(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 115003697 A
(43)申请公布日 2022.09.02

(21)申请号 202080081341.X
(22)申请日 2020.09.23
(30)优先权数据
62/904,595 2019.09.23 US
62/935,590 2019.11.14 US
62/945,752 2019.12.09 US
(85)PCT国际申请进入国家阶段日
2022.05.23
(86)PCT国际申请的申请数据
PCT/US2020/052291 2020.09.23
(87)PCT国际申请的公布数据
WO2021/061829 EN 2021.04.01
(71)申请人 特韦斯特生物科学公司
地址 美国加利福尼亚州
(72)发明人 雷·塔比比亚扎 亚伦·萨托 潘卡杰·加格 刘强
(74)专利代理机构 北京安信方达知识产权代理有限公司 11262
专利代理师 贺淑东 武晶晶
(51)Int.Cl.
C07K 16/28(2006.01)
C40B 40/08(2006.01)

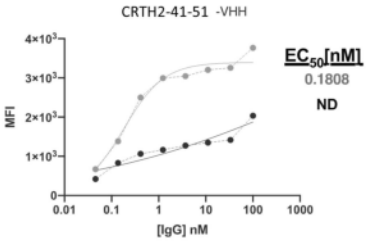
权利要求书5页 说明书84页 附图47页

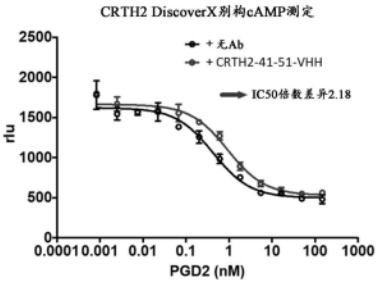
(54)发明名称

CRTH2的变异核酸文库

(57)摘要

本文提供了与具有编码包含CRTH2R结合域的支架的核酸的前列腺素D2受体2(DP2或CRTH2R)文库相关的方法和组合物。本文所述的CRTH2R文库编码包括抗体和单域抗体的免疫球蛋白。本文所述的文库包括斑驳化的文库,其包含各自编码至少一个预定参考核酸序列的预定变体的核酸。本文进一步描述了当翻译所述核酸文库时生成的蛋白质文库。本文进一步描述了表达本文所述的斑驳化核酸文库的细胞文库。





CN 115003697 A

1.一种核酸文库,其包含:多个核酸,其中每个所述核酸编码当翻译时编码CRTH2R结合免疫球蛋白的序列,其中所述CRTH2R结合免疫球蛋白包含CRTH2R结合域的变体,其中所述CRTH2R结合域是CRTH2R的配体,并且其中所述核酸文库包含至少10,000个变异免疫球蛋白重链和至少10,000个变异免疫球蛋白轻链。

2.根据权利要求1所述的核酸文库,其中所述核酸文库包含至少50,000个变异免疫球蛋白重链和至少50,000个变异免疫球蛋白轻链。

3.根据权利要求1所述的核酸文库,其中所述核酸文库包含至少100,000个变异免疫球蛋白重链和至少100,000个变异免疫球蛋白轻链。

4.根据权利要求1所述的核酸文库,其中所述核酸文库包含至少10^5个不相同的核酸。

5.根据权利要求1所述的核酸文库,其中所述免疫球蛋白重链当翻译时的长度为约90至约100个氨基酸。

6.根据权利要求1所述的核酸文库,其中所述免疫球蛋白重链当翻译时的长度为约100至约400个氨基酸。

7.根据权利要求1所述的核酸文库,其中所述变异免疫球蛋白重链当翻译时与SEQ ID NO:2338-SEQ ID NO:2360或SEQ ID NO:2403-SEQ ID NO:2405中的任一个具有至少80%的序列同一性。

8.根据权利要求1所述的核酸文库,其中所述变异免疫球蛋白轻链当翻译时与SEQ ID NO:2361-SEQ ID NO:2381或SEQ ID NO:2406-SEQ ID NO:2408中的任一个具有至少80%的序列同一性。

9.一种包含多个核酸的核酸文库,其中所述多个核酸中的每个核酸编码当翻译时编码抗体或其抗体片段的序列,其中所述抗体或其抗体片段包含重链可变区(VH),所述重链可变区包含CRTH2R结合域,其中所述多个核酸中的每个核酸包含编码所述CRTH2R结合域的序列变体的序列,并且其中所述抗体或抗体片段以小于100nM的K_D与其抗原结合。

10.根据权利要求9所述的核酸文库,其中所述VH的长度为约90至约100个氨基酸。

11.根据权利要求9所述的核酸文库,其中所述VH的长度为约100至约400个氨基酸。

12.根据权利要求9所述的核酸文库,其中所述VH的长度为约270至约300个碱基对。

13.根据权利要求9所述的核酸文库,其中所述VH的长度为约300至约1200个碱基对。

14.根据权利要求9所述的核酸文库,其中所述文库包含至少10^5个不同的核酸。

15.一种核酸文库,其包含:多个核酸,其中每个所述核酸编码当翻译时编码CRTH2R单域抗体的序列,其中所述多个序列中的每个序列包含编码重链可变区(VH)上的CDR1、CDR2或CDR3的变异序列;其中所述文库包含至少30,000个变异序列;并且其中所述CRTH2R单域抗体以小于100nM的K_D与其抗原结合。

16.根据权利要求15所述的核酸文库,其中当翻译时所述VH的长度为约90至约100个氨基酸。

17.根据权利要求15所述的核酸文库,其中当翻译时所述VH的长度为约100至约400个氨基酸。

18.根据权利要求15所述的核酸文库,其中所述VH的长度为约270至约300个碱基对。

19.根据权利要求15所述的核酸文库,其中所述VH的长度为约300至约1200个碱基对。

20.根据权利要求15所述的核酸文库,其中当翻译时所述VH包括与SEQ ID NO:2338-

SEQ ID NO:2360或SEQ ID NO:2403-SEQ ID NO:2405中的任一个具有至少80%的序列同一性。

21.一种结合CRTH2R的抗体或抗体片段,其包含免疫球蛋白重链和免疫球蛋白轻链:

a.其中所述免疫球蛋白重链包含与SEQ ID NO:2338-SEQ ID NO:2360或SEQ ID NO:2403-SEQ ID NO:2405中的任一个所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且

b.其中所述免疫球蛋白轻链包含与SEQ ID NO:2361-SEQ ID NO:2381或SEQ ID NO:2406-SEQ ID NO:2408中的任一个所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

22.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2338所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2361所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

23.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2339所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2362所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

24.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2340所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2363所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

25.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2341所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2364所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

26.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2342所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2365所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

27.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2343所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2366所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

28.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2344所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2367中所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

29.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2345中所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2368中所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

30.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2346所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2369中所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

31.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2347中所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球

蛋白轻链包含与SEQ ID NO:2370中所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

32.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2348所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2371所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

33.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2349所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2372中所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

34.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2350中所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2373中所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

35.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2351所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2374中所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

36.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2352中所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2375中所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

37.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2353所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2376中所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

38.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2354所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中免疫球蛋白轻链包含与SEQ ID NO:2377中所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

39.如权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2355中所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2378中所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

40.根据权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2356中所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2379中所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

41.如权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2357中所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2380中所示的氨基酸序列具有至少约90%同一性的氨基酸序列。

42.如权利要求21所述的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2358中所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球

蛋白轻链包含与SEQ ID NO:2381中所示的氨基酸序列具有至少约90％同一性的氨基酸序列。

43.根据权利要求21所述的抗体或抗体片段，其中所述免疫球蛋白重链包含与SEQ ID NO:2403所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2406中所示的氨基酸序列具有至少约90％同一性的氨基酸序列。

44.根据权利要求21所述的抗体或抗体片段，其中所述免疫球蛋白重链包含与SEQ ID NO:2404所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2407所示的氨基酸序列具有至少约90％同一性的氨基酸序列。

45.根据权利要求21所述的抗体或抗体片段，其中所述免疫球蛋白重链包含与SEQ ID NO:2405所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2408所示的氨基酸序列具有至少约90％同一性的氨基酸序列。

46.根据权利要求21所述的抗体或抗体片段，其中所述抗体是单克隆抗体、多克隆抗体、双特异性抗体、多特异性抗体、移植抗体、人抗体、人源化抗体、合成抗体、嵌合抗体、骆驼化抗体、单链Fv(scFv)、单链抗体、Fab片段、F(ab')2片段、Fd片段、Fv片段、单域抗体、分离的互补决定区(CDR)、双抗体、仅由单个单体可变域组成的片段、二硫键连接的Fv(sdFv)、胞内抗体、抗独特型(抗Id)抗体，或其ab抗原结合片段。

47.根据权利要求21所述的抗体或抗体片段，其中所述抗体或其抗体片段是嵌合的或人源化的。

48.根据权利要求21所述的抗体或抗体片段，其中所述抗体在cAMP测定中的EC50小于约25纳摩尔。

49.根据权利要求21所述的抗体或抗体片段，其中所述抗体在cAMP测定中的EC50小于约20纳摩尔。

50.根据权利要求21所述的抗体或抗体片段，其中所述抗体在cAMP测定中的EC50小于约10纳摩尔。

51.一种抗体或抗体片段，其中所述抗体或抗体片段包含互补决定区(CDR)，所述互补决定区(CDR)包含与SEQ ID NO:2382-SEQ ID NO:2402中的任一个所示的氨基酸序列具有至少约90％同一性的氨基酸序列。

52.一种抗体或抗体片段，其中所述抗体或抗体片段包含SEQ ID NO:2382-SEQ ID NO:2402中的任一个的序列并且其中所述抗体是单克隆抗体、多克隆抗体、双特异性抗体、多特异性抗体、移植抗体、人抗体、人源化抗体、合成抗体、嵌合抗体、骆驼化抗体、单链Fv(scFv)、单链抗体、Fab片段、F(ab')2片段、Fd片段、Fv片段、单域抗体、分离的互补决定区(CDR)、双抗体、仅由单个单体可变域组成的片段、二硫键连接的Fv(sdFv)、胞内抗体、抗独特型(抗Id)抗体，或其ab抗原结合片段。

53.一种治疗中枢神经系统、肾、肠、肺、毛发、皮肤、骨或软骨的疾病或病症的方法，其包括施用权利要求21-52中任一项的抗体或抗体片段。

54.一种治疗以炎症反应为特征的疾病或病症的方法，其包括施用权利要求21-52中任一项的抗体或抗体片段。

55.一种治疗过敏反应的方法，其包括施用权利要求21-52中任一项的抗体或抗体片段。

56.根据权利要求55所述的方法，其中所述过敏反应是慢性特发性荨麻疹或过敏性鼻炎。

57.一种治疗哮喘的方法，其包括施用权利要求21-52中任一项的抗体或抗体片段。

58.一种治疗脱发或秃顶的方法，其包括施用权利要求21-52中任一项的抗体或抗体片段。

59.一种用于生成编码CRTH2R抗体或其抗体片段的核酸文库的方法，其包括：

(a)提供编码以下i-iii的预定序列：

i.第一多个多核苷酸，其中所述第一多个多核苷酸中的每个多核苷酸编码至少1000个编码重链上的CDR1的变异序列；

ii.第二多个多核苷酸，其中所述第二多个多核苷酸中的每个多核苷酸编码至少1000个编码重链上的CDR2的变异序列；

iii.第三多个多核苷酸，其中所述第三多个多核苷酸中的每个多核苷酸编码至少1000个编码重链上的CDR3的变异序列；和

(b)将所述第一多个多核苷酸、所述第二多个多核苷酸和所述第三多个多核苷酸混合以形成编码所述CRTH2R抗体或其抗体片段的变异核酸的所述核酸文库，并且其中至少约70％的所述变异核酸编码以小于100nM的K_D与其抗原结合的抗体或抗体片段。

60.根据权利要求59所述的方法，其中所述CRTH2R抗体或其抗体片段是单域抗体。

61.根据权利要求60所述的方法，其中所述单域抗体包含一个重链可变域。

62.根据权利要求60所述的方法，其中所述单域抗体是VHH抗体。

63.根据权利要求59所述的方法，其中所述核酸文库包含至少50,000个变异序列。

64.根据权利要求59所述的方法，其中所述核酸文库包含至少100,000个变异序列。

65.根据权利要求59所述的方法，其中所述核酸文库包含至少10^5个不同的核酸。

66.根据权利要求59所述的方法，其中所述核酸文库包含至少一个编码CRTH2R抗体或抗体片段的序列，所述CRTH2R抗体或抗体片段以小于75nM的K_D与CRTH2R结合。

67.如权利要求59所述的方法，其中所述核酸文库包含至少一个编码CRTH2R抗体或抗体片段的序列，所述CRTH2R抗体或抗体片段以小于50nM的K_D与CRTH2R结合。

68.根据权利要求59所述的方法，其中所述核酸文库包含至少一个编码CRTH2R抗体或抗体片段的序列，所述CRTH2R抗体或抗体片段以小于10nM的K_D与CRTH2R结合。

69.根据权利要求68所述的方法，其中所述核酸文库包含至少500个变异序列。

70.根据权利要求59所述的方法，其中所述核酸文库包含至少五个编码CRTH2R抗体或抗体片段的序列，所述CRTH2R抗体或抗体片段以小于75nM的K_D与CRTH2R结合。

71.由权利要求1-20中任一项的核酸文库编码的蛋白质文库，其中所述蛋白质文库包含肽。

72.根据权利要求71所述的蛋白质文库，其中所述蛋白质文库包含免疫球蛋白。

73.根据权利要求71所述的蛋白质文库，其中所述蛋白质文库包含抗体。

74.根据权利要求71所述的蛋白质文库，其中所述蛋白质文库是拟肽文库。

75.一种载体文库，其包含权利要求1-20中任一项的核酸文库。

76.一种细胞文库，其包含权利要求1-20中任一项的核酸文库。

77.一种细胞文库，其包含权利要求71-74中任一项的蛋白质文库。

CRTH2的变异核酸文库

交叉引用

[0001] 本申请要求2019年9月23日提交的美国临时专利申请号62/904,595;2019年11月14日提交的美国临时专利申请号62/935,590;以及2019年12月9日提交的美国临时专利申请号62/945,752的权益,其每个均通过引用全文并入。

背景技术

[0002] G蛋白偶联受体(GPCR),如前列腺素D2受体2(DP2或CRTH2R)与多种疾病有关。因为GPCR通常在细胞中低水平表达,并且在纯化时非常不稳定,故而在获得合适的抗原中存在问题,因此难以产生针对GPCR的抗体。因此,需要靶向GPCR的用于治疗性干预的改良药剂。

援引并入

[0003] 本说明书中所提及的所有出版物、专利和专利申请均通过引用并入本文,其程度犹如具体地且单独地指出每个单独的出版物、专利或专利申请均通过引用而并入。

发明内容

[0004] 本文提供了核酸文库,包含:多个核酸,其中每个所述核酸编码当翻译时编码CRTH2R结合免疫球蛋白的序列,其中CRTH2R结合免疫球蛋白包含CRTH2R结合域的变体,其中CRTH2R结合域是CRTH2R的配体,并且其中所述核酸文库包含至少10,000个变异免疫球蛋白重链和至少10,000个变异免疫球蛋白轻链。本文进一步提供了核酸文库,其中所述核酸文库包含至少50,000个变异免疫球蛋白重链和至少50,000个变异免疫球蛋白轻链。本文进一步提供了核酸文库,其中所述核酸文库包含至少100,000个变异免疫球蛋白重链和至少100,000个变异免疫球蛋白轻链。本文进一步提供了核酸文库,其中所述核酸文库包含至少10^5个不相同的核酸。本文进一步提供了核酸文库,其中所述免疫球蛋白重链当翻译时的长度为约90至约100个氨基酸。本文进一步提供了核酸文库,其中所述免疫球蛋白重链当翻译时的长度为约100至约400个氨基酸。本文进一步提供了核酸文库,其中所述变异免疫球蛋白重链当翻译时与SEQ ID NO:2338-SEQ ID NO:2360或SEQ ID NO:2403-SEQ ID NO:2405中的任一个具有至少80%的序列同一性。本文进一步提供了核酸文库,其中所述变异免疫球蛋白轻链当翻译时与SEQ ID NO:2361-SEQ ID NO:2381或SEQ ID NO:2406-SEQ ID NO:2408中的任一个具有至少80%的序列同一性。

[0005] 本文提供了包含多个核酸的核酸文库,其中所述多个核酸中的每个核酸编码当翻译时编码抗体或其抗体片段的序列,其中所述抗体或其抗体片段包含重链可变区(VH),所述重链可变区包含CRTH2R结合域,其中所述多个核酸中的每个核酸包含编码所述CRTH2R结合域的序列变体的序列,并且其中所述抗体或抗体片段以小于100nM的K_D与其抗原结合。本文进一步提供了核酸文库,其中VH的长度为约90至约100个氨基酸。本文进一步提供了核酸文库,其中VH的长度为约100至约400个氨基酸。本文进一步提供了核酸文库,其中VH的长度为约270至约300个碱基对。本文进一步提供了核酸文库,其中VH的长度为约300至约1200个碱基对。本文进一步提供了核酸文库,其中所述文库包含至少10^5个不同的核酸。

[0006] 本文提供了核酸文库,其包含:多个核酸,其中每个所述核酸编码当翻译时编码CRTH2R单域抗体的序列,其中所述多个序列的每个序列包含编码重链可变区(VH)上的CDR1、CDR2或CDR3的变异序列;其中所述文库包含至少30,000个变异序列;并且其中所述CRTH2R单域抗体以小于100nM的K_D与其抗原结合。本文进一步提供了核酸文库,其中当翻译时所述VH的长度为约90至约100个氨基酸。本文进一步提供了核酸文库,其中当翻译时所述VH的长度为约100至约400个氨基酸。本文进一步提供了核酸文库,其中所述VH的长度为约270至约300个碱基对。本文进一步提供了核酸文库,其中所述VH的长度为约300至约1200个碱基对。本文进一步提供了核酸文库,其中当翻译时所述VH包括与SEQ ID NO:2338-SEQ ID NO:2360或SEQ ID NO:2403-SEQ ID NO:2405中的任一个具有至少80%的序列同一性。

[0007] 本文提供了结合CRTH2R的抗体或抗体片段,其包含免疫球蛋白重链和免疫球蛋白轻链:a.其中所述免疫球蛋白重链包含与SEQ ID NO:2338-SEQ ID NO:2360或SEQ ID NO:2403-SEQ ID NO:2405中的任一个所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且b.其中所述免疫球蛋白轻链包含与SEQ ID NO:2361-SEQ ID NO:2381或SEQ ID NO:2406-SEQ ID NO:2408中的任一个所示的氨基酸序列具有至少约90%同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2338所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2361所示的氨基酸序列具有至少约90%同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2339所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2362所示的氨基酸序列具有至少约90%同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2340所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2363所示的氨基酸序列具有至少约90%同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2341所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2364所示的氨基酸序列具有至少约90%同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2342所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2365所示的氨基酸序列具有至少约90%同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段,其中所述免疫球蛋白重链包含与SEQ ID NO:2343所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中免疫球蛋白轻链包含与SEQ ID NO:2366所示的氨基酸序列具有至少约90%同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段,其中免疫球蛋白重链包含与SEQ ID NO:2344所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中免疫球蛋白轻链包含与SEQ ID NO:2367所示的氨基酸序列具有至少约90%同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段,其中免疫球蛋白重链包含与SEQ ID NO:2345所示的氨基酸序列具有至少约90%同一性的氨基酸序列;并且其中免疫球蛋白轻链包含与SEQ ID NO:2368所示的氨基酸序列具有至少约90%同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段,其中免疫球蛋白重链包含与SEQ ID NO:2346所示的氨基酸序列具有至少约90%同一性的氨基酸序列;

并且其中免疫球蛋白轻链包含与SEQ ID NO:2369中所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2347所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中免疫球蛋白轻链包含与SEQ ID NO:2370所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2348所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中免疫球蛋白轻链包含与SEQ ID NO:2371所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2349所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中免疫球蛋白轻链包含与SEQ ID NO:2372所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2350所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2373所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中所述免疫球蛋白重链包含与SEQ ID NO:2351所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2374所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2352中所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中免疫球蛋白轻链包含与SEQ ID NO:2375中所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2353所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2376所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2354所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2377所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2355所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2378所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2356所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2379所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2357所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2380所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2358所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2381所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2403所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2406所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合

CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2404所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2407所示的氨基酸序列具有至少约90％同一性的氨基酸序列。进一步提供了结合CRTH2R的抗体或抗体片段，其中免疫球蛋白重链包含与SEQ ID NO:2405所示的氨基酸序列具有至少约90％同一性的氨基酸序列；并且其中所述免疫球蛋白轻链包含与SEQ ID NO:2408所示的氨基酸序列具有至少约90％同一性的氨基酸序列。本文进一步提供了结合CRTH2R的抗体或其抗体片段，其中所述抗体是单克隆抗体、多克隆抗体、双特异性抗体、多特异性抗体、移植抗体、人抗体、人源化抗体、合成抗体、嵌合抗体、骆驼化抗体、单链Fv(scFv)、单链抗体、Fab片段、F(ab')2片段、Fd片段、Fv片段、单域抗体、分离的互补决定区(CDR)、双抗体、仅由单个单体可变域组成的片段、二硫键连接的Fv(sdFv)、胞内抗体、抗独特型(抗Id)抗体，或其ab抗原结合片段。进一步提供了结合CRTH2R的抗体或抗体片段，其中所述抗体或其抗体片段是嵌合的或人源化的。进一步提供了结合CRTH2R的抗体或抗体片段，其中所述抗体在cAMP测定中的EC50小于约25纳摩尔。进一步提供了结合CRTH2R的抗体或抗体片段，其中所述抗体在cAMP测定中的EC50小于约20纳摩尔。进一步提供了结合CRTH2R的抗体或抗体片段，其中所述抗体在cAMP测定中的EC50小于约10纳摩尔。

[0008] 本文提供了抗体或抗体片段，其中所述抗体或抗体片段包含互补决定区(CDR)，所述互补决定区包含与SEQ ID NO:2382-2402中的任一个所示的氨基酸序列具有至少约90％同一性的氨基酸序列。

[0009] 本文提供了抗体或抗体片段，其中所述抗体或抗体片段包含SEQ ID NO:2382-SEQ ID NO:2402中的任一个的序列并且其中所述抗体是单克隆抗体、多克隆抗体、双特异性抗体、多特异性抗体、移植抗体、人抗体、人源化抗体、合成抗体、嵌合抗体、骆驼化抗体、单链Fv(scFv)、单链抗体、Fab片段、F(ab')2片段、Fd片段、Fv片段、单域抗体、分离的互补决定区(CDR)、双抗体、仅由单个单体可变域组成的片段、二硫键连接的Fv(sdFv)、胞内抗体、抗独特型(抗Id)抗体，或其ab抗原结合片段。

[0010] 本文提供了治疗中枢神经系统、肾、肠、肺、毛发、皮肤、骨或软骨的疾病或病症的方法，其包括施用本文所述的抗体或抗体片段。

[0011] 本文提供了治疗以炎症反应为特征的疾病或病症的方法，其包括施用本文所述的抗体或抗体片段。

[0012] 本文提供了治疗过敏反应的方法，其包括施用本文所述的抗体或抗体片段。本文进一步提供了方法，其中所述过敏反应是慢性特发性荨麻疹或过敏性鼻炎。

[0013] 本文提供了治疗哮喘的方法，其包括施用本文所述的抗体或抗体片段。

[0014] 本文提供了治疗脱发或秃顶的方法，其包括施用本文所述的抗体或抗体片段。

[0015] 本文提供了用于生成编码CRTH2R抗体或其抗体片段的核酸文库的方法，其包括：(a)提供编码以下i-iii的预定序列：i.第一多个多核苷酸，其中所述第一多个多核苷酸中的每个多核苷酸编码至少1000个编码重链上的CDR1的变异序列；ii.第二多个多核苷酸，其中所述第二多个多核苷酸中的每个多核苷酸编码至少1000个编码重链上的CDR2的变异序列；iii.第三多个多核苷酸，其中所述第三多个多核苷酸中的每个多核苷酸编码至少1000个编码重链上的CDR3的变异序列；(b)将所述第一多个多核苷酸、所述第二多个多核苷酸和所述第三多个多核苷酸混合以形成编码所述CRTH2R抗体或其抗体片段的变异核酸的所述

核酸文库，并且其中至少约70％的所述变异核酸编码以小于100nM的K_D与其抗原结合的抗体或抗体片段。本文进一步提供了用于生成核酸文库的方法，其中所述CRTH2R抗体或其抗体片段是单域抗体。本文进一步提供了用于生成核酸文库的方法，其中所述单域抗体包含一个重链可变域。本文进一步提供了用于生成核酸文库的方法，其中所述单域抗体是VHH抗体。本文进一步提供了用于生成核酸文库的方法，其中所述核酸文库包含至少50,000个变异序列。本文进一步提供了用于生成核酸文库的方法，其中所述核酸文库包含至少100,000个变异序列。本文进一步提供了用于生成核酸文库的方法，其中所述核酸文库包含至少10^5个不同的核酸。本文进一步提供了用于生成核酸文库的方法，其中所述核酸文库包含至少一个编码CRTH2R抗体或抗体片段的序列，所述CRTH2R抗体或抗体片段以小于75nM的K_D与CRTH2R结合。本文进一步提供了用于生成核酸文库的方法，其中所述核酸文库包含至少一个编码CRTH2R抗体或抗体片段的序列，所述CRTH2R抗体或抗体片段以小于50nM的K_D与CRTH2R结合。本文进一步提供了用于生成核酸文库的方法，其中所述核酸文库包含至少一个编码CRTH2R抗体或抗体片段的序列，所述CRTH2R抗体或抗体片段以小于10nM的K_D与CRTH2R结合。本文进一步提供了用于生成核酸文库的方法，其中所述核酸文库包含至少500个变异序列。本文进一步提供了用于生成核酸文库的方法，其中所述核酸文库包含至少五个编码CRTH2R抗体或抗体片段的序列，所述CRTH2R抗体或抗体片段以小于75nM的K_D与CRTH2R结合。

[0016] 本文提供了由本文所述的核酸文库编码的蛋白质文库，其中所述蛋白质文库包含肽。本文进一步提供了蛋白质文库，其中所述蛋白质文库包含免疫球蛋白。本文进一步提供了蛋白质文库，其中所述蛋白质文库包含抗体。本文进一步提供了蛋白质文库，其中所述蛋白质文库是拟肽文库。

[0017] 本文提供了包含本文所述的核酸文库的载体文库。

[0018] 本文提供了包含本文所述的核酸文库的细胞文库。

[0019] 本文提供了包含本文所述的蛋白质文库的细胞文库。

附图说明

[0020] 图1描绘了G蛋白偶联受体(GPCR)配体相互作用表面的示意图。

[0021] 图2A描绘了免疫球蛋白支架的第一示意图。

[0022] 图2B描绘了免疫球蛋白支架的第二示意图。

[0023] 图3描绘了用于放置在支架中的基序的示意图。

[0024] 图4描绘了GPCR的示意图。

[0025] 图5描绘了用于组装克隆片段和非克隆片段的区段的示意图。

[0026] 图6描绘了用于组装克隆片段和非克隆片段的区段的示意图。

[0027] 图7呈现了说明如本文所公开的基因合成的示例性处理工作流程的步骤图。

[0028] 图8示出了计算机系统的示例。

[0029] 图9是示出计算机系统的架构的框图。

[0030] 图10是说明网络的示图，该网络被配置用于并入多个计算机系统、多个蜂窝电话和个人数据助理，以及网络附加存储(NAS)。

[0031] 图11是使用共享虚拟地址存储空间的多处理器计算机系统的框图。

[0032]

[0033] 图12A描绘了包含使用连接体连接至VL结构域的VH结构域的免疫球蛋白支架的示意图。

[0034] 图12B描绘了免疫球蛋白支架的全结构域架构的示意图，该免疫球蛋白支架包含使用连接体连接至VL结构域的VH结构域、前导序列和pIII序列。

[0035] 图12C描绘了VL或VH结构域的四个框架元件(FW1、FW2、FW3、FW4)和可变的3个CDR(L1、L2、L3)元件的示意图。

[0036] 图13描绘了纯化的GPCR免疫球蛋白的产量图。

[0037] 图14描绘了GPCR结合蛋白表达的FACS数据。

[0038] 图15A-图15C描绘了纯化的免疫球蛋白的细胞结合亲和力。

[0039] 图15D描绘了纯化的免疫球蛋白的cAMP活性。

[0040] 图16描绘了纯化的免疫球蛋白的BVP ELISA数据。

[0041] 图17A-图17B描绘了CRTH2-41-51的FACS分析图(图17A)以及剂量曲线和cAMP活性图(图17B)。

[0042] 图18描绘了CRTH2-44-59的剂量曲线。

[0043] 图19描绘了CRTH2-44-59的FACS分析图。

[0044] 图20A-图20E描绘了细胞结合的FACS分析图，是通过平均荧光强度(MFI)与CRTH2R IgG的8点滴定使用CRTH2-74、CRTH2-24、CRTH2-28、CRTH2-39、CRTH2-19、CRTH2-9、CRTH2-8、CRTH2-27、CRTH2-45、CRTH2-35、CRTH2-50、CRTH2-66、CRTH2-57、CRTH2-32、CRTH2-15、CRTH2-25、CRTH2-42、CRTH2-55、CRTH2-60和CRTH2-70测量的。

[0045] 图21A描绘了示例性门控点图，显示了在100nM的CRTH2-27结合。

[0046] 图21B描绘了示例性APC直方图，显示了在100nM的CRTH2-27结合。

[0047] 图22A描绘了使用比较物抗体gPCR-51的如先前图中的结合分析。

[0048] 图22B描绘了使用比较物抗体gPCR-52的如先前图中的结合分析。

[0049] 图23A-图23B描绘了IgG与CRTH2-9、CRTH2-27、CRTH2-50、CRTH2-32和CRTH2-42的结合曲线，它们在cAMP测定中具有功能效应。

[0050] 图24A描绘了在300、100和33nM测试的所有抗体的CRTH2R cAMP测定结果。

[0051] 图24B描绘了在33nM测试的所有抗体的CRTH2R cAMP测定结果。

[0052] 图25显示了在五种CRTH2R IgG(CRTH2-9、CRTH2-27、CRTH2-50、CRTH2-32和CRTH2-42)中看到的负别构效应。

[0053] 图26A-图26C描绘了别构调节剂的对照实验，显示比较物抗体52是正别构调节剂。

[0054] 图27A-图27D描绘了β-抑制蛋白(β-arrestin)募集测定中的CRTH2R的活性。

[0055] 图28描绘了本文生成的噬菌体展示的超免疫文库的设计示意图。

[0056] 图29A-图29F描绘了CRTH2R免疫球蛋白CRTH2-48-03(图29A)、CRTH2-48-21(图29B)和CRTH2-48-27(图29C)的结合亲和力图以及CRTH2-48-03(图29D)、CRTH2-48-21(图29E)和CRTH2-48-27(图29F)的cAMP测定图。

[0057] 图30A描绘了重链IGHV3-23设计的示意图。

[0058] 图30B描绘了重链IGHV1-69设计的示意图。

[0059] 图30C描绘了轻链IGKV 2-28和IGLV 1-51设计的示意图。

[0060] 图30D描绘了GPCR文库的理论多样性和最终多样性的示意图。

具体实施方式

[0061] 除非另有说明，否则本公开采用在本领域技术范围内的常规分子生物学技术。除非另有定义，否则本文使用的所有技术和科学术语具有与本领域普通技术人员通常理解的相同的含义。

[0062] 定义

[0063] 贯穿本公开内容，多个实施方案以范围格式给出。应当理解，范围格式的描述只是为了方便和简明，而不应被解释为对任何实施方案的范围的硬性限制。因此，除非上下文另有明确规定，否则对范围的描述应被认为明确公开了所有可能的子范围以及该范围内精确到下限单位十分之一的各个数值。例如，对诸如从1至6的范围的描述应被认为已经明确公开了诸如从1至3、从1至4、从1至5、从2至4、从2至6、从3至6等子范围，以及该范围内的各个值，例如，1.1、2、2.3、5和5.9。无论范围的宽度如何，这都是适用的。这些中间范围的上限和下限可独立地包括在更小的范围内，并且也被涵盖于本公开内容中，受所述范围中任何具体排除的限值所约束。除非上下文另有明确规定，否则在所述范围包括限值之一或两者的情况下，排除了这些所包含的限值中的任一个或两者的范围也被包括在本公开内容中。

[0064] 本文使用的术语仅用于描述特定实施方案的目的，而非旨在限制任何实施方案。除非上下文另有明确规定，否则如本文所用的单数形式“一个”、“一种”和“该”也意欲包括复数形式。进一步应当理解，术语“包括”和/或“包含”在本说明书中使用时指代所述特征、整数、步骤、操作、元件和/或组分的存在，但不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组分和/或其群体。如本文所用的，术语“和/或”包括一个或多个相关所示项目的任何及所有组合。

[0065] 除非具体说明或从上下文中可以明显看出，否则如本文所用的，关于数字或数字范围的术语“约”应被理解为表示所述数字及其+/-10％的数字，或者对于范围列出的值，表示低于所示下限的10％至高于所示上限的10％。

[0066] 除非具体说明，否则如本文所用的，术语“核酸”涵盖双链或三链核酸以及单链分子。在双链或三链核酸中，核酸链不必共同延伸(即，双链核酸不必沿两条链的全长都是双链的)。当提供时，核酸序列以5’至3’的方向列出，除非另有说明。本文所述的方法提供了分离的核酸的生成。本文所述的方法另外提供了分离并纯化的核酸的生成。本文提及的“核酸”在长度上可包含至少5、10、20、30、40、50、60、70、80、90、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500、600、700、800、900、1000、1100、1200、1300、1400、1500、1600、1700、1800、1900、2000个或更多个碱基。而且，本文提供了合成任意数目的编码多肽区段的核苷酸序列的方法，该序列包括编码非核糖体肽(NRP)的序列，编码以下物质的序列：非核糖肽合成酶(NRPS)模块和合成变体、其它模块化蛋白质如抗体的多肽区段、来自其它蛋白质家族的多肽区段，包括非编码DNA或RNA，如调节序列，例如启动子、转录因子、增强子、siRNA、shRNA、RNAi、miRNA、衍生自微小RNA的核仁小RNA，或任何感兴趣的功能性或结构性DNA或RNA单元。以下是多核苷酸的非限制性实例：基因或基因片段的编码区或非编码区、基因间DNA、由连锁分析限定的基因座(多个基因座)、外显子、内含子、信使RNA(mRNA)、转移RNA、核糖体RNA、短干扰RNA(siRNA)、短发夹RNA(shRNA)、微小RNA

(miRNA)、核仁小RNA、核酶、互补DNA(cDNA)(其为mRNA的DNA呈现形式,通常通过信使RNA(mRNA)的逆转录或通过扩增来获得);经合成或通过扩增产生的DNA分子、基因组DNA、重组多核苷酸、支链多核苷酸、质粒、载体、任何序列的分离的DNA、任何序列的分离的RNA、核酸探针和引物。编码本文提及的基因或基因片段的cDNA可包含至少一个编码外显子序列的区域,而没有基因组等同序列中的间插内含子序列。

[0067] 如本文所用,关于序列的术语"百分比(%)序列同一性"定义为在比对序列且在必要时引入空位以达到最大百分比序列同一性,以及不考虑将任何保守取代作为序列同一性的部分之后,候选序列中氨基酸残基与特定序列中氨基酸残基相同的百分比。用于确定百分比氨基酸序列同一性目的的比对可以以本领域技术范围内的各种方式来实现,例如,使用公开可用的计算机软件,例如EMBOSS MATCHER、EMBOSS WATER、EMBOSS STRETCHER、EMBOSS NEEDLE、EMBOSS LALIGN、BLAST、BLAST-2、ALIGN或Megalign(DNASTAR)软件。本领域技术人员可以确定用于测量比对的适当参数,包括在被比较的序列的全长上实现最大比对所需的任何算法。

[0068] CRTH2R文库

[0069] 本文提供了涉及用于前列腺素D2受体2(DP2或CRTH2R)的G蛋白偶联受体(GPCR)结合文库的方法和组合物,所述文库包含编码包含CRTH2R结合域的支架的核酸。如本文所述的支架可以稳定地支持CRTH2R结合域。可以基于CRTH2R配体与CRTH2R的表面相互作用来设计CRTH2R结合域。如本文所述的文库可以进一步被斑驳化(variegated),以提供包含各自编码至少一个预定参考核酸序列的预定变体的核酸的变体文库。本文进一步描述了可以在翻译所述核酸文库时生成的蛋白质文库。在某些情况下,将如本文所述的核酸文库转移到细胞中以生成细胞文库。本文还提供了使用本文所述的方法合成的文库的下游应用。下游应用包括鉴定具有增强的生物学相关功能(例如,改善的稳定性、亲和力、结合、功能活性)和用于治疗或预防与CRTH2R信号传导相关的疾病状态的变异核酸或蛋白质序列。

[0070] 本文所述的用于优化CRTH2R免疫球蛋白或抗体的方法、组合物和系统包括反映抗体序列的天然多样性的比率-变体方法。在某些情况下,优化的CRTH2R免疫球蛋白或抗体文库包含变异CRTH2R免疫球蛋白或抗体序列。在某些情况下,变异CRTH2R免疫球蛋白或抗体序列被设计为包含变异CDR区。在某些情况下,包含变异CDR区的变异CRTH2R免疫球蛋白或抗体序列是通过在美洲驼、人源化或嵌合框架中改组天然CDR序列而产生的。在某些情况下,合成此类文库,将其克隆到表达载体中,并评估翻译产物(抗体)的活性。在某些情况下,合成并随后组装序列片段。在某些情况下,表达载体用于展示和富集所需抗体,例如噬菌体展示。在某些情况下,噬菌体载体是Fab噬菌粒载体。在某些情况下,在富集过程中使用的选择压力包括结合亲和力、毒性、免疫耐受性、稳定性或其他因素。此类表达载体允许选择具有特定特性的抗体("淘选"),并且此类序列的后续增殖或扩增使具有这些序列的文库富集。淘选轮可以重复任意次数,例如1、2、3、4、5、6、7轮或超过7轮。在某些情况下,每一轮淘选都涉及多次洗涤。在某些情况下,每轮淘选涉及至少或约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16次或超过16次洗涤。

[0071] 本文描述了计算机文库设计的方法和系统。在某些情况下,本文描述的文库是基于包含多种抗体序列的数据库设计的。在某些情况下,数据库包含针对各种靶标的多个变异抗体序列。在某些情况下,数据库包含至少100、500、1000、1500、2000、2500、3000、3500、

4000、4500、5000个或超过5000个抗体序列。示例性数据库是iCAN数据库。在某些情况下，数据库包含幼稚和记忆B细胞受体序列。在某些情况下，幼稚和记忆B细胞受体序列是人、小鼠或灵长类动物的序列。在某些情况下，幼稚和记忆B细胞受体序列是人序列。在某些情况下，针对位置特异性变异分析数据库。在某些情况下，本文所述的抗体包含CDR区中的位置特异性变异。在某些情况下，CDR区包含多个变异位点。

[0072] 支架文库

[0073] 本文提供了包含编码支架的核酸的文库，其中编码CRTH2R结合域的序列被置于该支架中。本文所述的支架允许一定范围的CRTH2R结合域编码序列在插入该支架中时与插入未修饰的支架时相比改善的稳定性。示例性的支架包括但不限于蛋白质、肽、免疫球蛋白、其衍生物或它们的组合。在某些情况下，该支架是免疫球蛋白。本文所述的支架具有改善的功能活性、结构稳定性、表达、特异性或其组合。在某些情况下，支架包含用于支撑CRTH2R结合域的长区域。

[0074] 本文提供了包含编码支架的核酸的文库，其中该支架是免疫球蛋白。在某些情况下，该免疫球蛋白是抗体。如本文所用的，术语抗体将被理解为包括具有典型抗体分子的特征性双臂、Y形的蛋白质以及保留与抗原特异性结合的能力的抗体的一个或多个片段。示例性抗体包括但不限于单克隆抗体、多克隆抗体、双特异性抗体、多特异性抗体、移植抗体、人抗体、人源化抗体、合成抗体、嵌合抗体、骆驼化抗体、单链Fv(scFv)(包括这样的片段，其中VL和VH使用重组方法通过合成或天然连接体连接，该连接体使它们能够成为单蛋白质链，其中VL和VH区配对形式单价分子，包括单链Fab和scFab)、单链抗体、Fab片段(包括包含VL、VH、CL和CH1结构域的单价片段)、F(ab′)2片段(包括包含在铰链区通过二硫键连接的两个Fab片段的二价片段)、Fd片段(包括包含VH和CH1片段的片段)、Fv片段(包括包含抗体单臂的VL和VH结构域的片段)、单域抗体(dAb或sdAb)(包括包含VH结构域的片段)、分离的互补决定区(CDR)、双抗体(包括包含二价二聚体的片段，该二聚体例如是彼此结合并识别两种不同抗原的两个VL和VH结构域)、仅由单个单体可变域组成的片段、二硫键连接的Fv(sdFv)、胞内抗体、抗独特型(抗Id)抗体，或它们的抗原结合片段。在某些情况下，本文公开的文库包含编码支架的核酸，其中该支架是Fv抗体，包括由包含完整抗原识别和抗原结合位点的最小抗体片段组成的Fv抗体。在一些实施方案中，该Fv抗体由紧密、非共价缔合的一个重链可变域和一个轻链可变域的二聚体组成，并且每个可变域的三个高变区相互作用以在该VH-VL二聚体的表面上限定抗原结合位点。在一些实施方案中，这六个高变区为抗体赋予抗原结合特异性。在一些实施方案中，单个可变域(或仅包含对抗原为特异性的三个高变区的Fv的一半，包括包含一个重链可变域或者重链可变区的从骆驼科动物分离的单域抗体，如VHH抗体或纳米抗体)具有识别并结合抗原的能力。在某些情况下，本文公开的文库包含编码支架的核酸，其中该支架是单链Fv或scFv，包括包含VH、VL或VH和VL结构域两者的抗体片段，其中这两个结构域均存在于单个多肽链中。在一些实施方案中，Fv多肽在VH与VL结构域之间进一步包含多肽连接体，从而允许scFv形成用于抗原结合的所需结构。在某些情况下，scFv连接至Fc片段，或VHH连接至Fc片段(包括小抗体)。在某些情况下，所述抗体包含免疫球蛋白分子和免疫球蛋白分子的免疫活性片段，例如，含有抗原结合位点的分子。免疫球蛋白分子是任何类型(例如IgG、IgE、IgM、IgD、IgA和IgY)、类别(例如IgG 1、IgG 2、IgG 3、IgG 4、IgA 1和IgA 2)或亚类的。

[0075] 在一些实施方案中，文库包含适用于预期治疗靶标的种类的免疫球蛋白。通常，这些方法包括“哺乳动物化”，并且包括将供体抗原结合信息转移到免疫原性较低的哺乳动物抗体接受体以生成有用的治疗性处理的方法。在某些情况下，该哺乳动物是小鼠、大鼠、马、绵羊、牛、灵长类动物(例如黑猩猩、狒狒、大猩猩、猩猩、猴)、狗、猫、猪、驴、兔和人。在某些情况下，本文提供了用于抗体的猫科动物化和犬科动物化的文库和方法。

[0076] 非人抗体的“人源化”形式可以是含有衍生自非人抗体的最小序列的嵌合抗体。人源化抗体通常是人抗体(接受体抗体)，其中来自一个或多个CDR的残基被替换为来自非人抗体(供体抗体)的一个或多个CDR的残基。供体抗体可以是任何合适的非人抗体，如具有所需特异性、亲和力或生物学效应的小鼠、大鼠、兔、鸡或非人灵长类动物抗体。在某些情况下，接受体抗体的选定框架区残基被替换为来自供体抗体的相应框架区残基。人源化抗体还可包含在接受体抗体或供体抗体中均未发现的残基。在某些情况下，进行这些修饰以进一步改进抗体性能。

[0077] “犬科动物化”可包括将非犬抗原结合信息从供体抗体转移到免疫原性较低的犬抗体接受体以生成可在狗中用作治疗剂的治疗的方法。在某些情况下，本文提供的非犬抗体的犬科动物化形式是含有衍生自非犬抗体的最小序列的嵌合抗体。在某些情况下，犬科动物化抗体是犬抗体序列(“接受体”或“受体”抗体)，其中接受体的高变区残基被替换为来自非犬物种(“供体”抗体)如小鼠、大鼠、兔、猫、狗、山羊、鸡、牛、马、美洲驼、骆驼、单峰驼、鲨鱼、非人灵长类动物、人、人源化、重组序列或具有所需性质的工程化序列的高变区残基。在某些情况下，犬抗体的框架区(FR)残基被替换为相应的非犬FR残基。在某些情况下，犬科动物化抗体包括在接受体抗体或供体抗体中未发现的残基。在某些情况下，进行这些修饰以进一步改进抗体性能。犬科动物化抗体还可包含犬抗体的免疫球蛋白恒定区(Fc)的至少一部分。

[0078] “猫科动物化”可包括将非猫抗原结合信息从供体抗体转移到免疫原性较低的猫抗体接受体以生成可在猫中用作治疗剂的治疗的方法。在某些情况下，本文提供的非猫抗体的猫科动物化形式是含有衍生自非猫抗体的最小序列的嵌合抗体。在某些情况下，猫科动物化抗体是猫抗体序列(“接受体”或“受体”抗体)，其中接受体的高变区残基被替换为来自非猫物种(“供体”抗体)如小鼠、大鼠、兔、猫、狗、山羊、鸡、牛、马、美洲驼、骆驼、单峰驼、鲨鱼、非人灵长类动物、人、人源化、重组序列或具有所需性质的工程化序列的高变区残基。在某些情况下，猫抗体的框架区(FR)残基被替换为相应的非猫FR残基。在某些情况下，猫科动物化抗体包括在接受体抗体或供体抗体中未发现的残基。在某些情况下，进行这些修饰以进一步改进抗体性能。猫科动物化抗体还可包含猫抗体的免疫球蛋白恒定区(Fc)的至少一部分。

[0079] 本文提供了包含编码支架的核酸的文库，其中该支架是非免疫球蛋白。在某些情况下，该支架是非免疫球蛋白结合域。例如，该支架是抗体模拟物。示例性抗体模拟物包括但不限于anticalins、affilins、affibody分子、affimers、affitins、alphabodies、avimers、atrimers、DARPins、fynomers、基于Kunitz结构域的蛋白质、单抗体(monobodies)、anticalins、knottins、基于犰狳重复蛋白的蛋白质和双环肽。

[0080] 本文所述的包含编码支架的核酸的文库(其中该支架是免疫球蛋白)在该免疫球蛋白的至少一个区域中包含变异。用于变异的抗体的示例性区域包括但不限于互补决定区

(CDR)、可变域或恒定域。在某些情况下，该CDR是CDR1、CDR2或CDR3。在某些情况下，该CDR是重链结构域，包括但不限于CDRH1、CDRH2和CDRH3。在某些情况下，该CDR是轻链结构域，包括但不限于CDRL1、CDRL2和CDRL3。在某些情况下，该可变域是轻链可变域(VL)或重链可变域(VH)。在某些情况下，CDR1、CDR2或CDR3属于轻链可变域(VL)。轻链可变域(VL)的CDR1、CDR2或CDR3可分别称为CDRL1、CDRL2或CDRL3。重链可变域(VH)的CDR1、CDR2或CDR3可分别称为CDRH1、CDRH2或CDRH3。在某些情况下，该VL结构域包含κ或λ链。在某些情况下，该恒定域是轻链恒定域(CL)或重链恒定域(CH)。

[0081] 本文所述的方法提供了合成包含编码支架的核酸的文库，其中每个核酸编码至少一个预定参考核酸序列的预定变体。在某些情况下，该预定参考序列是编码蛋白质的核酸序列，并且该变体文库包含编码至少单个密码子的变异的序列，使得由合成核酸编码的后续蛋白质中单个残基的多个不同变体通过标准翻译过程生成。在某些情况下，该支架文库包含共同编码多个位置处的变异的变化的核酸。在某些情况下，该变体文库包含编码CDRH1、CDRH2、CDRH3、CDRL1、CDRL2、CDRL3、VL或VH结构域的至少单个密码子的变异的序列。在某些情况下，该变体文库包含编码CDRH1、CDRH2、CDRH3、CDRL1、CDRL2、CDRL3、VL或VH结构域的多个密码子的变异的序列。在某些情况下，该变体文库包含编码框架元件1(FW1)、框架元件2(FW2)、框架元件3(FW3)或框架元件4(FW4)的多个密码子的变异的序列。用于变异的密码子的示例性数目包括但不限于至少或约1、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、225、250、275、300个或超过300个密码子。

[0082] 在某些情况下，用于变异的免疫球蛋白的至少一个区域来自重链V基因家族、重链D基因家族、重链J基因家族、轻链V基因家族或轻链J基因家族。在某些情况下，轻链V基因家族包含免疫球蛋白κ(IGK)基因或免疫球蛋白λ(IGL)。示例性基因包括但不限于IGHV1-18、IGHV1-69、IGHV1-8、IGHV3-21、IGHV3-23、IGHV3-30/33rn、IGHV3-28、IGHV1-69、IGHV3-74、IGHV4-39、IGHV4-59/61、IGKV1-39、IGKV1-9、IGKV2-28、IGKV3-11、IGKV3-15、IGKV3-20、IGKV4-1、IGLV1-51、IGLV2-14、IGLV1-40和IGLV3-1。在某些情况下，该基因是IGHV1-69、IGHV3-30、IGHV3-23、IGHV3、IGHV1-46、IGHV3-7、IGHV1或IGHV1-8。在某些情况下，该基因是IGHV1-69和IGHV3-30。在某些情况下，该基因是IGHJ3、IGHJ6、IGHJ、IGHJ4、IGHJ5、IGHJ2或IGH1。在某些情况下，该基因是IGHJ3、IGHJ6、IGHJ或IGHJ4。

[0083] 本文提供了包含编码免疫球蛋白支架的核酸的文库，其中所述文库用各种数目的片段合成。在某些情况下，所述片段包含CDRH1、CDRH2、CDRH3、CDRL1、CDRL2、CDRL3、VL或VH结构域。在某些情况下，所述片段包含框架元件1(FW1)、框架元件2(FW2)、框架元件3(FW3)或框架元件4(FW4)。在某些情况下，用至少或约2个片段、3个片段、4个片段、5个片段或超过5个片段合成所述支架文库。每个核酸片段的长度或合成的核酸的平均长度可以是至少或约50、75、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500、525、550、575、600个或超过600个碱基对。在某些情况下，该长度是大约50至600、75至575、100至550、125至525、150至500、175至475、200至450、225至425、250至400、275至375或300至350个碱基对。

[0084] 当翻译时，如本文所述的包含编码免疫球蛋白支架的核酸的文库包含各种长度的氨基酸。在某些情况下，每个氨基酸片段的长度或合成的氨基酸的平均长度可以是至少或约15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、105、110、115、120、125、

130、135、140、145、150个或超过150个氨基酸。在某些情况下，该氨基酸的长度是大约15至150、20至145、25至140、30至135、35至130、40至125、45至120、50至115、55至110、60至110、65至105、70至100或75至95个氨基酸。在某些情况下，该氨基酸的长度是约22个氨基酸至约75个氨基酸。在某些情况下，该免疫球蛋白支架包含至少或约100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000个或超过5000个氨基酸。

[0085] 使用本文所述的方法为用于变异的免疫球蛋白的至少一个区域从头合成大量变异序列。在某些情况下，为CDRH1、CDRH2、CDRH3、CDRL1、CDRL2、CDRL3、VL、VH或其组合从头合成大量变异序列。在某些情况下，为框架元件1(FW1)、框架元件2(FW2)、框架元件3(FW3)或框架元件4(FW4)从头合成大量变异序列。变异序列的数目可以是至少或约5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500个或超过500个序列。在某些情况下，变异序列的数目是至少或约500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000个或超过8000个序列。在某些情况下，变异序列的数目是大约10至500、25至475、50至450、75至425、100至400、125至375、150至350、175至325、200至300、225至375、250至350或275至325个序列。

[0086] 在某些情况下，用于免疫球蛋白至少一个区域的变异序列在长度或序列上是各不相同的。在某些情况下，从头合成的至少一个区域是用于CDRH1、CDRH2、CDRH3、CDRL1、CDRL2、CDRL3、VL、VH或其组合。在某些情况下，从头合成的至少一个区域是用于框架元件1(FW1)、框架元件2(FW2)、框架元件3(FW3)或框架元件4(FW4)。在某些情况下，该变异序列与野生型相比包含至少或约1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50个或超过50个变异核苷酸或氨基酸。在某些情况下，该变异序列与野生型相比包含至少或约1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45或50个额外的核苷酸或氨基酸。在某些情况下，该变异序列包含比野生型少至少或约1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45或50个的核苷酸或氨基酸。在某些情况下，该文库包含至少或约10^1、10^2、10^3、10^4、10^5、10^6、10^7、10^8、10^9、10^{10}个或超过10^{10}个变体。

[0087] 在合成支架文库之后，可以将支架文库用于筛选和分析。例如，测定支架文库的文库可展示性和淘选。在某些情况下，使用选择性标签分析可展示性。示例性标签包括但不限于放射性标记、荧光标记、酶、化学发光标签、比色标签、亲和标签或本领域已知的其他标记或标签。在某些情况下，该标签是组氨酸、聚组氨酸、myc、血凝素(HA)或FLAG。在某些情况下，通过使用各种方法的测序对支架文库进行测定，所述方法包括但不限于单分子实时(SMRT)测序、聚合酶克隆(Polony)测序、连接测序、可逆终止子测序、质子检测测序、离子半导体测序、纳米孔测序、电子测序、焦磷酸测序、Maxam-Gilbert测序、链终止(例如，Sanger)测序、+S测序或合成测序。

[0088] 在某些情况下，测定支架文库的功能活性、结构稳定性(例如，热稳定或pH稳定)、表达、特异性或其组合。在某些情况下，测定支架文库中能够折叠的支架。在某些情况下，测定抗体区域的功能活性、结构稳定性、表达、特异性、折叠或其组合。例如，测定VH区或VL区的功能活性、结构稳定性、表达、特异性、折叠或其组合。

[0089] CRTH2R文库

[0090] 本文提供了CRTH2R结合文库，其包含编码支架的核酸，所述支架包含CRTH2R结合

域的序列。在某些情况下，该支架是免疫球蛋白。在某些情况下，通过CRTH2R结合域与CRTH2R之间的相互作用来确定包含CRTH2R结合域的序列的支架。

[0091] 本文提供了包含编码包含CRTH2R结合域的支架的核酸的文库，其中该CRTH2R结合域是基于CRTH2R上的表面相互作用而设计的。在某些情况下，该CRTH2R结合域包含如SEQ ID NO:1定义的序列。在某些情况下，CRTH2R结合域与CRTH2R的氨基或羧基端相互作用。在某些情况下，CRTH2R结合域与至少一个跨膜结构域相互作用，该跨膜结构域包括但不限于跨膜结构域1(TM1)、跨膜结构域2(TM2)、跨膜结构域3(TM3)、跨膜结构域4(TM4)、跨膜结构域5(TM5)、跨膜结构域6(TM6)和跨膜结构域7(TM7)。在某些情况下，CRTH2R结合域与CRTH2R的细胞内表面相互作用。例如，CRTH2R结合域与至少一个细胞内环相互作用，该细胞内环包括但不限于细胞内环1(ICL1)、细胞内环2(ICL2)和细胞内环3(ICL3)。在某些情况下，CRTH2R结合域与CRTH2R的细胞外表面相互作用。例如，CRTH2R结合域与CRTH2R的至少一个胞外域(ECD)或细胞外环(ECL)相互作用。该细胞外环包括但不限于细胞外环1(ECL1)、细胞外环2(ECL2)和细胞外环3(ECL3)。

[0092] 本文描述了CRTH2R结合域，其中基于CRTH2R配体与CRTH2R之间的表面相互作用设计该CRTH2R结合域。在某些情况下，该配体是肽。在某些情况下，该配体是CRTH2R激动剂。在某些情况下，该配体是CRTH2R拮抗剂。在某些情况下，该配体是CRTH2R别构调节剂。在某些情况下，该别构调节剂是负别构调节剂。在某些情况下，该别构调节剂是正别构调节剂。

[0093] 使用各种方法分析基于CRTH2R配体与CRTH2R之间表面相互作用的CRTH2R结合域的序列。例如，进行多物种计算分析。在某些情况下，进行结构分析。在某些情况下，进行序列分析。可以使用本领域已知的数据库进行序列分析。数据库的非限制性实例包括但不限于NCBI BLAST(blast.ncbi.nlm.nih.gov/Blast.cgi)、UCSC Genome Browser(genome.ucsc.edu/)、UniProt(www.uniprot.org/)和IUPHAR/BPS Guide to PHARMACOLOGY(guidetopharmacology.org/)。

[0094] 本文描述了基于各种生物体之间的序列分析设计的CRTH2R结合域。例如，进行序列分析以鉴定不同生物中的同源序列。示例性的生物包括但不限于小鼠、大鼠、马、绵羊、牛、灵长类动物(例如黑猩猩、狒狒、大猩猩、猩猩、猴)、狗、猫、猪、驴、兔、鱼、苍蝇和人。

[0095] 在鉴定CRTH2R结合域之后，可以生成包含编码CRTH2R结合域的核酸的文库。在某些情况下，CRTH2R结合域文库包含基于构象配体相互作用、肽配体相互作用、小分子配体相互作用、CRTH2R的胞外域或靶向CRTH2R的抗体设计的CRTH2R结合域序列。在某些情况下，CRTH2R结合域文库包含基于肽配体相互作用设计的CRTH2R结合域的序列。CRTH2R结合域的文库可以被翻译以生成蛋白质文库。在某些情况下，翻译CRTH2R结合域的文库以生成肽文库、免疫球蛋白文库、其衍生物或它们的组合。在某些情况下，翻译CRTH2R结合域的文库以生成蛋白质文库，进一步修饰所述蛋白质文库以生成拟肽文库。在某些情况下，翻译CRTH2R结合域的文库以生成用于生成小分子的蛋白质文库。

[0096] 本文所述的方法提供了CRTH2R结合域文库的合成，该文库包含各自编码至少一个预定参考核酸序列的预定变体的核酸。在某些情况下，该预定参考序列是编码蛋白质的核酸序列，并且该变体文库包含编码至少单个密码子的变异的序列，使得由合成核酸编码的后续蛋白质中单个残基的多个不同变体通过标准翻译过程生成。在某些情况下，该CRTH2R结合域文库包含共同编码多个位置处的变异的变化的核酸。在某些情况下，该变体文库包

含编码CRTH2R结合域中的至少单个密码子的变异的序列。在某些情况下，该变体文库包含编码CRTH2R结合域中的多个密码子的变异的序列。用于变异的密码子的示例性数目包括但不限于至少或约1、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、225、250、275、300个或超过300个密码子。

[0097] 本文所述的方法提供了合成包含编码CRTH2R结合域的核酸的文库，其中该文库包含编码该CRTH2R结合域的长度变异的序列。在某些情况下，该文库包含编码与预定参考序列相比少至少或约1、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、225、250、275、300个或超过300个密码子的长度变异的序列。在某些情况下，该文库包含编码与预定参考序列相比多至少或约1、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、200、225、250、275、300个或超过300个密码子的长度变异的序列。

[0098] 在鉴定CRTH2R结合域后，可以将CRTH2R结合域置于如本文所述的支架中。在某些情况下，该支架是免疫球蛋白。在某些情况下，将CRTH2R结合域置于CDRH3区域中。可以置于支架中的GPCR结合域也可以被称为基序。可以基于结合、特异性、稳定性、表达、折叠或下游活性来设计包含CRTH2R结合域的支架。在某些情况下，包含CRTH2R结合域的支架使得能够与CRTH2R接触。在某些情况下，包含CRTH2R结合域的支架使得能够与CRTH2R高亲和力结合。CRTH2R结合域的示例性氨基酸序列在表1A中描述。

表1A.CRTH2R结合域的氨基酸序列

SEQ ID NO	GPCR	氨基酸序列
1	CRTH2R	MSANATLKPLCPILEQMSRLQSHSNTSIRYIDHAAVLLHGLASLLGL VENGVILFVVGCRMRQTVVTTWVLHLALSDLLASASLPFFTYFLAV GHSWELGTTFCKLHSSIFFLNMFASGFLLSAISLDRCLQVVRPVWA QNHRTVAAAHKVCLVLWALAVLNTVPYFVFRDTISRLDGRIMCYY NVLLLNPGPDRDATCNSRQAALAVSKFLLAFLVPLAIIASSHAAVSL RLQHRGRRRPGRFVRLVAAVVAAFALCWGPYHVFSLLEARAHAN PGLRPLVWRGLPFVTSLAFFNSVANPVLYVLTCPDMLRKLRRSLRT VLESVLVDDSELGGAGSSRRRRTSSTARSASPLALCSRPEEPRGPAR LLGWLLGSCAASPQTGPLNRALSSTSS

[0099] 本文提供了包含CRTH2R结合域的支架或免疫球蛋白，其中该CRTH2R结合域的序列支持与CRTH2R的相互作用。该序列可以与CRTH2R配体的序列同源或相同。在某些情况下，该CRTH2R结合域序列与SEQ ID NO:1具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性。在某些情况下，该CRTH2R结合域序列与SEQ ID NO:1具有至少或约95％同源性。在某些情况下，该CRTH2R结合域序列与SEQ ID NO:1具有至少或约97％同源性。在某些情况下，该CRTH2R结合域序列与SEQ ID NO:1具有至少或约99％同源性。在某些情况下，该CRTH2R结合域序列与SEQ ID NO:1具有至少或约100％同源性。在某些情况下，该CRTH2R结合域序列至少包含具有SEQ ID NO:1的至少或约10、20、30、40、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400个或超过400个氨基酸的部分。

[0100] 在一些实施方案中，CRTH2R抗体或免疫球蛋白序列包含互补决定区(CDR)，该互补

决定区包含如表1B和表14B中所示的序列。在一些实施方案中，CRTH2R抗体或免疫球蛋白序列包含互补决定区(CDR)，该互补决定区与SEQ ID NO:2409-SEQ ID NO:2420或SEQ ID NO:2382-SEQ ID NO:2402中的任一个具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含互补决定区(CDR)，该互补决定区与SEQ ID NO:2409-SEQ ID NO:2420或SEQ ID NO:2382-SEQ ID NO:2402中的任一个具有至少或约95％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含互补决定区(CDR)，该互补决定区与SEQ ID NO:2409-SEQ ID NO:2420或SEQ ID NO:2382-SEQ ID NO:2402中的任一个具有至少或约97％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含互补决定区(CDR)，该互补决定区与SEQ ID NO:2409-SEQ ID NO:2420或SEQ ID NO:2382-SEQ ID NO:2402中的任一个具有至少或约99％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含互补决定区(CDR)，该互补决定区与SEQ ID NO:2409-SEQ ID NO:2420或SEQ ID NO:2382-SEQ ID NO:2402中的任一个具有至少或约100％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含互补决定区(CDR)，该互补决定区包含具有SEQ ID NO:2409-SEQ ID NO:2420或SEQ ID NO:2382-SEQ ID NO:2402中的任一个的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分。

表1B.

