(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 116949052 A
(43)申请公布日 2023.10.27

(21)申请号 202310964403.7

(22)申请日 2016.11.11

(30)优先权数据
62/255,317 2015.11.13 US

(62)分案原申请数据
201680067846.4 2016.11.11

(71)申请人 武田药品工业株式会社
地址 日本国大阪府大阪市中央区道修町四丁目1番1号

(72)发明人 J·伦勒 F·法尔科纳 F·霍林 H·罗滕施泰纳 F·沙伊夫灵格

(74)专利代理机构 上海华诚知识产权代理有限公司 31300
专利代理师 汤国华

(51)Int.Cl.
C12N 15/12(2006.01)
C12N 15/864(2006.01)
A61K 48/00(2006.01)
A61K 38/37(2006.01)
A61P 7/04(2006.01)

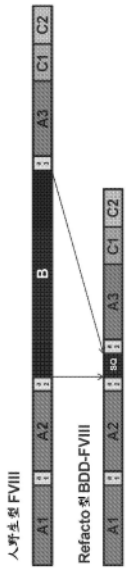
权利要求书6页 说明书62页
序列表（电子公布） 附图87页

(54)发明名称

用于血友病A的基因治疗的具有增加的表达的编码重组FVIII变体的病毒载体

(57)摘要

除其他方面外，本公开提供用于在哺乳动物细胞中表达的编码因子VIII变体的密码子改变的多核苷酸。在一些实施方案中，本公开还提供用于治疗血友病A的哺乳动物基因疗法载体和方法。



CN 116949052 A

1.一种包含与CS01-FL-NA(SEQ ID NO:13)具有至少95%同一性的核苷酸序列的多核苷酸,其中所述多核苷酸编码因子VIII多肽。

2.如权利要求1所述的多核苷酸,其中所述核苷酸序列与CS01-FL-NA(SEQ ID NO:13)具有至少99%的同一性。

3.如权利要求1所述的多核苷酸,其中所述核苷酸序列为CS01-FL-NA(SEQ ID NO:13)。

4.一种包含编码因子VIII多肽的核苷酸序列的多核苷酸,所述因子VIII多肽包含轻链、重链以及将所述重链的C端连接至所述轻链的N端的多肽接头,

其中所述因子VIII多肽的所述重链是由与CS01-HC-NA(SEQ ID NO:24)具有至少95%同一性的第一核苷酸序列编码;

其中所述因子FVIII多肽的所述轻链是由与CS01-LC-NA(SEQ ID NO:25)具有至少95%同一性的第二核苷酸序列编码;并且

其中所述多肽接头包含弗林蛋白酶裂解位点。

5.如权利要求4所述的多核苷酸,其中所述多肽接头是由与BDL001(SEQ ID NO:5)具有至少95%同一性的第三核苷酸序列编码。

6.如权利要求4或5所述的多核苷酸,其中:

所述第一核苷酸序列与CS01-HC-NA(SEQ ID NO:24)具有至少99%同一性;并且

所述第二核苷酸序列与CS01-LC-NA(SEQ ID NO:25)具有至少99%的同一性。

7.如权利要求4或5所述的多核苷酸,其中:

所述第一核苷酸序列为CS01-HC-NA(SEQ ID NO:24);并且

所述第二核苷酸序列CS01-LC-NA(SEQ ID NO:25)。

8.一种包含与CS04-FL-NA(SEQ ID NO:1)具有至少95%同一性的核苷酸序列的多核苷酸,其中所述多核苷酸编码因子VIII多肽。

9.如权利要求4所述的多核苷酸,其中所述核苷酸序列与CS04-FL-NA(SEQ ID NO:1)具有至少99%的同一性。

10.如权利要求4所述的多核苷酸,其中所述核苷酸序列为CS04-FL-NA(SEQ ID NO:1)。

11.一种包含编码因子VIII多肽的核苷酸序列的多核苷酸,所述因子VIII多肽包含轻链、重链以及将所述重链的C端连接至所述轻链的N端的多肽接头,

其中所述因子VIII多肽的所述重链是由与CS04-HC-NA(SEQ ID NO:3)具有至少95%同一性的第一核苷酸序列编码;

其中所述因子FVIII多肽的所述轻链是由与CS04-LC-NA(SEQ ID NO:4)具有至少95%同一性的第二核苷酸序列编码;并且

其中所述多肽接头包含弗林蛋白酶裂解位点。

12.如权利要求11所述的多核苷酸,其中所述多肽接头是由与BDL004(SEQ ID NO:6)具有至少95%同一性的第三核苷酸序列编码。

13.如权利要求11或12所述的多核苷酸,其中:

所述第一核苷酸序列与CS04-HC-NA(SEQ ID NO:3)具有至少99%的同一性;并且

所述第二核苷酸序列与CS04-LC-NA(SEQ ID NO:4)具有至少99%的同一性。

14.如权利要求11或12所述的多核苷酸,其中:

所述第一核苷酸序列为CS04-HC-NA(SEQ ID NO:3);并且

所述第二核苷酸序列CS04-LC-NA(SEQ ID NO:4)。

15.一种包含与CS23-FL-NA具有至少95%同一性的核苷酸序列的多核苷酸,其中所述多核苷酸编码因子VIII多肽。

16.如权利要求15所述的多核苷酸,其中所述核苷酸序列与CS23-FL-NA(SEQ ID NO:20)具有至少99%的同一性。

17.如权利要求15所述的多核苷酸,其中所述核苷酸序列为CS23-FL-NA(SEQ ID NO:20)。

18.一种包含编码因子VIII多肽的核苷酸序列的多核苷酸,所述因子VIII多肽包含轻链、重链以及将所述重链的C端连接至所述轻链的N端的多肽接头,

其中所述因子VIII多肽的所述重链是由与CS23-HC-NA(SEQ ID NO:22)具有至少95%同一性的第一核苷酸序列编码;

其中所述因子FVIII多肽的所述轻链是由与CS23-LC-NA(SEQ ID NO:23)具有至少95%同一性的第二核苷酸序列编码;并且

其中所述多肽接头包含弗林蛋白酶裂解位点。

19.如权利要求18所述的多核苷酸,其中所述多肽接头是由与BDLO23(SEQ ID NO:7)具有至少95%同一性的第三核苷酸序列编码。

20.如权利要求18或19所述的多核苷酸,其中:

所述第一核苷酸序列与CS23-HC-NA(SEQ ID NO:22)具有至少99%的同一性;并且

所述第二核苷酸序列与CS23-LC-NA(SEQ ID NO:23)具有至少99%的同一性。

21.如权利要求18或19所述的多核苷酸,其中:

所述第一核苷酸序列为CS23-HC-NA(SEQ ID NO:22);并且

所述第二核苷酸序列CS23-LC-NA(SEQ ID NO:23)。

22.如权利要求4、11以及18中任一项所述的多核苷酸,其中编码的多肽接头包含氨基酸序列与选自由以下各项组成的组的糖基化肽具有至少92%同一性的糖基化肽:NG1-AA、NG4-AA、NG5-AA、NG6-AA、NG7-AA、NG9-AA、NG10-AA、NG16-AA、NG17-AA、NG18-AA、NG19-AA、NG20-AA、NG21-AA以及NGV-AA。

23.如权利要求4、11以及18中任一项所述的多核苷酸,其中所述编码的多肽接头包含选自由以下各项组成的组的糖基化肽:NG1-AA、NG4-AA、NG5-AA、NG6-AA、NG7-AA、NG9-AA、NG10-AA、NG16-AA、NG17-AA、NG18-AA、NG19-AA、NG20-AA、NG21-AA以及NGV-AA。

24.如权利要求22或23所述的多核苷酸,其中所述糖基化肽是由核苷酸序列与选自由以下各项组成的组的序列具有至少95%同一性的多核苷酸编码:NG1-NA、NG4-NA、NG5-NA、NG6-NA、NG7-NA、NG9-NA、NG10-NA、NG16-NA、NG17-NA、NG18-NA、NG19-NA、NG20-NA、NG21-NA以及NGV-NA。

25.如权利要求22或23所述的多核苷酸,其中所述糖基化肽是由核苷酸序列选自以下各项中的一种的多核苷酸编码:NG1-NA、NG4-NA、NG5-NA、NG6-NA、NG7-NA、NG9-NA、NG10-NA、NG16-NA、NG17-NA、NG18-NA、NG19-NA、NG20-NA、NG21-NA以及NGV-NA。

26.如权利要求4、11以及18中任一项所述的多核苷酸,其中所述多肽接头是由与选自由以下各项组成的组的序列具有至少95%同一性的第三核苷酸序列编码:BDLNG1-NA、BDLNG3-NA、BDLNG5-NA、BDLNG6-NA、BDLNG9-NA、BDLNG10-NA、BDLNG16-NA、BDLNG17-NA、

BDLNG18-NA、BDLNG19-NA、BDLNG20-NA以及BDLNG21-NA。

27.如权利要求1至21中任一项所述的多核苷酸，其中所述多核苷酸编码包含与CS04-FL-AA(SEQ ID NO:2)具有至少95％同一性的氨基酸序列的因子VIII多肽。

28.如权利要求1至21中任一项所述的多核苷酸，其中所述多核苷酸编码包含CS04-FL-AA(SEQ ID NO:2)的氨基酸序列的因子VIII多肽。

29.一种包含与CS01-SC1-NA(SEQ ID NO:26)具有至少95％同一性的核苷酸序列的多核苷酸，其中所述多核苷酸编码单链因子VIII多肽。

30.如权利要求29所述的多核苷酸，其中所述核苷酸序列与CS01-SC1-NA(SEQ ID NO:26)具有至少99％的同一性。

31.如权利要求29所述的多核苷酸，其中所述核苷酸序列为CS01-SC1-NA(SEQ ID NO:26)。

32.一种包含与CS04-SC1-NA(SEQ ID NO:9)具有至少95％同一性的核苷酸序列的多核苷酸，其中所述多核苷酸编码单链因子VIII多肽。

33.如权利要求29所述的多核苷酸，其中所述核苷酸序列与CS04-SC1-NA(SEQ ID NO:9)具有至少99％的同一性。

34.如权利要求29所述的多核苷酸，其中所述核苷酸序列为CS04-SC1-NA(SEQ ID NO:9)。

35.一种包含与CS23-SC1-NA(SEQ ID NO:28)具有至少95％同一性的核苷酸序列的多核苷酸，其中所述多核苷酸编码单链因子VIII多肽。

36.如权利要求29所述的多核苷酸，其中所述核苷酸序列与CS23-SC1-NA(SEQ ID NO:28)具有至少99％的同一性。

37.如权利要求29所述的多核苷酸，其中所述核苷酸序列为CS23-SC1-NA(SEQ ID NO:28)。

38.如权利要求29至37中任一项所述的多核苷酸，其中所述多核苷酸编码包含与CS01-SC1-AA(SEQ ID NO:10)具有至少95％同一性的氨基酸序列的单链因子VIII多肽。

39.如权利要求29至37中任一项所述的多核苷酸，其中所述多核苷酸编码包含CS01-SC1-AA(SEQ ID NO:10)的氨基酸序列的单链因子VIII多肽。

40.一种包含与CS01-SC2-NA(SEQ ID NO:27)具有至少95％同一性的核苷酸序列的多核苷酸，其中所述多核苷酸编码单链因子VIII多肽。

41.如权利要求40所述的多核苷酸，其中所述核苷酸序列与CS01-SC2-NA(SEQ ID NO:27)具有至少99％的同一性。

42.如权利要求40所述的多核苷酸，其中所述核苷酸序列为CS01-SC2-NA(SEQ ID NO:27)。

43.一种包含与CS04-SC2-NA(SEQ ID NO:11)具有至少95％同一性的核苷酸序列的多核苷酸，其中所述多核苷酸编码单链因子VIII多肽。

44.如权利要求43所述的多核苷酸，其中所述核苷酸序列与CS04-SC2-NA(SEQ ID NO:11)具有至少99％的同一性。

45.如权利要求43所述的多核苷酸，其中所述核苷酸序列为CS04-SC2-NA(SEQ ID NO:11)。

46.一种包含与CS23-SC2-NA(SEQ ID NO:29)具有至少95%同一性的核苷酸序列的多核苷酸,其中所述多核苷酸编码单链因子VIII多肽。

47.如权利要求43所述的多核苷酸,其中所述核苷酸序列与CS23-SC2-NA(SEQ ID NO:29)具有至少99%的同一性。

48.如权利要求43所述的多核苷酸,其中所述核苷酸序列为CS23-SC2-NA(SEQ ID NO:29)。

49.如权利要求40至48中任一项所述的多核苷酸,其中所述多核苷酸编码包含与CS01-SC2-AA(SEQ ID NO:12)具有至少95%同一性的氨基酸序列的单链因子VIII多肽。

50.如权利要求40至48中任一项所述的多核苷酸,其中所述多核苷酸编码包含CS01-SC2-AA(SEQ ID NO:12)的氨基酸序列的单链因子VIII多肽。

51.一种包含与选自由以下各项组成的组的序列具有至少95%同一性的序列的多核苷酸:CS01-HC-NA、CS01-LC-NA、CS04-HC-NA、CS04-LC-NA、CS23-HC-NA、CS23-LC-NA。

52.一种包含与选自由以下各项组成的组的序列具有至少99%同一性的序列的多核苷酸:CS01-HC-NA、CS01-LC-NA、CS04-HC-NA、CS04-LC-NA、CS23-HC-NA、CS23-LC-NA。

53.一种包含选自由以下各项组成的组的序列的多核苷酸:CS01-HC-NA、CS01-LC-NA、CS04-HC-NA、CS04-LC-NA、CS23-HC-NA、CS23-LC-NA。

54.如权利要求51至53中任一项所述的多核苷酸,其中所述多核苷酸编码包含与CS01-FL-AA(SEQ ID NO:2)具有至少95%同一性的氨基酸序列的因子VIII多肽。

55.如权利要求51至53中任一项所述的多核苷酸,其中所述多核苷酸编码包含CS01-FL-AA(SEQ ID NO:2)的氨基酸序列的因子VIII多肽。

56.如权利要求1至55中任一项所述的多核苷酸,其中所述编码的因子VIII多肽相对于FVIII-FL-AA(SEQ ID NO:19)包含F328S氨基酸取代。

57.如权利要求1至55中任一项所述的多核苷酸,其中所述编码的因子VIII多肽相对于FVIII-FL-AA(SEQ ID NO:19)包含I105V、A127S、G151K、M166T以及L171P氨基酸取代。

58.如权利要求1至55中任一项所述的多核苷酸,其中所述编码的因子VIII多肽包含:

a)氨基酸缺失AIEPRSF755-761,相对于FVIII-FL-AA(SEQ ID NO:19);以及

b)N754后的氨基酸插入TTYVNRSL(SEQ ID NO:33),相对于FVIII-FL-AA(SEQ ID NO:19)。

59.如权利要求1至55中任一项所述的多核苷酸,其中所述编码的因子VIII多肽包含:

a)F328S氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19);以及

b)C1918G和C1922G氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19)。

60.如权利要求1至55中任一项所述的多核苷酸,其中所述编码的因子VIII多肽相对于FVIII-FL-AA(SEQ ID NO:19)包含I105V、A127S、G151K、M166T、L171P以及F328S氨基酸取代。

61.如权利要求1至55中任一项所述的多核苷酸,其中所述编码的因子VIII多肽包含:

a)F328S氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19);

b)氨基酸缺失AIEPRSF755-761,相对于FVIII-FL-AA(SEQ ID NO:19);以及

c)N754后的氨基酸插入TTYVNRSL(SEQ ID NO:33),相对于FVIII-FL-AA(SEQ ID NO:19)。

62.如权利要求1至55中任一项所述的多核苷酸,其中所述编码的因子VIII多肽包含:

a)I105V、A127S、G151K、M166T以及L171P氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19);

b)氨基酸缺失AIEPRSF755-761,相对于FVIII-FL-AA(SEQ ID NO:19);以及

c)N754后的氨基酸插入TTYVNRSL(SEQ ID NO:33),相对于FVIII-FL-AA(SEQ ID NO:19)。

63.如权利要求1至55中任一项所述的多核苷酸,其中所述编码的因子VIII多肽包含:

a)F328S氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19);

b)C1918G和C1922G氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19);以及

c)I105V、A127S、G151K、M166T以及L171P氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19)。

64.如权利要求1至55中任一项所述的多核苷酸,其中所述编码的因子VIII多肽包含:

a)F328S氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19);

b)C1918G和C1922G氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19);

c)氨基酸缺失AIEPRSF755-761,相对于FVIII-FL-AA(SEQ ID NO:19);以及

d)N754后的氨基酸插入TTYVNRSL(SEQ ID NO:33),相对于FVIII-FL-AA(SEQ ID NO:19)。

65.如权利要求1至55中任一项所述的多核苷酸,其中所述编码的因子VIII多肽包含:

a)I105V、A127S、G151K、M166T以及L171P氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19);

b)F328S氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19);

c)C1918G和C1922G氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19);

d)氨基酸缺失AIEPRSF755-761,相对于FVIII-FL-AA(SEQ ID NO:19);以及

e)N754后的氨基酸插入TTYVNRSL(SEQ ID NO:33),相对于FVIII-FL-AA(SEQ ID NO:19)。

66.如权利要求1至65中任一项所述的多核苷酸,其进一步包含可操作地连接至编码所述因子VIII多肽的所述多核苷酸的启动子元件。

67.如权利要求66所述的多核苷酸,其中所述启动子元件为编码所述因子VIII多肽的所述核苷酸序列上游的肝特异性启动子序列。

68.如权利要求67所述的多核苷酸,其进一步包含位于所述肝特异性启动子序列与编码所述因子VIII多肽的所述核苷酸序列之间的内含子序列。

69.一种腺相关病毒(AAV)载体,所述腺相关病毒载体包含如权利要求1至68中任一项所述的多核苷酸。

70.一种腺相关病毒(AAV)粒子,所述腺相关病毒粒子包含如权利要求1至68中任一项所述的多核苷酸。

71.一种宿主细胞,所述宿主细胞用包含如权利要求1至68中任一项所述的多核苷酸的腺相关病毒(AAV)粒子感染。

72.一种制备腺相关病毒(AAV)粒子的方法,所述方法包括将如权利要求1至68中任一项所述的多核苷酸引入哺乳动物宿主细胞中,其中所述多核苷酸能够在所述哺乳动物宿主

细胞中复制。

73.一种治疗血友病A的方法,所述方法包括向有需要的患者施用如权利要求70所述的腺相关病毒(AAV)粒子。

74.一种转导宿主细胞的方法,所述方法包括使所述宿主细胞与如权利要求70所述的腺相关病毒(AAV)粒子接触。

用于血友病A的基因治疗的具有增加的表达的编码重组FVIII变体的病毒载体

[0001] 本申请是申请日为2016年11月11日、申请号为201680067846.4(国际申请号为PCT/US2016/061684)、名称为“用于血友病A的基因治疗的具有增加的表达的编码重组FVIII变体的病毒载体”的发明专利申请的分案申请。

[0002] 相关申请的交叉引用

[0003] 本申请要求于2015年11月13日提交的美国临时专利申请号62/255,317的优先权，该临时专利申请的内容出于所有目的以全文引用的方式并入本文中。

[0004] 序列表

[0005] 本申请含有已经以ASCII格式电子提交并且以全文引用的方式并入本文中的序列表。所述ASCII副本创建于2016年11月9日，命名为008073_5107_WO_Sequence_Listing.txt并且大小为353,479字节。

[0006] 发明背景

[0007] 血液凝固通过称为凝血级联的相互依赖的生物化学反应的复杂且动态的生物途径进行。凝血因子VIII(FVIII)是级联中的关键组分。因子VIII被招募到出血位点，并且与活化的因子IX(FIXa)和因子X(FX)形成X酶复合物。X酶复合物活化FX，FX又将凝血酶原活化为凝血酶，然后凝血酶活化凝血级联中的其他组分以产生稳定的凝块(在Saenko等,Trends Cardiovasc.Med.,9:185-192(1999);Lenting等,Blood,92:3983-3996(1998)).

