来自(2)的元件加上带有'δ'终止盒和'FUT8'序列的3'UTR的mRNA(3′UTR对应于SEQ ID NO:140)。

[1319] 另外,评估了可以挽救突变基因的活性的两个阳性对照。

[1320] 图9显示了3'UTR中FUT8序列的作用。结果显示,具有FUT8增加靶蛋白活性。

[1321] 实施例9.v2.0 5′UTR增强小鼠中编码血凝素的mRNA的免疫原性

[1322] 向BALB/c小鼠注射编码四种血凝素(HA)蛋白(HA1、HA2、HA3和HA4)的四种mRNA的共配制物,所述四种mRNA具有v2.0 5′UTR和具有ι终止盒的v1.1 3′UTR(分别是SEQ ID NO:

50和180)。具体地,将带有UTR的HA mRNA在SM-102LNP中共配制。以3周初免-加强时间间隔肌肉内给与LNP。在指定天数评估血清中的抗体水平,代表剂量1后和剂量2后(图中的d21和d36)。评估以下剂量:1μg、0.2μg和0.04μg,以及单独的PBS作为对照。在剂量1后,主要观察到所有4种血凝素(HA)的抗体滴度增加。在剂量2后,在脾中未观察到对CD4 T细胞对H1和H3 HA的反应的影响。在剂量2后,在脾中观察到针对H1 HA的较高CD8 CD107a T细胞反应(参见例如图10)。在剂量1后和剂量2后6小时还观察到IP-10水平增加。一般来说,v2.0 5′UTR的存在在第一剂量后和中等剂量水平下增强免疫原性。

[1323] 实施例10.v2.0 5′UTR和κ 3′UTR的组合增强编码血凝素的mRNA的体外表达

[1324] 与实施例9类似,用编码四种血凝素(HA)蛋白(HA1、HA2、HA3和HA4)的四种mRNA的共配制物转染细胞,所述四种mRNA具有v1.1 5′UTR和v1.1 3′UTR(分别是SEQ ID NO:56和180)或v2.0 5′UTR和κ 3′UTR(分别是SEQ ID NO:50和SEQ ID NO:139)。再次,将带有任一组UTR的HA mRNA在SM-102 LNP中共配制。以3周初免-加强时间间隔肌肉内给与LNP。在指定天数评估血清中的抗体水平,代表剂量1后和剂量2后。还用四种不同的已知流感疫苗(CR9114、CR8059、5e04和2B06)处理细胞。在mRNA给药后24小时、48小时和72小时监测MFI阳性细胞的数量。如图11中所示,在低剂量的共配制mRNA下,主要在72小时观察到四种HA蛋白的表达增加。

[1325] 其它实施方案

[1326] 将了解,虽然本发明已经结合其详细描述进行了描述,但前述描述旨在说明而不是限制由所附权利要求书的范围限定的本发明的范围。其它方面、优点和修改在所附权利要求书的范围内。

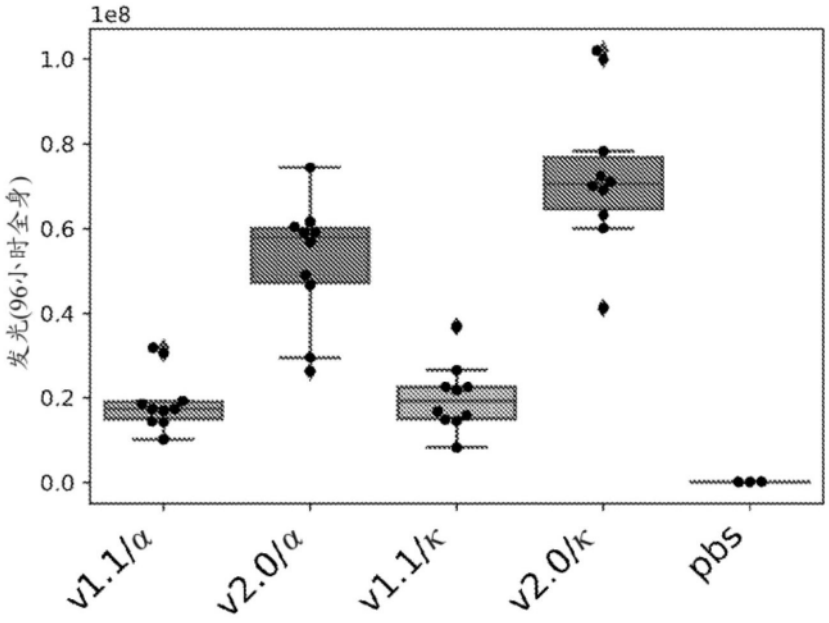

1e8
1.0
0.8
0.6
0.4
0.2
0.0
发光(96小时全身)
v1.1/α
v2.0/α
v1.1/κ
v2.0/κ
pbs

图1A

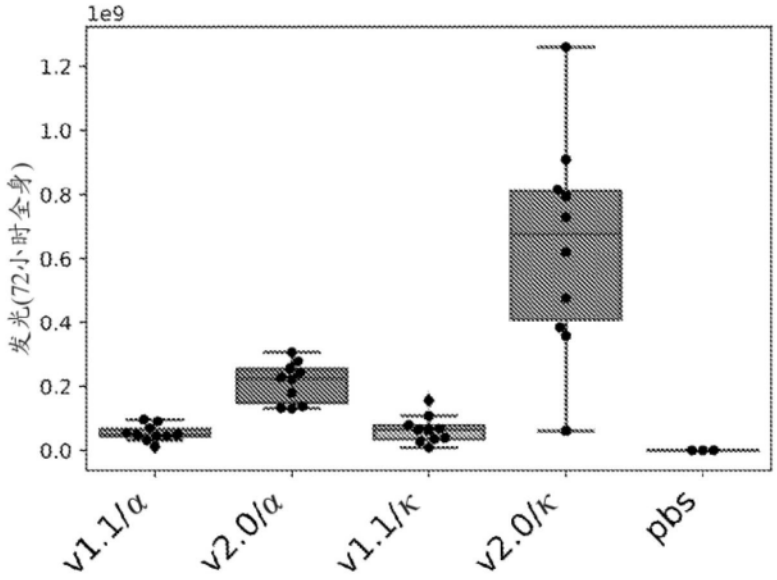

1e9
1.2
1.0
0.8
0.6
0.4
0.2
0.0
发光(72小时全身)
v1.1/α
v2.0/α
v1.1/κ
v2.0/κ
pbs

图1B

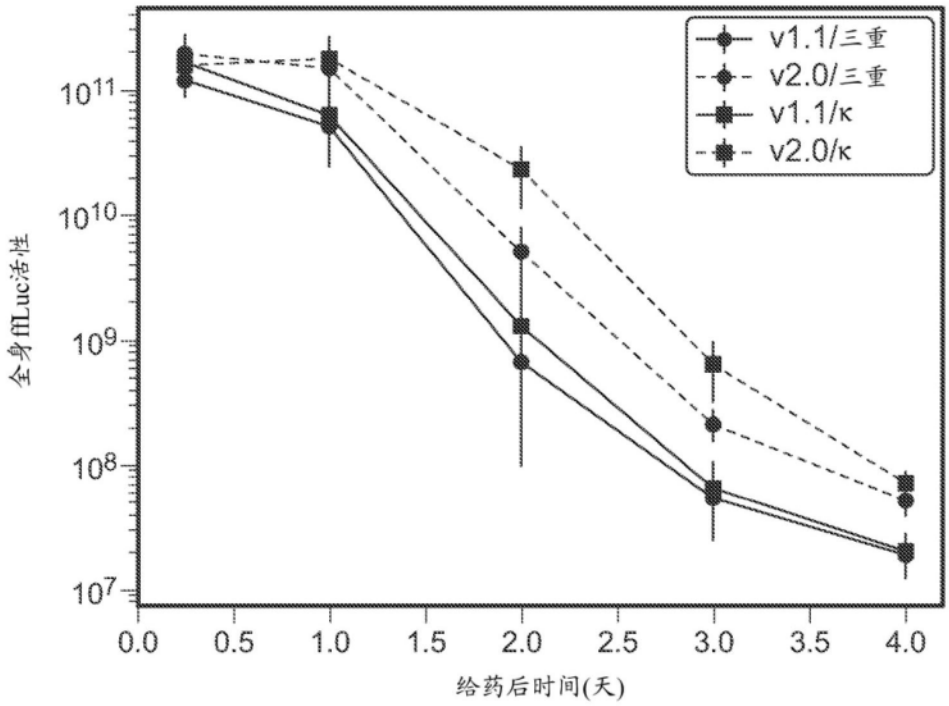

v1.1/三重
v2.0/三重
v1.1/κ
v2.0/κ
全身fLuc活性
10^{11}
10^{10}
10^{9}
10^{8}
10^{7}
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
给药后时间(天)