构建体说明	氨基酸序列	SEQ ID NO
IGHV1-69 CDR1	GGTFSSYA	2409
IGHV1-69 CDR2	IIPIFGTA	2410
IGHV1-69 CDR3	CARNNNNNNNNFDYW	2411
IGHV3-23 CDR1	GFTFSSYA	2412
IGHV3-23 CDR2	ISGSGGST	2413
IGHV3-23 CDR3	CAKNNNNNNNNFDYW	2414
IGKV1-39 CDR1	QSISSY	2415
IGKV1-39 CDR2	AAS	2416
IGKV1-39 CDR3	CQQSYSTPNTF	2417
IGKV3-20 CDR1	QSVSSSY	2418
IGKV3-20 CDR2	GAS	2419
IGKV3-20 CDR3	CQQYGSSPNTF	2420

[0101] 在一些实施方案中，CRTH2R抗体或免疫球蛋白序列包含CDR1，该CDR1与SEQ ID NO:2409、SEQ ID NO:2412、SEQ ID NO:2415或SEQ ID NO:2418中的任一个具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR1，该CDR1与SEQ ID NO:2409、SEQ ID NO:2412、SEQ ID NO:2415或SEQ ID NO:2418中的任一个具有至少或约95％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR1，该CDR1与SEQ ID NO:2409、SEQ ID NO:2412、SEQ ID NO:2415或SEQ ID NO:2418中的任一个具有至少或约97％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR1，该CDR1与SEQ ID NO:2409、SEQ ID NO:2412、SEQ ID NO:2415或SEQ ID NO:2418中的任一个具有至少或约99％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR1，该CDR1与SEQ ID NO:

2409、SEQ ID NO:2412、SEQ ID NO:2415或SEQ ID NO:2418中的任一个具有100％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR1，该CDR1包含具有SEQ ID NO:2409、SEQ ID NO:2412、SEQ ID NO:2415或SEQ ID NO:2418中的任一个的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分。

[0102] 在一些实施方案中，CRTH2R抗体或免疫球蛋白序列包含CDR2，该CDR2与SEQ ID NO:2410、SEQ ID NO:2413、SEQ ID NO:2416或SEQ ID NO:2419中的任一个具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR2，该CDR2与SEQ ID NO:2410、SEQ ID NO:2413、SEQ ID NO:2416或SEQ ID NO:2419中的任一个具有至少或约95％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR2，该CDR2与SEQ ID NO:2410、SEQ ID NO:2413、SEQ ID NO:2416或SEQ ID NO:2419中的任一个具有至少或约97％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR2，该CDR2与SEQ ID NO:2410、SEQ ID NO:2413、SEQ ID NO:2416或SEQ ID NO:2419中的任一个具有至少或约99％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR2，该CDR2与SEQ ID NO:2410、SEQ ID NO:2413、SEQ ID NO:2416或SEQ ID NO:2419中的任一个具有100％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR2，该CDR2包含具有SEQ ID NO:2410、SEQ ID NO:2413、SEQ ID NO:2416或SEQ ID NO:2419中的任一个的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分。

[0103] 在一些实施方案中，CRTH2R抗体或免疫球蛋白序列包含CDR3，该CDR3与SEQ ID NO:2411、SEQ ID NO:2414、SEQ ID NO:2417、SEQ ID NO:2420或SEQ ID NO:2382-SEQ ID NO:2402中的任一个具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR3，该CDR3与SEQ ID NO:2411、SEQ ID NO:2414、SEQ ID NO:2417、SEQ ID NO:2420或SEQ ID NO:2382-SEQ ID NO:2402中的任一个具有至少或约95％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR3，该CDR3与SEQ ID NO:2411、SEQ ID NO:2414、SEQ ID NO:2417、SEQ ID NO:2420或SEQ ID NO:2382-SEQ ID NO:2402中的任一个具有至少或约97％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR3，该CDR3与SEQ ID NO:2411、SEQ ID NO:2414、SEQ ID NO:2417、SEQ ID NO:2420或SEQ ID NO:2382-SEQ ID NO:2402中的任一个具有至少或约99％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR3，该CDR3与SEQ ID NO:2411、SEQ ID NO:2414、SEQ ID NO:2417、SEQ ID NO:2420或SEQ ID NO:2382-SEQ ID NO:2402中的任一个具有100％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDR3，该CDR3包含具有SEQ ID NO:2411、SEQ ID NO:2414、SEQ ID NO:2417、SEQ ID NO:2420或SEQ ID NO:2382-SEQ ID NO:2402中的任一个的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分。

[0104] 在一些实施方案中，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2以及CDRH3，该CDRH1与SEQ ID NO:2409具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRH2与SEQ ID NO:2410具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRH3与SEQ ID NO:2411具有至少或约70％、80％、85％、90％、91％、92％、

93％、94％、95％、96％、97％、98％、99％或100％的序列同一性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2以及CDRH3，该CDRH1与SEQ ID NO:2409具有至少或约95％、97％、99％或100％同源性，该CDRH2与SEQ ID NO:2410具有至少或约95％、97％、99％或100％同源性；该CDRH3与SEQ ID NO:2411具有至少或约95％、97％、99％或100％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2以及CDRH3，该CDRH1包含具有SEQ ID NO:2409的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分，该CDRH2包含具有SEQ ID NO:2410的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分，该CDRH3包含具有SEQ ID NO:2411的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分。

[0105] 在一些实施方案中，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2以及CDRH3，该CDRH1与SEQ ID NO:2412具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRH2与SEQ ID NO:2413具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRH3与SEQ ID NO:2414具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2以及CDRH3，该CDRH1与SEQ ID NO:2412具有至少或约95％、97％、99％或100％同源性；该CDRH2与SEQ ID NO:2413具有至少或约95％、97％、99％或100％同源性；该CDRH3与SEQ ID NO:2414具有至少或约95％、97％、99％或100％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2和CDRH3，该CDRH1包含具有SEQ ID NO:2412的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分，该CDRH2包含具有SEQ ID NO:2413的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分，该CDRH3包含具有SEQ ID NO:2414的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分。

[0106] 在一些实施方案中，CRTH2R抗体或免疫球蛋白序列包含CDRL1、CDRL2以及CDRL3，该CDRL1与SEQ ID NO:2415具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL2与SEQ ID NO:2416具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL3与SEQ ID NO:2417具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2以及CDRH3，该CDRH1与SEQ ID NO:2415具有至少或约95％、97％、99％或100％同源性；该CDRH2与SEQ ID NO:2416具有至少或约95％、97％、99％或100％同源性；该CDRH3与SEQ ID NO:2417具有至少或约95％、97％、99％或100％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRL1、CDRL2以及CDRL3，该CDRL1包含具有SEQ ID NO:2415的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL2包含具有SEQ ID NO:2416的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL3包含具有SEQ ID NO:2417的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分。

[0107] 在一些实施方案中，CRTH2R抗体或免疫球蛋白序列包含CDRL1、CDRL2以及CDRL3，该CDRL1与SEQ ID NO:2418具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、

95％、96％、97％、98％、99％或100％的序列同一性；该CDRL2与SEQ ID NO:2419具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL2与SEQ ID NO:2420具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRL1、CDRL2以及CDRL3，该CDRL1与SEQ ID NO:2418具有至少或约95％、97％、99％或100％同源性；该CDRL2与SEQ ID NO:2419具有至少或约95％、97％、99％或100％同源性；该CDRL3与SEQ ID NO:2420具有至少或约95％、97％、99％或100％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRL1、CDRL2以及CDRL3，该CDRL1包含具有SEQ ID NO:2418的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL2包含具有SEQ ID NO:2419的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL3包含具有SEQ ID NO:2420的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分。

[0108] 在一些实施方案中，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2、CDRH3、CDRL1、CDRL2以及CDRL3，该CDRH1与SEQ ID NO:2409具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRH2与SEQ ID NO:2410具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRH3与SEQ ID NO:2411具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL1与SEQ ID NO:2415具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL2与SEQ ID NO:2416具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL3与SEQ ID NO:2417具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2、CDRH3、CDRL1、CDRL2以及CDRL3，该CDRH1与SEQ ID NO:2409具有至少或约95％、97％、99％或100％同源性；该CDRH2与SEQ ID NO:2410具有至少或约95％、97％、99％或100％同源性；该CDRH3与SEQ ID NO:2411具有至少或约95％、97％、99％或100％同源性；该CDRL1与SEQ ID NO:2415具有至少或约95％、97％、99％或100％同源性；该CDRL2与SEQ ID NO:2416具有至少或约95％、97％、99％或100％同源性；该CDRL3与SEQ ID NO:2417具有至少或约95％、97％、99％或100％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2、CDRH3、CDRL1、CDRL2以及CDRL3，该CDRH1包含具有SEQ ID NO:2409的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRH2包含具有SEQ ID NO:2410的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRH3包含具有SEQ ID NO:2411的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL1包含具有SEQ ID NO:2415的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL2包含具有SEQ ID NO:2416的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL3包含具有SEQ ID NO:2417的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分。

[0109] 在一些实施方案中，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2、CDRH3、

CDRL1、CDRL2以及CDRL3，该CDRH1与SEQ ID NO:2409具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRH2与SEQ ID NO:2410具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRH3与SEQ ID NO:2411具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL1与SEQ ID NO:2418具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL2与SEQ ID NO:2419具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL3与SEQ ID NO:2420具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2、CDRH3、CDRL1、CDRL2以及CDRL3，该CDRH1与SEQ ID NO:2409具有至少或约95％、97％、99％或100％同源性；该CDRH2与SEQ ID NO:2410具有至少或约95％、97％、99％或100％同源性；该CDRH3与SEQ ID NO:2411具有至少或约95％、97％、99％或100％同源性；该CDRL1与SEQ ID NO:2418具有至少或约95％、97％、99％或100％同源性；该CDRL2与SEQ ID NO:2419具有至少或约95％、97％、99％或100％同源性；该CDRL3与SEQ ID NO:2420具有至少或约95％、97％、99％或100％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2、CDRH3、CDRL1、CDRL2以及CDRL3，该CDRH1包含具有SEQ ID NO:2409的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRH2包含具有SEQ ID NO:2410的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRH3包含具有SEQ ID NO:2411的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL1包含具有SEQ ID NO:2418的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL2包含具有SEQ ID NO:2419的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL3包含具有SEQ ID NO:2420的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分。

[0110] 在一些实施方案中，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2、CDRH3、CDRL1、CDRL2以及CDRL3，该CDRH1与SEQ ID NO:2412具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRH2与SEQ ID NO:2413具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRH3与SEQ ID NO:2414具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL1与SEQ ID NO:2415具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL2与SEQ ID NO:2416具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL3与SEQ ID NO:2417具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2、CDRH3、CDRL1、CDRL2以及CDRL3，该CDRH1与SEQ ID NO:2412具有至少或约95％、97％、99％或100％同源性；该CDRH2与SEQ ID NO:2413具有至少或约95％、97％、99％或100％同源性；该CDRH3与SEQ ID NO:2414具有至少或约95％、

97％、99％或100％同源性；该CDRL1与SEQ ID NO:2415具有至少或约95％、97％、99％或100％同源性；该CDRL2与SEQ ID NO:2416具有至少或约95％、97％、99％或100％同源性；该CDRL3与SEQ ID NO:2417具有至少或约95％、97％、99％或100％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2、CDRH3、CDRL1、CDRL2以及CDRL3，该CDRH1包含具有SEQ ID NO:2412的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRH2包含具有SEQ ID NO:2413的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRH3包含具有SEQ ID NO:2414的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL1包含具有SEQ ID NO:2415的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL2包含具有SEQ ID NO:2416的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL3包含具有SEQ ID NO:2417的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分。

[0111] 在一些实施方案中，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2、CDRH3、CDRL1、CDRL2以及CDRL3，该CDRH1与SEQ ID NO:2412具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRH2与SEQ ID NO:2413具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRH3与SEQ ID NO:2414具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL1与SEQ ID NO:2418具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL2与SEQ ID NO:2419具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性；该CDRL3与SEQ ID NO:2420具有至少或约70％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的序列同一性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2、CDRH3、CDRL1、CDRL2以及CDRL3，该CDRH1与SEQ ID NO:2412具有至少或约95％、97％、99％或100％同源性；该CDRH2与SEQ ID NO:2413具有至少或约95％、97％、99％或100％同源性；该CDRH3与SEQ ID NO:2414具有至少或约95％、97％、99％或100％同源性；该CDRL1与SEQ ID NO:2418具有至少或约95％、97％、99％或100％同源性；该CDRL2与SEQ ID NO:2419具有至少或约95％、97％、99％或100％同源性；该CDRL3与SEQ ID NO:2420具有至少或约95％、97％、99％或100％同源性。在某些情况下，CRTH2R抗体或免疫球蛋白序列包含CDRH1、CDRH2、CDRH3、CDRL1、CDRL2以及CDRL3，该CDRH1包含具有SEQ ID NO:2412的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRH2包含具有SEQ ID NO:2413的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRH3包含具有SEQ ID NO:2414的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL1包含具有SEQ ID NO:2418的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL2包含具有SEQ ID NO:2419的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分；该CDRL3包含具有SEQ ID NO:2420的至少或约3、4、5、6、7、8、9、10、12、14、16个或超过16个氨基酸的至少一部分。

[0112] 在一些实施方案中，本文描述了与CRTH2R结合的抗体或免疫球蛋白。在某些情况

下,CRTH2R抗体或免疫球蛋白序列包含重链可变域,该重链可变域与SEQ ID NO:2338-SEQ ID NO:2360或SEQ ID NO:2403-SEQ ID NO:2405中的任一个具有至少或约70%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的序列同一性。在某些情况下,CRTH2R抗体或免疫球蛋白序列包含重链可变域,该重链可变域与SEQ ID NO:2338-SEQ ID NO:2360或SEQ ID NO:2403-SEQ ID NO:2405中的任一个具有至少或约95%的序列同一性。在某些情况下,CRTH2R抗体或免疫球蛋白序列包含重链可变域,该重链可变域与SEQ ID NO:2338-SEQ ID NO:2360或SEQ ID NO:2403-SEQ ID NO:2405中的任一个具有至少或约97%的序列同一性。在某些情况下,CRTH2R抗体或免疫球蛋白序列包含重链可变域,该重链可变域与SEQ ID NO:2338-SEQ ID NO:2360或SEQ ID NO:2403-SEQ ID NO:2405中的任一个具有至少或约99%的序列同一性。在某些情况下,CRTH2R抗体或免疫球蛋白序列包含重链可变域,该重链可变域与SEQ ID NO:2338-SEQ ID NO:2360或SEQ ID NO:2403-SEQ ID NO:2405中的任一个具有至少或约100%的序列同一性。在某些情况下,CRTH2R抗体或免疫球蛋白序列包含重链可变域,该重链可变域包含具有SEQ ID NO:2338-SEQ ID NO:2360和SEQ ID NO:2403-SEQ ID NO:2405中的任一个的至少或约1、2、3、4、5、6、7、8、9、10、12、14、16、18、20、30、40、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400个或超过400个氨基酸的至少一部分。

[0113] 在某些情况下,CRTH2R抗体或免疫球蛋白序列包含轻链可变域,该轻链可变域与SEQ ID NO:2361-SEQ ID NO:2381或SEQ ID NO:2406-SEQ ID NO:2408中的任一个具有至少或约70%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的序列同一性。在某些情况下,CRTH2R抗体或免疫球蛋白序列包含轻链可变域,该轻链可变域与SEQ ID NO:2361-SEQ ID NO:2381和SEQ ID NO:2406-SEQ ID NO:2408中的任一个具有至少或约95%的序列同一性。在某些情况下,CRTH2R抗体或免疫球蛋白序列包含轻链可变域,该轻链可变域与SEQ ID NO:2361-SEQ ID NO:2381或SEQ ID NO:2406-SEQ ID NO:2408中的任一个具有至少或约97%的序列同一性。在某些情况下,CRTH2R抗体或免疫球蛋白序列包含轻链可变域,该轻链可变域与SEQ ID NO:2361-SEQ ID NO:2381或SEQ ID NO:2406-SEQ ID NO:2408中的任一个具有至少或约99%的序列同一性。在某些情况下,CRTH2R抗体或免疫球蛋白序列包含轻链可变域,该轻链可变域与SEQ ID NO:2361-SEQ ID NO:2381或SEQ ID NO:2406-SEQ ID NO:2408中的任一个具有至少或约100%的序列同一性。在某些情况下,CRTH2R抗体或免疫球蛋白序列包含轻链可变域,该轻链可变域包含具有SEQ ID NO:2361-SEQ ID NO:2381或SEQ ID NO:2406-SEQ ID NO:2408中的任一个的至少或约1、2、3、4、5、6、7、8、9、10、12、14、16、18、20、30、40、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400个或超过400个氨基酸的至少一部分。

[0114] 本文提供了CRTH2R结合文库,其包含编码包含CRTH2R结合域的支架或免疫球蛋白的核酸,该CRTH2R结合域包含结构域类型、结构域长度的变异或残基变异。在某些情况下,该结构域是包含CRTH2R结合域的支架中的区域。例如,该区域是VH、CDRH3或VL结构域。在某些情况下,该结构域是CRTH2R结合域。

[0115] 本文所述的方法提供了合成各自编码至少一个预定参考核酸序列的预定变体的

核酸的CRTH2R结合文库。在某些情况下，该预定参考序列是编码蛋白质的核酸序列，并且该变体文库包含编码至少单个密码子的变异的序列，使得由合成核酸编码的后续蛋白质中单个残基的多个不同变体通过标准翻译过程生成。在某些情况下，该CRTH2R结合文库包含共同编码多个位置处的变异的变化的核酸。在某些情况下，该变体文库包含编码VH、CDRH3或VL结构域的至少单个密码子的变异的序列。在某些情况下，该变体文库包含编码CRTH2R结合域中的至少单个密码子的变异的序列。例如，表1A中列出的CRTH2R结合域的至少一个单密码子是各不相同的。在某些情况下，该变体文库包含编码VH、CDRH3或VL结构域的多个密码子的变异的序列。在某些情况下，该变体文库包含编码CRTH2R结合域中的多个密码子的变异的序列。用于变异的密码子的示例性数目包括但不限于至少或约1、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、225、250、275、300个或超过300个密码子。

[0116] 本文所述的方法提供了合成各自编码至少一个预定参考核酸序列的预定变体的核酸的CRTH2R结合文库，其中该CRTH2R结合文库包含编码结构域的长度变异的序列。在某些情况下，该结构域是VH、CDRH3或VL结构域。在某些情况下，该结构域是CRTH2R结合域。在某些情况下，该文库包含编码与预定参考序列相比少至少或约1、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、225、250、275、300个或超过300个密码子的长度变异的序列。在某些情况下，该文库包含编码与预定参考序列相比多至少或约1、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、200、225、250、275、300个或超过300个密码子的长度变异的序列。

[0117] 本文提供了CRTH2R结合文库，其包含编码包含CRTH2R结合域的支架的核酸，其中所述CRTH2R结合文库用各种数目的片段合成。在某些情况下，所述片段包含VH、CDRH3或VL结构域。在某些情况下，所述CRTH2R结合文库用至少或约2个片段、3个片段、4个片段、5个片段或超过5个片段合成。每个核酸片段的长度或合成的核酸的平均长度可以是至少或约50、75、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500、525、550、575、600个或超过600个碱基对。在某些情况下，该长度是大约50至600、75至575、100至550、125至525、150至500、175至475、200至450、225至425、250至400、275至375或300至350个碱基对。

[0118] 当翻译时，如本文所述的包含编码含有CRTH2R结合域的支架的核酸的CRTH2R结合域文库包含各种长度的氨基酸。在某些情况下，每个氨基酸片段的长度或合成的氨基酸的平均长度可以是至少或约15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、105、110、115、120、125、130、135、140、145、150个或超过150个氨基酸。在某些情况下，该氨基酸长度是大约15至150、20至145、25至140、30至135、35至130、40至125、45至120、50至115、55至110、60至110、65至105、70至100或75至95个氨基酸。在某些情况下，该氨基酸长度是约22个至约75个氨基酸。

[0119] 包含从头合成的编码含有CRTH2R结合域的支架的变异序列的CRTH2R结合文库包含大量变异序列。在某些情况下，为CDRH1、CDRH2、CDRH3、CDRL1、CDRL2、CDRL3、VL、VH或其组合从头合成大量变异序列。在某些情况下，为框架元件1(FW1)、框架元件2(FW2)、框架元件3(FW3)或框架元件4(FW4)从头合成大量变异序列。在某些情况下，为GPCR结合域从头合成大量变异序列。例如，变异序列的数目对于VH结构域是约1个至约10个序列，对于CRTH2R结合

域是108个序列，对于VK结构域是约1个至约44个序列，详见图2A-图2B。变异序列的数目可以是至少或约5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500个或超过500个序列。在某些情况下，变异序列的数目是约10至300、25至275、50至250、75至225、100至200或125至150个序列。

[0120] 包含从头合成的编码含有CRTH2R结合域的支架的变异序列的CRTH2R结合文库具有改善的多样性。例如，通过将CRTH2R结合域变体置于包含N末端CDRH3变异和C末端CDRH3变异的免疫球蛋白支架变体中来生成变体。在某些情况下，变体包括亲和力成熟变体。备选地或组合地，变体包括免疫球蛋白其他区域中的变体，该区域包括但不限于CDRH1、CDRH2、CDRL1、CDRL2和CDRL3。在某些情况下，CRTH2R结合文库的变体的数目是至少或约10^4、10^5、10^6、10^7、10^8、10^9、10^{10}、10^{11}、10^{12}、10^{13}、10^{14}、10^{15}、10^{16}、10^{17}、10^{18}、10^{19}、10^{20}或超过10^{20}个不相同的序列不相同的序列。例如，包含约10个VH区变异序列、约237个CDRH3区变异序列和约43个VL和CDRL3区变异序列的文库包含105个不相同的序列(10x237x43)。

[0121] 本文提供了包含编码CRTH2R抗体的核酸的文库，该CRTH2R抗体在该抗体的至少一个区域中包含变异，其中该区域是CDR区。在某些情况下，该CRTH2R抗体是包含一个重链可变域的单域抗体，如VHH抗体。在某些情况下，该VHH抗体在一个或多个CDR区中包含变异。在某些情况下，本文所述的文库包含CDR1、CDR2或CDR3的至少或约1、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、1200、1400、1600、1800、2000、2400、2600、2800、3000个或超过3000个序列。在某些情况下，本文所述的文库包含CDR1、CDR2或CDR3的至少或约10^4、10^5、10^6、10^7、10^8、10^9、10^{10}、10^{11}、10^{12}、10^{13}、10^{14}、10^{15}、10^{16}、10^{17}、10^{18}、10^{19}、10^{20}个或超过10^{20}个序列。例如，所述文库包含至少2000个CDR1序列、至少1200个CDR2序列和至少1600个CDR3序列。在某些情况下，每个序列是不相同的。

[0122] 在某些情况下，所述CDR1、CDR2或CDR3属于轻链可变域(VL)。轻链可变域(VL)的CDR1、CDR2或CDR3可分别被称为CDRL1、CDRL2或CDRL3。在某些情况下，本文所述的文库包含VL的CDR1、CDR2或CDR3的至少或约1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、1200、1400、1600、1800、2000、2400、2600、2800、3000个或超过3000个序列。在某些情况下，本文所述的文库包含VL的CDR1、CDR2或CDR3的至少或约10^4、10^5、10^6、10^7、10^8、10^9、10^{10}、10^{11}、10^{12}、10^{13}、10^{14}、10^{15}、10^{16}、10^{17}、10^{18}、10^{19}、10^{20}个或超过10^{20}个序列。例如，所述文库包含至少20个VL的CDR1序列、至少4个VL的CDR2序列和至少140个VL的CDR3序列。在某些情况下，所述文库包含至少2个VL的CDR1序列、至少1个VL的CDR2序列和至少3000个VL的CDR3序列。在某些情况下，该VL是IGKV1-39、IGKV1-9、IGKV2-28、IGKV3-11、IGKV3-15、IGKV3-20、IGKV4-1、IGLV1-51、IGLV2-14、IGLV1-40或IGLV3-1。在某些情况下，该VL是IGKV2-28。在某些情况下，该VL是IGLV1-51。

[0123] 在某些情况下，所述CDR1、CDR2或CDR3属于重链可变域(VH)。重链可变域(VH)的CDR1、CDR2或CDR3可分别被称为CDRH1、CDRH2或CDRH3。在某些情况下，本文所述的文库包含VH的CDR1、CDR2或CDR3的至少或约1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、1200、1400、1600、1800、2000、2400、2600、2800、3000个或超过3000个序列。在某些情况下，本文所述的文库包含VH的CDR1、CDR2或CDR3的至少或约10^4、10^5、10^6、10^7、10^8、10^9、10^{10}、10^{11}、10^{12}、10^{13}、10^{14}、10^{15}、10^{16}、10^{17}、10^{18}、

10^{19}、10^{20}个或超过10^{20}个序列。例如，文库包含至少30个VH的CDR1序列、至少570个VH的CDR2序列和至少10^{8}个VH的CDR3序列。在某些情况下，该20或IGHV4-59/61。在某些情况下，该VH是IGHV1-69、IGHV3-30、IGHV3-23、IGHV3、IGHV1-46、IGHV3-7、IGHV1或IGHV1-8。在某些情况下，该VH是IGHV1-69和IGHV3-30。在某些情况下，该VH是IGHV3-23。

[0124] 在一些实施方案中，如本文所述的文库包含不同长度的CDRL1、CDRL2、CDRL3、CDRH1、CDRH2或CDRH3。在某些情况下，该CDRL1、CDRL2、CDRL3、CDRH1、CDRH2或CDRH3的长度包含至少或约5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、40、50、60、70、80、90个或超过90个氨基酸。例如，该CDRH3在长度上包含至少或约12、15、16、17、20、21或23个氨基酸。在某些情况下，该CDRL1、CDRL2、CDRL3、CDRH1、CDRH2或CDRH3在长度上包含约1至约10、约5至约15、约10至约20或约15至约30个范围内的氨基酸。

[0125] 当翻译时，如本文所述的包含编码具有变异CDR序列的抗体的核酸的文库包含各种长度的氨基酸。在某些情况下，每个氨基酸片段的长度或合成的氨基酸的平均长度可以是至少或约15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、105、110、115、120、125、130、135、140、145、150个或超过150个氨基酸。在某些情况下，该氨基酸长度是大约15至150、20至145、25至140、30至135、35至130、40至125、45至120、50至115、55至110、60至110、65至105、70至100或75至95个氨基酸。在某些情况下，该氨基酸长度是约22个氨基酸至约75个氨基酸。在某些情况下，该抗体包含至少或约100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000个或超过5000个氨基酸。

[0126] CDRL1、CDRL2、CDRL3、CDRH1、CDRH2或CDRH3的长度比可以在本文所述的文库中变化。在某些情况下，在长度上包含至少或约5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、40、50、60、70、80、90个或超过90个氨基酸的CDRL1、CDRL2、CDRL3、CDRH1、CDRH2或CDRH3占所述文库的约10％、20％、30％、40％、50％、60％、70％、80％、90％或超过90％。例如，在长度上包含约23个氨基酸的CDRH3以40％存在于文库中，在长度上包含约21个氨基酸的CDRH3以30％存在于文库中，在长度上包含约17个氨基酸的CDRH3以20％存在于文库中，在长度上包含约12个氨基酸的CDRH3以10％存在于文库中。在某些情况下，在长度上包含约20个氨基酸的CDRH3以40％存在于文库中，在长度上包含约16个氨基酸的CDRH3以30％存在于文库中，在长度上包含约15个氨基酸的CDRH3以20％存在于文库中，在长度上包含约12个氨基酸的CDRH3以10％存在于文库中。

[0127] 如本文所述的编码VHH抗体的文库包含经改组以生成理论多样性为至少或约10^{7}、10^{8}、10^{9}、10^{10}、10^{11}、10^{12}、10^{13}、10^{14}、10^{15}、10^{16}、10^{17}、10^{18}、10^{19}、10^{20}或超过10^{20}个序列的文库的变异CDR序列。在某些情况下，该文库的最终文库多样性为至少或约10^{7}、10^{8}、10^{9}、10^{10}、10^{11}、10^{12}、10^{13}、10^{14}、10^{15}、10^{16}、10^{17}、10^{18}、10^{19}、10^{20}或超过10^{20}个序列。

[0128] 本文提供了编码免疫球蛋白的CRTH2R结合文库。在某些情况下，该CRTH2R免疫球蛋白是抗体。在某些情况下，该CRTH2R免疫球蛋白是VHH抗体。在某些情况下，该CRTH2R免疫球蛋白对CRTH2R的结合亲和力(例如K_D)小于1nM、小于1.2nM、小于2nM、小于5nM、小于10nM、小于11nm、小于13.5nM、小于15nM、小于20nM、小于25nM或小于30nM。在某些情况下，该CRTH2R免疫球蛋白具有小于1nM的K_D。在某些情况下，该CRTH2R免疫球蛋白具有小于1.2nM的K_D。在某些情况下，该CRTH2R免疫球蛋白具有小于2nM的K_D。在某些情况下，该CRTH2R免疫球蛋白具有小于5nM的K_D。在某些情况下，该CRTH2R免疫球蛋白具有小于10nM的K_D。在某些情

况下，该CRTH2R免疫球蛋白具有小于13.5nM的K_D。在某些情况下，该CRTH2R免疫球蛋白具有小于15nM的K_D。在某些情况下，该CRTH2R免疫球蛋白具有小于20nM的K_D。在某些情况下，该CRTH2R免疫球蛋白具有小于25nM的K_D。在某些情况下，该CRTH2R免疫球蛋白具有小于30nM的K_D。

[0129] 在某些情况下，该CRTH2R免疫球蛋白是CRTH2R激动剂。在某些情况下，该CRTH2R免疫球蛋白是CRTH2R拮抗剂。在某些情况下，该CRTH2R免疫球蛋白是CRTH2R别构调节剂。在某些情况下，该别构调节剂是负别构调节剂。在某些情况下，该别构调节剂是正别构调节剂。在某些情况下，该CRTH2R免疫球蛋白在至少或约1nM、2nM、4nM、6nM、8nM、10nM、20nM、30nM、40nM、50nM、60nM、70nM、80nM、90nM、100nM、120nM、140nM、160nM、180nM、200nM、300nM、400nM、500nM、600nM、700nM、800nM、900nM、1000nM或超过1000nM的浓度下导致激动、拮抗或别构效应。在某些情况下，该CRTH2R免疫球蛋白是负别构调节剂。在某些情况下，该CRTH2R免疫球蛋白在至少或约0.001、0.005、0.01、0.05、0.1、0.5、1nM、2nM、4nM、6nM、8nM、10nM、20nM、30nM、40nM、50nM、60nM、70nM、80nM、90nM、100nM或超过100nM的浓度下是负别构调节剂。在某些情况下，该CRTH2R免疫球蛋白在约0.001至约100、0.01至约90、约0.1至约80、1至约50、约10至约40nM或约1至约10nM范围内的浓度下是负别构调节剂。在某些情况下，该CRTH2R免疫球蛋白的EC50或IC50为至少或约0.001、0.0025、0.005、0.01、0.025、0.05、0.06、0.07、0.08、0.9、0.1、0.5、1、2、3、4、5、6nM或超过6nM。在某些情况下，该CRTH2R免疫球蛋白的EC50或IC50为至少或约1nM、2nM、4nM、6nM、8nM、10nM、20nM、30nM、40nM、50nM、60nM、70nM、80nM、90nM、100nM或超过100nM。

[0130] 如本文所述的CRTH2R免疫球蛋白可具有改善的性质。在某些情况下，该CRTH2R免疫球蛋白是单体的。在某些情况下，该CRTH2R免疫球蛋白不易聚集。在某些情况下，至少或约70％、75％、80％、85％、90％、95％或99％的CRTH2R免疫球蛋白是单体的。在某些情况下，该CRTH2R免疫球蛋白是热稳定的。在某些情况下，该CRTH2R免疫球蛋白导致非特异性结合减少。

[0131] 在合成包含编码含有CRTH2R结合域的支架的核酸的CRTH2R结合文库之后，可将文库用于筛选和分析。例如，分析文库的文库可展示性和淘选。在某些情况下，使用选择性标签分析可展示性。示例性标签包括但不限于放射性标记、荧光标记、酶、化学发光标签、比色标签、亲和标签或本领域已知的其他标记或标签。在某些情况下，该标签是组氨酸、聚组氨酸、myc、血凝素(HA)或FLAG。例如在图3可见所述CRTH2R结合文库包含编码支架的核酸，所述支架包含具有多个标签如GFP、FLAG和Lucy以及DNA条形码的GPCR结合域。在某些情况下，通过使用各种方法的测序对文库进行测定，所述方法包括但不限于单分子实时测序(SMRT)、聚合酶克隆(Polony)测序、连接测序、可逆终止子测序、质子检测测序、离子半导体测序、纳米孔测序、电子测序、焦磷酸测序、Maxam-Gilbert测序、链终止(例如，Sanger)测序、+S测序或合成测序。

[0132] 表达系统

[0133] 本文提供了包含编码含有CRTH2R结合域的支架的核酸的文库，其中所述文库具有改善的特异性、稳定性、表达、折叠或下游活性。在某些情况下，本文所述的文库用于筛选和分析。

[0134] 本文提供了包含编码含有CRTH2R结合域的支架的核酸的文库，其中所述核酸文库

用于筛选和分析。在某些情况下，筛选和分析包括体外、体内或离体测定。用于筛选的细胞包括取自活对象的原代细胞或细胞系。细胞可以来自原核生物(例如细菌和真菌)或真核生物(例如动物和植物)。示例性的动物细胞包括但不限于来自小鼠、兔、灵长类动物和昆虫的动物细胞。在某些情况下，用于筛选的细胞包括细胞系，包括但不限于中国仓鼠卵巢(CHO)细胞系、人胚肾(HEK)细胞系或幼仓鼠肾(BHK)细胞系。在某些情况下，本文所述的核酸文库也可以被递送至多细胞生物体。示例性的多细胞生物体包括但不限于植物、小鼠、兔、灵长类动物和昆虫。

[0135] 可以针对各种药理学或药代动力学性质筛选本文所述的核酸文库或其编码的蛋白质文库。在某些情况下，使用体外测定、体内测定或离体测定来筛选文库。例如，所筛选的体外药理学或药代动力学性质包括但不限于结合亲和力、结合特异性和结合亲合力。所筛选的本文所述文库的示例性体内药理学或药代动力学性质包括但不限于治疗功效、活性、临床前毒性性质、临床功效性质、临床毒性性质、免疫原性、效力和临床安全性性质。

[0136] 可以筛选的药理学或药代动力学性质包括但不限于细胞结合亲和力和细胞活性。例如，进行细胞结合亲和力测定或细胞活性测定，以确定本文所述的文库的激动、拮抗或别构效应。在某些情况下，该细胞活性测定是cAMP测定。在某些情况下，将本文所述的文库与CRTH2R配体的细胞结合或细胞活性进行比较。

[0137] 可以在基于细胞的测定或在非基于细胞的测定中筛选本文所述的文库。非基于细胞的测定的实例包括但不限于使用病毒颗粒、使用体外翻译蛋白质和使用具有CRTH2R的脂蛋白体。

[0138] 可通过测序筛选如本文所述的核酸文库。在某些情况下，使用下一代测序来确定CRTH2R结合变体的序列富集。在某些情况下，确定V基因分布、J基因分布、V基因家族、每个长度的CDR3计数或其组合。在某些情况下，确定克隆频率、克隆积累、谱系积累或其组合。在某些情况下，确定序列数、具有VH克隆的序列、克隆、大于1的克隆、克隆型、大于1的克隆型、谱系、simpsons或其组合。在某些情况下，确定不相同的CDR3的百分比。例如，不相同的CDR3的百分比被计算为样品中不相同的CDR3的数目除以样品中具有CDR3的序列总数。

[0139] 本文提供了核酸文库，其中所述核酸文库可以在载体中表达。用于插入本文公开的核酸文库的表达载体可包括真核或原核表达载体。示例性的表达载体包括但不限于哺乳动物表达载体:pSF-CMV-NEO-NH2-PPT-3XFLAG、pSF-CMV-NEO-COOH-3XFLAG、pSF-CMV-PURO-NH2-GST-TEV、pSF-OXB20-COOH-TEV-FLAG(R)-6His、pCEP4 pDEST27、pSF-CMV-Ub-KrYFP、pSF-CMV-FMDV-daGFP、pEF1a-mCherry-N1载体、pEF1a-tdTomato载体、pSF-CMV-FMDV-Hygro、pSF-CMV-PGK-Puro、pMCP-tag(m)和pSF-CMV-PURO-NH2-CMYC;细菌表达载体:pSF-OXB20-BetaGal、pSF-OXB20-Fluc、pSF-OXB20和pSF-Tac;植物表达载体:pRI 101-AN DNA和pCambia2301;和酵母表达载体:pTYB21和pKLAC2，以及昆虫载体:pAc5.1/V5-His A和pDEST8。在某些情况下，该载体是pcDNA3或pcDNA3.1。

[0140] 本文描述了在载体中表达以生成包含支架的构建体的核酸文库，该支架包含CRTH2R结合域的序列。在某些情况下，该构建体的大小各不相同。在某些情况下，该构建体包含至少或约500、600、700、800、900、1000、1100、1300、1400、1500、1600、1700、1800、2000、2400、2600、2800、3000、3200、3400、3600、3800、4000、4200、4400、4600、4800、5000、6000、7000、8000、9000、10000个或超过10000个碱基。在某些情况下，该构建体包含约300至1，

000、300至2,000、300至3,000、300至4,000、300至5,000、300至6,000、300至7,000、300至8,000、300至9,000、300至10,000、1,000至2,000、1,000至3,000、1,000至4,000、1,000至5,000、1,000至6,000、1,000至7,000、1,000至8,000、1,000至9,000、1,000至10,000、2,000至3,000、2,000至4,000、2,000至5,000、2,000至6,000、2,000至7,000、2,000至8,000、2,000至9,000、2,000至10,000、3,000至4,000、3,000至5,000、3,000至6,000、3,000至7,000、3,000至8,000、3,000至9,000、3,000至10,000、4,000至5,000、4,000至6,000、4,000至7,000、4,000至8,000、4,000至9,000、4,000至10,000、5,000至6,000、5,000至7,000、5,000至8,000、5,000至9,000、5,000至10,000、6,000至7,000、6,000至8,000、6,000至9,000、6,000至10,000、7,000至8,000、7,000至9,000、7,000至10,000、8,000至9,000、8,000至10,000或9,000至10,000个碱基的范围。

[0141] 本文提供了包含编码含有CRTH2R结合域的支架的核酸的文库,其中所述核酸文库在细胞中表达。在某些情况下,合成所述文库以表达报告基因。示例性的报告基因包括但不限于乙酰羟酸合酶(AHAS)、碱性磷酸酶(AP)、β-半乳糖苷酶(LacZ)、β-葡萄糖醛酸酶(GUS)、氯霉素乙酰转移酶(CAT)、绿色荧光蛋白(GFP)、红色荧光蛋白(RFP)、黄色荧光蛋白(YFP)、青色荧光蛋白(CFP)、天蓝色荧光蛋白、黄水晶荧光蛋白、橙色荧光蛋白、樱桃荧光蛋白、绿松石荧光蛋白、蓝色荧光蛋白、辣根过氧化物酶(HRP)、萤光素酶(Luc)、胭脂碱合酶(NOS)、章鱼碱合酶(OCS)、萤光素酶及其衍生物。确定报告基因的调节的方法是本领域公知的、并且包括但不限于荧光光度法(例如,荧光光谱法、荧光激活细胞分选(FACS)、荧光显微术)和抗生素抗性确定。

[0142] 疾病和病症

[0143] 本文提供了包含编码含有CRTH2R结合域的支架的核酸的CRTH2R结合文库,其可具有治疗效果。在某些情况下,所述CRTH2R结合文库当翻译时产生蛋白质,该蛋白质用于治疗疾病或病症。在某些情况下,该蛋白质是免疫球蛋白。在某些情况下,该蛋白质是拟肽。示例性的疾病包括但不限于癌症、炎性疾病或病症、代谢疾病或紊乱、心血管疾病或病症、呼吸系统疾病或病症、疼痛、消化系统疾病或病症、生殖系统疾病或病症、内分泌疾病或病症或者神经系统疾病或病症。在某些情况下,该癌症是实体癌或血液系统癌症。在某些情况下,如本文所述的前列腺素D2受体2(DP2或CRTH2R)抑制剂用于治疗中枢神经系统、肾、肠、肺、头发、皮肤、骨或软骨的疾病或病症。在某些情况下,如本文所述的CRTH2R抑制剂或拮抗剂用于治疗以炎症反应为特征的疾病或病症。在某些情况下,如本文所述的CRTH2R抑制剂或拮抗剂用于治疗过敏反应。在某些情况下,过敏反应是慢性特发性荨麻疹。在某些情况下,过敏反应是过敏性鼻炎。在某些情况下,如本文所述的CRTH2R抑制剂或拮抗剂用于治疗哮喘。在某些情况下,如本文所述的CRTH2R抑制剂或拮抗剂用于治疗脱发或秃顶。在某些情况下,该对象是哺乳动物。在某些情况下,该对象是小鼠、兔、狗或人。通过本文描述的方法治疗的对象可以是婴儿、成人或儿童。包含如本文所述的抗体或抗体片段的药物组合物可以静脉内或皮下施用。

[0144] 变体文库

[0145] 密码子变异

[0146] 本文所述的变异核酸文库可包含多个核酸,其中每个核酸编码与参考核酸序列相比的变异密码子序列。在某些情况下,第一核酸群体中的每个核酸在单变异位点处含有变

体。在某些情况下，第一核酸群体在单变异位点处含有多个变体，使得第一核酸群体在相同变异位点处含有超过一个变体。第一核酸群体可包含在相同变异位点处共同编码多个密码子变体的核酸。第一核酸群体可包含在相同位置处共同编码多达19个或更多个密码子的核酸。第一核酸群体可包含在相同位置处共同编码多达60个变异三联体的核酸，或者第一核酸群体可包含在相同位置处共同编码多达61个不同密码子三联体的核酸。每个变体可编码在翻译过程中产生不同氨基酸的密码子。

表3提供了对于变异位点可能的每个密码子(和代表性氨基酸)的列表。

表2.密码子和氨基酸列表

氨基酸	单字母代码	三字母代码	密码子					
丙氨酸	A	Ala	GCA	GCC	GCG	GCT		
半胱氨酸	C	Cys	TGC	TGT				
天冬氨酸	D	Asp	GAC	GAT				
谷氨酸	E	Glu	GAA	GAG				
苯丙氨酸	F	Phe	TTC	TTT				
甘氨酸	G	Gly	GGA	GGC	GGG	GGT		
组氨酸	H	His	CAC	CAT				
异亮氨酸	I	Iso	ATA	ATC	ATT			
赖氨酸	K	Lys	AAA	AAG				
亮氨酸	L	Leu	TTA	TTG	CTA	CTC	CTG	CTT
甲硫氨酸	M	Met	ATG					
天冬酰胺	N	Asn	AAC	AAT				
脯氨酸	P	Pro	CCA	CCC	CCG	CCT		
谷氨酰胺	Q	Gln	CAA	CAG				
精氨酸	R	Arg	AGA	AGG	CGA	CGC	CGG	CGT
丝氨酸	S	Ser	AGC	AGT	TCA	TCC	TCG	TCT
苏氨酸	T	Thr	ACA	ACC	ACG	ACT		
缬氨酸	V	Val	GTA	GTC	GTG	GTT		
色氨酸	W	Trp	TGG					
酪氨酸	Y	Tyr	TAC	TAT				

[0147] 核酸群体可包含在多个位置处共同编码至多20个密码子变异的改变的核酸。在这类情况下，该群体中的每个核酸包含在相同核酸中超过一个位置处的密码子变异。在某些情况下，该群体中的每个核酸包含在单个核酸中的1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或更多个密码子处的密码子变异。在某些情况下，每个变异长核酸包含在单个长核酸中的1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30个或更多个密码子处的密码子变异。在某些情况下，该变异核酸群体包含在单个核酸中的1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30个或更多个密码子处的密码子变异。在某些情况下，该变异核酸群体包含在单个长核酸中的至少约10、20、30、40、50、60、70、80、90、100个或更多个密

码子处的密码子变异。

[0148] 高度平行的核酸合成

[0149] 本文提供了一种平台方法,其利用从多核苷酸合成到硅上纳米孔内基因装配的端到端过程的小型化、平行化及垂直整合来创建革命性的合成平台。本文所述的装置采用与96孔板相同的占地面积(footprint)提供了这样一种硅合成平台,与传统合成方法相比,该硅合成平台能够将通量提高高达1,000倍或更多,其中在单次高度平行化运行中产生高达约1,000,000个或更多个多核苷酸或10,000个或更多个基因。

[0150] 随着新一代测序的出现,高分辨率基因组数据已成为深入研究各种基因在正常生物学和疾病发病机理中的生物学作用的研究的重要因素。本研究的核心是分子生物学的中心法则和"连续信息的逐残基转移"的概念。将DNA中编码的基因组信息转录成信息,随后将其翻译成蛋白质,该蛋白质是给定生物学途径内的活性产物。

[0151] 另一个令人兴奋的研究领域是关于着眼于高度特异性细胞靶标的治疗性分子的发现、研发和制备。高度多样性的DNA序列文库是靶向治疗剂的开发流程的核心。在设计、构建和测试蛋白质工程循环中使用基因突变体表达蛋白质,在理想情况下该循环得到针对对其治疗靶标具有高亲和力的蛋白质的高度表达而优化的基因。作为实例,考虑受体的结合口袋。同时测试结合口袋内所有残基的所有序列排列的能力将允许进行彻底的探索,从而增加成功的可能性。饱和诱变(其中研究人员试图在受体内的特定位点处生成所有可能的突变)代表了针对这种开发挑战的一种方法。虽然其成本高、耗时且耗力,但它能够将每个变体引入到每个位置。相反,组合诱变(其中几个选定的位置或短DNA段可得到广泛修饰)生成具有偏向呈现的变体的不完全组库。

[0152] 为了加速药物开发流程,具有在可用于测试的正确位置处以预期频率可获得的所需变体的文库(换言之,精确文库)使得能够降低成本以及筛选的周转时间。本文提供了用于合成核酸合成变体文库的方法,其能够以所需的频率精确引入每种期望的变体。对于最终用户来说,这意味着不仅能够彻底对序列空间进行采样,而且能够以有效的方式查询这些假设,从而降低成本和筛选时间。全基因组编辑可以阐明重要的途径,可以检测每个变体和序列排列以获得最佳功能性的文库,并且可以使用数以千计的基因重建整个途径和基因组,以重新改造生物系统以供药物发现。

[0153] 在第一个实例中,药物本身可以使用本文所述的方法进行优化。例如,为了改善抗体的指定功能,设计并合成编码抗体一部分的变异多核苷酸文库。然后可以通过本文所述的过程(例如,PCR诱变之后插入载体中)生成抗体的变异核酸文库。然后在生产细胞系中表达该抗体,并针对增强的活性进行筛选。示例筛选包括检查对抗原的结合亲和力、稳定性或效应物功能(例如,ADCC、补体或凋亡)的调节。用来优化抗体的示例性区域包括但不限于Fc区、Fab区、Fab区的可变区、Fab区的恒定区、重链或轻链的可变域(VH或VL)以及VH或VL的特定互补决定区(CDR)。

[0154] 通过本文所述的方法合成的核酸文库可以在与疾病状态相关的各种细胞中表达。与疾病状态相关的细胞包括细胞系、组织样品、来自对象的原代细胞、从对象扩充的培养细胞或模型系统中的细胞。示例性的模型系统包括但不限于疾病状态的植物和动物模型。

[0155] 为了鉴定与疾病状态的预防、减轻或治疗相关的变异分子,本文所述的变异核酸文库在与疾病状态相关的细胞中表达,或者在可以诱发细胞疾病状态的细胞中表达。在某

些情况下，使用药剂在细胞中诱发疾病状态。用于疾病状态诱发的示例性工具包括但不限于Cre/Lox重组系统、LPS炎症诱发和用来诱发低血糖的链脲佐菌素。与疾病状态相关的细胞可以是来自模型系统的细胞或培养的细胞，以及来自具有特定疾病状况的对象的细胞。示例性疾病状况包括细菌、真菌、病毒、自身免疫性或增生性病症(例如，癌症)。在某些情况下，所述变异核酸文库在模型系统、细胞系或来源于对象的原代细胞中表达，并针对至少一种细胞活性的改变进行筛选。示例性的细胞活性包括但不限于增殖、周期进展、细胞死亡、粘附、迁移、复制、细胞信号传导、能量产生、氧利用、代谢活性和老化、对自由基损伤的响应或其任意组合。

[0156] 基底

[0157] 用作多核苷酸合成表面的装置可以是基底的形式，其包括但不限于均质阵列表面、图案化的阵列表面、通道、珠子、凝胶等。本文提供了包含多个簇的基底，其中每个簇包含多个支持多核苷酸附着和合成的座位。在某些情况下，基底包含均匀的阵列表面。例如，该均匀的阵列表面是均匀的板。如本文所用的术语“座位”是指结构上的离散区域，其提供了对编码单个预定序列的多核苷酸从该表面延伸的支持。在某些情况下，座位在二维表面(例如，基本上为平面的表面)上。在某些情况下，座位在三维表面(例如，孔、微孔、通道或柱杆)上。在某些情况下，座位的表面包含这样的材料，该材料被活化官能化，以附着至少一个核苷酸以供多核苷酸合成，或者优选地，附着相同核苷酸的群体以供多核苷酸群体合成。在某些情况下，多核苷酸是指编码相同核酸序列的多核苷酸群体。在某些情况下，基底的表面包括基底的一个或多个表面。使用所提供的系统和方法在本文所述的文库内合成的多核苷酸的平均错误率通常小于1/1000、小于约1/2000、小于约1/3000或更低，通常没有错误校正。

[0158] 本文提供了支持在共同支持物上的可寻址位置处平行合成具有不同预定序列的多个多核苷酸的表面。在某些情况下，基底为合成超过50、100、200、400、600、800、1000、1200、1400、1600、1800、2,000、5,000、10,000、20,000、50,000、100,000、200,000、300,000、400,000、500,000、600,000、700,000、800,000、900,000、1,000,000、1,200,000、1,400,000、1,600,000、1,800,000、2,000,000、2,500,000、3,000,000、3,500,000、4,000,000、4,500,000、5,000,000、10,000,000个或更多个不同的多核苷酸提供支持。在某些情况下，该表面为合成超过50、100、200、400、600、800、1000、1200、1400、1600、1800、2,000、5,000、10,000、20,000、50,000、100,000、200,000、300,000、400,000、500,000、600,000、700,000、800,000、900,000、1,000,000、1,200,000、1,400,000、1,600,000、1,800,000、2,000,000、2,500,000、3,000,000、3,500,000、4,000,000、4,500,000、5,000,000、10,000,000个或更多个编码不同序列的多核苷酸提供支持。在某些情况下，至少一部分多核苷酸具有相同的序列或被配置为用相同的序列合成。在某些情况下，该基底提供用于增长具有至少80、90、100、120、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500个或更多个碱基的多核苷酸的表面环境。

[0159] 本文提供了在基底的不同座位上合成多核苷酸的方法，其中每个座位支持合成多核苷酸群体。在某些情况下，每个座位支持合成与在另一座位上增长的多核苷酸群体具有不同序列的多核苷酸群体。在某些情况下，每个多核苷酸序列被合成为在用于多核苷酸合成的表面上同一座位簇内的不同座位上具有1、2、3、4、5、6、7、8、9或更多的冗余度。在某些

情况下，基底的座位位于多个簇内。在某些情况下，基底包含至少10、500、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、11000、12000、13000、14000、15000、20000、30000、40000、50000个或更多个簇。在某些情况下，基底包含超过2,000、5,000、10,000、100,000、200,000、300,000、400,000、500,000、600,000、700,000、800,000、900,000、1,000,000、1,100,000、1,200,000、1,300,000、1,400,000、1,500,000、1,600,000、1,700,000、1,800,000、1,900,000、2,000,000、300,000、400,000、500,000、600,000、700,000、800,000、900,000、1,000,000、1,200,000、1,400,000、1,600,000、1,800,000、2,000,000、2,500,000、3,000,000、3,500,000、4,000,000、4,500,000、5,000,000或10,000,000个或更多个不同的座位。在某些情况下，基底包含约10,000个不同的座位。单簇内的座位的量在不同情况下是不同的。在某些情况下，每个簇包含1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、120、130、150、200、300、400、500个或更多个座位。在某些情况下，每个簇包含约50-500个座位。在某些情况下，每个簇包含约100-200个座位。在某些情况下，每个簇包含约100-150个座位。在某些情况下，每个簇包含约109、121、130或137个座位。在某些情况下，每个簇包含约19、20、61、64个或更多个座位。备选地或组合地，多核苷酸合成在均匀的阵列表面上进行。

[0160] 在某些情况下，在基底上合成的不同多核苷酸的数目取决于基底中可用的不同座位的数目。在某些情况下，基底的簇或表面内的座位密度至少是或约为1、10、25、50、65、75、100、130、150、175、200、300、400、500、1,000个或更多个座位/mm^2。在某些情况下，基底包含10-500、25-400、50-500、100-500、150-500、10-250、50-250、10-200或50-200mm^2。在某些情况下，簇或表面内两个相邻座位的中心之间的距离为约10-500、约10-200或约10-100um。在某些情况下，相邻座位的两个中心之间的距离大于约10、20、30、40、50、60、70、80、90或100um。在某些情况下，两个相邻座位的中心之间的距离小于约200、150、100、80、70、60、40、30、20或10um。在某些情况下，每个座位具有约0.5、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90或100um的宽度。在某些情况下，每个座位具有约0.5-100、0.5-50、10-75或0.5-50um的宽度。

[0161] 在某些情况下，基底内的簇的密度是至少或约1个簇/100mm^2、1个簇/10mm^2、1个簇/5mm^2、1个簇/4mm^2、1个簇/3mm^2、1个簇/2mm^2、1个簇/1mm^2、2个簇/1mm^2、3个簇/1mm^2、4个簇/1mm^2、5个簇/1mm^2、10个簇/1mm^2、50个簇/1mm^2或更高。在某些情况下，基底包含约1个簇/10mm^2至约10个簇/1mm^2。在某些情况下，两个相邻簇的中心之间的距离至少为或约为50、100、200、500、1000、2000或5000um。在某些情况下，两个相邻簇的中心之间的距离约为50-100、50-200、50-300、50-500和100-2000um。在某些情况下，两个相邻簇的中心之间的距离约为0.05-50、0.05-10、0.05-5、0.05-4、0.05-3、0.05-2、0.1-10、0.2-10、0.3-10、0.4-10、0.5-10、0.5-5或0.5-2mm。在某些情况下，每个簇具有约0.5至约2、约0.5至约1或约1至约2mm的横截面。在某些情况下，每个簇具有约0.5、0.6、0.7、0.8、0.9、1、1.1、1.2、1.3、1.4、1.5、1.6、1.7、1.8、1.9或2mm的横截面。在某些情况下，每个簇具有约0.5、0.6、0.7、0.8、0.9、1、1.1、1.15、1.2、1.3、1.4、1.5、1.6、1.7、1.8、1.9或2mm的内部横截面。

[0162] 在某些情况下，基底是大约标准96孔板的大小，例如，约100至约200mm乘以约50至约150mm。在某些情况下，基底具有小于或等于约1000、500、450、400、300、250、200、150、100或50mm的直径。在某些情况下，基底的直径约为25-1000、25-800、25-600、25-500、25-400、

25-300或25-200mm。在某些情况下，基底具有至少约100、200、500、1,000、2,000、5,000、10,000、12,000、15,000、20,000、30,000、40,000、50,000mm^2或更大的平面表面积。在某些情况下，基底的厚度约为50-2000、50-1000、100-1000、200-1000或250-1000mm。

[0163] 表面材料

[0164] 本文提供的基底、装置和反应器由适合于本文描述的方法、组合物和系统的任何种类的材料制成。在某些情况下，将基底材料制造成表现出低水平的核苷酸结合。在某些情况下，修饰基底材料以生成表现出高水平的核苷酸结合的不同表面。在某些情况下，基底材料对可见光和/或紫外线是透明的。在某些情况下，基底材料具有足够的导电性，例如，能够跨整个基底或其一部分形成均匀的电场。在某些情况下，导电材料在电气上接地。在某些情况下，该基底是导热的或隔热的。在某些情况下，该材料是耐化学的且耐热的，以支持化学或生化反应，例如多核苷酸合成反应过程。在某些情况下，基底包含柔性材料。对于柔性材料而言，材料可包括但不限于：改性及未改性的尼龙、硝酸纤维素、聚丙烯等。在某些情况下，基底包含刚性材料。对于刚性材料而言，材料可包括但不限于：玻璃；熔融石英；硅、塑料(例如，聚四氟乙烯、聚丙烯、聚苯乙烯、聚碳酸脂，及其混合物等)；金属(例如，金、铂等)。基底、固体支持物或反应器可由选自硅、聚苯乙烯、琼脂糖、葡聚糖、纤维素聚合物、聚丙烯酰胺、聚二甲基硅氧烷(PDMS)和玻璃的材料制成。基底/固体支持物或者其中的微结构、反应器可使用本文所示材料或本领域中已知的任何其他适当材料的组合制成。

[0165] 表面架构

[0166] 本文提供了用于本文描述的方法、组合物和系统的基底，其中所述基底具有适合于本文描述的方法、组合物和系统的表面架构。

在某些情况下，基底包含凸起和/或凹陷特征。具有这类特征的一个益处是用来支持多核苷酸合成的表面积增大。在某些情况下，具有凸起和/或凹陷特征的基底被称为三维基底。在某些情况下，三维基底包含一个或多个通道。在某些情况下，一个或多个座位包含通道。在某些情况下，通道可通过沉积装置如材料沉积装置进行试剂沉积。在某些情况下，试剂和/或流体收集在与一个或多个通道流体连通的较大的孔中。例如，基底包含与具有簇的多个座位相对应的多个通道，并且所述多个通道与该簇的一个孔流体连通。在一些方法中，多核苷酸文库在簇的多个座位中合成。

[0167] 本文提供用于本文描述的方法、组合物和系统的基底，其中该基底被配置用于多核苷酸合成。在某些情况下，该结构被配制为允许用于表面上多核苷酸合成的受控的流动和质量传递路径。在某些情况下，基底的构造允许在多核苷酸合成过程中质量传递路径、化学暴露次数和/或洗涤功效的受控且均匀的分布。在某些情况下，基底的构造允许增加扫描效率，例如通过提供足以用于增长多核苷酸的体积，使得由增长的多核苷酸所排除的体积占可用于或适合于增长多核苷酸的初始可用体积的不超过50％、45％、40％、35％、30％、25％、20％、15％、14％、13％、12％、11％、10％、9％、8％、7％、6％、5％、4％、3％、2％、1％或更少。在某些情况下，三维结构允许流体的受管控的流动，从而允许化学暴露的快速交换。

[0168] 本文提供了用于本文描述的方法、组合物和系统的基底，其中所述基底包含适合于本文描述的方法、组合物和系统的结构。在某些情况下，通过物理结构实现隔离。在某些情况下，通过表面的差异官能化以生成用于多核苷酸合成的活化和钝化区域来实现隔离。在某些情况下，差异官能化通过在整个基底表面上交替呈现疏水性，从而造成可引起沉积

的试剂结珠或润湿的水接触角效应来实现。采用较大的结构可减少飞溅和邻近斑点的试剂对不同的多核苷酸合成位置的交叉污染。在某些情况下，使用装置如材料沉积装置将试剂沉积到不同的多核苷酸合成位置。具有三维特征的基底以允许以低错误率(例如，小于约1:500、1:1000、1:1500、1:2,000;1:3,000;1:5,000;或1:10,000)合成大量多核苷酸(例如，超过约10,000个)的方式配置。在某些情况下，基底包含密度为约或大于约1、5、10、20、30、40、50、60、70、80、100、110、120、130、140、150、160、170、180、190、200、300、400或500个特征/mm^2的特征。

[0169] 基底的孔可具有与基底的另一个孔相同或不同的宽度、高度和/或容积。基底的通道可具有与基底的另一个通道相同或不同的宽度、高度和/或容积。在某些情况下，簇的直径或包含簇的孔的直径或两者约为0.05-50、0.05-10、0.05-5、0.05-4、0.05-3、0.05-2、0.05-1、0.05-0.5、0.05-0.1、0.1-10、0.2-10、0.3-10、0.4-10、0.5-10、0.5-5或0.5-2mm。在某些情况下，簇或孔或两者的直径小于或约为5、4、3、2、1、0.5、0.1、0.09、0.08、0.07、0.06或0.05mm。在某些情况下，簇或孔或两者的直径为约1.0mm至1.3mm。在某些情况下，簇或孔或两者的直径约为1.150mm。在某些情况下，簇或孔或两者的直径约为0.08mm。簇的直径是指二维或三维基底内的簇。

[0170] 在某些情况下，孔的高度约为20-1000、50-1000、100-1000、200-1000、300-1000、400-1000或500-1000um。在某些情况下，孔的高度小于约1000、900、800、700或600um。

[0171] 在某些情况下，基底包含与簇内的多个座位相对应的多个通道，其中通道的高度或深度为5-500、5-400、5-300、5-200、5-100、5-50或10-50um。在某些情况下，通道的高度小于100、80、60、40或20um。

[0172] 在某些情况下，通道、座位(例如，在基本上平坦的基底中)或通道和座位两者(例如，在其中座位对应于通道的三维基底中)的直径约为1-1000、1-500、1-200、1-100、5-100或10-100um，例如约90、80、70、60、50、40、30、20或10um。在某些情况下，通道、座位或通道和座位两者的直径小于约100、90、80、70、60、50、40、30、20或10um。在某些情况下，两个相邻通道、座位或通道和座位的中心之间的距离约为1-500、1-200、1-100、5-200、5-100、5-50或5-30，例如约20um。

[0173] 表面修饰

[0174] 本文提供了用于在表面上合成多核苷酸的方法，其中该表面包含各种表面修饰。在某些情况下，采用表面修饰通过加成工艺或减成工艺对表面进行化学和/或物理改变，以改变基底表面或基底表面的选定位点或区域的一种或多种化学和/或物理性质。例如，表面修饰包括但不限于：(1)改变表面的润湿性质；(2)对表面进行官能化，即，提供、修改或取代表面官能团；(3)对表面进行去官能化，即，移除表面官能团；(4)以其他方式例如通过刻蚀来改变表面的化学组成；(5)增大或减小表面粗糙度；(6)在表面上提供涂层，例如，展现出与表面的润湿性质不同的润湿性质的涂层；和/或(7)在表面上沉积微粒。

[0175] 在某些情况下，在表面顶部添加化学层(被称为粘附促进剂)有利于基底表面上的座位的结构化图案化。用于施加粘附促进剂的示例性表面包括但不限于玻璃、硅、二氧化硅和氮化硅。在某些情况下，该粘附促进剂是具有高表面能的化学品。在某些情况下，在基底的表面上沉积第二化学层。在某些情况下，第二化学层具有低表面能。在某些情况下，涂覆在表面上的化学层的表面能支持小液滴在表面上的定位。根据所选择的图案化布置，座位

的接近度和/或在座位处的流体接触面积是可改变的。

[0176] 在某些情况下，(例如为了多核苷酸合成)核酸或其他部分所沉积到的基底表面或解析座位是光滑的或基本上为平面的(例如，二维的)，或者具有不规则性，诸如凸起或凹陷特征(例如，三维特征)。在某些情况下，用一个或多个不同的化合物层来修饰基底表面。感兴趣的此类修饰层包括但不限于无机层和有机层，如金属、金属氧化物，聚合物、有机小分子等。

[0177] 在某些情况下，使用增大和/或减小表面能的一个或多个部分对基底的解析座位进行官能化。在某些情况下，部分是化学惰性的。在某些情况下，部分被配置为支持所需的化学反应，例如在多核苷酸合成反应中的一个或多个过程。表面的表面能或疏水性是决定核苷酸附着到该表面上的亲和力的因素。在某些情况下，基底官能化方法包括：(a)提供具有包含二氧化硅的表面的基底；和(b)使用本文所述的或本领域已知的合适的硅烷化剂(例如，有机官能烷氧基硅烷分子)对所述表面进行硅烷化。方法和官能化剂在通过引用整体并入本文的美国专利5474796中有描述。

[0178] 在某些情况下，基底表面通常经由存在于基底表面上的反应性亲水部分，在有效地将硅烷偶联至基底表面的反应条件下，使基底表面与含有硅烷混合物的衍生化组合物相接触来进行官能化。硅烷化一般通过使用有机官能烷氧基硅烷分子自装配来覆盖表面。还可使用本领域当前已知的多种硅氧烷官能化试剂，例如用于降低或增大表面能。有机官能烷氧基硅烷根据其有机官能来分类。

[0179] 多核苷酸合成

[0180] 用于多核苷酸合成的本公开的方法可包括涉及亚磷酰胺化学法的过程。在某些情况下，多核苷酸合成包括将碱基与亚磷酰胺偶联。多核苷酸合成可包括通过在偶联条件下沉积亚磷酰胺来偶联碱基，其中相同的碱基任选地与亚磷酰胺沉积超过一次，即双偶联。多核苷酸合成可包括未反应位点的加帽。在某些情况下，加帽是可选的。多核苷酸合成还可包括氧化或氧化步骤或多个氧化步骤。多核苷酸合成可包括解封闭、脱三苯甲基化和硫化。在某些情况下，多核苷酸合成包括氧化或硫化。在某些情况下，在多核苷酸合成反应期间的一个步骤或每个步骤之间，例如使用四唑或乙腈来洗涤所述装置。亚磷酰胺合成方法中任一步骤的时间范围可小于约2min、1min、50sec、40sec、30sec、20sec和10sec。

[0181] 使用亚磷酰胺方法的多核苷酸合成可包括随后将亚磷酰胺构件(例如，核苷亚磷酰胺)添加至增长的多核苷酸链以形成亚磷酸三酯键。亚磷酰胺多核苷酸合成沿3’至5’方向进行。亚磷酰胺多核苷酸合成允许在每个合成循环中将一个核苷酸受控添加至增长的核酸链。在某些情况下，每个合成循环包括偶联步骤。亚磷酰胺偶联包括在活化的核苷亚磷酰胺与(例如通过连接体)结合至基底的核苷之间形成亚磷酸三酯键。在某些情况下，将活化的核苷亚磷酰胺提供给装置。在某些情况下，将核苷亚磷酰胺与活化剂一起提供给装置。在某些情况下，核苷亚磷酰胺以相对于与基底结合的核苷1.5、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、50、60、70、80、90、100倍或更多倍的过量来提供给装置。在某些情况下，核苷亚磷酰胺的添加在无水环境中(例如，在无水乙腈中)进行。添加核苷亚磷酰胺后，任选地洗涤该装置。在某些情况下，偶联步骤额外重复一次或多次，任选地在向基底添加核苷亚磷酰胺之间进行洗涤步骤。在某些情况下，本文使用的多核苷酸合成方法包括1、2、3个或更多个连续的偶联步骤。在许多情况下，在偶联之前，与装置结合