[0008] 血友病A是一种以缺乏因子VIII活性为特征的先天性X连锁性出血性病症。降低的因子VIII活性抑制了凝血级联中的正反馈回路。这会导致不完全凝固，表现为持续时间延长、大面积青紫、自发性口腔和鼻腔出血、关节僵硬和慢性疼痛以及严重情况下可能出现内出血和贫血的出血事件(Zhang等,Clinic.Rev.Allerg.Immunol.,37:114-124(2009))。

[0009] 常规地，通过因子VIII替代疗法来治疗血友病A，所述因子VIII替代疗法由向血友病A个体施用因子VIII蛋白(例如血浆衍生或重组产生的因子VIII)组成。预防性地施用因子VIII以预防出血事件或响应于急性出血事件降低其频率，和/或围手术期施用以管理手术期间的出血。然而，因子VIII替代疗法有几个不希望的特征。

[0010] 第一，因子VIII替代疗法是用于治疗或管理血友病A，但不能治愈潜在的因子VIII缺乏。由于这个原因，血友病A个体在其一生中都需要因子VIII替代疗法。持续治疗是昂贵的，并且要求个体保持严格的依从性，因为仅缺少少数预防性剂量就会对严重血友病A的个体产生严重后果。

[0011] 第二，因为因子VIII在体内的半衰期相对较短，常规预防性因子VIII替代疗法需要每两天或每三天进行施药。这给个体在其生活中保持依从性带来负担。尽管第三代“长效”因子VIII药物可能会降低施用频率，但使用这些药物的预防性因子FVIII替代疗法仍需要永久性地每月、每周或更频繁的进行施药。举例来说，用ELOCTATE™[抗血友病因子(重组)，Fc融合蛋白]预防性治疗需要每3到5天施药(ELOCTATE™Prescribing Information, Biogen Idec Inc.,(2015))。此外，尚未完全了解化学修饰的生物制剂(例如聚乙二醇化多肽)的长期影响。

[0012] 第三,接受因子VIII替代疗法的所有个体中有15%至30%形成抗因子VIII抑制剂抗体,从而导致疗法效率低。可使用因子VIII旁路疗法(例如施用血浆衍生或重组产生的凝血酶原复合物浓缩物)来治疗形成抑制剂抗体的个体的血友病。然而,因子VIII旁路疗法与因子VIII替代疗法相比不太有效(Mannucci P.M.,J Thromb Haemost.,1(7):1349-55(2003))并且可能与心血管并发症的风险增加有关(Luu和Ewenstein,Haemophilia,10增刊2:10-16(2004))。

[0013] 体细胞基因疗法为血友病A的治疗带来极大的希望,因为它可以补救潜在的低表达功能性因子VIII活性(例如归因于错义或无义突变),而不是提供一次剂量的因子VIII活性给个体。由于作用机制的这种差异,与因子VIII替代疗法相比,一次施用因子VIII基因疗法载体可为个体提供几年的因子VIII,从而降低治疗成本并且消除了对持续患者依从性的需要。

[0014] 凝血因子IX(FIX)基因疗法已被有效地用于治疗血友病B个体,血友病B是一种以因子IX活性降低为特征的相关凝血病状(Manno C.S.等,Nat Med.,12(3):342-47(2006))。然而,因子VIII基因疗法存在几个独特的挑战。举例来说,全长野生型因子VIII多肽(2351个氨基酸;UniProt登录号P00451)比全长野生型因子IX多肽(461个氨基酸;UniProt登录号P00740)大五倍。因此,野生型因子VIII的编码序列为7053个碱基对,这太大而不能包装在常规AAV基因疗法载体中。此外,报道的因子VIII的B-结构域缺失变体(BDD-FVIII)的重组表达是差的。因此,若干小组已试图改变BDD-FVIII构建体的密码子使用,但取得的成功有限。

发明概要

[0015] 因此,对编码序列更高效地包装到基因疗法载体中并且经由基因疗法载体递送的因子VIII变体存在需要。还对更高效地表达因子VIII的合成的密码子改变的核酸存在需要。此类因子VIII变体和密码子改变的核酸允许因子VIII缺乏(例如血友病A)的改善的治疗。公开的密码子改变的因子VIII变体减少或消除了上述缺乏和与治疗因子VIII缺乏(例如血友病A)相关的其他问题。

[0016] 根据一些实施方案,本公开提供编码因子VIII变体的核酸,所述核酸与公开的密码子改变的因子VIII重链(例如CS01-HC-NA、CS04-HC-NA或CS23-HC-NA)和轻链(CS01-LC-NA、CS04-LC-NA或CS23-LC-NA)的序列具有高序列同一性。在一些实施方案中,这些核酸在编码因子VIII重链和轻链的序列之间进一步包括编码接头序列的序列,所述接头序列替代天然因子VIII B-结构域(例如包含弗林蛋白酶裂解位点(furin cleavage site)的接头序列)。

[0017] 在一个方面中,本公开提供一种包括编码因子VIII多肽的核苷酸序列的多核苷酸。因子VIII多肽包括轻链、重链以及将重链的C端连接至轻链的N端的多肽接头。因子VIII多肽的重链是由与CS04-HC-NA(SEQ ID NO:3)具有至少95%同一性的第一核苷酸序列编码。因子FVIII多肽的轻链是由与CS04-LC-NA(SEQ ID NO:4)具有至少95%同一性的第二核苷酸序列编码。多肽接头包含弗林蛋白酶裂解位点。

[0018] 在上文所描述的多核苷酸的一个实施方案中,多肽接头由与BDL004(SEQ ID NO:6)具有至少95%同一性的第三核苷酸序列编码。

[0019] 在一个方面中，本公开提供一种包括编码因子VIII多肽的核苷酸序列的多核苷酸。因子VIII多肽包括轻链、重链以及将重链的C端连接至轻链的N端的多肽接头。因子VIII多肽的重链是由与CS01-HC-NA(SEQ ID NO:24)具有至少95％同一性的第一核苷酸序列编码。因子FVIII多肽的轻链是由与CS01-LC-NA(SEQ ID NO:25)具有至少95％同一性的第二核苷酸序列编码。多肽接头包含弗林蛋白酶裂解位点。

[0020] 在上文所描述的多核苷酸的一个实施方案中，多肽接头由与BDL001(SEQ ID NO:5)具有至少95％同一性的第三核苷酸序列编码。

[0021] 在一个方面中，本公开提供一种包括编码因子VIII多肽的核苷酸序列的多核苷酸。因子VIII多肽包括轻链、重链以及将重链的C端连接至轻链的N端的多肽接头。因子VIII多肽的重链是由与CS23-HC-NA(SEQ ID NO:22)具有至少95％同一性的第一核苷酸序列编码。因子FVIII多肽的轻链是由与CS23-LC-NA(SEQ ID NO:23)具有至少95％同一性的第二核苷酸序列编码。多肽接头包含弗林蛋白酶裂解位点。

[0022] 在上文所描述的多核苷酸的一个实施方案中，多肽接头由与BDL023(SEQ ID NO:7)具有至少95％同一性的第三核苷酸序列编码。

[0023] 在上文所描述的多核苷酸的一个实施方案中，编码因子VIII多肽的重链的第一核苷酸序列与相应重链序列(例如CS04-HC-NA(SEQ ID NO:3)、CS01-HC-NA(SEQ ID NO:24)或CS23-HC-NA(SEQ ID NO:22))具有至少96％的同一性，并且编码因子FVIII多肽的轻链的第二核苷酸序列与相应轻链序列(例如CS04-LC-NA(SEQ ID NO:4)、CS01-LC-NA(SEQ ID NO:25)或CS23-LC-NA(SEQ ID NO:23))具有至少96％的同一性。

[0024] 在上文所描述的多核苷酸的一个实施方案中，编码因子VIII多肽的重链的第一核苷酸序列与相应重链序列(例如CS04-HC-NA(SEQ ID NO:3)、CS01-HC-NA(SEQ ID NO:24)或CS23-HC-NA(SEQ ID NO:22))具有至少97％的同一性，并且编码因子FVIII多肽的轻链的第二核苷酸序列与相应轻链序列(例如CS04-LC-NA(SEQ ID NO:4)、CS01-LC-NA(SEQ ID NO:25)或CS23-LC-NA(SEQ ID NO:23))具有至少97％的同一性。

[0025] 在上文所描述的多核苷酸的一个实施方案中，编码因子VIII多肽的重链的第一核苷酸序列与相应重链序列(例如CS04-HC-NA(SEQ ID NO:3)、CS01-HC-NA(SEQ ID NO:24)或CS23-HC-NA(SEQ ID NO:22))具有至少98％的同一性，并且编码因子FVIII多肽的轻链的第二核苷酸序列与相应轻链序列(例如CS04-LC-NA(SEQ ID NO:4)、CS01-LC-NA(SEQ ID NO:25)或CS23-LC-NA(SEQ ID NO:23))具有至少98％的同一性。

[0026] 在上文所描述的多核苷酸的一个实施方案中，编码因子VIII多肽的重链的第一核苷酸序列与相应重链序列(例如CS04-HC-NA(SEQ ID NO:3)、CS01-HC-NA(SEQ ID NO:24)或CS23-HC-NA(SEQ ID NO:22))具有至少99％的同一性，并且编码因子FVIII多肽的轻链的第二核苷酸序列与相应轻链序列(例如CS04-LC-NA(SEQ ID NO:4)、CS01-LC-NA(SEQ ID NO:25)或CS23-LC-NA(SEQ ID NO:23))具有至少99％的同一性。

[0027] 在上文所描述的多核苷酸的一个实施方案中，编码因子VIII多肽的重链的第一核苷酸序列与相应重链序列(例如CS04-HC-NA(SEQ ID NO:3)、CS01-HC-NA(SEQ ID NO:24)或CS23-HC-NA(SEQ ID NO:22))具有至少99.5％的同一性，并且编码因子FVIII多肽的轻链的第二核苷酸序列与相应轻链序列(例如CS04-LC-NA(SEQ ID NO:4)、CS01-LC-NA(SEQ ID NO:25)或CS23-LC-NA(SEQ ID NO:23))具有至少99.5％的同一性。

[0028] 在上文所描述的多核苷酸的一个实施方案中,编码因子VIII多肽的重链的第一核苷酸序列与相应重链序列(例如CS04-HC-NA(SEQ ID NO:3)、CS01-HC-NA(SEQ ID NO:24)或CS23-HC-NA(SEQ ID NO:22))具有至少99.9%的同一性,并且编码因子FVIII多肽的轻链的第二核苷酸序列与相应轻链序列(例如CS04-LC-NA(SEQ ID NO:4)、CS01-LC-NA(SEQ ID NO:25)或CS23-LC-NA(SEQ ID NO:23))具有至少99.9%的同一性。

[0029] 在上文所描述的多核苷酸的一个实施方案中,编码因子VIII多肽的重链的第一核苷酸序列为CS04-HC-NA(SEQ ID NO:3),并且编码因子FVIII多肽的轻链的第二核苷酸序列为CS04-LC-NA(SEQ ID NO:4)。

[0030] 在上文所描述的多核苷酸的一个实施方案中,编码因子VIII多肽的重链的第一核苷酸序列为CS01-HC-NA(SEQ ID NO:24),并且编码因子FVIII多肽的轻链的第二核苷酸序列为CS01-LC-NA(SEQ ID NO:25)。

[0031] 在上文所描述的多核苷酸的一个实施方案中,编码因子VIII多肽的重链的第一核苷酸序列为CS23-HC-NA(SEQ ID NO:22),并且编码因子FVIII多肽的轻链的第二核苷酸序列为CS23-LC-NA(SEQ ID NO:23)。

[0032] 在一个方面中,本公开提供一种包含与CS04-FL-NA具有至少95%同一性的核苷酸序列的多核苷酸,其中所述多核苷酸编码因子VIII多肽。

[0033] 在一个方面中,本公开提供一种包含与CS01-FL-NA具有至少95%同一性的核苷酸序列的多核苷酸,其中所述多核苷酸编码因子VIII多肽。

[0034] 在一个方面中,本公开提供一种包含与CS23-FL-NA具有至少95%同一性的核苷酸序列的多核苷酸,其中所述多核苷酸编码因子VIII多肽。

[0035] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列与相应全长多核苷酸序列(例如CS04-FL-NA(SEQ ID NO:1)、CS01-FL-NA(SEQ ID NO:13)或CS23-FL-NA(SEQ ID NO:20))具有至少96%的同一性。

[0036] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列与相应全长多核苷酸序列(例如CS04-FL-NA(SEQ ID NO:1)、CS01-FL-NA(SEQ ID NO:13)或CS23-FL-NA(SEQ ID NO:20))具有至少97%的同一性。

[0037] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列与相应全长多核苷酸序列(例如CS04-FL-NA(SEQ ID NO:1)、CS01-FL-NA(SEQ ID NO:13)或CS23-FL-NA(SEQ ID NO:20))具有至少98%的同一性。

[0038] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列与相应全长多核苷酸序列(例如CS04-FL-NA(SEQ ID NO:1)、CS01-FL-NA(SEQ ID NO:13)或CS23-FL-NA(SEQ ID NO:20))具有至少99%的同一性。

[0039] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列与相应全长多核苷酸序列(例如CS04-FL-NA(SEQ ID NO:1)、CS01-FL-NA(SEQ ID NO:13)或CS23-FL-NA(SEQ ID NO:20))具有至少99.5%的同一性。

[0040] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列与相应全长多核苷酸序列(例如CS04-FL-NA(SEQ ID NO:1)、CS01-FL-NA(SEQ ID NO:13)或CS23-FL-NA(SEQ ID NO:20))具有至少99.9%的同一性。

[0041] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列为CS04-FL-NA(SEQ ID

NO:1)。

[0042] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列为CS01-FL-NA(SEQ ID NO:13)。

[0043] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列为CS23-FL-NA(SEQ ID NO:20)。

[0044] 在上文所描述的多核苷酸的一个实施方案中,多核苷酸编码包含与CS04-FL-AA(SEQ ID NO:2)具有至少95%同一性的氨基酸序列的因子VIII多肽。

[0045] 在上文所描述的多核苷酸的一个实施方案中,多核苷酸编码包含与CS04-FL-AA(SEQ ID NO:2)具有至少96%同一性的氨基酸序列的因子VIII多肽。

[0046] 在上文所描述的多核苷酸的一个实施方案中,多核苷酸编码包含与CS04-FL-AA(SEQ ID NO:2)具有至少97%同一性的氨基酸序列的因子VIII多肽。

[0047] 在上文所描述的多核苷酸的一个实施方案中,多核苷酸编码包含与CS04-FL-AA(SEQ ID NO:2)具有至少98%同一性的氨基酸序列的因子VIII多肽。

[0048] 在上文所描述的多核苷酸的一个实施方案中,多核苷酸编码包含与CS04-FL-AA(SEQ ID NO:2)具有至少99%同一性的氨基酸序列的因子VIII多肽。

[0049] 在上文所描述的多核苷酸的一个实施方案中,多核苷酸编码包含与CS04-FL-AA(SEQ ID NO:2)具有至少99.5%同一性的氨基酸序列的因子VIII多肽。

[0050] 在上文所描述的多核苷酸的一个实施方案中,多核苷酸编码包含与CS04-FL-AA(SEQ ID NO:2)具有至少99.9%同一性的氨基酸序列的因子VIII多肽。

[0051] 在上文所描述的多核苷酸的一个实施方案中,多核苷酸编码包含CS04-FL-AA(SEQ ID NO:2)的氨基酸序列的因子VIII多肽。

[0052] 在一个方面中,本公开提供一种包含与CS04-SC1-NA(SEQ ID NO:9)具有至少95%同一性的核苷酸序列的多核苷酸,其中所述多核苷酸编码单链因子VIII多肽。

[0053] 在一个方面中,本公开提供一种包含与CS04-SC2-NA(SEQ ID NO:11)具有至少95%同一性的核苷酸序列的多核苷酸,其中所述多核苷酸编码单链因子VIII多肽。

[0054] 在一个方面中,本公开提供一种包含与CS01-SC1-NA(SEQ ID NO:26)具有至少95%同一性的核苷酸序列的多核苷酸,其中所述多核苷酸编码单链因子VIII多肽。

[0055] 在一个方面中,本公开提供一种包含与CS01-SC2-NA(SEQ ID NO:27)具有至少95%同一性的核苷酸序列的多核苷酸,其中所述多核苷酸编码单链因子VIII多肽。

[0056] 在一个方面中,本公开提供一种包含与CS23-SC1-NA(SEQ ID NO:28)具有至少95%同一性的核苷酸序列的多核苷酸,其中所述多核苷酸编码单链因子VIII多肽。

[0057] 在一个方面中,本公开提供一种包含与CS23-SC2-NA(SEQ ID NO:29)具有至少95%同一性的核苷酸序列的多核苷酸,其中所述多核苷酸编码单链因子VIII多肽。

[0058] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列与相应全长多核苷酸序列(例如CS04-SC1-NA(SEQ ID NO:9)、CS04-SC2-NA(SEQ ID NO:11)、CS01-SC1-NA(SEQ ID NO:26)、CS01-SC2-NA(SEQ ID NO:27)、CS23-SC1-NA(SEQ ID NO:28)或CS23-SC2-NA(SEQ ID NO:29))具有至少96%的同一性。

[0059] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列与相应全长多核苷酸序列(例如CS04-SC1-NA(SEQ ID NO:9)、CS04-SC2-NA(SEQ ID NO:11)、CS01-SC1-NA(SEQ ID

NO:26)、CS01-SC2-NA(SEQ ID NO:27)、CS23-SC1-NA(SEQ ID NO:28)或CS23-SC2-NA(SEQ ID NO:29))具有至少97%的同一性。

[0060] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列与相应全长多核苷酸序列(例如CS04-SC1-NA(SEQ ID NO:9)、CS04-SC2-NA(SEQ ID NO:11)、CS01-SC1-NA(SEQ ID NO:26)、CS01-SC2-NA(SEQ ID NO:27)、CS23-SC1-NA(SEQ ID NO:28)或CS23-SC2-NA(SEQ ID NO:29))具有至少98%的同一性。

[0061] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列与相应全长多核苷酸序列(例如CS04-SC1-NA(SEQ ID NO:9)、CS04-SC2-NA(SEQ ID NO:11)、CS01-SC1-NA(SEQ ID NO:26)、CS01-SC2-NA(SEQ ID NO:27)、CS23-SC1-NA(SEQ ID NO:28)或CS23-SC2-NA(SEQ ID NO:29))具有至少99%的同一性。

[0062] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列与相应全长多核苷酸序列(例如CS04-SC1-NA(SEQ ID NO:9)、CS04-SC2-NA(SEQ ID NO:11)、CS01-SC1-NA(SEQ ID NO:26)、CS01-SC2-NA(SEQ ID NO:27)、CS23-SC1-NA(SEQ ID NO:28)或CS23-SC2-NA(SEQ ID NO:29))具有至少99.5%的同一性。

[0063] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列与相应全长多核苷酸序列(例如CS04-SC1-NA(SEQ ID NO:9)、CS04-SC2-NA(SEQ ID NO:11)、CS01-SC1-NA(SEQ ID NO:26)、CS01-SC2-NA(SEQ ID NO:27)、CS23-SC1-NA(SEQ ID NO:28)或CS23-SC2-NA(SEQ ID NO:29))具有至少99.9%的同一性。

[0064] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列为CS04-SC1-NA(SEQ ID NO:9)。

[0065] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列为CS04-SC2-NA(SEQ ID NO:11)。

[0066] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列为CS01-SC1-NA(SEQ ID NO:26)。

[0067] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列为CS01-SC2-NA(SEQ ID NO:27)。

[0068] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列为CS23-SC1-NA(SEQ ID NO:28)。

[0069] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列为CS23-SC2-NA(SEQ ID NO:29)。