图1C

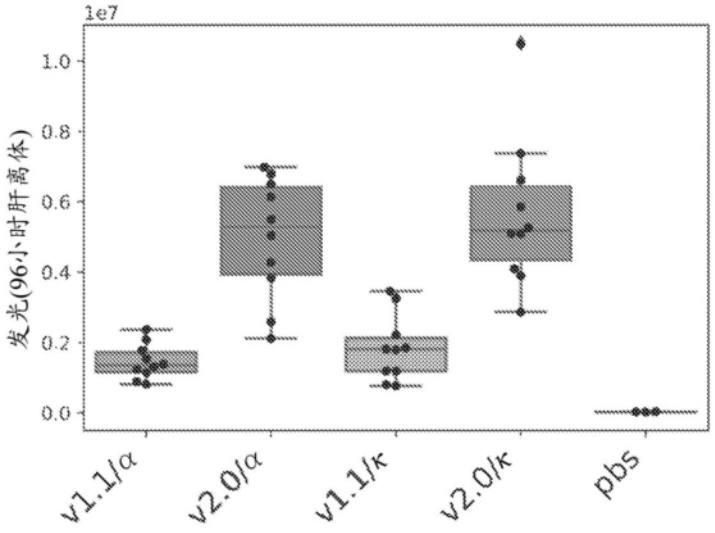

1e7
1.0
0.8
0.6
0.4
0.2
0.0
发光(96小时肝离体)
v1.1/α
v2.0/α
v1.1/κ
v2.0/κ
pbs

图1D

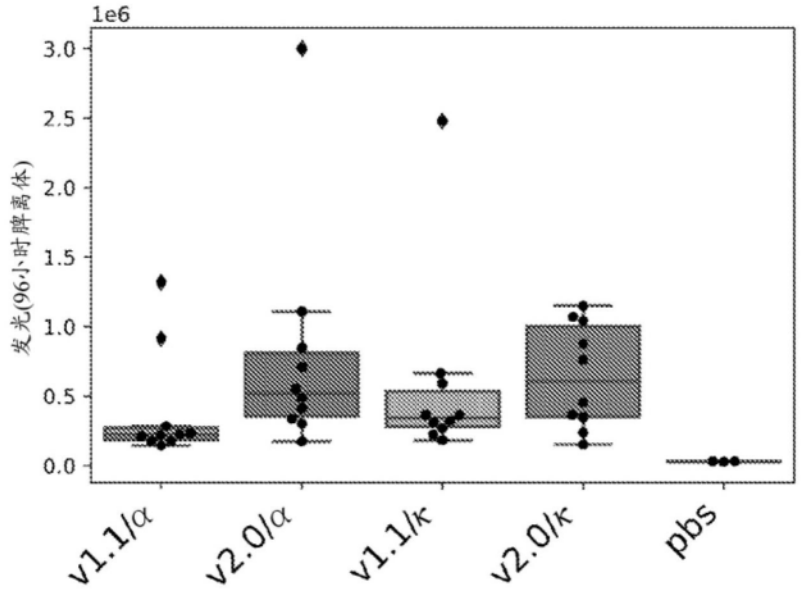

1e6
3.0
2.5
2.0
1.5
1.0
0.5
0.0
发光(96小时脾离体)
v1.1/α
v2.0/α
v1.1/κ
v2.0/κ
pbs

图1E

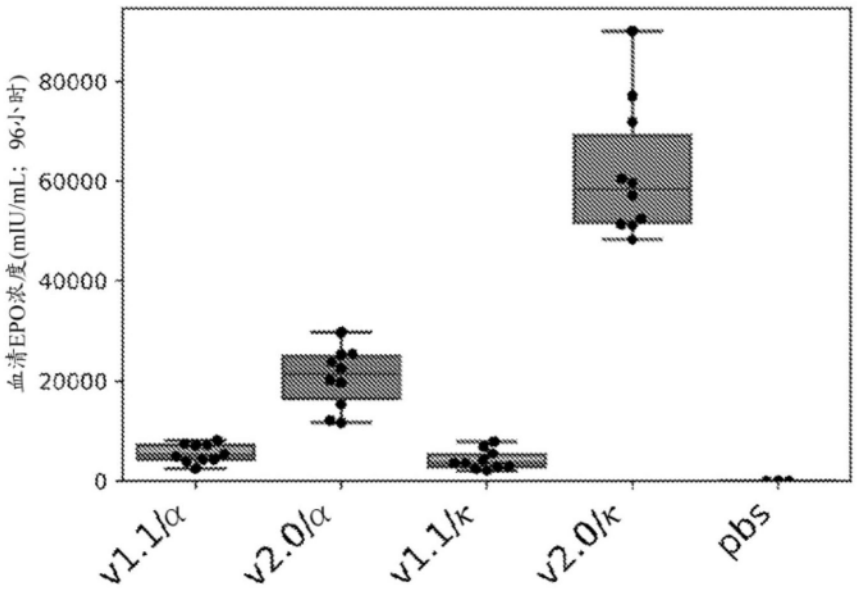

80000
60000
40000
20000
0
血清EPO浓度(mIU/mL；96小时)
v1.1/α
v2.0/α
v1.1/κ
v2.0/κ
pbs

图1F

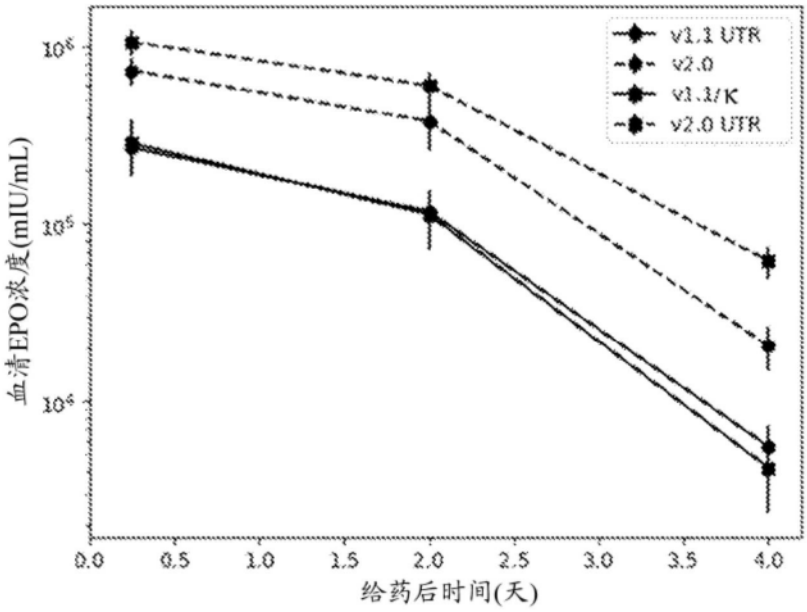

v1.1 UTR
v2.0
v1.1/K
v2.0 UTR
血清EPO浓度(mIU/mL)
给药后时间(天)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0

图1G

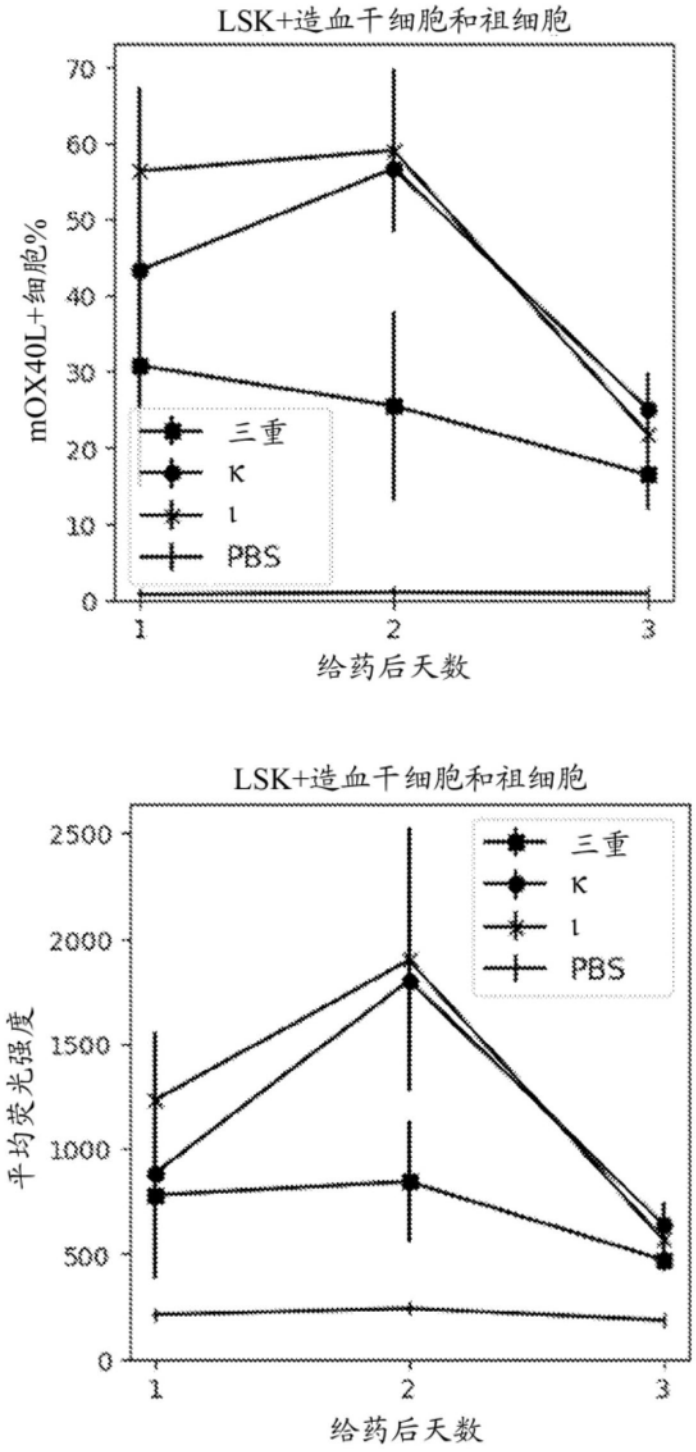

LSK+造血干细胞和祖细胞
mOX40L+细胞%
70
60
50
40
30
20
10
0
三重
κ
ι
PBS
1
2
3
给药后天数
LSK+造血干细胞和祖细胞
平均荧光强度
2500
2000
1500
1000
500
0
三重
κ
ι
PBS
1
2
3
给药后天数