的核苷通过去除保护基团来脱保护，其中该保护基团起到防止聚合的作用。常见的保护基团为4,4'-二甲氧基三苯甲基(DMT)。

[0182] 偶联后，亚磷酰胺多核苷酸合成方法任选地包括加帽步骤。在加帽步骤中，用加帽剂处理增长的多核苷酸。加帽步骤可用来在偶联后封闭未反应的与基底结合的5'-OH基团以防止进一步链延伸，从而防止形成具有内部碱基缺失的多核苷酸。此外，用1H-四唑活化的亚磷酰胺可以在很小的程度上与鸟苷的O6位置反应。不受理论的束缚，在用I2/水氧化后，该副产物(可能经由O6-N7迁移)可经历脱嘌呤。无嘌呤位点可终止在多核苷酸的最终脱保护过程中被切割，从而降低全长产物的产率。O6修饰可通过在用I2/水氧化之前用加帽试剂处理而去除。在某些情况下，与没有加帽的合成相比，在多核苷酸合成过程中包括加帽步骤会降低错误率。作为实例，加帽步骤包括用乙酸酐和1-甲基咪唑的混合物处理与基底结合的多核苷酸。在加帽步骤之后，任选地洗涤所述装置。

[0183] 在某些情况下，在添加核苷亚磷酰胺之后，并且任选地在加帽和一个或多个洗涤步骤之后，对与装置结合的增长的核酸进行氧化。氧化步骤包括将亚磷酸三酯氧化成四配位的磷酸三酯——天然存在的磷酸二酯核苷间连接的受保护的前体。在某些情况下，增长的多核苷酸的氧化通过任选地在弱碱(例如，吡啶、二甲基吡啶、三甲吡啶)的存在下用碘和水处理来实现。氧化可在无水条件下采用例如叔丁基过氧化氢或(1S)-(+)-(10-樟脑磺酰基)-氧杂叶丙啶(CSO)进行。在一些方法中，在氧化之后进行加帽步骤。第二个加帽步骤允许装置干燥，因为可能持续存在的来自氧化的残余水可以抑制随后的偶联。氧化后，任选地洗涤装置和增长的多核苷酸。在某些情况下，氧化步骤用硫化步骤来代替，以获得多核苷酸硫代磷酸，其中任何加帽步骤均可在硫化之后进行。许多试剂能够进行有效的硫转移，包括但不限于3-(二甲基氨基亚甲基)氨基)-3H-1,2,4-二噻唑-3-硫酮、DDTT、3H-1,2-苯并二噻戊环-3-酮1,1-二氧化物(也被称为Beaucage试剂)和N,N,N'N'-四乙基秋兰姆二硫化物(TETD)。

[0184] 为了使后续核苷掺入循环通过偶联而发生，除去与装置结合的增长的多核苷酸的受保护的5'末端，使得伯羟基与下一个核苷亚磷酰胺反应。在某些情况下，保护基团为DMT，并且用在二氯甲烷中的三氯乙酸进行解封闭。进行延长时间的脱三苯甲基化或者使用比推荐的酸溶液更强的酸溶液进行脱三苯甲基化可导致与固体支持物结合的多核苷酸的脱嘌呤增加，并因此降低了所需全长产物的产率。本文所述的本公开的方法和组合物提供了受控的解封闭条件，从而限制不希望的脱嘌呤反应。在某些情况下，与装置结合的多核苷酸在解封闭后洗涤。在某些情况下，解封闭后的有效洗涤有助于以低错误率合成多核苷酸。

[0185] 多核苷酸合成方法一般包括一系列迭代的以下步骤:将受保护的单体施加至活化官能化的表面(例如，座位)以与活化的表面、连接体或与预先脱保护的单体连接;使所施加的单体脱保护，使其可与随后施加的受保护的单体反应;以及施加另一种受保护的单体以供连接。一个或多个中间步骤包括氧化或硫化。在某些情况下，在一个或全部步骤之前或之后有一个或多个洗涤步骤。

[0186] 基于亚磷酰胺的多核苷酸合成方法包括一系列化学步骤。在某些情况下，合成方法的一个或多个步骤涉及试剂循环，其中该方法的一个或多个步骤包括向该装置施加对该步骤有用的试剂。例如，试剂通过一系列液相沉积和真空干燥步骤进行循环。对于包含诸如孔、微孔、通道等三维特征的基底，试剂任选地经由孔和/或通道穿过该装置的一个或多个

区域。

[0187] 本文所述的方法和系统涉及用于合成多核苷酸的多核苷酸合成装置。该合成可以是平行的。例如，可以平行合成至少或约至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、30、35、40、45、50、100、150、200、250、300、350、400、450、500、550、600、650、700、750、800、850、900、1000、10000、50000、75000、100000个或更多个多核苷酸。可以平行合成的多核苷酸的总数可以是2-100000、3-50000、4-10000、5-1000、6-900、7-850、8-800、9-750、10-700、11-650、12-600、13-550、14-500、15-450、16-400、17-350、18-300、19-250、20-200、21-150、22-100、23-50、24-45、25-40、30-35个。本领域技术人员知晓，平行合成的多核苷酸的总数可处于由这些值中的任何值所限定的任何范围内，例如25-100。平行合成的多核苷酸的总数可处于由充当范围端点的任何值所限定的任何范围内。在装置内合成的多核苷酸的总摩尔质量或每种多核苷酸的摩尔质量可以是至少或至少约10、20、30、40、50、100、250、500、750、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、25000、50000、75000、100000皮摩尔或更大。每种多核苷酸的长度或装置内多核苷酸的平均长度可以是至少或约至少10、15、20、25、30、35、40、45、50、100、150、200、300、400、500个或更多个核苷酸。每种多核苷酸的长度或装置内多核苷酸的平均长度可以是至多或大约至多500、400、300、200、150、100、50、45、35、30、25、20、19、18、17、16、15、14、13、12、11、10个或更少的核苷酸。每种多核苷酸的长度或装置内多核苷酸的平均长度可以处于10-500、9-400、11-300、12-200、13-150、14-100、15-50、16-45、17-40、18-35、19-25之间。本领域技术人员知晓，每种多核苷酸的长度或装置内多核苷酸的平均长度可处于由这些值中的任何值所限定的任何范围内，例如100-300。每种多核苷酸的长度或装置内多核苷酸的平均长度可处于由充当范围端点的任何值所限定的任何范围内。

[0188] 本文提供的在表面上合成多核苷酸的方法允许以较快的速度合成。作为实例，每小时合成至少3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、55、60、70、80、90、100、125、150、175、200个或更多个核苷酸。核苷酸包括腺嘌呤、鸟嘌呤、胸腺嘧啶、胞嘧啶、尿苷构件，或其类似物/修饰形式。在某些情况下，多核苷酸文库在基底上平行合成。例如，包含大约或至少约100、1,000、10,000、30,000、75,000、100,000、1,000,000、2,000,000、3,000,000、4,000,000或5,000,000个解析座位的装置能够支持合成至少相同数目的不同的多核苷酸，其中编码不同序列的多核苷酸在解析座位上合成。在某些情况下，在少于约三个月、两个月、一个月、三周、15天、14天、13天、12天、11天、10天、9天、8天、7天、6天、5天、4天、3天、2天、24小时或更短的时间内，以本文所述的低错误率在装置上合成多核苷酸文库。在某些情况下，使用本文所述的基底和方法从以低错误率合成的多核苷酸文库装配的较大核酸在少于约三个月、两个月、一个月、三周、15天、14天、13天、12天、11天、10天、9天、8天、7天、6天、5天、4天、3天、2天、24小时或更短的时间内制备。

[0189] 在某些情况下，本文所述的方法提供了生成包含在多个密码子位点处不同的变异核酸的核酸文库。在某些情况下，核酸可具有1个位点、2个位点、3个位点、4个位点、5个位点、6个位点、7个位点、8个位点、9个位点、10个位点、11个位点、12个位点、13个位点、14个位点、15个位点、16个位点、17个位点、18个位点、19个位点、20个位点、30个位点、40个位点、50个位点或更多个变异密码子位点。

[0190] 在某些情况下，变异密码子位点的一个或多个位点可以是相邻的。在某些情况下，变异密码子位点的一个或多个位点可以是不相邻的，并且由1、2、3、4、5、6、7、8、9、10个或更多个密码子隔开。

[0191] 在某些情况下，核酸可包含变异密码子位点的多个位点，其中所有变异密码子位点彼此相邻，形成一段变异密码子位点。在某些情况下，核酸可包含变异密码子位点的多个位点，其中所述变异密码子位点彼此均不相邻。在某些情况下，核酸可包含变异密码子位点的多个位点，其中一些变异密码子位点彼此相邻，形成一段变异密码子位点，而一些变异密码子位点彼此不相邻。

[0192] 参见附图，图7示出了用于从较短核酸合成核酸(例如，基因)的示例性处理工作流程。该工作流程大致分为以下阶段：(1)从头合成单链核酸文库，(2)连接核酸以形成更大的片段，(3)错误校正，(4)质量控制，以及(5)运输。在从头合成之前，预先选择预期的核酸序列或一组核酸序列。例如，预先选择一组基因用于生成。

[0193] 一旦选择用于生成的大核酸，则针对从头合成来设计预定的核酸文库。用于生成高密度多核苷酸阵列的各种合适的方法是已知的。在该工作流程示例中，提供了装置表面层。在该示例中，改变表面的化学性质，以改进多核苷酸合成过程。生成低表面能区域以排斥液体，同时生成高表面能区域以吸引液体。表面本身可以是平面表面的形式或者包含形状的变化，例如增加表面积的突起或微孔。在该工作流程示例中，如在通过引用整体并入本文的国际专利申请公开WO/2015/021080中所公开的，所选择的高表面能分子发挥支持DNA化学过程的双重功能。

[0194] 多核苷酸阵列的原位制备在固体支持物上进行701，并利用单核苷酸延伸过程平行延伸多个寡聚物。沉积装置如材料沉积装置被设计为以逐步方式释放试剂，使得多个多核苷酸平行地一次延伸一个残基，以生成具有预定核酸序列的寡聚物702。在某些情况下，多核苷酸在该阶段从表面上切下。切割包括例如采用氨或甲胺的气体切割。

[0195] 将生成的多核苷酸文库放置于反应室中。在该示例性工作流程中，反应室(也被称为“纳米反应器”)为硅涂覆的孔，其含有PCR试剂并下降到多核苷酸文库上703。在多核苷酸密封704之前或之后，添加试剂以从基底释放多核苷酸。在该示例性工作流程中，多核苷酸在纳米反应器密封705之后释放。一旦释放，单链多核苷酸的片段即发生杂交，以跨越整个长程DNA序列。部分杂交705是可能的，因为每个合成的多核苷酸被设计为具有与池中的至少一个其他多核苷酸重叠的一小部分。

[0196] 杂交后，开始PCA反应。在聚合酶循环过程中，多核苷酸与互补片段退火，并且用聚合酶补平缺口。根据哪些多核苷酸彼此发现，每个循环随机增加各个片段的长度。片段之间的互补性允许形成完整的大跨度的双链DNA 706。

[0197] 在PCA完成之后，将纳米反应器与装置分开707，并定位成与具有PCR引物的装置相互作用708。密封后，纳米反应器经历PCR 709并扩增较大的核酸。在PCR之后710，打开纳米室711，添加错误校正试剂712，将腔室密封713并进行错误校正反应，以从双链PCR扩增产物中去除具有较差互补性的错配碱基对和/或链714。打开并分离纳米反应器715。错误校正产物接下来经历另外的处理步骤，如PCR和分子条形码化，随后包装722以供运输723。

[0198] 在某些情况下，采取质量控制措施。在错误校正之后，质量控制步骤包括例如与具有用于扩增错误校正产物的测序引物的晶片进行相互作用716，将晶片密封到含有错误校

正扩增产物的腔室中717，并进行另一轮扩增718。打开纳米反应器719，合并产物720并进行测序721。在得到可接受的质量控制结果之后，包装的产物722准许运输723。

[0199] 在某些情况下，通过诸如图7中的工作流程生成的核酸使用本文公开的重叠引物进行诱变。在某些情况下，通过在固体支持物上原位制备来生成引物文库，并利用单核苷酸延伸过程平行延伸多个寡聚物。沉积装置如材料沉积装置被设计为以逐步方式释放试剂，使得多个多核苷酸平行地一次延伸一个残基，以生成具有预定核酸序列的寡聚物702。

[0200] 计算机系统

[0201] 本文所述的任何系统均可以可操作地连接至计算机，并且可以本地或远程地通过计算机进行自动化。在各种情况下，本公开的方法和系统可进一步包括计算机系统上的软件程序及其使用。因此，对于分配/抽真空/再填充功能的同步(如编排和同步材料沉积装置运动、分配动作和真空致动)的计算机化控制处于本公开内容的范围内。计算机系统可被编程为在用户指定的碱基序列与材料沉积装置的位置之间接合，以将正确的试剂递送至基底的指定区域。

[0202] 图8中示出的计算机系统800可被理解为能够从介质811和/或网络端口805读取指令的逻辑设备，其可任选地连接至具有固定介质812的服务器809。诸如图8示出的系统可包括CPU 801、磁盘驱动器803、可选的输入设备如键盘815和/或鼠标816以及可选的监视器807。可通过示出的通信媒介实现与本地或远程位置处的服务器的数据通信。通信媒介可包括传输和/或接收数据的任何手段。例如，通信媒介可以是网络连接、无线连接或因特网连接。这样的连接可提供经由万维网的通信。可以设想有关本公开的数据可经过这样的网络或连接而传输，以便由图8所示的用户方822接收和/或审阅。

[0203] 如图9所示，高速缓冲存储器904可连接至或并入处理器902，以提供由处理器902新近或频繁使用的指令或数据的高速存储器。处理器902通过处理器总线908连接至北桥906。北桥906通过存储器总线912连接至随机存取存储器(RAM)910，并管理处理器902对RAM 910的访问。北桥906还通过芯片集总线916连接至南桥914。南桥914又连接至外围总线918。外围总线可以是例如PCI、PCI-X、PCI Express或其他外围总线。北桥和南桥通常被称为处理器芯片集，并管理在处理器、RAM与外围总线918上的外围组件之间的数据传送。在一些备选的架构中，北桥的功能性可以并入处理器中，而不是使用单独的北桥芯片。在某些情况下，系统900可包括附接至外围总线918的加速器卡922。加速器可包括现场可编程门阵列(FPGA)或用于加速某个处理的其他硬件。例如，加速器可用于适应性数据重建或用来评价在扩展集处理中使用的代数表达式。

[0204] 软件和数据存储在外部存储器924中，并可加载至RAM 910和/或高速缓冲存储器904中，以供处理器使用。系统900包括用于管理系统资源的操作系统；操作系统的非限制性实例包括：Linux、WindowsTM、MACOSTM、BlackBerry OSTM、iOSTM和其他功能上等效的操作系统，以及在操作系统顶部运行的、用于根据本公开的示例情况管理数据存储和优化的应用软件。在该实例中，系统900还包括与外围总线连接的网络接口卡(NIC)920和921，以提供与外部存储如网络附加存储(NAS)和可用于分布式并行处理的其他计算机系统的网络接口。

[0205] 图10是显示了具有多个计算机系统1002a和1002b、多个蜂窝电话和个人数据助理1002c以及网络附加存储(NAS)1004a和1004b的网络1000的示图。在示例实例中，系统

1002a、1002b和1002c可管理数据存储并优化对存储在网络附加存储(NAS)1004a和1004b中的数据的数据访问。数学模型可用于该数据，并使用跨计算机系统1002a和1002b和蜂窝电话以及个人数据助理系统1002c的分布式并行处理进行评价。计算机系统1002a和1002b和蜂窝电话以及个人数据助理系统1002c也可提供对存储在网络附加存储(NAS)1004a和1004b中的数据的适应性数据重建的并行处理。图10仅示出了一个实例，而多种多样的其他计算机架构和系统可与本公开的多个实例一起使用。例如，刀片服务器可以用来提供并行处理。处理器刀片可通过背板连接，以提供并行处理。存储还可通过单独的网络接口连接至背板或作为网络附加存储(NAS)。在一些示例实例中，处理器可维持单独的存储空间，并通过网络接口、背板或其他连接器传输数据以便由其他处理器并行处理。在其他情况下，部分或全部处理器可使用共享的虚拟地址存储空间。

[0206] 图11是根据示例情况使用共享虚拟地址存储空间的多处理器计算机系统的框图。该系统包括可访问共享的存储器子系统1104的多个处理器1102a-f。该系统中在存储器子系统1104中并入多个可编程硬件存储算法处理器(MAP)1106a-f。每个MAP 1106a-f可以包含存储器1108a-f以及一个或多个现场可编程门阵列(FPGA)1110a-f。MAP提供可配置的功能单元，并且可以向FPGA 1110a-f提供特定算法或算法的部分，以供与相应的处理器密切协同地进行处理。例如，在示例情况中，MAP可用来评价与数据模型相关的代数表达式以及用来进行适应性数据重建。在该示例中，每个MAP可被用于这些目的的所有处理器全局访问。在一种配置中，每个MAP可使用直接存储器访问(DMA)来访问相关联的存储器1108a-f，使其独立于且异步于各自的微处理器1102a-f而执行任务。在这一配置中，MAP可将结果直接馈送至另一MAP以用于流水处理和并行执行算法。

[0207] 以上计算机架构和系统仅为实例，并且多种多样的其他计算机、蜂窝电话和个人数据助理架构和系统可与示例实例结合使用，包括使用通用处理器、协处理器、FPGA和其他可编程逻辑设备、芯片上系统(SOC)、专用集成电路(ASIC)和其他处理和逻辑元件的任何组合的系统。在某些情况下，全部或部分计算机系统可用软件或硬件来实现。任何种类的数据存储介质可与示例实例结合使用，包括随机存取存储器、硬盘驱动器、闪速存储器、磁带驱动器、磁盘阵列、网络附加存储(NAS)和其他的本地或分布式数据存储设备和系统。

[0208] 在示例性情况中，计算机系统可使用在任何上述或其他计算机架构和系统上执行的软件模块来实现。在其他实例中，该系统的功能可部分或完全地在固件、可编程逻辑设备如图9提到的现场可编程门阵列(FPGA)、芯片上系统(SOC)、专用集成电路(ASIC)或其他处理和逻辑元件中实现。例如，集处理器(Set Processor)和优化器可通过使用硬件加速器卡如图9所示的加速器卡922用硬件加速方式实现。

[0209] 阐述以下实施例是为了向本领域技术人员更清楚地说明本文所公开的实施方案的原理和实践，而不应解释为限制任何请求保护的实施方案的范围。除非另有说明，否则所有份数和百分比均以重量计。

实施例

[0210] 给出以下实施例是为了说明本公开的多个实施方案的目的，而并非意图以任何方式限制本发明。这些实施例以及目前代表优选实施方案的本文所述方法是示例性的，而非旨在限制本公开的范围。本领域技术人员将会想到其变化以及包含在由权利要求的范围所限定的本公开的精神之内的其他用途。

[0211] 实施例1:装置表面的官能化

[0212] 对装置进行官能化以支持多核苷酸文库的附接和合成。首先使用包含90%H_2SO_4和10%H_2O_2的食人鱼溶液(piranha solution)将装置表面润湿清洗20分钟。将该装置在含有去离子水的数个烧杯中漂洗,在去离子水鹅颈旋塞下保持5min,并用N_2干燥。随后将该装置在NH_4OH(1:100;3mL:300mL)中浸泡5min,使用手持式喷枪(handgun)用去离子水漂洗,在连续三个含有去离子水的烧杯中各浸泡1min,然后再使用手持式喷枪用去离子水漂洗。然后通过将装置表面暴露于O_2来等离子体清洗该装置。使用SAMCO PC-300仪器在下游模式下以250瓦进行O_2等离子体蚀刻1min。

[0213] 使用具有以下参数的YES-1224P汽相沉积烘箱系统,用包含N-(3-三乙氧基甲硅烷基丙基)-4-羟基丁酰胺的溶液对清洁的装置表面进行活化官能化:0.5至1托,60min,70℃,135℃汽化器。使用Brewer Science 200X旋涂仪对装置表面进行抗蚀剂涂覆。将SPR™ 3612光致抗蚀剂以2500rpm旋涂在装置上40sec。该装置在Brewer热板上以90℃预烘30min。使用Karl Suss MA6掩模对准仪对装置进行光刻。将该装置暴露2.2sec并在MSF 26A中显影1min。剩余的显影剂用手持式喷枪漂洗,并将装置在水中浸泡5min。该装置在烘箱中以100℃烘烤30min,随后使用Nikon L200目视检查光刻缺陷。采用预清除(descum)工艺利用SAMCO PC-300仪器以250瓦进行O_2等离子体蚀刻1min来去除残余抗蚀剂。

[0214] 用与10μL轻质矿物油混合的100μL全氟辛基三氯硅烷溶液对装置表面进行钝化官能化。将该装置放置于腔室中,泵送10min,随后关闭通往泵的阀门并静置10min。使该腔室排气。该装置通过在70℃下在500mL NMP中进行两次5min浸泡并同时以最大功率(在Crest系统上的9)进行超声波处理来剥离抗蚀剂。然后将该装置在室温下在500mL异丙醇中浸泡5min,同时以最大功率进行超声波处理。将该装置浸入300mL的200标准酒精度(proof)的乙醇中并用N_2吹干。活化该官能化表面以充当多核苷酸合成的支持物。

[0215] 实施例2:在寡核苷酸合成装置上合成50-聚体序列

[0216] 将二维寡核苷酸合成装置组装至流动池中,其与流动池(Applied Biosystems(ABI394 DNA合成仪")连接。该二维寡核苷酸合成装置用N-(3-三乙氧基甲硅烷基丙基)-4-羟基丁酰胺(Gelest)均匀地官能化,并用来使用本文所述的多核苷酸合成方法合成50bp的示例性多核苷酸("50-聚体多核苷酸")。

[0217] 所述50-聚体的序列如SEQ ID NO.:2所示。5'AGACAATCAACCATTTGGGGTGGACAGCCTTGACCTCTAGACTTCGGCAT##TTTTTTTTT3'(SEQ ID NO.:2),其中#表示胸苷-琥珀酰基己酰胺CED亚磷酰胺(来自ChemGenes的CLP-2244),它是允许在脱保护过程中从表面上释放多核苷酸的可切割的连接体。

[0218] 根据表3中的方案和ABI合成仪,使用标准DNA合成化学法(偶联、加帽、氧化和解封闭)完成合成。

表3:合成方案

表 3

通用 DNA 合成工艺名称	工艺步骤	时间（sec）
洗涤（乙腈洗液流）	乙腈系统冲洗	4
	乙腈至流动池	23
	N2 系统冲洗	4
	乙腈系统冲洗	4
DNA 碱基添加（亚磷酰胺+活化剂流）	活化剂歧管冲洗	2
	活化剂至流动池	6
	活化剂+亚磷酰胺至流动池	6
	活化剂至流动池	0.5
	活化剂+亚磷酰胺至流动池	5
	活化剂至流动池	0.5
	活化剂+亚磷酰胺至流动池	5
	活化剂至流动池	0.5
	活化剂+亚磷酰胺至流动池	5
	孵育 25sec	25
洗涤（乙腈洗液流）	乙腈系统冲洗	4
	乙腈至流动池	15
	N2 系统冲洗	4
	乙腈系统冲洗	4
DNA 碱基添加（亚磷酰胺+活化剂流）	活化剂歧管冲洗	2
	活化剂至流动池	5
	活化剂+亚磷酰胺至流动池	18
	孵育 25sec	25
洗涤（乙腈洗液流）	乙腈系统冲洗	4

表 3

通用 DNA 合成工艺名称	工艺步骤	时间（sec）
	乙腈至流动池	15
	N2 系统冲洗	4
	乙腈系统冲洗	4
加帽（帽 A+B，1：1，流）	帽 A+B 至流动池	15
洗涤（乙腈洗液流）	乙腈系统冲洗	4
	乙腈至流动池	15
	乙腈系统冲洗	4
氧化（氧化剂流）	氧化剂至流动池	18
洗涤（乙腈洗液流）	乙腈系统冲洗	4
	N2 系统冲洗	4
	乙腈系统冲洗	4
	乙腈至流动池	15
	乙腈系统冲洗	4
	乙腈至流动池	15
	N2 系统冲洗	4
	乙腈系统冲洗	4
	乙腈至流动池	23
	N2 系统冲洗	4
	乙腈系统冲洗	4
解封闭（解封闭流）	解封闭剂至流动池	36
洗涤（乙腈洗液流）	乙腈系统冲洗	4
	N2 系统冲洗	4
	乙腈系统冲洗	4
	乙腈至流动池	18
	N2 系统冲洗	4.13
	乙腈系统冲洗	4.13
	乙腈至流动池	15

[0219] 亚磷酰胺/活化剂组合以类似于本体试剂通过流动池递送的方式进行递送。当在全部时间内保持环境被试剂“润湿”时，不进行干燥步骤。

[0220] 从ABI 394合成仪中去除限流器，以使得能够更快速流动。在没有限流器的情况下，酰胺类(amidites)(在ACN中0.1M)、活化剂(在ACN中的0.25M苯甲酰基硫基四唑(“BTT”；来自GlenResearch的30-3070-xx))和Ox(在20％吡啶、10％水和70％THF中的0.02M I2)的流速大致为～100uL/sec，乙腈(“ACN”)和加帽试剂(帽A和帽B的1:1混合物，其中帽A是在THF/吡啶中的乙酸酐，帽B是在THF中的16％1-甲基咪唑(1-methylimidizole))的流速大致为～200uL/sec，而解封闭剂(在甲苯中的3％二氯乙酸)的流速大致为～300uL/sec(相比之下，在有限流器的情况下，所有试剂的流速均为～50uL/sec)。观测完全排出氧化剂的时间，相应地调整化学品流动时间的时间选择，并在不同的化学品之间引入额外的ACN洗涤。在多核苷酸合成后，将芯片在75psi下在气态氨中脱保护过夜。将五滴水施加到表面上以回收多核苷酸。然后在BioAnalyzer小RNA芯片上分析所回收的多核苷酸。

[0221] 实施例3:在寡核苷酸合成装置上合成100-聚体序列

[0222] 使用实施例2中描述的用于合成50-聚体序列的相同过程，在两个不同的硅芯片上合成100-聚体多核苷酸(“100-聚体多核苷酸”;5’CGGGATCCTTATCGTCATCGTCGTACAGATCCCGACCCATTTGCTGTCCACCAGTCATGCTAGCCATACCATGATGATGATGATGATGAGAACCCCGCAT##TTTTTTTTTT3’，其中#表示胸苷-琥珀酰基己酰胺CED亚磷酰胺(来自ChemGenes的CLP-2244);SEQ ID NO.:3)，第一个用N-(3-三乙氧基甲硅烷基丙基)-4-羟基丁酰胺均匀地官能化，而第二个用11-乙酰氧基十一烷基三乙氧基硅烷和正癸基三乙氧基硅烷的5/95混合物官能化，并在BioAnalyzer仪器上分析从表面提取的多核苷酸。

[0223] 使用下列热循环程序，在50uL PCR混合物(25uL NEB Q5主混合物，2.5uL 10uM正向引物，2.5uL 10uM反向引物，1uL从表面提取的多核苷酸，用水加至50uL)中使用正向引物(5’ATGCGGGGTTCTCATCATC3’;SEQ ID NO.:4)和反向引物(5’CGGGATCCTTATCGTCATCG3’;SEQ ID NO.:5)进一步PCR扩增来自两个芯片的全部十个样品：

98℃，30sec

98℃，10sec;63℃，10sec;72℃，10sec;重复12个循环

72℃，2min

[0224] PCR产物还在BioAnalyzer上运行，在100-聚体位置处显示出尖锐峰。然后，对PCR扩增的样品进行克隆，并进行Sanger测序。

表4总结了从来自芯片1的斑点1-5采集的样品和从来自芯片2的斑点6-10采集的样品的Sanger测序结果。

表4:测序结果

斑点	错误率	循环效率
1	1/763bp	99.87％
2	1/824bp	99.88％
3	1/780bp	99.87％
4	1/429bp	99.77％
5	1/1525bp	99.93％
6	1/1615bp	99.94％
7	1/531bp	99.81％
8	1/1769bp	99.94％

9	1/854bp	99.88%
10	1/1451bp	99.93%

[0225] 因此,合成的多核苷酸的高质量和均匀度在具有不同表面化学的两个芯片上重现。总体上,所测序的100-聚体中有89%是没有错误的完美的序列,对应于262个中的233个。

[0226] 表5总结了从来自斑点1-10的多核苷酸样品中获得的序列的错误特征。

表5:错误特征

样品ID/斑点号	OSA_0046/1	OSA_0047/2	OSA_0048/3	OSA_0049/4	OSA_0050/5	OSA_0051/6	OSA_0052/7	OSA_0053/8	OSA_0054/9	OSA_0055/10
全序列	32	32	32	32	32	32	32	32	32	32
测序质量	25/28	27/27	26/30	21/23	25/26	29/30	27/31	29/31	28/29	25/28
寡核苷酸质量	23/25	25/27	22/26	18/21	24/25	25/29	22/27	28/29	26/28	20/25
ROI匹配计数	2500	2698	2561	2122	2499	2666	2625	2899	2798	2348
ROI突变	2	2	1	3	1	0	2	1	2	1
ROI多碱基缺失	0	0	0	0	0	0	0	0	0	0
ROI小插入	1	0	0	0	0	0	0	0	0	0
ROI单碱基缺失	0	0	0	0	0	0	0	0	0	0
大缺失计数	0	0	1	0	0	1	1	0	0	0
突变:G>A	2	2	1	2	1	0	2	1	2	1
突变:T>C	0	0	0	1	0	0	0	0	0	0
ROI错误计数	3	2	2	3	1	1	3	1	2	1
ROI错误率	错误:~1/834	错误:~1/1350	错误:~1/1282	错误:~1/708	错误:~1/2500	错误:~1/2667	错误:~1/876	错误:~1/2900	错误:~1/1400	错误:~1/2349
ROI减引物错误率	MP错误:~1/763	MP错误:~1/824	MP错误:~1/780	MP错误:~1/429	MP错误:~1/1525	MP错误:~1/1615	MP错误:~1/531	MP错误:~1/1769	MP错误:~1/854	MP错误:~1/1451

[0227] 实施例4:抗体支架的设计

[0228] 为了生成支架,进行了结构分析、重链组库(repertoire)测序分析以及异二聚体高通量测序数据集的特异性分析。每条重链与每个轻链支架相缔合。每个重链支架被分配了5个不同的长CDRH3环选项。每个轻链支架被分配了5个不同的L3支架。重链CDRH3茎选自在个体和V基因区段之间发现的经常观察到的长H3环茎(N末端和C末端上的10个氨基酸)。轻链支架L3选自包含长H3的异二聚体。使用基于来自Protein Data Bank(PDB)和深度测序数据集的信息的直接异二聚体,其中CDR H1、H2、L1、L2、L3和CDRH3茎被固定。然后对各种支架进行格式化,以供在噬菌体上展示以评估表达。

[0229] 结构分析

[0230] 分析了大约2,017个抗体结构,从中观察到22个具有至少25个氨基酸长度的CDRH3的结构。重链包括:IGHV1-69、IGHV3-30、IGHV4-49和IGHV3-21。鉴定出的轻链包括:IGLV3-21、IGKV3-11、IGKV2-28、IGKV1-5、IGLV1-51、IGLV1-44和IGKV1-13。在分析中,多次观察到四个异二聚体组合,包括:IGHV4-59/61-IGLV3-21、IGHV3-21-IGKV2-28、IGHV1-69-IGKV3-11和IGHV1-69-IGKV1-5。序列和结构分析确定了一些结构中的CDRH3内部二硫键,茎中堆积了大体积的侧链,如酪氨酸,从而为长期的H3稳定性提供了支持。还观察到包括β-转角-β片层和"锤头型"子结构域的二级结构。

[0231] 组库分析

[0232] 对通过无偏5'RACE从12名健康对照获得的1,083,875个IgM+/CD27-幼稚B细胞受体(BCR)序列和1,433,011个CD27+序列进行了组库分析。这12名健康对照包括人数相等的男性和女性,由4名白人、4名亚裔和4名西班牙裔个体组成。组库分析表明,人组库中的不到1%包含具有长于21个氨基酸的CDRH3的BCR。在长CDR3子组库中观察到V基因偏倚(bias),其中IGHV1-69、IGHV4-34、IGHV1-18和IGHV1-8在具有长H3环的BCR中显示出优先富集。对于IGHV3-23、IGHV4-59/61、IGHV5-51、IGHV3-48、IGHV3-53/66、IGHV3-15、IGHV3-74、IGHV3-73、IGHV3-72和IGHV2-70观察到背离长环的偏倚。IGHV4-34支架被证明是自反应性的,并且半衰期短。

[0233] 还基于5'RACE参考组库设计了用于长环的可行的N末端和C末端CDRH3支架变异。观察到约81,065个氨基酸长度为22个氨基酸或更长的CDRH3。通过在V基因支架中进行比较,避免了支架特定的H3茎变异,从而允许将支架多样性克隆到多个支架参考中。

[0234] 异二聚体分析

[0235] 对支架进行了异二聚体分析,并测定了支架的变异序列和长度。

[0236] 结构分析

[0237] 使用变异序列的GPCR支架进行了结构分析,并测定了长度。

[0238] 实施例5:GPCR抗体文库的生成

[0239] 基于GPCR-配体相互作用表面和支架排列,设计并从头合成了文库。参见实施例4。关于图5,为重链503可变域设计了10个变异序列,为重链互补决定区3 507设计了237个变异序列,并为轻链513可变域设计了44个变异序列。按照类似于实施例1-3中所述的方法,将这些片段合成为三个片段,详见图6。

[0240] 从头合成后,为重链602可变域生成了10个变异序列,为重链互补决定区3 604生成了236个变异序列,并为包含轻链可变域606和CDRL3的区域设计了43个变异序列,其中9个变体是为轻链可变域设计的。这产生了具有约10^5多样性(10x236x43)的文库。使用具有1600万个读取的下一代测序(NGS)对其进行了确认。针对重链可变域的10个变体中每一个的归一化测序读取约为1。针对轻链可变域的43个变体中每一个的归一化测序读取约为1。针对重链互补决定区3的236个变异序列的归一化测序读取约为1。

[0241] 然后测试了各种轻链和重链的表达和蛋白质折叠。针对重链可变域的10个变异序列包括:IGHV1-18、IGHV1-69、IGHV1-8、IGHV3-21、IGHV3-23、IGHV3-30/33rn、IGHV3-28、IGHV3-74、IGHV4-39和IGHV4-59/61。在这10个变异序列中,IGHV1-18、IGHV1-69和IGHV3-30/33rn表现出改善的特性,如改善的热稳定性。针对轻链可变域的9个变异序列包括:

IGKV1-39、IGKV1-9、IGKV2-28、IGKV3-11、IGKV3-15、IGKV3-20、IGKV4-1、IGLV1-51和IGLV2-14。在这9个变异序列中，IGKV1-39、IGKV3-15、IGLV1-51和IGLV2-14表现出改善的特性，如改善的热稳定性。

[0242] 实施例6:GPCR抗体文库在HEK293细胞中的表达

[0243] 在生成GPCR抗体文库后，选择约47个GPCR进行筛选。在pCDNA3.1载体中设计了大小为约1.8kb至约4.5kb的GPCR构建体。然后按照与实施例4-5中所述类似的方法，包括层级式装配，合成了GPCR构建体。在这47个GPCR构建体中，合成了46个GPCR构建体。

[0244] 将合成的GPCR构建体转染到HEK293中，并使用免疫荧光法检测其表达。用包含N末端血凝素(HA)标记的人Y1受体的GPCR构建体转染HEK293细胞(数据未显示)。转染24-48小时后，用磷酸盐缓冲盐水(PBS)洗涤细胞，并用4％低聚甲醛固定。使用针对HA标签的荧光第一抗体或包含荧光团和DAPI的第二抗体对细胞进行染色，以使细胞核呈现为蓝色。人Y1受体在未透化细胞的细胞表面上以及在透化细胞的细胞表面和细胞内呈现出来(数据未显示)。

[0245] 还通过设计包含自体荧光蛋白的GPCR构建体对GPCR构建体进行了可视化。人Y1受体包含与其C末端融合的EYFP，而人Y5受体包含与其C末端融合的ECFP(数据未显示)。HEK293细胞用人Y1受体转染，或用人Y1受体和人Y5受体共转染。转染后，将细胞洗涤并用4％低聚甲醛固定。将细胞用DAPI染色。通过荧光显微术对人Y1受体和人Y5受体的定位进行可视化。

[0246] 实施例7免疫球蛋白文库的设计

[0247] 设计了一种免疫球蛋白支架文库，以供放置GPCR结合域并提高一系列GPCR结合域编码序列的稳定性。该免疫球蛋白支架包括用连接体附接有VL结构域的VH结构域。为VH结构域和VL结构域的框架元件和CDR元件生成了变异核酸序列。该设计的结构在图12A中示出。全结构域架构在图12B中示出。前导序列、连接体和pIII的序列在表7中列出。

表7.核苷酸序列

SEQ ID NO	结构域	序列
6	前导序列	GCAGCCGCTGGCTTGCTGCTGCTGGCAGCTCAGCCGGCCATGGCC
7	连接体	GCTAGCGGTGGAGGCGGTTCAGGCGGAGGTGGCTCTGGCGGTGGCGGATCGCATGCATCC
8	pIII	CGCGCGGCCGCTGGAAGCGGCTCCCACCATCACCATCACCAT

[0248] 设计的VL结构域包括IGKV1-39、IGKV3-15、IGLV1-51和IGLV2-14。四个VL结构域中的每一个与它们各自的不变的四个框架元件(FW1、FW2、FW3、FW4)和可变的3个CDR(L1、L2、L3)元件组装在一起。对于IGKV1-39，存在为L1设计的490个变体，为L2设计的420个变体，和为L3设计的824个变体，从而导致多样性为$1.7x10^8$(490*420*824)。对于IGKV3-15，存在为L1设计的490个变体，为L2设计的265个变体，和为L3设计的907个变体，从而导致多样性为$1.2x10^8$(490*265*907)。对于IGLV1-51，存在为L1设计的184个变体，为L2设计的151个变体，和为L3设计的824个变体，从而导致多样性为$2.3x10^7$(184*151*824)。对于IGLV2-14，存

在为L1设计的967个变体，为L2设计的535个变体，和为L3设计的922个变体，从而导致多样性为4.8 10^8(967*535*922)。表8列出了IGLV1-51的四个框架元件(FW1、FW2、FW3、FW4)的氨基酸序列和核苷酸序列。表9列出了IGLV1-51的可变的3个CDR(L1、L2、L3)元件。还为IGKV1-39、IGKV3-15和IGLV2-14设计了四个框架元件(FW1、FW2、FW3、FW4)和可变的3个CDR(L1、L2、L3)元件的变异氨基酸序列和核苷酸序列。

表8.IGLV1-51框架元件的序列

元件	SEQ ID NO	氨基酸序列	SEQ ID NO	核苷酸序列
IGLV1-51				
FW1	9	QSVLTQPPSVSAAPGQKVTISC	10	CAGTCTGTGTTGACGCAGCCGCCCTCAGTGTCTGCGGCCCCAGGACAGAAGGTCACCATCTCCTGC
FW2	11	WYQQLPGTAPKLLIY	12	TGGTATCAGCAGCTCCCAGGAACAGCCCCCAAACTCCTCATTTAT
FW3	13	GIPDRFSGSKSGTSATLGITGLQTGDEADYY	14	GGGATTCCTGACCGATTCTCTGGCTCCAAGTCTGGCACGTCAGCCACCCTGGGCATCACCGGACTCCAGACTGGGGACGAGGCCGATTATTAC
FW4	15	GGGTKLTVL	16	GGCGGAGGGACCAAGCTGACCGTCCTA

表9.IGLV1-51 CDR元件的序列

SEQ ID NO	氨基酸序列	SEQ ID NO	核苷酸序列
IGLV1-51-L1			
17	SGSSSNIGSNHVS	200	TCTGGAAGCAGCTCCAACATTGGGAGTAATCATGTATCC
18	SGSSSNIGNNYLS	201	TCTGGAAGCAGCTCCAACATTGGGAATAATTATCTATCC
19	SGSSSNIANNYVS	202	TCTGGAAGCAGCTCCAACATTGCGAATAATTATGTATCC
20	SGSSPNIGNNYVS	203	TCTGGAAGCAGCCCCAACATTGGGAATAATTATGTATCG
21	SGSRSNIGSNYVS	204	TCTGGAAGCAGATCCAATATTGGGAGTAATTATGTTTCG
22	SGSSSNVGDNYVS	205	TCTGGAAGCAGCTCCAACGTTGGCGATAATTATGTTTCC
23	SGSSSNIGIQYVS	206	TCTGGAAGCAGCTCCAACATTGGGATTCAATATGTATCC
24	SGSSSNVGNNFVS	207	TCTGGAAGCAGCTCCAATGTTGGTAACAATTTTGTCTCC
25	SGSASNIGNNYVS	208	TCTGGAAGCGCCTCCAACATTGGGAATAATTATGTATCC
26	SGSGSNIGNNDVS	209	TCTGGAAGCGGCTCCAATATTGGGAATAATGATGTGTCC
27	SGSISNIGNNYVS	210	TCTGGAAGCATCTCCAACATTGGTAATAATTATGTATCC
28	SGSISNIGKNYVS	211	TCTGGAAGCATCTCCAACATTGGGAAAAATTATGTGTCG
29	SGSSSNIGHNYVS	212	TCTGGAAGCAGCTCCAACATTGGGCATAATTATGTATCG
30	PGSSSNIGNNYVS	213	CCTGGAAGCAGCTCCAACATTGGGAATAATTATGTATCC
31	SGSTSNIGIHYVS	214	TCTGGAAGCACCTCCAACATTGGAATTCATTATGTATCC
32	SGSSSNIGSHYVS	215	TCTGGAAGCAGCTCCAACATTGGCAGTCATTATGTTTCC
33	SGSSSNIGNEYVS	216	TCCGGAAGCAGCTCCAACATTGGAAATGAATATGTATCC
34	SGSTSNIGNNYIS	217	TCTGGAAGCACCTCCAACATTGGAAATAATTATATATCG
35	SGSSSNIGNHFVS	218	TCTGGAAGCAGCTCCAATATTGGGAATCATTTTGTATCG
36	SGSSSNIGNNYVA	219	TCTGGAAGCAGCTCCAACATTGGGAATAATTATGTGGCC
37	SGSSSNIGSYYVS	220	TCTGGAAGCAGCTCCAACATTGGAAGTTATTATGTATCC
38	SGSGFNIGNNYVS	221	TCTGGAAGTGGTTTCAACATTGGGAATAATTATGTCTCT
39	SGSTSNIGNNYVS	222	TCTGGAAGCACCTCCAACATTGGGAATAATTATGTGTCC
40	SGSSSDIGNNYVS	223	TCTGGAAGCAGCTCCGACATTGGCAATAATTATGTATCC
41	SGSSSNIGNNVVS	224	TCTGGAAGCAGCTCCAACATTGGGAATAATGTTGTATCC
42	SGSKSNIGKNYVS	225	TCTGGAAGCAAGTCTAACATTGGGAAAAATTATGTATCC
43	SGSSTNIGNNYVS	226	TCTGGAAGCAGCACCAACATTGGGAATAATTATGTATCC

44	SGSISNIGDNYVS	227	TCTGGAAGCATCTCCAACATTGGGGATAATTATGTATCC
45	SGSSSNIGSKDVS	228	TCTGGAAGCAGCTCCAACATTGGGAGTAAGGATGTATCA
46	SGSSSNIENNDVS	229	TCTGGAAGCAGCTCCAACATTGAGAATAATGATGTATCG
47	SGSSSNIGNHYVS	230	TCTGGAAGCAGCTCCAACATTGGGAATCATTATGTATCC
48	SGSSSNIGKDFVS	231	TCTGGAAGCAGCTCCAACATTGGGAAGGATTTTGTCTCC
49	SGSTSNIGSNFVS	232	TCTGGCAGTACTTCCAACATCGGAAGTAATTTTGTTTCC
50	SGSTSNIGHNYVS	233	TCTGGAAGCACCTCCAACATTGGGCATAATTATGTATCC
51	SASSSNIGNNYVS	234	TCTGCAAGCAGCTCCAACATTGGGAATAATTATGTATCC
52	SGSSSSIGNNYVS	235	TCTGGAAGCAGCTCCAGCATTGGCAATAATTATGTATCC
53	SGSSSTIGNNYVS	236	TCTGGAAGCAGCTCCACCATTGGGAATAATTATGTATCC
54	SGSSSNIENNYVS	237	TCTGGAAGCAGCTCCAACATTGAAAATAATTATGTATCC
55	SGSSSNIGNQYVS	238	TCTGGAAGCAGCTCCAACATTGGGAATCAGTATGTATCC
56	SGSSSNIGNNYVF	239	TCTGGAAGCAGCTCCAACATTGGGAATAATTATGTATTC
57	SGSSSNIGRNYVS	240	TCTGGAAGCAGCTCCAACATTGGGAGGAATTATGTCTCC
58	SGGSSNIGNYYVS	241	TCTGGAGGCAGCTCCAACATTGGAAATTATTATGTATCG
59	SGSSSNIGDNYVS	242	TCTGGAAGCAGCTCCAACATTGGAGATAATTATGTCTCC
60	SGGSSNIGINYVS	243	TCTGGAGGCAGCTCCAACATTGGAATTAATTATGTATCC
61	SGGSSNIGKNYVS	244	TCTGGAGGCAGCTCCAACATTGGGAAGAATTATGTATCC
62	SGSSSNIGKRSVS	245	TCTGGAAGCAGCTCCAACATTGGGAAGAGATCTGTATCG
63	SGSRSNIGNNYVS	246	TCTGGAAGCAGATCCAACATTGGGAATAACTATGTATCC
64	SGSSSNIGNNLVS	247	TCGGGAAGCAGCTCCAACATTGGGAATAATCTTGTTTCC
65	SGSSSNIGINYVS	248	TCTGGAAGCAGCTCCAACATTGGGATCAATTATGTATCC
66	SGSSSNIGNNFVS	249	TCTGGAAGCAGCTCCAACATCGGGAATAATTTTGTATCC
67	SGTSSNIGRNFVS	250	TCTGGAACCAGCTCCAACATTGGCAGAAATTTTGTATCC
68	SGRRSNIGNNYVS	251	TCTGGAAGGAGGTCCAACATTGGAAATAATTATGTGTCC
69	SGGSFNIGNNYVS	252	TCTGGAGGCAGCTTCAATATTGGGAATAATTATGTATCC
70	SGSTSNIGENYVS	253	TCTGGAAGCACTTCCAACATTGGGGAGAATTATGTGTCC
71	SGSSSNIGSDYVS	254	TCTGGAAGCAGCTCCAATATTGGGAGTGATTATGTATCC
72	SGTSSNIGSNYVS	255	TCTGGAACCAGCTCCAACATTGGGAGTAATTATGTATCC
73	SGSSSNIGTNFVS	256	TCTGGAAGCAGCTCCAACATTGGGACTAATTTTGTATCC
74	SGSSSNFGNNYVS	257	TCTGGAAGCAGCTCCAACTTTGGGAATAATTATGTATCC
75	SGSTSNIGNNHVS	258	TCTGGAAGCACCTCCAACATTGGGAATAATCATGTATCC
76	SGSSSNIGNDFVS	259	TCTGGAAGCAGCTCCAACATTGGGAATGATTTTGTATCC
77	SGSSSDIGDNYVS	260	TCTGGAAGCAGCTCCGACATTGGCGATAATTATGTGTCC
78	SGSSSNIGKYYVS	261	TCTGGAAGCAGCTCCAACATTGGGAAATATTATGTATCC
79	SGSSSNIGGNYVS	262	TCTGGAAGCAGCTCCAACATTGGCGGTAATTATGTATCC
80	SGSSSNTGNNYVS	263	TCTGGAAGCAGCTCCAACACTGGGAATAATTATGTATCC
81	SGSSSNVGNNYVS	264	TCTGGAAGCAGCTCCAACGTTGGGAATAATTATGTGTCT
82	SGSSSNIANNFVS	265	TCTGGAAGCAGCTCCAACATTGCGAATAATTTTGTATCC
83	SGSSSNIGNDYVS	266	TCTGGAAGCAGCTCCAACATTGGGAATGATTATGTATCC
84	SGSTSNIENNYVS	267	TCTGGAAGCACCTCCAATATTGAGAATAATTATGTTTCC
85	SGGSSNIGNNDVS	268	TCTGGAGGCAGCTCCAATATTGGCAATAATGATGTGTCC
86	SGSTSNIGNHYVS	269	TCTGGAAGCACCTCCAACATTGGGAATCATTATGTATCC
87	SGSSSNIGDNDVS	270	TCAGGAAGCAGCTCCAATATTGGGGATAATGATGTATCC
88	SGYSSNIGNNYVS	271	TCTGGATACAGCTCCAACATTGGGAATAATTATGTATCC

89	SGSGSNIGNNFVS	272	TCTGGAAGCGGCTCCAACATTGGAAATAATTTTGTATCC
90	SGSSSNIWNNYVS	273	TCTGGAAGCAGCTCCAACATTTGGAATAATTATGTATCC
91	FGSSSNIGNNYVS	274	TTTGGAAGCAGCTCCAACATTGGGAATAATTATGTATCC
92	SGSSSNIEKNYVS	275	TCTGGAAGCAGCTCCAACATTGAGAAGAATTATGTATCC
93	SGSRSNIGNYYVS	276	TCTGGAAGTAGATCCAATATTGGAAATTATTATGTATCC
94	SGTKSNIGNNYVS	277	TCTGGAACCAAGTCAAACATTGGGAATAATTATGTATCT
95	SGSTSNIGNYYVS	278	TCTGGAAGCACCTCCAACATTGGGAATTATTATGTATCC
96	SGTSSNIGNNYVA	279	TCTGGAACCAGCTCCAACATTGGGAATAATTATGTGGCC
97	PGTSSNIGNNYVS	280	CCTGGAACCAGCTCCAACATTGGGAATAATTATGTATCC
98	SGSTSNIGINYVS	281	TCCGGAAGCACCTCCAACATTGGGATTAATTATGTATCC
99	SGSSSNIGSNLVS	282	TCTGGAAGCAGCTCCAACATTGGGAGTAATCTGGTATCC
100	SGSSSNIENNHVS	283	TCTGGAAGCAGCTCCAACATTGAGAATAATCATGTATCC
101	SGTRSNIGNNYVS	284	TCTGGAACCAGGTCCAACATCGGCAATAATTATGTTTCG
102	SGSTSNIGDNYVS	285	TCTGGAAGCACCTCCAACATTGGGGACAATTATGTTTCC
103	SGGSSNIGKNFVS	286	TCTGGAGGCAGTTCCAACATTGGGAAGAATTTTGTATCC
104	SGSRSDIGNNYVS	287	TCTGGAAGCAGGTCCGACATTGGGAATAATTATGTATCC
105	SGTSSNIGNNDVS	288	TCTGGAACTAGCTCCAACATTGGGAATAATGATGTATCC
106	SGSSSNIGSKYVS	289	TCTGGAAGCAGCTCCAACATTGGGAGTAAATATGTATCA
107	SGSSFNIGNNYVS	290	TCTGGAAGCAGCTTCAACATTGGGAATAATTATGTATCC
108	SGSSSNIGNTYVS	291	TCTGGAAGCAGCTCCAACATTGGGAATACTTATGTATCC
109	SGSSSNIGDNHVS	292	TCTGGAAGCAGCTCCAATATTGGGGATAATCATGTATCC
110	SGSSSNIGNNHVS	293	TCTGGAAGCAGCTCCAACATTGGCAATAATCATGTTTCC
111	SGSTSNIGNNDVS	294	TCTGGAAGCACCTCCAACATTGGGAATAATGATGTATCC
112	SGSRSNVGNNYVS	295	TCTGGAAGCAGATCCAACGTTGGCAATAATTATGTTTCA
113	SGGTSNIGKNYVS	296	TCCGGAGGCACCTCCAACATTGGGAAGAATTATGTGTCT
114	SGSSSNIADNYVS	297	TCTGGAAGCAGCTCCAACATTGCCGATAATTATGTTTCC
115	SGSSSNIGANYVS	298	TCTGGAAGCAGCTCCAACATTGGCGCCAATTATGTATCC
116	SGSSSNIGSNYVA	299	TCTGGAAGCAGCTCCAACATTGGGAGTAATTATGTGGCC
117	SGSSSNIGNNFLS	300	TCTGGAAGCAGCTCCAACATTGGGAACAATTTTCTCTCC
118	SGRSSNIGKNYVS	301	TCTGGAAGAAGCTCCAACATTGGGAAGAATTATGTATCC
119	SGSSPNIGANYVS	302	TCTGGAAGCAGCCCCAACATTGGGGCTAATTATGTATCC
120	SGSSSNIGPNYVS	303	TCCGGAAGCAGCTCCAACATTGGGCCTAATTATGTGTCC
121	SGSSSTIGNNYIS	304	TCTGGAAGCAGCTCCACCATTGGGAATAATTATATATCC
122	SGSSSNIGNYFVS	305	TCTGGAAGCAGCTCCAACATTGGGAATTATTTTGTATCC
123	SGSRSNIGNNFVS	306	TCTGGAAGCCGCTCCAACATTGGTAATAATTTTGTATCC
124	SGGSSNIGSNFVS	307	TCTGGAGGCAGCTCCAACATTGGGAGTAATTTTGTATCC
125	SGSSSNIGYNYVS	308	TCTGGAAGCAGCTCCAACATTGGGTATAATTATGTATCC
126	SGTSSNIENNYVS	309	TCTGGAACCAGCTCGAACATTGAGAACAATTATGTATCC
127	SGSSSNIGNYYVS	310	TCTGGAAGTAGCTCCAACATTGGGAATTATTATGTATCC
128	SGSTSNIGKNYVS	311	TCTGGAAGCACCTCCAACATTGGGAAGAATTATGTATCC
129	SGSSSNIGTYYVS	312	TCTGGAAGCAGTTCCAACATTGGGACTTATTATGTCTCT
130	SGSSSNVGKNYVS	313	TCTGGAAGCAGCTCCAACGTTGGGAAAAATTATGTATCT
131	SGSTSNIGDNFVS	314	TCTGGAAGCACCTCCAACATTGGGGATAATTTTGTATCC
132	SGSTSNIGTNYVS	315	TCTGGAAGCACCTCCAACATTGGAACTAATTATGTTTCC
133	SGGTSNIGNNYVS	316	TCTGGAGGTACTTCCAACATTGGGAATAATTATGTCTCC

134	SGSYSNIGNNYVS	317	TCTGGAAGCTACTCCAATATTGGGAATAATTATGTATCC
135	SGSSSNIEDNYVS	318	TCTGGAAGCAGCTCCAACATTGAAGATAATTATGTATCC
136	SGSSSNIGKHYVS	319	TCTGGAAGCAGCTCCAACATTGGGAAACATTATGTATCC
137	SGSGSNIGSNYVS	320	TCCGGTTCCGGCTCAAACATTGGAAGTAATTATGTCTCC
138	SGSSSNIGNNYIS	321	TCTGGAAGCAGCTCCAACATTGGAAATAATTATATATCA
139	SGASSNIGNNYVS	322	TCTGGAGCCAGTTCCAACATTGGGAATAATTATGTTTCC
140	SGRTSNIGNNYVS	323	TCTGGACGCACCTCCAACATCGGGAACAATTATGTATCC
141	SGGSSNIGSNYVS	324	TCTGGAGGCAGCTCCAATATTGGGAGTAATTACGTATCC
142	SGSGSNIGNNYVS	325	TCTGGAAGCGGCTCCAACATTGGGAATAATTATGTATCC
143	SGSTSNIGSNYVS	326	TCTGGAAGCACCTCCAACATTGGGAGTAATTATGTATCC
144	SGSSSSIGNNYVA	327	TCTGGAAGCAGCTCCAGCATTGGGAATAATTATGTGGCG
145	SGSSSNLGNNYVS	328	TCTGGAAGCAGTTCCAACCTTGGAAATAATTATGTATCC
146	SGTSSNIGKNYVS	329	TCTGGAACCAGCTCCAACATTGGGAAAAATTATGTATCC
147	SGSSSDIGNKYIS	330	TCTGGAAGCAGCTCCGATATTGGGAACAAGTATATATCC
148	SGSSSNIGSNYIS	331	TCTGGAAGCAGCTCCAACATTGGAAGTAATTACATATCC
149	SGSTSNIGANYVS	332	TCTGGAAGCACCTCCAACATTGGGGCTAACTATGTGTCC
150	SGSSSNIGNKYVS	333	TCTGGAAGCAGCTCCAACATTGGGAATAAGTATGTATCC
151	SGSSSNIGNNYGS	334	TCTGGAAGCAGCTCCAACATTGGGAATAATTATGGATCC
152	SGSTSNIANNYVS	335	TCTGGAAGCACCTCCAACATTGCGAATAATTATGTATCC
153	SGSYSNIGSNYVS	336	TCTGGAAGCTACTCCAATATTGGGAGTAATTATGTATCC
154	SGSSSNIGSNFVS	337	TCTGGAAGCAGCTCCAACATTGGGAGTAATTTTGTATCC
155	SGSSSNLENNYVS	338	TCTGGAAGCAGCTCCAATCTTGAGAATAATTATGTATCC
156	SGSISNIGSNYVS	339	TCTGGAAGCATCTCCAATATTGGCAGTAATTATGTATCC
157	SGSSSDIGSNYVS	340	TCTGGAAGCAGCTCCGACATTGGGAGTAATTATGTATCC
158	SGSSSNIGTNYVS	341	TCTGGAAGCAGCTCCAACATTGGGACTAATTATGTATCC
159	SGSSSNIGKNFVS	342	TCTGGAAGCAGCTCCAACATTGGGAAGAATTTTGTATCC
160	SGSSSNIGNNFIS	343	TCTGGAAGCAGCTCCAACATTGGGAATAATTTTATATCC
161	SGGSSNIGNNYVS	344	TCTGGAGGCAGCTCCAACATTGGCAATAATTATGTTTCC
162	SGSSSNIGENYVS	345	TCTGGAAGCAGCTCCAACATTGGGGAGAATTATGTATCC
163	SGSSSNIGNNFVA	346	TCTGGAAGCAGCTCCAATATTGGGAATAATTTTGTGGCC
164	SGGSSNIGNNYVA	347	TCTGGAGGCAGCTCCAACATTGGGAATAATTATGTAGCC
165	SGSSSHIGNNYVS	348	TCTGGAAGCAGCTCCCACATTGGAAATAATTATGTATCC
166	SGSSSNIGSNDVS	349	TCTGGAAGCAGCTCCAATATTGGAAGTAATGATGTATCG
167	SGSSSNIGNNYVT	350	TCTGGAAGCAGCTCCAACATTGGGAATAATTATGTAACC
168	SGSSSNIGNNPVS	351	TCTGGAAGCAGCTCCAACATTGGGAATAATCCTGTATCC
169	SGGSSNIGNHYVS	352	TCTGGAGGCAGCTCCAATATTGGGAATCATTATGTATCC
170	SGTSSNIGNNYVS	353	TCTGGAACCAGCTCCAACATTGGGAATAATTATGTATCC
171	SGSSSNIGSNYVS	354	TCTGGAAGCAGCTCCAACATTGGAAGTAATTATGTCTCG
172	SGGTSNIGSNYVS	355	TCTGGAGGCACCTCCAACATTGGAAGTAATTATGTATCC
173	SGSKSNIGNNYVS	356	TCTGGAAGCAAGTCCAACATTGGGAATAATTATGTATCC
174	SGRSSNIGNNYVS	357	TCTGGAAGAAGCTCCAACATTGGGAATAATTATGTATCG
175	SGSSSNVGSNYVS	358	TCTGGAAGCAGCTCCAACGTTGGGAGTAATTATGTTTCC
176	SGSTSNIGNNFVS	359	TCTGGAAGCACCTCCAATATTGGGAATAATTTTGTATCC
177	SGSNFNIGNNYVS	360	TCTGGAAGCAACTTCAACATTGGGAATAATTATGTCTCC
178	SGSTSNIGYNYVS	361	TCTGGAAGCACCTCCAATATTGGATATAATTATGTATCC

179	SGSSSNIVSNYVS	362	TCTGGAAGCAGCTCCAATATTGTAAGTAATTATGTATCC
180	SGTSSNIGNNFVS	363	TCTGGAACCAGCTCCAACATTGGGAATAATTTTGTATCC
181	SGSSSNIGRNFVS	364	TCTGGAAGCAGCTCCAACATTGGGAGGAATTTTGTGTCC
182	SGTTSNIGNNYVS	365	TCTGGAACGACCTCCAACATTGGGAATAATTATGTCTCC
183	SGSSSNIGNNDVS	366	TCTGGAAGCAGCTCCAACATTGGGAATAATGATGTATCC
184	SGSSSNIGNHDVS	367	TCTGGAAGCAGCTCCAACATTGGGAATCATGATGTATCC
185	SGSSSNIGSSHVS	368	TCTGGAAGCAGCTCCAACATTGGAAGTAGTCATGTATCC
186	SGSSSNIGIHYVS	369	TCTGGAAGCAGCTCCAACATTGGGATTCATTATGTATCC
187	SGGGSNIGYNYVS	370	TCTGGAGGCGGCTCCAACATTGGCTATAATTATGTCTCC
188	SGSSSNIGDHYVS	371	TCTGGAAGCAGCTCCAACATTGGGGATCATTATGTGTCG
189	SGSSSNLGKNYVS	372	TCTGGAAGCAGCTCCAACCTTGGGAAGAATTATGTATCT
190	SGSSSNIGDNFVS	373	TCTGGAAGCAGCTCCAACATTGGCGATAATTTTGTATCC
191	SGSTSNIEKNYVS	374	TCTGGAAGCACCTCCAACATTGAGAAAAACTATGTATCG
192	SGSSSNIGKDYVS	375	TCTGGAAGCAGCTCCAACATTGGGAAGGATTATGTATCC
193	SGSSSNIGKNYVS	376	TCTGGAAGCAGCTCCAACATTGGGAAGAATTATGTATCC
194	SGSSSNIGNNYVS	377	TCTGGAAGCAGCTCCAACATTGGGAATAATTATGTATCC
195	SGSSSNIGNNYAS	378	TCTGGAAGCAGCTCCAACATTGGGAATAATTATGCCTCC
196	SGISSNIGNNYVS	379	TCTGGAATCAGCTCCAACATTGGGAATAATTATGTATCC
197	TGSSSNIGNNYVS	380	ACTGGAAGCAGCTCCAACATTGGGAATAATTATGTATCC
198	SGTSSNIGNNHVS	381	TCTGGAACCAGCTCCAACATTGGGAATAATCATGTTTCC
199	SGSRSNIGKNYVS	382	TCTGGAAGTCGTTCCAACATTGGGAAAAATTATGTATCC
IGLV1-51-L2			
383	DNNKRPP	534	GACAATAATAAGCGACCCCCA
384	ENNRRPS	535	GAGAATAATAGGCGACCCTCA
385	DNNKQPS	536	GACAATAATAAGCAACCCTCA
386	DNNKRPL	537	GACAATAACAAGCGACCCTTG
387	DNDKRPA	538	GACAATGATAAGCGACCCGCA
388	DNHERPS	539	GACAATCATGAGCGACCCTCA
389	ENRKRPS	540	GAAAACCGTAAGCGACCCTCA
390	DNDQRPS	541	GACAATGATCAGCGACCCTCA
391	ENYKRPS	542	GAGAATTATAAGCGACCCTCA
392	ENTKRPS	543	GAAAATACTAAGCGACCCTCA
393	DTEKRPS	544	GACACTGAGAAGAGGCCCTCA
394	DNDKRPP	545	GACAATGATAAGCGACCCCCA
395	DHNKRPS	546	GACCATAATAAGCGACCCTCA
396	GNNERPS	547	GGCAATAATGAGCGACCCTCA
397	DTSKRPS	548	GACACTAGTAAGCGACCCTCA
398	EYNKRPS	549	GAATATAATAAGCGCCCCTCA
399	ENIKRPS	550	GAAAATATTAAGCGACCCTCA
400	DNVKRPS	551	GACAATGTTAAGCGACCCTCA
401	ENDKRSS	552	GAAAACGATAAACGATCCTCA
402	ENNKRHS	553	GAAAATAATAAGCGACACTCA
403	GNDQRPS	554	GGAAATGATCAGCGACCCTCA
404	DNDRRPS	555	GACAATGATAGGCGACCCTCA
405	DNHKRPS	556	GACAATCATAAGCGGCCCTCA