[0070] 在上文所描述的多核苷酸的一个实施方案中,核苷酸序列与选自由以下各项组成的组的序列具有至少95%的同一性:CS01-FL-NA、CS01-HC-NA、CS01-LC-NA、CS04-FL-NA、CS04-HC-NA、CS04-LC-NA、CS23-FL-NA、CS23-HC-NA、CS23-LC-NA、CS01m1-FL-NA、CS01m2-FL-NA、CS01m3-FL-NA、CS01m4-FL-NA、CS01m12-FL-NA、CS01m13-FL-NA、CS01m23-FL-NA、CS01m24-FL-NA、CS01m34-FL-NA、CS01m123-FL-NA、CS01m234-FL-NA、CS04m1-FL-NA、CS04m2-FL-NA、CS04m3-FL-NA、CS04m4-FL-NA、CS04m12-FL-NA、CS04m13-FL-NA、CS04m23-FL-NA、CS04m24-FL-NA、CS04m34-FL-NA、CS04m123-FL-NA、CS04m234-FL-NA、CS23m1-FL-NA、CS23m2-FL-NA、CS23m3-FL-NA、CS23m4-FL-NA、CS23m12-FL-NA、CS23m13-FL-NA、CS23m23-FL-NA、CS23m24-FL-NA、CS23m34-FL-NA、CS23m123-FL-NA、CS23m234-FL-NA、CS01-SC1-NA、

CS04-SC1-NA、CS23-SC1-NA、CS01m1-SC1-NA、CS01m2-SC1-NA、CS01m3-SC1-NA、CS01m4-SC1-NA、CS01m12-SC1-NA、CS01m13-SC1-NA、CS01m23-SC1-NA、CS01m24-SC1-NA、CS01m34-SC1-NA、CS01m123-SC1-NA、CS01m234-SC1-NA、CS04m1-SC1-NA、CS04m2-SC1-NA、CS04m3-SC1-NA、CS04m4-SC1-NA、CS04m12-SC1-NA、CS04m13-SC1-NA、CS04m23-SC1-NA、CS04m24-SC1-NA、CS04m34-SC1-NA、CS04m123-SC1-NA、CS04m234-SC1-NA、CS23m1-SC1-NA、CS23m2-SC1-NA、CS23m3-SC1-NA、CS23m4-SC1-NA、CS23m12-SC1-NA、CS23m13-SC1-NA、CS23m23-SC1-NA、CS23m24-SC1-NA、CS23m34-SC1-NA、CS23m123-SC1-NA、CS23m234-SC1-NA、CS01-SC2-NA、CS04-SC2-NA、CS23-SC2-NA、CS01m1-SC2-NA、CS01m2-SC2-NA、CS01m3-SC2-NA、CS01m4-SC2-NA、CS01m12-SC2-NA、CS01m13-SC2-NA、CS01m23-SC2-NA、CS01m24-SC2-NA、CS01m34-SC2-NA、CS01m123-SC2-NA、CS01m234-SC2-NA、CS04m1-SC2-NA、CS04m2-SC2-NA、CS04m3-SC2-NA、CS04m4-SC2-NA、CS04m12-SC2-NA、CS04m13-SC2-NA、CS04m23-SC2-NA、CS04m24-SC2-NA、CS04m34-SC2-NA、CS04m123-SC2-NA、CS04m234-SC2-NA、CS23m1-SC2-NA、CS23m2-SC2-NA、CS23m3-SC2-NA、CS23m4-SC2-NA、CS23m12-SC2-NA、CS23m13-SC2-NA、CS23m23-SC2-NA、CS23m24-SC2-NA、CS23m34-SC2-NA、CS23m123-SC2-NA以及CS23m234-SC2-NA。

[0071] 在上文所描述的多核苷酸的一个实施方案中，核苷酸序列与选自由以下各项组成的组的序列具有至少96％的同一性：CS01-FL-NA、CS01-HC-NA、CS01-LC-NA、CS04-FL-NA、CS04-HC-NA、CS04-LC-NA、CS23-FL-NA、CS23-HC-NA、CS23-LC-NA、CS01m1-FL-NA、CS01m2-FL-NA、CS01m3-FL-NA、CS01m4-FL-NA、CS01m12-FL-NA、CS01m13-FL-NA、CS01m23-FL-NA、CS01m24-FL-NA、CS01m34-FL-NA、CS01m123-FL-NA、CS01m234-FL-NA、CS04m1-FL-NA、CS04m2-FL-NA、CS04m3-FL-NA、CS04m4-FL-NA、CS04m12-FL-NA、CS04m13-FL-NA、CS04m23-FL-NA、CS04m24-FL-NA、CS04m34-FL-NA、CS04m123-FL-NA、CS04m234-FL-NA、CS23m1-FL-NA、CS23m2-FL-NA、CS23m3-FL-NA、CS23m4-FL-NA、CS23m12-FL-NA、CS23m13-FL-NA、CS23m23-FL-NA、CS23m24-FL-NA、CS23m34-FL-NA、CS23m123-FL-NA、CS23m234-FL-NA、CS01-SC1-NA、CS04-SC1-NA、CS23-SC1-NA、CS01m1-SC1-NA、CS01m2-SC1-NA、CS01m3-SC1-NA、CS01m4-SC1-NA、CS01m12-SC1-NA、CS01m13-SC1-NA、CS01m23-SC1-NA、CS01m24-SC1-NA、CS01m34-SC1-NA、CS01m123-SC1-NA、CS01m234-SC1-NA、CS04m1-SC1-NA、CS04m2-SC1-NA、CS04m3-SC1-NA、CS04m4-SC1-NA、CS04m12-SC1-NA、CS04m13-SC1-NA、CS04m23-SC1-NA、CS04m24-SC1-NA、CS04m34-SC1-NA、CS04m123-SC1-NA、CS04m234-SC1-NA、CS23m1-SC1-NA、CS23m2-SC1-NA、CS23m3-SC1-NA、CS23m4-SC1-NA、CS23m12-SC1-NA、CS23m13-SC1-NA、CS23m23-SC1-NA、CS23m24-SC1-NA、CS23m34-SC1-NA、CS23m123-SC1-NA、CS23m234-SC1-NA、CS01-SC2-NA、CS04-SC2-NA、CS23-SC2-NA、CS01m1-SC2-NA、CS01m2-SC2-NA、CS01m3-SC2-NA、CS01m4-SC2-NA、CS01m12-SC2-NA、CS01m13-SC2-NA、CS01m23-SC2-NA、CS01m24-SC2-NA、CS01m34-SC2-NA、CS01m123-SC2-NA、CS01m234-SC2-NA、CS04m1-SC2-NA、CS04m2-SC2-NA、CS04m3-SC2-NA、CS04m4-SC2-NA、CS04m12-SC2-NA、CS04m13-SC2-NA、CS04m23-SC2-NA、CS04m24-SC2-NA、CS04m34-SC2-NA、CS04m123-SC2-NA、CS04m234-SC2-NA、CS23m1-SC2-NA、CS23m2-SC2-NA、CS23m3-SC2-NA、CS23m4-SC2-NA、CS23m12-SC2-NA、CS23m13-SC2-NA、CS23m23-SC2-NA、CS23m24-SC2-NA、CS23m34-SC2-NA、CS23m123-SC2-NA以及CS23m234-SC2-NA。

[0072] 在上文所描述的多核苷酸的一个实施方案中，核苷酸序列与选自由以下各项组成的组的序列具有至少97％的同一性：CS01-FL-NA、CS01-HC-NA、CS01-LC-NA、CS04-FL-NA、

CS04-HC-NA、CS04-LC-NA、CS23-FL-NA、CS23-HC-NA、CS23-LC-NA、CS01m1-FL-NA、CS01m2-FL-NA、CS01m3-FL-NA、CS01m4-FL-NA、CS01m12-FL-NA、CS01m13-FL-NA、CS01m23-FL-NA、CS01m24-FL-NA、CS01m34-FL-NA、CS01m123-FL-NA、CS01m234-FL-NA、CS04m1-FL-NA、CS04m2-FL-NA、CS04m3-FL-NA、CS04m4-FL-NA、CS04m12-FL-NA、CS04m13-FL-NA、CS04m23-FL-NA、CS04m24-FL-NA、CS04m34-FL-NA、CS04m123-FL-NA、CS04m234-FL-NA、CS23m1-FL-NA、CS23m2-FL-NA、CS23m3-FL-NA、CS23m4-FL-NA、CS23m12-FL-NA、CS23m13-FL-NA、CS23m23-FL-NA、CS23m24-FL-NA、CS23m34-FL-NA、CS23m123-FL-NA、CS23m234-FL-NA、CS01-SC1-NA、CS04-SC1-NA、CS23-SC1-NA、CS01m1-SC1-NA、CS01m2-SC1-NA、CS01m3-SC1-NA、CS01m4-SC1-NA、CS01m12-SC1-NA、CS01m13-SC1-NA、CS01m23-SC1-NA、CS01m24-SC1-NA、CS01m34-SC1-NA、CS01m123-SC1-NA、CS01m234-SC1-NA、CS04m1-SC1-NA、CS04m2-SC1-NA、CS04m3-SC1-NA、CS04m4-SC1-NA、CS04m12-SC1-NA、CS04m13-SC1-NA、CS04m23-SC1-NA、CS04m24-SC1-NA、CS04m34-SC1-NA、CS04m123-SC1-NA、CS04m234-SC1-NA、CS23m1-SC1-NA、CS23m2-SC1-NA、CS23m3-SC1-NA、CS23m4-SC1-NA、CS23m12-SC1-NA、CS23m13-SC1-NA、CS23m23-SC1-NA、CS23m24-SC1-NA、CS23m34-SC1-NA、CS23m123-SC1-NA、CS23m234-SC1-NA、CS01-SC2-NA、CS04-SC2-NA、CS23-SC2-NA、CS01m1-SC2-NA、CS01m2-SC2-NA、CS01m3-SC2-NA、CS01m4-SC2-NA、CS01m12-SC2-NA、CS01m13-SC2-NA、CS01m23-SC2-NA、CS01m24-SC2-NA、CS01m34-SC2-NA、CS01m123-SC2-NA、CS01m234-SC2-NA、CS04m1-SC2-NA、CS04m2-SC2-NA、CS04m3-SC2-NA、CS04m4-SC2-NA、CS04m12-SC2-NA、CS04m13-SC2-NA、CS04m23-SC2-NA、CS04m24-SC2-NA、CS04m34-SC2-NA、CS04m123-SC2-NA、CS04m234-SC2-NA、CS23m1-SC2-NA、CS23m2-SC2-NA、CS23m3-SC2-NA、CS23m4-SC2-NA、CS23m12-SC2-NA、CS23m13-SC2-NA、CS23m23-SC2-NA、CS23m24-SC2-NA、CS23m34-SC2-NA、CS23m123-SC2-NA以及CS23m234-SC2-NA。

[0073] 在上文所描述的多核苷酸的一个实施方案中，核苷酸序列与选自由以下各项组成的组的序列具有至少98％的同一性：CS01-FL-NA、CS01-HC-NA、CS01-LC-NA、CS04-FL-NA、CS04-HC-NA、CS04-LC-NA、CS23-FL-NA、CS23-HC-NA、CS23-LC-NA、CS01m1-FL-NA、CS01m2-FL-NA、CS01m3-FL-NA、CS01m4-FL-NA、CS01m12-FL-NA、CS01m13-FL-NA、CS01m23-FL-NA、CS01m24-FL-NA、CS01m34-FL-NA、CS01m123-FL-NA、CS01m234-FL-NA、CS04m1-FL-NA、CS04m2-FL-NA、CS04m3-FL-NA、CS04m4-FL-NA、CS04m12-FL-NA、CS04m13-FL-NA、CS04m23-FL-NA、CS04m24-FL-NA、CS04m34-FL-NA、CS04m123-FL-NA、CS04m234-FL-NA、CS23m1-FL-NA、CS23m2-FL-NA、CS23m3-FL-NA、CS23m4-FL-NA、CS23m12-FL-NA、CS23m13-FL-NA、CS23m23-FL-NA、CS23m24-FL-NA、CS23m34-FL-NA、CS23m123-FL-NA、CS23m234-FL-NA、CS01-SC1-NA、CS04-SC1-NA、CS23-SC1-NA、CS01m1-SC1-NA、CS01m2-SC1-NA、CS01m3-SC1-NA、CS01m4-SC1-NA、CS01m12-SC1-NA、CS01m13-SC1-NA、CS01m23-SC1-NA、CS01m24-SC1-NA、CS01m34-SC1-NA、CS01m123-SC1-NA、CS01m234-SC1-NA、CS04m1-SC1-NA、CS04m2-SC1-NA、CS04m3-SC1-NA、CS04m4-SC1-NA、CS04m12-SC1-NA、CS04m13-SC1-NA、CS04m23-SC1-NA、CS04m24-SC1-NA、CS04m34-SC1-NA、CS04m123-SC1-NA、CS04m234-SC1-NA、CS23m1-SC1-NA、CS23m2-SC1-NA、CS23m3-SC1-NA、CS23m4-SC1-NA、CS23m12-SC1-NA、CS23m13-SC1-NA、CS23m23-SC1-NA、CS23m24-SC1-NA、CS23m34-SC1-NA、CS23m123-SC1-NA、CS23m234-SC1-NA、CS01-SC2-NA、CS04-SC2-NA、CS23-SC2-NA、CS01m1-SC2-NA、CS01m2-SC2-NA、CS01m3-SC2-NA、CS01m4-SC2-NA、CS01m12-SC2-NA、CS01m13-SC2-NA、CS01m23-SC2-NA、CS01m24-SC2-NA、CS01m34-SC2-

NA、CS01m123-SC2-NA、CS01m234-SC2-NA、CS04m1-SC2-NA、CS04m2-SC2-NA、CS04m3-SC2-NA、CS04m4-SC2-NA、CS04m12-SC2-NA、CS04m13-SC2-NA、CS04m23-SC2-NA、CS04m24-SC2-NA、CS04m34-SC2-NA、CS04m123-SC2-NA、CS04m234-SC2-NA、CS23m1-SC2-NA、CS23m2-SC2-NA、CS23m3-SC2-NA、CS23m4-SC2-NA、CS23m12-SC2-NA、CS23m13-SC2-NA、CS23m23-SC2-NA、CS23m24-SC2-NA、CS23m34-SC2-NA、CS23m123-SC2-NA以及CS23m234-SC2-NA。

[0074] 在上文所描述的多核苷酸的一个实施方案中，核苷酸序列与选自由以下各项组成的组的序列具有至少99％的同一性：CS01-FL-NA、CS01-HC-NA、CS01-LC-NA、CS04-FL-NA、CS04-HC-NA、CS04-LC-NA、CS23-FL-NA、CS23-HC-NA、CS23-LC-NA、CS01m1-FL-NA、CS01m2-FL-NA、CS01m3-FL-NA、CS01m4-FL-NA、CS01m12-FL-NA、CS01m13-FL-NA、CS01m23-FL-NA、CS01m24-FL-NA、CS01m34-FL-NA、CS01m123-FL-NA、CS01m234-FL-NA、CS04m1-FL-NA、CS04m2-FL-NA、CS04m3-FL-NA、CS04m4-FL-NA、CS04m12-FL-NA、CS04m13-FL-NA、CS04m23-FL-NA、CS04m24-FL-NA、CS04m34-FL-NA、CS04m123-FL-NA、CS04m234-FL-NA、CS23m1-FL-NA、CS23m2-FL-NA、CS23m3-FL-NA、CS23m4-FL-NA、CS23m12-FL-NA、CS23m13-FL-NA、CS23m23-FL-NA、CS23m24-FL-NA、CS23m34-FL-NA、CS23m123-FL-NA、CS23m234-FL-NA、CS01-SC1-NA、CS04-SC1-NA、CS23-SC1-NA、CS01m1-SC1-NA、CS01m2-SC1-NA、CS01m3-SC1-NA、CS01m4-SC1-NA、CS01m12-SC1-NA、CS01m13-SC1-NA、CS01m23-SC1-NA、CS01m24-SC1-NA、CS01m34-SC1-NA、CS01m123-SC1-NA、CS01m234-SC1-NA、CS04m1-SC1-NA、CS04m2-SC1-NA、CS04m3-SC1-NA、CS04m4-SC1-NA、CS04m12-SC1-NA、CS04m13-SC1-NA、CS04m23-SC1-NA、CS04m24-SC1-NA、CS04m34-SC1-NA、CS04m123-SC1-NA、CS04m234-SC1-NA、CS23m1-SC1-NA、CS23m2-SC1-NA、CS23m3-SC1-NA、CS23m4-SC1-NA、CS23m12-SC1-NA、CS23m13-SC1-NA、CS23m23-SC1-NA、CS23m24-SC1-NA、CS23m34-SC1-NA、CS23m123-SC1-NA、CS23m234-SC1-NA、CS01-SC2-NA、CS04-SC2-NA、CS23-SC2-NA、CS01m1-SC2-NA、CS01m2-SC2-NA、CS01m3-SC2-NA、CS01m4-SC2-NA、CS01m12-SC2-NA、CS01m13-SC2-NA、CS01m23-SC2-NA、CS01m24-SC2-NA、CS01m34-SC2-NA、CS01m123-SC2-NA、CS01m234-SC2-NA、CS04m1-SC2-NA、CS04m2-SC2-NA、CS04m3-SC2-NA、CS04m4-SC2-NA、CS04m12-SC2-NA、CS04m13-SC2-NA、CS04m23-SC2-NA、CS04m24-SC2-NA、CS04m34-SC2-NA、CS04m123-SC2-NA、CS04m234-SC2-NA、CS23m1-SC2-NA、CS23m2-SC2-NA、CS23m3-SC2-NA、CS23m4-SC2-NA、CS23m12-SC2-NA、CS23m13-SC2-NA、CS23m23-SC2-NA、CS23m24-SC2-NA、CS23m34-SC2-NA、CS23m123-SC2-NA以及CS23m234-SC2-NA。

[0075] 在上文所描述的多核苷酸的一个实施方案中，核苷酸序列与选自由以下各项组成的组的序列具有至少99.5％的同一性：CS01-FL-NA、CS01-HC-NA、CS01-LC-NA、CS04-FL-NA、CS04-HC-NA、CS04-LC-NA、CS23-FL-NA、CS23-HC-NA、CS23-LC-NA、CS01m1-FL-NA、CS01m2-FL-NA、CS01m3-FL-NA、CS01m4-FL-NA、CS01m12-FL-NA、CS01m13-FL-NA、CS01m23-FL-NA、CS01m24-FL-NA、CS01m34-FL-NA、CS01m123-FL-NA、CS01m234-FL-NA、CS04m1-FL-NA、CS04m2-FL-NA、CS04m3-FL-NA、CS04m4-FL-NA、CS04m12-FL-NA、CS04m13-FL-NA、CS04m23-FL-NA、CS04m24-FL-NA、CS04m34-FL-NA、CS04m123-FL-NA、CS04m234-FL-NA、CS23m1-FL-NA、CS23m2-FL-NA、CS23m3-FL-NA、CS23m4-FL-NA、CS23m12-FL-NA、CS23m13-FL-NA、CS23m23-FL-NA、CS23m24-FL-NA、CS23m34-FL-NA、CS23m123-FL-NA、CS23m234-FL-NA、CS01-SC1-NA、CS04-SC1-NA、CS23-SC1-NA、CS01m1-SC1-NA、CS01m2-SC1-NA、CS01m3-SC1-NA、CS01m4-SC1-NA、CS01m12-SC1-NA、CS01m13-SC1-NA、CS01m23-SC1-NA、CS01m24-SC1-NA、CS01m34-SC1-