图2A

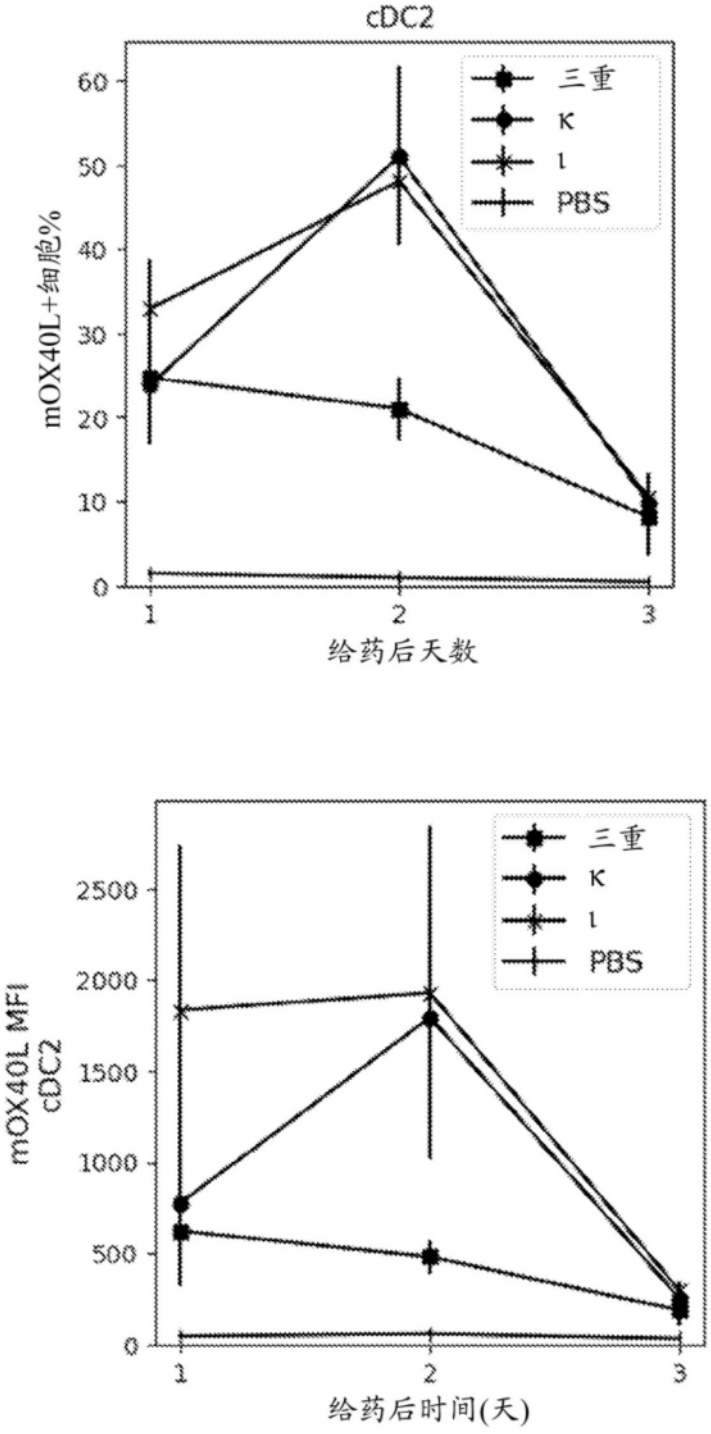


图2B

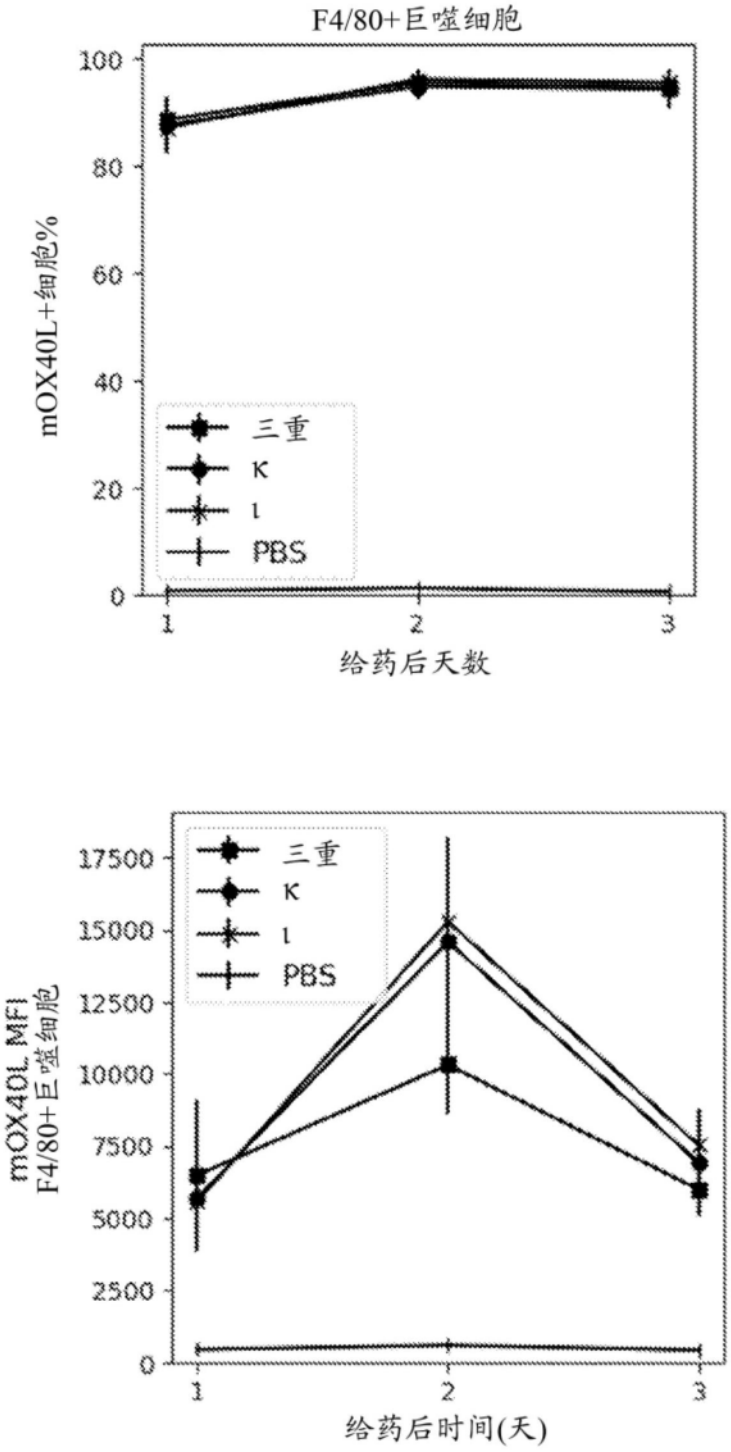

F4/80+巨噬细胞
mOX40L+细胞%
100
80
60
40
20
0
1
2
3
给药后天数
三重
κ
ι
PBS
mOX40L MFI
F4/80+巨噬细胞
17500
15000
12500
10000
7500
5000
2500
0
1
2
3
给药后时间(天)
三重
κ
ι
PBS

图2C

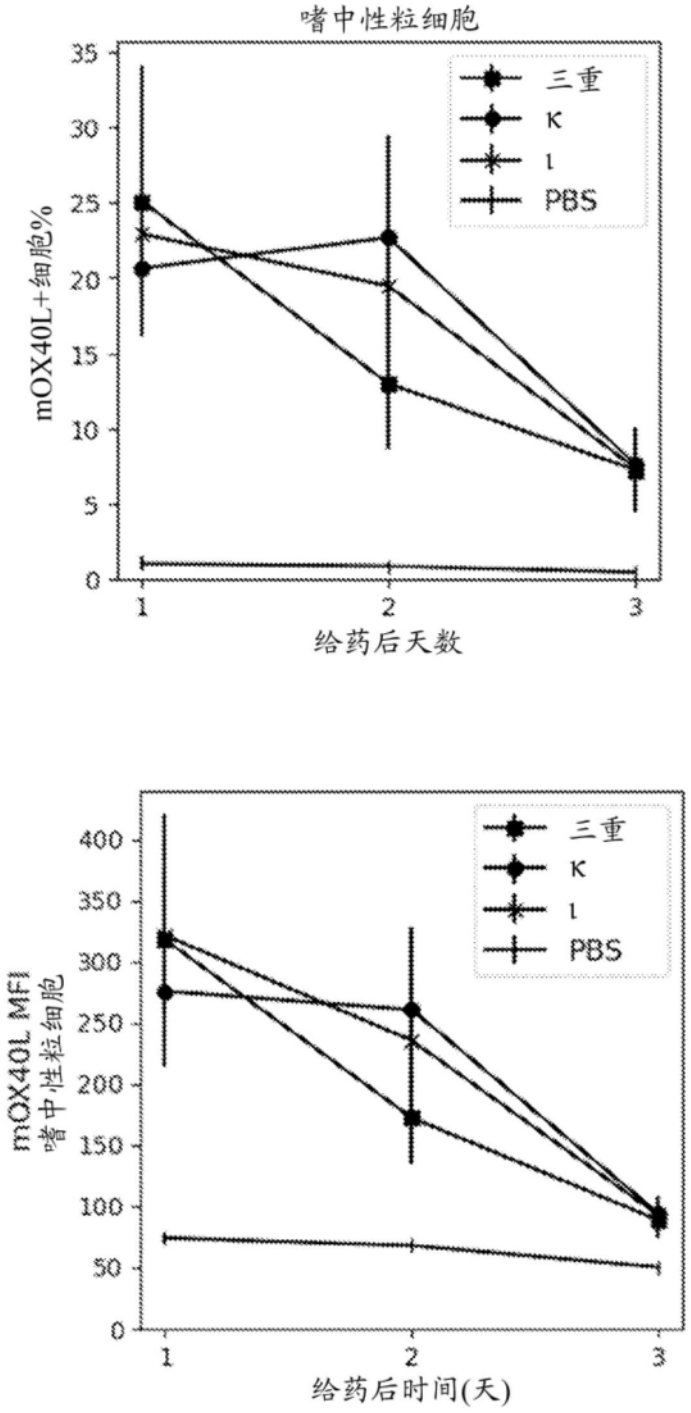


图2D

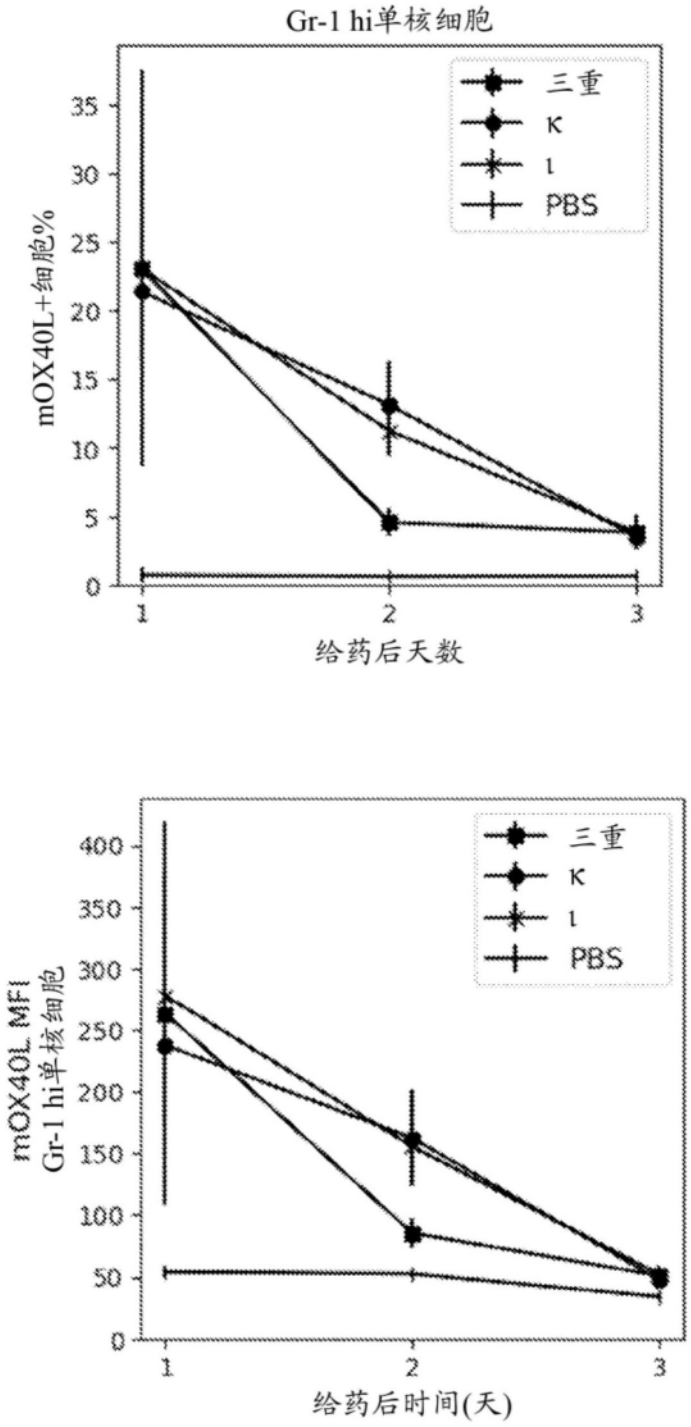

Gr-1 hi单核细胞
三重
κ
ι
PBS
mOX40L+细胞%
0
5
10
15
20
25
30
35
1
2
3
给药后天数
三重
κ
ι
PBS
mOX40L MFI
Gr-1 hi单核细胞
0
50
100
150
200
250
300
350
400
1
2
3
给药后时间(天)

图2E

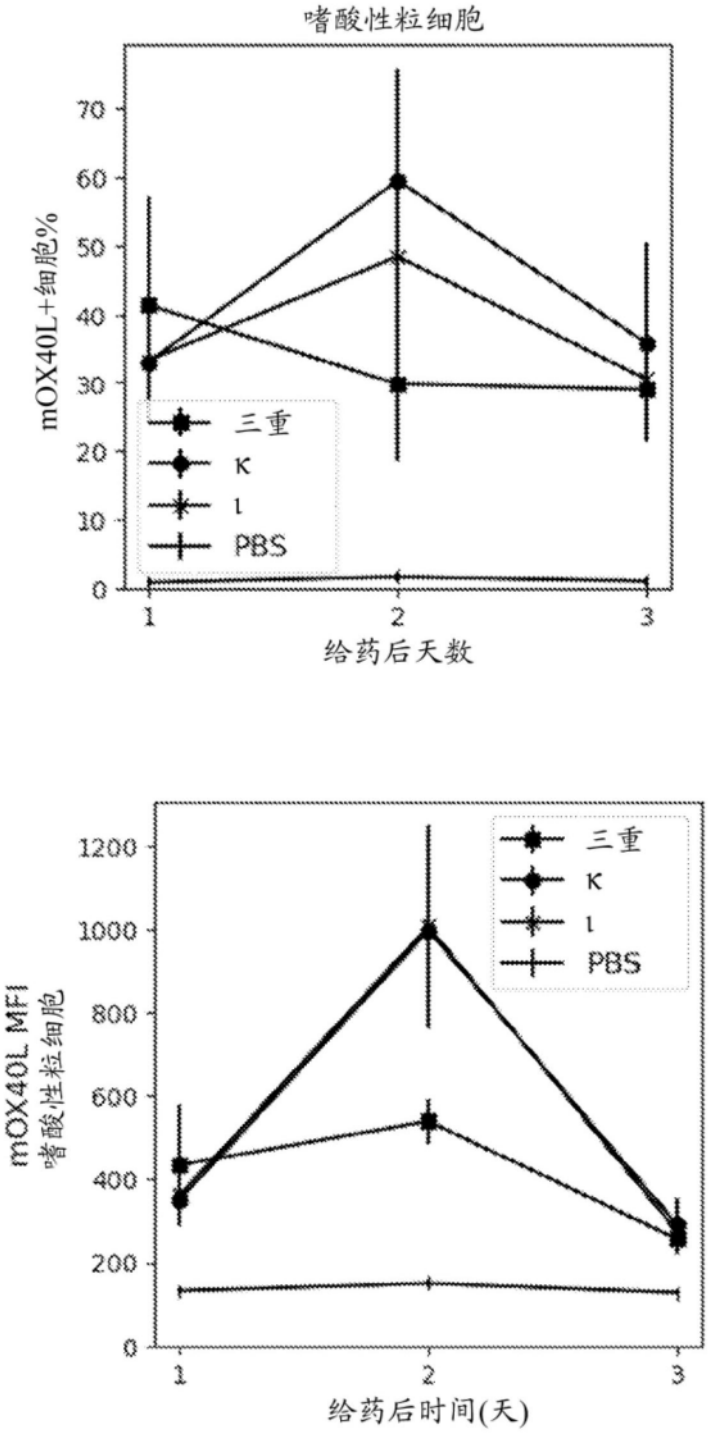

嗜酸性粒细胞
mOX40L+细胞%
70
60
50
40
30
20
10
0
1
2
3
给药后天数
三重
κ
λ
PBS
mOX40L MFI
嗜酸性粒细胞
1200
1000
800
600
400
200
0
1
2
3
给药后时间(天)
三重
κ
λ
PBS