406	DNNDRPS	557	GACAATAATGACCGACCCTCA
407	ENNQRPS	558	GAGAATAATCAGCGACCCTCA
408	DNNQRPS	559	GACAATAATCAGCGACCCTCA
409	ENVKRPS	560	GAGAATGTTAAGCGACCCTCA
410	DTYKRPS	561	GACACTTATAAGAGACCCTCA
411	NNNNRPS	562	AACAATAATAACCGACCCTCA
412	GNNNRPS	563	GGCAATAATAATCGACCCTCA
413	ENDQRPS	564	GAAAATGATCAGCGACCCTCA
414	DNNKRAS	565	GACAATAATAAGCGAGCCTCA
415	DNDKRPL	566	GACAATGATAAGCGACCCTTA
416	DTDERPS	567	GACACTGATGAGCGACCTTCA
417	DNRKRPS	568	GACAATAGGAAGCGACCCTCA
418	DNDARPS	569	GACAATGATGCTCGACCCTCA
419	DNNKRLS	570	GACAATAATAAGCGACTCTCA
420	DNDKRAS	571	GACAATGATAAGCGAGCCTCA
421	DNTERPS	572	GACAATACTGAGCGACCCTCA
422	DNNIRPS	573	GACAATAATATTCGACCCTCA
423	DNKRRPS	574	GACAATAAGAGGCGACCCTCA
424	DDNNRPS	575	GACGATAATAACCGACCCTCA
425	ANNRRPS	576	GCGAATAATCGACGACCCTCA
426	DNDKRLS	577	GACAATGATAAGCGACTGTCA
427	DNNKRPA	578	GACAATAATAAGCGACCCGCA
428	DNYRRPS	579	GACAATTATAGACGTCCCTCA
429	ANDQRPS	580	GCCAATGATCAGCGACCCTCA
430	DNDKRRS	581	GACAATGATAAGCGACGCTCA
431	DKNERPS	582	GACAAGAATGAGCGACCCTCA
432	DNKERPS	583	GACAATAAGGAGCGACCCTCA
433	DNNKGPS	584	GACAATAATAAGGGACCCTCA
434	ENDRRPS	585	GAAAATGATAGACGACCCTCA
435	ENDERPS	586	GAAAATGATGAGCGACCCTCA
436	QNNKRPS	587	CAAAATAATAAGCGACCCTCA
437	DNRERPS	588	GACAATCGTGAGCGACCCTCA
438	DNNRRPS	589	GACAATAATAGACGACCCTCA
439	GNNRRPS	590	GGAAATAATAGGCGACCCTCA
440	DNDNRPS	591	GACAATGATAACCGACCCTCA
441	EDNKRPS	592	GAAGATAATAAGCGACCCTCA
442	DDDERPS	593	GACGATGATGAGCGGCCCTCA
443	ASNKRPS	594	GCAAGTAATAAGCGACCCTCA
444	DNNKRSS	595	GACAATAATAAGCGATCCTCA
445	QNNERPS	596	CAAAATAATGAGCGACCCTCA
446	DDDRRPS	597	GACGATGATAGGCGACCCTCA
447	NNDKRPS	598	AACAATGATAAGCGACCCTCA
448	DNNNRPS	599	GACAATAATAACCGACCCTCA
449	DNNVRPS	600	GACAATAATGTGCGACCCTCA
450	ENNERPS	601	GAAAATAATGAGCGACCCTCA

451	DNNHRPS	602	GACAATAATCACCGACCCTCA
452	DNDERPS	603	GACAATGATGAGCGCCCCTCG
453	DNIRRPS	604	GACAATATCCGGCGACCCTCA
454	DFNKRPS	605	GACTTTAATAAGCGACCCTCA
455	ETNKRPS	606	GAAACTAATAAGCGACCCTCA
456	NDNKRPS	607	AACGATAATAAGCGACCCTCA
457	DDNKRPS	608	GACGATAATAAGCGACCCTCA
458	DNYKRPS	609	GACAATTATAAGCGACCCTCA
459	HNNKRPS	610	CACAATAATAAGCGACCCTCA
460	DNHQRPS	611	GACAATCATCAGCGACCCTCA
461	DNYKRAS	612	GACAATTATAAGCGAGCCTCA
462	DNIKRPS	613	GACAATATTAAGCGACCCTCA
463	DTHKRPS	614	GACACTCATAAGCGACCCTCA
464	DTNRRPS	615	GACACTAATAGGCGACCCTCT
465	DTNQRPS	616	GACACTAATCAGCGACCCTCA
466	ESDKRPS	617	GAAAGTGATAAGCGACCCTCA
467	DNDKRSS	618	GACAATGATAAGCGATCTTCG
468	GSNKRPS	619	GGCAGTAATAAGCGACCCTCA
469	DNNKRVS	620	GACAATAACAAGCGAGTTTCA
470	NNNRRPS	621	AACAATAATAGGCGACCCTCA
471	DNFKRPS	622	GACAATTTTAAGCGACCCTCA
472	ENDKRPS	623	GAAAATGATAAACGACCCTCA
473	ENNKRLS	624	GAAAATAATAAGCGACTCTCA
474	ADNKRPS	625	GCAGATAATAAGCGACCCTCA
475	EDNERPS	626	GAAGATAATGAGCGCCCCTCA
476	DTDQRPS	627	GACACTGATCAGCGACCCTCA
477	DNYQRPS	628	GACAATTATCAGCGACCCTCA
478	DENKRPS	629	GACGAGAATAAGCGACCCTCA
479	DTNKRPS	630	GACACTAATAAGCGACCCTCA
480	DDYRRPS	631	GACGATTATCGGCGACCCTCA
481	DNDKRHS	632	GACAACGATAAGCGGCACTCA
482	ENDNRPS	633	GAAAATGATAATCGACCCTCA
483	DDNERPS	634	GACGATAATGAGCGCCCCTCA
484	DNKKRPS	635	GACAATAAGAAGCGACCCTCA
485	DVDKRPS	636	GACGTTGATAAGCGACCCTCA
486	ENKKRPS	637	GAAAATAAAAAACGACCCTCT
487	VNDKRPS	638	GTCAATGATAAGCGACCCTCA
488	DNDHRPS	639	GACAATGATCACCGACCCTCA
489	DINKRPS	640	GACATTAATAAGCGACCCTCA
490	ANNERPS	641	GCCAATAATGAGCGACCCTCA
491	DNENRPS	642	GACAATGAAAACCGACCGTCA
492	GDDKRPS	643	GGCGATGATAAGCGACCCTCA
493	ANNQRPS	644	GCCAATAATCAGCGACCTTCA
494	DDDKRPS	645	GACGATGATAAGCGACCCTCA
495	YNNKRPS	646	TACAATAATAAGCGGCCCTCA

496	EDDKRPS	647	GAAGATGATAAGCGACCCTCA
497	ENNNRPS	648	GAAAACAATAACCGACCCTCG
498	DNNLRPS	649	GACAATAATCTGCGACCCTCA
499	ESNKRPS	650	GAGAGTAACAAGCGACCCTCA
500	DTDKRPS	651	GACACTGATAAGCGGCCCTCA
501	DDDQRPS	652	GACGATGATCAGCGACCCTCA
502	VNNKRPS	653	GTGAATAATAAGAGACCCTCC
503	DDYKRPS	654	GACGATTATAAGCGACCCTCA
504	DNTKRPS	655	GACAATACTAAGCGACCCTCA
505	DDTERPS	656	GACGATACTGAGCGACCCTCA
506	GNDKRPS	657	GGCAATGATAAGCGACCCTCA
507	DNEKRPS	658	GACAATGAAAAGCGACCCTCA
508	DNDDRPS	659	GACAATGATGACCGACCCTCA
509	DDNRRPS	660	GACGATAATAGGCGTCCCTCA
510	GNNKRPS	661	GGCAATAATAAGCGACCCTCA
511	ANDKRPS	662	GCCAATGATAAGCGACCCTCA
512	DNNKRHS	663	GACAATAATAAGCGACACTCA
513	DDNQRPS	664	GACGACAATCAGCGACCCTCA
514	GNDRRPS	665	GGCAATGATAGGCGACCCTCA
515	DNHNRPS	666	GACAATCATAACCGACCCTCA
516	DNYERPS	667	GACAATTATGAGCGACCCTCA
517	ENNKRSS	668	GAAAATAATAAGCGATCCTCA
518	DDHKRPS	669	GACGATCATAAGCGGCCCTCA
519	DNNKRRS	670	GACAATAATAAACGACGTTCA
520	DNDKRPS	671	GACAATGATAAGCGACCGTCA
521	DKNKRPS	672	GACAAGAATAAGCGACCCTCA
522	DNNKRPS	673	GACAATAATAAGCGACCCTCA
523	DIDKRPS	674	GACATTGATAAGCGACCCTCA
524	DDKKRPS	675	GACGATAAGAAGCGACCCTCA
525	ANNKRPS	676	GCCAATAATAAGCGACCCTCA
526	DNDKGPS	677	GACAATGATAAGGGACCCTCA
527	EDNRRPS	678	GAAGATAATAGGCGACCCTCA
528	ENNKRPS	679	GAGAATAATAAGCGACCCTCA
529	NNNKRPS	680	AACAATAATAAGCGACCCTCA
530	DNNERPS	681	GACAATAATGAGCGACCCTCA
531	DNIQRPS	682	GACAATATTCAGCGACCCTCA
532	DNNYRPS	683	GACAATAATTACCGACCCTCA
533	DNYNRPS	684	GACAATTATAACCGACCCTCA
IGLV1-51-L3			
685	CGTWDTSLSAVVF	1509	TGCGGAACATGGGATACCAGCCTGAGTGCTGTGGTGTTC
686	CGTWDTSLSAGVF	1510	TGCGGAACATGGGATACCAGCCTGAGTGCTGGGGTGTTC
687	CGTWDTSLSAWVF	1511	TGCGGAACATGGGATACCAGCCTGAGTGCTTGGGTGTTC
688	CGTWDRSLSAGVF	1512	TGCGGAACATGGGATAGGAGCCTGAGTGCGGGGGTGTTC
689	CGTWDRSLSAWVF	1513	TGCGGAACATGGGATAGGAGCCTGAGTGCTTGGGTATTT
690	CGTWDTSLSGGVF	1514	TGCGGAACATGGGATACCAGCCTGAGTGGTGGGGTGTTC

691	CGTWDTSLRAGVF	1515	TGCGGAACATGGGATACTAGCCTGCGTGCTGGCGTCTTC
692	CGTWDRSLSVWVF	1516	TGCGGAACATGGGATAGGAGCCTGAGTGTTTGGGTGTTC
693	CGTWDTSLSVVVF	1517	TGCGGAACATGGGATACCAGTCTGAGTGTTGTGGTCTTC
694	CGTWDTSLSAAVF	1518	TGCGGAACGTGGGATACCAGCCTGAGTGCTGCGGTGTTC
695	CGAWDTSLSAGVF	1519	TGCGGAGCATGGGATACCAGCCTGAGTGCTGGAGTGTTC
696	CATWDTSLSAVVF	1520	TGCGCAACATGGGATACCAGCCTGAGTGCTGTGGTATTC
697	CATWDTSLSAGVF	1521	TGCGCAACATGGGATACCAGCCTGAGTGCTGGTGTGTTC
698	CGTWESSLSAWVF	1522	TGTGGAACATGGGAGAGCAGCCTGAGTGCTTGGGTGTTC
699	CGTWDTTLSAGVF	1523	TGCGGAACATGGGATACCACCCTGAGTGCGGGTGTCTTC
700	CGTWDTSLSVWVF	1524	TGCGGAACATGGGATACTAGCCTGAGTGTGTGGGTGTTC
701	CGTWDTSLSVGVF	1525	TGCGGAACATGGGATACTAGCCTGAGTGTTGGGGTGTTC
702	CGTWDTSLSTGVF	1526	TGCGGAACATGGGACACCAGTCTGAGCACTGGCGTCTTC
703	CGTWDTSLSGVVF	1527	TGCGGAACATGGGATACCAGCCTGAGTGGTGTGGTCTTC
704	CGTWDTSLSAYVF	1528	TGCGGAACATGGGATACCAGCCTGAGTGCTTATGTCTTC
705	CGTWDTSLSAEVF	1529	TGCGGAACATGGGATACCAGCCTGAGTGCTGAGGTGTTC
706	CGTWDTGLSAGVF	1530	TGCGGAACATGGGATACCGGCCTGAGTGCTGGGGTATTC
707	CGTWDRSLSAYVF	1531	TGCGGAACGTGGGATAGGAGCCTGAGTGCTTATGTCTTC
708	CGTWDRSLSAVVF	1532	TGCGGAACATGGGATAGGAGCCTCAGTGCCGTGGTATTC
709	CGTWDNTLSAWVF	1533	TGCGGAACATGGGATAACACCCTGAGTGCGTGGGTGTTC
710	CGTWDNRLSAGVF	1534	TGCGGAACATGGGATAACAGGCTGAGTGCTGGGGTGTTC
711	CGTWDISLSAWVF	1535	TGCGGAACATGGGACATCAGCCTGAGTGCTTGGGTGTTC
712	CGTWHSSLSAGVF	1536	TGCGGAACATGGCATAGCAGCCTGAGTGCTGGGGTATTC
713	CGTWGSSLSAWVF	1537	TGCGGAACATGGGGTAGCAGTTTGAGTGCTTGGGTGTTC
714	CGTWESSLSGWVF	1538	TGCGGAACATGGGAGAGCAGCCTGAGTGGTTGGGTGTTC
715	CGTWESSLSAVVF	1539	TGCGGAACATGGGAGAGCAGCCTGAGTGCTGTGGTTTTC
716	CGTWDYSLSAVVF	1540	TGCGGAACATGGGATTACAGCCTGAGTGCTGTGGTATTC
717	CGTWDYSLSAGVF	1541	TGCGGAACATGGGATTACAGCCTGAGTGCTGGGGTATTC
718	CGTWDVSLSVGVF	1542	TGCGGAACATGGGATGTCAGCCTGAGTGTTGGAGTGTTC
719	CGTWDTTLSAVVF	1543	TGCGGAACATGGGATACCACCCTGAGTGCTGTGGTTTTC
720	CGTWDTTLNIGVF	1544	TGCGGAACATGGGATACCACTCTGAATATTGGGGTGTTC
721	CGTWDTSLTAVVF	1545	TGCGGAACATGGGATACCAGCCTGACTGCTGTGGTATTC
722	CGTWDTSLTAAVF	1546	TGCGGAACCTGGGATACCAGCCTGACTGCTGCTGTGTTC
723	CGTWDTSLSVGLF	1547	TGCGGCACATGGGATACCAGCCTGAGTGTGGGGCTATTC
724	CGTWDTSLSGRVF	1548	TGCGGAACCTGGGATACCAGCCTGAGTGGTAGGGTGTTC
725	CGTWDTSLSGAVF	1549	TGCGGAACATGGGATACCAGCCTGAGTGGTGCAGTGTTC
726	CGTWDTSLSAGLF	1550	TGCGGAACATGGGATACCAGCCTGAGTGCTGGCCTGTTC
727	CGTWDTSLSAGGVF	1551	TGCGGAACATGGGATACCAGCCTGAGTGCTGGAGGGGTCTTC
728	CGTWDTSLRAYVF	1552	TGCGGAACATGGGATACCAGCCTGCGTGCTTATGTCTTC
729	CGTWDTSLRAWVF	1553	TGCGGAACATGGGATACTAGTTTGCGTGCTTGGGTATTC
730	CGTWDTSLNTGVF	1554	TGCGGAACATGGGATACCAGCCTGAATACTGGGGTATTC
731	CGTWDTSLNIWVF	1555	TGCGGAACATGGGATACCAGCCTGAATATTTGGGTGTTC
732	CGTWDTSLNIGVF	1556	TGCGGAACATGGGATACAAGCCTGAATATTGGGGTGTTC
733	CGTWDTSLIAVVF	1557	TGCGGAACATGGGATACCAGCCTGATTGCTGTGGTGTTC
734	CGTWDRSLSGWVF	1558	TGCGGAACGTGGGATAGGAGCCTGAGTGGTTGGGTGTTC
735	CGTWDNRLSGWVF	1559	TGCGGAACATGGGATAACAGGCTGAGTGGTTGGGTGTTC

736	CGTWDKSLSAVVF	1560	TGCGGAACGTGGGATAAGAGCCTGAGTGCTGTGGTCTTC
737	CGTWDKGLSAWVF	1561	TGCGGAACATGGGATAAAGGCCTGAGTGCTTGGGTGTTC
738	CGTWDISLSAGVF	1562	TGCGGAACATGGGATATCAGCCTGAGTGCTGGGGTGTTC
739	CGTWDESLSGGEVVF	1563	TGCGGAACATGGGATGAGAGCCTGAGTGGTGGCGAGGTGGTCTTC
740	CGTWDASLSAWVF	1564	TGCGGAACATGGGATGCCAGCCTGAGTGCCTGGGTGTTC
741	CGTWDAGLSAWVF	1565	TGCGGAACTTGGGATGCCGGCCTGAGTGCTTGGGTGTTC
742	CGAWDTSLSAWVF	1566	TGCGGAGCATGGGATACCAGCCTGAGTGCTTGGGTGTTC
743	CGAWDTSLSAVVF	1567	TGCGGAGCATGGGATACCAGCCTGAGTGCTGTGGTGTTC
744	CGAWDTSLRAGVF	1568	TGCGGAGCATGGGATACCAGCCTGCGTGCTGGGGTTTTC
745	CATWDTSVSAWVF	1569	TGCGCAACATGGGATACCAGCGTGAGTGCTTGGGTGTTC
746	CATWDTSLSAWVF	1570	TGCGCAACATGGGATACCAGCCTGAGTGCGTGGGTGTTC
747	CATWDNTLSAGVF	1571	TGCGCAACATGGGACAACACCCTGAGTGCTGGGGTGTTC
748	CAAWDRSLSVWVF	1572	TGCGCAGCATGGGATAGGAGCCTGAGTGTTTGGGTGTTC
749	CYTWHSSLRGGVF	1573	TGCTACACATGGCATTCCAGTCTGCGTGGTGGGGTGTTC
750	CVTWTSSPSAWVF	1574	TGCGTAACGTGGACTAGTAGCCCGAGTGCTTGGGTGTTC
751	CVTWRGGLVLF	1575	TGCGTGACATGGCGTGGTGGCCTTGTGTTGTTC
752	CVTWDTSLTSVVL	1576	TGCGTAACATGGGATACCAGCCTGACTTCTGTGGTACTC
753	CVTWDTSLSVYWVF	1577	TGCGTAACATGGGATACCAGCCTGAGTGTTTATTGGGTGTTC
754	CVTWDTSLSAWVF	1578	TGCGTTACATGGGATACCAGCCTGAGTGCCTGGGTGTTC
755	CVTWDTDLSVALF	1579	TGCGTCACATGGGATACCGACCTCAGCGTTGCGCTCTTC
756	CVTWDRSLSGWVF	1580	TGCGTAACATGGGATAGGAGCCTGAGTGGTTGGGTGTTC
757	CVTWDRSLREVLF	1581	TGCGTAACATGGGATCGCAGCCTGAGAGAGGTGTTATTC
758	CVTWDRSLRAVVF	1582	TGCGTAACATGGGATCGCAGCCTGAGAGCGGTGGTATTC
759	CVTWDRSLDAGVF	1583	TGCGTAACATGGGACAGGAGCCTCGATGCTGGGGTTTTC
760	CVTWDNTLSAGVF	1584	TGCGTGACATGGGATAACACCCTGAGTGCTGGGGTCTTC
761	CVTWDNNLFGVVF	1585	TGCGTAACATGGGATAACAACCTGTTTGGTGTGGTCTTC
762	CVSWDTSLSGAVF	1586	TGCGTATCATGGGATACCAGCCTGAGTGGTGCGGTATTC
763	CVSWDTSLSAGVF	1587	TGCGTCTCATGGGATACCAGCCTGAGTGCTGGGGTATTC
764	CTTWFRTPSDVVF	1588	TGCACAACATGGTTTAGGACTCCGAGTGATGTGGTCTTC
765	CTTWFRTASDVVF	1589	TGCACAACATGGTTTAGGACTGCGAGTGATGTGGTCTTC
766	CTTWDYGLSVVF	1590	TGCACAACGTGGGATTACGGTCTGAGTGTCGTCTTC
767	CTARDTSLSPGGVF	1591	TGCACAGCAAGGGATACCAGCCTGAGTCCTGGCGGGGTCTTC
768	CSTWNTRPSDVVF	1592	TGCTCAACATGGAATACGAGGCCGAGTGATGTGGTGTTC
769	CSTWESSLTTVVF	1593	TGTTCAACATGGGAGAGCAGTTTGACTACTGTGGTCTTC
770	CSTWDTSLTNVLF	1594	TGCTCAACATGGGATACCAGCCTCACTAATGTGCTATTC
771	CSTWDTSLSGVVF	1595	TGCTCAACATGGGATACCAGCCTGAGTGGAGTAGTCTTC
772	CSTWDHSLKAALF	1596	TGCTCAACATGGGATCACAGCCTGAAAGCTGCACTGTTC
773	CSTWDARLSVRVF	1597	TGCTCAACCTGGGATGCGAGGCTGAGTGTCCGGGTGTTC
774	CSSYTSSSTWVF	1598	TGCTCCTCATATACAAGCAGCAGCACTTGGGTGTTC
775	CSSYATRGLRVLF	1599	TGCAGCTCATACGCAACCCGCGGCCTTCGTGTGTTGTTC
776	CSSWDATLSVRIF	1600	TGTTCATCATGGGACGCCACCCTGAGTGTTCGCATATTC
777	CQVWEGSSDHWVF	1601	TGTCAGGTGTGGGAGGGTAGTAGTGATCATTGGGTGTTC
778	CQTWDNRLSAVVF	1602	TGCCAAACCTGGGATAACAGACTGAGTGCTGTGGTGTTC
779	CQTWDHSLHVGVF	1603	TGTCAAACGTGGGATCACAGCCTGCATGTTGGGGTGTTC
780	CQSYDDILNVWVL	1604	TGCCAGTCCTATGACGACATCTTGAATGTTTGGGTCCTT

781	CNTWDKSLTSELF	1605	TGCAATACATGGGATAAGAGTTTGACTTCTGAACTCTTC
782	CLTWDRSLNVRVF	1606	TGCTTAACATGGGATCGCAGCCTGAATGTGAGGGTGTTC
783	CLTWDHSLTAYVF	1607	TGCCTAACATGGGACCACAGCCTGACTGCTTATGTCTTC
784	CLTRDTSLSAPVF	1608	TGCTTAACAAGGGATACCAGTCTGAGTGCCCCTGTGTTC
785	CKTWESGLNFGHVF	1609	TGCAAAACATGGGAAAGTGGCCTTAATTTTGGCCACGTCTTC
786	CKTWDTSLSAVVF	1610	TGCAAAACATGGGATACCAGCCTGAGTGCTGTGGTCTTC
787	CGVWDVSLGAGVF	1611	TGCGGAGTCTGGGATGTCAGTCTGGGTGCTGGGGTGTTC
788	CGVWDTTPSAVLF	1612	TGCGGAGTCTGGGATACCACCCCGAGTGCCGTTCTTTTC
789	CGVWDTTLSAVLF	1613	TGCGGAGTCTGGGATACCACCCTGAGTGCCGTTCTTTTC
790	CGVWDTSLGVF	1614	TGCGGAGTATGGGATACCAGCCTGGGGGTCTTC
791	CGVWDTNLGKWVF	1615	TGCGGGGTATGGGATACCAACCTGGGTAAATGGGTTTTC
792	CGVWDTGLDAGWVF	1616	TGTGGAGTTTGGGATACTGGCCTGGATGCTGGTTGGGTGTTC
793	CGVWDNVLEAYVF	1617	TGCGGAGTGTGGGATAACGTCCTGGAGGCCTATGTCTTC
794	CGVWDISLSANWVF	1618	TGCGGAGTCTGGGATATCAGCCTGAGTGCTAATTGGGTGTTC
795	CGVWDHSLGIWAF	1619	TGCGGAGTATGGGATCACAGCCTGGGGATTTGGGCCTTC
796	CGVWDDILTAEVF	1620	TGCGGAGTTTGGGATGATATTCTGACTGCTGAAGTGTTC
797	CGVRDTSLGVF	1621	TGCGGAGTTCGGGATACCAGCCTGGGGGTCTTC
798	CGTYDTSLPAWVF	1622	TGCGGAACATACGATACGAGCCTGCCTGCTTGGGTGTTT
799	CGTYDNLVFGYVF	1623	TGCGGAACTTACGATAATCTTGTATTTGGTTATGTCTTC
800	CGTYDDRLREVF	1624	TGCGGAACATACGATGATAGACTCAGAGAGGTGTTC
801	CGTWVTSLSAGVF	1625	TGCGGAACGTGGGTTACCAGCCTGAGTGCTGGGGTGTTC
802	CGTWVSSLTTVVF	1626	TGCGGAACATGGGTTAGCAGCCTGACTACTGTAGTATTC
803	CGTWVSSLNVWVF	1627	TGCGGAACATGGGTTAGCAGCCTGAACGTCTGGGTGTTC
804	CGTWVGRFWVF	1628	TGCGGAACATGGGTTGGCAGGTTTTGGGTATTC
805	CGTWSGGPSGHWLF	1629	TGCGGAACATGGTCTGGCGGCCCGAGTGGCCATTGGTTGTTC
806	CGTWSGGLSGHWLF	1630	TGCGGAACATGGTCTGGCGGCCTGAGTGGCCATTGGTTGTTC
807	CGTWQTGREAVLF	1631	TGCGGAACGTGGCAGACCGGCCGGGAGGCTGTCCTATTT
808	CGTWQSRLRWVF	1632	TGCGGAACGTGGCAGAGCAGGCTGAGGTGGGTGTTC
809	CGTWQSRLGWVF	1633	TGCGGAACGTGGCAGAGCAGGCTGGGGTGGGTGTTC
810	CGTWPRSLSAVWVF	1634	TGCGGAACATGGCCTAGGAGCCTGAGTGCTGTTTGGGTGTTC
811	CGTWNNYLSAGDVVF	1635	TGCGGAACATGGAATAACTACCTGAGTGCTGGCGATGTGGTTTTC
812	CGTWLGSQSPYWVF	1636	TGCGGAACATGGCTTGGCAGCCAGAGTCCTTATTGGGTCTTC
813	CGTWHTGLSAYVF	1637	TGCGGAACATGGCATACCGGCCTGAGTGCTTATGTCTTC
814	CGTWHSTLSAGHWVF	1638	TGCGGAACATGGCATAGTACCCTGAGTGCTGGCCATTGGGTGTTC
815	CGTWHSSLSTWVF	1639	TGCGGAACATGGCATAGTAGCCTGAGTACTTGGGTGTTC
816	CGTWHSSLSAYVF	1640	TGCGGAACATGGCATAGCAGCCTGAGTGCCTATGTCTTC
817	CGTWHSSLSAVVF	1641	TGCGGAACATGGCATAGCAGCCTGAGTGCTGTGGTATTC
818	CGTWHSGLSGWVF	1642	TGCGGAACGTGGCATTCCGGCCTGAGTGGGTGGGTTTTC
819	CGTWHNTLRNVIF	1643	TGCGGAACATGGCATAACACCCTGCGTAATGTGATATTC
820	CGTWHASLTAVF	1644	TGCGGAACATGGCATGCCAGCCTGACTGCTGTGTTC
821	CGTWGWYGSQRGVVF	1645	TGCGGGACATGGGGATGGTATGGCAGCCAGAGAGGCGTCGTCTTC
822	CGTWGWYGGQRGVVF	1646	TGCGGGACATGGGGATGGTATGGCGGCCAGAGAGGCGTCGTCTTC
823	CGTWGTSLSAWVF	1647	TGCGGAACCTGGGGAACCAGCCTGAGTGCTTGGGTGTTC

824	CGTWGSSLTTGLF	1648	TGCGGAACCTGGGGTAGCAGCCTGACTACTGGCCTGTTC
825	CGTWGSSLTAYVF	1649	TGCGGAACATGGGGTAGCAGCCTGACTGCCTATGTCTTC
826	CGTWGSSLSVVF	1650	TGCGGAACATGGGGTAGCAGCCTGAGTGTTGTGTTC
827	CGTWGSSLSGGVF	1651	TGCGGAACATGGGGTAGCAGCCTGAGTGGTGGGGTGTTC
828	CGTWGSSLSAYWVF	1652	TGCGGAACATGGGGTAGCAGCCTGAGTGCTTATTGGGTGTTC
829	CGTWGSSLSAYVVF	1653	TGCGGAACATGGGGTAGCAGCCTGAGTGCTTATGTGGTGTTC
830	CGTWGSSLSAYVF	1654	TGCGGAACATGGGGTAGCAGCCTGAGTGCTTATGTCTTC
831	CGTWGSSLSAVVF	1655	TGCGGAACGTGGGGTAGTAGCCTGAGTGCTGTGGTGTTC
832	CGTWGSSLSAPYVF	1656	TGCGGAACATGGGGTAGCAGCCTGAGTGCTCCTTATGTCTTC
833	CGTWGSSLSAPVF	1657	TGCGGAACATGGGGTAGCAGCCTGAGTGCCCCGGTGTTC
834	CGTWGSSLSAGVF	1658	TGCGGAACATGGGGTAGCAGCCTGAGTGCTGGGGTGTTC
835	CGTWGSSLSAGLF	1659	TGCGGAACTTGGGGTAGCAGCCTGAGTGCTGGACTGTTC
836	CGTWGSSLSAGALF	1660	TGCGGAACATGGGGTAGCAGCCTGAGTGCTGGGGCACTCTTC
837	CGTWGSSLRAWVF	1661	TGCGGAACATGGGGCAGTAGCCTGCGTGCTTGGGTGTTC
838	CGTWFTSLASGVF	1662	TGCGGAACCTGGTTTACTAGTCTGGCTAGTGGGGTTTTC
839	CGTWETSLSVVVI	1663	TGCGGAACTTGGGAGACCAGTCTGAGTGTCGTGGTCATC
840	CGTWETSLSGVF	1664	TGCGGAACATGGGAGACCAGCCTGAGTGGTGTCTTC
841	CGTWETSLSDWVF	1665	TGCGGAACATGGGAAACCAGCCTGAGTGATTGGGTATTC
842	CGTWETSLSAGVF	1666	TGCGGAACATGGGAGACCAGCCTGAGTGCTGGGGTATTC
843	CGTWETSLNYVAF	1667	TGCGGAACATGGGAAACCAGCCTTAATTATGTGGCCTTC
844	CGTWETSLNTWLL	1668	TGCGGAACATGGGAGACCAGCCTGAATACTTGGTTGCTC
845	CGTWETSESGNYIF	1669	TGCGGAACATGGGAGACCAGCGAGAGTGGTAATTACATCTTC
846	CGTWETRLGTWVI	1670	TGCGGAACATGGGAAACCAGACTGGGTACTTGGGTGATC
847	CGTWETQLYWVF	1671	TGCGGAACATGGGAGACCCAGTTATATTGGGTGTTC
848	CGTWETGLSAGEVF	1672	TGCGGAACATGGGAGACTGGCCTAAGTGCTGGAGAGGTGTTC
849	CGTWESTLSVFLF	1673	TGCGGAACTTGGGAAAGCACCCTGAGTGTTTTCCTATTC
850	CGTWESSLTVVVF	1674	TGCGGGACATGGGAAAGTAGCCTGACTGTTGTGGTCTTC
851	CGTWESSLTGVVF	1675	TGCGGAACATGGGAAAGTAGCCTGACTGGAGTGGTATTC
852	CGTWESSLTGFVF	1676	TGCGGAACATGGGAAAGCAGCCTGACTGGTTTTGTCTTC
853	CGTWESSLSVGVF	1677	TGTGGAACATGGGAGAGCAGCCTGAGTGTTGGGGTGTTC
854	CGTWESSLSEWVF	1678	TGCGGAACCTGGGAAAGTAGCCTCAGTGAATGGGTGTTC
855	CGTWESSLSAVF	1679	TGCGGAACATGGGAGAGCAGCCTGAGTGCTGTATTC
856	CGTWESSLSAGYIF	1680	TGCGGAACATGGGAGAGCAGCCTGAGTGCTGGTTATATCTTC
857	CGTWESSLSAGVF	1681	TGCGGAACATGGGAGAGCAGCCTGAGTGCTGGAGTGTTC
858	CGTWESSLSAGPVF	1682	TGCGGAACATGGGAAAGCAGCCTGAGCGCTGGCCCGGTGTTC
859	CGTWESSLSAGGQVF	1683	TGCGGAACATGGGAAAGCAGCCTGAGTGCTGGAGGCCAGGTGTTC
860	CGTWESSLSAFGGYVF	1684	TGCGGAACATGGGAGAGCAGCCTGAGTGCCTTCGGCGGTTATGTCTTC
861	CGTWESSLRVWVF	1685	TGCGGAACATGGGAAAGCAGCCTGAGGGTTTGGGTGTTC
862	CGTWESSLFTGPWVF	1686	TGCGGAACATGGGAAAGCAGCCTCTTTACTGGGCCTTGGGTGTTC
863	CGTWESLSATYVF	1687	TGCGGAACATGGGAGAGCCTGAGTGCCACCTATGTCTTC
864	CGTWESGLSAGVF	1688	TGCGGAACATGGGAGAGCGGCCTGAGTGCTGGTGTCTTC
865	CGTWESDFWVF	1689	TGCGGAACATGGGAAAGCGACTTTTGGGTGTTT
866	CGTWENRLSAVVF	1690	TGCGGTACATGGGAAAACAGACTGAGTGCTGTGGTCTTC
867	CGTWENRLSAGVF	1691	TGCGGAACATGGGAAAACAGACTGAGTGCCGGGGTATTC

868	CGTWEISLTTSVVF	1692	TGCGGAACATGGGAAATCAGCCTGACTACTTCTGTGGTATTC
869	CGTWEISLSTSVVF	1693	TGCGGAACATGGGAAATCAGCCTGAGTACTTCTGTGGTATTC
870	CGTWEGSLSVVF	1694	TGCGGAACATGGGAAGGCAGCCTCAGTGTTGTTTTC
871	CGTWEGSLRVF	1695	TGCGGAACATGGGAAGGCAGCCTGAGGGTGTTC
872	CGTWEGSLRHVF	1696	TGCGGAACATGGGAGGGCAGCCTGAGGCACGTGTTC
873	CGTWDYSPVRAGVF	1697	TGCGGAACATGGGATTACAGCCCTGTACGTGCTGGGGTGTTC
874	CGTWDYSLSVYLF	1698	TGCGGAACGTGGGATTACAGCCTGAGTGTTTATCTCTTC
875	CGTWDYSLSSGVVF	1699	TGCGGAACATGGGATTACAGCCTGAGTTCTGGCGTGGTATTC
876	CGTWDYSLSAWVF	1700	TGCGGAACATGGGATTACAGCCTGAGTGCCTGGGTGTTC
877	CGTWDYSLSAEVF	1701	TGCGGAACATGGGATTACAGTCTGAGTGCTGAGGTGTTC
878	CGTWDYSLRRAIF	1702	TGCGGAACATGGGATTACAGCCTGCGTCGTGCGATATTC
879	CGTWDWSLILQLF	1703	TGCGGAACATGGGATTGGAGCCTCATTCTTCAATTGTTC
880	CGTWDVTLHTGVF	1704	TGCGGAACATGGGATGTCACCTTGCATACTGGGGTGTTC
881	CGTWDVTLHIGVF	1705	TGCGGAACATGGGATGTCACCTTGCATATTGGGGTGTTC
882	CGTWDVTLHAGVF	1706	TGCGGAACATGGGATGTCACCTTGCATGCTGGGGTGTTC
883	CGTWDVSLYSGGVF	1707	TGCGGAACATGGGATGTCAGTTTGTATAGTGGCGGGGTCTTC
884	CGTWDVSLTSFVF	1708	TGTGGAACATGGGATGTCAGCCTGACTTCTTTCGTCTTC
885	CGTWDVSLSVGVL	1709	TGCGGAACATGGGATGTCAGCCTGAGTGTTGGGGTGCTC
886	CGTWDVSLSAGDVVF	1710	TGCGGAACGTGGGATGTCAGCCTGAGTGCTGGCGATGTAGTTTTC
887	CGTWDVSLNVVVF	1711	TGCGGAACATGGGATGTCAGCCTGAATGTCGTGGTTTTC
888	CGTWDVSLNTQVF	1712	TGCGGAACATGGGATGTCAGCCTGAATACTCAGGTGTTC
889	CGTWDVSLGALF	1713	TGCGGCACATGGGATGTGAGCCTGGGTGCGCTGTTC
890	CGTWDVNLKTVVF	1714	TGCGGAACGTGGGACGTTAATCTGAAAACTGTCGTTTTC
891	CGTWDVILSAEVF	1715	TGCGGAACATGGGATGTCATCCTGAGTGCTGAGGTATTC
892	CGTWDTTVSAVVF	1716	TGCGGAACATGGGATACCACCGTGAGTGCTGTGGTTTTC
893	CGTWDTTLTAWVF	1717	TGCGGAACATGGGATACCACCCTGACTGCCTGGGTGTTC
894	CGTWDTTLSVFLF	1718	TGCGGAACATGGGACACCACCTTGAGTGTTTTCCTATTC
895	CGTWDTSVSAGVF	1719	TGCGGGACTTGGGATACCAGTGTGAGTGCTGGGGTGTTC
896	CGTWDTSVISWVF	1720	TGCGGAACATGGGATACCAGTGTGATTTCTTGGGTTTTC
897	CGTWDTSRSSLYVVF	1721	TGCGGAACATGGGATACCAGTCGGAGTTCTCTCTATGTGGTCTTC
898	CGTWDTSRSAWVF	1722	TGCGGAACATGGGATACCAGCCGGAGTGCTTGGGTATTC
899	CGTWDTSRNPGGIF	1723	TGCGGAACATGGGATACCAGCCGGAATCCTGGAGGAATTTTC
900	CGTWDTSRGHVF	1724	TGCGGAACATGGGACACCAGTCGGGGTCATGTTTTC
901	CGTWDTSPSTGQVLF	1725	TGCGGAACATGGGATACCAGCCCGAGTACTGGCCAGGTGCTTTTC
902	CGTWDTSPSAWVF	1726	TGCGGAACATGGGATACCAGCCCGAGTGCCTGGGTGTTC
903	CGTWDTSLTWVF	1727	TGCGGAACATGGGATACTAGCCTGACCTGGGTGTTC
904	CGTWDTSLTWFAVF	1728	TGCGGAACATGGGATACCAGCCTGACGTGGTTCGCAGTGTTC
905	CGTWDTSLTVVVF	1729	TGCGGAACATGGGATACCAGCCTGACTGTTGTGGTATTC
906	CGTWDTSLTTSWVF	1730	TGCGGAACATGGGATACCAGCCTGACTACTTCTTGGGTGTTC
907	CGTWDTSLTTGPFWVF	1731	TGCGGAACATGGGATACCAGCCTGACCACTGGTCCTTTTTGGGTGTTC
908	CGTWDTSLTPFYVF	1732	TGCGGAACATGGGATACCAGCCTGACTCCTTTTTATGTCTTC
909	CGTWDTSLTAYVF	1733	TGCGGAACATGGGATACCAGCCTGACTGCTTATGTCTTC
910	CGTWDTSLTAWVF	1734	TGCGGAACATGGGATACCAGCCTGACTGCTTGGGTGTTC

911	CGTWDTSLTAWGVF	1735	TGCGGAACATGGGATACCAGCCTGACTGCGTGGGGGGTGTTC
912	CGTWDTSLTAVVL	1736	TGCGGCACATGGGATACCAGCCTGACTGCGGTGGTTCTC
913	CGTWDTSLTARVF	1737	TGCGGAACCTGGGATACCAGCCTGACTGCTCGGGTTTTC
914	CGTWDTSLTAIVF	1738	TGCGGAACATGGGATACCAGCCTGACTGCGATTGTCTTC
915	CGTWDTSLTAGVF	1739	TGCGGAACATGGGATACCAGCCTGACTGCTGGTGTCTTC
916	CGTWDTSLSVYVF	1740	TGCGGAACATGGGATACCAGCCTGAGTGTTTATGTCTTC
917	CGTWDTSLSVVF	1741	TGCGGAACATGGGATACCAGCCTGAGTGTGGTGTTC
918	CGTWDTSLSVGEF	1742	TGCGGGACATGGGATACCAGCCTGAGTGTTGGGGAATTC
919	CGTWDTSLSTWVF	1743	TGCGGAACATGGGATACCAGCCTGAGTACTTGGGTGTTC
920	CGTWDTSLSTVVF	1744	TGCGGAACATGGGATACCAGCCTGAGTACTGTGGTATTC
921	CGTWDTSLSTGQVLF	1745	TGCGGAACATGGGATACCAGCCTGAGTACTGGCCAGGTGCTTTTC
922	CGTWDTSLSTGPLWVF	1746	TGCGGCACATGGGATACCAGCCTGAGCACTGGTCCTCTTTGGGTGTTC
923	CGTWDTSLSSYVF	1747	TGCGGAACTTGGGATACCAGCCTGAGTTCTTATGTCTTC
924	CGTWDTSLSSVVF	1748	TGCGGAACATGGGATACCAGCCTGAGTTCTGTGGTCTTC
925	CGTWDTSLSSRYIF	1749	TGCGGAACATGGGATACCAGCCTGAGTTCTAGATACATATTC
926	CGTWDTSLSSRFIF	1750	TGCGGAACATGGGATACCAGCCTGAGTTCTAGATTCATATTC
927	CGTWDTSLSSGWVF	1751	TGCGGAACATGGGATACCAGCCTGAGTTCTGGGTGGGTGTTC
928	CGTWDTSLSRYVF	1752	TGCGGAACATGGGATACCAGCCTGAGTCGGTATGTGTTC
929	CGTWDTSLSQWLF	1753	TGCGGAACTTGGGATACCAGTCTGAGTCAATGGCTGTTC
930	CGTWDTSLSPGLWVF	1754	TGCGGAACATGGGATACCAGCCTGAGTCCTGGCCTTTGGGTGTTC
931	CGTWDTSLSNYVF	1755	TGCGGAACATGGGATACCAGCCTGAGTAATTATGTCTTC
932	CGTWDTSLSIWVF	1756	TGCGGAACATGGGATACCAGCCTAAGTATTTGGGTGTTC
933	CGTWDTSLSIGPFWVF	1757	TGCGGCACATGGGATACCAGCCTGAGCATTGGTCCTTTTTGGGTGTTC
934	CGTWDTSLSGWVF	1758	TGCGGAACATGGGATACCAGCCTGAGTGGTTGGGTGTTC
935	CGTWDTSLSGTVF	1759	TGCGGAACATGGGATACCAGCCTGAGTGGTACAGTGTTC
936	CGTWDTSLSGGQVF	1760	TGCGGAACATGGGATACTAGTCTGAGTGGTGGCCAGGTGTTC
937	CGTWDTSLSGGIF	1761	TGCGGAACATGGGATACCAGCCTGAGTGGTGGGATATTC
938	CGTWDTSLSGEDVVI	1762	TGCGGAACATGGGATACCAGCCTGAGTGGTGAGGATGTGGTAATC
939	CGTWDTSLSFLYAF	1763	TGCGGAACATGGGATACCAGCCTGAGTTTCCTTTATGCTTTC
940	CGTWDTSLSEVVF	1764	TGCGGAACATGGGATACCAGCCTGAGTGAGGTCGTATTC
941	CGTWDTSLSEVF	1765	TGCGGAACATGGGATACCAGCCTGAGTGAAGTGTTC
942	CGTWDTSLSENWVF	1766	TGCGGAACATGGGATACTAGCCTGAGTGAAAATTGGGTGTTC
943	CGTWDTSLSAYIF	1767	TGCGGAACATGGGATACCAGCCTGAGTGCCTACATATTC
944	CGTWDTSLSAVVL	1768	TGCGGAACATGGGATACCAGCCTGAGTGCTGTGGTACTC
945	CGTWDTSLSAVF	1769	TGCGGAACATGGGATACCAGCCTGAGTGCTGTTTTC
946	CGTWDTSLSARVF	1770	TGCGGAACATGGGATACCAGCCTGAGTGCCCGGGTGTTC
947	CGTWDTSLSARQVF	1771	TGCGGCACATGGGATACCAGCCTGAGTGCCCGCCAGGTATTC
948	CGTWDTSLSALVF	1772	TGCGGAACATGGGATACCAGCCTGAGTGCTTTGGTTTTC
949	CGTWDTSLSAKVF	1773	TGCGGAACATGGGATACCAGCCTGAGTGCTAAGGTGTTC
950	CGTWDTSLSAKIF	1774	TGCGGAACATGGGATACCAGCCTGAGTGCGAAAATCTTC
951	CGTWDTSLSAKAVF	1775	TGCGGAACATGGGATACCAGCCTGAGTGCCAAGGCGGTATTC
952	CGTWDTSLSAHAVF	1776	TGCGGAACATGGGATACCAGCCTGAGTGCCCATGCTGTGTTC
953	CGTWDTSLSAGYVF	1777	TGCGGAACATGGGATACCAGCCTGAGTGCTGGCTATGTCTTC

954	CGTWDTSLSAGRWVF	1778	TGCGGAACATGGGACACCAGTCTGAGTGCTGGCCGCTGGGTGTTC
955	CGTWDTSLSAGIF	1779	TGCGGAACATGGGATACCAGCCTGAGTGCTGGGATATTC
956	CGTWDTSLSAGGFRVF	1780	TGCGGAACATGGGATACCAGCCTGAGTGCTGGTGGGTTCCGGGTCTTC
957	CGTWDTSLSAGAF	1781	TGCGGAACATGGGATACCAGCCTGAGTGCTGGGGCATTC
958	CGTWDTSLSADWFF	1782	TGCGGAACATGGGATACCAGTCTGAGTGCTGATTGGTTTTTC
959	CGTWDTSLSADEYVF	1783	TGCGGAACATGGGATACCAGCCTGAGTGCTGATGAATATGTCTTC
960	CGTWDTSLSAAWVF	1784	TGCGGCACATGGGATACCAGCCTGAGTGCGGCTTGGGTGTTC
961	CGTWDTSLSAALF	1785	TGCGGAACATGGGATACCAGCCTGAGTGCTGCGCTATTC
962	CGTWDTSLSAAGVF	1786	TGCGGAACATGGGATACCAGCCTGAGTGCTGCGGGGGTTTTC
963	CGTWDTSLRVVVF	1787	TGCGGAACATGGGATACCAGCCTGAGAGTTGTGGTTTTC
964	CGTWDTSLRTWVF	1788	TGCGGAACATGGGATACCAGCCTGAGAACCTGGGTATTC
965	CGTWDTSLRGAVF	1789	TGCGGAACGTGGGATACCAGCCTGAGGGGTGCAGTGTTC
966	CGTWDTSLRAVVF	1790	TGCGGAACATGGGATACCAGCCTGCGTGCTGTGGTATTC
967	CGTWDTSLNVVYVF	1791	TGCGGAACATGGGATACAAGCCTGAATGTAGTTTATGTCTTC
968	CGTWDTSLNTYLF	1792	TGCGGAACATGGGATACCAGCCTCAACACCTACCTGTTC
969	CGTWDTSLNFAWLF	1793	TGCGGAACATGGGATACTAGCCTGAACTTCGCTTGGCTGTTC
970	CGTWDTSLLVWLF	1794	TGCGGCACATGGGATACCAGCCTTCTTGTGTGGCTTTTC
971	CGTWDTSLKTWVF	1795	TGCGGAACATGGGATACCAGTCTGAAGACGTGGGTGTTC
972	CGTWDTSLIVWVF	1796	TGCGGAACATGGGATACCAGTCTGATTGTCTGGGTGTTC
973	CGTWDTSLITGVF	1797	TGCGGAACATGGGATACCAGCCTAATTACTGGGGTGTTC
974	CGTWDTSLISVVF	1798	TGCGGAACATGGGATACCAGCCTGATTAGCGTGGTATTC
975	CGTWDTSLIAYVF	1799	TGCGGAACATGGGATACCAGCCTGATTGCTTATGTCTTC
976	CGTWDTSLHTELF	1800	TGCGGAACATGGGATACCAGCCTGCACACTGAGTTGTTC
977	CGTWDTSLGSYVF	1801	TGCGGAACTTGGGATACCAGCCTGGGTTCTTATGTCTTC
978	CGTWDTSLGSLWVF	1802	TGCGGAACATGGGATACCAGCCTGGGTTCTCTTTGGGTGTTC
979	CGTWDTSLGSGVF	1803	TGCGGTACATGGGATACCAGCCTGGGTTCTGGGGTATTC
980	CGTWDTSLGGRGVF	1804	TGCGGAACTTGGGATACCAGTCTGGGTGGTAGAGGGGTCTTC
981	CGTWDTSLGAWVF	1805	TGCGGAACATGGGATACCAGCCTGGGTGCTTGGGTGTTC
982	CGTWDTSLGAVVF	1806	TGCGGAACATGGGATACCAGCCTGGGTGCCGTGGTATTC
983	CGTWDTSLGAGVF	1807	TGCGGAACATGGGATACCAGCCTGGGTGCTGGGGTATTC
984	CGTWDTSLGAGLF	1808	TGCGGAACATGGGATACCAGCCTGGGTGCTGGCCTATTC
985	CGTWDTSLDAVVF	1809	TGCGGAACATGGGATACCAGTCTGGATGCTGTGGTTTTC
986	CGTWDTSLDAVLF	1810	TGCGGGACTTGGGATACCAGCCTGGATGCTGTGCTGTTC
987	CGTWDTSLAWVF	1811	TGCGGAACATGGGATACCAGCCTGGCTTGGGTGTTC
988	CGTWDTSLATGLF	1812	TGCGGAACATGGGATACCAGCCTGGCGACTGGACTGTTC
989	CGTWDTSLAPVVF	1813	TGCGGGACATGGGATACCAGCCTGGCCCCTGTAGTCTTC
990	CGTWDTRLTIVIF	1814	TGCGGAACATGGGACACCCGCCTGACTATTGTGATCTTC
991	CGTWDTRLSVWLF	1815	TGTGGAACATGGGACACCAGGCTGAGTGTTTGGCTGTTC
992	CGTWDTRLSVGVF	1816	TGCGGAACGTGGGACACCAGACTGAGTGTTGGGGTTTTC
993	CGTWDTRLSTVIF	1817	TGCGGCACATGGGATACCAGACTGAGTACTGTAATTTTC
994	CGTWDTRLSSVVF	1818	TGCGGAACATGGGATACCCGCCTGAGTTCTGTGGTCTTC
995	CGTWDTRLSIVVF	1819	TGCGGAACATGGGATACCCGCCTGAGTATTGTGGTTTTC
996	CGTWDTRLSAYVVF	1820	TGCGGAACATGGGATACCAGACTGAGTGCCTATGTGGTATTC
997	CGTWDTRLSAWVF	1821	TGCGGAACCTGGGACACCCGCCTGAGTGCGTGGGTGTTC

998	CGTWDTRLSAVVF	1822	TGCGGAACATGGGATACCAGACTGAGTGCTGTGGTGTTC
999	CGTWDTRLSAGLF	1823	TGCGGAACATGGGATACCCGCCTGAGTGCTGGGTTGTTC
1000	CGTWDTRLSAGGVF	1824	TGCGGAACATGGGATACCAGACTGAGTGCTGGTGGGGTGTTC
1001	CGTWDTRLNVWLF	1825	TGCGGAACATGGGATACCAGATTGAATGTGTGGCTATTC
1002	CGTWDTNREVVLL	1826	TGCGGAACATGGGATACCAACCGGGAAGTTGTGCTCCTC
1003	CGTWDTNLRAHVF	1827	TGCGGAACATGGGATACCAACCTGCGTGCCCATGTCTTC
1004	CGTWDTNLPAVVF	1828	TGCGGAACATGGGATACTAATCTGCCCGCTGTAGTGTTC
1005	CGTWDTNLGGVF	1829	TGCGGAACATGGGACACCAATTTGGGTGGGGTGTTC
1006	CGTWDTIVSIGVF	1830	TGCGGAACATGGGATACCATCGTGAGTATTGGGGTGTTC
1007	CGTWDTILSAVVF	1831	TGCGGAACATGGGATACCATCCTGAGTGCGGTGGTGTTC
1008	CGTWDTILSAEVF	1832	TGCGGCACATGGGATACCATCCTGAGTGCTGAGGTGTTC
1009	CGTWDTHLGVVF	1833	TGCGGAACATGGGATACCCACCTGGGTGTGGTTTTC
1010	CGTWDTGPSPHWLF	1834	TGCGGAACATGGGATACCGGCCCGAGCCCTCATTGGCTGTTC
1011	CGTWDTGLTFGGVF	1835	TGCGGAACATGGGATACCGGCCTGACTTTTGGAGGCGTGTTC
1012	CGTWDTGLTAFVF	1836	TGCGGAACATGGGATACCGGCCTGACTGCTTTTGTCTTC
1013	CGTWDTGLSVWVF	1837	TGCGGAACATGGGATACCGGCCTGAGTGTTTGGGTGTTC
1014	CGTWDTGLSTGIF	1838	TGCGGAACATGGGATACCGGCCTGAGTACTGGGATTTTC
1015	CGTWDTGLSSLLF	1839	TGCGGAACATGGGATACCGGCCTGAGTTCCCTGCTCTTC
1016	CGTWDTGLSIVVF	1840	TGCGGAACGTGGGACACCGGCCTGAGTATTGTGGTGTTC
1017	CGTWDTGLSFVVF	1841	TGCGGAACGTGGGACACCGGCCTGAGTTTTGTGGTGTTC
1018	CGTWDTGLSAWVF	1842	TGCGGAACATGGGATACCGGCCTGAGTGCTTGGGTGTTC
1019	CGTWDTGLSAGVVF	1843	TGCGGAACATGGGATACCGGCCTGAGTGCTGGTGTGGTATTC
1020	CGTWDTGLRGWIF	1844	TGCGGAACATGGGATACCGGTCTGAGGGGTTGGATTTTC
1021	CGTWDTELSAGVF	1845	TGCGGAACATGGGATACCGAGCTAAGTGCGGGGGTCTTC
1022	CGTWDTALTAGVF	1846	TGCGGAACGTGGGATACCGCCCTGACTGCTGGGGTGTTC
1023	CGTWDTALSLVVF	1847	TGCGGAACATGGGATACTGCCCTGAGTCTTGTGGTCTTC
1024	CGTWDTALSAWLF	1848	TGCGGAACATGGGATACCGCCCTGAGTGCCTGGCTGTTC
1025	CGTWDTALSAGVF	1849	TGCGGCACATGGGATACCGCCCTGAGTGCTGGGGTGTTC
1026	CGTWDTALRGVLF	1850	TGCGGAACATGGGATACCGCCCTGCGTGGCGTGCTGTTC
1027	CGTWDTALKEWLF	1851	TGCGGAACATGGGATACCGCCCTGAAAGAATGGCTGTTC
1028	CGTWDRTLTAGDVLF	1852	TGCGGAACATGGGATAGGACCCTGACTGCTGGCGATGTGCTCTTC
1029	CGTWDRSVTYVF	1853	TGCGGAACATGGGATAGAAGCGTGACTTATGTCTTC
1030	CGTWDRSRNEWVF	1854	TGCGGAACATGGGATCGCAGCCGAAATGAATGGGTGTTC
1031	CGTWDRSLTVWVF	1855	TGCGGAACATGGGATCGCAGTCTGACTGTTTGGGTCTTC
1032	CGTWDRSLTPGWLF	1856	TGCGGAACATGGGATCGCAGCCTGACTCCTGGGTGGTTGTTC
1033	CGTWDRSLTAWVF	1857	TGCGGAACATGGGATAGAAGCCTGACTGCTTGGGTGTTC
1034	CGTWDRSLSVVVF	1858	TGCGGAACATGGGACCGCAGCCTGAGTGTTGTGGTATTC
1035	CGTWDRSLSVVF	1859	TGCGGCACATGGGATCGCAGCCTGAGTGTAGTCTTC
1036	CGTWDRSLSVQLF	1860	TGCGGAACATGGGATAGGAGCCTGAGTGTTCAATTGTTC
1037	CGTWDRSLSVLWVF	1861	TGCGGAACATGGGATCGCAGCCTCAGTGTTCTTTGGGTGTTC
1038	CGTWDRSLSVGLF	1862	TGCGGAACATGGGATCGCAGCCTGAGTGTTGGATTATTC
1039	CGTWDRSLSTWVF	1863	TGCGGAACATGGGATCGCAGCCTGAGTACTTGGGTGTTC
1040	CGTWDRSLSTHWVL	1864	TGCGGAACATGGGATAGAAGCCTGAGTACTCATTGGGTGCTC
1041	CGTWDRSLSTHWVF	1865	TGCGGAACATGGGATAGAAGCCTGAGTACTCATTGGGTGTTC
1042	CGTWDRSLSSAVF	1866	TGCGGAACCTGGGATCGAAGCCTGAGTTCTGCGGTGTTC

1043	CGTWDRSLSPSYVF	1867	TGCGGAACATGGGACAGAAGCCTGAGTCCCTCTTATGTCTTC
1044	CGTWDRSLSGEVF	1868	TGCGGAACATGGGATAGGAGCCTGAGTGGTGAGGTGTTC
1045	CGTWDRSLSGAVF	1869	TGCGGAACATGGGATAGGAGCCTGAGTGGTGCGGTGTTC
1046	CGTWDRSLSAVAF	1870	TGCGGAACATGGGATCGCAGCCTGAGTGCTGTGGCATTC
1047	CGTWDRSLSAGGEF	1871	TGCGGAACATGGGATAGGAGCCTGAGTGCCGGGGGGGAATTC
1048	CGTWDRSLSAFWVF	1872	TGCGGAACATGGGATCGCAGCCTGAGTGCTTTTTGGGTGTTC
1049	CGTWDRSLSAAVF	1873	TGCGGAACATGGGATAGGAGCCTGAGTGCTGCGGTGTTC
1050	CGTWDRSLSAALF	1874	TGCGGAACATGGGATAGGAGCCTGAGTGCTGCACTCTTC
1051	CGTWDRSLRVF	1875	TGCGGAACATGGGATCGCAGCCTGAGAGTGTTC
1052	CGTWDRSLNWVF	1876	TGCGGTACATGGGACAGAAGCCTTAATTGGGTGTTC
1053	CGTWDRSLNVYVF	1877	TGCGGAACATGGGATCGCAGCCTGAATGTTTATGTCTTC
1054	CGTWDRSLNVGVF	1878	TGCGGAACATGGGATAGGAGCCTGAATGTTGGGGTGTTC
1055	CGTWDRSLHVVF	1879	TGCGGAACATGGGATCGGAGCCTGCATGTGGTCTTC
1056	CGTWDRSLGGWVF	1880	TGTGGAACATGGGATCGCAGCCTGGGTGGTTGGGTGTTC
1057	CGTWDRSLGAFWVF	1881	TGCGGAACATGGGATCGCAGCCTGGGTGCTTTTTGGGTGTTC
1058	CGTWDRSLFWVF	1882	TGCGGAACATGGGATAGAAGCCTGTTTTGGGTGTTC
1059	CGTWDRSLAAGVF	1883	TGCGGAACGTGGGATCGCAGCCTGGCTGCTGGGGTGTTC
1060	CGTWDRRLSGVVF	1884	TGCGGAACATGGGATAGGAGGTTGAGTGGTGTCGTATTC
1061	CGTWDRRLSDVVF	1885	TGCGGAACGTGGGATCGCCGCCTAAGTGATGTGGTATTC
1062	CGTWDRRLSAVVF	1886	TGCGGAACATGGGATAGGAGGCTGAGTGCTGTGGTATTC
1063	CGTWDRRLNVAFF	1887	TGCGGAACATGGGATAGACGCCTGAATGTTGCGTTCTTC
1064	CGTWDRRLLAVF	1888	TGTGGAACATGGGATAGGAGGCTGCTTGCTGTTTTC
1065	CGTWDRNLRAVVF	1889	TGCGGAACTTGGGATAGGAACCTGCGCGCCGTGGTCTTC
1066	CGTWDRLSAGVF	1890	TGCGGAACATGGGATAGGCTGAGTGCTGGGGTGTTC
1067	CGTWDRGPNTGVF	1891	TGCGGAACATGGGATAGAGGCCCGAATACTGGGGTATTC
1068	CGTWDRGLNTVYVF	1892	TGCGGAACATGGGATAGAGGCCTGAATACTGTTTACGTCTTC
1069	CGTWDNYVSAPWVF	1893	TGCGGAACATGGGATAACTATGTGAGTGCCCCTTGGGTGTTC
1070	CGTWDNYLSAGDVVF	1894	TGCGGAACATGGGATAACTACCTGAGTGCTGGCGATGTGGTTTTC
1071	CGTWDNYLRAGVF	1895	TGCGGAACATGGGATAACTACCTGAGAGCTGGGGTCTTC
1072	CGTWDNYLGAVVF	1896	TGCGGAACATGGGACAATTATCTGGGTGCCGTGGTTTTC
1073	CGTWDNYLGAGVF	1897	TGCGGAACATGGGATAACTACCTGGGTGCGGGGGTGTTC
1074	CGTWDNTVSAPWVF	1898	TGCGGAACATGGGATAACACCGTGAGTGCCCCTTGGGTTTTC
1075	CGTWDNTLSLWVF	1899	TGCGGAACATGGGATAACACCCTGAGTCTTTGGGTGTTC
1076	CGTWDNTLSAGVF	1900	TGCGGAACATGGGATAACACCCTGAGTGCTGGGGTCTTC
1077	CGTWDNTLLTVLF	1901	TGCGGAACATGGGACAACACTCTGCTTACTGTGTTATTC
1078	CGTWDNRLSSVIF	1902	TGCGGAACATGGGATAACAGACTGAGTAGTGTGATTTTC
1079	CGTWDNRLSAVVF	1903	TGCGGAACATGGGATAACAGGTTGAGTGCTGTGGTCTTC
1080	CGTWDNRLSAGGIF	1904	TGCGGAACATGGGATAACAGGCTGAGTGCTGGTGGGATATTC
1081	CGTWDNRLSAEVF	1905	TGCGGAACATGGGATAACAGACTGAGTGCTGAGGTGTTC
1082	CGTWDNRLRVGVL	1906	TGTGGAACATGGGATAACAGACTGCGTGTTGGGGTTCTC
1083	CGTWDNRLLENVF	1907	TGCGGAACATGGGATAATCGCCTGCTTGAGAATGTCTTC
1084	CGTWDNNLRAVF	1908	TGCGGAACATGGGATAACAACCTGCGTGCTGTCTTC
1085	CGTWDNNLRAGVF	1909	TGCGGAACTTGGGATAATAACCTGCGTGCTGGAGTGTTC

1086	CGTWDNNLGGGRVF	1910	TGCGGAACATGGGACAACAATTTGGGCGGTGGCCGGGTGTTC
1087	CGTWDNNLGAGVL	1911	TGCGGAACATGGGATAACAACCTGGGTGCTGGCGTCCTC
1088	CGTWDNNLGAGVF	1912	TGCGGAACATGGGATAACAACCTGGGTGCTGGCGTCTTC
1089	CGTWDNILSAAVF	1913	TGCGGAACTTGGGATAACATCCTGAGCGCTGCGGTGTTC
1090	CGTWDNILDAGVF	1914	TGCGGAACCTGGGATAACATCTTGGATGCAGGGGTTTTC
1091	CGTWDNDLSGWLF	1915	TGCGGAACATGGGATAACGACCTGAGTGGTTGGCTGTTC
1092	CGTWDNDLSAWVF	1916	TGCGGAACATGGGATAACGACCTGAGTGCCTGGGTGTTC
1093	CGTWDLTLGGVVF	1917	TGCGGAACATGGGATCTCACCCTGGGTGGTGTGGTGTTC
1094	CGTWDLSLSAGVF	1918	TGCGGAACATGGGATCTCAGCCTGAGTGCTGGGGTATTC
1095	CGTWDLSLKEWVF	1919	TGCGGAACATGGGATCTCAGCCTGAAAGAATGGGTGTTC
1096	CGTWDLSLDAVVF	1920	TGCGGAACGTGGGATCTCAGCCTGGATGCTGTTGTTTTC
1097	CGTWDLKVF	1921	TGCGGAACCTGGGACCTGAAGGTTTTC
1098	CGTWDKTLSVWVF	1922	TGCGGAACATGGGATAAGACTCTGAGTGTTTGGGTGTTC
1099	CGTWDKSLSVWVF	1923	TGCGGAACATGGGATAAGAGCCTGAGTGTTTGGGTGTTC
1100	CGTWDKSLSGVVF	1924	TGCGGAACATGGGATAAGAGCCTGAGTGGTGTGGTATTT
1101	CGTWDKSLSDWVF	1925	TGCGGAACATGGGATAAGAGCCTGAGTGATTGGGTGTTC
1102	CGTWDKSLSALVF	1926	TGCGGAACATGGGATAAGAGCCTGAGTGCTTTGGTTTTC
1103	CGTWDKSLSAGVF	1927	TGCGGAACATGGGATAAGAGCCTGAGTGCTGGCGTCTTC
1104	CGTWDKSLSADVF	1928	TGCGGAACATGGGATAAGAGCCTGAGTGCCGACGTCTTC
1105	CGTWDKRLTIVVF	1929	TGCGGAACATGGGATAAACGCCTGACTATTGTGGTCTTC
1106	CGTWDKRLSAWVL	1930	TGCGGAACATGGGATAAACGCCTGAGTGCCTGGGTGCTC
1107	CGTWDKNLRAVVF	1931	TGCGGAACATGGGATAAGAACCTGCGTGCTGTGGTCTTC
1108	CGTWDITLSGFVF	1932	TGCGGAACATGGGATATCACCCTGAGTGGGTTTGTCTTC
1109	CGTWDITLHTGVF	1933	TGCGGAACATGGGATATCACCTTGCATACTGGAGTATTC
1110	CGTWDISVTVVF	1934	TGCGGAACATGGGATATCAGTGTGACTGTGGTGTTC
1111	CGTWDISVRGYAF	1935	TGCGGAACATGGGATATCAGTGTGAGGGGTTATGCCTTC
1112	CGTWDISRWVF	1936	TGCGGAACATGGGATATCAGCCGTTGGGTTTTC
1113	CGTWDISPSAWVF	1937	TGCGGAACATGGGATATCAGCCCGAGTGCTTGGGTGTTC
1114	CGTWDISLSVWVF	1938	TGCGGAACATGGGATATTAGCCTGAGTGTCTGGGTGTTC
1115	CGTWDISLSVVF	1939	TGCGGAACATGGGATATCAGCCTGAGTGTTGTATTC
1116	CGTWDISLSSVVF	1940	TGCGGAACTTGGGATATCAGCCTGAGTTCTGTGGTGTTC
1117	CGTWDISLSHWLF	1941	TGCGGAACATGGGATATCAGCCTGAGTCACTGGTTGTTC
1118	CGTWDISLSGWVF	1942	TGCGGAACATGGGATATCAGTCTGAGTGGTTGGGTGTTC
1119	CGTWDISLSGRVF	1943	TGCGGAACATGGGATATCAGCCTGAGTGGTCGAGTGTTC
1120	CGTWDISLSAWAF	1944	TGCGGAACATGGGACATCAGCCTGAGTGCTTGGGCGTTC
1121	CGTWDISLSAVVF	1945	TGCGGAACATGGGATATCAGCCTGAGTGCTGTGGTTTTC
1122	CGTWDISLSAVIF	1946	TGCGGGACATGGGACATCAGCCTGAGTGCTGTGATATTC
1123	CGTWDISLSAVF	1947	TGCGGAACATGGGATATCAGCCTGAGTGCTGTGTTC
1124	CGTWDISLSARVF	1948	TGCGGAACATGGGATATCAGCCTGAGTGCCCGGGTGTTC
1125	CGTWDISLSALVF	1949	TGCGGAACATGGGATATCAGCCTGAGTGCCCTGGTGTTC
1126	CGTWDISLSAHVF	1950	TGCGGAACATGGGATATTAGCCTGAGTGCCCATGTCTTC
1127	CGTWDISLSAGVVF	1951	TGCGGAACATGGGATATCAGCCTGAGTGCTGGGGTGGTATTC
1128	CGTWDISLSAGPYVF	1952	TGCGGAACATGGGATATCAGCCTGAGTGCCGGCCCTTATGTCTTC
1129	CGTWDISLSAGGVF	1953	TGCGGCACATGGGATATCAGCCTGAGTGCTGGAGGGGTGTTC