NA、CS01m123-SC1-NA、CS01m234-SC1-NA、CS04m1-SC1-NA、CS04m2-SC1-NA、CS04m3-SC1-NA、CS04m4-SC1-NA、CS04m12-SC1-NA、CS04m13-SC1-NA、CS04m23-SC1-NA、CS04m24-SC1-NA、CS04m34-SC1-NA、CS04m123-SC1-NA、CS04m234-SC1-NA、CS23m1-SC1-NA、CS23m2-SC1-NA、CS23m3-SC1-NA、CS23m4-SC1-NA、CS23m12-SC1-NA、CS23m13-SC1-NA、CS23m23-SC1-NA、CS23m24-SC1-NA、CS23m34-SC1-NA、CS23m123-SC1-NA、CS23m234-SC1-NA、CS01-SC2-NA、CS04-SC2-NA、CS23-SC2-NA、CS01m1-SC2-NA、CS01m2-SC2-NA、CS01m3-SC2-NA、CS01m4-SC2-NA、CS01m12-SC2-NA、CS01m13-SC2-NA、CS01m23-SC2-NA、CS01m24-SC2-NA、CS01m34-SC2-NA、CS01m123-SC2-NA、CS01m234-SC2-NA、CS04m1-SC2-NA、CS04m2-SC2-NA、CS04m3-SC2-NA、CS04m4-SC2-NA、CS04m12-SC2-NA、CS04m13-SC2-NA、CS04m23-SC2-NA、CS04m24-SC2-NA、CS04m34-SC2-NA、CS04m123-SC2-NA、CS04m234-SC2-NA、CS23m1-SC2-NA、CS23m2-SC2-NA、CS23m3-SC2-NA、CS23m4-SC2-NA、CS23m12-SC2-NA、CS23m13-SC2-NA、CS23m23-SC2-NA、CS23m24-SC2-NA、CS23m34-SC2-NA、CS23m123-SC2-NA以及CS23m234-SC2-NA。

[0076] 在上文所描述的多核苷酸的一个实施方案中，核苷酸序列与选自由以下各项组成的组的序列具有至少99.5％的同一性：CS01-FL-NA、CS01-HC-NA、CS01-LC-NA、CS04-FL-NA、CS04-HC-NA、CS04-LC-NA、CS23-FL-NA、CS23-HC-NA、CS23-LC-NA、CS01m1-FL-NA、CS01m2-FL-NA、CS01m3-FL-NA、CS01m4-FL-NA、CS01m12-FL-NA、CS01m13-FL-NA、CS01m23-FL-NA、CS01m24-FL-NA、CS01m34-FL-NA、CS01m123-FL-NA、CS01m234-FL-NA、CS04m1-FL-NA、CS04m2-FL-NA、CS04m3-FL-NA、CS04m4-FL-NA、CS04m12-FL-NA、CS04m13-FL-NA、CS04m23-FL-NA、CS04m24-FL-NA、CS04m34-FL-NA、CS04m123-FL-NA、CS04m234-FL-NA、CS23m1-FL-NA、CS23m2-FL-NA、CS23m3-FL-NA、CS23m4-FL-NA、CS23m12-FL-NA、CS23m13-FL-NA、CS23m23-FL-NA、CS23m24-FL-NA、CS23m34-FL-NA、CS23m123-FL-NA、CS23m234-FL-NA、CS01-SC1-NA、CS04-SC1-NA、CS23-SC1-NA、CS01m1-SC1-NA、CS01m2-SC1-NA、CS01m3-SC1-NA、CS01m4-SC1-NA、CS01m12-SC1-NA、CS01m13-SC1-NA、CS01m23-SC1-NA、CS01m24-SC1-NA、CS01m34-SC1-NA、CS01m123-SC1-NA、CS01m234-SC1-NA、CS04m1-SC1-NA、CS04m2-SC1-NA、CS04m3-SC1-NA、CS04m4-SC1-NA、CS04m12-SC1-NA、CS04m13-SC1-NA、CS04m23-SC1-NA、CS04m24-SC1-NA、CS04m34-SC1-NA、CS04m123-SC1-NA、CS04m234-SC1-NA、CS23m1-SC1-NA、CS23m2-SC1-NA、CS23m3-SC1-NA、CS23m4-SC1-NA、CS23m12-SC1-NA、CS23m13-SC1-NA、CS23m23-SC1-NA、CS23m24-SC1-NA、CS23m34-SC1-NA、CS23m123-SC1-NA、CS23m234-SC1-NA、CS01-SC2-NA、CS04-SC2-NA、CS23-SC2-NA、CS01m1-SC2-NA、CS01m2-SC2-NA、CS01m3-SC2-NA、CS01m4-SC2-NA、CS01m12-SC2-NA、CS01m13-SC2-NA、CS01m23-SC2-NA、CS01m24-SC2-NA、CS01m34-SC2-NA、CS01m123-SC2-NA、CS01m234-SC2-NA、CS04m1-SC2-NA、CS04m2-SC2-NA、CS04m3-SC2-NA、CS04m4-SC2-NA、CS04m12-SC2-NA、CS04m13-SC2-NA、CS04m23-SC2-NA、CS04m24-SC2-NA、CS04m34-SC2-NA、CS04m123-SC2-NA、CS04m234-SC2-NA、CS23m1-SC2-NA、CS23m2-SC2-NA、CS23m3-SC2-NA、CS23m4-SC2-NA、CS23m12-SC2-NA、CS23m13-SC2-NA、CS23m23-SC2-NA、CS23m24-SC2-NA、CS23m34-SC2-NA、CS23m123-SC2-NA以及CS23m234-SC2-NA。

[0077] 在上文所描述的多核苷酸的一个实施方案中，核苷酸序列是选自由以下各项组成的组：CS01-FL-NA、CS01-HC-NA、CS01-LC-NA、CS04-FL-NA、CS04-HC-NA、CS04-LC-NA、CS23-FL-NA、CS23-HC-NA、CS23-LC-NA、CS01m1-FL-NA、CS01m2-FL-NA、CS01m3-FL-NA、CS01m4-FL-NA、CS01m12-FL-NA、CS01m13-FL-NA、CS01m23-FL-NA、CS01m24-FL-NA、CS01m34-FL-NA、

CS01m123-FL-NA、CS01m234-FL-NA、CS04m1-FL-NA、CS04m2-FL-NA、CS04m3-FL-NA、CS04m4-FL-NA、CS04m12-FL-NA、CS04m13-FL-NA、CS04m23-FL-NA、CS04m24-FL-NA、CS04m34-FL-NA、CS04m123-FL-NA、CS04m234-FL-NA、CS23m1-FL-NA、CS23m2-FL-NA、CS23m3-FL-NA、CS23m4-FL-NA、CS23m12-FL-NA、CS23m13-FL-NA、CS23m23-FL-NA、CS23m24-FL-NA、CS23m34-FL-NA、CS23m123-FL-NA、CS23m234-FL-NA、CS01-SC1-NA、CS04-SC1-NA、CS23-SC1-NA、CS01m1-SC1-NA、CS01m2-SC1-NA、CS01m3-SC1-NA、CS01m4-SC1-NA、CS01m12-SC1-NA、CS01m13-SC1-NA、CS01m23-SC1-NA、CS01m24-SC1-NA、CS01m34-SC1-NA、CS01m123-SC1-NA、CS01m234-SC1-NA、CS04m1-SC1-NA、CS04m2-SC1-NA、CS04m3-SC1-NA、CS04m4-SC1-NA、CS04m12-SC1-NA、CS04m13-SC1-NA、CS04m23-SC1-NA、CS04m24-SC1-NA、CS04m34-SC1-NA、CS04m123-SC1-NA、CS04m234-SC1-NA、CS23m1-SC1-NA、CS23m2-SC1-NA、CS23m3-SC1-NA、CS23m4-SC1-NA、CS23m12-SC1-NA、CS23m13-SC1-NA、CS23m23-SC1-NA、CS23m24-SC1-NA、CS23m34-SC1-NA、CS23m123-SC1-NA、CS23m234-SC1-NA、CS01-SC2-NA、CS04-SC2-NA、CS23-SC2-NA、CS01m1-SC2-NA、CS01m2-SC2-NA、CS01m3-SC2-NA、CS01m4-SC2-NA、CS01m12-SC2-NA、CS01m13-SC2-NA、CS01m23-SC2-NA、CS01m24-SC2-NA、CS01m34-SC2-NA、CS01m123-SC2-NA、CS01m234-SC2-NA、CS04m1-SC2-NA、CS04m2-SC2-NA、CS04m3-SC2-NA、CS04m4-SC2-NA、CS04m12-SC2-NA、CS04m13-SC2-NA、CS04m23-SC2-NA、CS04m24-SC2-NA、CS04m34-SC2-NA、CS04m123-SC2-NA、CS04m234-SC2-NA、CS23m1-SC2-NA、CS23m2-SC2-NA、CS23m3-SC2-NA、CS23m4-SC2-NA、CS23m12-SC2-NA、CS23m13-SC2-NA、CS23m23-SC2-NA、CS23m24-SC2-NA、CS23m34-SC2-NA、CS23m123-SC2-NA以及CS23m234-SC2-NA。

[0078] 在上文所描述的多核苷酸的一个实施方案中，编码的因子VIII多肽包含位于两个连续氨基酸之间的糖基化多肽。

[0079] 在上文所描述的多核苷酸的一个实施方案中，编码的多肽接头包括氨基酸序列与选自由以下各项组成的组的糖基化肽具有至少92％同一性的糖基化肽：NG1-AA、NG4-AA、NG5-AA、NG6-AA、NG7-AA、NG9-AA、NG10-AA、NG16-AA、NG17-AA、NG18-AA、NG19-AA、NG20-AA、NG21-AA以及NGV-AA。

[0080] 在上文所描述的多核苷酸的一个实施方案中，编码的多肽接头包含氨基酸序列选自由以下各项组成的组的糖基化肽：NG1-AA、NG4-AA、NG5-AA、NG6-AA、NG7-AA、NG9-AA、NG10-AA、NG16-AA、NG17-AA、NG18-AA、NG19-AA、NG20-AA、NG21-AA以及NGV-AA。

[0081] 在上文所描述的多核苷酸的一个实施方案中，糖基化肽是由核苷酸序列与选自由以下各项组成的组的序列具有至少95％同一性的多核苷酸编码：NG1-NA、NG4-NA、NG5-NA、NG6-NA、NG7-NA、NG9-NA、NG10-NA、NG16-NA、NG17-NA、NG18-NA、NG19-NA、NG20-NA、NG21-NA以及NGV-NA。

[0082] 在上文所描述的多核苷酸的一个实施方案中，糖基化肽是由核苷酸序列选自以下各项中的一种的多核苷酸编码：NG1-NA、NG4-NA、NG5-NA、NG6-NA、NG7-NA、NG9-NA、NG10-NA、NG16-NA、NG17-NA、NG18-NA、NG19-NA、NG20-NA、NG21-NA以及NGV-NA。

[0083] 在上文所描述的多核苷酸的一个实施方案中，多肽接头是由与选自由以下各项组成的组的序列具有至少95％同一性的第三核苷酸序列编码：BDLNG1-NA、BDLNG3-NA、BDLNG5-NA、BDLNG6-NA、BDLNG9-NA、BDLNG10-NA、BDLNG16-NA、BDLNG17-NA、BDLNG18-NA、BDLNG19-NA、BDLNG20-NA以及BDLNG21-NA。

[0084] 在上文所描述的多核苷酸的一个实施方案中,编码的因子VIII多肽相对于FVIII-FL-AA(SEQ ID NO:19)包括F328S(SPI,F309SSPE)氨基酸取代。

[0085] 在上文所描述的多核苷酸的一个实施方案中,编码的因子VIII多肽相对于FVIII-FL-AA(SEQ ID NO:19)包括I105V、A127S、G151K、M166T以及L171P(SPI;分别为I86V、A108S、G132K、M147T以及L152P,SPE)氨基酸取代。

[0086] 在上文所描述的多核苷酸的一个实施方案中,编码的因子VIII多肽包括a)氨基酸AIEPR755-759的缺失,相对于FVIII-FL-AA(SEQ ID NO:19);以及b)N754后的氨基酸TTYVNRSL(SEQ ID NO:33)的插入,相对于FVIII-FL-AA(SEQ ID NO:19)。在一些实施方案(例如其中编码的FVIII分子包括野生型B-结构域的N端区的一部分)中,编码的因子VIII多肽还相对于FVIII-FL-AA(SEQ ID NO:19)包括氨基酸SF760-761的缺失。

[0087] 在上文所描述的多核苷酸的一个实施方案中,编码的因子VIII多肽包括a)F328S(SPI;F309S SPE)氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19),以及b)C1918G和C1922G(SPI;分别为C1899G和C1903 SPE)氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19)。

[0088] 在上文所描述的多核苷酸的一个实施方案中,编码的因子VIII多肽包括a)F328S(SPI;F309S SPE)氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19),以及b)I105V、A127S、G151K、M166T以及L171P(SPI;分别为I86V、A108S、G132K、M147T以及L152P,SPE)氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19)。

[0089] 在上文所描述的多核苷酸的一个实施方案中,编码的因子VIII多肽包括a)F328S氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19),b)氨基酸AIEPR755-759的缺失,相对于FVIII-FL-AA(SEQ ID NO:19)以及c)N754后的氨基酸TTYVNRSL(SEQ ID NO:33)的插入,相对于FVIII-FL-AA(SEQ ID NO:19)。在一些实施方案(例如其中编码的FVIII分子包括野生型B-结构域的N端区的一部分)中,编码的因子VIII多肽还相对于FVIII-FL-AA(SEQ ID NO:19)包括氨基酸SF760-761的缺失。

[0090] 在上文所描述的多核苷酸的一个实施方案中,编码的因子VIII多肽包括a)I105V、A127S、G151K、M166T以及L171P氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19),b)氨基酸AIEPR755-759的缺失,相对于FVIII-FL-AA(SEQ ID NO:19)以及c)N754后的氨基酸TTYVNRSL(SEQ ID NO:33)的插入,相对于FVIII-FL-AA(SEQ ID NO:19)。在一些实施方案(例如其中编码的FVIII分子包括野生型B-结构域的N端区的一部分)中,编码的因子VIII多肽还相对于FVIII-FL-AA(SEQ ID NO:19)包括氨基酸SF760-761的缺失。

[0091] 在上文所描述的多核苷酸的一个实施方案中,编码的因子VIII多肽包括a)F328S氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19),b)C1918G和C1922G氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19)以及c)I105V、A127S、G151K、M166T以及L171P氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19)。

[0092] 在上文所描述的多核苷酸的一个实施方案中,编码的因子VIII多肽包括a)F328S氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19),b)C1918G和C1922G氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19),c)氨基酸AIEPR755-759的缺失,相对于FVIII-FL-AA(SEQ ID NO:19)以及d)N754后的氨基酸TTYVNRSL(SEQ ID NO:33)的插入,相对于FVIII-FL-AA(SEQ ID NO:19)。在一些实施方案(例如其中编码的FVIII分子包括野生型B-结构域的N端区的一

部分)中,编码的因子VIII多肽还相对于FVIII-FL-AA(SEQ ID NO:19)包括氨基酸SF760-761的缺失。

[0093] 在上文所描述的多核苷酸的一个实施方案中,编码的因子VIII多肽包括a)I105V、A127S、G151K、M166T以及L171P氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19),b)F328S氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19),c)C1918G和C1922G氨基酸取代,相对于FVIII-FL-AA(SEQ ID NO:19),d)氨基酸AIEPR755-759的缺失,相对于FVIII-FL-AA(SEQ ID NO:19)以及e)N754后的氨基酸TTYVNRSL(SEQ ID NO:33)的插入,相对于FVIII-FL-AA(SEQ ID NO:19)。在一些实施方案(例如其中编码的FVIII分子包括野生型B-结构域的N端区的一部分)中,编码的因子VIII多肽还相对于FVIII-FL-AA(SEQ ID NO:19)包括氨基酸SF760-761的缺失。

[0094] 在上文所描述的多核苷酸的一个实施方案中,多核苷酸还包括可操作地连接至编码因子VIII多肽的多核苷酸的启动子元件。

[0095] 在上文所描述的多核苷酸的一个实施方案中,多核苷酸还包括可操作地连接至编码因子VIII多肽的多核苷酸的增强子元件。

[0096] 在上文所描述的多核苷酸的一个实施方案中,多核苷酸还包括可操作地连接至编码因子VIII多肽的多核苷酸的聚腺苷酸化元件。

[0097] 在上文所描述的多核苷酸的一个实施方案中,多核苷酸还包括可操作地连接至编码因子VIII多肽的核苷酸序列的内含子。

[0098] 在上文所描述的多核苷酸的一个实施方案中,内含子是位于启动子元件与编码因子VIII多肽的核苷酸序列的翻译起始位点(例如第一个编码ATG)之间。

[0099] 在另一方面中,本公开提供一种包括如上文所描述的多核苷酸的哺乳动物基因疗法载体。

[0100] 在上文所描述的哺乳动物基因疗法载体的一个实施方案中,哺乳动物基因疗法载体为腺相关病毒(AAV)载体。

[0101] 在上文所描述的哺乳动物基因疗法载体的一个实施方案中,AAV载体为AAV-8载体。

[0102] 在另一方面中,本公开提供一种治疗血友病A的方法,所述方法包括向有需要的患者施用如上文所描述的哺乳动物基因疗法载体。

[0103] 在另一方面中,本公开提供一种用于治疗血友病A的如上文所描述的哺乳动物基因疗法载体。

[0104] 在另一方面中,本公开提供如上文所描述的哺乳动物基因疗法载体用于制造用于治疗血友病A的药物的用途。

[0105] 在另一方面中,本公开提供一种因子VIII多肽,所述因子VIII多肽包括轻链、重链以及将重链的C端连接至轻链的N端的多肽接头。因子VIII多肽的重链具有与序列CSO1-HC-AAm23至少95%同一的序列。因子VIII多肽的轻链具有与序列CSO1-LC-AAm23至少95%同一的序列。因子VIII多肽的多肽接头包括弗林蛋白酶裂解位点。因子VIII多肽包括i)I105V、A127S、G151K、M166T以及L171P氨基酸取代,ii)氨基酸AIEPR755-759的缺失,相对于FVIII-FL-AA(SEQ ID NO:19)以及iii)N754后的氨基酸TTYVNRSL(SEQ ID NO:33)的插入,相对于FVIII-FL-AA(SEQ ID NO:19)。

[0106] 在另一方面中，本公开提供一种因子VIII多肽，所述因子VIII多肽包括轻链、重链以及将重链的C端连接至轻链的N端的多肽接头。因子VIII多肽的重链具有与序列CS01-HC-AAm123至少95％同一的序列。因子VIII多肽的轻链具有与序列CS01-LC-AAm123至少95％同一的序列。因子VIII多肽的多肽接头包括弗林蛋白酶裂解位点。因子VIII多肽包括i)I105V、A127S、G151K、M166T以及L171P氨基酸取代，ii)氨基酸AIEPR755-759的缺失，相对于FVIII-FL-AA(SEQ ID NO:19)，iii)N754后的氨基酸TTYVNRSL(SEQ ID NO:33)的插入，相对于FVIII-FL-AA(SEQ ID NO:19)以及iv)F328S氨基酸取代。

[0107] 在另一方面中，本公开提供一种因子VIII多肽，所述因子VIII多肽包括轻链、重链以及将重链的C端连接至轻链的N端的多肽接头。因子VIII多肽的重链具有与序列CS01-HC-AAm234至少95％同一的序列。因子VIII多肽的轻链具有与序列CS01-LC-AAm234至少95％同一的序列。因子VIII多肽的多肽接头包括弗林蛋白酶裂解位点。因子VIII多肽包括i)I105V、A127S、G151K、M166T以及L171P氨基酸取代，ii)氨基酸AIEPR755-759的缺失，相对于FVIII-FL-AA(SEQ ID NO:19)，iii)N754后的氨基酸TTYVNRSL(SEQ ID NO:33)的插入，相对于FVIII-FL-AA(SEQ ID NO:19)以及iv)F328S/C1918G/C1922G氨基酸取代。

[0108] 在所描述的因子VIII多肽的一个实施方案中，因子VIII多肽的重链具有与相应重链序列(例如CS01-HC-AAm23、CS01-HC-AAm123或CS01-HC-AAm234)至少96％同一的序列，并且因子FVIII多肽的轻链具有与相应轻链序列(例如CS01-LC-AAm23、CS01-LC-AAm123或CS01-LC-AAm234)至少96％同一的序列。