图2F

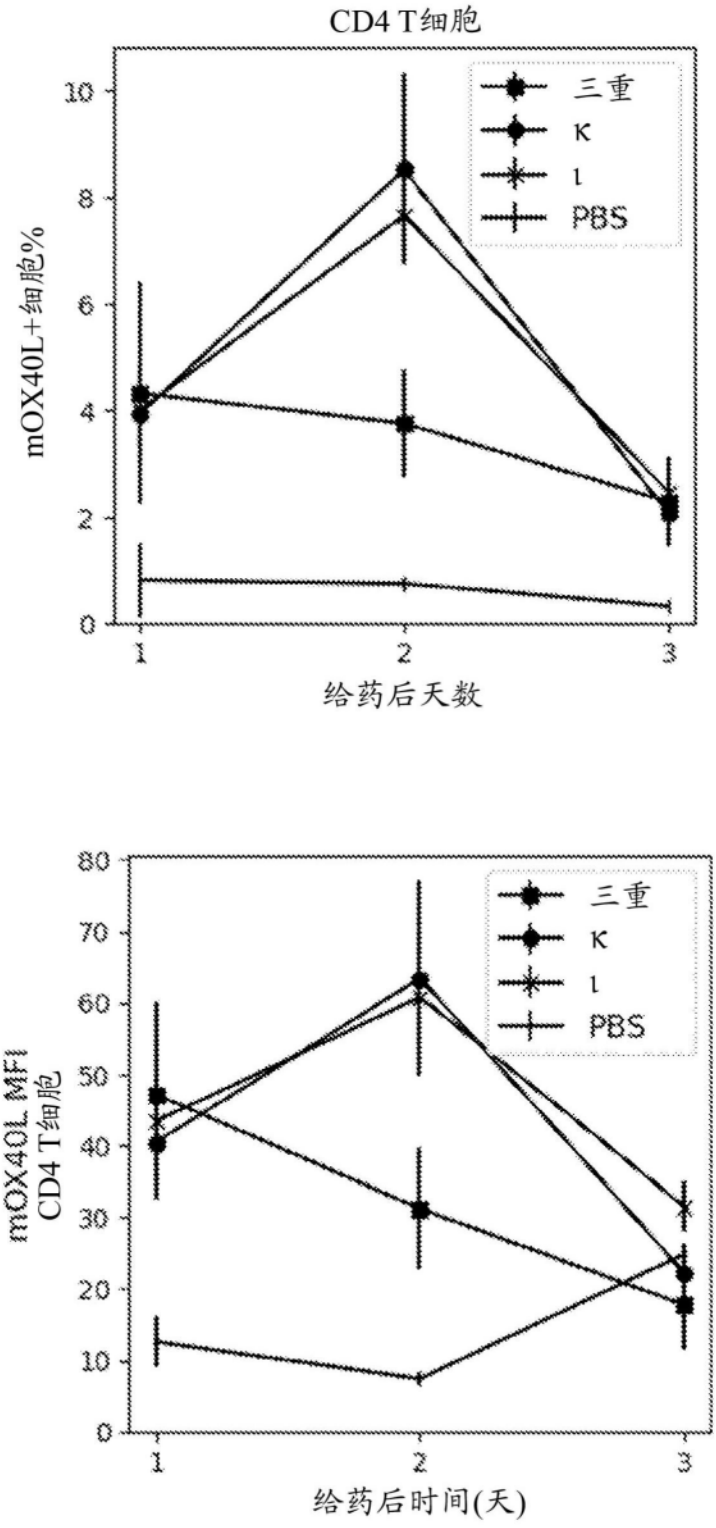

CD4 T细胞
三重
κ
ι
PBS
mOX40L+细胞%
0
2
4
6
8
10
1
2
3
给药后天数
三重
κ
ι
PBS
mOX40L MFI
CD4 T细胞
0
10
20
30
40
50
60
70
80
1
2
3
给药后时间(天)

图2G

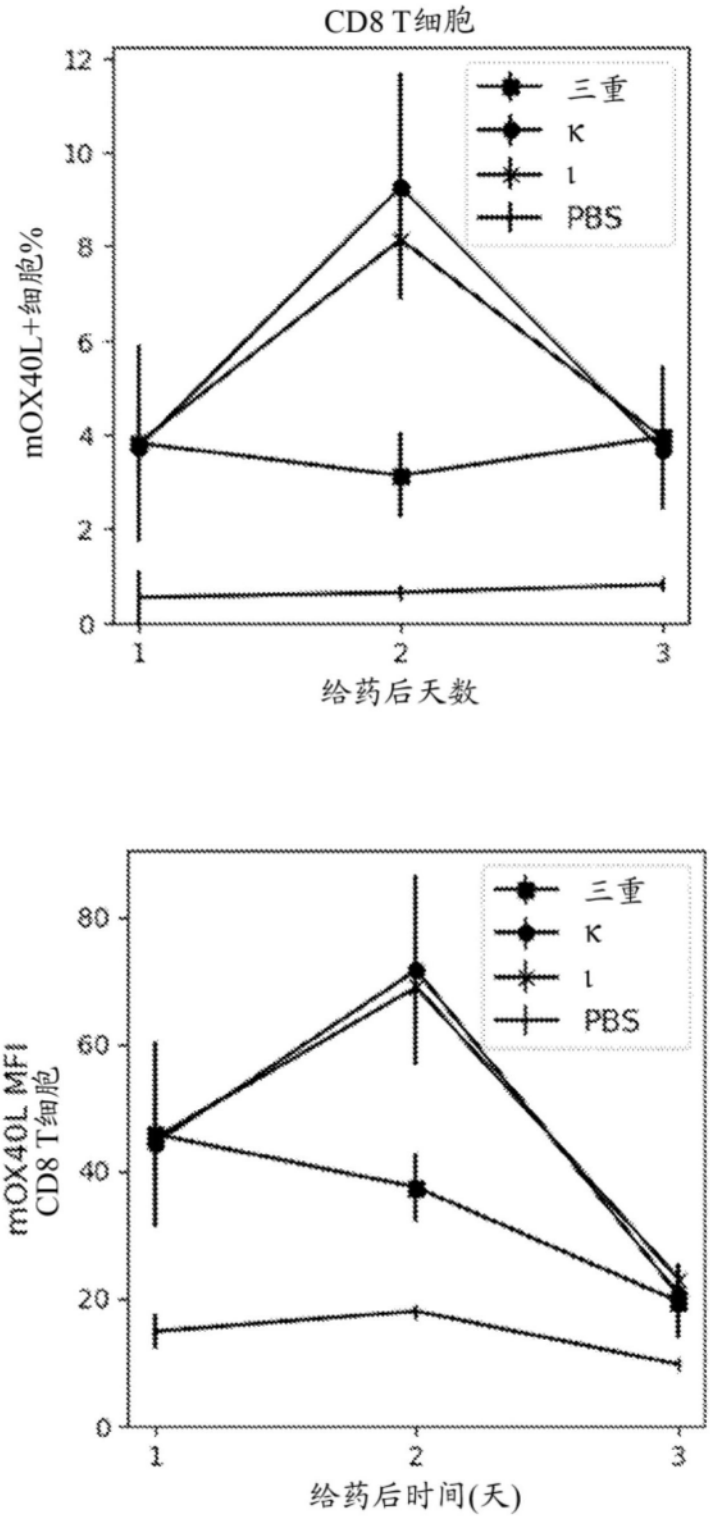


图2H

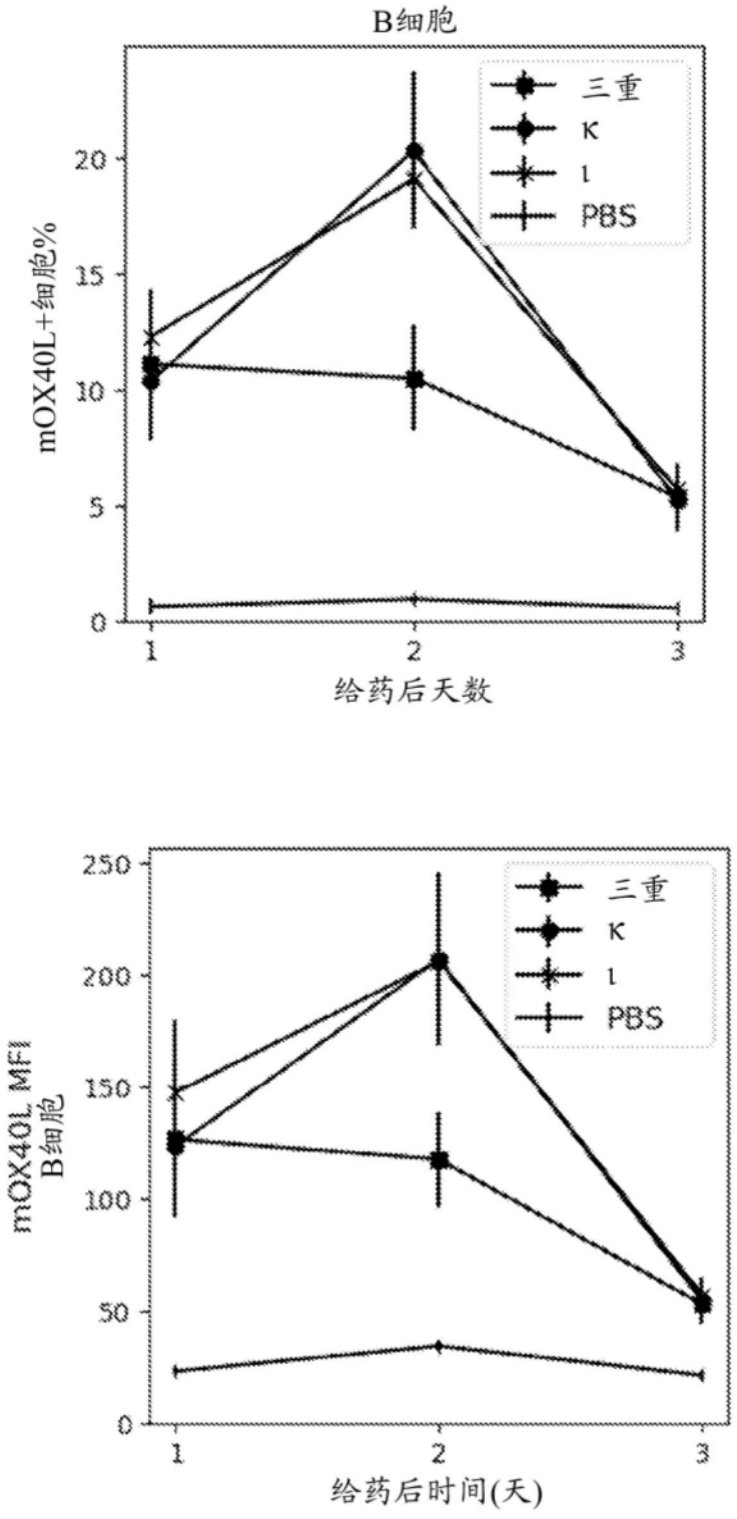

B细胞
三重
κ
λ
PBS
mOX40L+细胞%
0
5
10
15
20
1
2
3
给药后天数
三重
κ
λ
PBS
mOX40L MFI
B细胞
0
50
100
150
200
250
1
2
3
给药后时间(天)