1130	CGTWDISLSAEVF	1954	TGCGGAACATGGGATATCAGCCTGAGTGCTGAGGTTTTC
1131	CGTWDISLSAAVF	1955	TGCGGAACATGGGATATCAGCCTGAGTGCTGCTGTGTTC
1132	CGTWDISLRAVF	1956	TGCGGAACATGGGATATCAGCCTGCGTGCTGTGTTC
1133	CGTWDISLNTGVF	1957	TGCGGAACATGGGATATTAGCCTGAATACTGGGGTGTTC
1134	CGTWDISLNNYVF	1958	TGCGGAACATGGGATATCAGCCTAAATAATTATGTCTTC
1135	CGTWDISLIAGVF	1959	TGCGGAACATGGGATATCAGCCTAATTGCTGGGGTATTC
1136	CGTWDISLHTWLF	1960	TGCGGAACATGGGATATCAGCCTGCATACTTGGCTGTTC
1137	CGTWDIRLTDELLF	1961	TGCGGAACATGGGATATCCGCCTGACCGATGAGCTGTTATTC
1138	CGTWDIRLSGFVF	1962	TGCGGAACATGGGATATCAGACTGAGCGGTTTTGTTTTC
1139	CGTWDINLGAGGLYVF	1963	TGCGGAACATGGGATATCAACCTGGGTGCTGGGGGCCTTTATGTCTTC
1140	CGTWDIILSAEVF	1964	TGCGGAACATGGGATATCATCCTGAGTGCTGAGGTATTC
1141	CGTWDHTLSAVF	1965	TGCGGAACATGGGATCACACCCTGAGTGCTGTCTTC
1142	CGTWDHTLLTVLF	1966	TGCGGAACATGGGACCACACTCTGCTTACTGTGTTATTC
1143	CGTWDHSLTAVVF	1967	TGCGGAACATGGGATCACAGCCTGACTGCTGTGGTATTC
1144	CGTWDHSLTAGIF	1968	TGCGGAACCTGGGATCACAGCCTGACTGCTGGGATATTC
1145	CGTWDHSLSVVLF	1969	TGCGGAACATGGGATCACAGCCTGAGTGTTGTATTATTC
1146	CGTWDHSLSLVF	1970	TGCGGAACATGGGATCACAGCCTGAGTTTGGTATTC
1147	CGTWDHSLSIGVF	1971	TGCGGAACATGGGATCACAGCCTGTCTATTGGGGTTTTC
1148	CGTWDHSLSAGVF	1972	TGCGGAACATGGGATCACAGCCTGAGTGCTGGGGTGTTC
1149	CGTWDHSLSAFVF	1973	TGTGGAACTTGGGATCACAGCCTGAGTGCTTTCGTGTTC
1150	CGTWDHSLSAAVF	1974	TGCGGAACATGGGATCACAGTCTGAGTGCTGCTGTTTTC
1151	CGTWDHNLRAVF	1975	TGCGGAACATGGGACCACAATCTGCGTGCTGTCTTC
1152	CGTWDFTLSVGRF	1976	TGCGGGACATGGGATTTCACCCTGAGTGTTGGGCGCTTC
1153	CGTWDFTLSAPVF	1977	TGCGGAACATGGGATTTCACCCTGAGTGCTCCTGTCTTC
1154	CGTWDFSVSAGWVF	1978	TGCGGAACGTGGGATTTCAGCGTGAGTGCTGGGTGGGTGTTC
1155	CGTWDFSLTTWLF	1979	TGCGGAACGTGGGATTTCAGTCTTACTACCTGGTTATTC
1156	CGTWDFSLSVWVF	1980	TGCGGAACATGGGATTTCAGCCTGAGTGTTTGGGTGTTC
1157	CGTWDFSLSTGVF	1981	TGCGGAACATGGGATTTCAGCCTGAGTACTGGGGTTTTC
1158	CGTWDFSLSGVVF	1982	TGCGGCACATGGGATTTCAGCCTGAGTGGTGTGGTATTC
1159	CGTWDFSLSGFVF	1983	TGCGGAACATGGGATTTCAGCCTGAGTGGTTTCGTGTTC
1160	CGTWDFSLSAGVF	1984	TGCGGAACATGGGATTTCAGCCTGAGTGCTGGGGTGTTC
1161	CGTWDETVRGWVF	1985	TGCGGAACATGGGATGAAACCGTGAGAGGTTGGGTGTTC
1162	CGTWDESLRSWVF	1986	TGCGGAACATGGGATGAAAGTCTGAGAAGCTGGGTGTTC
1163	CGTWDERQTDESYVF	1987	TGCGGAACTTGGGATGAGAGGCAGACTGATGAGTCCTATGTCTTC
1164	CGTWDERLVAGQVF	1988	TGCGGAACATGGGATGAGAGACTCGTTGCTGGCCAGGTCTTC
1165	CGTWDERLSPGAFF	1989	TGCGGAACATGGGATGAGAGACTGAGTCCTGGAGCTTTTTTC
1166	CGTWDEKVF	1990	TGCGGAACATGGGATGAGAAGGTGTTC
1167	CGTWDEGQTTDFFVF	1991	TGCGGAACCTGGGATGAAGGCCAGACTACTGATTTCTTTGTCTTC
1168	CGTWDDTLAGVVF	1992	TGCGGAACATGGGATGACACCCTGGCTGGTGTGGTCTTC
1169	CGTWDDRLTSAVF	1993	TGCGGAACATGGGATGACAGGCTGACTTCTGCGGTCTTC
1170	CGTWDDRLFVVVF	1994	TGCGGAACATGGGATGACAGACTGTTTGTTGTGGTATTC
1171	CGTWDDNLRGWVF	1995	TGCGGAACATGGGATGATAACCTGAGAGGTTGGGTGTTC
1172	CGTWDDNLRGVVF	1996	TGCGGAACATGGGATGACAACCTGCGTGGTGTCGTGTTC
1173	CGTWDDNLNIGRVF	1997	TGCGGAACCTGGGATGACAATTTGAATATTGGAAGGGTGTTC

1174	CGTWDDILSAVIF	1998	TGCGGAACATGGGATGACATCCTGAGTGCTGTGATATTC
1175	CGTWDDILRGWVF	1999	TGCGGAACATGGGATGATATCCTGAGAGGTTGGGTGTTC
1176	CGTWDATLSPGWLF	2000	TGCGGAACATGGGATGCCACCCTGAGTCCTGGGTGGTTATTC
1177	CGTWDASVTSWVF	2001	TGCGGAACATGGGATGCCAGCGTGACTTCTTGGGTGTTC
1178	CGTWDASLTSVVF	2002	TGCGGAACATGGGATGCCAGCCTGACTTCTGTGGTCTTC
1179	CGTWDASLSVWVF	2003	TGCGGAACATGGGATGCCAGCCTGAGTGTTTGGGTGTTC
1180	CGTWDASLSVPWVF	2004	TGCGGAACATGGGATGCCAGCCTGAGTGTTCCTTGGGTGTTC
1181	CGTWDASLSVAVF	2005	TGCGGAACATGGGATGCCAGCCTGAGTGTGGCGGTATTC
1182	CGTWDASLSTWVF	2006	TGCGGAACATGGGATGCCAGCCTGAGTACCTGGGTATTC
1183	CGTWDASLSGVVF	2007	TGCGGAACATGGGATGCCAGCCTGAGTGGTGTGGTATTC
1184	CGTWDASLSGGGEF	2008	TGCGGAACATGGGATGCCAGCCTGAGTGGTGGGGGAGAATTC
1185	CGTWDASLSAGVF	2009	TGCGGAACATGGGATGCCAGCCTGAGTGCTGGGGTGTTC
1186	CGTWDASLSAGLF	2010	TGCGGAACATGGGATGCCAGCCTGAGTGCTGGGCTTTTC
1187	CGTWDASLSAEVF	2011	TGTGGCACATGGGATGCCAGCCTGAGTGCTGAAGTCTTC
1188	CGTWDASLSADFWVF	2012	TGCGGAACATGGGATGCCAGCCTGAGTGCTGACTTTTGGGTGTTC
1189	CGTWDASLRVFF	2013	TGCGGAACATGGGATGCCAGCCTGAGAGTCTTCTTC
1190	CGTWDASLRAVVL	2014	TGCGGAACATGGGATGCCAGTCTGAGGGCTGTGGTACTC
1191	CGTWDASLNIWVF	2015	TGCGGAACATGGGATGCCAGCCTGAATATTTGGGTTTTC
1192	CGTWDASLKNLVF	2016	TGCGGGACATGGGATGCCAGCCTGAAGAATCTGGTCTTC
1193	CGTWDASLGAWVF	2017	TGCGGAACATGGGATGCCAGCCTGGGTGCCTGGGTATTC
1194	CGTWDASLGAVVF	2018	TGCGGAACATGGGATGCCAGCCTGGGTGCTGTGGTCTTC
1195	CGTWDASLGAGVF	2019	TGCGGAACATGGGATGCCAGCCTGGGTGCGGGGGTCTTC
1196	CGTWDARLSGLYVF	2020	TGCGGAACATGGGATGCTAGGCTGAGTGGCCTTTATGTCTTC
1197	CGTWDARLGGAVF	2021	TGTGGAACCTGGGATGCGAGACTGGGTGGTGCAGTCTTC
1198	CGTWDANLRAGVF	2022	TGCGGAACATGGGATGCCAATCTGCGTGCTGGGGTCTTC
1199	CGTWDAIISGWVF	2023	TGCGGAACATGGGATGCTATCATAAGTGGTTGGGTGTTC
1200	CGTWDAGQSVWVF	2024	TGCGGAACATGGGATGCCGGCCAGAGTGTTTGGGTGTTC
1201	CGTWDAGLTGLYVF	2025	TGCGGCACATGGGATGCCGGGCTGACTGGCCTTTATGTCTTC
1202	CGTWDAGLSVYVF	2026	TGCGGAACTTGGGATGCCGGTCTGAGTGTTTATGTCTTC
1203	CGTWDAGLSTGVF	2027	TGCGGGACATGGGATGCCGGCCTGAGTACTGGGGTCTTC
1204	CGTWDAGLSGDVF	2028	TGCGGAACATGGGATGCCGGCCTGAGTGGGGACGTTTTC
1205	CGTWDAGLSAGYVF	2029	TGCGGAACATGGGATGCCGGCCTGAGTGCTGGTTATGTCTTC
1206	CGTWDAGLRVWVF	2030	TGCGGAACATGGGATGCCGGCCTGCGTGTTTGGGTGTTC
1207	CGTWDAGLREIF	2031	TGCGGAACATGGGATGCCGGCCTGAGGGAAATTTTC
1208	CGTWASSLSSWVF	2032	TGCGGAACATGGGCCAGCAGCCTGAGTTCTTGGGTGTTC
1209	CGTWAGSLSGHVF	2033	TGCGGAACATGGGCTGGCAGCCTGAGTGGTCATGTCTTC
1210	CGTWAGSLSAAWVF	2034	TGCGGAACATGGGCTGGCAGCCTGAGTGCCGCTTGGGTGTTC
1211	CGTWAGSLNVYWVF	2035	TGCGGAACATGGGCTGGCAGCCTGAATGTTTATTGGGTGTTC
1212	CGTWAGNLRPNWVF	2036	TGCGGAACATGGGCTGGCAACCTGAGACCTAATTGGGTGTTC
1213	CGTRGSLGGAVF	2037	TGCGGAACAAGGGGTAGCCTGGGTGGTGCGGTGTTC
1214	CGTRDTTLSVPVF	2038	TGCGGAACAAGGGATACCACCCTGAGTGTCCCGGTGTTC
1215	CGTRDTSLNIEIF	2039	TGCGGAACACGGGATACCAGCCTCAATATTGAAATCTTC
1216	CGTRDTSLNDVF	2040	TGTGGAACACGGGATACCAGCCTGAATGATGTCTTC

1217	CGTRDTRLSIVVF	2041	TGCGGAACACGGGATACCCGCCTGAGTATTGTGGTTTTC
1218	CGTRDTILSAEVF	2042	TGCGGCACACGGGATACCATCCTGAGTGCTGAGGTGTTC
1219	CGTRDRSLSGWVF	2043	TGCGGAACACGGGATAGAAGCCTGAGTGGTTGGGTGTTC
1220	CGSWYYNVFLF	2044	TGCGGATCATGGTATTACAATGTCTTCCTTTTC
1221	CGSWHSSLNLVVF	2045	TGCGGATCTTGGCATAGCAGCCTCAACCTTGTCGTCTTC
1222	CGSWGSGLSAPYVF	2046	TGCGGATCATGGGGTAGTGGCCTGAGTGCCCCTTATGTCTTC
1223	CGSWESGLGAWLF	2047	TGCGGTTCGTGGGAAAGCGGCCTGGGTGCTTGGCTGTTC
1224	CGSWDYGLLLF	2048	TGCGGATCCTGGGATTACGGCCTCCTACTCTTC
1225	CGSWDVSLTAVF	2049	TGCGGTTCATGGGATGTCAGCCTGACTGCTGTTTTC
1226	CGSWDVSLNVGIF	2050	TGCGGATCCTGGGATGTCAGTCTCAATGTTGGCATTTTC
1227	CGSWDTTLRAWVF	2051	TGCGGATCATGGGATACCACCCTGCGTGCTTGGGTGTTC
1228	CGSWDTSPVRAWVF	2052	TGCGGCTCGTGGGATACCAGCCCTGTCCGTGCTTGGGTGTTC
1229	CGSWDTSLSVWVF	2053	TGCGGATCATGGGATACCAGCCTGAGTGTTTGGGTGTTC
1230	CGSWDTSLSAEVF	2054	TGCGGATCATGGGATACCAGCCTGAGTGCTGAGGTGTTC
1231	CGSWDTSLRAWVF	2055	TGCGGCTCGTGGGATACCAGCCTGCGTGCTTGGGTGTTC
1232	CGSWDTSLRAWAF	2056	TGCGGCTCGTGGGATACCAGCCTGCGTGCTTGGGCGTTC
1233	CGSWDTSLDARLF	2057	TGCGGATCATGGGATACCAGCCTGGATGCTAGGCTGTTC
1234	CGSWDTILLVYVF	2058	TGCGGATCATGGGATACCATCCTGCTTGTCTATGTCTTC
1235	CGSWDRWQAAVF	2059	TGCGGATCATGGGATCGCTGGCAGGCTGCTGTCTTC
1236	CGSWDRSLSGYVF	2060	TGCGGATCATGGGATAGGAGCCTGAGTGGGTATGTCTTC
1237	CGSWDRSLSAYVF	2061	TGCGGATCATGGGATAGAAGCCTGAGTGCTTATGTCTTC
1238	CGSWDRSLSAVVF	2062	TGCGGATCATGGGATAGGAGCCTGAGTGCCGTGGTTTTC
1239	CGSWDNTLGVVLF	2063	TGCGGATCATGGGATAACACCTTGGGTGTTGTTCTCTTC
1240	CGSWDNRLSTVIF	2064	TGCGGATCGTGGGATAACAGACTAAGTACTGTCATCTTC
1241	CGSWDNRLNTVIF	2065	TGCGGAAGCTGGGATAATCGATTGAACACTGTGATTTTC
1242	CGSWDLSPVRVLVF	2066	TGCGGTTCATGGGATCTCAGCCCTGTACGTGTCCTTGTGTTC
1243	CGSWDLSLSAVVF	2067	TGCGGATCATGGGATCTCAGCCTGAGTGCTGTCGTTTTC
1244	CGSWDKNLRAVLF	2068	TGCGGATCATGGGATAAAAACCTGCGTGCTGTGCTGTTC
1245	CGSWDISLSAGVF	2069	TGCGGCTCATGGGATATCAGCCTGAGTGCTGGGGTGTTC
1246	CGSWDIRLSAEVF	2070	TGCGGATCATGGGATATCAGACTGAGTGCAGAGGTCTTC
1247	CGSWDIKLNIGVF	2071	TGCGGATCATGGGACATCAAACTGAATATTGGGGTATTC
1248	CGSWDFSLNYFVF	2072	TGCGGATCATGGGATTTCAGTCTCAATTATTTTGTCTTC
1249	CGSWDASLSTEVF	2073	TGCGGATCATGGGATGCCAGCCTGAGTACTGAGGTGTTC
1250	CGSWDAGLRGWVF	2074	TGCGGATCCTGGGATGCCGGCCTGCGTGGCTGGGTTTTC
1251	CGRWESSLGAVVF	2075	TGCGGAAGATGGGAGAGCAGCCTGGGTGCTGTGGTTTTC
1252	CGRWDFSLSAYVF	2076	TGCGGAAGATGGGATTTTAGTCTGAGTGCTTATGTCTTC
1253	CGQWDNDLSVWVF	2077	TGCGGACAATGGGATAACGACCTGAGTGTTTGGGTGTTC
1254	CGPWHSSVTSGHVL	2078	TGCGGACCCTGGCATAGCAGCGTGACTAGTGGCCACGTGCTC
1255	CGLWDASLSAPTWVF	2079	TGCGGATTATGGGATGCCAGCCTGAGTGCTCCTACTTGGGTGTTC
1256	CGIWHTSLSAWVF	2080	TGTGGAATATGGCACACTAGCCTGAGTGCTTGGGTGTTC
1257	CGIWDYSLDTWVF	2081	TGCGGAATATGGGATTACAGCCTGGATACTTGGGTGTTC
1258	CGIWDTSLSAWVF	2082	TGCGGCATATGGGATACCAGCCTGAGTGCTTGGGTGTTC
1259	CGIWDTRLSVYVF	2083	TGCGGAATTTGGGATACCAGGCTGAGTGTTTATGTCTTC
1260	CGIWDTRLSVYIF	2084	TGCGGAATTTGGGATACCAGGCTGAGTGTTTATATCTTC
1261	CGIWDTNLGYLF	2085	TGTGGAATATGGGATACGAATCTGGGTTATCTCTTC

1262	CGIWDTGLSAVVF	2086	TGCGGTATATGGGATACCGGCCTGAGTGCTGTGGTATTC
1263	CGIWDRSLSAWVF	2087	TGCGGAATATGGGATCGCAGCCTGAGTGCTTGGGTGTTT
1264	CGIRDTRLSVYVF	2088	TGCGGAATTCGGGATACCAGGCTGAGTGTTTATGTCTTC
1265	CGGWSSRLGVGPVF	2089	TGCGGAGGATGGAGTAGCAGACTGGGTGTTGGCCCAGTGTTT
1266	CGGWGSGLSAWVF	2090	TGCGGAGGATGGGGTAGCGGCCTGAGTGCTTGGGTGTTC
1267	CGGWDTSLSAWVF	2091	TGCGGAGGATGGGATACCAGCCTGAGTGCTTGGGTGTTC
1268	CGGWDRGLDAWVF	2092	TGCGGAGGATGGGATAGGGGCCTGGATGCTTGGGTTTTC
1269	CGAWRNNVWVF	2093	TGCGGAGCATGGCGTAATAACGTGTGGGTGTTC
1270	CGAWNRRLNPHSHWVF	2094	TGCGGAGCATGGAACAGGCGCCTGAATCCTCATTCTCATTGGGTGTTC
1271	CGAWHNKLSAVF	2095	TGCGGAGCCTGGCACAACAAACTGAGCGCGGTCTTC
1272	CGAWGSSLRASVF	2096	TGCGGAGCATGGGGTAGCAGCCTGAGAGCTAGTGTCTTC
1273	CGAWGSGLSAWVF	2097	TGCGGAGCATGGGGTAGCGGCCTGAGTGCTTGGGTGTTC
1274	CGAWESSLSAPYVF	2098	TGCGGAGCATGGGAAAGTAGCCTGAGTGCCCCTTATGTCTTC
1275	CGAWESSLNVGLI	2099	TGCGGAGCATGGGAGAGCAGCCTCAATGTTGGACTGATC
1276	CGAWESGRSAGVVF	2100	TGCGGAGCATGGGAGAGCGGCCGGAGTGCTGGGGTGGTGTTC
1277	CGAWDYSVSGWVF	2101	TGCGGAGCTTGGGATTACAGTGTGAGTGGTTGGGTGTTC
1278	CGAWDYSLTAGVF	2102	TGCGGAGCATGGGATTACAGCCTGACTGCCGGAGTATTC
1279	CGAWDYRLSAVLF	2103	TGCGGAGCCTGGGATTACAGACTGAGTGCCGTGCTATTC
1280	CGAWDVRLDVGVF	2104	TGCGGAGCGTGGGATGTTCGTCTGGATGTTGGGGTGTTC
1281	CGAWDTYSYVF	2105	TGCGGAGCATGGGATACCTACAGTTATGTCTTC
1282	CGAWDTTLSGVVF	2106	TGCGGAGCATGGGATACGACCCTGAGTGGTGTGGTATTC
1283	CGAWDTTLSAVIF	2107	TGCGGAGCGTGGGATACTACCCTGAGTGCTGTGATATTC
1284	CGAWDTSQGASYVF	2108	TGCGGCGCATGGGATACCAGCCAGGGTGCGTCTTATGTCTTT
1285	CGAWDTSPVRAGVF	2109	TGCGGAGCATGGGATACCAGCCCTGTACGTGCTGGGGTGTTC
1286	CGAWDTSLWLF	2110	TGCGGAGCATGGGATACCAGCCTGTGGCTTTTC
1287	CGAWDTSLTVYVF	2111	TGCGGAGCATGGGATACCAGCCTGACTGTTTATGTCTTC
1288	CGAWDTSLTAGVF	2112	TGCGGAGCATGGGACACCAGTCTGACTGCTGGGGTGTTC
1289	CGAWDTSLSTVVF	2113	TGCGGAGCTTGGGATACCAGCCTGAGTACTGTGGTTTTC
1290	CGAWDTSLSSRYIF	2114	TGCGGAGCATGGGATACCAGCCTGAGTTCTAGATACATATTC
1291	CGAWDTSLSGYVF	2115	TGCGGAGCATGGGATACCAGCCTGAGTGGTTATGTCTTC
1292	CGAWDTSLSGWVF	2116	TGCGGAGCCTGGGATACCAGCCTGAGTGGCTGGGTGTTC
1293	CGAWDTSLSGVLF	2117	TGCGGAGCATGGGATACCAGTCTGAGTGGTGTGCTATTC
1294	CGAWDTSLSGLVF	2118	TGCGGAGCTTGGGATACCAGCTTGAGTGGTCTTGTTTTC
1295	CGAWDTSLSGFVF	2119	TGCGGAGCTTGGGATACCAGCTTGAGTGGTTTTGTTTTC
1296	CGAWDTSLSGEVF	2120	TGCGGAGCATGGGATACCAGCCTGAGTGGTGAGGTCTTT
1297	CGAWDTSLSDFVF	2121	TGCGGAGCTTGGGATACCAGCTTGAGTGATTTTGTTTTC
1298	CGAWDTSLRTAIF	2122	TGCGGAGCATGGGATACCAGCCTGCGAACTGCGATATTC
1299	CGAWDTSLRLF	2123	TGCGGAGCATGGGATACCAGCCTGCGGCTTTTC
1300	CGAWDTSLNVHVF	2124	TGCGGAGCATGGGATACCAGCCTGAATGTTCATGTCTTC
1301	CGAWDTSLNKWVF	2125	TGCGGAGCATGGGATACCAGCCTCAATAAATGGGTGTTC
1302	CGAWDTRLSARLF	2126	TGCGGAGCATGGGATACCCGCCTCAGTGCGCGGCTGTTC
1303	CGAWDTRLRGFIF	2127	TGCGGAGCATGGGATACCAGACTGAGGGGTTTTATTTTC
1304	CGAWDTNLGNVLL	2128	TGCGGAGCATGGGATACTAATTTGGGGAATGTTCTCCTC
1305	CGAWDTNLGKWVF	2129	TGCGGGGCATGGGATACCAACCTGGGTAAATGGGTTTTC

1306	CGAWDTGLEWYVF	2130	TGCGGAGCATGGGATACCGGCCTTGAGTGGTATGTTTTT
1307	CGAWDRTSGLWLF	2131	TGCGGAGCATGGGATAGGACTTCTGGATTGTGGCTTTTC
1308	CGAWDRSLVAGLF	2132	TGCGGAGCGTGGGATCGTAGCCTGGTTGCTGGACTCTTC
1309	CGAWDRSLTVYVF	2133	TGCGGAGCGTGGGATAGAAGCCTGACTGTTTATGTCTTC
1310	CGAWDRSLSGYVF	2134	TGCGGAGCATGGGATAGAAGCCTGAGTGGTTATGTCTTC
1311	CGAWDRSLSAYVF	2135	TGCGGAGCATGGGATAGAAGCCTGAGTGCTTATGTCTTC
1312	CGAWDRSLSAVVF	2136	TGCGGAGCATGGGATAGAAGCCTGAGTGCGGTGGTATTC
1313	CGAWDRSLSAGVF	2137	TGCGGAGCATGGGATCGCAGCCTGAGTGCTGGGGTTTTC
1314	CGAWDRSLRIVVF	2138	TGCGGAGCGTGGGATCGCAGCCTGCGTATTGTGGTATTC
1315	CGAWDRSLRAYVF	2139	TGCGGAGCATGGGATAGAAGTCTGAGGGCTTACGTCTTC
1316	CGAWDRSLNVWLF	2140	TGCGGAGCATGGGATAGAAGTCTGAATGTTTGGCTGTTC
1317	CGAWDRGLNVGWLF	2141	TGCGGCGCCTGGGATAGGGGCCTGAATGTCGGTTGGCTTTTC
1318	CGAWDNRLSILAF	2142	TGCGGCGCATGGGATAATAGACTGAGTATTTTGGCCTTC
1319	CGAWDNDLTAYVF	2143	TGCGGAGCTTGGGATAATGACCTGACAGCTTATGTCTTC
1320	CGAWDFSLTPLF	2144	TGCGGGGCATGGGATTTCAGCCTGACTCCTCTCTTC
1321	CGAWDDYRGVSIYVF	2145	TGCGGAGCCTGGGATGACTATCGGGGTGTGAGTATTTATGTCTTC
1322	CGAWDDRPSSAVVF	2146	TGTGGAGCATGGGATGACCGGCCTTCGAGTGCCGTGGTTTTC
1323	CGAWDDRLTVVVF	2147	TGCGGAGCATGGGATGACAGACTGACTGTCGTTGTTTTC
1324	CGAWDDRLGAVF	2148	TGCGGAGCGTGGGATGACAGGCTGGGTGCTGTGTTC
1325	CGAWDASLNPGRAF	2149	TGCGGAGCGTGGGATGCCAGCCTGAATCCTGGCCGGGCATTC
1326	CGAWDAGLREIF	2150	TGCGGAGCATGGGATGCCGGCCTGAGGGAAATTTTC
1327	CGAWAGSPSPWVF	2151	TGCGGAGCTTGGGCTGGCAGTCCGAGTCCTTGGGTTTTC
1328	CGAFDTTLSAGVF	2152	TGCGGAGCATTCGACACCACCCTGAGTGCTGGCGTTTTC
1329	CETWESSLSVGVF	2153	TGCGAAACATGGGAGAGCAGCCTGAGTGTTGGGGTCTTC
1330	CETWESSLRVWVF	2154	TGCGAAACATGGGAAAGCAGCCTGAGGGTTTGGGTGTTC
1331	CETWDTSLSGGVF	2155	TGCGAAACGTGGGATACCAGCCTGAGTGGTGGGGTGTTC
1332	CETWDTSLSDFYVF	2156	TGCGAAACATGGGATACCAGCCTGAGTGACTTTTATGTCTTC
1333	CETWDTSLSALF	2157	TGCGAAACATGGGATACCAGCCTGAGTGCCCTCTTC
1334	CETWDTSLRAEVF	2158	TGCGAAACATGGGATACCAGCCTGCGTGCTGAAGTCTTC
1335	CETWDTSLNVVVF	2159	TGCGAAACATGGGATACCAGCCTGAATGTTGTGGTATTC
1336	CETWDTSLGAVVF	2160	TGCGAAACATGGGATACCAGCCTGGGTGCCGTGGTGTTC
1337	CETWDRSLSGVVF	2161	TGCGAAACATGGGATAGAAGCCTGAGTGGTGTGGTATTC
1338	CETWDRSLSAWVF	2162	TGCGAAACATGGGATAGGAGCCTGAGTGCTTGGGTGTTT
1339	CETWDRSLSAVVF	2163	TGCGAAACATGGGATCGCAGCCTGAGTGCTGTGGTCTTC
1340	CETWDRGLSVVVF	2164	TGCGAGACGTGGGATAGAGGCCTGAGTGTTGTGGTTTTC
1341	CETWDRGLSAVVF	2165	TGCGAAACATGGGATAGGGGCCTGAGTGCAGTGGTATTC
1342	CETWDHTLSVVIF	2166	TGCGAAACATGGGATCACACCCTGAGTGTTGTGATATTC
1343	CETWDASLTVVLF	2167	TGCGAAACATGGGATGCCAGCCTGACTGTTGTGTTATTC
1344	CETWDASLSAGVF	2168	TGCGAAACATGGGATGCCAGCCTGAGTGCTGGGGTGTTC
1345	CETWDAGLSEVVF	2169	TGCGAAACGTGGGATGCCGGCCTGAGTGAGGTGGTGTTC
1346	CETFDTSLSVVVF	2170	TGCGAAACATTTGATACCAGCCTGAGTGTTGTAGTCTTC
1347	CETFDTSLNIVVF	2171	TGCGAAACATTTGATACCAGCCTAAATATTGTAGTCTTT
1348	CESWDRSRIGVVF	2172	TGCGAATCATGGGATAGAAGCCGGATTGGTGTGGTCTTC
1349	CESWDRSLSARVY	2173	TGCGAAAGTTGGGACAGGAGTCTGAGTGCCCGGGTGTAC

1350	CESWDRSLRAVVF	2174	TGCGAATCCTGGGATAGGAGCCTGCGTGCCGTGGTCTTC
1351	CESWDRSLIVVF	2175	TGCGAATCTTGGGATCGTAGTTTGATTGTGGTGTTC
1352	CESWDNNLNEVVF	2176	TGCGAAAGTTGGGATAACAATTTAAATGAGGTGGTTTTC
1353	CEIWESSPSADDLVF	2177	TGCGAAATATGGGAGAGCAGCCCGAGTGCTGACGATTTGGTGTTC
1354	CEAWDTSLSGAVF	2178	TGCGAAGCATGGGATACCAGCCTGAGTGGTGCGGTGTTC
1355	CEAWDTSLSAGVF	2179	TGCGAAGCATGGGATACCAGCCTGAGTGCCGGGGTGTTC
1356	CEAWDTSLGGGVF	2180	TGCGAAGCATGGGATACCAGCCTGGGTGGTGGGGTGTTC
1357	CEAWDRSLTGSLF	2181	TGCGAAGCATGGGATCGCAGCCTGACTGGTAGCCTGTTC
1358	CEAWDRGLSAVVF	2182	TGCGAAGCGTGGGATAGGGGCCTGAGTGCAGTGGTATTC
1359	CEAWDNILSTVVF	2183	TGCGAAGCCTGGGATAACATCCTGAGTACTGTGGTGTTC
1360	CEAWDISLSAGVF	2184	TGCGAAGCATGGGACATCAGCCTGAGTGCTGGGGTGTTC
1361	CEAWDADLSGAVF	2185	TGCGAAGCATGGGATGCCGACCTGAGTGGTGCGGTGTTC
1362	CATWTGSFRTGHYVF	2186	TGCGCAACATGGACTGGTAGTTTCAGAACTGGCCATTATGTCTTC
1363	CATWSSSPRGWVF	2187	TGCGCAACATGGAGTAGCAGTCCCAGGGGGTGGGTGTTC
1364	CATWHYSLSAGRVF	2188	TGCGCAACATGGCATTACAGCCTGAGTGCTGGCCGAGTGTTC
1365	CATWHTSLSIVQF	2189	TGCGCAACATGGCATACCAGCCTGAGTATTGTGCAGTTC
1366	CATWHSTLSADVLF	2190	TGCGCAACATGGCATAGCACCCTGAGTGCTGATGTGCTTTTC
1367	CATWHSSLSAGRLF	2191	TGCGCAACATGGCATAGCAGCCTGAGTGCTGGCCGACTCTTC
1368	CATWHIARSAWVF	2192	TGCGCAACATGGCATATCGCTCGGAGTGCCTGGGTGTTC
1369	CATWGSSQSAVVF	2193	TGCGCAACATGGGGTAGTAGTCAGAGTGCCGTGGTATTC
1370	CATWGSSLSAGGVF	2194	TGCGCAACATGGGGTAGCAGCCTGAGTGCTGGGGGTGTTTTC
1371	CATWEYSLSVVLF	2195	TGTGCAACATGGGAATACAGCCTGAGTGTTGTGCTGTTC
1372	CATWETTRRASFVF	2196	TGCGCAACATGGGAGACCACCCGACGTGCCTCTTTTGTCTTC
1373	CATWETSLNVYVF	2197	TGCGCAACATGGGAGACCAGCCTGAATGTTTATGTCTTC
1374	CATWETSLNVVVF	2198	TGCGCAACATGGGAAACTAGCCTGAATGTTGTGGTCTTC
1375	CATWETSLNLYVF	2199	TGCGCAACATGGGAGACCAGCCTGAATCTTTATGTCTTC
1376	CATWETGLSAGEVF	2200	TGCGCAACATGGGAGACTGGCCTAAGTGCTGGAGAGGTGTTC
1377	CATWESTLSVVVF	2201	TGCGCGACGTGGGAGAGTACCCTAAGTGTTGTGGTTTTC
1378	CATWESSLSIFVF	2202	TGCGCAACGTGGGAGAGCAGCCTGAGTATTTTTGTCTTC
1379	CATWESSLNTFYVF	2203	TGCGCAACATGGGAAAGCAGCCTCAACACTTTTTATGTCTTC
1380	CATWESRVDTRGLLF	2204	TGCGCAACATGGGAGAGTAGGGTGGATACTCGAGGGTTGTTATTC
1381	CATWESGLSGAGVF	2205	TGCGCAACATGGGAGAGCGGCCTGAGTGGTGCGGGGGTGTTC
1382	CATWEGSLNTFYVF	2206	TGCGCAACATGGGAAGGCAGCCTCAACACTTTTTATGTCTTC
1383	CATWDYSLSAVVF	2207	TGCGCAACTTGGGATTATAGCCTGAGTGCTGTGGTGTTC
1384	CATWDYRLSIVVF	2208	TGCGCAACATGGGATTACAGACTGAGTATTGTGGTATTC
1385	CATWDYNLGAAVF	2209	TGCGCAACATGGGATTATAACCTGGGAGCTGCGGTGTTC
1386	CATWDVTLGVLHF	2210	TGCGCCACATGGGATGTCACCCTGGGTGTCTTGCATTTC
1387	CATWDTTLSVWVF	2211	TGCGCAACATGGGATACAACACTGAGTGTCTGGGTCTTC
1388	CATWDTTLSVVLF	2212	TGCGCAACATGGGATACCACCCTGAGTGTAGTACTTTTC
1389	CATWDTTLSVEVF	2213	TGCGCAACATGGGATACCACCCTGAGTGTTGAGGTCTTC
1390	CATWDTSPSLSGFWVF	2214	TGCGCAACATGGGATACCAGCCCCAGCCTGAGTGGTTTTTGGGTGTTC
1391	CATWDTSLTGVVF	2215	TGCGCAACATGGGATACCAGCCTGACTGGTGTGGTATTC
1392	CATWDTSLTGAVF	2216	TGCGCAACATGGGATACCAGCCTGACTGGTGCGGTGTTC

1393	CATWDTSLTAWVF	2217	TGCGCAACATGGGATACCAGCCTGACTGCCTGGGTATTC
1394	CATWDTSLTAVVF	2218	TGCGCAACATGGGATACCAGCCTGACTGCTGTGGTTTTC
1395	CATWDTSLTAKVF	2219	TGCGCAACATGGGATACTAGCCTGACTGCTAAGGTGTTC
1396	CATWDTSLSVVVF	2220	TGCGCAACATGGGACACCAGCCTGAGTGTTGTGGTTTTC
1397	CATWDTSLSVGVF	2221	TGCGCTACTTGGGATACCAGCCTGAGTGTTGGGGTATTT
1398	CATWDTSLSSWVF	2222	TGCGCAACATGGGATACCAGCCTGAGTTCTTGGGTGTTC
1399	CATWDTSLSGGVL	2223	TGCGCAACATGGGATACCAGCCTGAGTGGTGGGGTACTC
1400	CATWDTSLSGGVF	2224	TGCGCAACATGGGATACCAGCCTGAGTGGTGGGGTGTTC
1401	CATWDTSLSGGRVF	2225	TGCGCAACATGGGATACCAGCCTGAGTGGTGGCCGAGTGTTC
1402	CATWDTSLSGDRVF	2226	TGCGCAACATGGGATACCAGCCTGAGTGGTGACCGAGTGTTC
1403	CATWDTSLSEGVF	2227	TGCGCAACGTGGGATACTAGCCTGAGTGAAGGGGTGTTC
1404	CATWDTSLSAVVL	2228	TGCGCAACCTGGGATACCAGCCTGAGTGCCGTGGTGCTC
1405	CATWDTSLSAVF	2229	TGCGCAACATGGGATACCAGCCTGAGTGCTGTCTTC
1406	CATWDTSLSARVF	2230	TGCGCGACATGGGATACCAGCCTGAGTGCTCGGGTGTTC
1407	CATWDTSLSALF	2231	TGCGCAACATGGGATACCAGCCTGAGTGCCTTATTC
1408	CATWDTSLSAHVF	2232	TGCGCAACATGGGATACCAGCCTGAGTGCTCATGTCTTC
1409	CATWDTSLSAGRVF	2233	TGCGCAACATGGGATACCAGCCTGAGTGCTGGCCGGGTGTTC
1410	CATWDTSLSAEVF	2234	TGCGCAACATGGGATACCAGCCTGAGTGCGGAGGTCTTC
1411	CATWDTSLSADAGGGVF	2235	TGCGCAACATGGGATACCAGCCTGAGTGCTGATGCTGGTGGGGGGGTCTTC
1412	CATWDTSLRVVVF	2236	TGCGCAACATGGGATACCAGCCTGCGTGTCGTGGTATTC
1413	CATWDTSLRGVF	2237	TGCGCAACATGGGATACCAGCCTGAGAGGGGTGTTC
1414	CATWDTSLPAWVF	2238	TGCGCAACATGGGATACCAGCCTGCCTGCGTGGGTGTTC
1415	CATWDTSLNVGVF	2239	TGTGCAACATGGGATACCAGCCTGAATGTTGGGGTATTC
1416	CATWDTSLGIVLF	2240	TGCGCAACATGGGATACCAGCCTGGGTATTGTGTTATTT
1417	CATWDTSLGARVVF	2241	TGCGCAACATGGGACACCAGCCTGGGTGCGCGTGTGGTCTTC
1418	CATWDTSLGALF	2242	TGTGCAACGTGGGATACCAGTCTAGGTGCCTTGTTC
1419	CATWDTSLATGLF	2243	TGCGCAACATGGGATACCAGCCTGGCGACTGGACTGTTC
1420	CATWDTSLAAWVF	2244	TGCGCAACATGGGATACCAGCCTGGCTGCCTGGGTATTC
1421	CATWDTRLSAVVF	2245	TGCGCAACCTGGGATACCAGGCTGAGTGCTGTGGTCTTC
1422	CATWDTRLSAGVF	2246	TGCGCAACATGGGATACCAGGCTGAGTGCTGGGGTGTTC
1423	CATWDTRLLITVF	2247	TGTGCAACGTGGGACACACGTCTACTTATTACGGTTTTC
1424	CATWDTLLSVELF	2248	TGCGCAACATGGGACACCCTCCTGAGTGTTGAACTCTTC
1425	CATWDTGRNPHVVF	2249	TGCGCAACATGGGATACTGGCCGCAATCCTCATGTGGTCTTC
1426	CATWDTGLSSVLF	2250	TGCGCAACATGGGATACCGGCCTGTCTTCGGTGTTGTTC
1427	CATWDTGLSAVF	2251	TGCGCAACGTGGGATACCGGCCTGAGTGCGGTTTTC
1428	CATWDRTLSIGVF	2252	TGCGCTACGTGGGATAGGACCCTGAGTATTGGAGTCTTC
1429	CATWDRSVTAVLF	2253	TGCGCAACGTGGGATCGCAGTGTGACTGCTGTGCTCTTC
1430	CATWDRSLSGVVF	2254	TGCGCAACCTGGGATAGGAGCCTGAGTGGTGTGGTGTTC
1431	CATWDRSLSAVVF	2255	TGCGCAACATGGGATAGAAGCCTGAGTGCTGTGGTCTTC
1432	CATWDRSLSAVPWVF	2256	TGCGCAACATGGGATAGAAGCCTGAGTGCTGTTCCTTGGGTGTTC
1433	CATWDRSLSAGVF	2257	TGCGCAACATGGGATCGCAGCCTGAGTGCTGGGGTGTTC
1434	CATWDRSLRAGVF	2258	TGCGCAACGTGGGATAGGAGCCTGCGTGCTGGGGTGTTC
1435	CATWDRSLNVYVL	2259	TGCGCAACATGGGATCGCAGTCTGAATGTTTATGTCCTC
1436	CATWDRILSAEVF	2260	TGCGCAACGTGGGATCGCATCCTGAGCGCTGAGGTGTTC

1437	CATWDRGLSTGVF	2261	TGCGCAACGTGGGATAGAGGCCTGAGTACTGGGGTGTTC
1438	CATWDNYLGAAVF	2262	TGCGCAACATGGGATAACTACCTGGGTGCTGCCGTGTTC
1439	CATWDNTPSNIVVF	2263	TGCGCAACATGGGATAACACGCCTTCGAATATTGTGGTATTC
1440	CATWDNTLSVWVF	2264	TGCGCAACATGGGATAATACACTGAGTGTGTGGGTCTTC
1441	CATWDNTLSVNWVF	2265	TGCGCAACATGGGATAACACCCTGAGTGTCAATTGGGTGTTC
1442	CATWDNTLNVFYVF	2266	TGCGCAACCTGGGATAACACACTGAATGTCTTTTATGTTTTC
1443	CATWDNRLSSVVF	2267	TGTGCGACATGGGATAATCGGCTCAGTTCTGTGGTCTTC
1444	CATWDNRLSAGVL	2268	TGCGCAACATGGGATAACCGCCTGAGTGCTGGGGTGCTC
1445	CATWDNRLSAGVF	2269	TGCGCAACGTGGGATAACAGGCTGAGTGCTGGGGTGTTC
1446	CATWDNRDWVF	2270	TGCGCAACATGGGATAACAGGGATTGGGTCTTC
1447	CATWDNNLGAGVF	2271	TGCGCAACATGGGATAACAACCTGGGTGCTGGGGTGTTC
1448	CATWDNKLTSGVF	2272	TGCGCAACATGGGATAACAAGCTGACTTCTGGGGTCTTC
1449	CATWDNILSAWVF	2273	TGCGCAACATGGGATAACATCCTGAGTGCCTGGGTGTTT
1450	CATWDNDIHSGLF	2274	TGCGCAACCTGGGACAACGATATACATTCTGGGCTGTTC
1451	CATWDLSLSALF	2275	TGCGCAACTTGGGATCTCAGCCTGAGTGCCCTGTTC
1452	CATWDITLSAEVF	2276	TGCGCAACATGGGATATCACCCTGAGTGCTGAGGTGTTC
1453	CATWDISPSAGGVF	2277	TGCGCAACGTGGGATATCAGCCCGAGTGCTGGCGGGGTGTTC
1454	CATWDISLSTGRAVF	2278	TGCGCAACATGGGATATCAGTCTAAGTACTGGCCGGGCTGTGTTC
1455	CATWDISLSQVF	2279	TGCGCAACATGGGATATCAGTCTGAGTCAGGTATTC
1456	CATWDIRLSSGVF	2280	TGCGCAACATGGGATATCAGGCTGAGTAGTGGAGTGTTC
1457	CATWDIGPSAGGVF	2281	TGCGCAACGTGGGATATCGGCCCGAGTGCTGGCGGGGTGTTC
1458	CATWDHSRAGVLF	2282	TGCGCAACATGGGATCACAGCCGGGCTGGTGTGCTATTC
1459	CATWDHSPSVGEVF	2283	TGCGCAACATGGGATCACAGTCCGAGTGTTGGAGAAGTCTTC
1460	CATWDHSLRVGVF	2284	TGCGCAACATGGGATCACAGCCTGCGTGTTGGGGTGTTC
1461	CATWDHSLNIGVF	2285	TGCGCAACATGGGATCACAGCCTGAACATTGGGGTGTTC
1462	CATWDHSLGLWAF	2286	TGCGCAACATGGGATCACAGCCTGGGTCTTTGGGCATTC
1463	CATWDHNLRLVF	2287	TGCGCCACATGGGATCACAATCTGCGTCTTGTTTTC
1464	CATWDHILASGVF	2288	TGCGCGACTTGGGATCACATCCTGGCTTCTGGGGTGTTC
1465	CATWDFSLSVWVF	2289	TGCGCAACATGGGATTTCAGCCTGAGTGTTTGGGTGTTC
1466	CATWDFSLSAWVF	2290	TGCGCAACATGGGATTTCAGCCTGAGTGCTTGGGTGTTC
1467	CATWDDTLTAGVF	2291	TGCGCAACATGGGATGACACCCTCACTGCTGGTGTGTTC
1468	CATWDDRLSAVLF	2292	TGCGCAACATGGGACGACAGGCTGAGTGCTGTGCTTTTC
1469	CATWDDRLDAAVF	2293	TGCGCAACATGGGATGACAGGCTGGATGCTGCGGTGTTC
1470	CATWDATLNTGVF	2294	TGCGCAACATGGGATGCGACCCTGAATACTGGGGTGTTC
1471	CATWDASLSVWLL	2295	TGCGCAACATGGGATGCCAGCCTGAGTGTTTGGCTGCTC
1472	CATWDASLSGGVF	2296	TGCGCGACATGGGATGCCAGCCTGAGTGGTGGGGTGTTC
1473	CATRDTTLSAVLF	2297	TGCGCAACACGGGATACCACCCTCAGCGCCGTTCTGTTC
1474	CATLGSSLSLWVF	2298	TGCGCTACATTGGGTAGTAGCCTGAGTCTCTGGGTGTTC
1475	CATIETSLPAWVF	2299	TGCGCAACAATCGAAACTAGCCTGCCTGCCTGGGTATTC
1476	CATGDRSLTVEVF	2300	TGCGCAACAGGGGACAGAAGCCTGACTGTTGAGGTATTC
1477	CATGDLGLTIVF	2301	TGCGCTACAGGGGATCTCGGCCTGACCATAGTCTTC
1478	CASWDYRGRSGWVF	2302	TGCGCATCATGGGATTACAGGGGGAGATCTGGTTGGGTGTTC
1479	CASWDTTLNVGVF	2303	TGCGCATCATGGGATACCACCCTGAATGTTGGGGTGTTC
1480	CASWDTTLGFVLF	2304	TGCGCTTCATGGGATACCACCCTGGGTTTTGTGTTATTC

1481	CASWDTSLSGGYVF	2305	TGCGCATCATGGGATACCAGCCTGAGTGGTGGTTATGTCTTC
1482	CASWDTSLRAGVF	2306	TGCGCATCATGGGATACCAGCCTCCGTGCTGGGGTGTTC
1483	CASWDTSLGAGVF	2307	TGCGCATCATGGGATACCAGCCTGGGTGCTGGGGTGTTC
1484	CASWDRGLSAVVF	2308	TGCGCATCATGGGACAGAGGCCTGAGTGCAGTGGTGTTC
1485	CASWDNVLRGVVF	2309	TGTGCTAGTTGGGATAACGTCCTGCGTGGTGTGGTATTC
1486	CASWDNRLTAVVF	2310	TGCGCGTCATGGGATAACAGGCTGACTGCCGTGGTTTTC
1487	CASWDASLSVAF	2311	TGCGCATCATGGGATGCAAGCCTGTCCGTCGCTTTC
1488	CASWDAGLSSYVF	2312	TGCGCTTCGTGGGATGCCGGCCTGAGTTCTTATGTCTTC
1489	CASGDTSLSGVIF	2313	TGCGCATCCGGGGATACCAGCCTGAGTGGTGTGATATTC
1490	CARWHTSLSIWVF	2314	TGCGCAAGATGGCATACGAGCCTAAGTATTTGGGTCTTC
1491	CAIWDTGLSPGQVAF	2315	TGCGCAATATGGGATACCGGCCTGAGTCCTGGCCAAGTTGCCTTC
1492	CAAWHSGLGLPVF	2316	TGCGCAGCATGGCATAGCGGCCTGGGTCTCCCGGTCTTC
1493	CAAWDYSLSAGVF	2317	TGCGCAGCATGGGATTACAGCCTGAGTGCTGGGGTGTTC
1494	CAAWDTTLRVRLF	2318	TGCGCAGCCTGGGATACTACCCTGCGTGTTAGGCTGTTC
1495	CAAWDTSLTAWVF	2319	TGCGCAGCATGGGATACCAGCCTGACTGCCTGGGTTTTC
1496	CAAWDTSLSGGVF	2320	TGCGCAGCATGGGATACCAGCTTGAGTGGTGGGGTGTTC
1497	CAAWDTSLSGEAVF	2321	TGCGCAGCATGGGATACCAGCCTGAGTGGCGAGGCTGTGTTC
1498	CAAWDTSLSGAVF	2322	TGCGCAGCATGGGATACCAGCTTGAGTGGTGCGGTGTTC
1499	CAAWDTSLSAWVF	2323	TGCGCAGCATGGGATACCAGCCTGAGTGCCTGGGTGTTC
1500	CAAWDTSLSAGVF	2324	TGCGCAGCATGGGATACCAGCCTGAGTGCTGGGGTATTC
1501	CAAWDTSLDTYVF	2325	TGCGCAGCATGGGATACCAGCCTGGATACTTATGTCTTC
1502	CAAWDTRLSGVLF	2326	TGCGCTGCATGGGATACCCGTCTGAGTGGTGTGTTATTC
1503	CAAWDTRLSAGVF	2327	TGCGCAGCATGGGATACCAGGCTGAGTGCTGGGGTGTTC
1504	CAAWDRSLSTGVF	2328	TGCGCAGCATGGGATCGCAGTCTGAGTACTGGAGTTTTC
1505	CAAWDIRRSVLF	2329	TGCGCAGCGTGGGATATCCGCCGGTCTGTCCTTTTC
1506	CAAWDHTQRLSF	2330	TGCGCTGCGTGGGATCACACTCAGCGTCTTTCCTTC
1507	CAAWDHSLSAGQVF	2331	TGCGCAGCATGGGATCACAGCCTGAGTGCTGGCCAGGTGTTC
1508	CAAVDTGLKEWVF	2332	TGCGCAGCAGTCGATACTGGTCTGAAAGAATGGGTGTTC

[0249] 预筛选CDR,使其不含氨基酸倾向(liabilities)、隐蔽剪接位点或核苷酸限制位点。CDR变异在至少两个个体中观察到,并且包含单突变、双突变和三突变的近种系空间。组装的顺序在图12C中示出。

[0250] 设计的VH结构域包括IGHV1-69和IGHV3-30。两个重链VH结构域中的每一个都与它们各自的不变的4个框架元件(FW1、FW2、FW3、FW4)和可变的3个CDR(H1、H2、H3)元件组装在一起。对于IGHV1-69,为H1设计了417个变体,为H2设计了258个变体。对于IGHV3-30,为H1设计了535个变体,为H2设计了165个变体。对于CDR H3,在IGHV1-69和IGHV-30中都使用相同的盒,因为两者均设计为使用相同的FW4,并且对于IGHV1-69和IGHV3-30,FW3的边缘也相同。CDR H3包含N末端和C末端元件,这些元件组合地接合至中央中间元件以生成1×10^{10}多样性。N末端和中间元件与“GGG”甘氨酸密码子重叠。中间和C末端元件与“GGT”甘氨酸密码子重叠。CDR H3包含5个单独组装的子池。各个N末端和C末端元件包含如表10所示的序列。

表10.N末端和C末端元件的序列

元件	SEQ ID NO	序列
茎 A	2333	CARDLRELECEEWT XXX SRGPCVDPRGVAGSFDVW
茎 B	2334	CARDMYYDF XXX EVVPADDAFDIW
茎 C	2335	CARDGRGSLPRPKGGP XXX YDSSEDSGGAFDIW
茎 D	2336	CARANQHF XXX GYHYYGMDVW
茎 E	2337	CAKHMSMQ XXX RADLVGDAFDVW

[0251] 实施例8.GPCR结合蛋白功能性

[0252] 对于GPCR结合蛋白,将来自噬菌体选择的前100-200个scFv转化为全长免疫球蛋白。在免疫球蛋白转化后,将克隆在ExpiCHO中瞬时转染以产生免疫球蛋白。Kingfisher和Hamilton用于批量IgG纯化,随后是实验室芯片以收集所有纯化的免疫球蛋白的纯度数据。如图13和表11所示,从10mL培养物中获得高收率和纯度。

表11免疫球蛋白纯度百分比

名称	IgG %纯度
mAb1	100
mAb2	100
mAb3	100
mAb4	100
mAb5	98
mAb6	100
mAb7	97
mAb8	100
mAb9	100
mAb10	100
mAb11	100
mAb12	100
mAb13	100
mAb14	100
mAb15	100

[0253] 然后生成表达GPCR靶标的稳定细胞系。图14显示通过FACS确认靶标表达。然后将表达>80%的靶标的细胞直接用于基于细胞的选择。针对过表达目的靶标的细胞进行五轮选择。每轮选择使用10^8个细胞。在对表达靶标的细胞进行选择之前,首先在10^8个CHO背景细胞上耗尽来自每一轮的噬菌体。通过增加在后续选择轮次中的洗涤次数来增加选择的严格性。监测每轮选择的富集比。

[0254] 使用FACS(图15A-图15C)和cAMP活性(图15D)测试纯化的IgG的细胞结合亲和力。观察到别构抑制。

[0255] 使用BVP ELISA测试纯化的IgG。详见图16,BVP ELISA显示一些克隆具有与比较物抗体相当的BVP评分。

[0256] 实施例9:VHH文库

[0257] 开发了合成VHH文库。对于具有定制的CDR多样性的“VHH比率”文库，使用Clustal Omega比对2391个VHH序列(iCAN数据库)，以确定每个位置的共有序列，并且框架衍生自每个位置的共有序列。对所有2391个序列的CDR进行了位置特异性变异分析，并将这种多样性引入到文库设计中。对于具有改组的CDR多样性的“VHH改组”文库，对iCAN数据库扫描纳米抗体序列中的独特CDR。鉴定了1239个独特的CDR1、1600个独特的CDR2和1608个独特的CDR3，并且框架衍生自iCAN数据库中2391个序列中每个框架位置的共有序列。每个独特的CDR都在共有框架中单独合成并改组，以生成理论多样性为3.2x10^9的文库。然后使用限制酶消化将该文库克隆到噬菌粒载体中。对于“VHH h改组”文库(合成的“人”VHH文库，具有改组的CDR多样性)，对iCAN数据库扫描纳米抗体序列中的独特CDR。鉴定了1239个独特的CDR1、1600个独特的CDR2和1608个独特的CDR3，并且框架1、3和4衍生自人种系DP-47框架。框架2衍生自iCAN数据库中2391个序列中每个框架位置的共有序列。每个独特的CDR都使用NUGE工具在部分人源化的框架中单独合成并改组，以生成理论多样性为3.2x10^9的文库。然后使用NUGE工具将该文库克隆到噬菌粒载体中。

[0258] 使用Carterra SPR系统评估VHH-Fc变体的结合亲和力和亲和力分布。VHH-Fc显示出对TIGIT的一定范围的亲和力，下限为12nM K_D，上限为1685nM K_D(在顶行突出显示；分别从左侧数第六个和从右侧数第三个)。表12提供了用于ELISA的VHH-Fc克隆、蛋白A(mg/ml)和K_D(nM)的具体值。

表12

克隆	ELISA	文库	ProA（mg/ml）	K_D（nM）
31-1	5.7	VHH h 改组	0.29	12
31-6	9.6	VHH h 改组	0.29	14
31-26	5.1	VHH h 改组	0.31	19
30-30	8.0	VHH 改组	0.11	23
31-32	8.0	VHH h 改组	0.25	27
29-10	5.0	VHH 比率	0.19	32
29-7	7.3	VHH 比率	0.28	41
30-43	13.5	VHH 改组	0.18	44
31-8	12.7	VHH h 改组	0.29	45
31-56	11.7	VHH h 改组	0.26	46
30-52	4.2	VHH 改组	0.22	49
31-47	8.8	VHH h 改组	0.23	53
30-15	9.3	VHH 改组	0.26	55
30-54	5.5	VHH 改组	0.30	58
30-49	10.3	VHH 改组	0.26	62
29-22	3.4	VHH 比率	0.27	65
29-30	9.2	VHH 比率	0.28	65
31-35	5.7	VHH h 改组	0.24	66
29-1	10.4	VHH 比率	0.09	68
29-6	6.8	VHH 比率	0.29	69
31-34	6.0	VHH h 改组	0.32	70
29-12	6.2	VHH 比率	0.23	70
30-1	5.4	VHH 改组	0.39	71
29-33	3.9	VHH 比率	0.15	74
30-20	4.6	VHH 改组	0.19	74
31-20	6.6	VHH h 改组	0.37	74
31-24	3.1	VHH h 改组	0.15	75
30-14	9.9	VHH 改组	0.19	75

30-53	7.6	VHH 改组	0.24	78
31-39	9.9	VHH h 改组	0.32	78
29-18	10.9	VHH 比率	0.19	78
30-9	8.0	VHH 改组	0.40	79
29-34	8.6	VHH 比率	0.21	80
29-27	8.6	VHH 比率	0.18	82
29-20	5.9	VHH 比率	0.26	83
30-55	6.0	VHH 改组	0.41	85
30-39	6.1	VHH 改组	0.07	88
31-15	6.2	VHH h 改组	0.32	88
29-21	4.3	VHH 比率	0.23	88
29-37	5.3	VHH 比率	0.26	89
29-40	6.6	VHH 比率	0.31	90
31-30	3.2	VHH h 改组	0.33	93
31-10	12.3	VHH h 改组	0.31	94
29-3	13.6	VHH 比率	0.11	94
30-57	5.2	VHH 改组	0.24	95
29-31	4.4	VHH 比率	0.18	96
31-27	8.1	VHH h 改组	0.31	96
31-33	6.0	VHH h 改组	0.32	96
30-40	7.1	VHH 改组	0.21	99
31-18	4.1	VHH h 改组	0.36	99
30-5	9.3	VHH 改组	0.05	100

[0259] 实施例10.CRTH2R的VHH文库

[0260] 通过类似于实施例9所述的方法开发了CRTH2R的VHH文库。简言之，产生表达CRTH2R的稳定细胞系，并通过FACS确认靶标表达。然后将表达>80％的靶标的细胞用于基于细胞的选择。针对稳定过表达目的靶标的细胞进行五轮基于细胞的选择。每轮选择使用10^8个细胞。在对表达靶标的细胞进行选择之前，首先在10^8个CHO背景细胞上耗尽来自每一轮的噬菌体。通过增加随后几轮选择中的洗涤次数来增加选择的严格性。然后使用胰蛋白酶从噬菌体中洗脱细胞，并扩增噬菌体用于下一轮淘选。来自第4轮和第5轮的总共1000个克隆通过NGS进行测序，以鉴定用于重新格式化为VHH-Fc的独特克隆。

[0261] 在175个独特CRTH2R VHH Fc结合物中，有26个结合物的靶细胞平均荧光强度(MFI)值是亲本细胞的2倍。变体CRTH2-41-51的数据见于图17A-图17B和表13A-表13B。表13A-13B显示了用RL1-A通道检测到的流式细胞术数据。变体CRTH2-44-59的数据见图18和图19。

表13A.淘选概述

文库	独特噬菌体	VHH-Fc FACS 结合物（亲本细胞的 2 倍的 MFI 值）
VHH h 改组	99	16
VHH 比率/改组	76	10

表13B.CRTH2-41-51数据

样品名称	子集名称	计数	中值:RL1-A
样品C7.fcs	CRTH2R细胞	8663	7441
样品E10.fcs	亲本细胞	11589	2120

[0262] 实施例11.用于CRTH2R的IgG的鉴定

[0263] 通过在CHO CRTH2R阳性细胞(GFP+)和亲本CHO细胞(GFP-)上进行测试并且比较亲本阴性和靶阳性细胞以排除假阳性来确定抗CRTH2R抗体的细胞结合。表14A中列出的抗体从100nM(15ug/mL)开始用3倍滴定总共8个点进行滴定。CRTH2R IgG抗体的重链和轻链序列示于表14B中。通过浓度通过平均荧光强度(MFI)检测的结合示于图20A-图20E中。与CRTH2-27在100nM的示例性门控点图和APC直方图示于图21A-图21B中。两种抗体(gPCR-51和gPCR-52)用作阳性对照。两个阳性对照的结合概况示于图22A-图22B中。