[0109] 在所描述的因子VIII多肽的一个实施方案中，因子VIII多肽的重链具有与相应重链序列(例如CS01-HC-AAm23、CS01-HC-AAm123或CS01-HC-AAm234)至少97％同一的序列，并且因子FVIII多肽的轻链具有与相应轻链序列(例如CS01-LC-AAm23、CS01-LC-AAm123或CS01-LC-AAm234)至少97％同一的序列。

[0110] 在所描述的因子VIII多肽的一个实施方案中，因子VIII多肽的重链具有与相应重链序列(例如CS01-HC-AAm23、CS01-HC-AAm123或CS01-HC-AAm234)至少98％同一的序列，并且因子FVIII多肽的轻链具有与相应轻链序列(例如CS01-LC-AAm23、CS01-LC-AAm123或CS01-LC-AAm234)至少98％同一的序列。

[0111] 在所描述的因子VIII多肽的一个实施方案中，因子VIII多肽的重链具有与相应重链序列(例如CS01-HC-AAm23、CS01-HC-AAm123或CS01-HC-AAm234)至少99％同一的序列，并且因子FVIII多肽的轻链具有与相应轻链序列(例如CS01-LC-AAm23、CS01-LC-AAm123或CS01-LC-AAm234)至少99％同一的序列。

[0112] 在所描述的因子VIII多肽的一个实施方案中，因子VIII多肽的重链具有与相应重链序列(例如CS01-HC-AAm23、CS01-HC-AAm123或CS01-HC-AAm234)至少99.5％同一的序列，并且因子FVIII多肽的轻链具有与相应轻链序列(例如CS01-LC-AAm23、CS01-LC-AAm123或CS01-LC-AAm234)至少99.5％同一的序列。

[0113] 在所描述的因子VIII多肽的一个实施方案中，因子VIII多肽的重链具有与相应重链序列(例如CS01-HC-AAm23、CS01-HC-AAm123或CS01-HC-AAm234)相同的序列，并且因子FVIII多肽的轻链具有与相应轻链序列(例如CS01-LC-AAm23、CS01-LC-AAm123或CS01-LC-AAm234)相同的序列。

[0114] 在上文所描述的因子VIII多肽的一个实施方案中，多肽接头与BDL-SQ-AA(SEQ ID

NO:30)具有至少95%的同一性。

[0115] 在上文所描述的因子VIII多肽的一个实施方案中,多肽接头具有BDL-SQ-AA(SEQ ID NO:30)的氨基酸序列。

[0116] 在上文所描述的因子VIII多肽的一个实施方案中,多肽接头包括氨基酸序列与选自由以下各项组成的组的糖基化肽具有至少92%同一性的糖基化肽:NG1-AA、NG4-AA、NG5-AA、NG6-AA、NG7-AA、NG9-AA、NG10-AA、NG16-AA、NG17-AA、NG18-AA、NG19-AA、NG20-AA、NG21-AA以及NGV-AA。

[0117] 在上文所描述的因子VIII多肽的一个实施方案中,多肽接头包括选自由以下各项组成的组的糖基化肽:NG1-AA、NG4-AA、NG5-AA、NG6-AA、NG7-AA、NG9-AA、NG10-AA、NG16-AA、NG17-AA、NG18-AA、NG19-AA、NG20-AA、NG21-AA以及NGV-AA。

[0118] 在上文所描述的因子VIII多肽的一个实施方案中,多肽接头具有与选自由以下各项组成的组的序列具有至少95%同一性的氨基酸序列:BDLNG1-AA、BDLNG3-AA、BDLNG5-AA、BDLNG6-AA、BDLNG9-AA、BDLNG10-AA、BDLNG16-AA、BDLNG17-AA、BDLNG18-AA、BDLNG19-AA、BDLNG20-AA以及BDLNG21-AA。

[0119] 在上文所描述的因子VIII多肽的一个实施方案中,多肽接头具有选自由以下各项组成的组的氨基酸序列:BDLNG1-AA、BDLNG3-AA、BDLNG5-AA、BDLNG6-AA、BDLNG9-AA、BDLNG10-NA、BDLNG16-AA、BDLNG17-AA、BDLNG18-AA、BDLNG19-AA、BDLNG20-AA以及BDLNG21-AA。

[0120] 在另一方面中,本公开提供一种具有与CS40-FL-AAm23(SEQ ID NO:104)具有至少95%同一性的氨基酸序列的因子VIII多肽。因子VIII多肽包括i)I105V、A127S、G151K、M166T以及L171P氨基酸取代,ii)氨基酸AIEPR755-759的缺失,相对于FVIII-FL-AA(SEQ ID NO:19)以及iii)N754后的氨基酸TTYVNRSL(SEQ ID NO:33)的插入,相对于FVIII-FL-AA(SEQ ID NO:19)。

[0121] 在另一方面中,本公开提供一种具有与CS40-FL-AAm123具有至少95%同一性的氨基酸序列的因子VIII多肽。因子VIII多肽包括i)I105V、A127S、G151K、M166T以及L171P氨基酸取代,ii)氨基酸AIEPR755-759的缺失,相对于FVIII-FL-AA(SEQ ID NO:19),iii)N754后的氨基酸TTYVNRSL(SEQ ID NO:33)的插入,相对于FVIII-FL-AA(SEQ ID NO:19)以及iv)F328S氨基酸取代。

[0122] 在另一方面中,本公开提供一种具有与CS40-FL-AAm234具有至少95%同一性的氨基酸序列的因子VIII多肽。因子VIII多肽包括i)I105V、A127S、G151K、M166T以及L171P氨基酸取代,ii)氨基酸AIEPR755-759的缺失,相对于FVIII-FL-AA(SEQ ID NO:19),iii)N754后的氨基酸TTYVNRSL(SEQ ID NO:33)的插入,相对于FVIII-FL-AA(SEQ ID NO:19)以及iv)F328S/C1918G/C1922G氨基酸取代。

[0123] 在所描述的因子VIII多肽的一个实施方案中,因子VIII多肽具有与相应全长序列(例如CS40-FL-AAm23(SEQ ID NO:104)、CS40-FL-AAm123或CS40-FL-AAm234)至少96%同一的序列。

[0124] 在所描述的因子VIII多肽的一个实施方案中,因子VIII多肽具有与相应全长序列(例如CS40-FL-AAm23(SEQ ID NO:104)、CS40-FL-AAm123或CS40-FL-AAm234)至少97%同一的序列。

[0125] 在所描述的因子VIII多肽的一个实施方案中,因子VIII多肽具有与相应全长序列(例如CS40-FL-AAm23(SEQ ID NO:104)、CS40-FL-AAm123或CS40-FL-AAm234)至少98%同一的序列。

[0126] 在所描述的因子VIII多肽的一个实施方案中,因子VIII多肽具有与相应全长序列(例如CS40-FL-AAm23(SEQ ID NO:104)、CS40-FL-AAm123或CS40-FL-AAm234)至少99%同一的序列。

[0127] 在所描述的因子VIII多肽的一个实施方案中,因子VIII多肽具有与相应全长序列(例如CS40-FL-AAm23(SEQ ID NO:104)、CS40-FL-AAm123或CS40-FL-AAm234)至少99.5%同一的序列。

[0128] 在所描述的因子VIII多肽的一个实施方案中,因子VIII多肽具有与相应全长序列(例如CS40-FL-AAm23(SEQ ID NO:104)、CS40-FL-AAm123或CS40-FL-AAm234)相同的序列。

[0129] 附图简述

[0130] 图1示出了野生型和ReFacto型人因子VIII蛋白构建体的示意图。

[0131] 图2A和2B示出了根据一些实施方案编码因子VIII变体的CS04密码子改变的核苷酸序列(SEQ ID NO:1)(对于全长编码序列为“CS04-FL-NA”)。

[0132] 图3示出了根据一些实施方案由CS04密码子改变的核苷酸序列编码的因子VIII变体氨基酸序列(SEQ ID NO:2)(对于全长氨基酸序列为“CS04-FL-AA”)。

[0133] 图4示出了根据一些实施方案编码因子VIII变体的重链的CS04密码子改变的核苷酸序列的一部分(SEQ ID NO:3)(“CS04-HC-NA”)。

[0134] 图5示出了根据一些实施方案编码因子VIII变体的轻链的CS04密码子改变的核苷酸序列的一部分(SEQ ID NO:4)(“CS04-LC-NA”)。

[0135] 图6示出了根据一些实施方案的B-结构域取代的接头的示例性编码序列(按出现顺序分别为SEQ ID NOS 5-7和36-48)。BDL001(SEQ ID NO:5)、BDL004(SEQ ID NO:6)以及BDL023(SEQ ID NO:7)分别为CS01、CS04以及CS23密码子改变的核苷酸序列中编码B-结构域取代的接头的相应部分。

[0136] 图7A、7B以及7C示出了根据一些实施方案含有CS04密码子改变的核苷酸序列的AAV载体序列(SEQ ID NO:8)(“CS04-AV-NA”)。

[0137] 图8A和8B示出了根据一些实施方案编码具有F328S氨基酸取代的因子VIII变体的CS01m1密码子改变的核苷酸序列(SEQ ID NO:49)(“CS01m1-FL-NA”)。

[0138] 图9A和9B示出了根据一些实施方案编码单链因子VIII变体的CS04Δ(760-1667)(SPI;CS04Δ(741-1648),SPE)密码子改变的核苷酸序列(SEQ ID NO:9)(“CS04-SC1-NA”)。

[0139] 图10示出了根据一些实施方案由CS01Δ(760-1667)(SPI;CS01Δ(741-1648),SPE)、CS04Δ(760-1667)(SPI;CS04Δ(741-1648),SPE)以及CS23Δ(760-1667)(SPI;CS23Δ(741-1648),SPE)密码子改变的核苷酸序列编码的因子VIII变体氨基酸序列(SEQ ID NO:10)(分别为“CS01-SC1-AA”、“CS04-SC1-AA”以及“CS23-SC1-AA”)。

[0140] 图11A和11B示出了根据一些实施方案编码单链因子VIII变体的CS04Δ(772-1667)(SPI;CS04Δ(753-1648),SPE)密码子改变的核苷酸序列(SEQ ID NO:11)(“CS04-SC2-NA”)。

[0141] 图12示出了根据一些实施方案由CS01Δ(772-1667)(SPI;CS01Δ(753-1648),

SPE)、CS04Δ(772-1667)(SPI;CS04Δ(753-1648),SPE)以及CS23Δ(772-1667)(SPI;CS23Δ(753-1648),SPE)密码子改变的核苷酸序列编码的因子VIII变体氨基酸序列(SEQ ID NO:12)(分别为“CS01-SC2-AA”、“CS04-SC2-AA”以及“CS23-SC2-AA”)。

[0142] 图13A和13B示出了根据一些实施方案插入B-结构域取代的接头中的示例性糖基化肽的氨基酸和核苷酸序列。“NG1”或NG1-AA”为氨基酸序列的密码,在上面一行中示出。“NG1-NA”为核酸序列的密码,在每一组的下面一行中示出。图13A和13B以SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75的形式公开氨基酸序列,并且按出现顺序全部分别以SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74的形式公开核苷酸序列。

[0143] 图14示出了野生型因子VIII B-结构域的体内N-糖基化的计算机模拟预测的结果。该图按出现顺序分别公开了SEQ ID NO 76和76-82。

[0144] 图15示出了V3肽接头的体内N-糖基化的计算机模拟预测的结果。该图按出现顺序分别公开了SEQ ID NO 83和83-89。

[0145] 图16A和16B示出了根据一些实施方案编码因子VIII变体的CS01密码子改变的核苷酸序列(SEQ ID NO:13)(“CS01-FL-NA”)。

[0146] 图17A和17B示出了根据一些实施方案编码因子VIII变体的CS08密码子改变的核苷酸序列(SEQ ID NO:14)(“CS08-FL-NA”)。

[0147] 图18A和18B示出了根据一些实施方案编码因子VIII变体的CS10密码子改变的核苷酸序列(SEQ ID NO:15)(“CS10-FL-NA”)。

[0148] 图19A和19B示出了根据一些实施方案编码因子VIII变体的CS11密码子改变的核苷酸序列(SEQ ID NO:16)(“CS11-FL-NA”)。

[0149] 图20A和20B示出了根据一些实施方案的CS40野生型ReFacto编码序列(SEQ ID NO:17)(“CS40-FL-NA”)。

[0150] 图21A和21B示出了根据一些实施方案编码因子VIII变体的CH25密码子改变的核苷酸序列(SEQ ID NO:18)(“CH25-FL-NA”)。

[0151] 图22示出了根据一些实施方案的野生型人因子VIII氨基酸序列(SEQ ID NO:19)(“FVIII-FL-AA”)。

[0152] 图23说明了克隆pCS40、pCS01、pCS04、pCS08、pCS10、pCS11以及pCh25构建体的方案,所述方案是通过将合成的Refacto型BDD-FVIIIDNA序列经由AscI和NotI限制位点插入载体骨架pCh-BB01中来进行。

[0153] 图24示出了如通过琼脂糖凝胶电泳所分析的AAV载体基因组制剂的完整性。泳道1,DNA标记物;泳道2,vCS40;泳道3,vCS01;泳道4,vCS04。AAV载体全部具有相同尺寸的基因组,以约5kb进行迁移(箭头,右侧)。左侧的刻度指示DNA片段的大小(千碱基,kb)。

[0154] 图25示出了通过PAGE和银染色对AAV载体制剂进行的蛋白质分析。泳道1,蛋白质标记物(M);泳道2,vCS40,泳道3,vCS01;以及泳道4,vCS04。构建体全部具有相同的由VP1、VP2以及VP3(右侧的箭头)组成的AAV8衣壳。左侧的刻度指示蛋白质标记物的大小(千道尔顿,kDa)。

[0155] 图26A和26B示出了根据一些实施方案编码因子VIII变体的CS23密码子改变的核苷酸序列(SEQ ID NO:20)(“CS23-FL-NA”)。

[0156] 图27示出了根据一些实施方案由CS23密码子改变的核苷酸序列编码的因子VIII变体氨基酸序列(SEQ ID NO:21)(“CS23-FL-AA”)。

[0157] 图28示出了根据一些实施方案编码因子VIII变体的重链的CS23密码子改变的核苷酸序列的一部分(SEQ ID NO:22)(“CS23-HC-NA”)。

[0158] 图29示出了根据一些实施方案编码因子VIII变体的轻链的CS23密码子改变的核苷酸序列的一部分(SEQ ID NO:23)(“CS23-LC-NA”)。

[0159] 图30A和30B示出了根据一些实施方案编码具有m1(F328S)和m3氨基酸取代的因子VIII变体的CS01m13密码子改变的核苷酸序列(SEQ ID NO:90)(“CS01-FL-NA-m13”)。

[0160] 图31A和31B示出了根据一些实施方案编码具有m2和m3突变集合的因子VIII变体的CS01m23密码子改变的核苷酸序列(SEQ ID NO:91)(“CS01-FL-NA-m23”)。

[0161] 图32A和32B示出了根据一些实施方案编码具有m3氨基酸取代的因子VIII变体的CS01m3密码子改变的核苷酸序列(SEQ ID NO:92)(“CS01-FL-NA-m3”)。

[0162] 图33A和33B示出了根据一些实施方案编码具有m2突变集合(I105V/A127S/G151K/M166T/L171P(SPI))氨基酸取代的因子VIII变体的CS01m2密码子改变的核苷酸序列(SEQ ID NO:93)(“CS01-FL-NA-m2”)。

[0163] 图34A和34B示出了根据一些实施方案编码具有m2突变体(I105V/A127S/G151K/M166T/L171P(SPI))氨基酸取代的因子VIII变体的CS04m2密码子改变的核苷酸序列(SEQ ID NO:94)(“CS01-FL-NA-m2”)。

[0164] 图35A和35B示出了根据一些实施方案编码具有m3氨基酸取代的因子VIII变体的CS04m3密码子改变的核苷酸序列(SEQ ID NO:95)(“CS04-FL-NA-m3”)。

[0165] 图36A和36B示出了根据一些实施方案编码具有m2突变体集合(I105V/A127S/G151K/M166T/L171P(SPI))和m3氨基酸取代的因子VIII变体的CS04m23密码子改变的核苷酸序列(SEQ ID NO:96)(“CS04-FL-NA-m23”)。

[0166] 图37A和37B示出了根据一些实施方案编码具有m1(F328S)氨基酸取代的因子VIII变体的CS04m1密码子改变的核苷酸序列(SEQ ID NO:97)(“CS04-FL-NA-m1”)。

[0167] 图38A和38B示出了根据一些实施方案编码具有m1和m3氨基酸取代的因子VIII变体的CS04m13密码子改变的核苷酸序列(SEQ ID NO:98)(“CS04-FL-NA-m13”)。

[0168] 图39A和39B示出了根据一些实施方案编码具有m1和m3氨基酸取代的因子VIII变体的CS23m13密码子改变的核苷酸序列(SEQ ID NO:99)(“CS23m13-FL-NA”)。

[0169] 图40A和40B示出了根据一些实施方案编码具有m3氨基酸取代的因子VIII变体的CS23m3密码子改变的核苷酸序列(SEQ ID NO:100)(“CS23-FL-NA-m3”)。

[0170] 图41A和41B示出了根据一些实施方案编码具有m2突变体集合(I105V/A127S/G151K/M166T/L171P氨基酸取代)的因子VIII变体的CS23m2密码子改变的核苷酸序列(SEQ ID NO:101)(“CS23-FL-NA-m2”)。

[0171] 图42A和42B示出了根据一些实施方案编码具有m1(F328S)氨基酸取代的因子VIII变体的CS23m1密码子改变的核苷酸序列(SEQ ID NO:102)(“CS23-FL-NA-m1”)。

[0172] 图43A和43B示出了根据一些实施方案编码具有m2突变体集合(I105V/A127S/G151K/M166T/L171P)和m3氨基酸取代的因子VIII变体的CS23m23密码子改变的核苷酸序列(SEQ ID NO:103)(“CS23-FL-NA-m23”)。

[0173] 图44展示了pCS构建体的克隆,这是通过将携带不同突变(参见插入的表格)的合成的Refacto型BDD-FVIII经由AscI和NotI限制位点插入载体骨架pCh-BB01中来实现。

[0174] 图45展示了通过PAGE和银染色对AAV载体制剂进行的蛋白质分析。泳道1,蛋白质标记物(M);泳道2,vCS01,泳道3,vCS17;泳道4,vCS19;泳道5,vCS20;泳道6,vCS40;泳道7,vCS04;泳道8,vCS17;泳道9,vCS24构建体。构建体全部具有由VP1、VP2以及VP3(右侧的箭头)组成的相同AAV8衣壳。左侧的刻度指示蛋白质标记物的大小(千道尔顿,kDa)。

[0175] 图46展示了通过琼脂糖凝胶电泳分析的AAV载体基因组制剂的完整性。泳道1,DNA标记物(M);泳道2,vCS04,泳道3,vCS17;泳道4,vCS20;泳道5,vCS24;泳道6,vCS16;泳道7,vCS40构建体。载体负荷为每一泳道1.5E10 vg。AAV载体具有相同尺寸的基因组,以约5kb进行迁移(箭头,右侧)。左侧的刻度指示DNA片段的大小(千碱基,kb)。

[0176] 图47示出了根据一些实施方案编码因子VIII变体的重链的CS01密码子改变的核苷酸序列的一部分(SEQ ID NO:24)(“CS01-HC-NA”)。

[0177] 图48示出了根据一些实施方案编码因子VIII变体的轻链的CS01密码子改变的核苷酸序列的一部分(SEQ ID NO:25)(“CS01-LC-NA”)。

[0178] 图49A和49B示出了根据一些实施方案编码单链因子VIII变体的CS01Δ(760-1667)(SPI;CS01Δ(741-1648),SPE)密码子改变的核苷酸序列(SEQ ID NO:26)(“CS01-SC1-NA”)。

[0179] 图50A和50B示出了根据一些实施方案编码单链因子VIII变体的CS01Δ(772-1667)(SPI;CS01Δ(753-1648),SPE)密码子改变的核苷酸序列(SEQ ID NO:27)(“CS01-SC2-NA”)。