图2I

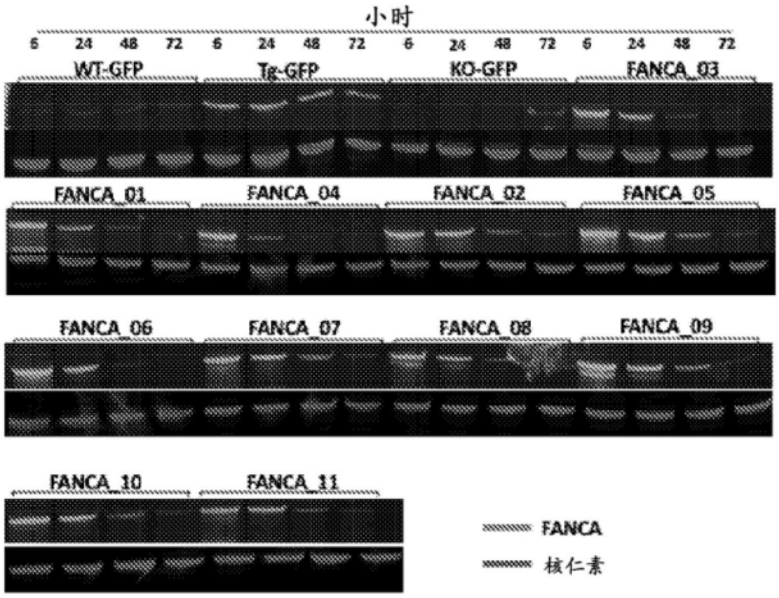

小时
6 24 48 72 6 24 48 72 6 24 48 72 6 24 48 72
WT-GFP
Tg-GFP
KO-GFP
FANCA_03
FANCA_01
FANCA_04
FANCA_02
FANCA_05
FANCA_06
FANCA_07
FANCA_08
FANCA_09
FANCA_10
FANCA_11
FANCA
核仁素

图3A

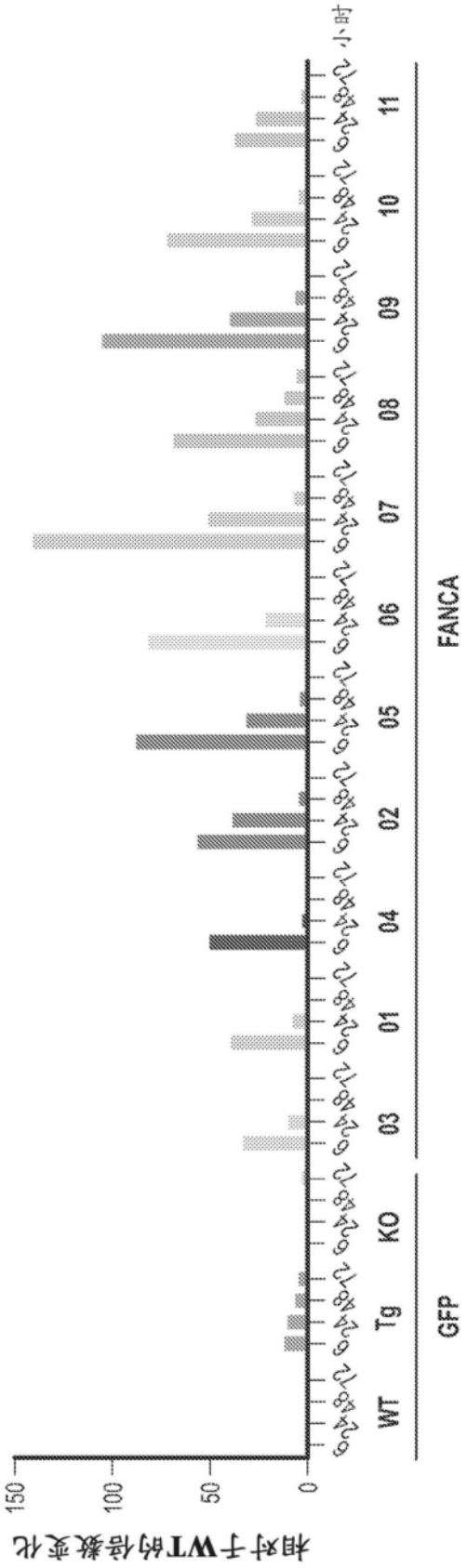

相对于WT的倍数变化
150
100
50
0
WT
Tg
KO
GFP
01
02
03
04
05
06
07
08
09
10
11
FANCA

图3B

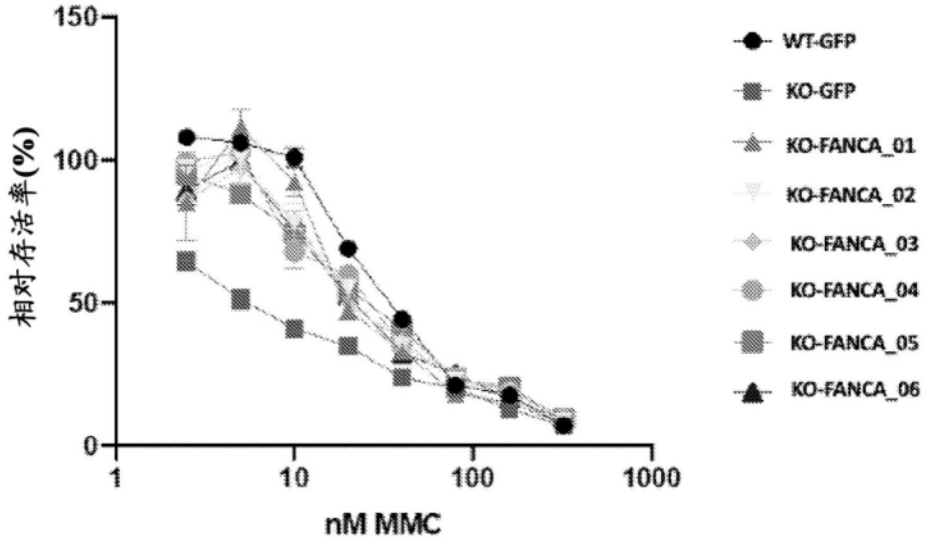

150
100
50
0
相对存活率(%)
1
10
100
1000
nM MMC
WT-GFP
KO-GFP
KO-FANCA_01
KO-FANCA_02
KO-FANCA_03
KO-FANCA_04
KO-FANCA_05
KO-FANCA_06

图4

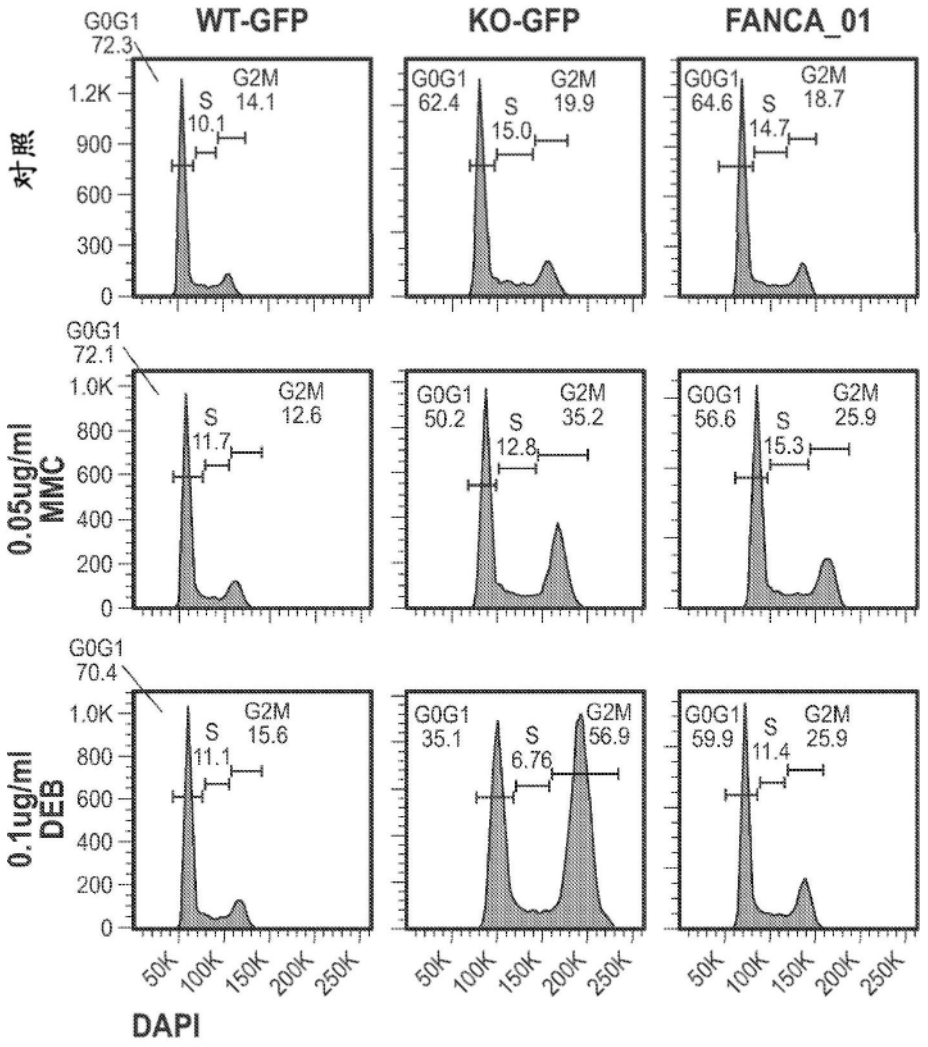

WT-GFP
KO-GFP
FANCA_01
对照
0.05ug/ml MMC
0.1ug/ml DEB
G0G1 72.3
S 10.1
G2M 14.1
G0G1 62.4
S 15.0
G2M 19.9
G0G1 64.6
S 14.7
G2M 18.7
G0G1 72.1
S 11.7
G2M 12.6
G0G1 50.2
S 12.8
G2M 35.2
G0G1 56.6
S 15.3
G2M 25.9
G0G1 70.4
S 11.1
G2M 15.6
G0G1 35.1
S 6.76
G2M 56.9
G0G1 59.9
S 11.4
G2M 25.9
1.2K
900
600
300
0
1.0K
800
600
400
200
0
50K
100K
150K
200K
250K
DAPI