表14A.CRTH2R抗体可变重链和轻链序列

SEQ ID NO	CRTH2R 抗体	重链
2338	CRTH2-74	QVQLVESGGGVVQPGRSLRLSCAASGFSFSEYGIHWVRQAPGKGLEWVAVISYEG SNEYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARANQHFGPVAGG ATPSEEPGSQLTRAELGWDAPPGQESLADELLQLGTEHGYHYYGMDVWGQGTLV TVSS
2339	CRTH2-24	QVQLVQSGAEVKKPGSSVKVSCKASGGSFSNYGISWVRQAPGQGLEWMGGIIPLI GTANYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDMYYDFTLGPQS IGPLGEVVPADDAFDIWGQGTLVTVSS
2340	CRTH2-28	QVQLVQSGAEVKKPGSSVNVSCKASGGTFSDYAFSWVRQAPGQGLEWMGAIIPFF GTVNYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDMYYDFATGTG GPEDDLYPQGELNDGYRIEVVPADDAFDIWGQGTLVTVSS
2341	CRTH2-39	QVQLVQSGAEVKKPGSSVKVSCKASVDTFSRYSISWVRQAPGQGLEWMGGIIPVF DTTNYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDMYYDFGVILGG TAVGTNNGSANEVVPADDAFDIWGQGTLVTVSS
2342	CRTH2-19	QVQLVQSGAEVKKPGSSVKVSCKASGGTFSSHAINWVRQAPGQGLEWMGRIIPIV GTTTYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDMYYDFDYFGLT LTGDRNDDEVVPADDAFDIWGQGTLVTVSS
2343	CRTH2-9	QVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWMGGIIPIFG TANYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDMYYDFWLGDQS TGSLIGAEVVPADDAFDIWGQGTLVTVSS
2344	CRTH2-8	QVQLVQSGAEVKKPGSSVKVSCKASGGTFTDYAISWVRQAPGQGLEWMGGIIPFF GSPNYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDMYYDFAAGLEG TITEVFDEEGHQGGTEVVPADDAFDIWGQGTLVTVSS
2345	CRTH2-27	QVQLVESGGGVVQPGRSLRLSCAASGFTFDNYGMHWVRQAPGKGLEWVAVISYE GSNKKYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARDMYYDFGSIY GEDVVGELPEVVPADDAFDIWGQGTLVTVSS
2346	CRTH2-45	QVQLVESGGGVVQPGRSLRLSCAASGFTFSHYAMHWVRQAPGKGLEWVADISHE GSNKYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARDGRGSLPRPK GGPTSGGGFSTNIGYGFVVQSYDSSEDSGGAFDIWGQGTLVTVSS
2347	CRTH2-35	QVQLVQSGAEVKKPGSSVKVSCKASGGTFRSYAISWVRQAPGQGLEWMGGIIPIS GTTNYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARANQHFTRIFGNY QIYFGHFGYHYYGMDVWGQGTLVTVSS
2348	CRTH2-50	QVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYALSWVRKAPGQGLEWMGGTIPIF GTVNYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARANQHFTRVIGQP SPAVPSRGYIYHGYHYYGMDVWGQGTLVTVSS
2349	CRTH2-66	QVQLVESGGGVVQPGRSLRLSCAASGFDFSGYGMHWVRQAPGKGLEWVAVISYE GSNKFYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARDLRELECEEWT IEVHGQEFAVHQDRGGVFSRGPCVDPRGVAGSFDVWGQGTLVTVSS
2350	CRTH2-57	QVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAMSWVRQAPGQGLEWMGGIIPL FGTTDYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARANQHFVKIQGA PVSTPVPGFGTTGYHYYGMDVWGQGTLVTVSS
2351	CRTH2-32	QVQLVESGGGVVQPGRSLRLSCAASGFTFSKHGMHWVRQAPGKGLEWVAFISYE GSEKYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARDMYYDFHYST VGATYYYYLGSETEVVPADDAFDIWGQGTLVTVSS
2352	CRTH2-15	QVQLVQSGAEVKKPGSSVKVSCKASGGTFSTYAIDWVRQAPGQGLEWMGGIIPLF GSPNYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARANQHFFLYEGTSS SWLHVGHARYGYHYYGMDVWGQGTLVTVSS
2353	CRTH2-25	QVQLVQSGAEVKKPGSSVKVSCKASGGSFRSYGISWVRQAPGQGLEWMGRIIPLF GTPDYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDMYYDFEDVDEG SLYLDMGRTFEVVPADDAFDIWGQGTLVTVSS
2354	CRTH2-42	QVQLVESGGGVVQPGRSLRLSCAASGFAFSSYAMHWVRQAPGKGLEWVAVISYE GSNEYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARDLRELECEEWT VLQYGKFHMRWAESGEGSLSRGPCVDPRGVAGSFDVWGQGTLVTVSS
2355	CRTH2-55	QVQLVESGGGVVQPGRSLRLSCAASGFTFRSYDMHWVRQAPGKGLEWVAVISYE GSEKYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAKHMSMQASTEG

		DFGLEEVTGEGVDDRADLVGDAFDVWGQGTLVTVSS
2356	CRTH2-60	QVQLVQSGAEVKKPGSSVKVSCKASGGTFKNYAINWVRQAPGQGLEWMGAIIPK FGAANYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARANQHFSAVRGL AFGYGYRIGGYHYYGMDVWGQGTLVTVSS
2357	CRTH2-70	QVQLVQSGAEVKKPGSSVKVSCKASGGTFSNHAIIWVRQAPGQGLEWMGGIIPIFG TPSYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDMYYDFDVISAGV VGAGNPEVVPADDAFDIWGQGTLVTVSS
2358	CRTH2-48-9	EVQLLESGGGLVQPGGSLRLSCAASGFSFSTHAMSWVRQAPGKGLEWVSTIGGSG GSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARAHGDSSSWYFS YYYMDVWGQGTLVTVSS
2359	CRTH2-41-51	EVQLVESGGGLVQPGGSLRLSCAASGGIFRFNAMGWFRQAPGKERELVAGISGSG GDTYYADSVKGRFTISADNSKNTAYLQMNSLKPEDTAVYYCAAFRGIMRPDWGQ GTLVTVSS
2360	CRTH2-44-6	EVQLVESGGGLVQPGGSLRLSCAASGPTFDTYVMGWFRQAPGKEREFVAAISMSG DDTAYADSVKGRFTISADNSKNTAYLQMNSLKPEDTAVYYCATDLRGRGDVSEY EYDWGQGTLVTVSS
SEQ ID NO	**CRTH2R 抗体**	**轻链**
2361	CRTH2-74	QSVLTQPPSVSAAPGQKVTISCSGSTSNIGKNYVSWYQQLPGTAPKLLIYDDDERPS GIPDRFSGSMSGTSATLGITGLQTGDEADYYCEAWDADLSGAVFGGGTKLTVL
2362	CRTH2-24	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNFVSWYQQLPGTAPKLLIYDNIQRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDTSLSAVVFGGGTKLTVLRTV AAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTE QDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
2363	CRTH2-28	QSVLTQPPSVSAAPGQKVTISCSGSISNIGKNYVSWYQQLPGTAPKLLIYDDHKRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCATWDRGLSAAVFGGGTKLTVL
2364	CRTH2-39	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGDNDVSWYQQLPGTAPKLLIYDDDKRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCASWDTSLSGGYVFGGGTKLTVL
2365	CRTH2-19	QSALTQPASVSGSPGQSITISCTGTSSDVGGYDYVTWYQQHPGKAPKLMIYDVDTR PSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCSSYTTSTSYVFGGGTKLTVL
2366	CRTH2-9	QSVLTQPPSVSAAPGQKVTISCSGSTSNIGNNYVSWYQQLPGTAPKLLIYENDERPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCATWDTRLSAVVFGGGTKLTVL
2367	CRTH2-8	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGKNYVSWYQQLPGTAPKLLIYDNNQRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDTSLTSVVFGGGTKLTVL
2368	CRTH2-27	QSALTQPASVSGSPGQSITISCTGTSNDVGAYNFVSWYQQHPGKAPKLMIYDISNR PSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCSSYTRSNTRVFGGGTKLTVL
2369	CRTH2-45	QSVLTQPPSVSAAPGQKVTISCSGTSSNIENNYVSWYQQLPGTAPKLLIYDNVKRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDNTVSAPWVFGGGTKLTVL
2370	CRTH2-35	QSALTQPASVSGSPGQSITISCTGTSSDIGGYEFVSWYQQHPGKAPKLMIYGVSRRP SGVSNRFSGSKSGNTASLTISGLQAEDEADYYCGSYTSSSTPYVFGGGTKLTVL
2371	CRTH2-50	QSALTQPASVSGSPGQSITISCTGTSSDIGGYNFVSWYQQHPGKAPKLMIYDVSNRP QGVSNRFSGSKSGNTASLTISGLQAEDEADYYCSSYTSSNTYVVFGGGTKLTVL
2372	CRTH2-66	EIVMTQSPATLSVSPGERATLSCRASQGVGSNLAWYQQKPGQAPRLLIYRTSIRAT GIPARFSGSGSGTEFTLTISSLQSEDFAVYYCQQYYSWPPLTFGGGTKVEIK
2373	CRTH2-57	QSVLTQPPSVSAAPGQKVTISCSGSSSNIEDNYVSWYQQLPGTAPKLLIYDNFKRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDTSLSAALFGGGTKLTVL
2374	CRTH2-32	QSALTQPASVSGSPGQSITISCTGTSSGVGGYDYVSWYQQHPGKAPKLMIYDDNN RPSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCSSYTGSSTLYVFGGGTKLTVL
2375	CRTH2-15	QSVLTQPPSVSAAPGQKVTISCSGSGSNIGSNYVSWYQQLPGTAPKLLIYDNIRRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCAAWDTRLSAGVFGGGTKLTVL
2376	CRTH2-25	DIQMTQSPSSLSASVGDRVTITCRASQGISTYLNWYQQKPGKAPKLLIYATSSLQSG VPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSTP-WTFGGGTKVEIK
2377	CRTH2-42	QSALTQPASVSGSPGQSITISCTGTSSDVGGYRYVSWYQQHPGKAPKLMIYNVNYR PSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCSSYRSSSTLGVFGGGTKLTVL

2378	CRTH2-55	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGDNFVSWYQQLPGTAPKLLIYDDDERPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCGAWDRSLSAVVFGGGTKLTVL
2379	CRTH2-60	QSVLTQPPSVSAAPGQKVTISCSGSTSNIGINYVSWYQQLPGTAPKLLIYENRKRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCATWDASLKNLVFGGGTKLTVL
2380	CRTH2-70	QSVLTQPPSVSAAPGQKVTISCSGSTSNIGNNFVSWYQQLPGTAPKLLIYDNEKRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDERQTDESYVFGGGTKLTVL
2381	CRTH2-48-9	DIQMTQSPSSLSASVGDRVTITCRASQSISDYVNWYQQKPGKAPKLLIYGASILQTG VPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSFTTPWTFGGGTKVEIK

表14B.可变重链CDR3序列

SEQ ID NO	CRTH2R 抗体	CDRH3
2382	CRTH2-74	CARANQHFGPVAGGATPSEEPGSQLTRAELGWDAPPGQESLADELLQLGTEH GYHHYYGMDVW
2383	CRTH2-24	CARDMYYDFTLGPQSIGPLGEVVPADDAFDIW
2384	CRTH2-28	CARDMYYDFATGTGGPEDDLYPQGELNDGYRIEVVPADDAFDIW
2385	CRTH2-39	CARDMYYDFGVILGGTAVGTNNGSANEVVPADDAFDIW
2386	CRTH2-19	CARDMYYDFDYFGLTLTGDRNDDEVVPADDAFDIW
2387	CRTH2-9	CARDMYYDFWLGDQSTGSLIGAEVVPADDAFDIW
2388	CRTH2-8	CARDMYYDFAAGLEGTITEVFDEEGHQGGTEVVPADDAFDIW
2389	CRTH2-27	CARDMYYDFGSIYGEDVVGELPEVVPADDAFDIW
2390	CRTH2-45	CARDGRGSLPRPKGGPTSGGGFSTNIGYGFVVQSYDSSEDSGGAFDIW
2391	CRTH2-35	CARANQHFTRIFGNYQIYFGHFGYHYYGMDVW
2392	CRTH2-50	CARANQHFTRVIGQPSPAVPSRGYIYHGYHYYGMDVW
2393	CRTH2-66	CARDLRELECEEWTIEVHGQEFAVHQDRGGVFSRGPCVDPRGVAGGSFDVW
2394	CRTH2-57	CARANQHFVKIQGAPVSTPVPGFGTTGYHYYGMDVW
2395	CRTH2-32	CARDMYYDFHYSTVGATYYYYLGSETEVVPADDAFDIW
2396	CRTH2-15	CARANQHFFLYEGTSSSWLHVGHARYGYHYYYGMDVW
2397	CRTH2-25	CARDMYYDFEDVDEGSLYLDMGRTFEVVPADDAFDIW
2398	CRTH2-42	CARDLRELECEEWTVLQYGKFHMRWAESGEGSLSRGPCVDPRGVAGSFDVW
2399	CRTH2-55	CAKHMSMQASTEGDFGLEEVTGEGVDDRADLVGDAFDVM
2400	CRTH2-60	CARANQHFSAVRGLAFGYGYRIGGYHYYGMDVW
2401	CRTH2-70	CARDMYYDFDVISAGVVGAGNPEVVPADDAFDIW
2402	CRTH2-74	CARDMYYDFDVISAGVVGAGNPEVVPADDAFDIW

[0264] 在随后的示例中,五种抗体在cAMP测定中显示出具有功能效应;CRTH2-9、CRTH2-27、CRTH2-50、CRTH2-32和CRTH2-42。在图23A-图23B中对这些抗体的结合曲线进行了比较。

[0265] 实施例12.使用cAMP测定的拮抗剂活性

[0266] 测定CRTH2R IgG抗体文库以确定PGD2诱导的cAMP信号中的拮抗剂功能。简而言之,将细胞用IgG(滴定1:3)在室温下预孵育1小时。随后,由于CRTH2R是$G\alpha_i$偶联的,因此在毛喉素存在下,在37℃用PGD2(0.59nM)刺激细胞30分钟。

[0267] 示出了显示抗体对以相对光单位(rlu)的检测信号的影响的结果。在最高测试浓度(300nM)下,一些CRTH2R IgG引起信号向上偏转,表明由PGD2刺激诱导的cAMP信号的抑制。为了比较,条形图显示了对于所测试的三个最高IgG浓度,IgG处理的与对照处理的比率,如图24A所示。图24B中描绘的抗体示出了在33nM导致超过20%的拮抗剂活性的CRTH2R IgG抗体,特别是CRTH2-74、CRTH2-24、CRTH2-28、CRTH2-19、CRTH2-45、CRTH2-9、CRTH2-8、CRTH2-15、CRTH2-42、CRTH2-60和CRTH2-70。

[0268] 实施例13.PGD2诱导的cAMP信号的别构调节

[0269] 测定CRTH2R IgG抗体的别构活性。通过测定PGD2诱导的cAMP信号中的CRTH2R IgG抗体来确定别构调节。简而言之,在没有IgG抗体或有100nM CRTH2R IgG抗体的情况下重新孵育细胞。随后,在毛喉素存在下用不同浓度的PGD2刺激细胞,然后测定cAMP活性。

[0270] cAMP测定的结果见于图25。PGD2剂量响应曲线右移(和IC_{50}值增加)表明负别构效应。如图25所示,与单独的PGD2相比,5种CRTH2R IgG(CRTH2-9、CRTH2-27、CRTH2-50、CRTH2-32和CRTH2-42)导致IC50倍数差异>2.0,表明它们是负别构调节剂。

[0271] 实施例14.PGD2诱导的cAMP信号的激动剂活性

[0272] 测定CRTH2R IgG抗体的激动剂功能。通过在PGD2诱导的cAMP信号中测定实施例11中描述的CRTH2R IgG抗体来确定激动剂活性。

[0273] 简而言之,在毛喉素存在下,用PGD2或CRTH2R IgG抗体处理细胞。CRTH2R IgG抗体

包括CRTH2-74、CRTH2-24、CRTH2-28、CRTH2-39、CRTH2-19、CRTH2-9、CRTH2-8、CRTH2-27、CRTH2-45、CRTH2-35、CRTH2-50、CRTH2-66、CRTH2-57、CRTH2-32、CRTH2-15、CRTH2-25、CRTH2-42、CRTH2-55、CRTH2-60和CRTH2-70。在37℃进行30分钟的处理刺激。然后进行cAMP测定(数据未显示)。

[0274] 实施例15.显示别构调节剂的对照实验

[0275] 针对已知的CRTH2R拮抗剂(小分子OC000459)和两种对照抗体测定了别构调节。进行类似于实施例13中所述的实验。简而言之,用OC000459、比较物CRTH2R AB51抗体或比较物CRTH2R AB52抗体处理细胞。然后在毛喉素存在下用PGD2刺激细胞。

[0276] 结果显示在图26A-图26C中。OC000459导致曲线强烈右移和IC50值增加到459倍(图26A)。用CRTH2R AB51孵育不会导致IC50值发生变化(图26B)。用比较物抗体#52孵育导致IC50值降低到之前的1/3.5,表明它是一种正别构调节剂,即它具有激动作用(图26C)。

[0277] 实施例16.用于拮抗剂调节的CRTH2Rβ-抑制蛋白募集测定

[0278] 确定了九种CRTH2R IgG抗体的拮抗剂调节作用。九种CRTH2R IgG抗体包括CRTH2-9、CRTH2-27、CRTH2-50、CRTH2-32、CRTH2-42、CRTH2-74、CRTH2-55、CRTH2-28和CRTH2-39。这九种抗体与OC000459相比的拮抗剂功能使用PGD2诱导的β-抑制蛋白募集来确定。包括使用小分子OC000459的阳性对照的结果示于图27A-图27D中。

[0279] 实施例17.用于别构调节的CRTH2Rβ-抑制蛋白募集测定

[0280] 测定了九种CRTH2R IgG的别构调节。九种CRTH2R IgG包括CRTH2-9、CRTH2-27、CRTH2-50、CRTH2-32、CRTH2-42、CRTH2-74、CRTH2-55、CRTH2-28和CRTH2-39。这九种抗体与OC000459相比的别构调节是使用PGD2诱导的β-抑制蛋白募集来确定的。

[0281] 简而言之,将细胞用IgG(100nM)在室温下预孵育1小时,然后在37℃进行PGD2刺激90分钟。数据针对每个图中的第一个数据点(最低PGD2和零Ab)进行了归一化。包括使用小分子OC000459的阳性对照在内的结果示于图27A-图27D中。

[0282] 实施例18.CRTH2R超免疫免疫球蛋白文库

[0283] 使用类似于实施例9中所述的方法创建超免疫免疫球蛋白(IgG)文库。简而言之,超免疫IgG文库是通过分析由来自3名健康供体各自的超过3700万个独特IgH序列组成的人幼稚和记忆B细胞受体序列的数据库生成的。从分析中收集并使用类似于实施例1-3的方法单独构建了超过200万个CDRH3序列。在文库合成步骤中,去除了在开发中经常造成问题的任何重复的CDRH3和潜在的责任基序。然后将这些CDRH3序列多样性组合组装并整合到DP47人框架中,以构建具有$1x10^{10}$大小的高功能抗体Fab文库。该设计的示意图可见于图28中。

[0284] 生成了CRTH2R超免疫免疫球蛋白文库。简而言之,针对稳定过表达感兴趣靶标的细胞进行了五轮基于细胞的选择。每轮选择使用10^8个细胞。在选择靶标表达细胞之前,首先在10^8个CHO背景细胞上耗尽来自每一轮的噬菌体。通过在随后的几轮选择中增加洗涤次数来增加选择的严格性。然后使用胰蛋白酶将细胞从噬菌体中洗脱下来,并且扩增噬菌体以进行下一轮淘选。

[0285] 评估了CRTH2R免疫球蛋白的结合亲和力和PGD2诱导的cAMP的别构调节剂功能。如图29A-图29F中所见,三种特异性CRTH2R免疫球蛋白被鉴定为对hCRTH2R具有亚纳摩尔至个位数纳摩尔的细胞结合亲和力并且在别构cAMP测定中具有抑制活性。此三种CRTH2R免疫球蛋白CRTH2-48-3、CRTH2-48-21和CRTH2-48-27的序列见于表15中。

表15.CRTH2R序列

可变重链		
SEQ ID NO:	IgG	氨基酸序列
2403	CRTH2-48-3	EVQLVESGGGLVQAGGSLRLSCAASGSIFRINAMGWFRQAPGKEREGVAAINN FGTTKYADSVKGRFTISADNAKNTVYLQMNSLKPEDTAVYYCAAVRWGPRN DDRYDWGQGTQVTVSS
2404	CRTH2-48-21	EVQLVESGGGLVQAGGSLRLSCAASGSFFSINAMGWFRQAPGKEREFVAGITR SGVSTSYADSVKGRFTISADNAKNTVYLQMNSLKPEDTAVYYCAAHRIVVGG TSVGDWRWGQGTQVTVSS
2405	CRTH2-48-27	EVQLVESGGGLVQAGGSLRLSCAASGSIFHINAMGWFRQAPGKEREGVAAINN FGTTKYADSVKGRFTISANNAKNTVYLQMNSLKPEDTAVYYCAAVRWGPRN DDRYDWGQGTLVTVSS
可变轻链		
2406	CRTH2-48-3	DIQMTQSPSSLSASVGDRVTITCRASQSISSDLNWYQQKPGKAPKLLIYFASGL QSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSSPLTFGGGTKVEIKR
2407	CRTH2-48-21	DIQMTQSPSSLSASVGDRVTITCRTSQSISNYLNWYQQKPGKAPKLLIYATSSLE SGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSTLLTFGGGTKVEIKR
2408	CRTH2-48-27	DIQMTQSPSSLSASVGDRVTITCRASQSISRYLHWYQQKPGKAPKLLIYGASRL ESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCRQSYSTPWTFGGGTKVEIKR

[0286] 实施例19.具有不同CDR的GPCR文库

[0287] 使用CDR随机化方案创建GPCR文库。

[0288] 简言之,基于GPCR抗体序列设计GPCR文库。分析了超过60种不同的GPCR抗体,并使用CDR随机化方案修饰了来自这些GPCR的序列。

[0289] 重链IGHV3-23设计见图30A。如图30A所示,IGHV3-23CDRH3具有四种不同的长度:23个氨基酸、21个氨基酸、17个氨基酸和12个氨基酸,每个长度都有其残基多样性。四种长度的比例如下:CDRH3 23个氨基酸长度为40%,CDRH3 21个氨基酸长度为30%,CDRH3 17个氨基酸长度为20%,CDRH3 12个氨基酸长度为10%。CDRH3多样性被确定为9.3×10^{8},全重链IGHV3-23多样性为$1.9x10^{13}$。

[0290] 重链IGHV1-69设计见图30B。如图30B所示,IGHV1-69CDRH3具有四种不同的长度:20个氨基酸、16个氨基酸、15个氨基酸和12个氨基酸,每个长度都有其残基多样性。四种长度的比例如下:CDRH3 20个氨基酸长度为40%,CDRH3 16个氨基酸长度为30%,CDRH3 15个氨基酸长度为20%,CDRH3 12个氨基酸长度为10%。CDRH3多样性被确定为9×10^{7},全重链IGHV-69多样性为$4.1x10^{12}$。

[0291] 轻链IGKV 2-28和IGLV 1-51设计见图30C。分析抗体轻链CDR序列的位置特异性变异。选择了两个具有固定CDR长度的轻链框架。κ链和轻链的理论多样性分别被确定为13800和5180。

[0292] 最终的理论多样性被确定为$4.7x10^{17}$,最终生成的Fab文库的多样性为$6x10^{9}$。参见图30D。

[0293] 实施例20.具有不同CDR的CRTH2R库

[0294] CRTH2R文库是使用类似于实施例19中所述的CDR随机化方案创建的。

[0295] 简而言之,CRTH2R文库是基于GPCR抗体序列设计的。分析了超过60种不同的GPCR抗体,并使用CDR随机化方案修饰了来自这些GPCR的序列。使用CDR随机化方案设计的CRTH2R变体IgG被纯化并测定以确定基于细胞的亲和力测量以及用于功能分析。

[0296] 虽然本文已经示出并描述了本公开的优选实施方案,但对于本领域技术人员明显的是,这些实施方案仅以示例的方式提供。在不脱离本公开内容的情况下,本领域技术人员

将会想到许多变化、改变和替换。应当理解，在实施本公开时可以采用本文所述本公开的实施方案的各种替代方案。旨在以所附权利要求书限定本公开的范围，并且由此涵盖这些权利要求范围内的方法和结构及其等同物。

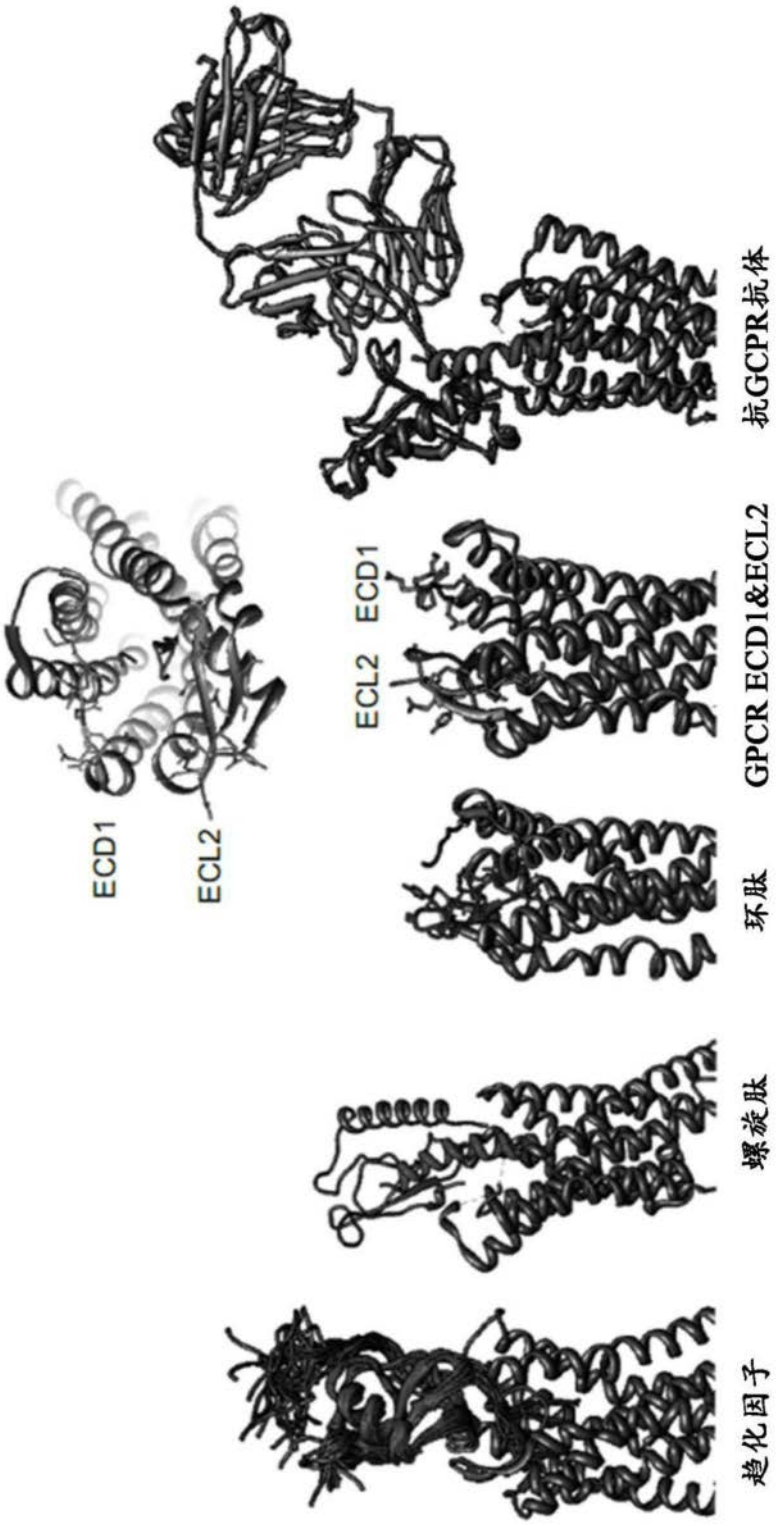


图1

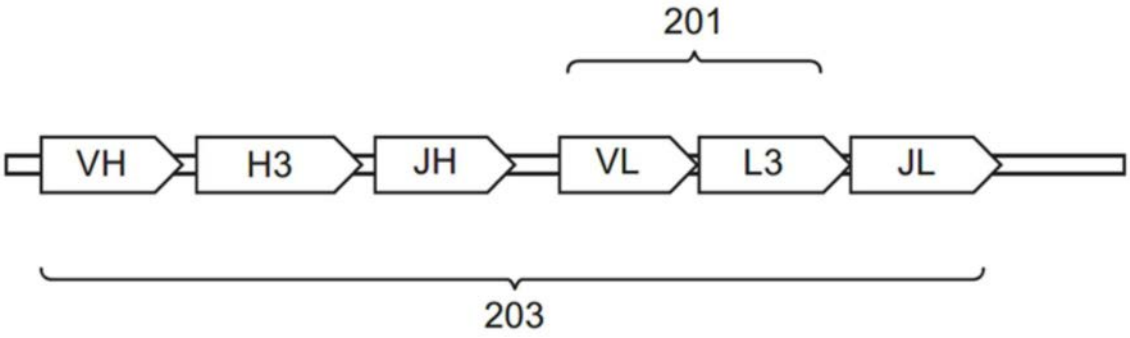


图2A

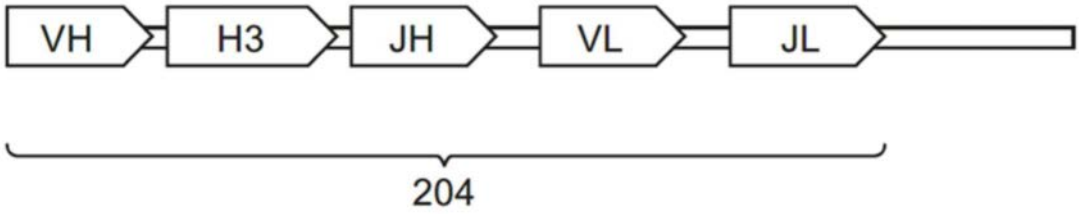


图2B

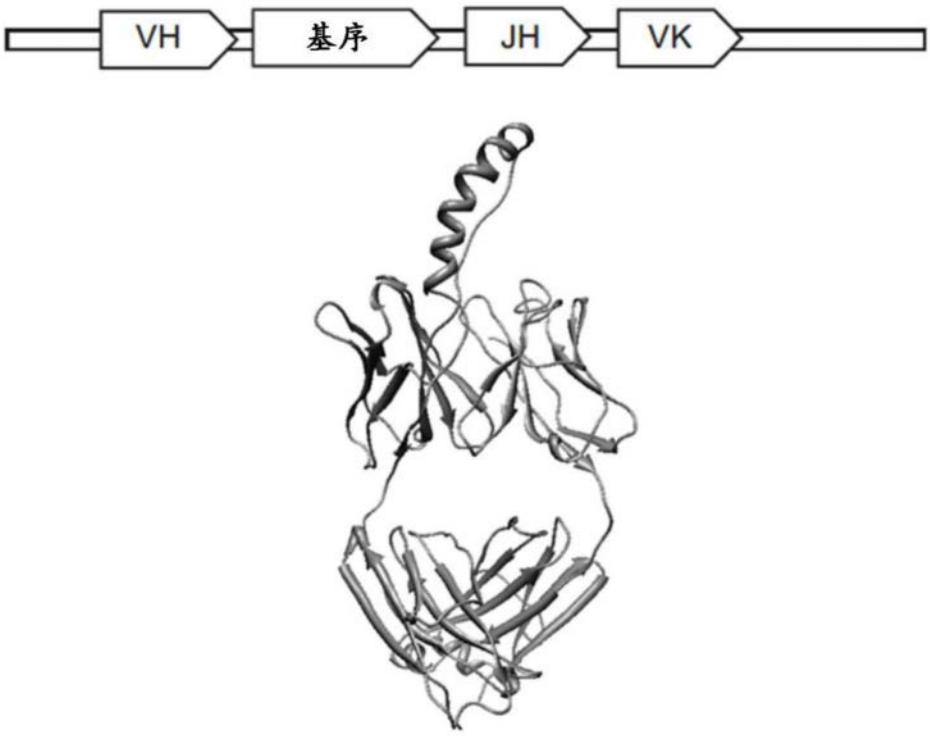


图3

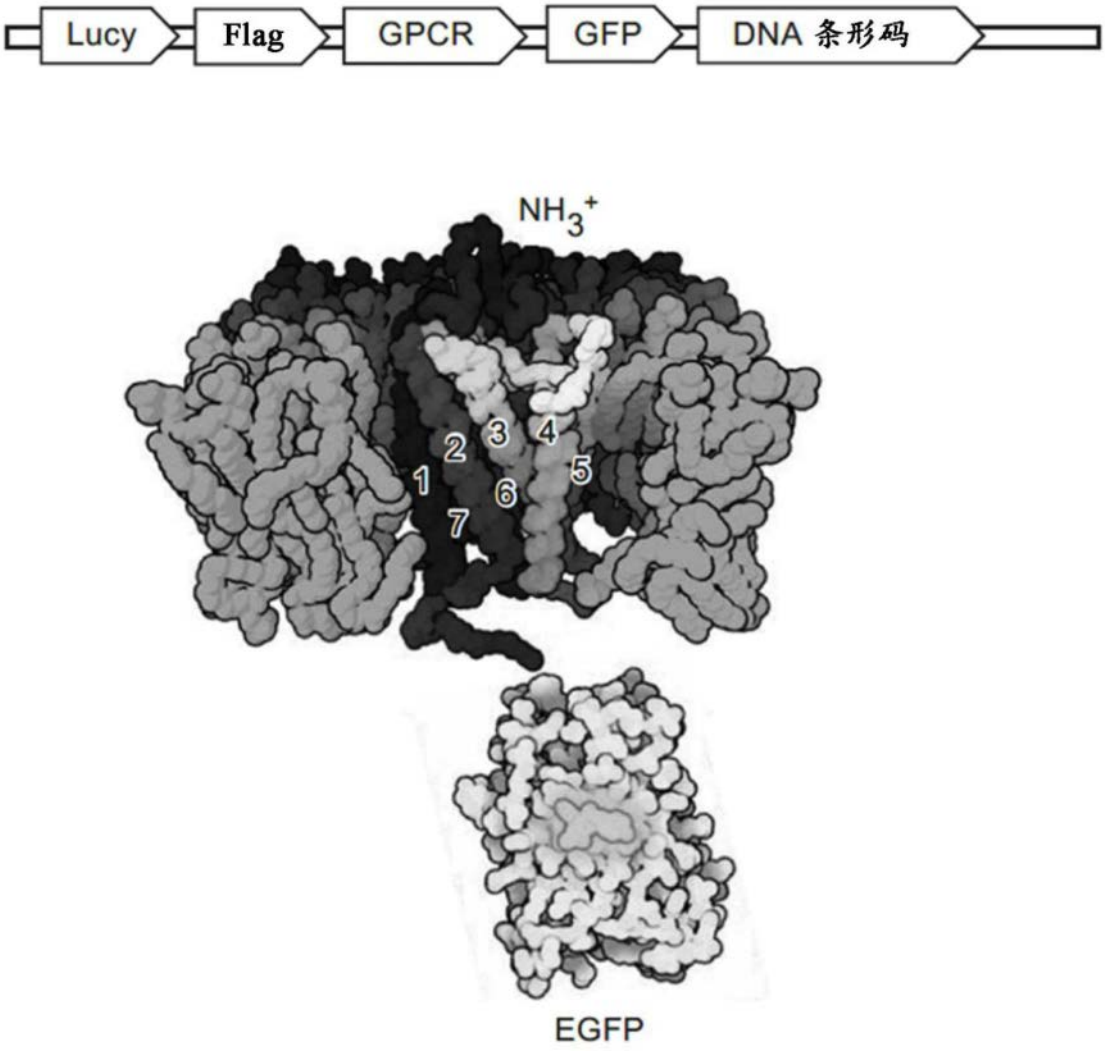


图4

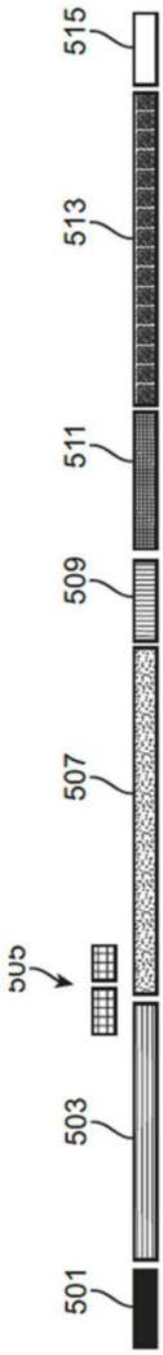


图5

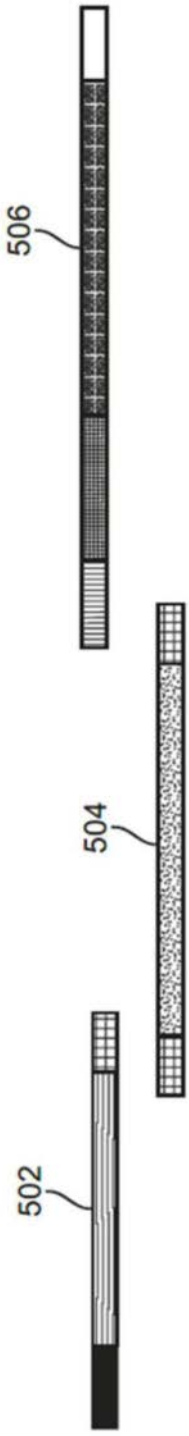


图6

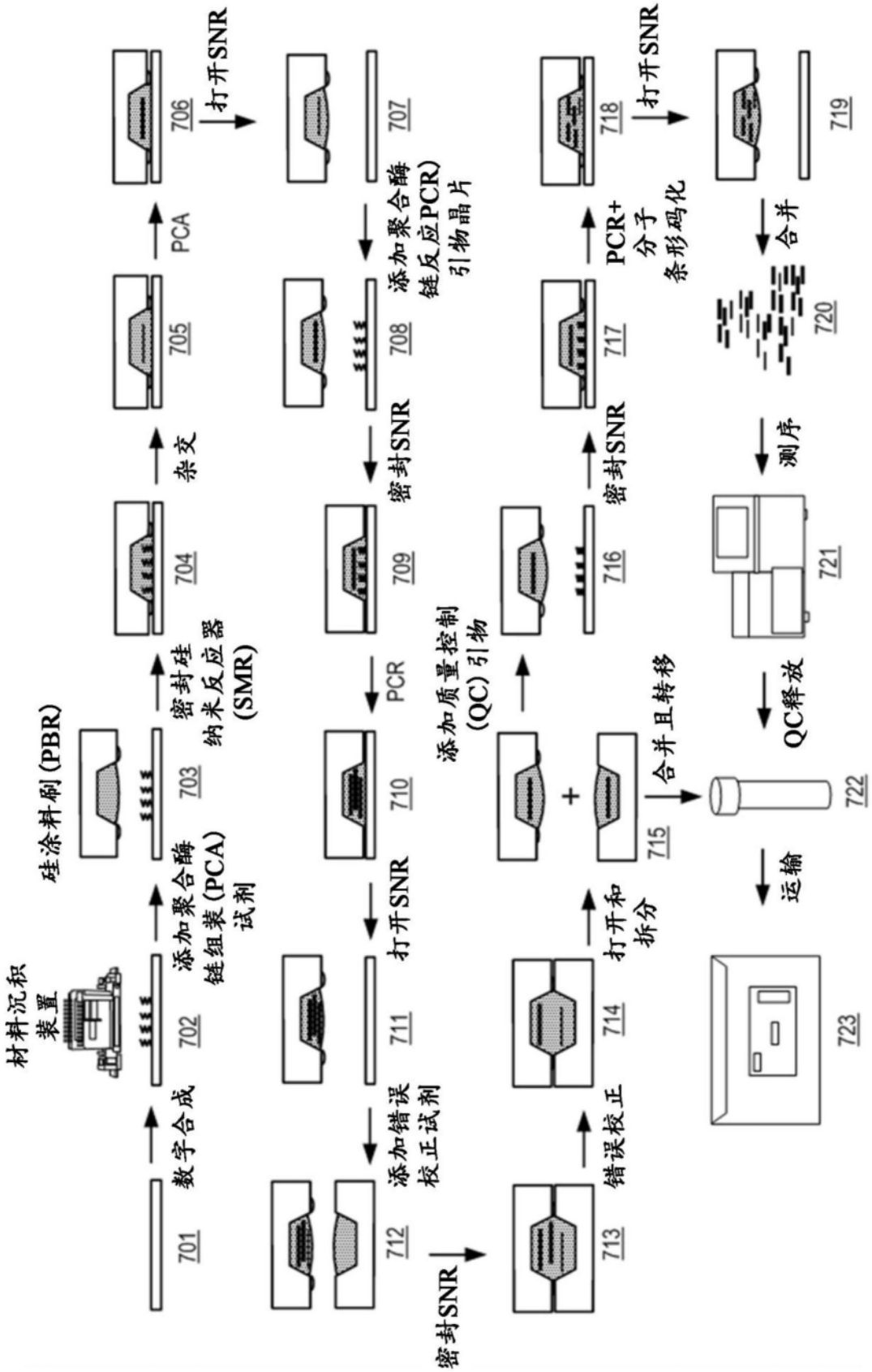

材料沉积装置
硅涂料刷 (PBR)
701
数字合成
702
添加聚合酶链组装 (PCA)试剂
703
密封硅纳米反应器 (SMR)
704
杂交
705
PCA
706
打开SNR
712
添加错误校正试剂
711
打开SNR
710
PCR
709
密封SNR
708
添加聚合酶链反应PCR)引物晶片
707
密封SNR
713
错误校正
714
打开和拆分
715
添加质量控制(QC)引物
716
密封SNR
717
PCR+分子条形码化
718
打开SNR
719
合并
720
测序
721
QC释放
722
运输
723
合并且转移

图7

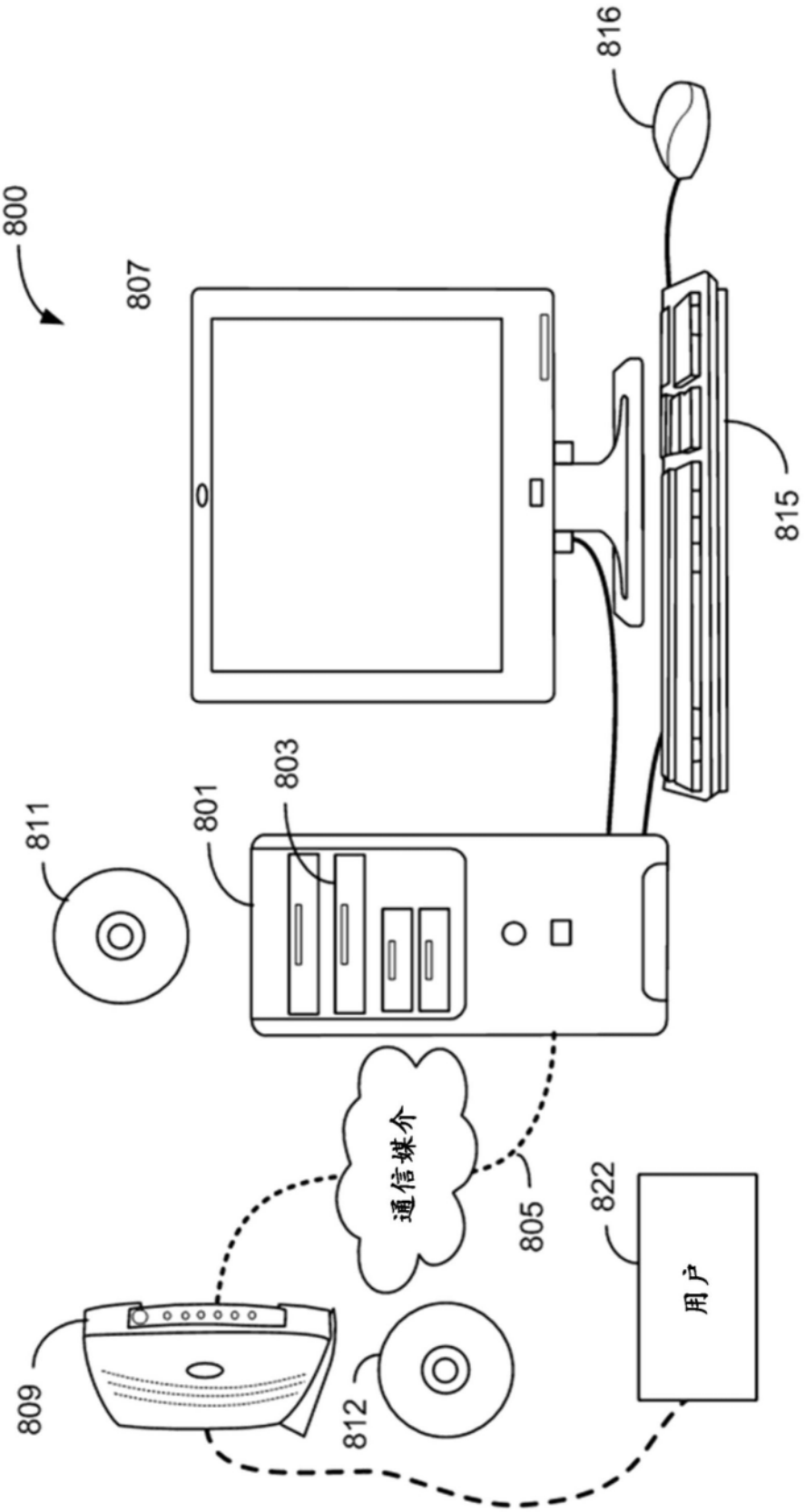


图8

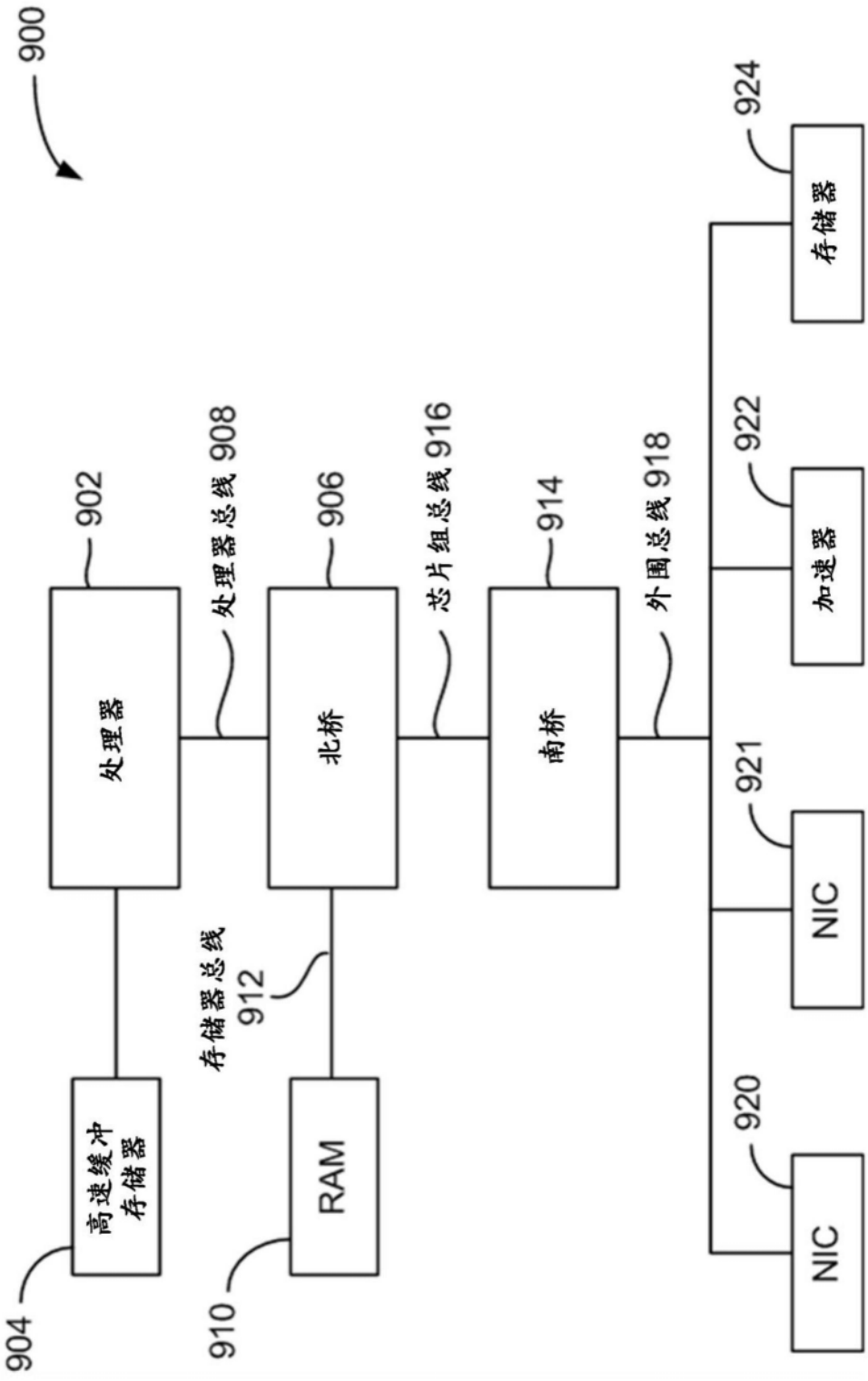

900
处理器
902
处理器总线 908
北桥
906
芯片组总线 916
南桥
914
外围总线 918
存储器
924
加速器
922
NIC
921
NIC
920
高速缓冲存储器
904
存储器总线 912
RAM
910

图9

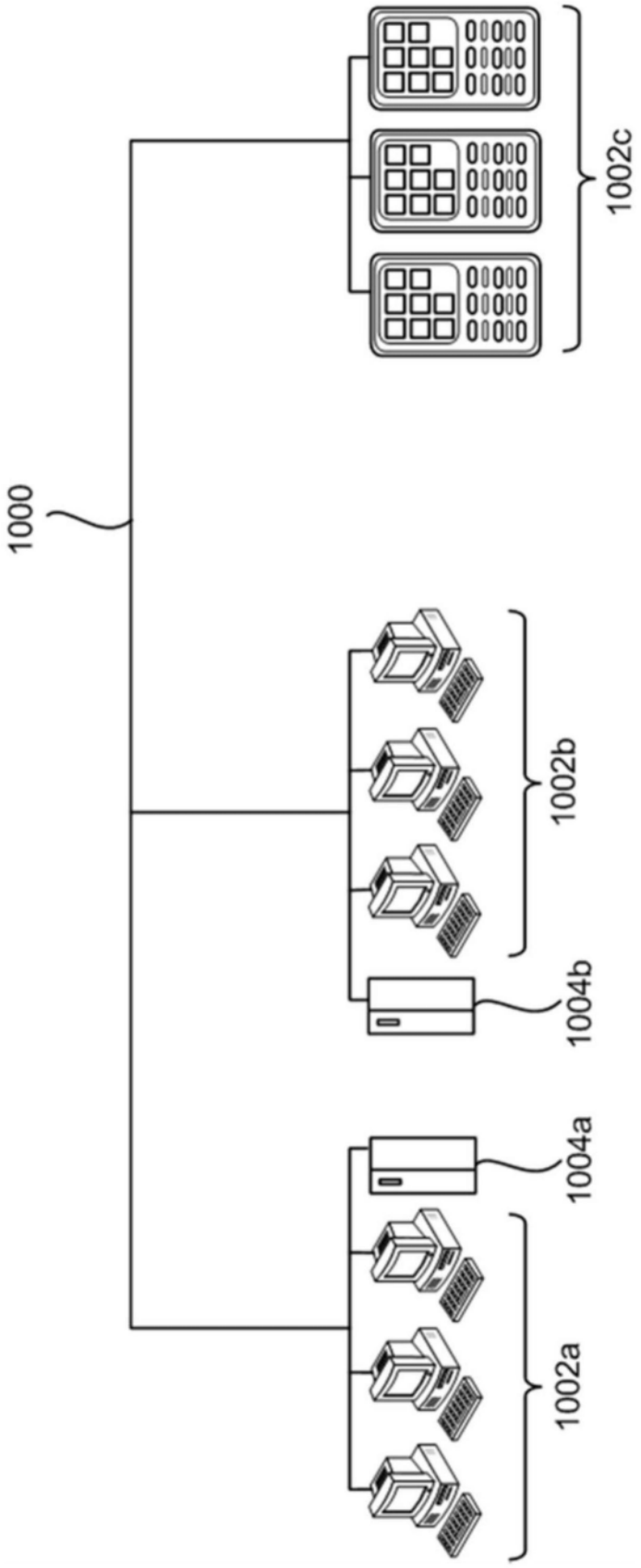


图10

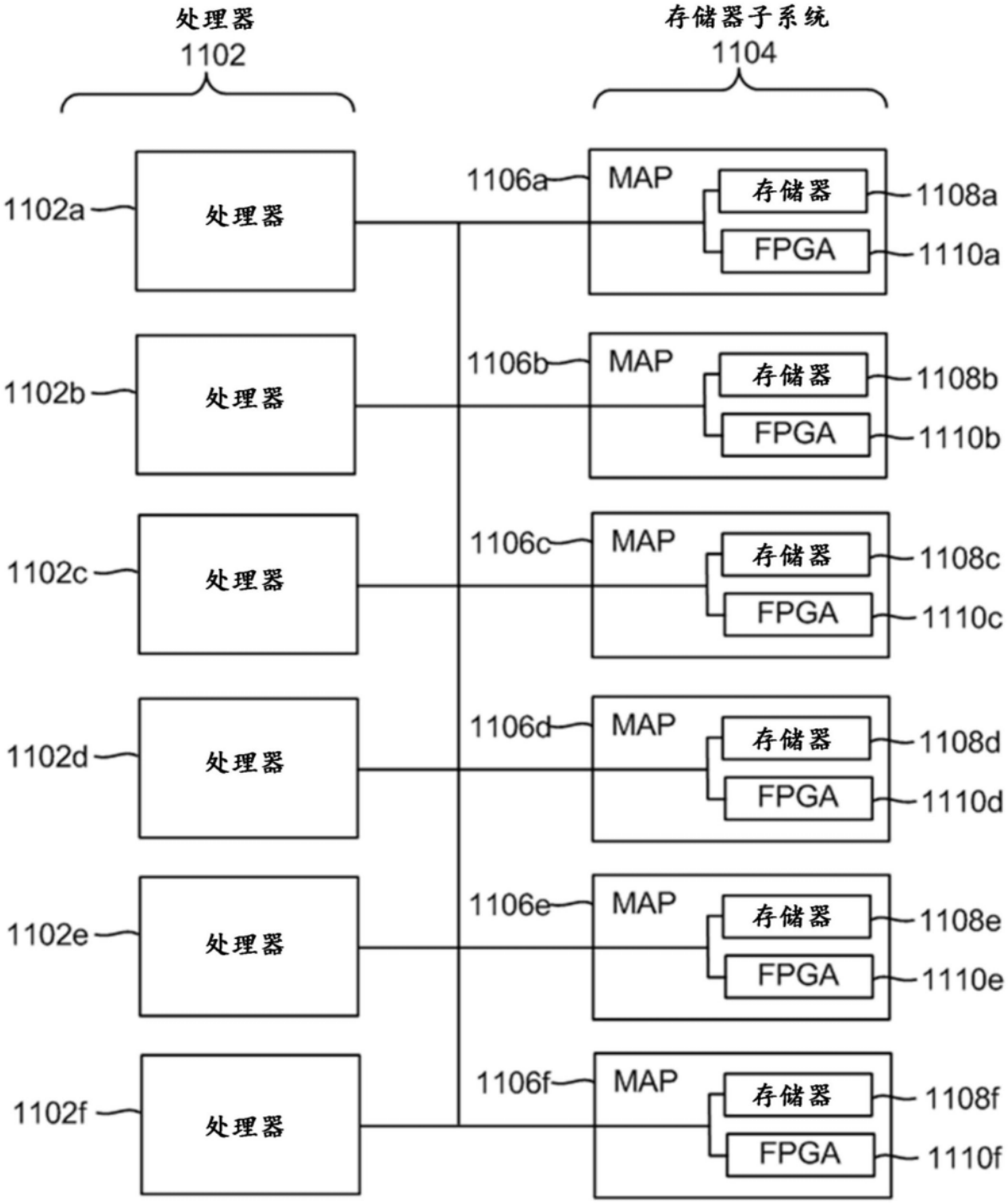

处理器
1102
存储器子系统
1104
1102a
处理器
1106a
MAP
存储器
FPGA
1108a
1110a
1102b
处理器
1106b
MAP
存储器
FPGA
1108b
1110b
1102c
处理器
1106c
MAP
存储器
FPGA
1108c
1110c
1102d
处理器
1106d
MAP
存储器
FPGA
1108d
1110d
1102e
处理器
1106e
MAP
存储器
FPGA
1108e
1110e
1102f
处理器
1106f
MAP
存储器
FPGA
1108f
1110f

图11

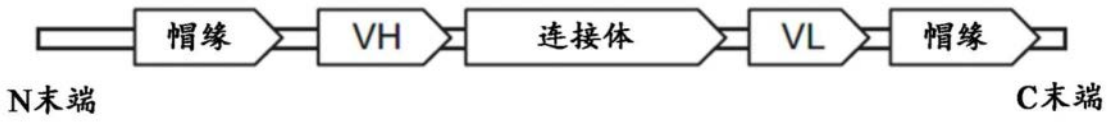

帽缘
VH
连接体
VL
帽缘
N末端
C末端

图12A

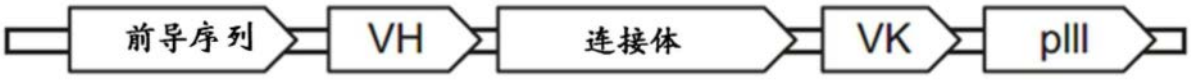


图12B



图12C

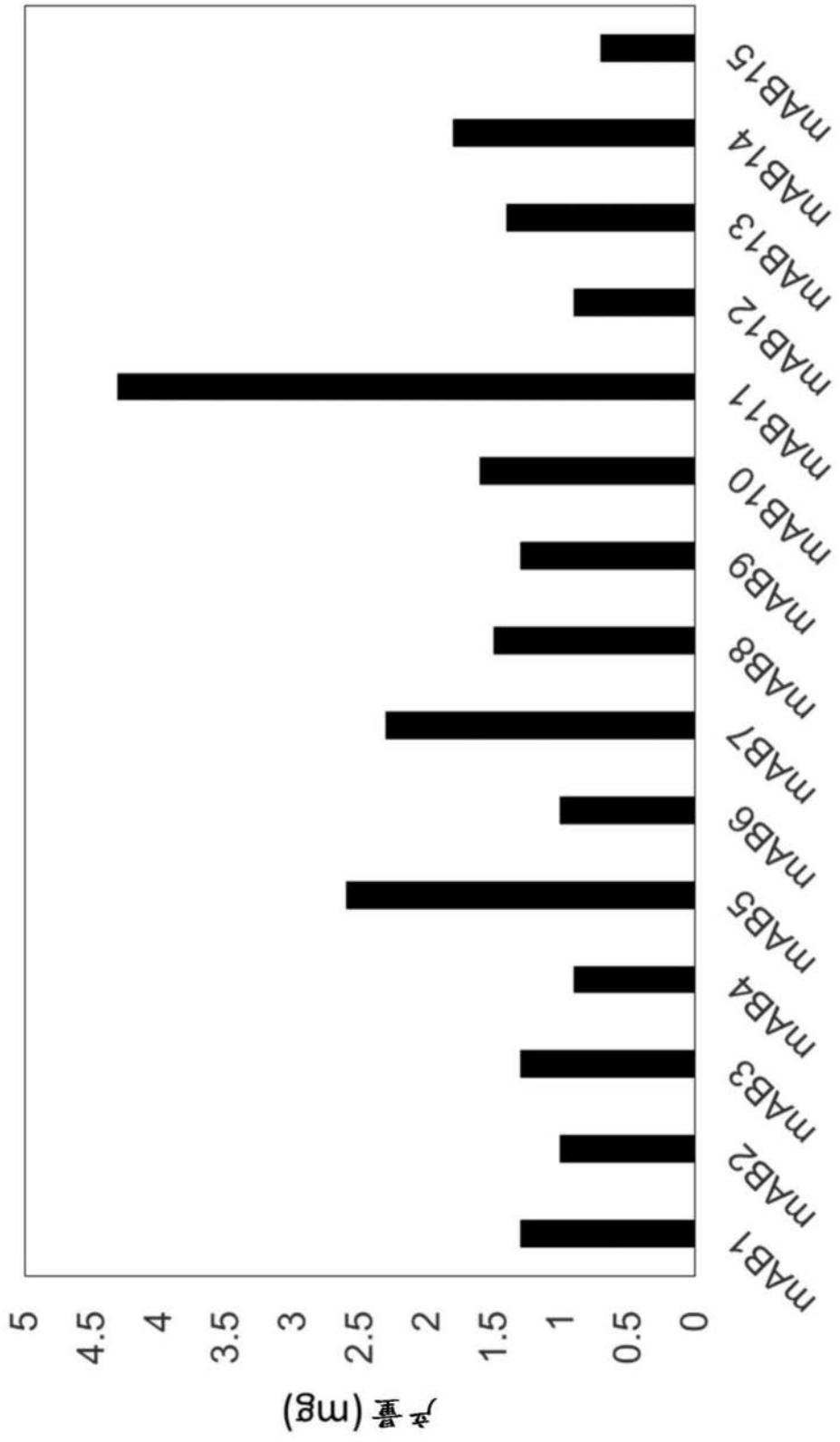

mAB15
mAB14
mAB13
mAB12
mAB11
mAB10
mAB9
mAB8
mAB7
mAB6
mAB5
mAB4
mAB3
mAB2
mAB1
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
产量 (mg)

图13

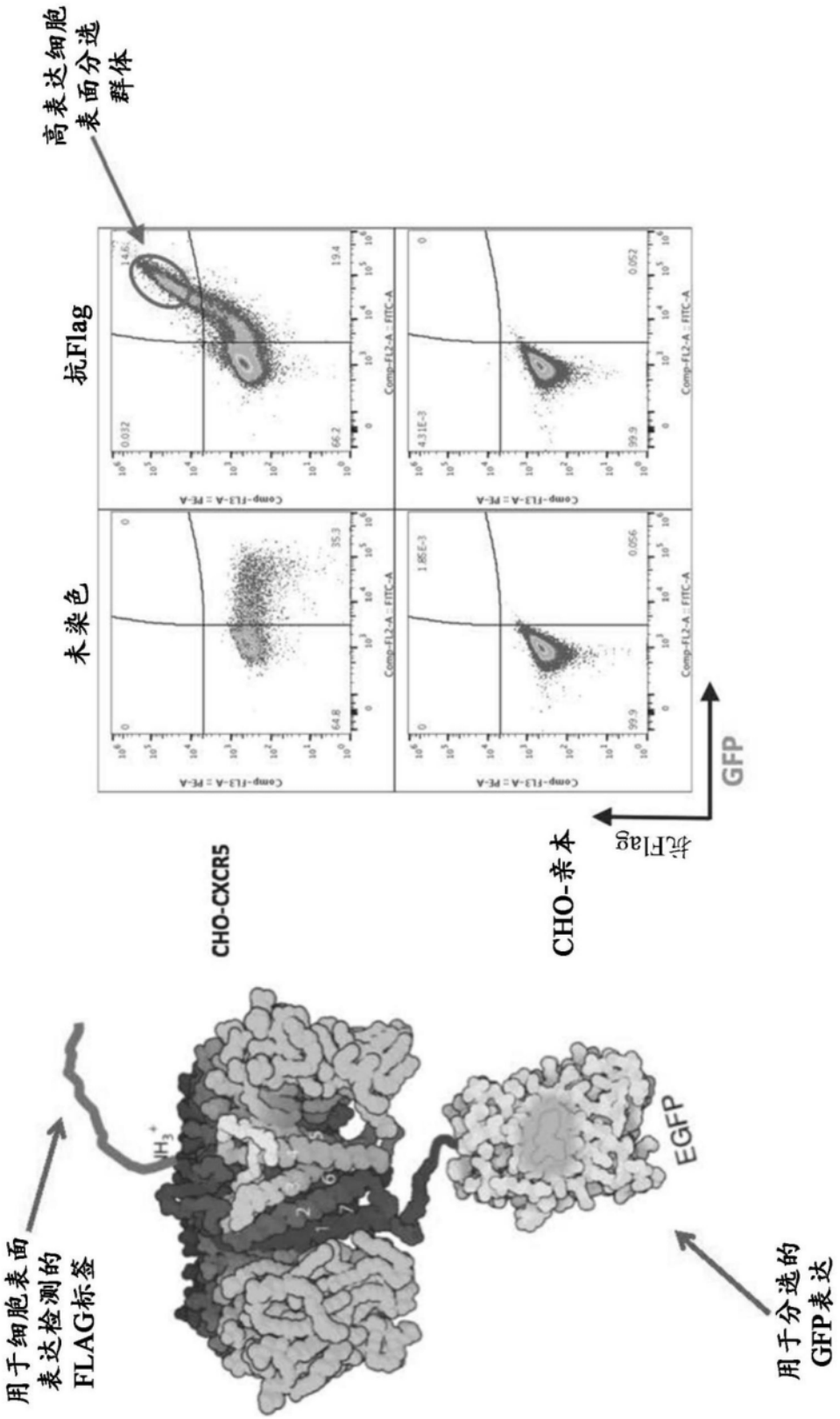

高表达细胞
表面分选
群体
抗Flag
未染色
CHO-CXCR5
CHO-亲本
抗Flag
GFP
用于细胞表面
表达检测的
FLAG标签
EGFP
用于分选的
GFP表达