[0180] 图51A和51B示出了根据一些实施方案编码单链因子VIII变体的CS23Δ(760-1667)(SPI;CS23Δ(741-1648),SPE)密码子改变的核苷酸序列(SEQ ID NO:28)(“CS23-SC1-NA”)。

[0181] 图52A和52B示出了根据一些实施方案编码单链因子VIII变体的CS23Δ(772-1667)(SPI;CS23Δ(753-1648),SPE)密码子改变的核苷酸序列(SEQ ID NO:29)(“CS23-SC2-NA”)。

[0182] 图53示出了根据一些实施方案由CS01m23密码子改变的核苷酸序列编码的因子VIII变体氨基酸序列(SEQ ID NO:104)(“CS01m23-FL-AA”)。

[0183] 图54示出了根据一些实施方案由CS04m3密码子改变的核苷酸序列编码的因子VIII变体氨基酸序列(SEQ ID NO:105)(“CS01m23-FL-AA”)。

[0184] 图55示出了根据一些实施方案由CS01m12密码子改变的核苷酸序列编码的因子VIII变体氨基酸序列(SEQ ID NO:106)(“CS01m12-FL-AA”)。

[0185] 图56示出了根据一些实施方案由CS04m12密码子改变的核苷酸序列编码的因子VIII变体氨基酸序列(SEQ ID NO:107)(“CS04m12-FL-AA”)。

[0186] 图57A和57B示出了根据一些实施方案编码具有m1(F328S)和m2氨基酸取代的因子VIII变体的CS01m12密码子改变的核苷酸序列(SEQ ID NO:108)(“CS01-FL-NAm12”)。

[0187] 图58A和58B示出了根据一些实施方案编码具有m1(F328S)和m2氨基酸取代的因子VIII变体的CS04m12密码子改变的核苷酸序列(SEQ ID NO:109)(“CS04-FL-NAm12”)。

[0188] 发明详述

[0189] I.引言

[0190] 基于AAV的基因疗法为血友病患者的治疗带来极大的希望。对于血友病B,最早的临床数据是令人鼓舞的,因为至少在有些患者中FIX水平可在约10%保持1年以上。然而,对于血友病A,由于各种原因,用AAV载体实现5-10%的治疗剂表达水平仍然是具有挑战性的。第一,因子VIII编码序列对于常规的基于AAV的载体来说太大。第二,工程化的B-结构域缺失或截短的因子VIII构建体在体内具有不良表达,即使在密码子优化时也是如此。第三,这些B-结构域缺失或截短的因子VIII变体构建体在体内的半衰期短,从而加剧了不良表达的效果。第四,如同诸如因子IX等其他凝血因子,即使在表达时,FVIII也不能高效地从细胞分泌。

[0191] 此外,这些挑战不能通过简单地施用更高剂量的基因疗法构建体来解决。根据目前的知识,基于AAV的基因疗法载体的载体剂量应增加到超过每千克体重$2x10^{12}$ vg。这是因为在如此高的剂量下,T细胞免疫应答被触发,从而破坏转导的细胞,并且因此降低或甚至消除转基因表达。因此,需要改善FVIII表达的策略以使FVIII基因疗法成为血友病A患者的可行治疗选择。

[0192] 本公开涉及发生解决与因子VIII基因疗法相关的这些和其他问题的密码子改变的因子VIII变体编码序列。举例来说,本文所公开的多核苷酸在哺乳动物细胞中提供显著改善的表达,并且因稳定的堆积相互作用而展示改善的病毒体包装。在一些实施方案中,通过使用与密码子改变的CS01、CS04以及CS23构建体具有高序列同一性(例如与CS01-HC、CS04-HC以及CS23-HC重链编码序列中的一种具有高序列同一性并且与CS01-LC、CS04-LC以及CS23-LC轻链编码序列中的一种具有高序列同一性)的因子VIII的重链和轻链的编码序列实现了这些优点。

[0193] 在一些实施方案中,已通过使野生型B结构域截短、缺失或替换来缩短由本文所描述的多核苷酸编码的因子VIII分子。因此,所述多核苷酸更适合于经由常规基因疗法载体表达因子VIII,所述常规基因疗法载体表达诸如野生型因子VIII等较大多肽的效率低。

[0194] 有利地,本文显示CS01、CS04以及CS23密码子改变的因子VIII变体编码序列在体内提供B-结构域缺失的因子VIII构建体的优良表达。举例来说,在实施例2和实施例4中证实在因子VIII敲除小鼠中,静脉内施用具有CS01(SEQ ID NO:13)、CS04(SEQ ID NO:1)以及CS23(SEQ ID NO:20)编码序列的基于AAV的基因疗法载体使得因子VIII表达相对于用野生型多核苷酸序列(SEQ ID NO:17)编码的相应CS40构建体增加18倍、74倍以及30倍(表4和表7)。

[0195] 此外,本文还显示CS01和CS04密码子改变的因子VIII变体编码序列提供优良的病毒体包装和病毒产量。举例来说,实施例1中证实当从相同量的细胞沉淀分离时,含有CS01和CS04构建体的AAV载体构建体提供相对于用野生型多核苷酸序列编码的相应CS40构建体高5和7倍的病毒产量。

[0196] 有利地,申请人还发现,可通过向基础因子VIII多肽序列中引入突变而使由CS01、CS04以及CS23密码子改变的序列产生的改善的因子VIII活性进一步增强。举例来说,如实施例4中证实,单独的和相互组合的F328S、X5以及X1突变当在体内在CS01或CS04密码子改变的背景中表达时使FVIII活性相对于野生型密码子改变的构建体进一步增加2至7倍(表7)。更显著的是,编码突变体因子VIII突变体的这些密码子改变的序列与用野生型多核苷

酸序列编码的相应CS40构建体相比提供高达246倍的增加(表7)。

[0197] II.定义

[0198] 除非另外指明,否则如本文所用,以下术语具有属于它们的含义。

[0199] 如本文所用,术语“因子VIII”和“FVIII”可互换使用,并且指具有因子VIII活性的任何蛋白质(例如活性FVIII,通常称为FVIIIa)或具有因子VIII活性,特别是因子IXa辅因子活性的蛋白质的蛋白质前体(例如前蛋白或前蛋白原)。在一个示例性实施方案中,因子VIII多肽是指具有与野生型因子VIII多肽的重链和轻链具有高序列同一性(例如至少70%、75%、80%、85%、90%、95%、99%或更高)的序列的多肽。在一些实施方案中,因子VIII多肽的B-结构域为缺失的、截短的或用接头多肽替换以减小编码因子VIII多肽的多核苷酸的尺寸。在一个示例性实施方案中,SEQ ID NO:2的氨基酸20-1457构成因子VIII多肽。

[0200] 野生型因子VIII多肽的非限制性实例包括人前因子VIII原(例如GenBank登录号AAA52485、CAA25619、AAA58466、AAA52484、AAA52420、AAV85964、BAF82636、BAG36452、CAI41660、CAI41666、CAI41672、CAI43241、CAO03404、EAW72645、AAH22513、AAH64380、AAH98389、AAI11968、AAI11970或AAB61261)、相应前因子VIII以及其天然变体;猪前因子VIII原(例如UniProt登录号F1RZ36或K7GSZ5)、相应前因子VIII以及其天然变体;小鼠前因子VIII原(例如GenBank登录号AAA37385、CAM15581、CAM26492或EDL29229)、相应前因子VIII以及其天然变体;大鼠前因子VIII原(例如GenBank登录号AAQ21580)、相应前因子VIII以及其天然变体;大鼠前因子VIII原;以及其他哺乳动物因子VIII同源物(例如猴、猿、仓鼠、豚鼠等)。

[0201] 如本文所用,因子VIII多肽包括具有因子IX辅因子活性的天然变体和人工构建体。如本公开中所用,因子VIII涵盖保留一定的基础因子IX辅因子活性(例如相应野生型活性的至少5%、10%、25%、50%、75%或更多)的任何天然变体、替代序列、同种型或突变体蛋白。存在于人类群体中的因子VIII氨基酸变异(相对于FVIII-FL-AA(SEQ ID NO:19))的实例包括但不限于S19R、R22T、Y24C、Y25C、L26P/R、E30V、W33G、Y35C/H、G41C、R48C/K、K67E/N、L69P、E72K、D75E/V/Y、P83R、G89D/V、G92A/V、A97P、E98K、V99D、D101G/H/V、V104D、K108T、M110V、A111T/V、H113R/Y、L117F/R、G121S、E129V、G130R、E132D、Y133C、D135G/Y、T137A/I、S138R、E141K、D145H、V147D、Y155H、V159A、N163K、G164D/V、P165S、C172W、S176P、S179P、V181E/M、K185T、D186G/N/Y、S189L、L191F、G193R、L195P、C198G、S202N/R、F214V、L217H、A219D/T、V220G、D222V、E223K、G224W、T252I、V253F、N254I、G255V、L261P、P262L、G263S、G266F、C267Y、W274C、H275L、G278R、G280D、E284K、V285G、E291G/K、T294I、F295L、V297A、N299I、R301C/H/L、A303E/P、I307S、S308L、F312S、T314A/I、A315V、G323E、L326P、L327P/V、C329F、I331V、M339T、E340K、V345A/L、C348R/S/Y、Y365C、R391C/H/P、S392L/P、A394S、W401G、I405F/S、E409G、W412G/R、K427I、L431F/S、R437P/W、I438F、G439D/S/V、Y442C、K444R、Y450D/N、T454I、F455C、G466E、P470L/R/T、G474E/R/V、E475K、G477V、D478N、T479R、F484C、A488G、R490G、Y492C/H、Y492H、I494T、P496R、G498R、R503H、G513S/V、I522Y、K529E、W532G、P540T、T541S、D544N、R546W、R550C/G/H、S553P、S554C/G、V556D、R560T、D561G/H/Y、I567T、P569R、S577F、V578A、D579A/H、N583S、Q584H/K/R、I585R/T、M586V、D588G/Y、L594Q、S596P、N601D/K、R602G、S603I/R、W604C、Y605H/S、N609I、R612C、N631K/S、M633I、S635N、N637D/I/S、Y639C、L644V、L650F、V653A/M、L659P、A663V、Q664P、F677L、M681I、V682F、

Y683C/N、T686R、F698L、M699T/V、M701I、G705V、G710W、N713I、R717L/W、G720D/S、M721I/L、A723T、L725Q、V727F、E739K、Y742C、R795G、P947R、V1012L、E1057K、H1066Y、D1260E、K1289Q、Q1336K、N1460K、L1481P、A1610S、I1698T、Y1699C/F、E1701K、Q1705H、R1708C/H、T1714S、R1715G、A1720V、E1723K、D1727V、Y1728C、R1740G、K1751Q、F1762L、R1768H、G1769R、L1771P、L1775F/V、L1777P、G1779E/R、P1780L、I1782R、D1788H、M1791T、A1798P、S1799H、R1800C/G/H、P1801A、Y1802C、S1803Y、F1804S、L1808F、M1842I、P1844S、T1845P、E1848G、A1853T/V、S1858C、K1864E、D1865N/Y、H1867P/R、G1869D/V、G1872E、P1873R、L1875P、V1876L、C1877R/Y、L1882P、R1888I、E1894G、I1901F、E1904D/K、S1907C/R、W1908L、Y1909C、A1939T/V、N1941D/S、G1942A、M1945V、L1951F、R1960L/Q、L1963P、S1965I、M1966I/V、G1967D、S1968R、N1971T、H1973L、G1979V、H1980P/Y、F1982I、R1985Q、L1994P、Y1998C、G2000A、T2004R、M2007I、G2013R、W2015C、R2016P/W、E2018G、G2022D、G2028R、S2030N、V2035A、Y2036C、N2038S、2040Y、G2045E/V、I2051S、I2056N、A2058P、W2065R、P2067L、A2070V、S2082N、S2088F、D2093G/Y、H2101D、T2105N、Q2106E/P/R、G2107S、R2109C、I2117F/S、Q2119R、F2120C/L、Y2124C、R2135P、S2138Y、T2141N、M2143V、F2145C、N2148S、N2157D、P2162L、R2169C/H、P2172L/Q/R、T2173A/I、H2174D、R2178C/H/L、R2182C/H/P、M2183R/V、L2185S/W、S2192I、C2193G、P2196R、G2198V、E2200D、I2204T、I2209N、A2211P、A2220P、P2224L、R2228G/L/P/Q、L2229F、V2242M、W2248C/S、V2251A/E、M2257V、T2264A、Q2265R、F2279C/I、I2281T、D2286G、W2290L、G2304V、D2307A、P2319L/S、R2323C/G/H/L、R2326G/L/P/Q、Q2330P、W2332R、I2336F、R2339T、G2344C/D/S以及C2345S/Y。因子VIII蛋白还包括含有翻译后修饰的多肽。

[0202] 通常，编码因子VIII的多核苷酸编码非活性单链多肽(例如前蛋白原)，所述非活性单链多肽经历翻译后加工以形成活性因子VIII蛋白(例如FVIIIa)。举例来说，参考图1，首先使野生型人因子VIII前蛋白原裂解以释放编码的信号肽(未显示)，从而形成第一单链前蛋白(显示为“人野生型FVIII)。然后，使前蛋白在B结构域与A3结构域之间裂解以形成包括因子VIII重链(例如A1结构域和A2结构域)和B-结构域的第一多肽，以及包括因子VIII轻链(例如包括A3、C1以及C3结构域)的第二多肽。使第一多肽进一步裂解以去除B-结构域，并且还分离A1结构域和A2结构域，所述A1结构域和A2结构域在成熟因子VIIIa蛋白中保持与缔合因子VIII轻链。关于因子VIII成熟过程的综述，参见Graw等，Nat Rev Genet.，6(6):488-501(2005)，该文献的内容出于所有目的以全文引用的方式并入本文中。

[0203] 然而，在一些实施方案中，因子VIII多肽为单链因子VIII多肽。对单链因子VIII多肽进行工程化以去除天然裂解位点，并且任选去除、截短或替换因子VIII的B-结构域。因此，它们不会因裂解而成熟(除了任选的信号和/或前导肽的裂解)，并且作为单链具有活性。Zollner等(Thromb Res,134(1):125-31(2014))和Donath等(Biochem J.,312(1):49-55(1995))中描述了单链因子VIII多肽的非限制性实例，这些文献的公开内容在此出于所有目的以全文引用的方式并入本文中。

[0204] 如本文所用，术语“因子VIII重链”或简单地“重链”是指因子VIII多肽的A1和A2结构域的聚集体。在一个示例性实施方案中，CSO4-FL-AA(SEQ ID NO:2)的氨基酸20-759构成因子VIII重链。

[0205] 如本文所用，术语“因子VIII轻链”或简单地“轻链”是指因子VIII多肽的A3、C1以及C2结构域的聚集体。在一个示例性实施方案中，CSO4-FL-AA(SEQ ID NO:2)的氨基酸774-

1457构成因子VIII轻链。在一些实施方案中，因子VIII轻链排除在体内成熟期间释放的酸性a3肽。

[0206] 通常，因子VIII重链和轻链例如与任选的B-结构域或B-结构域取代的接头一起表达为单一多肽链。然而，在一些实施方案中，因子VIII重链和因子VIII轻链表达为单独的多肽链(例如共表达)，并且重构以形成因子VIII蛋白(例如体内或体外)。

[0207] 如本文所用，术语“B-结构域取代的接头”和“因子VIII接头”可互换使用，并且指野生型因子VIII B-结构域的截短型式(例如FVIII-FL-AA(SEQ ID NO:19)的氨基酸760-1667)或经过工程化以替换因子VIII多肽的B-结构域的肽。如本文所用，根据一些实施方案，在因子VIII变体多肽中，因子VIII接头是定位于因子VIII重链的C端与因子VIII轻链的N端之间。美国专利号4,868,112、5,112,950、5,171,844、5,543,502、5,595,886、5,610,278、5,789,203、5,972,885、6,048,720、6,060,447、6,114,148、6,228,620、6,316,226、6,346,513、6,458,563、6,924,365、7,041,635以及7,943,374;美国专利申请公布号2013/024960、2015/0071883以及2015/0158930;以及PCT公布号WO 2014/064277和WO 2014/127215中公开了B-结构域取代的接头的非限制性实例，这些专利的公开内容在此出于所有目的以全文引用的方式并入本文中。

[0208] 除非本文另外指明，否则因子VIII氨基酸的编号是指在图22中以SEQ ID NO:19形式呈现的全长野生型人因子VIII序列(FVIII-FL-AA)中的相应氨基酸。因此，当提到本文所公开的因子VIII变体蛋白中的氨基酸取代时，所叙述的氨基酸编号是指在全长野生型因子VIII序列中类似的(例如结构上或功能上等同的)和/或同源的(例如在一级氨基酸序列中进化保守的)氨基酸。举例来说，T2105N氨基酸取代是指全长野生型人因子VIII序列(FVIII-FL-AA;SEQ ID NO:19)的位置2105的T到N的取代、由CS04编码的因子VIII变体蛋白(CS04-FL-AA;SEQ ID NO:2)的位置1211的T到N的取代以及由CS04m3编码的因子VIII变体(CS04m3-FL-AA;SEQ ID NO:105)的位置1212的T到N的取代。

[0209] 如本文所描述，因子VIII氨基酸编号系统是取决于是否包括因子VIII信号肽(例如全长野生型人因子VIII序列的氨基酸1-19)。在包括信号肽的情况下，编号被称为“信号肽包括性的”或“SPI”。在不包括信号肽的情况下，编号被称为“信号肽排除性的”或“SPE”。举例来说，F328S为与SPE编号中的F309S相同的氨基酸的SPI编号。除非本文另外陈述，否则所有氨基酸编号是指在图22中以SEQ ID NO:19形式呈现的全长野生型人因子VIII序列(FVIII-FL-AA)中的相应氨基酸。

[0210] 如本文所描述，与由天然编码的因子VIII构建体(例如使用野生型人密码子编码相同因子VIII构建体的多核苷酸)提供的因子VIII表达水平相比，密码子改变的多核苷酸在体内(例如当作为基因疗法载体的一部分施用时)提供增加的转基因因子VIII表达。如本文所用，术语“增加的表达”是指与施用天然编码的因子VIII构建体的动物的血液中的转基因因子VIII活性水平相比，施用编码因子VIII的密码子改变的多核苷酸的动物的血液中的转基因因子VIII活性水平增加。可使用本领域中已知的任何因子VIII活性来测量活性水平。用于测定因子VIII活性的示例性分析为Technochrome FVIII分析(Technoclone,Vienna,Austria)。

[0211] 在一些实施方案中，增加的表达是指与施用天然编码的因子VIII多核苷酸的动物的血液中的转基因因子VIII活性水平相比，施用密码子改变的因子VIII多核苷酸的动物的

血液中的转基因因子VIII活性高至少25％。在一些实施方案中，增加的表达是指与施用天然编码的因子VIII多核苷酸的动物的血液中的转基因因子VIII活性水平相比，施用密码子改变的因子VIII多核苷酸的动物的血液中的转基因因子VIII活性高至少50％、高至少75％、高至少100％、高至少3倍、高至少4倍、高至少5倍、高至少6倍、高至少7倍、高至少8倍、高至少9倍、高至少10倍、高至少15倍、高至少20倍、高至少25倍、高至少30倍、高至少40倍、高至少50倍、高至少60倍、高至少70倍、高至少80倍、高至少90倍、高至少100倍、高至少125倍、高至少150倍、高至少175倍、高至少200倍、高至少225倍或高至少250倍。

[0212] 如本文所描述，与由天然编码的因子VIII构建体(例如使用野生型人密码子编码相同因子VIII构建体的多核苷酸)提供的载体产量水平相比，密码子改变的多核苷酸提供增加的载体产量。如本文所用，术语“增加的病毒产量”是指与接种了天然编码的因子VIII构建体的细胞培养物中的载体产量(例如每升培养物的效价)相比，接种了编码因子VIII的密码子改变的多核苷酸的细胞培养物中的载体产量增加。可使用本领域中已知的任何载体效价分析来测量载体产量。用于测定载体(例如AAV载体)产量的示例性分析为靶向AAV2反向末端重复序列的qPCR(Aurnhammer,Human Gene Therapy Methods:B部分23:18-28(2012))。