图5A

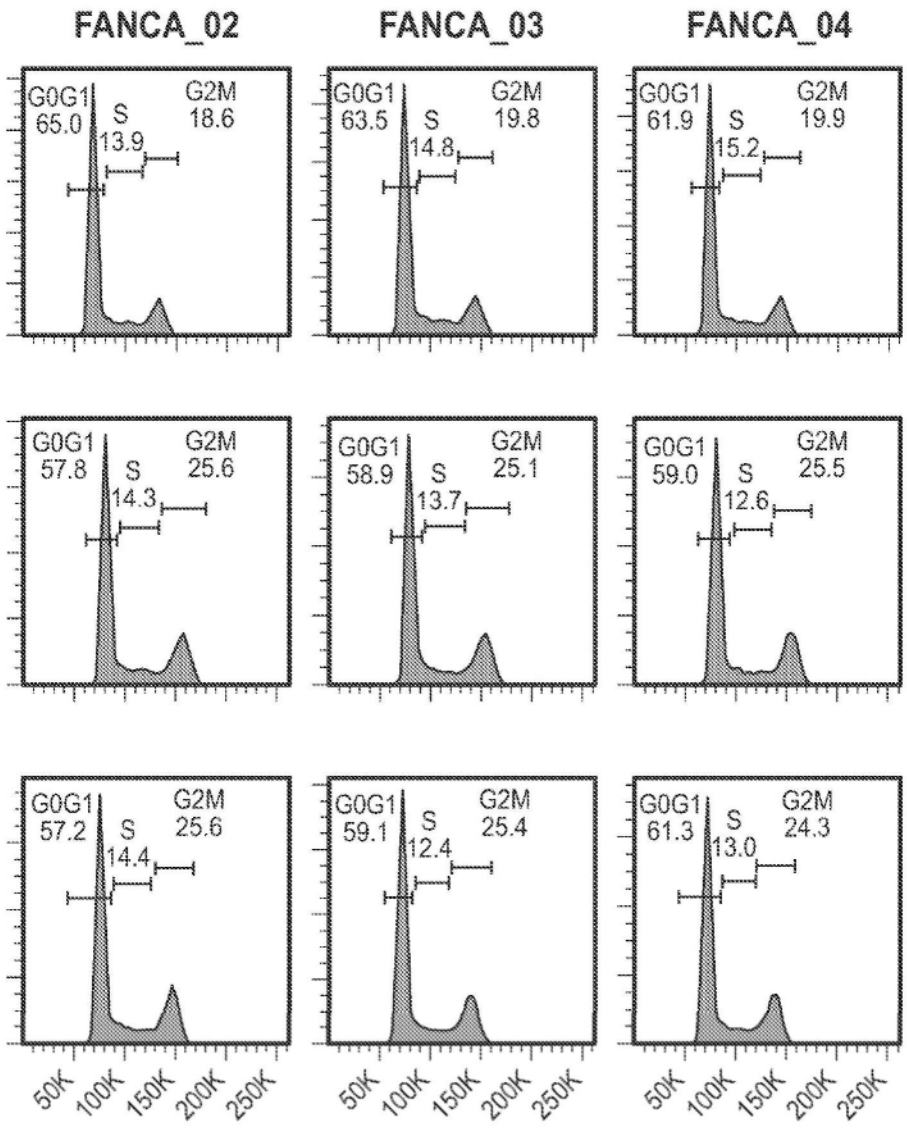

FANCA_02
FANCA_03
FANCA_04
G0G1
S
G2M
65.0
13.9
18.6
63.5
14.8
19.8
61.9
15.2
19.9
57.8
14.3
25.6
58.9
13.7
25.1
59.0
12.6
25.5
57.2
14.4
25.6
59.1
12.4
25.4
61.3
13.0
24.3
50K
100K
150K
200K
250K

图5A(续)

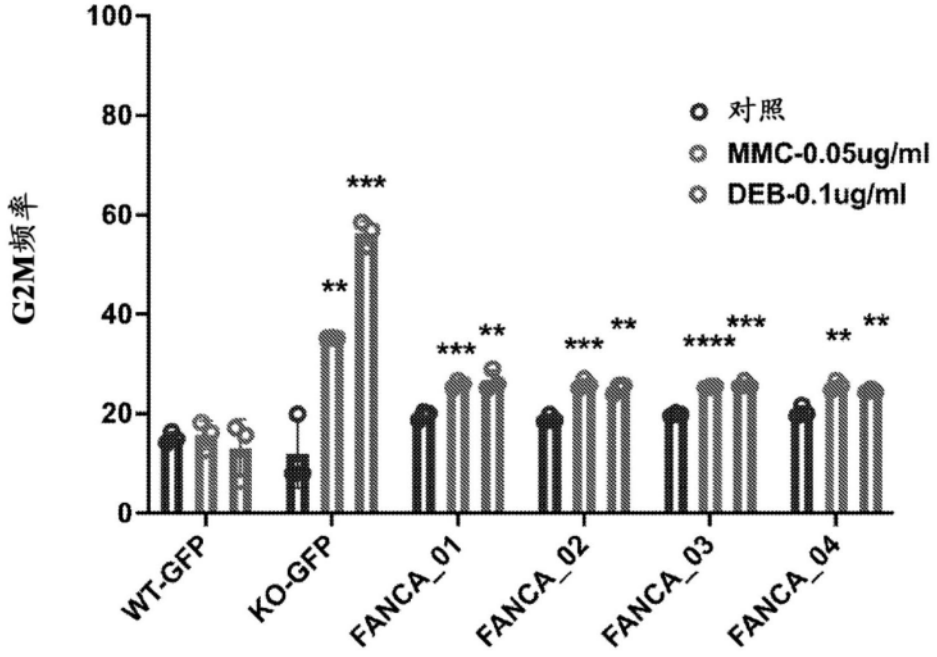

100
80
60
40
20
0
G2M频率
对照
MMC-0.05ug/ml
DEB-0.1ug/ml
**
***
***
**
***
**
****
***
**
**
WT-GFP
KO-GFP
FANCA_01
FANCA_02
FANCA_03
FANCA_04

图5B

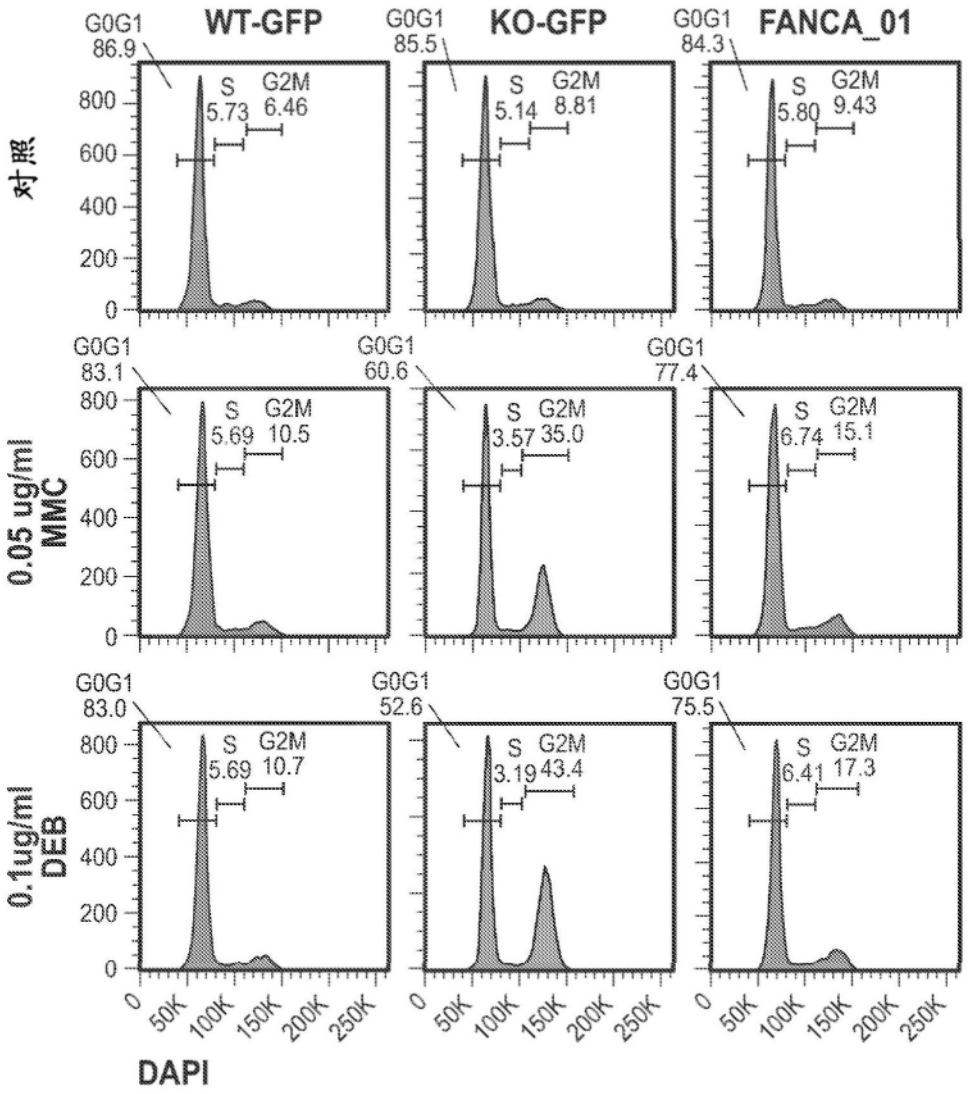

WT-GFP
KO-GFP
FANCA_01
对照
0.05 ug/ml MMC
0.1ug/ml DEB
G0G1
S
G2M
86.9
5.73
6.46
85.5
5.14
8.81
84.3
5.80
9.43
83.1
5.69
10.5
60.6
3.57
35.0
77.4
6.74
15.1
83.0
5.69
10.7
52.6
3.19
43.4
75.5
6.41
17.3
0
200
400
600
800
50K
100K
150K
200K
250K
DAPI

图6A

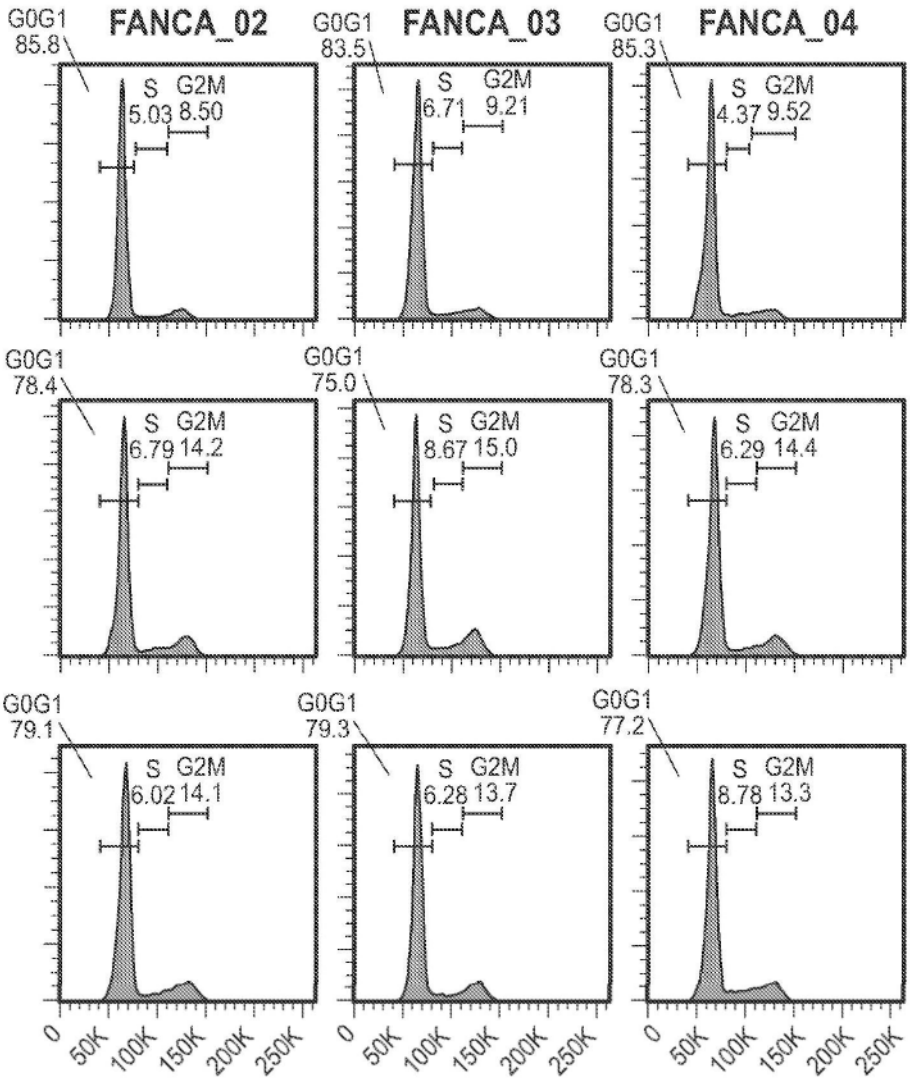

FANCA_02
G0G1
85.8
S G2M
5.03 8.50
FANCA_03
G0G1
83.5
S G2M
6.71 9.21
FANCA_04
G0G1
85.3
S G2M
4.37 9.52
G0G1
78.4
S G2M
6.79 14.2
G0G1
75.0
S G2M
8.67 15.0
G0G1
78.3
S G2M
6.29 14.4
G0G1
79.1
S G2M
6.02 14.1
G0G1
79.3
S G2M
6.28 13.7
G0G1
77.2
S G2M
8.78 13.3
0
50K
100K
150K
200K
250K