图14

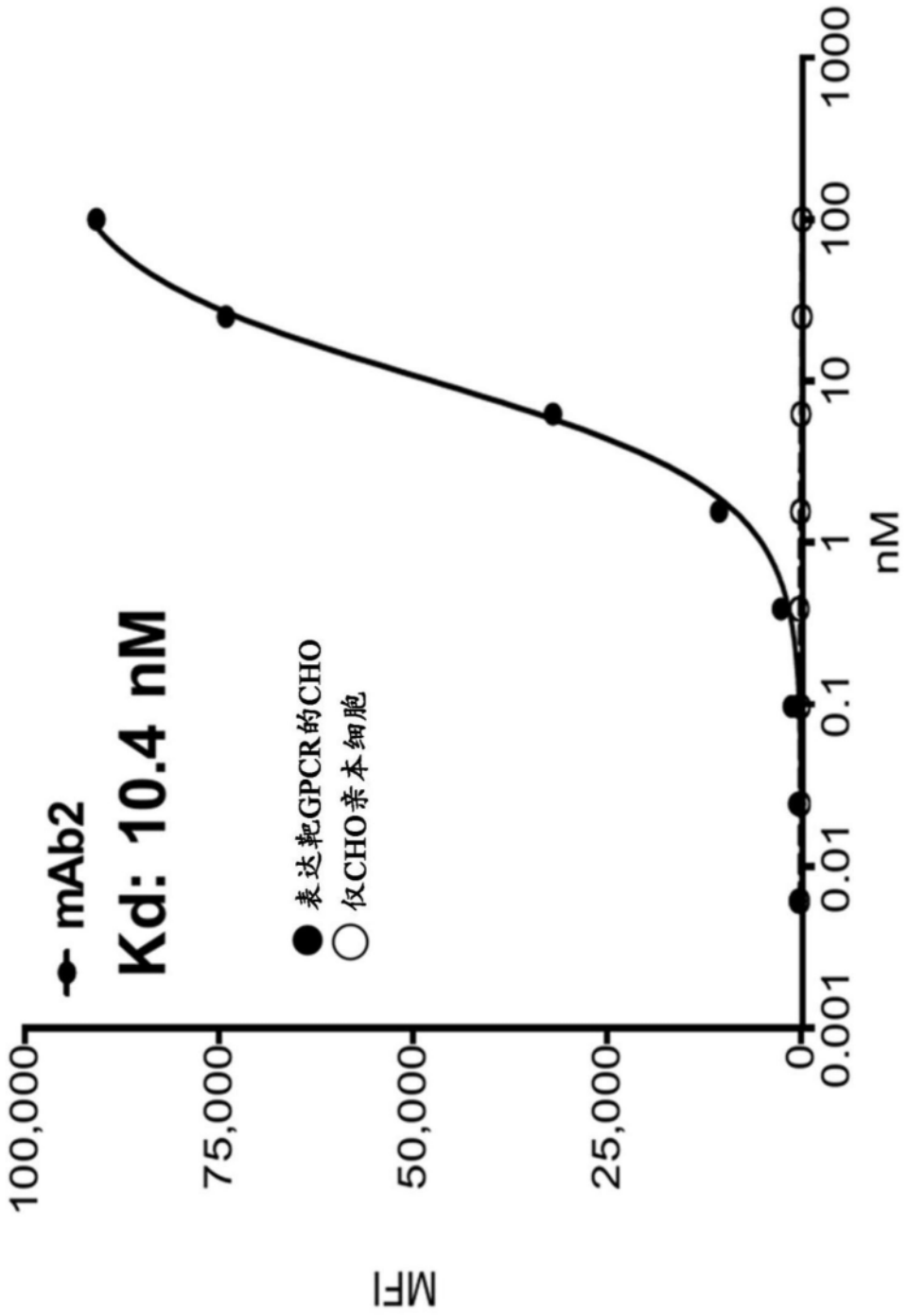

mAb2
Kd: 10.4 nM
表达靶GPCR的CHO
仅CHO亲本细胞
MFI
100,000
75,000
50,000
25,000
0
0.001
0.01
0.1
1
10
100
1000
nM

图15A

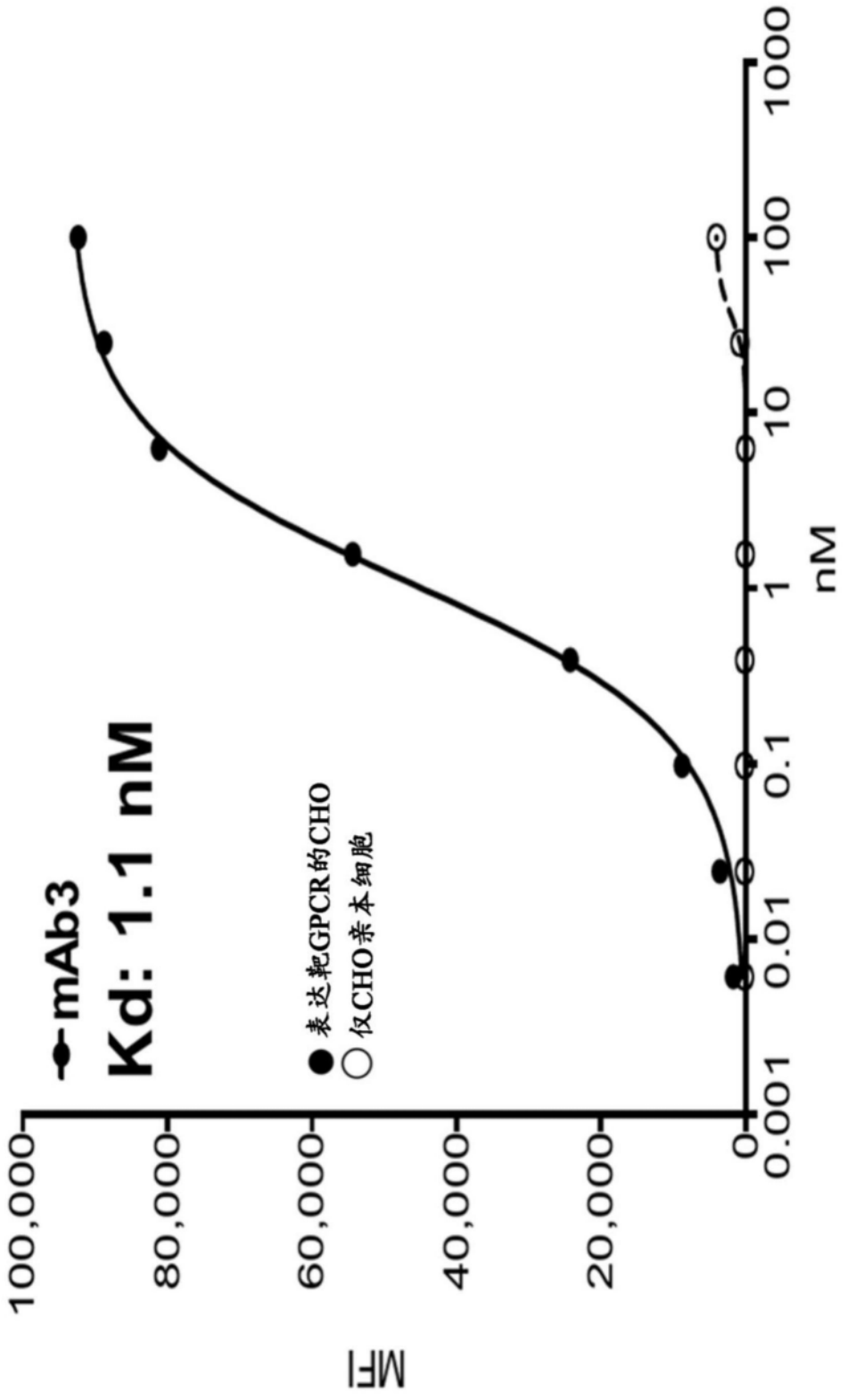

mAb3
Kd: 1.1 nM
表达靶GPCR的CHO
仅CHO亲本细胞
MFI
nM
100,000
80,000
60,000
40,000
20,000
0
0.001
0.01
0.1
1
10
100
1000

图15B

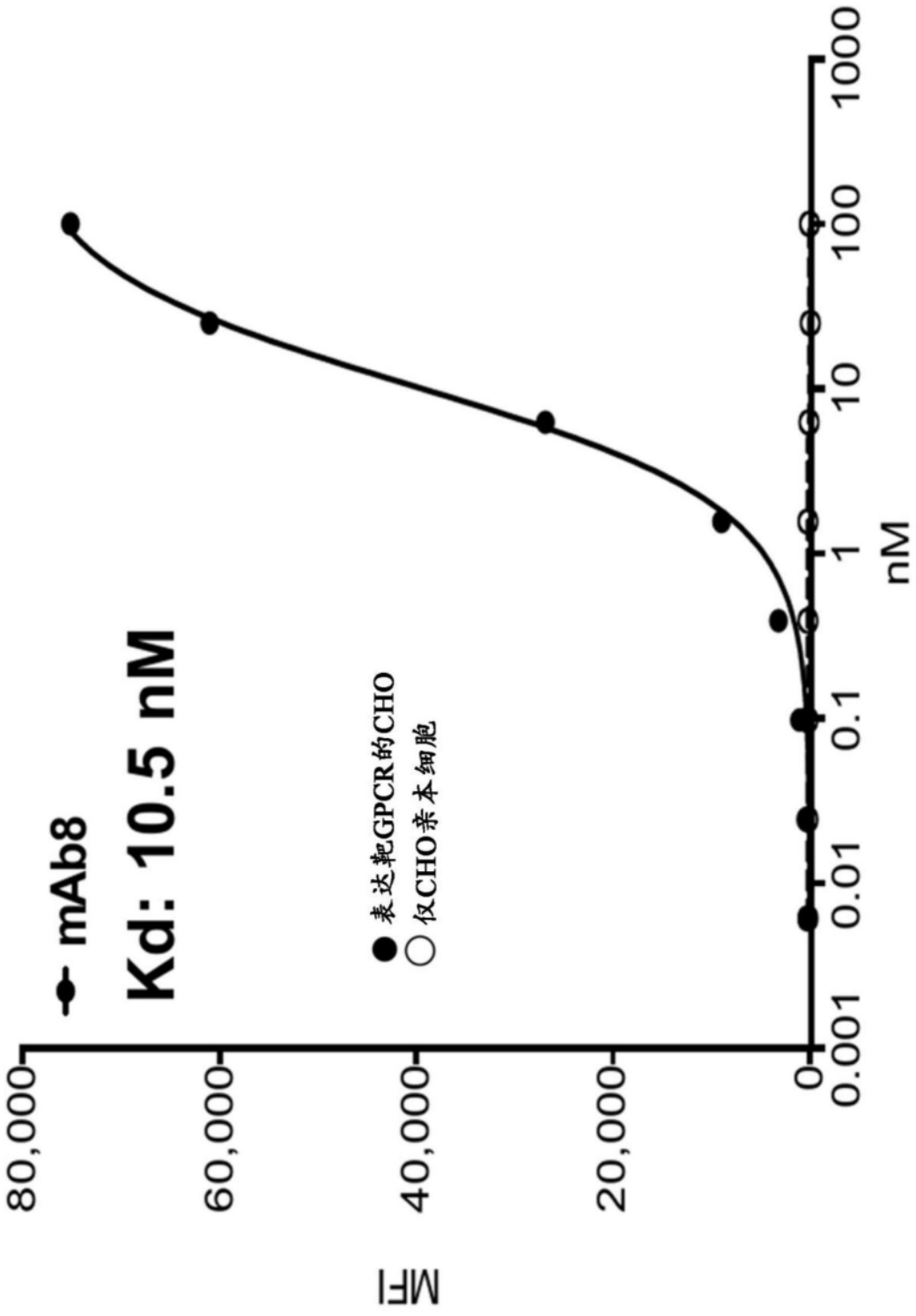

mAb8
Kd: 10.5 nM
● 表达靶GPCR的CHO
○ 仅CHO亲本细胞
MFI
80,000
60,000
40,000
20,000
0
0.001
0.01
0.1
1
10
100
1000
nM

图15C

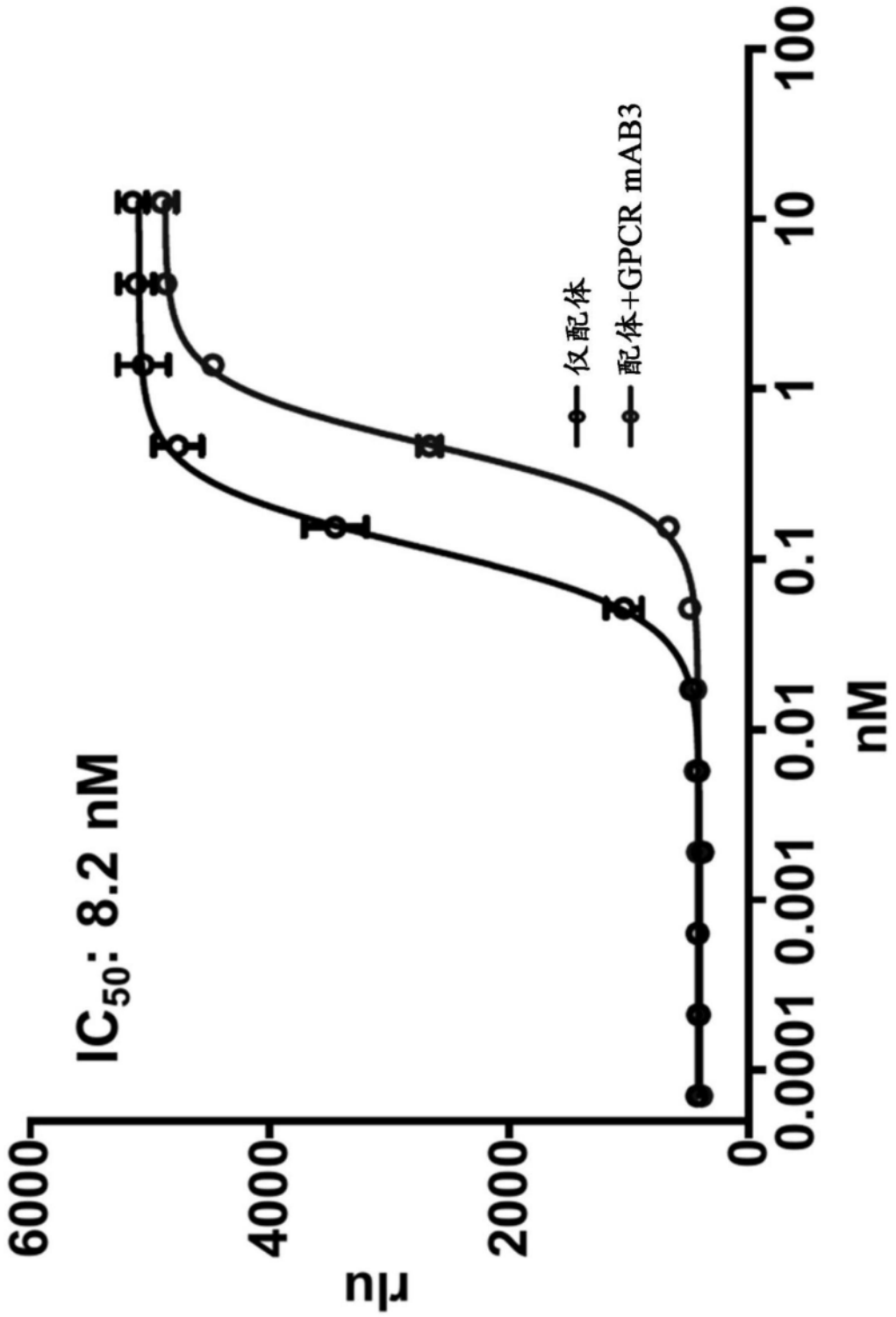

仅配体
配体+GPCR mAB3
IC50: 8.2 nM
rlu
nM
6000
4000
2000
0
0.0001
0.001
0.01
0.1
1
10
100

图15D

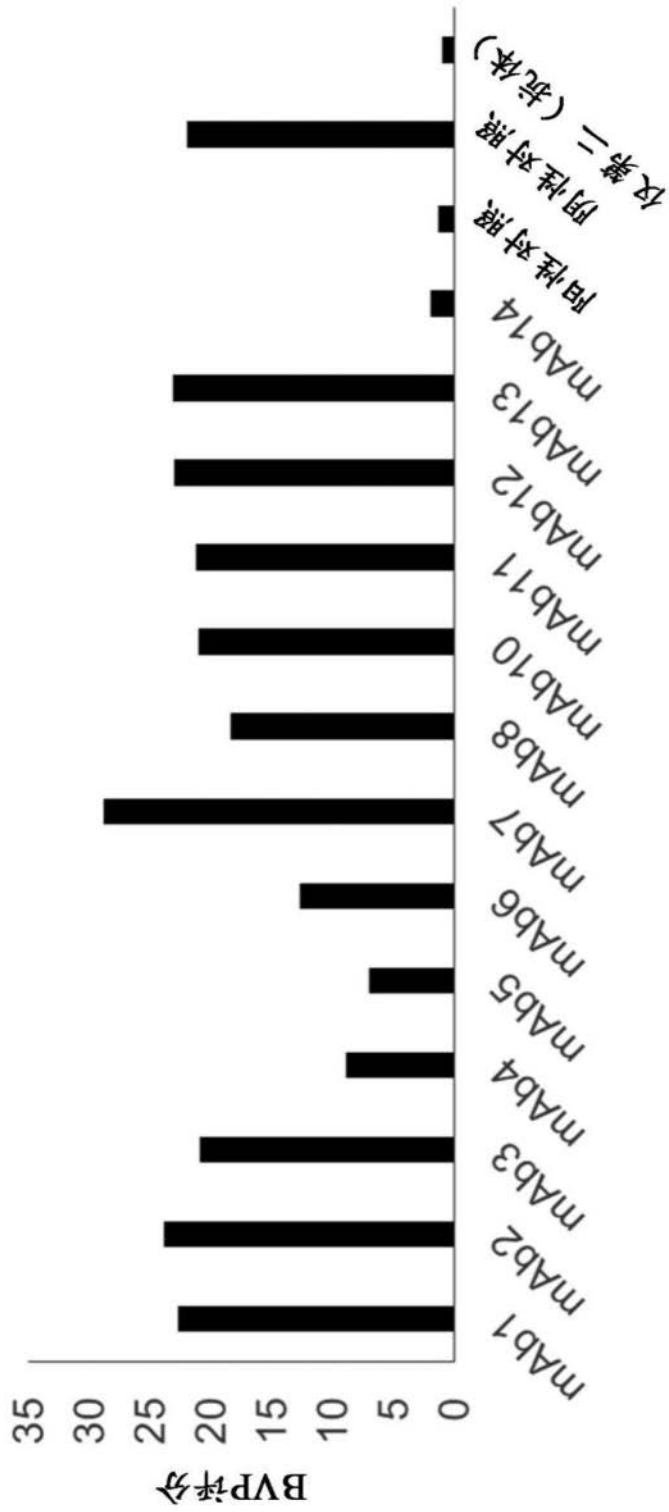

mAb1
mAb2
mAb3
mAb4
mAb5
mAb6
mAb7
mAb8
mAb10
mAb11
mAb12
mAb13
mAb14
阳性对照
阴性对照
仅第二（抗体）
0
5
10
15
20
25
30
35
BVP评分

图16

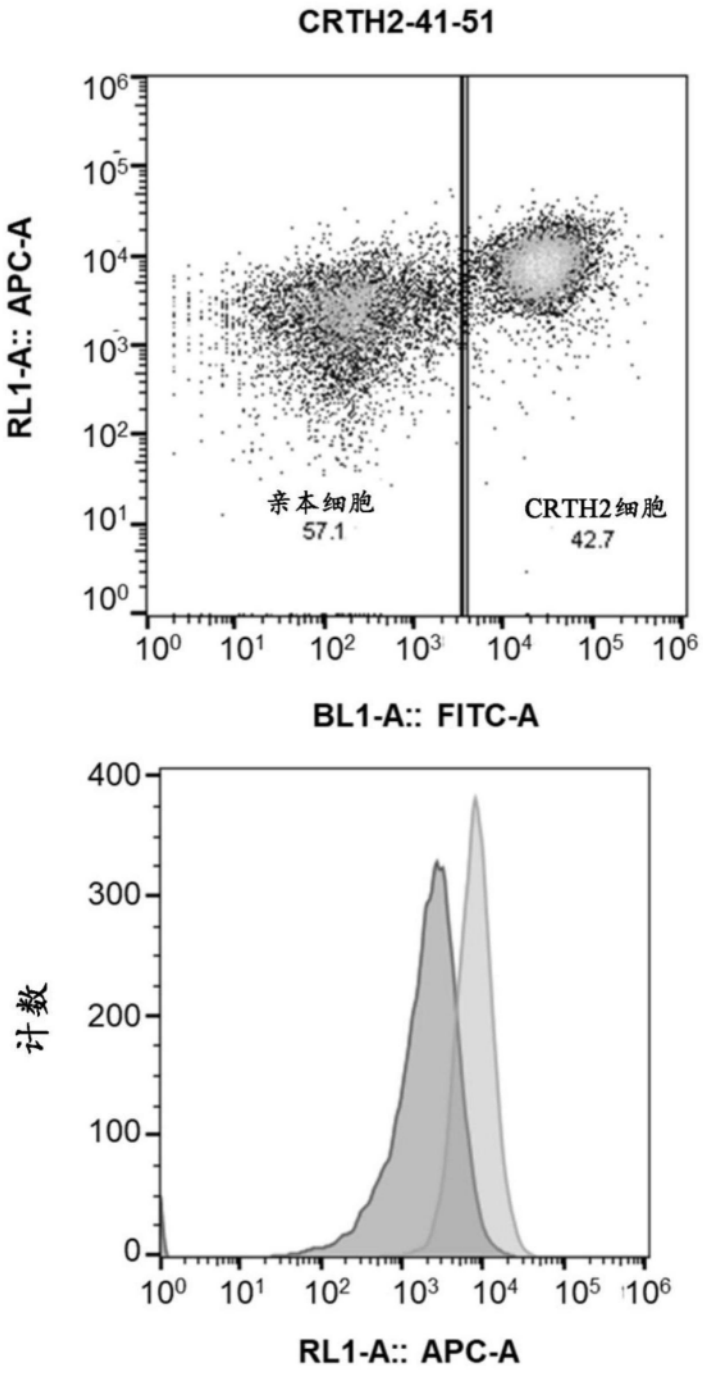
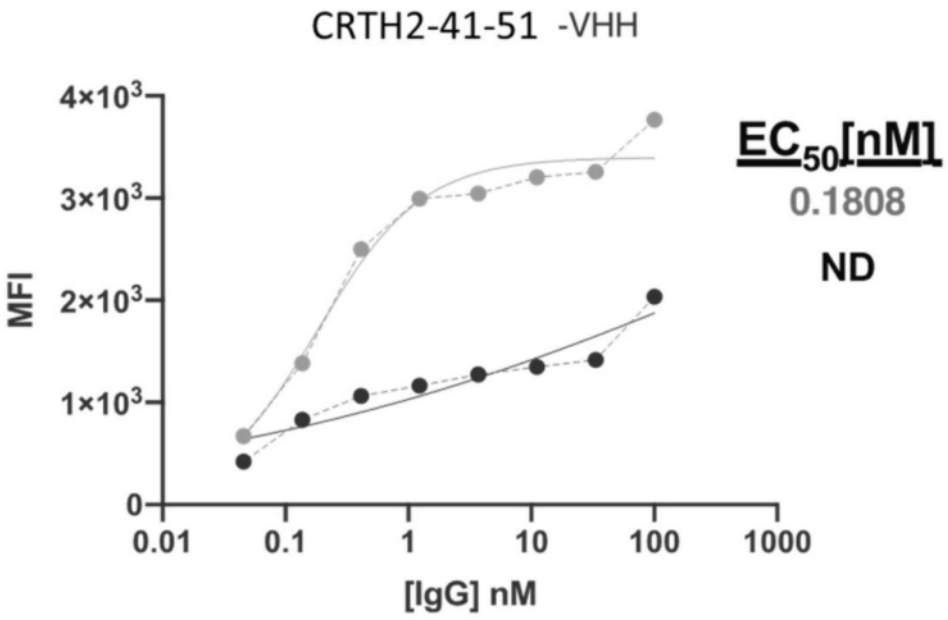

CRTH2-41-51
RL1-A:: APC-A
BL1-A:: FITC-A
亲本细胞
57.1
CRTH2细胞
42.7
计数
RL1-A:: APC-A

图17A


CRTH2-41-51 -VHH
MFI
4×10³
3×10³
2×10³
1×10³
0
0.01
0.1
1
10
100
1000
[IgG] nM
EC50[nM]
0.1808
ND

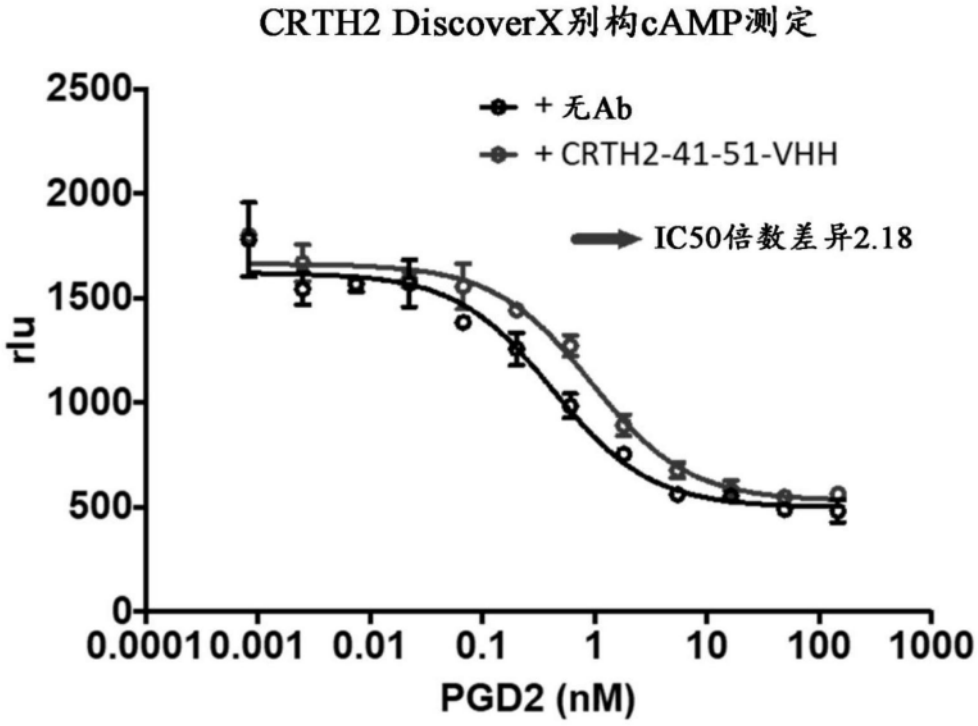

CRTH2 DiscoverX别构cAMP测定
+ 无Ab
+ CRTH2-41-51-VHH
IC50倍数差异2.18
rlu
2500
2000
1500
1000
500
0
0.0001
0.001
0.01
0.1
1
10
100
1000
PGD2 (nM)

图17B

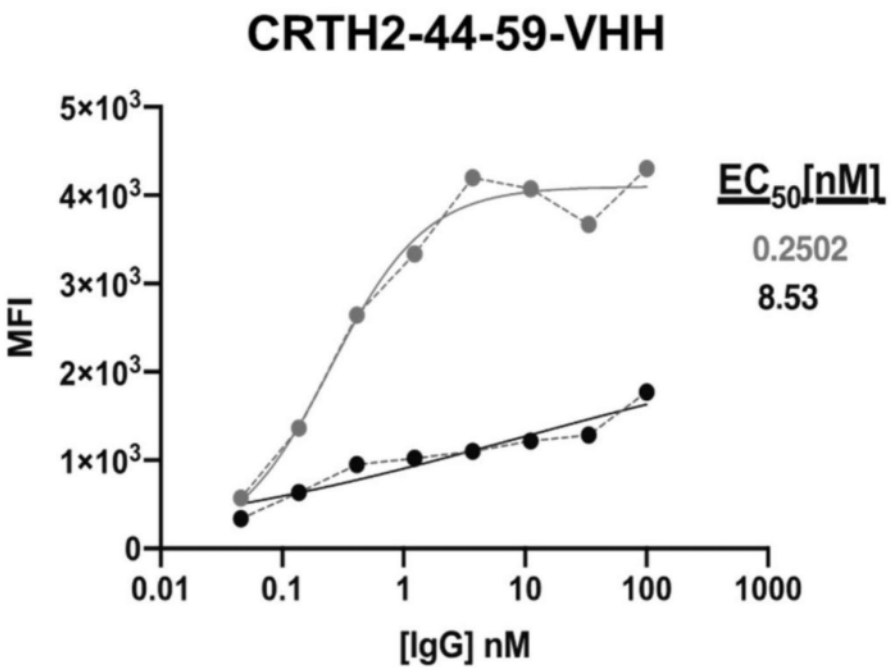

CRTH2-44-59-VHH
MFI
5×10³
4×10³
3×10³
2×10³
1×10³
0
0.01
0.1
1
10
100
1000
[IgG] nM
EC50[nM]
0.2502
8.53

图18

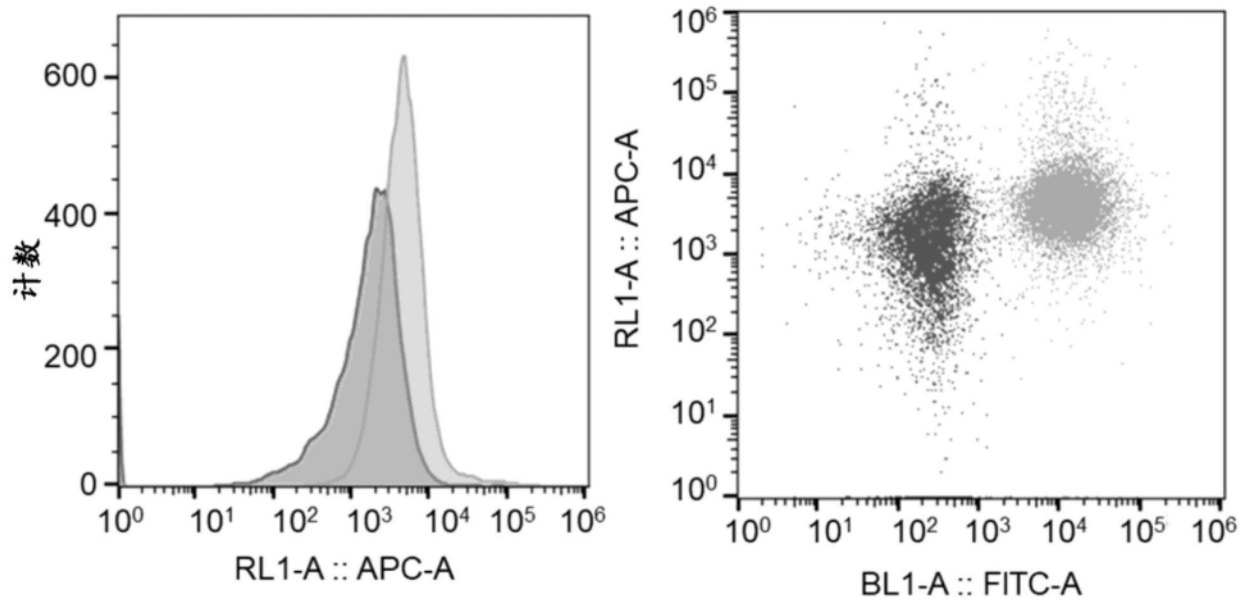

计数
600
400
200
0
10⁰ 10¹ 10² 10³ 10⁴ 10⁵ 10⁶
RL1-A :: APC-A
RL1-A :: APC-A
10⁶
10⁵
10⁴
10³
10²
10¹
10⁰
10⁰ 10¹ 10² 10³ 10⁴ 10⁵ 10⁶
BL1-A :: FITC-A

图19

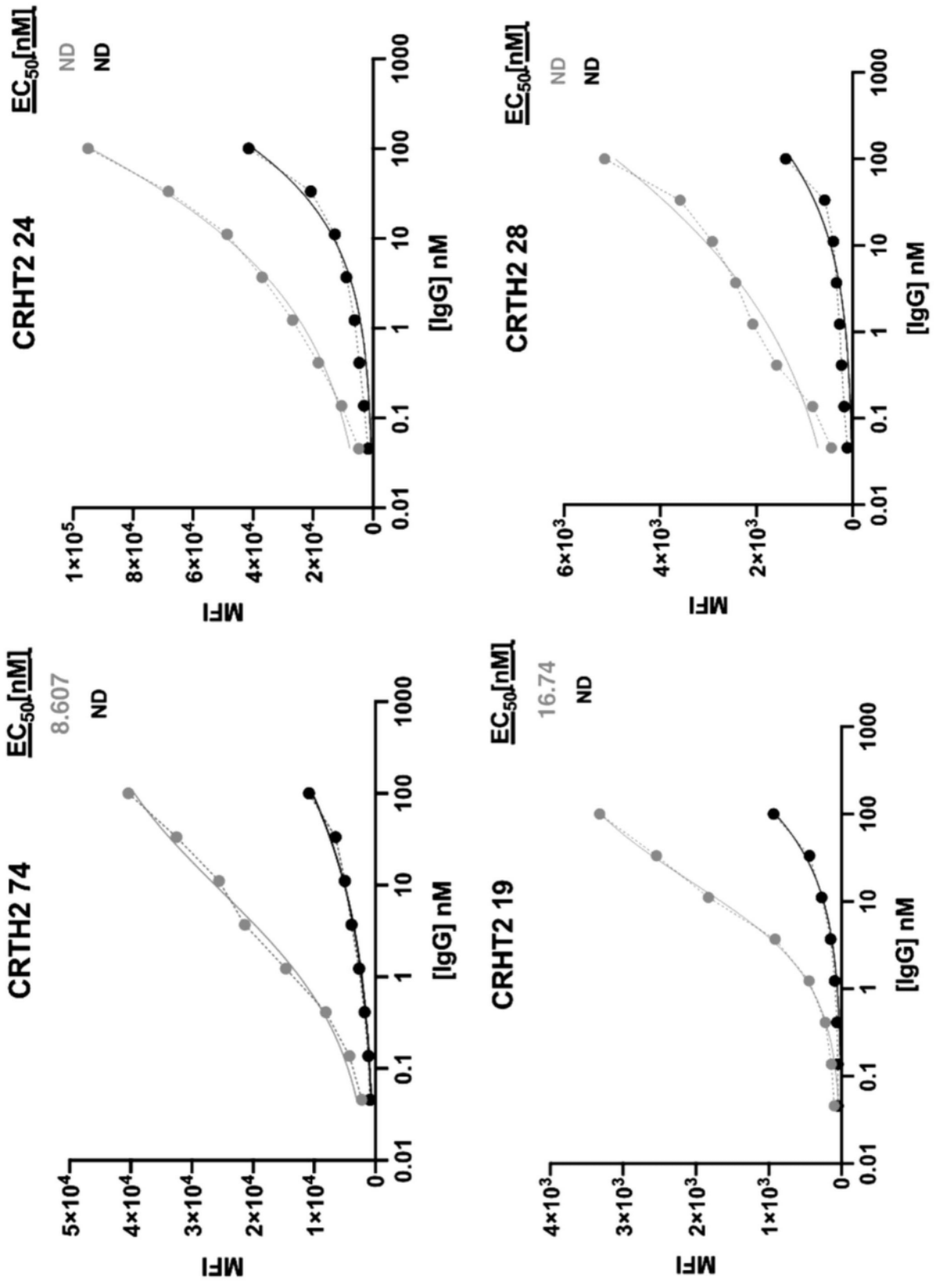

CRHT2 24
EC50[nM]
ND
ND
MFI
[IgG] nM
CRTH2 28
EC50[nM]
ND
ND
MFI
[IgG] nM
CRTH2 74
EC50[nM]
8.607
ND
MFI
[IgG] nM
CRHT2 19
EC50[nM]
16.74
ND
MFI
[IgG] nM

图20A

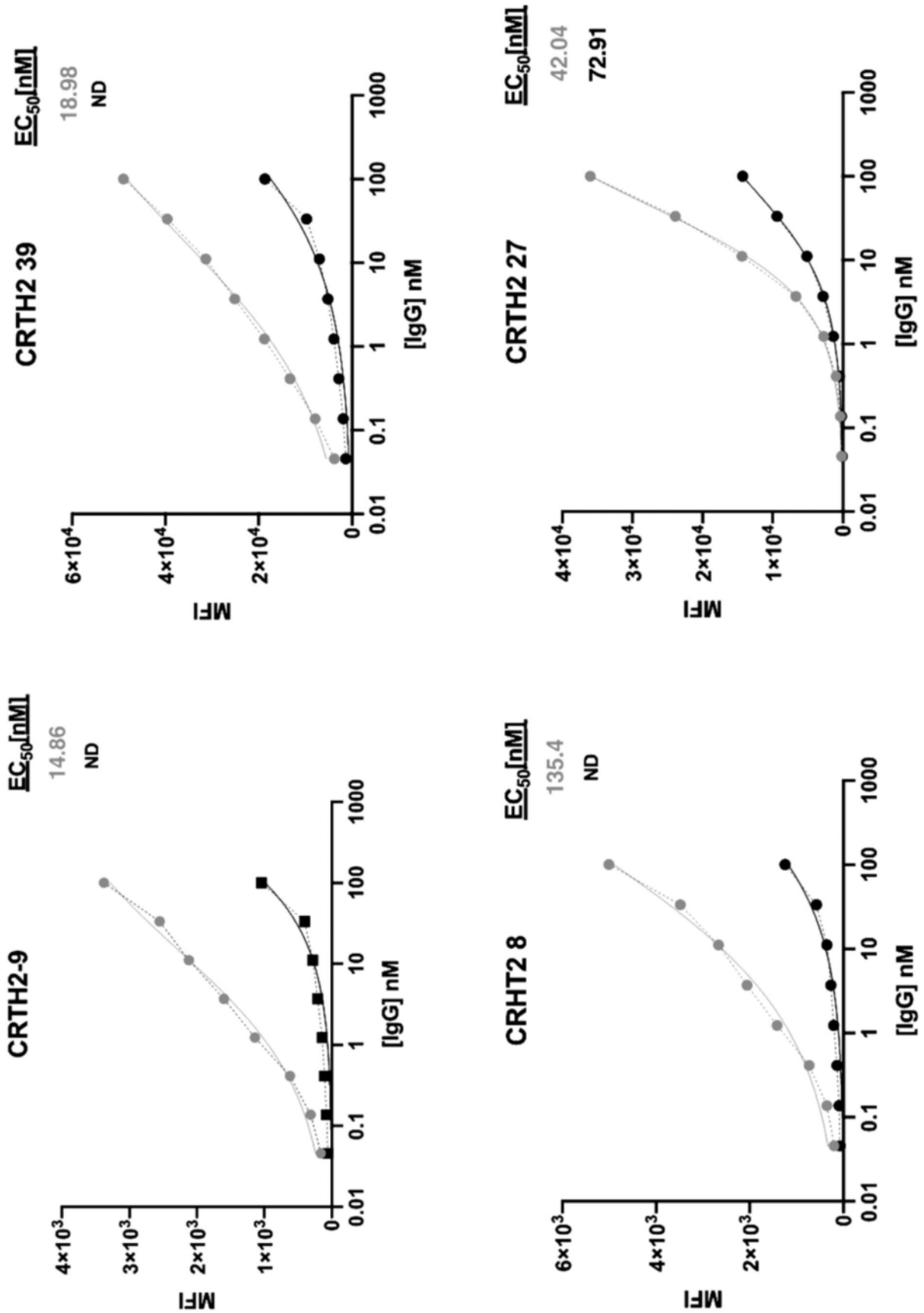

CRTH2 39
EC50[nM]
18.98
ND
CRTH2 27
42.04
72.91
CRTH2-9
14.86
ND
CRHT2 8
135.4
ND
MFI
[IgG] nM
0
0.01
0.1
1
10
100
1000
6×10^4
4×10^4
2×10^4
4×10^3
3×10^3
2×10^3
1×10^3
6×10^3

图20B

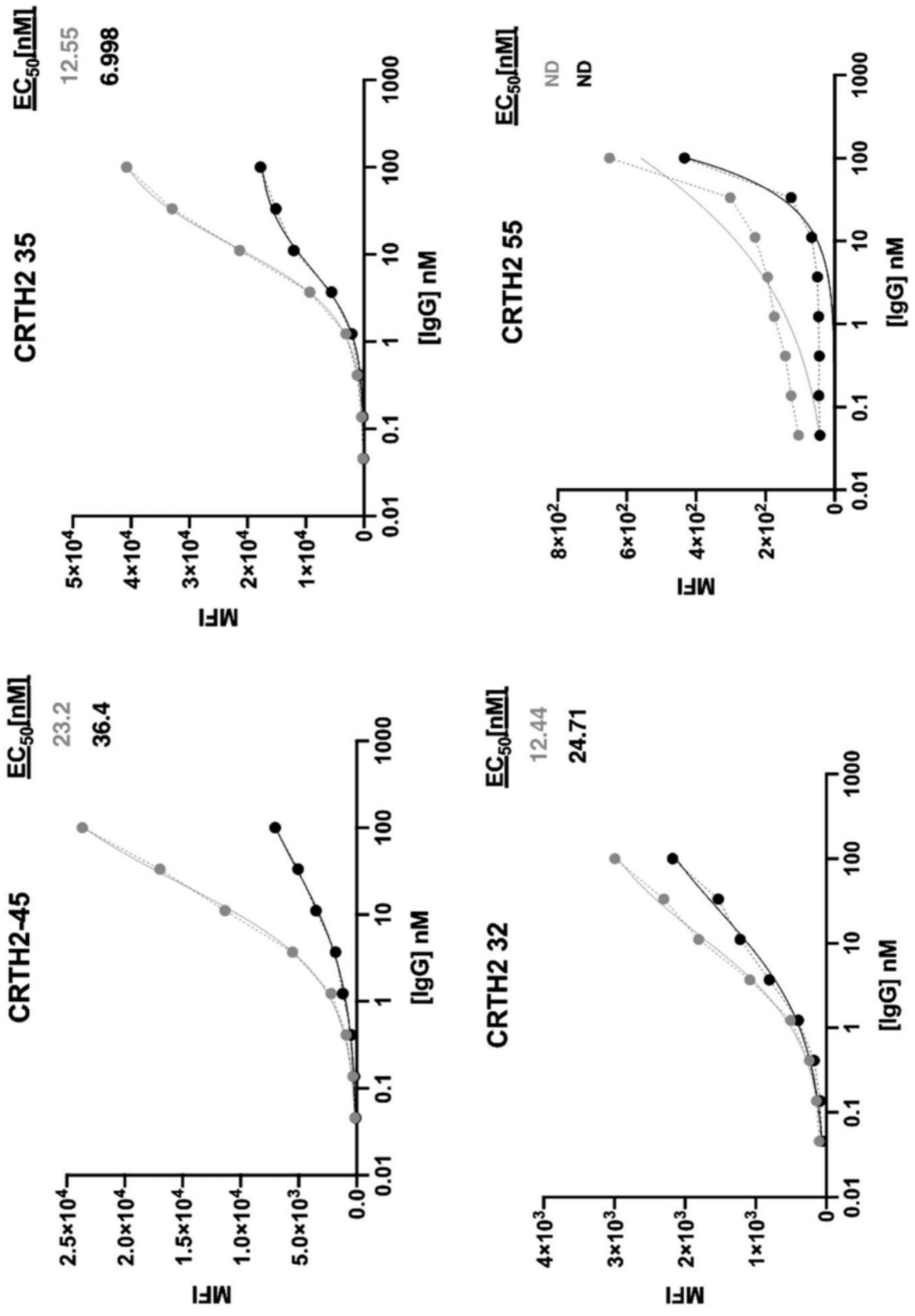

CRTH2 35
EC50[nM]
12.55
6.998
MFI
[IgG] nM
CRTH2 55
EC50[nM]
ND
ND
MFI
[IgG] nM
CRTH2-45
EC50[nM]
23.2
36.4
MFI
[IgG] nM
CRTH2 32
EC50[nM]
12.44
24.71
MFI
[IgG] nM

图20C

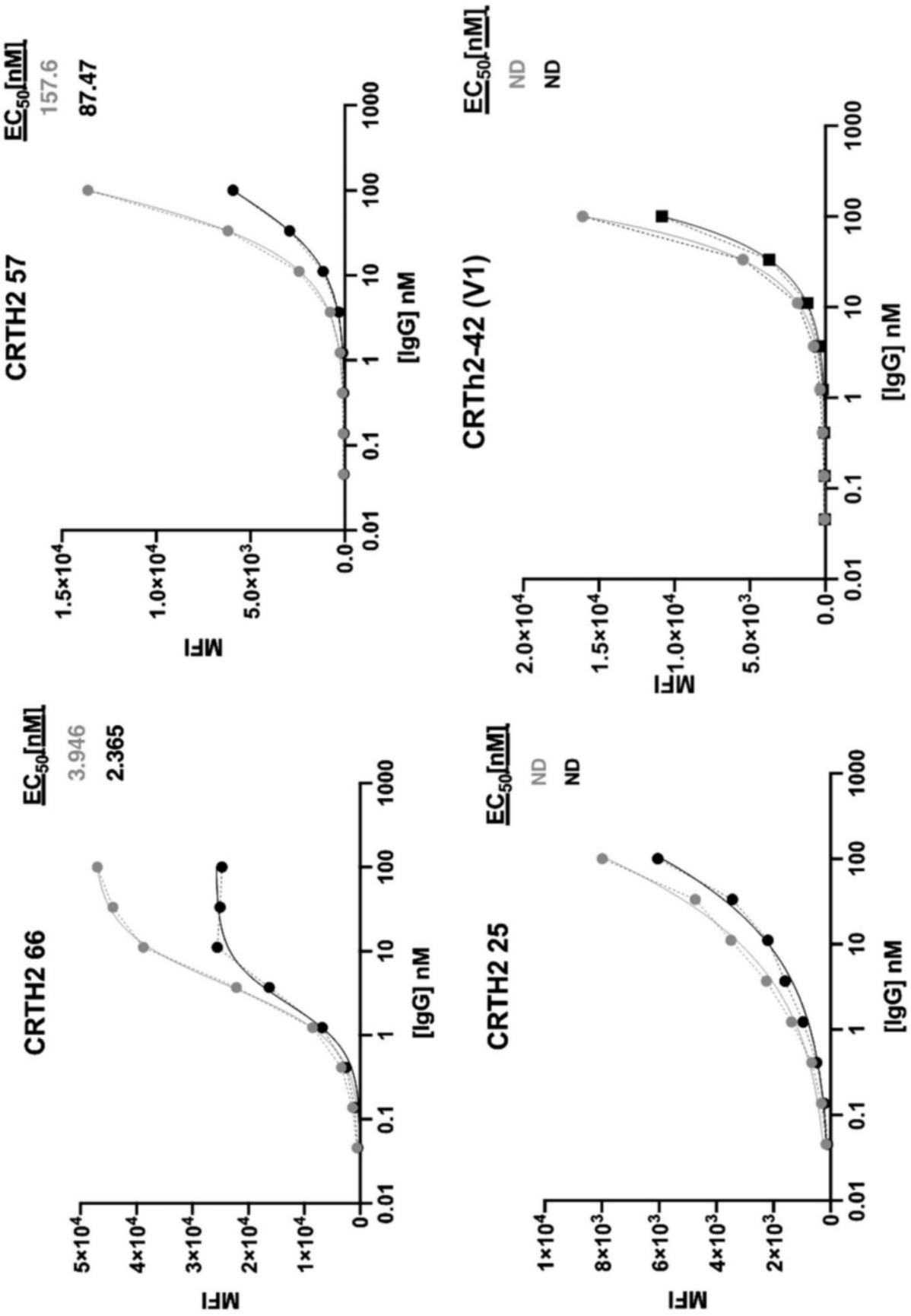

CRTH2 57
EC50[nM]
157.6
87.47
CRTh2-42 (V1)
EC50[nM]
ND
ND
CRTH2 66
EC50[nM]
3.946
2.365
CRTH2 25
EC50[nM]
ND
ND
MFI
[IgG] nM

图20D

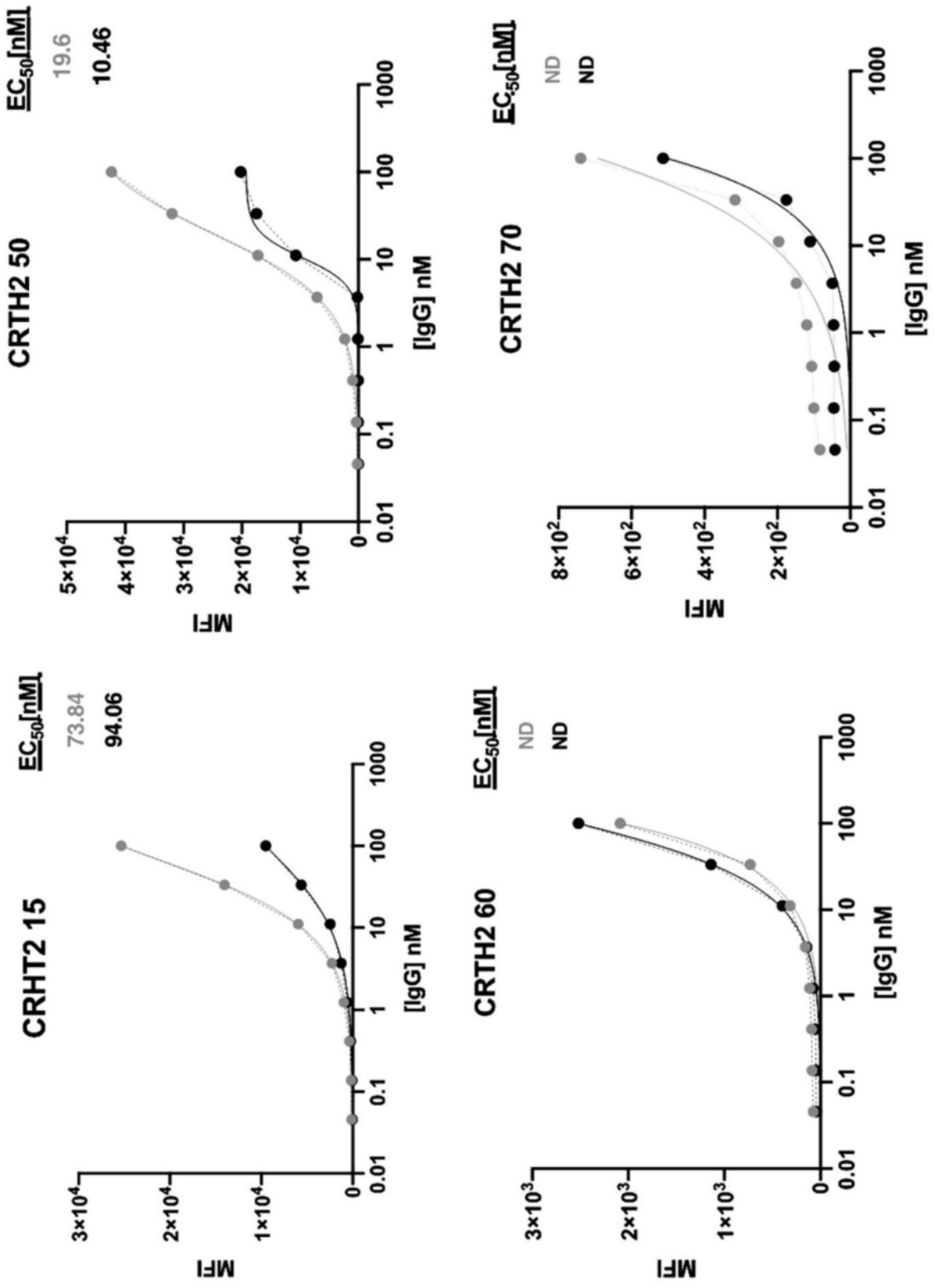

CRTH2 50
EC50[nM]
19.6
10.46
CRTH2 70
EC50[nM]
ND
ND
CRHT2 15
EC50[nM]
73.84
94.06
CRTH2 60
EC50[nM]
ND
ND
MFI
[IgG] nM

图20E

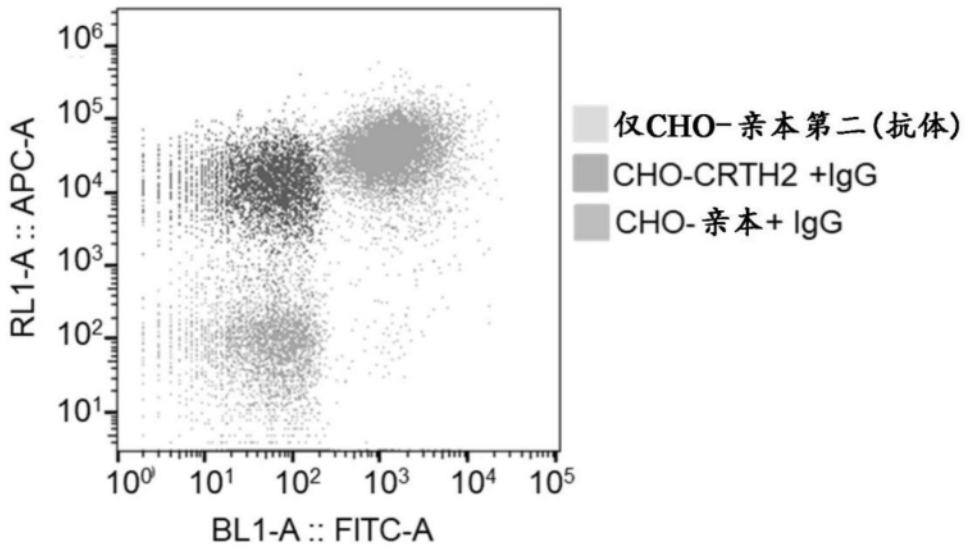

RL1-A :: APC-A
BL1-A :: FITC-A
仅CHO-亲本第二(抗体)
CHO-CRTH2 +IgG
CHO-亲本+ IgG

图21A

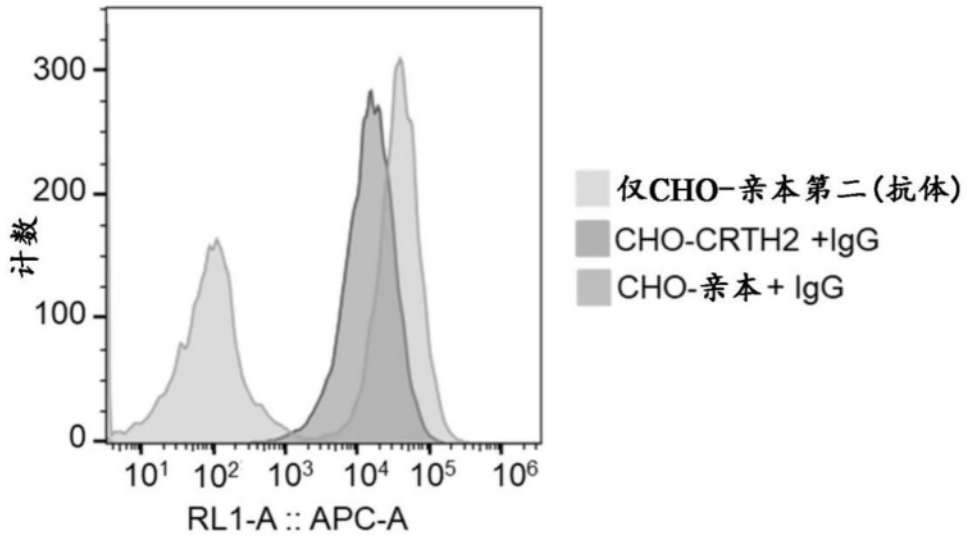

计数
RL1-A :: APC-A
仅CHO-亲本第二(抗体)
CHO-CRTH2 +IgG
CHO-亲本+ IgG

图21B

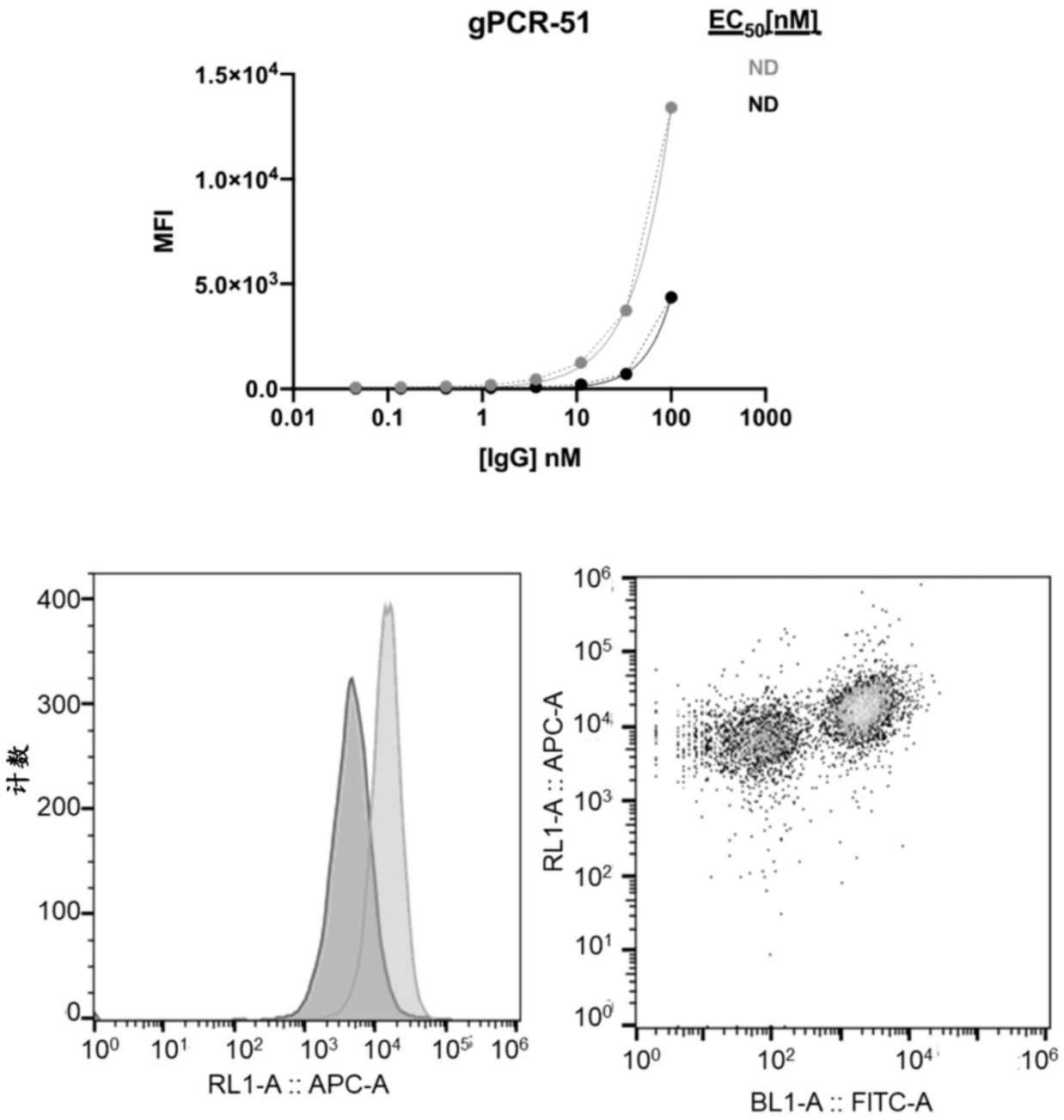

gPCR-51
EC50[nM]
ND
ND
1.5×10^4
1.0×10^4
5.0×10^3
0.0
MFI
0.01
0.1
1
10
100
1000
[IgG] nM
400
300
200
100
0
计数
10^0
10^1
10^2
10^3
10^4
10^5
10^6
RL1-A :: APC-A
RL1-A :: APC-A
BL1-A :: FITC-A

图22A

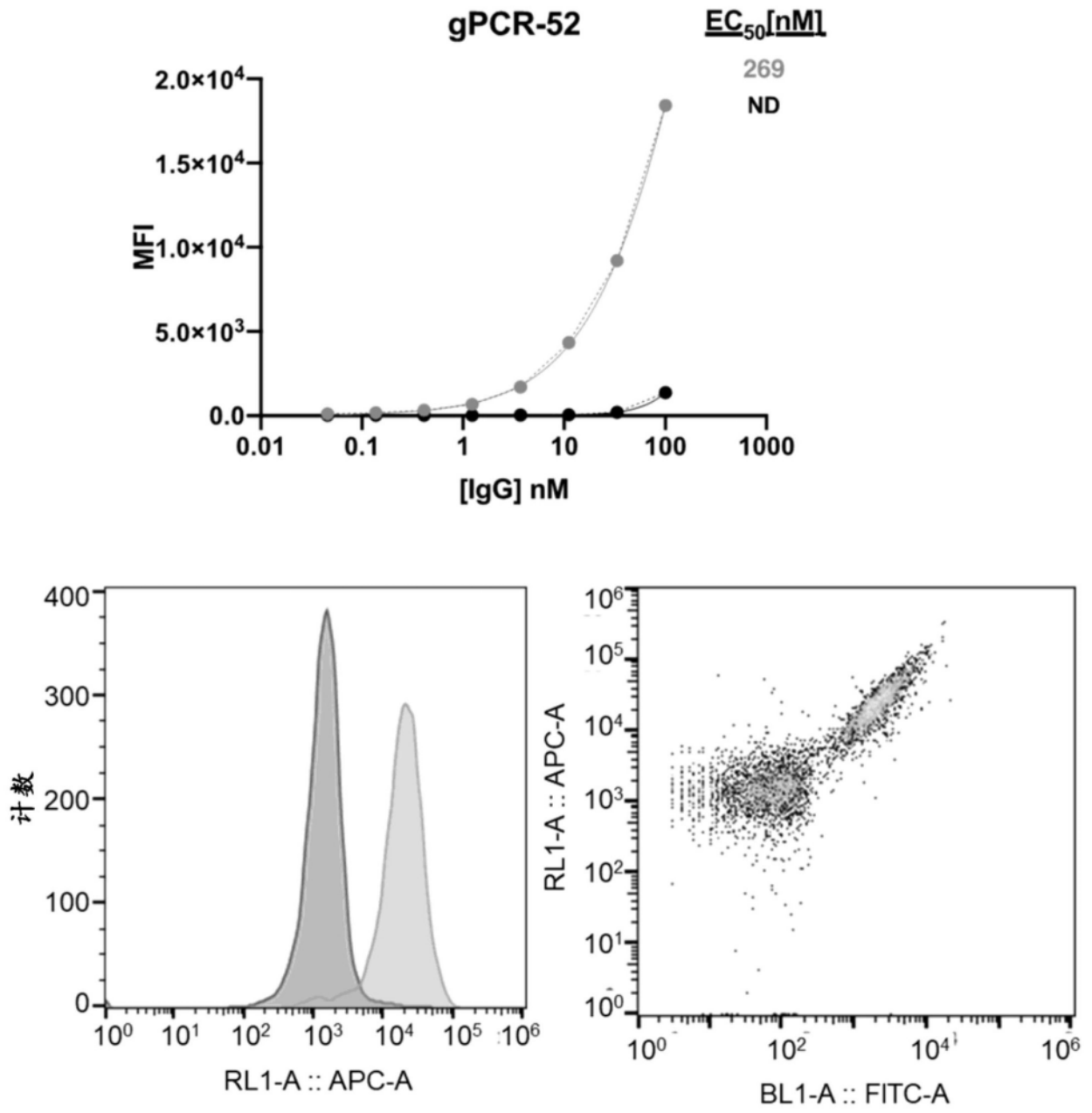


图22B

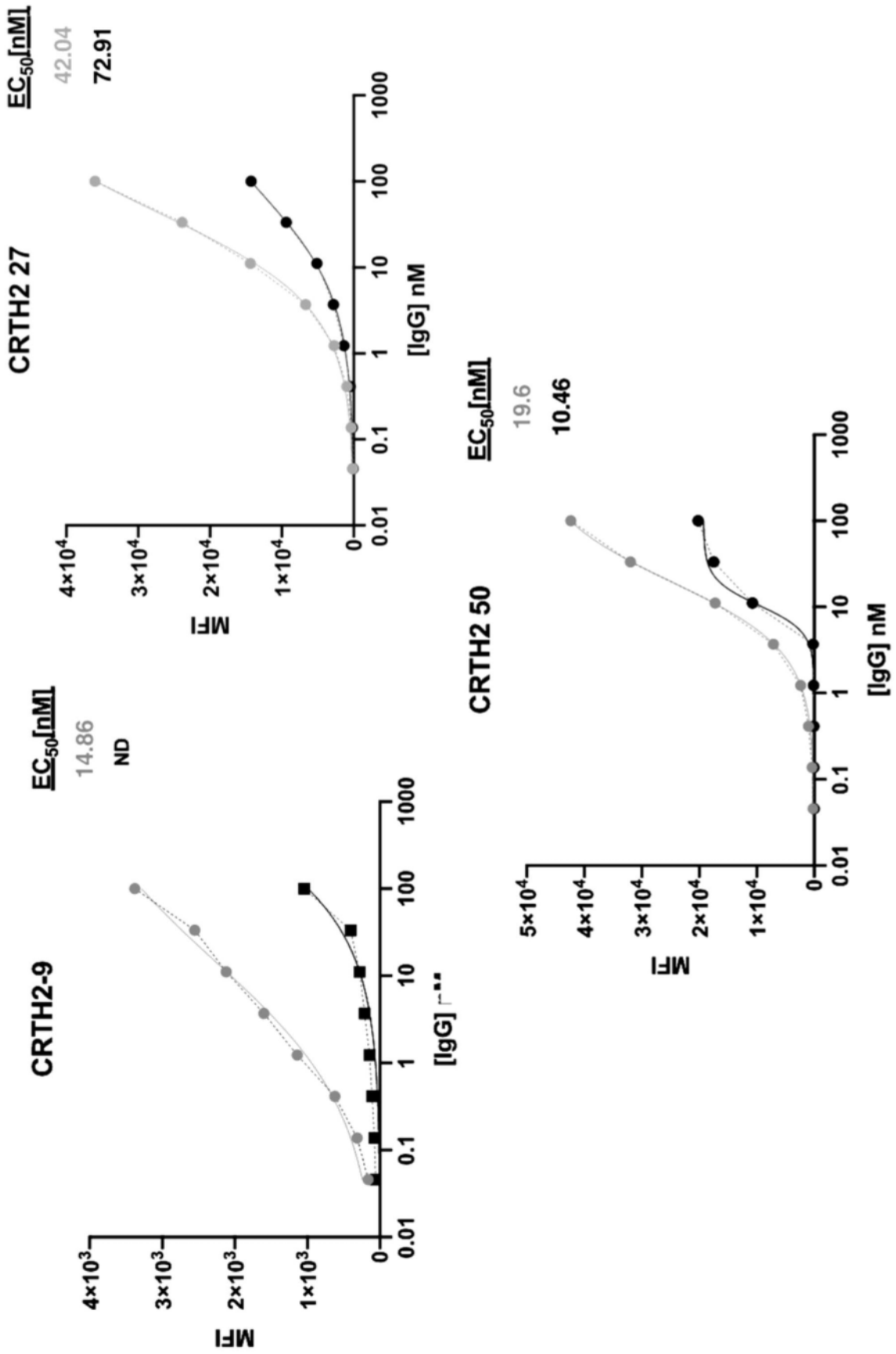

CRTH2 27
EC50[nM]
42.04
72.91
MFI
[IgG] nM
CRTH2-9
EC50[nM]
14.86
ND
MFI
[IgG]
CRTH2 50
EC50[nM]
19.6
10.46
MFI
[IgG] nM

图23A

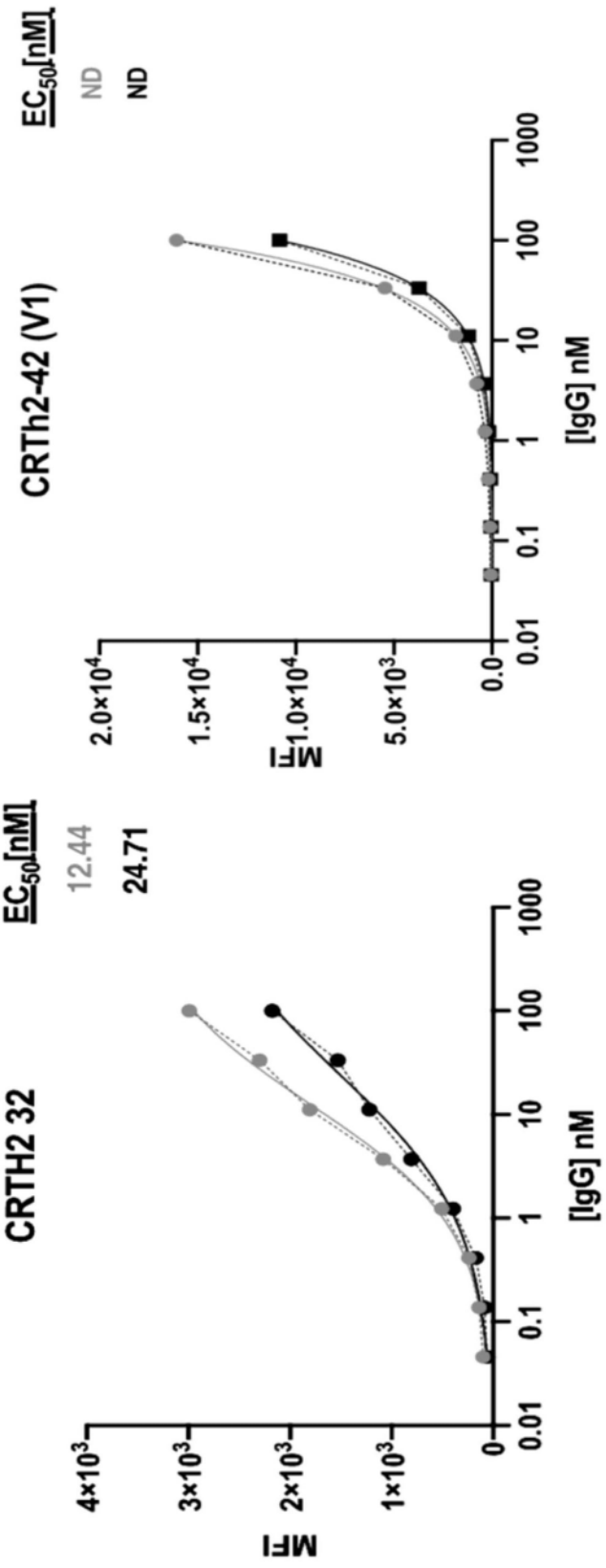

CRTh2-42 (V1)
EC50[nM]
ND
ND
MFI
2.0×10^4
1.5×10^4
1.0×10^4
5.0×10^3
0.0
0.01
0.1
1
10
100
1000
[IgG] nM
CRTH2 32
EC50[nM]
12.44
24.71
MFI
4×10^3
3×10^3
2×10^3
1×10^3
0
0.01
0.1
1
10
100
1000
[IgG] nM