[0213] 在一些实施方案中，增加的病毒产量是指与相同类型培养物中的天然编码的因子VIII构建体的产量相比，密码子改变的载体产量高至少25％。在一些实施方案中，增加的载体产量是指与相同类型培养物中的天然编码的因子VIII构建体的产量相比，密码子改变的载体产量高至少50％、高至少75％、高至少100％、高至少3倍、高至少4倍、高至少5倍、高至少6倍、高至少7倍、高至少8倍、高至少9倍、高至少10倍、高至少15倍或高至少20倍。

[0214] 如本文所用，术语“血友病”是指普遍特征为血液凝结或凝固减少的一组疾病状态。血友病可指A型、B型或C型血友病，或指所有三种疾病类型的复合疾病。A型血友病(血友病A)是由因子VIII(FVIII)活性降低或丧失造成的，并且在血友病亚型中为最突出的。B型血友病(血友病B)是由因子IX(FIX)凝结功能丧失或降低引起。C型血友病(血友病C)是因子XI(FXI)凝结活性丧失或降低的结果。血友病A和B为X连锁疾病，而血友病C是常染色体病。血友病的常规治疗包括预防性施用和按需施用凝结因子(诸如FVIII、FIX，包括Bebulin®-VH和FXI以及FEIBA-VH)、去氨加压素以及血浆输注。

[0215] 如本文所用，术语“FVIII基因疗法”包括向患者提供编码因子VIII的核酸以缓解、减少或预防与血友病相关的一种或多种症状(例如临床因素)的再次发生的任何治疗方法。该术语涵盖施用包含编码因子VIII分子，包括任何修饰形式的因子VIII(例如因子VIII变体)的核酸的任何化合物、药物、程序或方案以维持或改善患有血友病个体的健康。本领域技术人员将了解，可例如基于根据本公开所获得的结果改变FVIII疗法的过程或FVIII治疗剂的剂量。

[0216] 如本文所用，术语“旁路疗法”包括向患者提供非因子VIII止血剂、化合物或凝血因子以缓解、减少或预防与血友病相关的一种或多种症状(例如临床因素)的再次发生的任何治疗方法。非因子VIII化合物和凝血因子包括但不限于因子VIII抑制剂旁路活性(FEIBA)、重组活化因子VII(FVIIa)、凝血酶原复合物浓缩物以及活化的凝血酶原复合物浓缩物。这些非因子VIII化合物和凝血因子可为重组的或血浆衍生的。本领域技术人员将了解，可例如基于根据本公开所获得的结果改变旁路疗法的过程或旁路疗法的剂量。

[0217] 如本文所用,包括施用编码因子VIII分子的核酸和常规血友病A治疗剂的“组合疗法”包括向患者提供编码因子VIII分子的核酸与因子VIII分子和/或非因子VIII止血剂(例如旁路治疗剂)以缓解、减少或预防与血友病相关的一种或多种症状(例如临床因素)的再次发生的任何治疗方法。该术语涵盖施用包括编码因子VIII分子,包括任何修饰形式的因子VIII(例如因子VIII变体)的核酸的适用于维持或改善患有血友病个体的健康并且包括任何本文所描述的治疗剂的任何化合物、药物、程序或方案。

[0218] 术语“治疗有效量或剂量”或“治疗充足量或剂量”或“有效或充足量或剂量”是指产生施用它要达到的治疗效果的剂量。举例来说,适用于治疗血友病的药物的治疗有效量可为能够预防或缓解与血友病相关的一种或多种症状的量。确切的剂量将取决于治疗目的,并且将为本领域技术人员使用已知技术可确定的(参见例如Lieberman,Pharmaceutical Dosage Forms(第1-3卷,1992);Lloyd,The Art,Science and Technology of Pharmaceutical Compounding(1999);Pickar,Dosage Calculations(1999);以及Remington:The Science and Practice of Pharmacy,第20版,2003,Gennaro编,Lippincott,Williams&Wilkins)。

[0219] 如本文所用,术语“基因”是指编码多肽链的DNA分子片段(例如编码区)。在一些实施方案中,基因是按照紧接在产生多肽链时所涉及的编码区之前、之后和/或插入其中的区域进行定位(例如调控元件,诸如启动子、增强子、聚腺苷酸化序列、5’-非翻译区、3’-非翻译区或内含子)。

[0220] 如本文所用,术语“调控元件”是指在细胞中提供编码序列的表达的核苷酸序列,诸如启动子、增强子、终止子、聚腺苷酸化序列、内含子等。

[0221] 如本文所用,术语“启动子元件”是指帮助控制编码序列的表达的核苷酸序列。通常,将启动子元件定位于基因的翻译起始位点的5’。然而,在某些实施方案中,可将启动子元件定位于内含子序列内或编码序列的3’。在一些实施方案中,适用于基因疗法载体的启动子是来源于靶蛋白的天然基因(例如因子VIII启动子)。在一些实施方案中,适用于基因疗法载体的启动子关于在目标有机体的特定细胞或组织中的表达具特异性(例如肝特异性启动子)。在其他实施方案中,将多种充分表征的启动子元件中的一种用于本文所描述的基因疗法载体中。充分表征的启动子元件的非限制性实例包括CMV早期启动子、β-肌动蛋白启动子以及甲基CpG结合蛋白2(MeCP2)启动子。在一些实施方案中,启动子为组成型启动子,驱动靶蛋白的基本上恒定的表达。在其他实施方案中,启动子为诱导型启动子,响应于特定刺激(例如暴露于特定治疗或药剂)而驱动靶蛋白的表达。关于设计用于AAV介导的基因疗法的启动子的综述,参见Gray等(Human Gene Therapy22:1143-53(2011)),该文献的内容出于所有目的以全文引用的方式明确并入本文中。

[0222] 如本文所用,术语“载体”是指用于将核酸(例如编码因子VIII基因疗法构建体的核酸)转移到宿主细胞中的任何媒介物。在一些实施方案中,载体包括用于复制媒介物以及靶核酸的复制子。适用于基因疗法的载体的非限制性实例包括在体内用作自主复制单元的质粒、噬菌体、粘粒、人工染色体以及病毒。在一些实施方案中,载体为用于引入靶核酸(例如编码因子VIII变体的密码子改变的多核苷酸)的病毒媒介物。适用于基因疗法的许多修饰的真核病毒为本领域中已知的。举例来说,腺相关病毒(AAV)特别适合用于人类基因疗法,因为人是病毒的天然宿主,已知天然病毒不会导致任何疾病,并且这些病毒引起轻度免

疫应答。

[0223] 如本文所用，术语“CpG岛”是指多核苷酸内具有统计学上升高的CpG二核苷酸密度的区域。如本文所用，如果在200-碱基对窗口中：(i)多核苷酸(例如编码密码子改变的因子VIII蛋白的多核苷酸)的区域的GC含量高于50％，并且(ii)如通过以下关系所定义，每单位预期CpG二核苷酸观测到的CpG二核苷酸的比率至少为0.6：

[0224] $\frac{N[CpG]*N[窗长度]}{N[C]*N[G]} \geq 0.6$，那么该区域为CpG岛。

[0225] 关于鉴定CpG岛的方法的额外信息，参见Gardiner-Garden M.等，J Mol Biol.，196(2):261-82(1987)，该文献的内容出于所有目的以全文引用的方式明确并入本文中。

[0226] 如本文所用，术语“核酸”是指呈单链或双链形式的脱氧核糖核苷酸或核糖核苷酸和其聚合物以及其互补物。该术语涵盖含有已知核苷酸类似物或修饰的主链残基或键联的核酸，所述核酸为合成的、天然存在的以及非天然存在的，具有与参考核酸相似的结合特性，并且以与参考核酸相似的方式进行代谢。此类类似物的实例包括但不限于硫代磷酸酯、氨基磷酸酯、甲基膦酸酯、手性-甲基膦酸酯、2-O-甲基核糖核苷酸以及肽-核酸(PNA)。

[0227] 术语“氨基酸”是指天然存在和非天然存在的氨基酸，包括以与天然存在的氨基酸相似的方式起作用的氨基酸类似物和氨基酸模拟物。天然存在的氨基酸包括由遗传密码编码的那些氨基酸以及后来被修饰的那些氨基酸，例如羟脯氨酸、y-羧基谷氨酸以及O-磷酸丝氨酸。天然存在的氨基酸可包括例如D-氨基酸和L-氨基酸。本文中所用的氨基酸还可包括非天然氨基酸。氨基酸类似物是指与天然存在的氨基酸具有相同的基本化学结构的化合物，所述基本化学结构为结合至氢、羧基、氨基和R基团(例如高丝氨酸、正亮氨酸、蛋氨酸亚砜或蛋氨酸甲基锍)的任何碳。此类类似物具有修饰的R基团(例如正亮氨酸)或修饰的肽主链，但保留了与天然存在的氨基酸相同的基本化学结构。氨基酸模拟物是指具有与氨基酸的一般化学结构不同的结构，但以与天然存在的氨基酸相似的方式起作用的化合物。在本文中可通过通常已知的三字母符号或通过IUPAC-IUB生物化学命名委员会推荐的单字母符号来提到氨基酸。同样地，可通过普遍接受的单字母代码来提到核苷酸。

[0228] 编码本文中的突变体因子VIII构建体的核苷酸序列可与本文中提供的编码序列相同或可为不同的编码序列，其序列因遗传密码的冗余性或简并性而编码与本文中提供的编码序列相同的多肽。本领域的一般技术人员将认识到，可对核酸中的各密码子(除了通常是甲硫氨酸的唯一密码子的AUG和通常是色氨酸的唯一密码子的TGG)进行修饰以得到功能相同的分子。因此，关于表达产物而不是关于实际的基因疗法构建体的每个所描述的序列中隐含编码相同多肽的核酸的每种变异型式。

[0229] 关于氨基酸序列，本领域的一般技术人员将认识到，改变、添加或缺失所编码序列中的单一氨基酸或较小百分比氨基酸的对核酸或肽序列的个别取代、缺失或添加是“保守修饰的变体”，其中所述改变导致氨基酸被化学上相似的氨基酸取代。提供功能上相似的氨基酸的保守取代表是本领域中熟知的。此类保守修饰的变体是除了本公开的多态变体、种间同系物以及等位基因以外的变体并且不排除本公开的多态变体、种间同系物以及等位基因。

[0230] 提供功能上相似的氨基酸的保守氨基酸取代是本领域中熟知的。视特定氨基酸的功能性，例如催化性、结构性或空间上重要的氨基酸而定，不同的氨基酸分组可以被认为是

彼此的保守取代。表1提供基于氨基酸的电荷和极性、氨基酸的疏水性、氨基酸的表面暴露/结构性质以及氨基酸的二级结构倾向而视为保守取代的氨基酸的分组。

[0231] 表1.基于蛋白质中残基的功能性的保守氨基酸取代分组.

[0232]

重要特征	保守分组
电荷/极性	1. H、R 以及 K 2. D 和 E 3. C、T、S、G、N、Q 以及 Y 4. A、P、M、L、I、V、F 以及 W
疏水性	1. D、E、N、Q、R 以及 K 2. C、S、T、P、G、H 以及 Y 3. A、M、I、L、V、F 以及 W
结构/表面暴露	1. D、E、N、Q、H、R 以及 K 2. C、S、T、P、A、G、W 以及 Y 3. M、I、L、V 以及 F
二级结构倾向	1. A、E、Q、H、K、M、L 以及 R 2. C、T、I、V、F、Y 以及 W 3. S、G、P、D 以及 N
进化保守性	1. D 和 E 2. H、K 以及 R 3. N 和 Q 4. S 和 T 5. L、I 以及 V 6. F、Y 以及 W 7. A 和 G 8. M 和 C

[0233] 在两个或更多个核酸或肽序列的背景下，术语“同一”或“同一性”百分比是指如使用BLAST或BLAST 2.0序列比较算法用以下所描述的默认参数或通过人工比对和目视检查所测量，两个或更多个序列或子序列为相同的或具有指定百分比的相同的氨基酸残基或核苷酸(即当比较和比对相对于比较窗或指定区的最大对应性时，在指定区内具有约60％的同一性，优选为65％、70％、75％、80％、85％、90％、91％、92％、93％、94％、95％、96％、

97％、98％、99％或更高同一性)。

[0234] 如本领域中已知的,可使用许多不同程序来鉴定蛋白质(或如下文所论述的核酸)是否与已知序列具有序列同一性或相似性。使用本领域已知的标准技术来测定序列同一性和/或相似性,所述标准技术包括包括但不限于Smith和Waterman,Adv.Appl.Math.,2:482(1981)的局部序列同一性算法、通过Needleman和Wunsch,J.Mol.Biol.,48:443(1970)的序列同一性比对算法、通过Pearson和Lipman,Proc.Natl.Acad.Sci.U.S.A.,85:2444(1988)的检索相似性方法、通过这些算法的计算机实施方式(Wisconsin遗传学软件包中的GAP、BESTFIT、FASTA以及TFASTA,Genetics Computer Group,575Science Drive,Madison,WI)、由Devereux等,Nucl.Acid Res.,12:387-395(1984)描述的Best Fit序列程序,优选使用默认设置或通过检查来进行。优选地,通过FastDB基于以下参数来计算同一性百分比:错配罚分为1;空位罚分为1;空位尺寸罚分为0.33;并且连接罚分为30,“Current Methods in Sequence Comparison and Analysis,”Macromolecule Sequencing and Synthesis,Selected Methods and Applications,第127-149页(1988),Alan R.Liss,Inc,所有这些文献都以引用的方式并入本文中。

[0235] 适用算法的实例为PILEUP。PILEUP使用渐进性成对比对从一组相关序列形成多重序列比对。它还可绘制示出用于形成比对的聚类关系的树形图。PILEUP使用Feng和Doolittle,J.Mol.Evol.35:351-360(1987)的渐进性比对方法的简化型式;所述方法与Higgins和Sharp CABIOS 5:151-153(1989)所描述的方法相似,两个文献均以引用的方式并入本文中。适用的PILEUP参数包括默认空位权重3.00、默认空位长度权重0.10以及加权末端空位。

[0236] 适用算法的另一实例为以下文献中描述的BLAST算法:Altschul等,J.Mol.Biol.215,403-410,(1990);Altschul等,Nucleic Acids Res.25:3389-3402(1997);以及Karlin等,Proc.Natl.Acad.Sci.U.S.A.90:5873-5787(1993),这些文献均以引用的方式并入本文中。特别适用的BLAST程序为从Altschul等,Methods in Enzymology,266:460-480(1996);http://blast.wustl/edu/blast/README.html]获得的WU-BLAST-2程序。WU-BLAST-2使用若干检索参数,其中大部分被设置为默认值。可调节参数被设置为以下数值:重叠跨度＝1,重叠分数＝0.125,字串阈值(T)＝11。HSP S和HSP S2参数为动态值并且通过程序自身根据特定序列的组成和在其中检索所关注的序列的特定数据库的组成来建立;然而,可对所述值进行调节以增加灵敏度。

[0237] 另一适用算法为如由Altschul等人,Nucl.Acids Res.,25:3389-3402所报道的带空位的BLAST,该文献以引用的方式并入本文中。带空位的BLAST使用BLOSUM-62取代评分;阈值T参数设置为9;引发无空位延伸的双击法;使空位长度k承担10+k的代价;Xu设置为16,并且对于数据库检索阶段,Xg设置为40,并且对于算法的输出阶段设置为67。通过对应于约22位的评分来引发带空位的比对。

[0238] 通过用匹配的相同残基数目除以比对区中“较长”序列的残基总数目来测定氨基酸序列同一性％值。“较长”序列为具有比对区中的大部分实际残基的序列(忽略由WU-Blast-2为使比对评分最大化而引入的空位)。以类似方式,将关于所鉴定多肽的编码序列的“核酸序列同一性百分比(％)”定义为候选序列中与细胞周期蛋白的编码序列中的核苷酸残基相同的核苷酸残基的百分比。优选方法利用设置为默认参数的WU-BLAST-2的BLASTN

模块,其中重叠跨度和重叠分数分别设置为1和0.125。

[0239] 比对可包括在所要比对的序列中引入空位。另外,对于含有比由图2的序列(SEQ ID NO:1)编码的蛋白质更多或更少的氨基酸的序列,应了解,在一个实施方案中,序列同一性百分比将是基于相对于氨基酸或核苷酸总数目的相同氨基酸或核苷酸数目来确定。因此,举例来说,在一个实施方案中,将使用较短的序列中的核苷酸数目来测定如下文所论述比图2中所示的序列(SEQ ID NO:1)更短的序列的序列同一性。在同一性百分比计算中,未将相对权重分配给序列变异的不同表现形式,诸如插入、缺失、取代等。

[0240] 在一个实施方案中,仅同一性为正评分(+1)并且为所有形式的序列变异(包括空位)分配值“0”,这排除了对如下文关于序列相似性计算所描述的加权标度或参数的需要。可例如通过用匹配的相同残基的数目除以比对区内“较短”序列的残基总数目并且乘以100来计算序列同一性百分比。“较长”序列为具有比对区内大部分实际残基的序列。

[0241] 术语“等位基因变体”是指特定遗传基因座处的基因的多态形式,以及来源于所述基因的mRNA转录物的cDNA和由其编码的多肽。术语“优选的哺乳动物密码子”是指如选自以下列表的在哺乳动物细胞中表达的蛋白质中最常用的编码氨基酸的密码子集合中的密码子的子集:Gly(GGC,GGG);Glu(GAG);Asp(GAC);Val(GTG,GTC);Ala(GCC,GCT);Ser(AGC,TCC);Lys(AAG);Asn(AAC);Met(ATG);Ile(ATC);Thr(ACC);Trp(TGG);Cys(TGC);Tyr(TAT,TAC);Leu(CTG);Phe(TTC);Arg(CGC,AGG,AGA);Gln(CAG);His(CAC);以及Pro(CCC)。

[0242] 如本文所用,术语密码子改变的是指编码多肽(例如因子VIII变体蛋白)的多核苷酸序列,其中编码多肽的天然多核苷酸中的至少一个密码子已被改变以改善多核苷酸序列的特性。在一些实施方案中,改善的特性促使编码多肽的mRNA的转录增加、mRNA的稳定性增加(例如改善的mRNA半衰期)、多肽的翻译增加和/或多核苷酸包装于载体内增加。可用于实现改善的特性的改变的非限制性实例包括改变用于特定氨基酸的密码子的使用和/或分布、调整全局和/或局部GC含量、去除含AT的序列、去除重复的序列元件、调整全局和/或局部CpG二核苷酸含量、去除隐蔽的调控元件(例如TATA盒和CCAAT盒元件)、去除内含子/外显子剪接位点、改善调控序列(例如引入Kozak共有序列)以及去除能够在转录的mRNA中形成二级结构(例如茎环)的序列元件。

[0243] 如本文所论述,存在各种名称来指本文所公开的组分。“CS编号”(例如“CS04”、“CS01”、“CS23”等)是指编码FVIII多肽的密码子改变的多核苷酸和/或所编码的多肽,包括变体。举例来说,CS01-FL是指全长(Full Length)密码子改变的CS01多核苷酸序列或由CS01多核苷酸序列编码的氨基酸序列(对于氨基酸(Amino Acid)序列在本文中有时称为“CS01-FL-AA”,而对于核酸(Nucleic Acid)序列则为“CS01-FL-NA”)。类似地,“CS01-LC”是指编码FVIII多肽的轻链的密码子改变的核酸序列(“CS01-LC-NA”)或由CS01多核苷酸序列编码的FVIII轻链的氨基酸序列(在本文中有时也称为“CS01-LC-AA”)。同样地,CS01-HC、CS01-HC-AA以及CS01-HC-NA对于FVIII重链来说也是一样的情况。如本领域技术人员将了解,对于仅发生密码子改变(例如与Refacto相比不含额外氨基酸取代)的诸如CS01、CS04、CS23等构建体,氨基酸序列将为相同的,因为氨基酸序列不因密码子优化而改变。因此,本公开的序列构建体包括但不限于CS01-FL-NA、CS01-FL-AA、CS01-LC-NA、CS01-LC-AA、CS01-HC-AA、CS01-HC-NA、CS04-FL-NA、CS04-FL-AA、CS04-LC-NA、CS04-LC-AA、CS04-HC-AA、CS04-HC-NA、CS23-FL-NA、CS23-FL-AA、CS23-LC-NA、CS23-LC-AA、CS23-HC-AA以及CS23-HC-NA。