图6A(续)

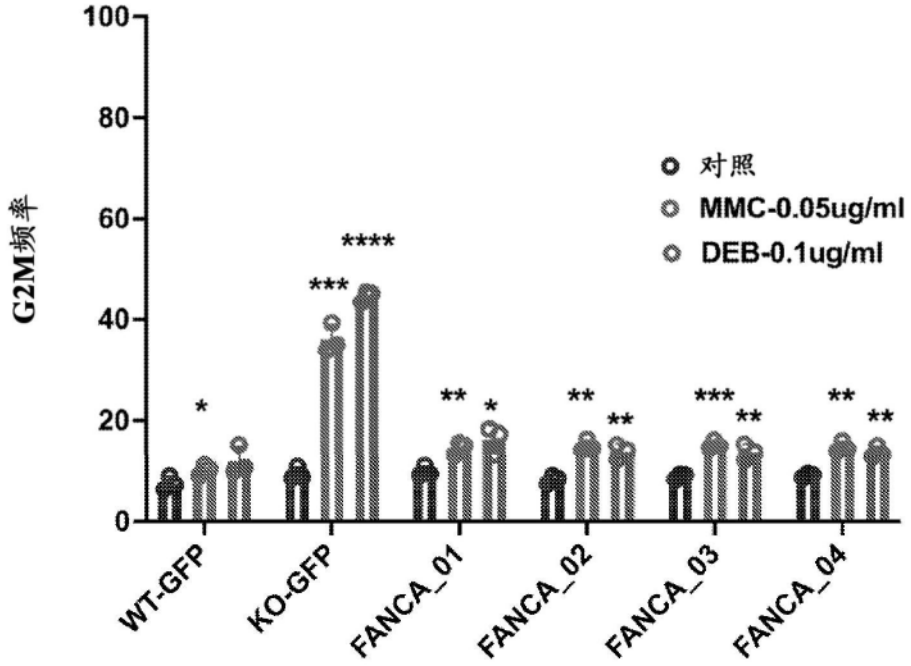

G2M频率
100
80
60
40
20
0
对照
MMC-0.05ug/ml
DEB-0.1ug/ml
*
***
****
**
*
**
**
***
**
**
**
WT-GFP
KO-GFP
FANCA_01
FANCA_02
FANCA_03
FANCA_04

图6B

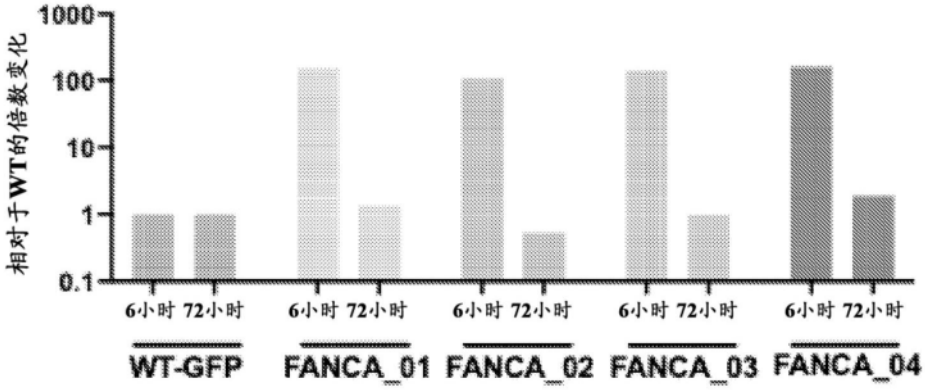

相对于WT的倍数变化
1000
100
10
1
0.1
6小时 72小时
6小时 72小时
6小时 72小时
6小时 72小时
6小时 72小时
WT-GFP
FANCA_01
FANCA_02
FANCA_03
FANCA_04

图6C

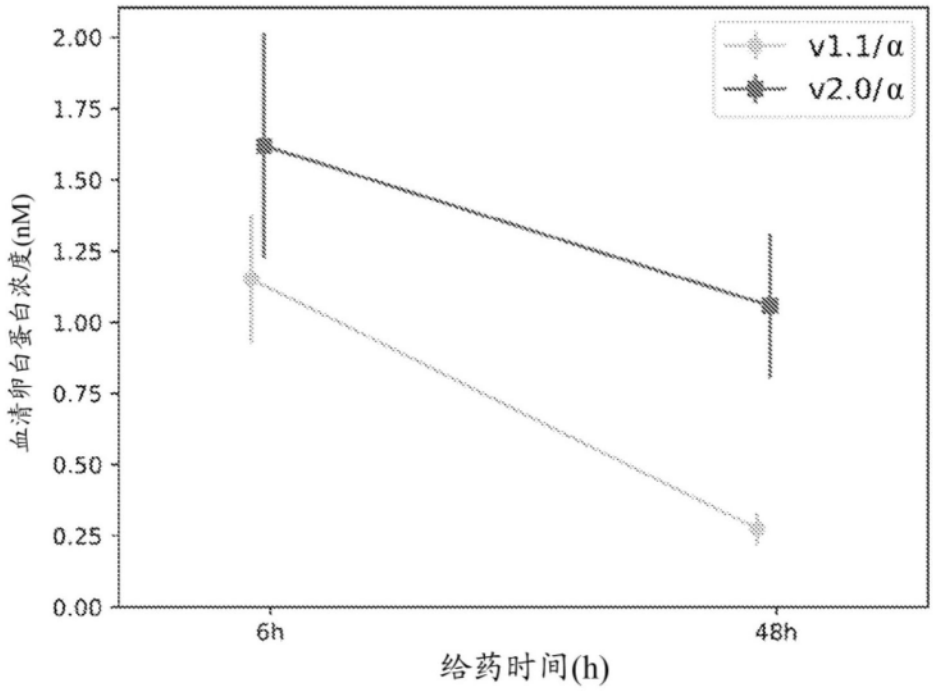

v1.1/α
v2.0/α
血清卵白蛋白浓度(nM)
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
6h
48h
给药时间(h)

图7

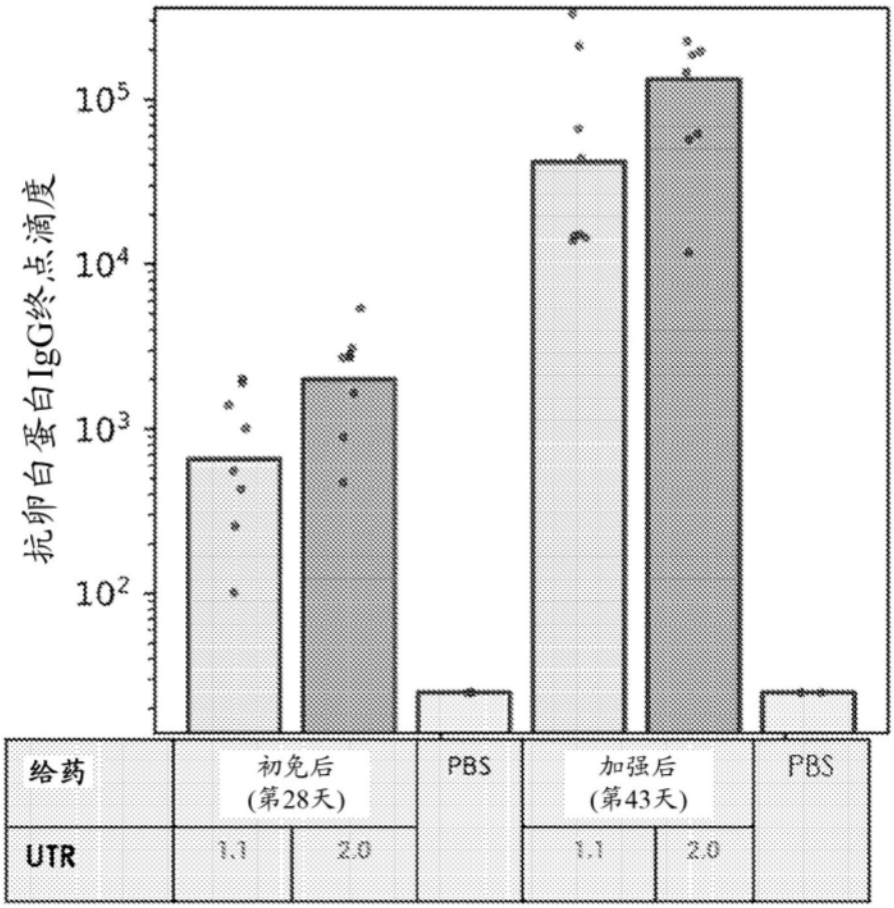

抗卵白蛋白IgG终点滴度
10^5
10^4
10^3
10^2
给药
初免后
(第28天)
PBS
加强后
(第43天)
PBS
UTR
1.1
2.0
1.1
2.0

图8

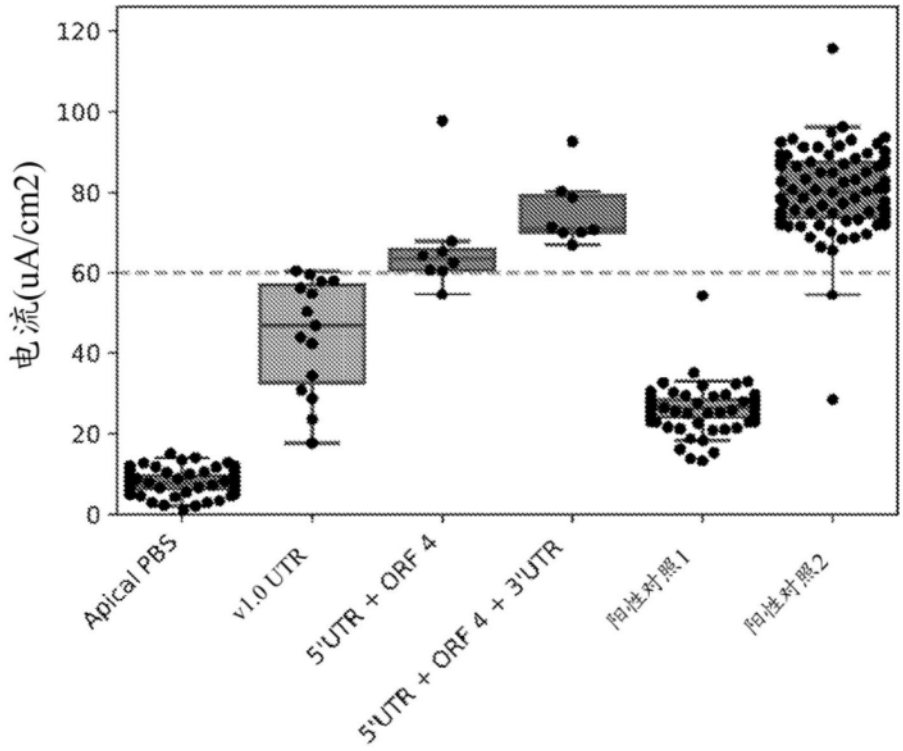

120
100
80
60
40
20
0
电流(uA/cm2)
Apical PBS
v1.0 UTR
5'UTR + ORF 4
5'UTR + ORF 4 + 3'UTR
阳性对照1
阳性对照2