图23B

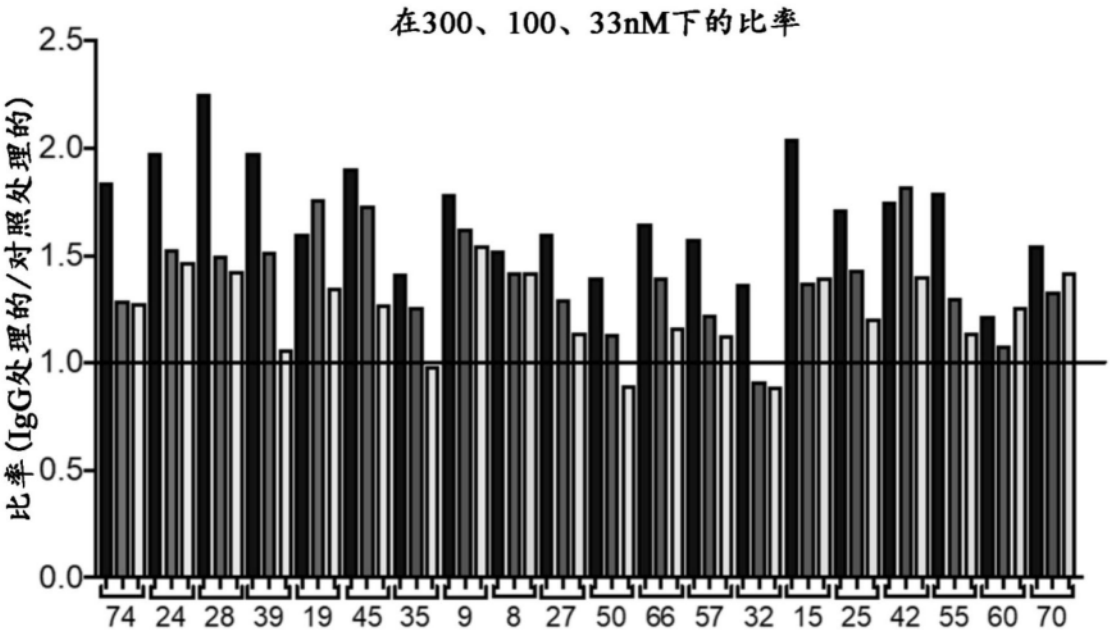

在300、100、33nM下的比率
比率(IgG处理的/对照处理的)
2.5
2.0
1.5
1.0
0.5
0.0
74 24 28 39 19 45 35 9 8 27 50 66 57 32 15 25 42 55 60 70

图24A

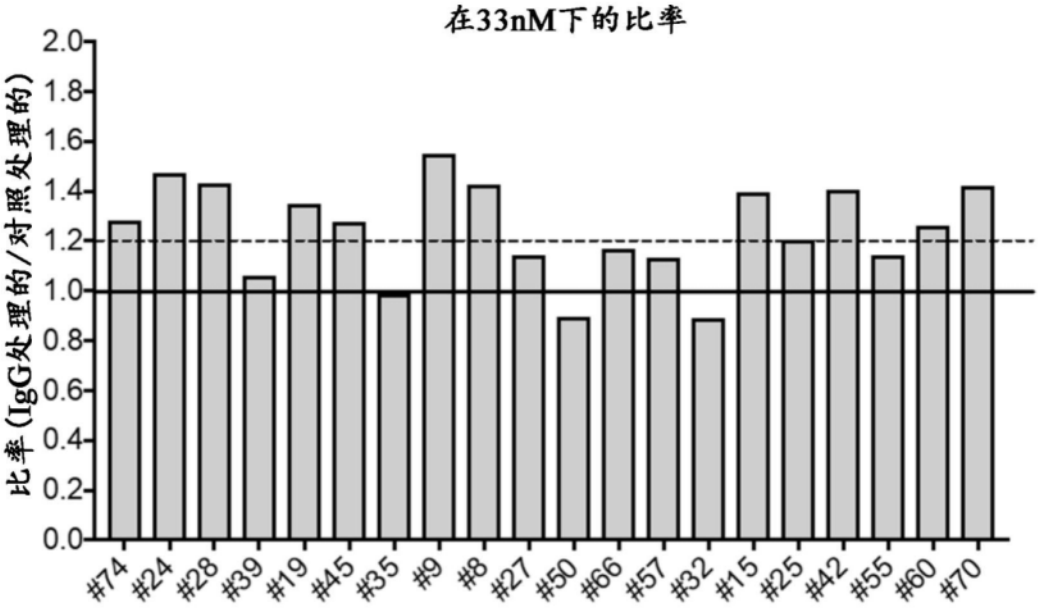
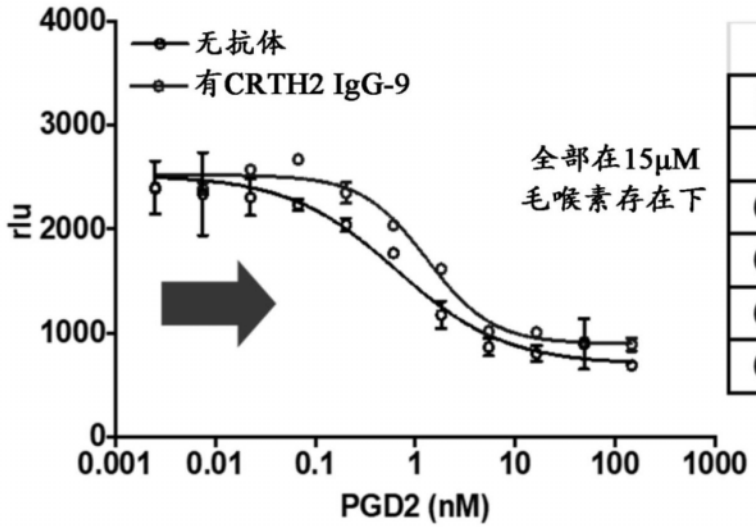

在33nM下的比率
比率(IgG处理的/对照处理的)
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
#74 #24 #28 #39 #19 #45 #35 #9 #8 #27 #50 #66 #57 #32 #15 #25 #42 #55 #60 #70

图24B



	IC50倍数差异
PGD2单独	-
CRTH2-9	2.03
CRTH2-27	2.02
CRTH2-50	2.01
CRTH2-32	4.28
CRTH2-42	4.99

图25

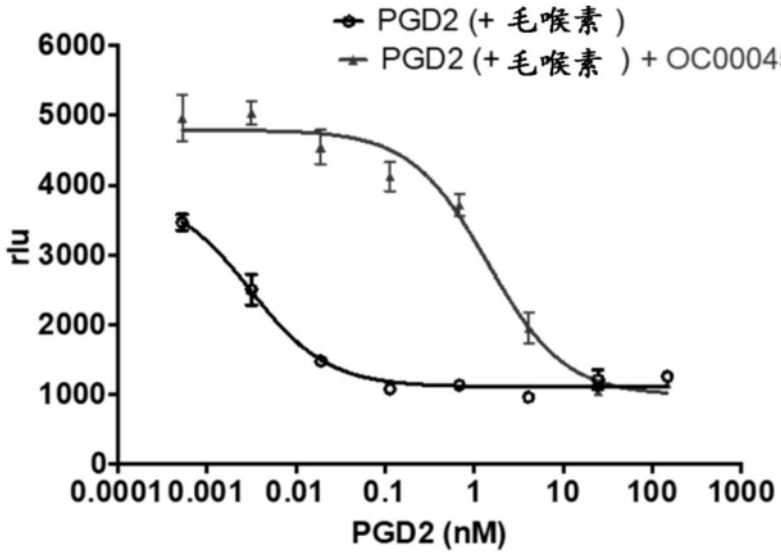


图26A

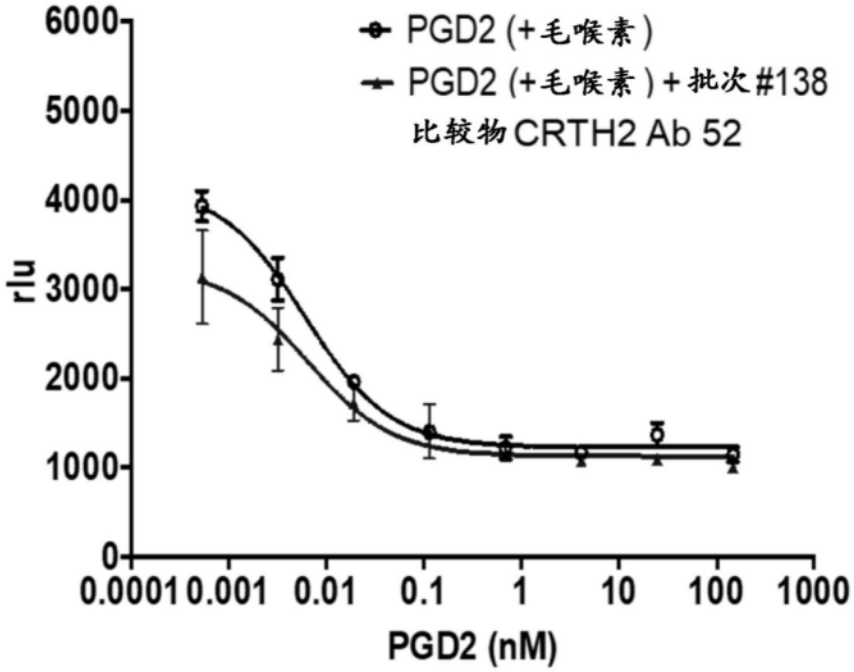

PGD2 (+毛喉素)
PGD2 (+毛喉素) + 批次#138
比较物CRTH2 Ab 52
rlu
6000
5000
4000
3000
2000
1000
0
0.0001 0.001 0.01 0.1 1 10 100 1000
PGD2 (nM)

图26B

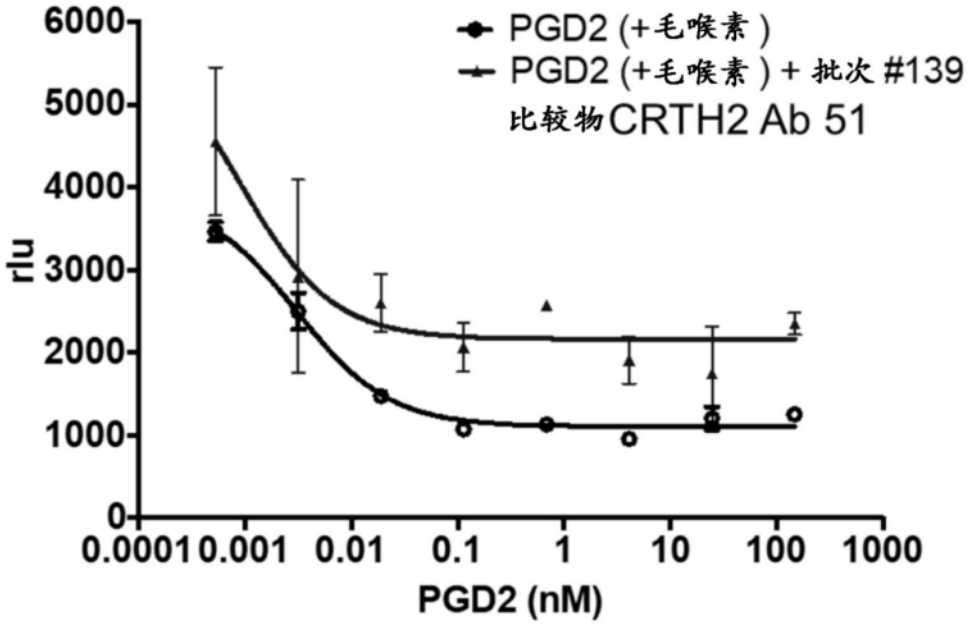

PGD2 (+毛喉素)
PGD2 (+毛喉素) + 批次 #139
比较物CRTH2 Ab 51
rlu
6000
5000
4000
3000
2000
1000
0
0.0001 0.001 0.01 0.1 1 10 100 1000
PGD2 (nM)

图26C

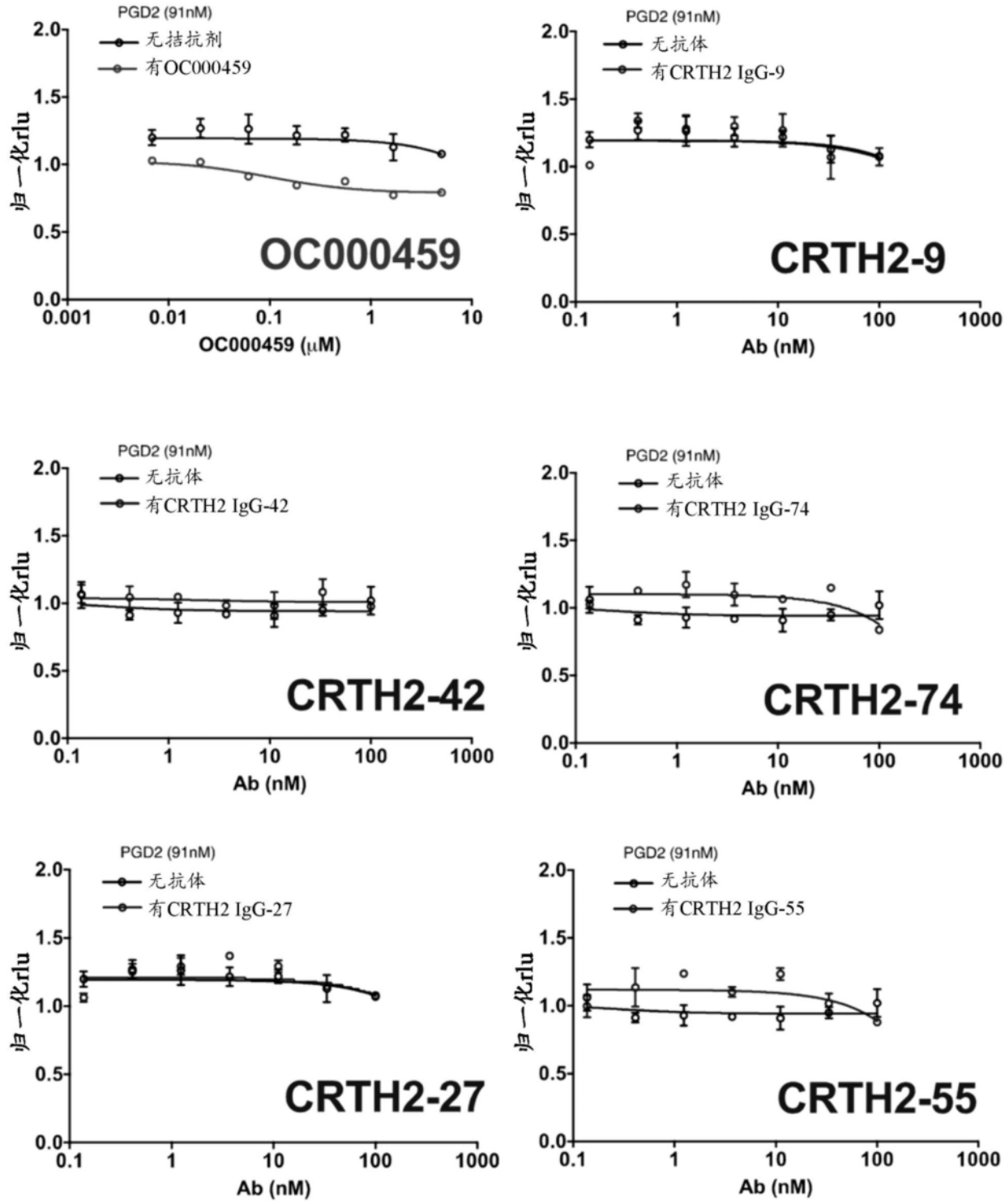

PGD2 (91nM)
无拮抗剂
有OC000459
OC000459
归一化rlu
OC000459 (μM)
PGD2 (91nM)
无抗体
有CRTH2 IgG-9
CRTH2-9
Ab (nM)
PGD2 (91nM)
无抗体
有CRTH2 IgG-42
CRTH2-42
Ab (nM)
PGD2 (91nM)
无抗体
有CRTH2 IgG-74
CRTH2-74
Ab (nM)
PGD2 (91nM)
无抗体
有CRTH2 IgG-27
CRTH2-27
Ab (nM)
PGD2 (91nM)
无抗体
有CRTH2 IgG-55
CRTH2-55
Ab (nM)

图27A

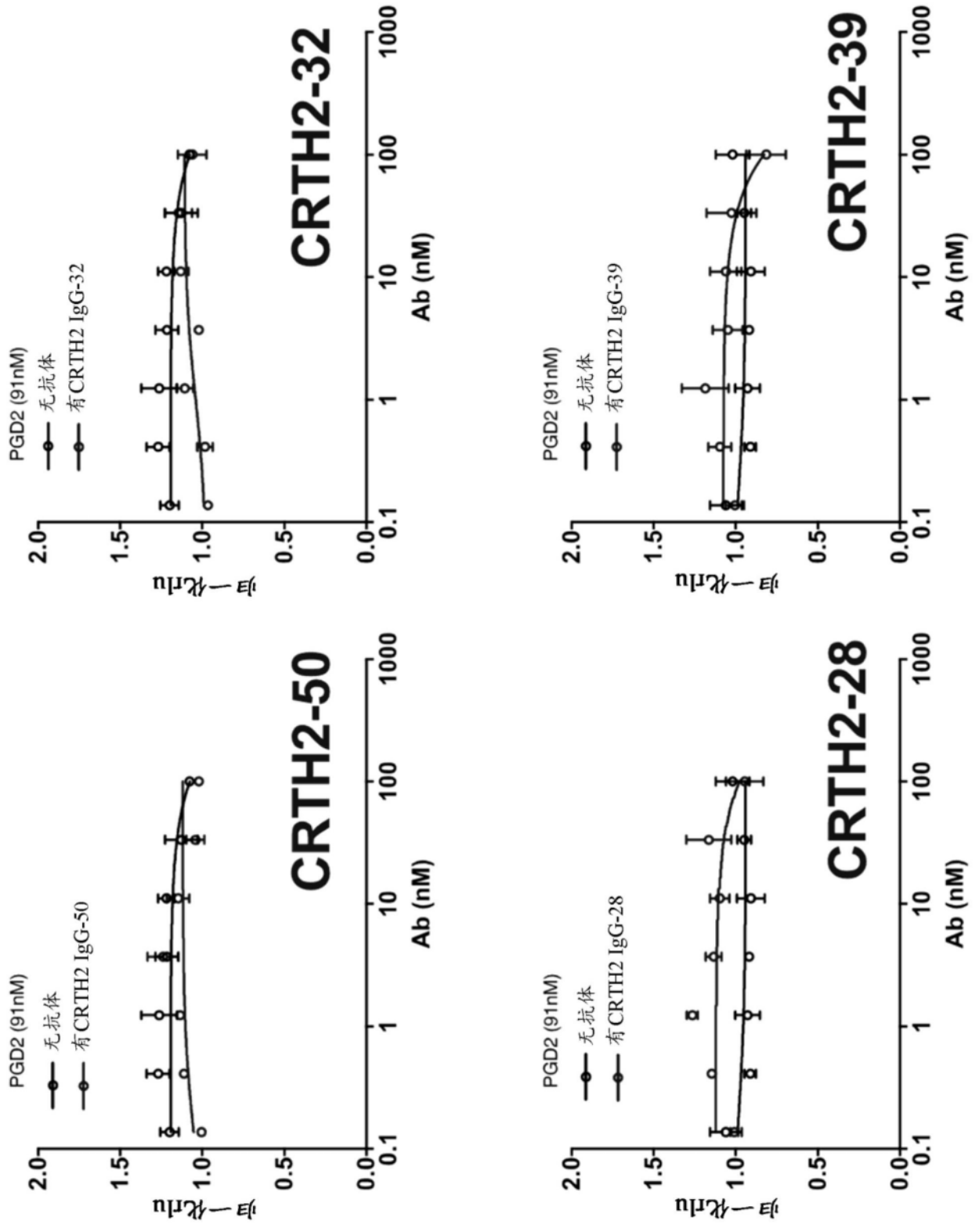

CRTH2-32
CRTH2-39
CRTH2-50
CRTH2-28
PGD2 (91nM)
无抗体
有CRTH2 IgG-32
有CRTH2 IgG-39
有CRTH2 IgG-50
有CRTH2 IgG-28
Ab (nM)
0.0
0.5
1.0
1.5
2.0
0.1
1
10
100
1000

图27B

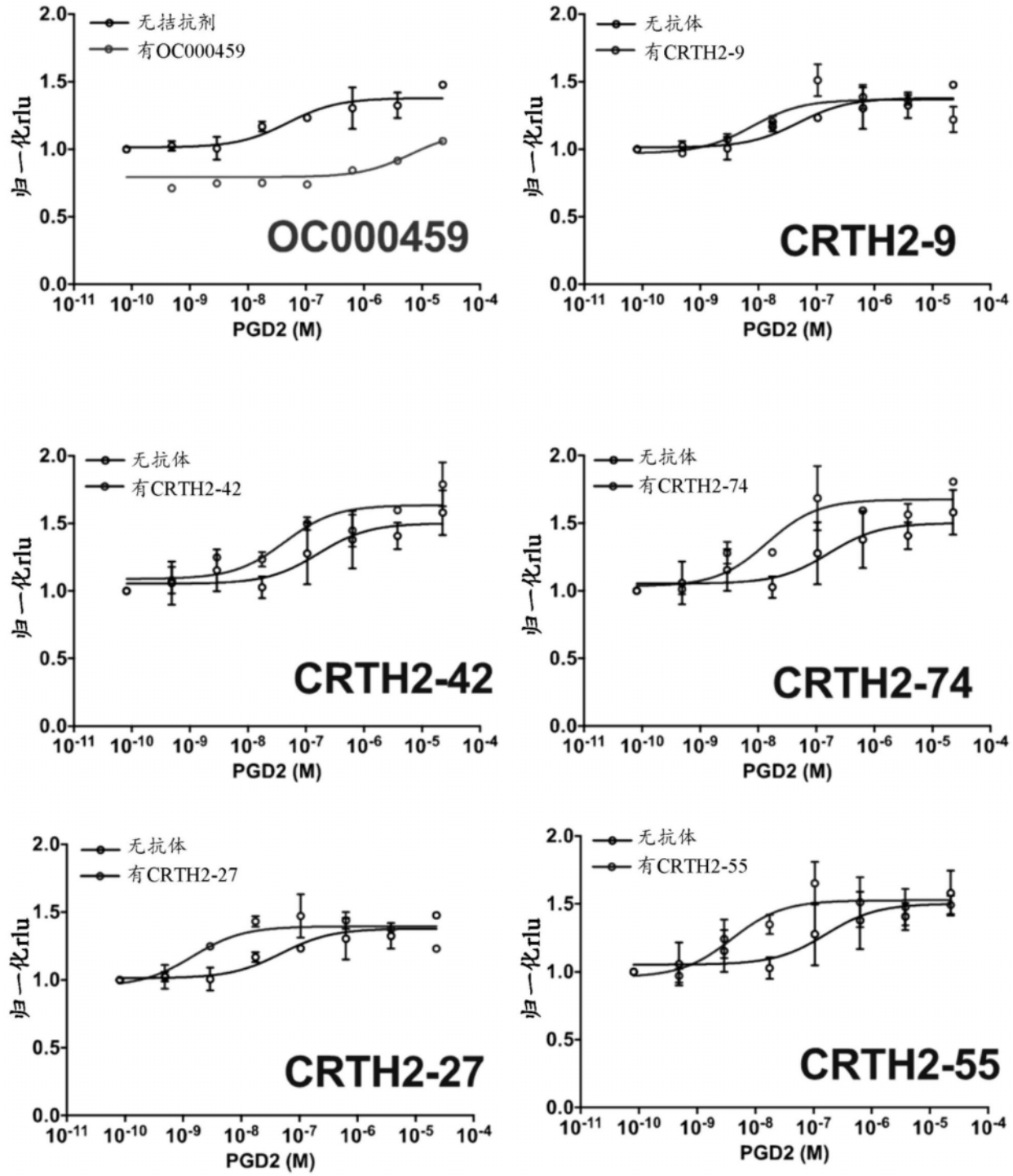

无拮抗剂
有OC000459
OC000459
无抗体
有CRTH2-9
CRTH2-9
无抗体
有CRTH2-42
CRTH2-42
无抗体
有CRTH2-74
CRTH2-74
无抗体
有CRTH2-27
CRTH2-27
无抗体
有CRTH2-55
CRTH2-55
归一化rlu
PGD2 (M)

图27C

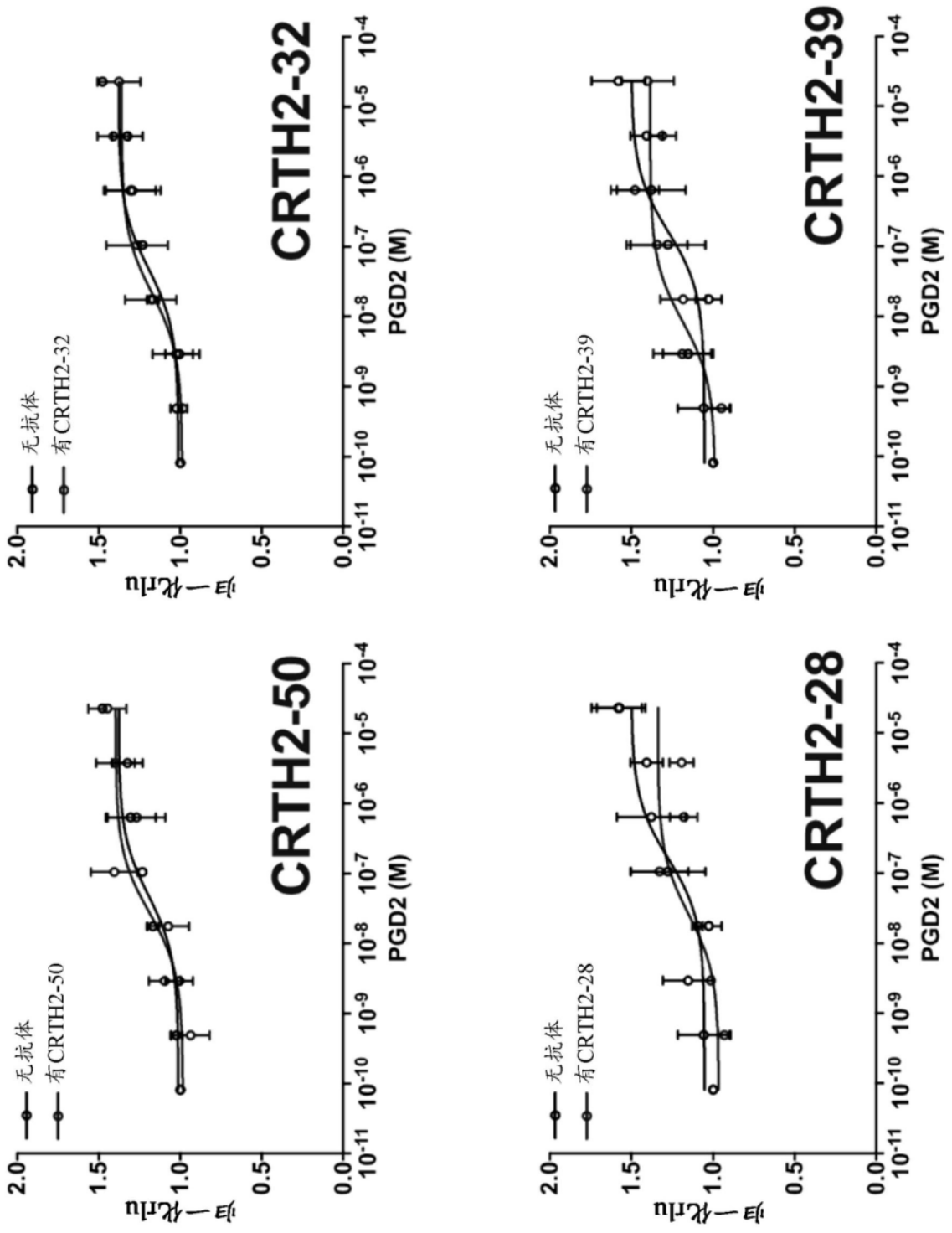
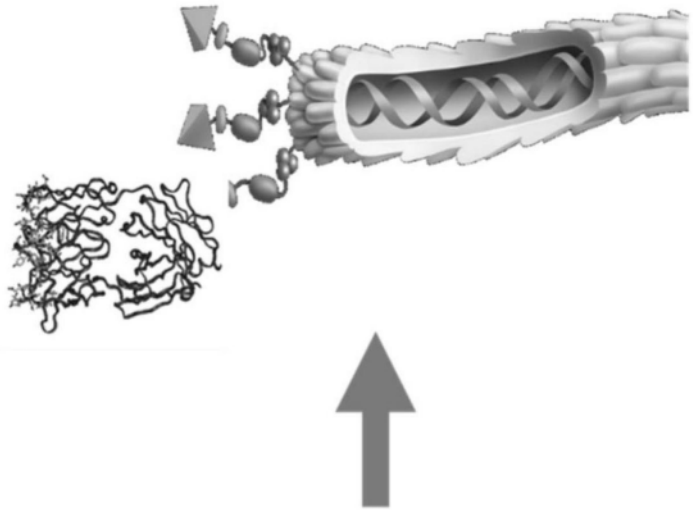

CRTH2-32
CRTH2-39
CRTH2-50
CRTH2-28
无抗体
有CRTH2-32
有CRTH2-39
有CRTH2-50
有CRTH2-28
PGD2 (M)
2.0
1.5
1.0
0.5
0.0
10-11
10-10
10-9
10-8
10-7
10-6
10-5
10-4

图27D



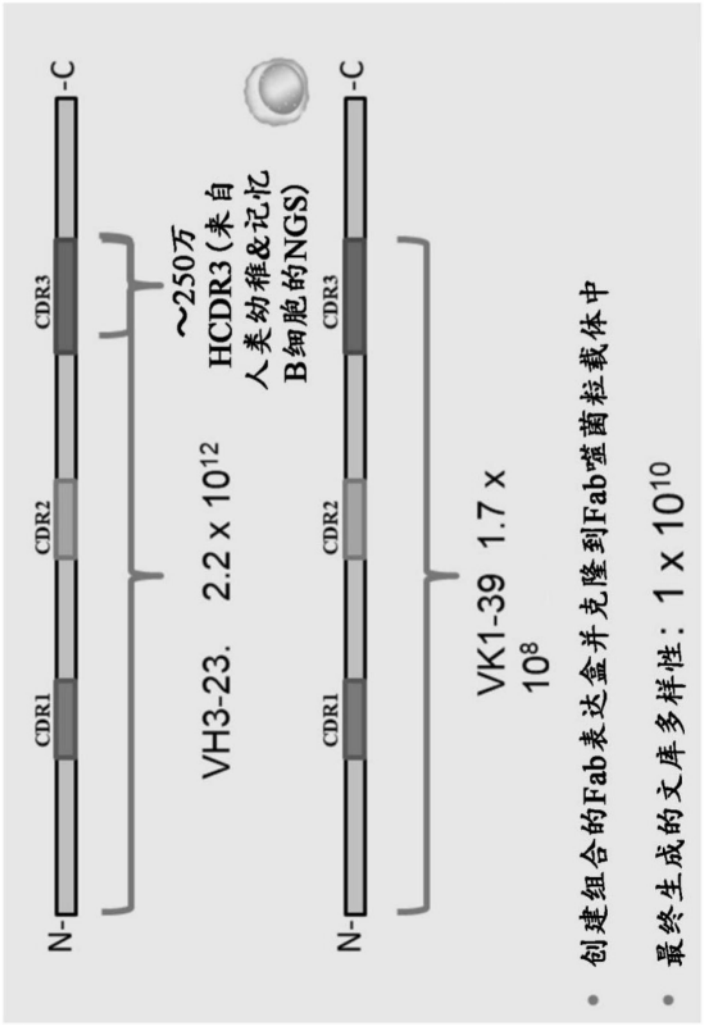

N-
CDR1
CDR2
CDR3
-C
VH3-23. 2.2 x 10^12
~250万
HCDR3(来自
人类幼稚&记忆
B细胞的NGS)
N-
CDR1
CDR2
CDR3
-C
VK1-39 1.7 x
10^8
• 创建组合的Fab表达盒并克隆到Fab噬菌粒载体中
• 最终生成的文库多样性：1 x 10^10

图28

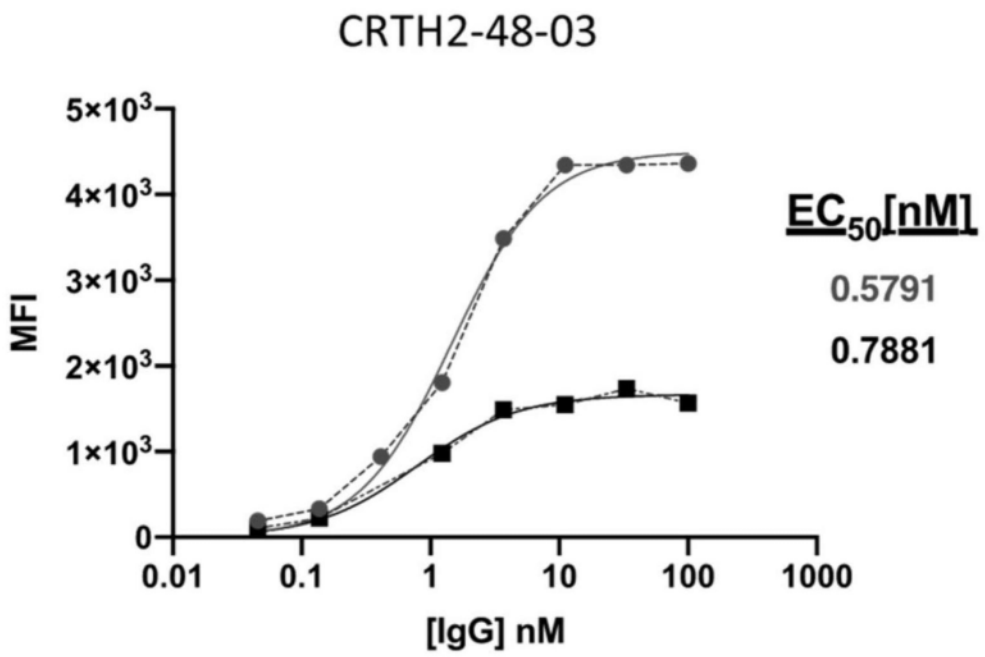

CRTH2-48-03
5×10³
4×10³
3×10³
2×10³
1×10³
0
MFI
0.01
0.1
1
10
100
1000
[IgG] nM
EC50[nM]
0.5791
0.7881

图29A

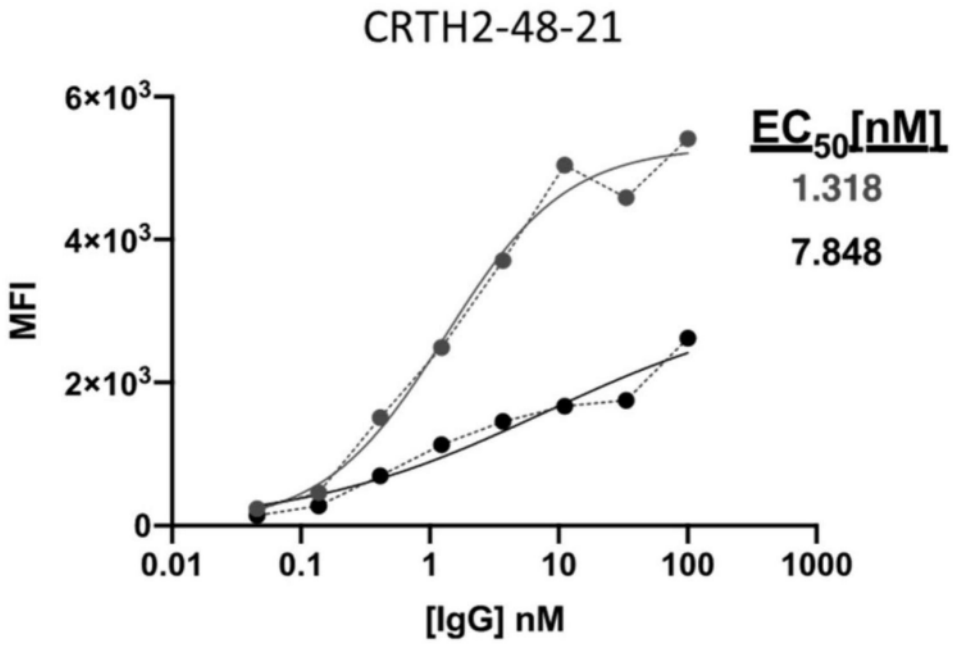

CRTH2-48-21
6×10³
4×10³
2×10³
0
MFI
0.01
0.1
1
10
100
1000
[IgG] nM
EC50[nM]
1.318
7.848

图29B

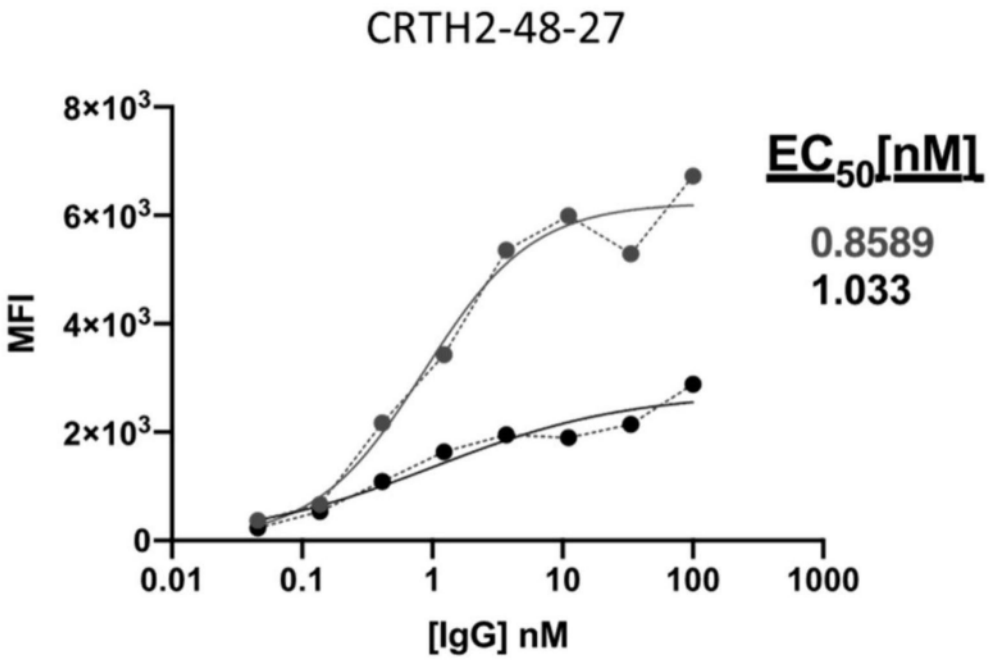

CRTH2-48-27
EC50[nM]
0.8589
1.033
MFI
8×10³
6×10³
4×10³
2×10³
0
0.01
0.1
1
10
100
1000
[IgG] nM

图29C

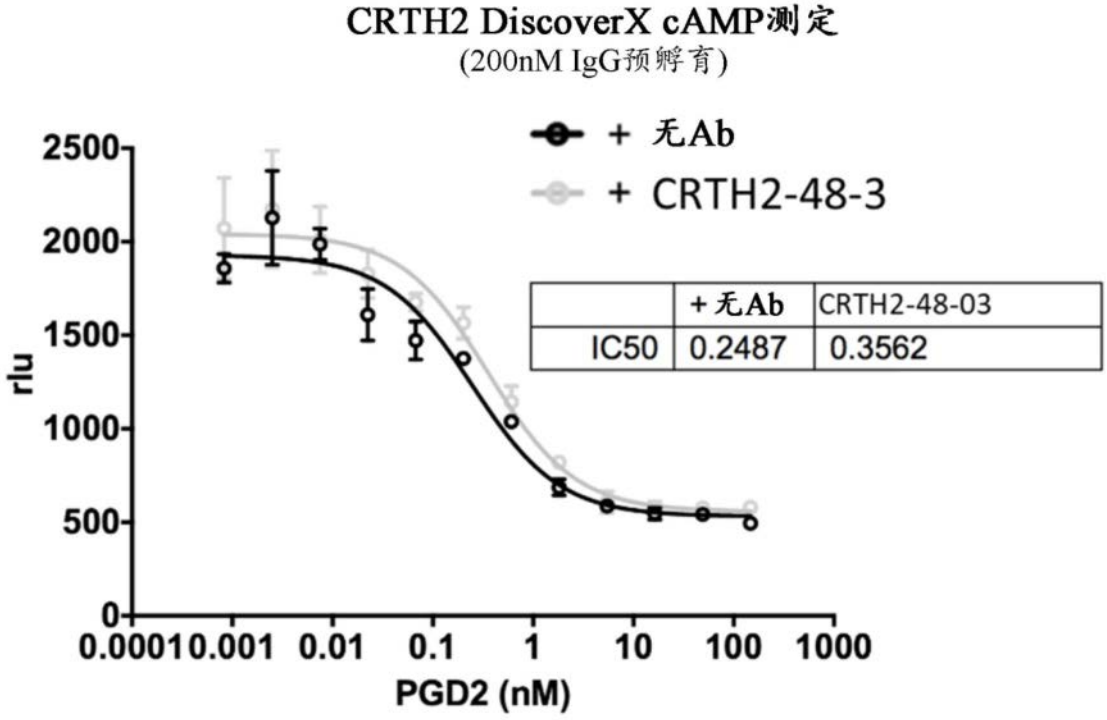

CRTH2 DiscoverX cAMP测定
(200nM IgG预孵育)
+ 无Ab
+ CRTH2-48-3
+无Ab
CRTH2-48-03
IC50
0.2487
0.3562
rlu
2500
2000
1500
1000
500
0
0.0001
0.001
0.01
0.1
1
10
100
1000
PGD2 (nM)

图29D

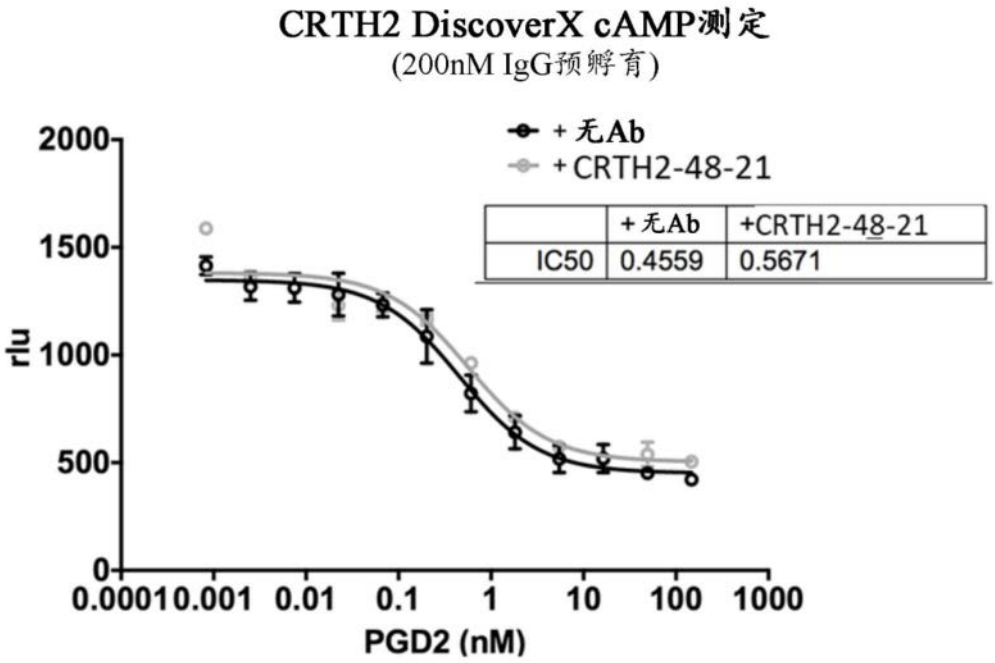

CRTH2 DiscoverX cAMP测定
(200nM IgG预孵育)
+无Ab
+CRTH2-48-21
+无Ab
+CRTH2-48-21
IC50
0.4559
0.5671
rlu
2000
1500
1000
500
0
0.0001
0.001
0.01
0.1
1
10
100
1000
PGD2 (nM)

图29E

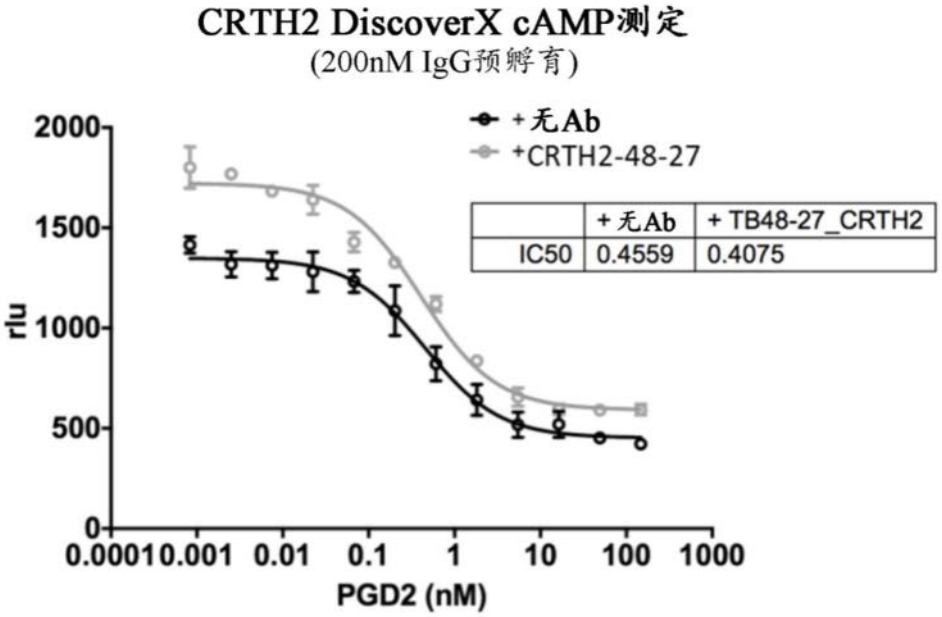

CRTH2 DiscoverX cAMP测定
(200nM IgG预孵育)
+无Ab
+CRTH2-48-27
+无Ab
+ TB48-27_CRTH2
IC50
0.4559
0.4075
rlu
2000
1500
1000
500
0
0.0001
0.001
0.01
0.1
1
10
100
1000
PGD2 (nM)

图29F

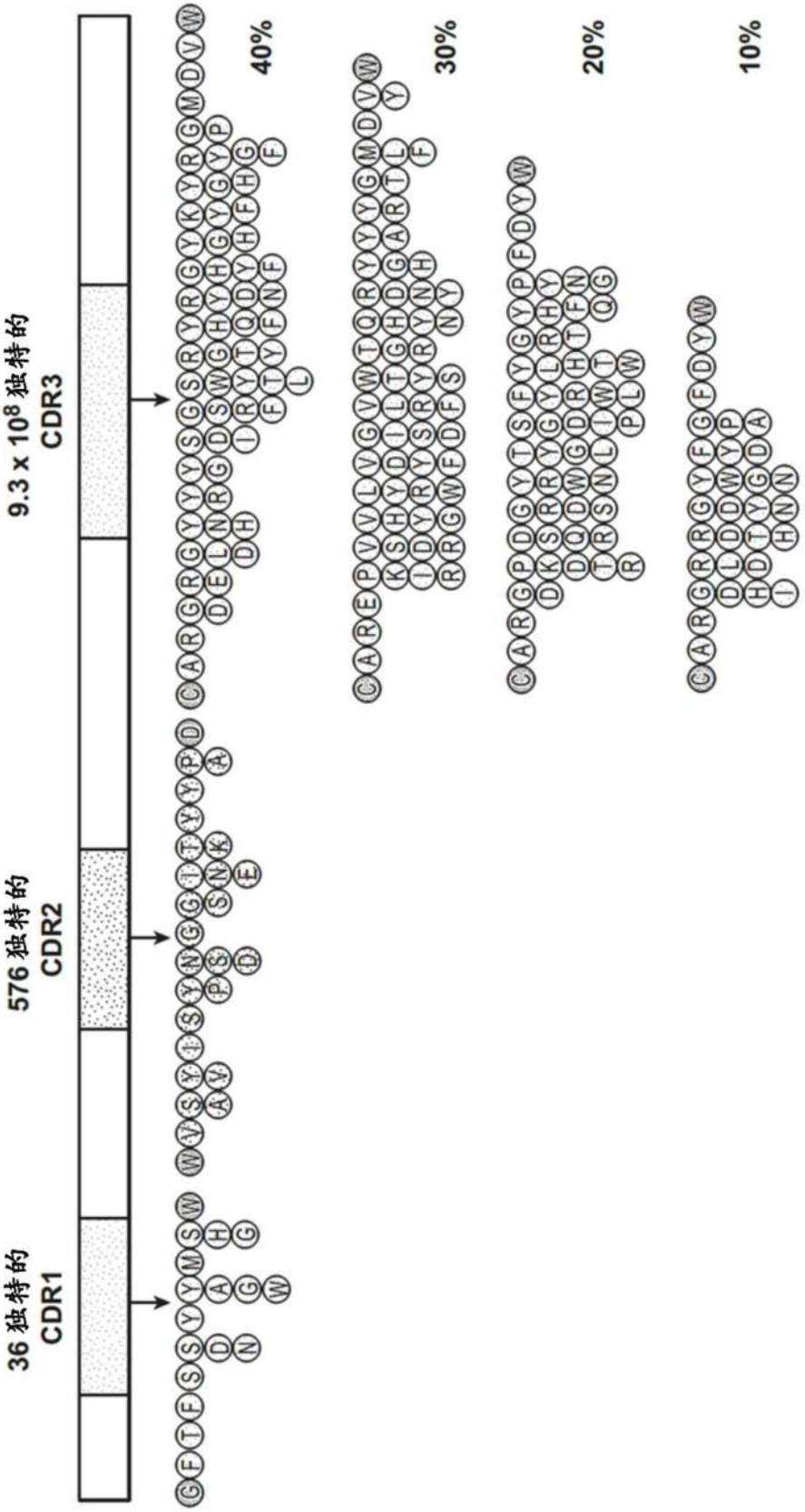


图30A

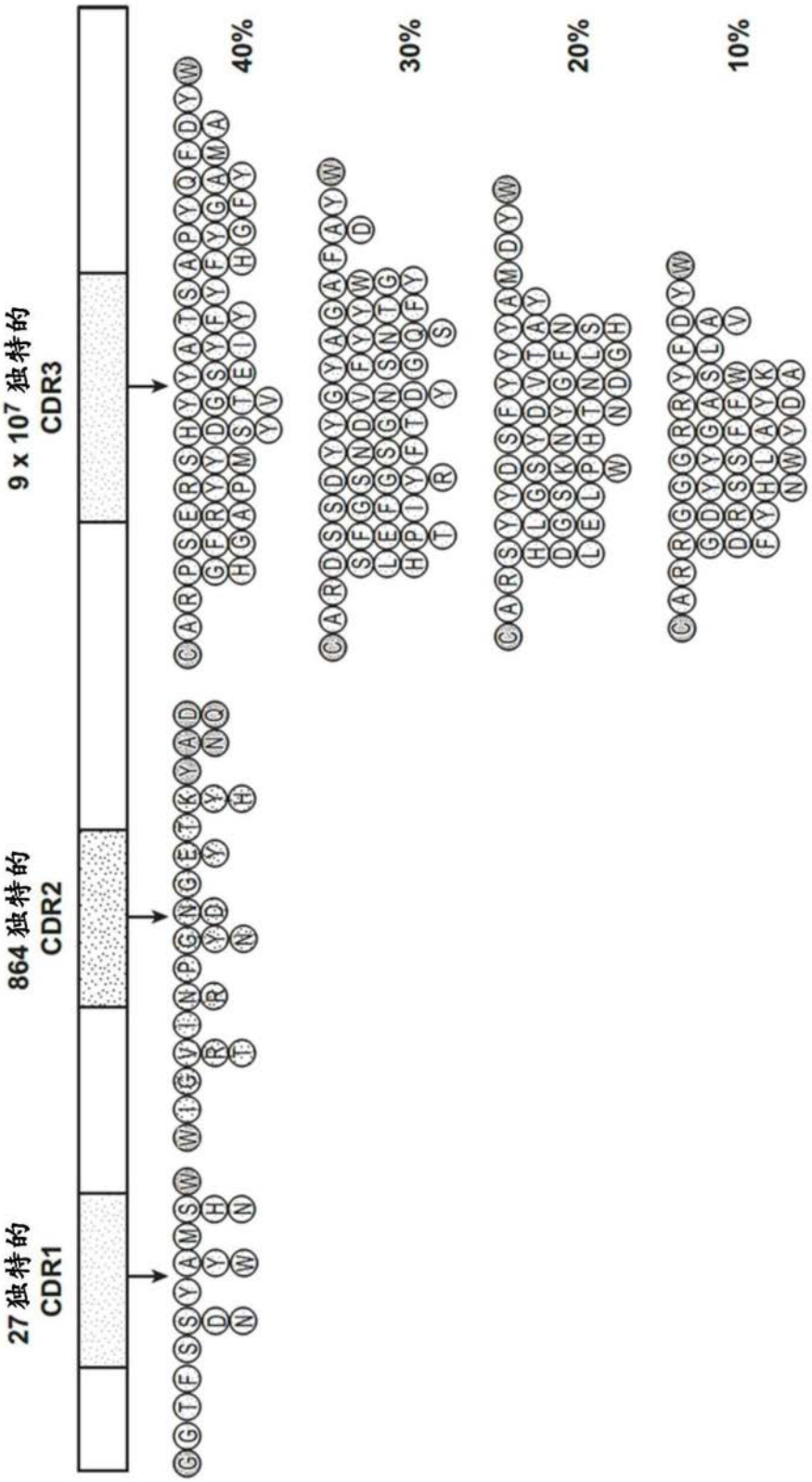


图30B

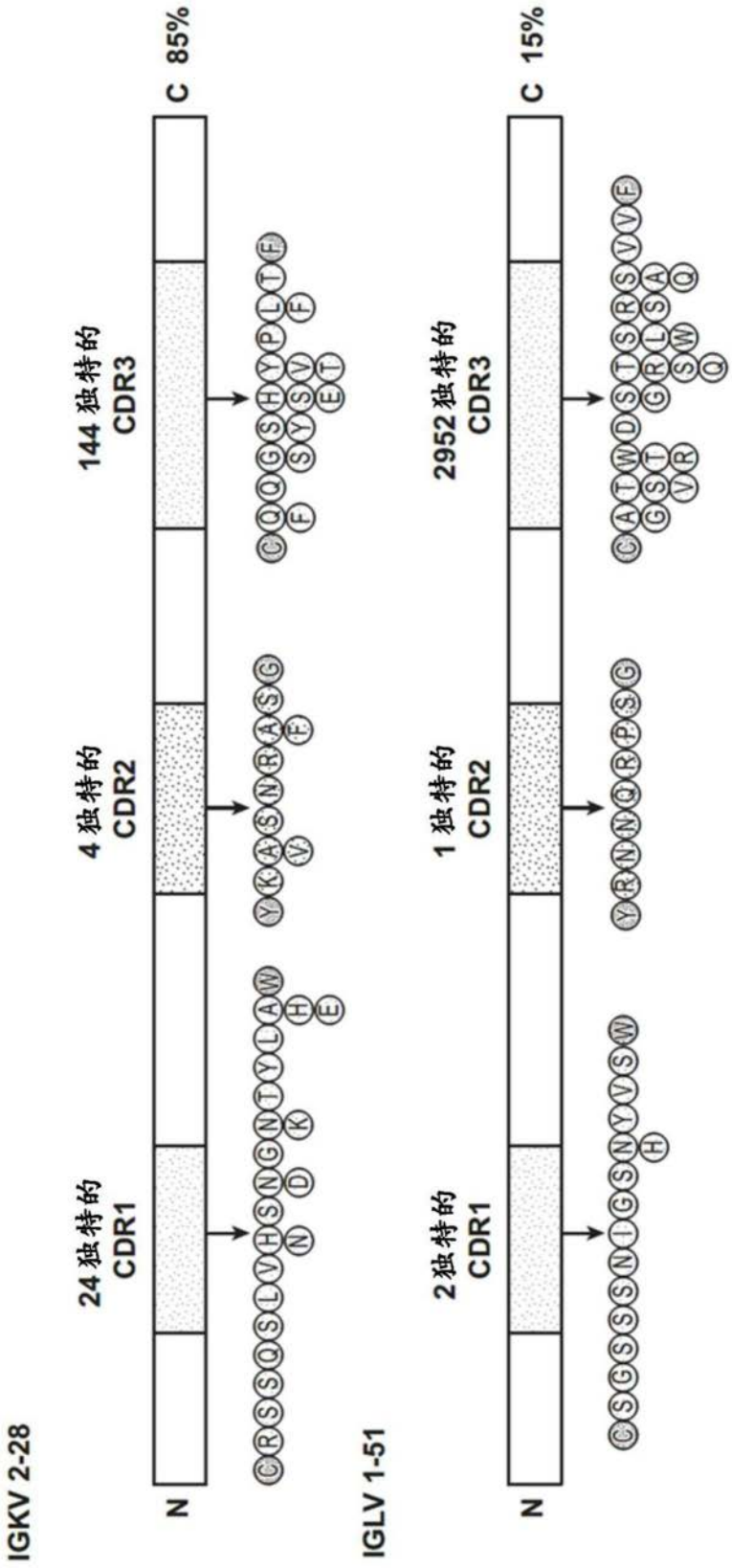

IGKV 2-28
N
C 85%
24 独特的 CDR1
4 独特的 CDR2
144 独特的 CDR3
IGLV 1-51
N
C 15%
2 独特的 CDR1
1 独特的 CDR2
2952 独特的 CDR3

图30C

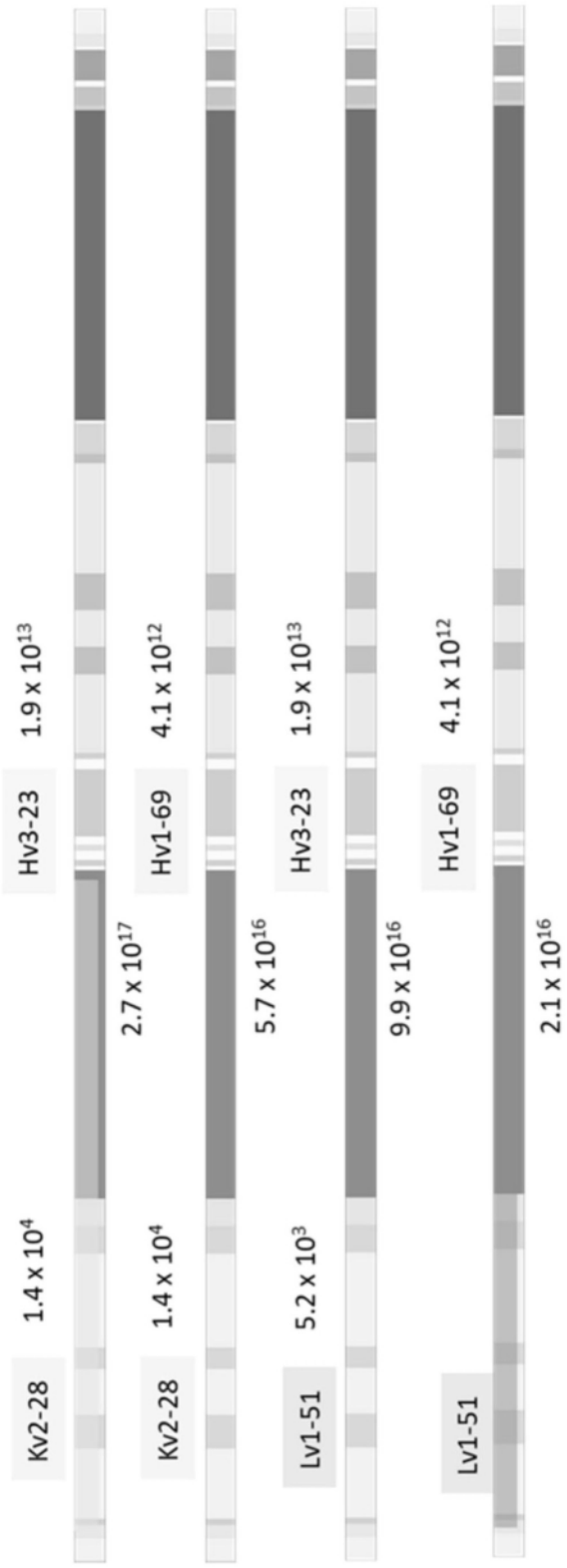

Kv2-28
1.4 x 10^4
Hv3-23
1.9 x 10^13
2.7 x 10^17
Kv2-28
1.4 x 10^4
Hv1-69
4.1 x 10^12
5.7 x 10^16
Lv1-51
5.2 x 10^3
Hv3-23
1.9 x 10^13
9.9 x 10^16
Lv1-51
Hv1-69
4.1 x 10^12
2.1 x 10^16

图30D