图9

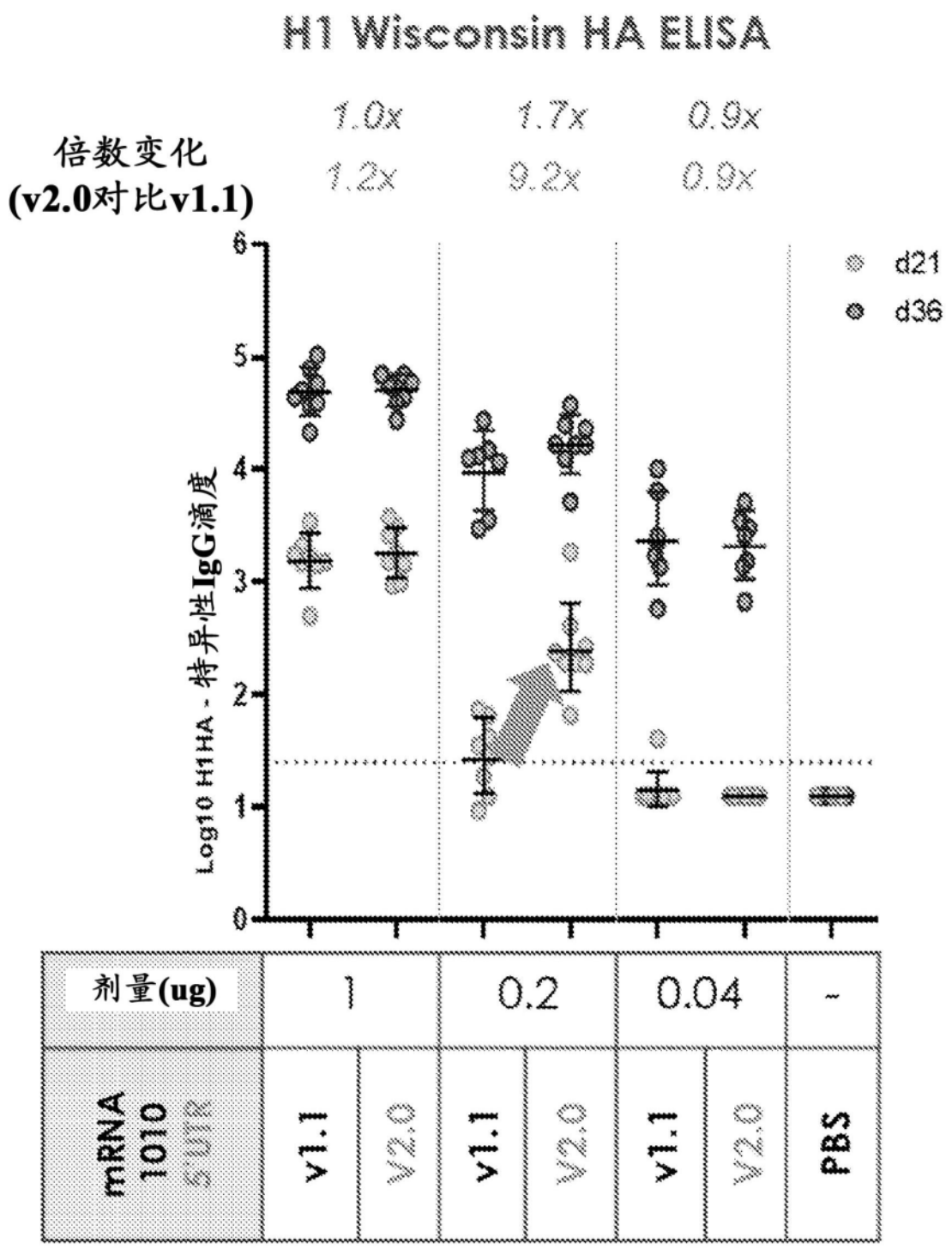

H1 Wisconsin HA ELISA
倍数变化
(v2.0对比v1.1)
1.0x
1.2x
1.7x
9.2x
0.9x
0.9x
d21
d36
Log10 H1HA - 特异性IgG滴度
0
1
2
3
4
5
6
剂量(ug)
1
0.2
0.04
-
mRNA 1010 5'UTR
v1.1
V2.0
v1.1
V2.0
v1.1
V2.0
PBS

图10

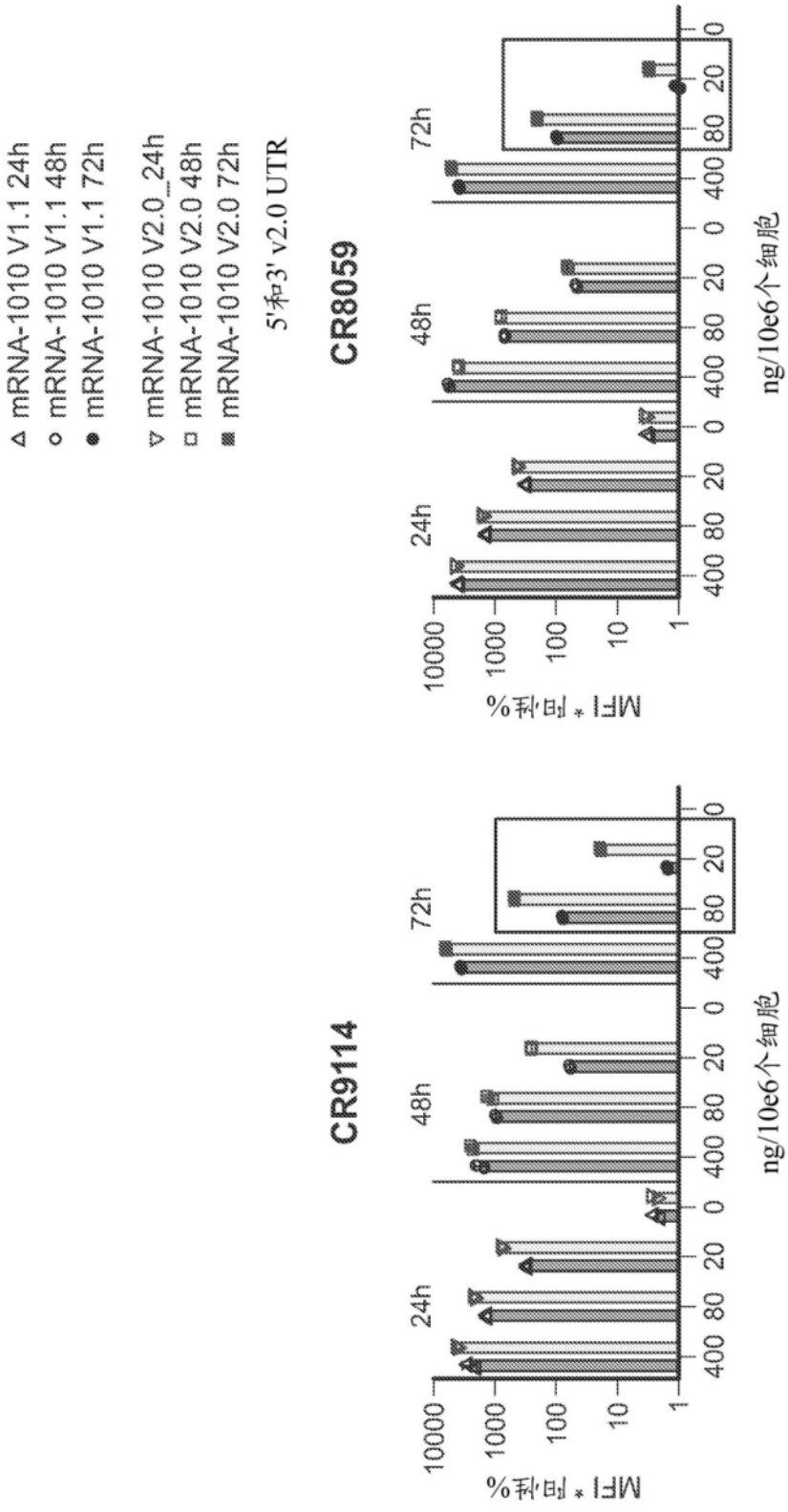

△ mRNA-1010 V1.1 24h
○ mRNA-1010 V1.1 48h
● mRNA-1010 V1.1 72h
▽ mRNA-1010 V2.0_24h
□ mRNA-1010 V2.0 48h
■ mRNA-1010 V2.0 72h
5'和3' v2.0 UTR
CR8059
CR9114
MFI * 阳性%
10000
1000
100
10
1
24h
48h
72h
400 80 20 0 400 80 20 0 400 80 20 0
ng/10e6个细胞

图11

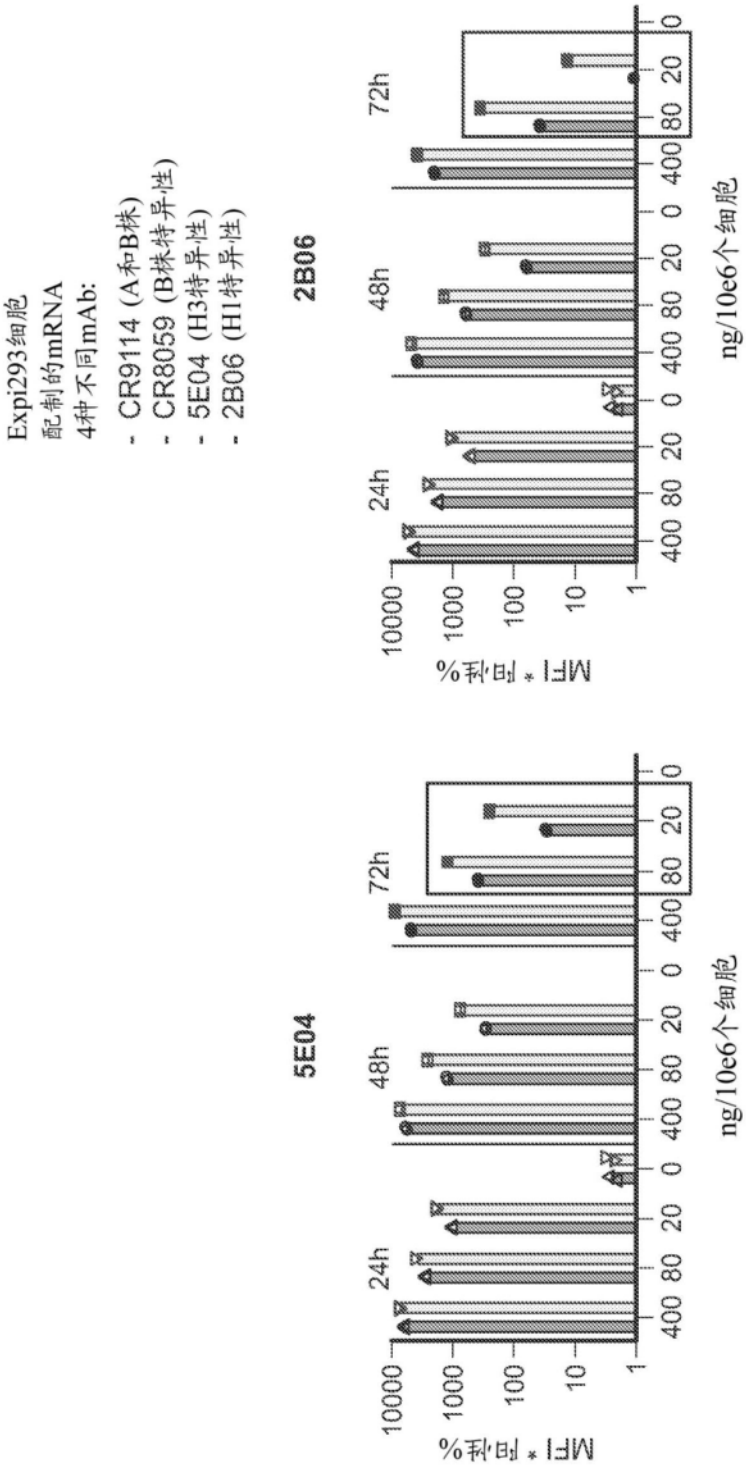

Expi293细胞
配制的mRNA
4种不同mAb:
- CR9114 (A和B株)
- CR8059 (B株特异性)
- 5E04 (H3特异性)
- 2B06 (H1特异性)
2B06
5E04
MFI * 阳性%
ng/10e6个细胞
24h
48h
72h
10000
1000
100
10
1
400 80 20 0 400 80 20 0

图11(续)