[0244] 此命名法还适用于如图13所示的糖基化肽，使得“NGA1-AA”指氨基酸序列，而NGA1-NA指核酸序列。

[0245] 本公开还包括如下文所描述的使用适当命名法的额外新因子VIII变体。

[0246] III.密码子改变的因子VIII变体

[0247] 在一些实施方案中，本公开提供编码因子VIII变体的密码子改变的多核苷酸。这些密码子改变的多核苷酸当在基于AAV的基因疗法构建体中施用时提供显著改善的因子VIII表达。与常规密码子优化构建体相比，密码子改变的多核苷酸还展示改善的AAV-病毒体包装。如实施例2和实施例4中所证实，申请人已通过发现三种编码因子VIII多肽的密码子改变的多核苷酸(CS01-FL-NA、CS04-FL-NA以及CS23-FL-NA)实现了这些优点，所述因子VIII多肽具有人野生型因子VIII重链和轻链以及短的14氨基酸的B-结构域取代的接头(“SQ”接头)，所述接头含有弗林蛋白酶裂解位点以促进活性FVIIIa蛋白质在体内成熟。如实施例4中进一步证实，将F328S、X5以及X1氨基酸突变的各种组合合并到编码的因子VIII分子中使因子VIII活性的体内表达进一步增加。

[0248] 在一个实施方案中，本文所提供的密码子改变的多核苷酸具有至少与CS01、CS04或CS23内编码因子VIII重链和因子VIII轻链的序列(分别为SEQ ID NO 13、1以及20)具有高序列同一性的核苷酸序列。如本领域中已知，因子VIII的B-结构域对于体内活性来说不是必需的。因此，在一些实施方案中，本文所提供的密码子改变的多核苷酸完全缺乏因子VIII B-结构域。在一些实施方案中，将天然因子VIII B-结构域替换为含有弗林蛋白酶裂解位点的短氨基酸接头，例如由CS01、CS04或CS23(分别为SEQ ID NO 2、2以及21)构建体的氨基酸760-773组成的“SQ”接头。“SQ”接头也称为BDLO04(如图6中所示，对于氨基酸序列为-AA，而对于核苷酸序列为-NA)。

[0249] 在一个实施方案中，由密码子改变的多核苷酸编码的因子VIII重链和轻链分别为人因子VIII重链和轻链。在其他实施方案中，由密码子改变的多核苷酸编码的因子VIII重链和轻链为来自另一种哺乳动物(例如猪因子VIII)的重链和轻链序列。在其他实施方案中，因子VIII重链和轻链为嵌合重链和轻链(例如人和第二哺乳动物序列的组合)。在其他实施方案中，因子VIII重链和轻链为来自另一种哺乳动物的重链和轻链的人源化型式，例如其中人残基在所选位置被取代以降低在向人施用时所得肽的免疫原性的来自另一种哺乳动物的重链和轻链序列。

[0250] 人基因的GC含量变化很大，从低于25％到高于90％。然而，一般来说，具有更高GC含量的人基因以更高的水平表达。举例来说，Kudla等(PLoS Biol.,4(6):80(2006))证实增加基因的GC含量会增加所编码多肽的表达，这主要是通过增加转录和实现mRNA转录物的更高稳态水平来实现。通常，密码子优化的基因构建体的所需GC含量等于或高于60％。然而，天然AAV基因组的GC含量约为56％。

[0251] 因此，在一些实施方案中，本文所提供的密码子改变的多核苷酸的CG含量更接近于天然AAV病毒体的GC含量(例如约56％ GC)，所述GC含量低于针对在哺乳动物细胞中表达而进行常规密码子优化的多核苷酸的优选CG含量(例如等于或超过60％ GC)。如实施例1中所概述，GC含量为约56％的CS04-FL-NA(SEQ ID NO:1)与GC含量更高的类似密码子改变的编码序列相比具有改善的病毒体包装。

[0252] 因此，在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含

量低于60％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量低于59％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量低于58％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量低于57％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量不超过56％。

[0253] 在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为54％至59％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为55％至59％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为56％至59％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为54％至58％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为55％至58％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为56％至58％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为54％至57％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为55％至57％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为56％至57％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为54％至56％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为55％至56％。

[0254] 在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为56±0.5％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为56±0.4％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为56±0.3％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为56±0.2％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为56±0.1％。在一些实施方案中，编码因子VIII多肽的密码子改变的多核苷酸的总GC含量为56％。

[0255] A.因子VIII氨基酸取代

[0256] 为使本文所描述的因子VIII构建体的基于AAV-载体的表达的效率进一步增加，在一些实施方式中进一步合并了已知改善分泌、增加比活性和/或增强因子VIII稳定性的氨基酸取代。鉴定出许多潜在变体，这些潜在变体在给定载体剂量下增加血浆FVIII活性水平。这些变体包括具有更高效的信号肽、阻止BiP相互作用的氨基酸取代、类似于更高效分泌的因子VIII直系同源物(例如猪因子VIII)的氨基酸取代、单链因子VIII变体以及稳定因子VIII和/或减少亚基解离的氨基酸取代的那些变体。

[0257] 定位于A1结构域与C2结构域之间的界面处的残基A108、R121以及L2302(SPE)的突变使因子VIII的稳定性增加。举例来说，A108I氨基酸取代引入了疏水性残基，所述疏水性残基更好地填充结构域间的空间，从而稳定相互作用。同样地，R121C/L2302C(SPE)双重氨基酸取代引入了跨越A1-C2结构域的二硫键，进一步稳定相互作用。总之，所有三种氨基酸取代使因子VIII的热稳定性增加3至4倍。关于综述，参见Wakabayashi等，J Biol Chem.286(29):25748-55(2011)和Wakabayashi等，Thromb Haemost.10(3):492-95(2012)。因此，在一些实施方案中，编码的因子VIII多肽包括A108I和/或R121C/L2302C氨基酸取代。

[0258] 定位于因子VIII的结合钙的结构域内的E113(SPE)的突变使FVIII凝结比活性增

加。举例来说,E113A似乎通过增加的FVIII对因子IXa的亲和力增加了FX酶形成。具体地说,E113A氨基酸取代使FVIII凝结比活性增加两倍,并且使对因子IXa的亲和力增加四倍(Biochemistry,41:8485(2002);J.Biol.Chem.,279:12677(2004);以及Biochemistry,44:10298(2005))。因此,在一些实施方案中,编码的因子VIII多肽包括E113A氨基酸取代。

[0259] 因子VIII APC裂解位点(残基331-341(SPE))周围的一个或多个氨基酸残基的取代使因子VIIIa因活化的蛋白C而造成的失活减少,而不会影响FVIII活性。举例来说,PQL333-335VDQ(SPE)氨基酸取代使因子VIII失活减少到1/16。同样地,MKN336-339GNQ氨基酸取代使因子VIII失活减少到1/9。当组合时,两种三重氨基酸取代(例如PQLRMKN333-339VDQRGNQ)(分别为SEQ ID NO 34和35)使因子VIII失活减少到1/100(J.Biol.Chem.,282:20264(2007)。因此,在一些实施方案中,编码的因子VIII多肽包括PQL333-335VDQ和/或MKN337-339GNQ(SPE)氨基酸取代。

[0260] A2结构域界面内的突变也增加因子VIII稳定性。具体地说,使A1-A2和A2-A3结构域界面中的带电荷残基突变使因子VIIIa中A2亚基的稳定性和保留增加。举例来说,D519、E665以及E1984突变为V或A使因子VIII的稳定性增加高达2倍,并且使VIIIa的稳定性增加高达5倍。具体地说,D519A/E665V氨基酸取代使稳定性增加3倍;D519V/E665V氨基酸取代使稳定性增加2倍,使A2解离减少到1/8,并且使产生凝血酶的潜力增加2-4倍;D519V/E1984A氨基酸取代使稳定性增加2倍;并且D519V/E665V/E1984A氨基酸取代使稳定性增加2倍(Blood 112:2761-69(2008);J.Thromb.Haemost.,7:438-44(2009))。因此,在一些实施方案中,编码的因子VIII多肽包括D519A/V、E665A/V以及E1984A/V氨基酸取代中的一种或多种。

[0261] 与本公开特别相关的是可单独地或与本文所描述的其他变体组合包括在内的许多特定突变。这些变体在本文中如下被编码为集合:“m1”是指单个氨基酸改变,“m2”是5个氨基酸变体的集合,“m3”是跨越多肽接头和重链之间的连接的7个氨基酸的缺失与六个氨基酸的插入的组合,“m4”是m1单突变与m5双突变的组合,而“m5”是两处半胱氨酸切除的集合。下文描述这些突变。可将这些突变单独或与其他变体组合包括在任何特定构建体中,并且它们被相应地编码。举例来说,如本文所概述,“m23”为m2和m3变体组合到特定支架上;因此,“CS01m23-FL-NA”或“CS01-FL-NAm23”是指包括编码m2和m3突变的核苷酸的CS01密码子改变的多核苷酸序列,而“CS01m23-FL-AA”或“CS01-FL-AAm23”是指氨基酸序列。由于CS01为密码子改变的,但不改变Refacto的氨基酸序列,这些变体可被认为与CS01-FL-AA(SEQ ID NO:2)的Refacto氨基酸序列相比在氨基酸层面上为突变。

[0262] 在许多实施方案中,将本公开的多肽制备成包括“m1”变体。A1结构域中与BiP相互作用的11氨基酸疏水β-折叠内的突变使因子VIII的分泌增加。举例来说,口袋内的F328S(SPI,F309S SPE)氨基酸取代使因子VIII分泌增加3倍。F328S变体在本文中被称为“m1”突变并且位于重链内。再次,如本文所描述,可包括信号肽(“信号肽包括性(Signal Peptide Inclusive)”或“SPI”)或者从所加工的最终蛋白序列开始(“信号肽排除性(Signal Peptide Exclusive)”或“SPE”)来进行变体的编号。因此,使用SPI编号,突变F328S与F309 SPE突变体相同。一般来说,本说明书使用SPI编号,但如本领域技术人员将了解,任一编号系统均产生相同的突变。

[0263] 因此,本公开中包括包含m1突变的多肽,包括CS01-FL-AAm1、CS01-HC-AAm1、CS04-

FL-AAm1、CS04-HC-AAm1 CS23-FL-AAm1、CS23-HC-AAm1、CS40-FL-AAm1以及CS40-HC-AAm1(所有这些均编码相同的相应蛋白序列)。

[0264] 另外,本公开中不仅包括包含m1突变的多肽序列,而且包括那些编码具有m1突变的蛋白质的密码子改变的多核苷酸序列,诸如CS01-FL-NAm1、CS01-HC-NAm1、CS04-FL-NAm1、CS04-HC-NAm1、CS23-FL-NAm1、CS23-HC-NA-m1、CS40-FL-NAm1以及CS40-HC-NAm1。

[0265] 在许多实施方案中,将本公开的多肽制备成包括“m2”变体集合,所述变体集合为I105V/A127S/G151K/M166T/L171P突变(SPI编号;(SPE编号分别为V86I/S108A/K132G/T147M/P152L)。m2突变集合是基于以下事实:猪氨基酸82-176取代B-结构域缺失的基因疗法构建体中的相应人氨基酸当在HEK293细胞中表达时使因子VIII活性增加(W.Xiao,通迅)。Id.单个猪氨基酸突变回人BDD-FVIII构建体在A1结构域内鉴定出促成该现象的五种氨基酸:I105V、A127S、G151K、M166T以及L171P(SPI)。将这些突变的组合引入人构建体中重现了更大的猪取代的改善的活性。Id.因此,在一些实施方案中,编码的因子VIII多肽包括一个或多个选自I105V、A127S、G151K、M166T以及L171P的氨基酸取代加上整个5氨基酸集合m2,从而在许多实施方案中具有特定用途。如同关于m1突变,m2变体位于重链中,并且因此本公开包括包含m2突变的多肽,包括CS01-FL-AAm2、CS01-HC-AAm2、CS04-FL-AAm2、CS04-HC-AAm2、CS23-FL-AAm2、CS23-HC-AAm2、CS40-FL-AAm2以及CS40-HC-AAm2(所有这些均编码相同的相应蛋白序列)。

[0266] 另外,本公开中不仅包括包含m2突变的多肽序列,而且包括编码具有m2突变的蛋白质的那些密码子改变的多核苷酸序列,诸如CS01-FL-NAm2、CS01-HC-NAm2、CS04-FL-NAm2、CS04-HC-NAm2、CS23-FL-NAm2、CS23-HC-NA-m2、CS40-FL-NAm2以及CS40-HC-NAm2。

[0267] 在其他实施方案中,本公开的多肽和多核苷酸包括m3突变。m3为在HC-B结构域界面上七个氨基酸取代六个,从而引入靠近界面引入的额外的糖基化位点。因此,在一些实施方案中,m3为相对于FVIII-FL-AA(SEQ ID NO:19)的氨基酸AIEPRSF755-761的缺失和N754后的氨基酸TTYVNRSL(SEQ ID NO:33)的插入(例如AIEPRSF755-761TTYVNRSL)(“TTYVNRSL”以SEQ ID NO:33形式公开)。相对于SEQ ID NO:19,残基AIEPR755-759落在重链的末端内,而残基S760和F761落在B-结构域内。在FVIII B-结构域缺失、截短或替换的一些实施方案中,残基S760和F761可能不存在于发生突变的基础氨基酸序列中。因此,在一些实施方案中,m3为相对于FVIII-FL-AA(SEQ ID NO:19)的氨基酸AIEPR755-759的缺失和N754后的氨基酸TTYVNRSL(SEQ ID NO:33)的插入(例如AIEPR755-759TTYVNRSL(“TTYVNRSL”以SEQ ID NO:33形式公开)。

[0268] m3变体位于重链与B结构域之间的连接处,并且因此本公开包括包含m3突变的多肽,包括CS01-FL-AAm3、CS01-HC-AAm3、CS04-FL-AAm3、CS04-HC-AAm3、CS23-FL-AAm3、CS23-HC-AAm3、CS40-FL-AAm3以及CS40-HC-AAm3(所有这些均编码相同的相应蛋白序列)。

[0269] 另外,本公开中不仅包括包含m3突变的多肽序列,而且包括那些编码具有m3突变的蛋白质的密码子改变的多核苷酸序列,诸如CS01-FL-NAm3、CS01-HC-NAm3、CS04-FL-NAm3、CS04-HC-NAm3、CS23-FL-NAm3、CS23-HC-NA-m3、CS40-FL-NAm3以及CS40-HC-NAm3。

[0270] 在其他实施方案中,本公开的多肽和多核苷酸包括m4突变。消除因子VIII中的C1899-C1903二硫键也使分泌增加。此外,因子VIII分泌的增加对于F328S(SPI,F309S SPE)和C1918G/C1922G氨基酸取代的组合为相加性的(Miao等,Blood,103:3412-19(2004);

Selvaraj等,J.Thromb.Haemost.,10:107-15(2012))。因此,在一些实施方案中,编码的因子VIII多肽包括m4突变,所述m4突变为F328S(SPI,F309SSPE)和C1918G/C1922G(SPI)氨基酸取代。因为F328S变体位于重链中,而两个半胱氨酸变体位于轻链中,所以包括m4突变的多肽序列为CS01-FL-AAm4、CS01-HC-AAm4、CS01-LC-AAm4、CS04-FL-AAm4、CS04-HC-AAm4、CS04-LC-AAm4、CS23-FL-AAm4、CS23-HC-AAm4以及CS23-LC-AAm4。

[0271] 另外,本公开中不仅包括包含m4突变的多肽序列,而且包括编码具有m4突变的蛋白质的那些密码子改变的多核苷酸序列,诸如CS01-FL-NAm4、CS01-HC-NAm4、CS01-LC-NAm4、CS04-FL-NAm4、CS04-HC-NAm4、CS04-LC-NAm4、CS23-FL-NAm4、CS23-HC-NAm4、CS23-LC-NAm4、CS40-FL-NA-m4、CS40-HC-NA-m4以及CS40-LC-NA-m4。

[0272] 在其他实施方案中,本公开的多肽和多核苷酸包括m5突变。如上所述,消除因子VIII中的C1899-C1903二硫键也使分泌增加。本文中将轻链内所含的C1918G/C1922G(SPI)氨基酸取代称为m5突变集合。

[0273] m5变体位于轻链中,并且因此本公开包括包含m5突变的多肽,包括CS01-FL-AAm5、CS01-LC-AAm5、CS04-FL-AAm5、CS04-LC-AAm5、CS23-FL-AAm5、CS23-LC-AAm5、CS40-FL-AAm5以及CS40-LC-AAm5(所有这些均编码相同的相应蛋白序列)。

[0274] 另外,本公开中不仅包括包含m5突变的多肽序列,而且包括编码具有m5突变的蛋白质的那些密码子改变的多核苷酸序列,诸如CS01-FL-NAm5、CS01-LC-NAm5、CS04-FL-NAm5、CS04-LC-NAm5、CS23-FL-NA-m5、CS23-LC-NA-m5、CS40-FL-NA-m5以及CS40-LC-NA-m5。

[0275] 除了个别地包括m1、m2、m3、m4以及m5的特定构建体(氨基酸与核酸),可如本文所概述形成突变集合的组合。如本文所指出的,这些组合被指定为m1与m2集合的组合“m12”或m1、m2以及m3集合的组合“m123”。因此,本公开中包括双重组合,包括m12、m13、m14、m15、m23、m24、m25、m34、m35以及m45。还包括三重组合,m123、m124、m125、m234、m235以及m345。进一步包括四重组合,m1234、m1235、m1345以及m12345组合。

[0276] 在一些实施方案中特别关注的是以下突变集合:m1、m2、m3和m4、m23、m123以及m234。

[0277] B.因子VIII B-结构域取代的接头

[0278] 在一些实施方案中,进一步改变FVIII重链与轻链之间的键联(例如野生型因子VIII中的B-结构域)。由于AAV包装容量的尺寸限制,B-结构域缺失、截短和/或接头取代的变体会改善FVIII基因疗法构建体的功效。最常用的B-结构域取代的接头是SQ FVIII的接头,它仅保留B结构域的14个氨基酸作为接头序列。猪VIII的另一变体(美国专利号6,458,563中所描述的“OBI-1”)在CHO细胞中很好地表达,并且具有稍长的24个氨基酸的接头。在一些实施方案中,由本文所描述的密码子改变的多核苷酸编码的因子VIII构建体包括SQ型B-结构域接头序列。在其他实施方案中,由本文所描述的密码子改变的多核苷酸编码的因子VIII构建体包括OBI-1型B-结构域接头序列。

[0279] 在一些实施方案中,本文所描述的编码的因子VIII多肽包括SQ型B-结构域接头,所述SQ型B-结构域接头包括野生型人因子VIII B-结构域(FVIII-FL-AA;SEQ ID NO:19)的氨基酸760-762/1657-1667(Sandberg等Thromb.Haemost.85:93(2001))。在一些实施方案中,SQ型B-结构域接头相对于相应野生型序列具有一个氨基酸取代。在一些实施方案中,SQ型B-结构域接头相对于相应野生型序列具有两个氨基酸取代。在一些实施方案中,将糖基

化肽插入SQ型B-结构域接头中。在一些实施方案中，糖基化肽是选自图13中所示的那些糖基化肽(按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75)。

[0280] 在一些实施方案中，本文所描述的编码的因子VIII多肽包括Greengene型B-结构域接头，所述Greengene型B-结构域接头包括野生型人因子VIII B-结构域(FVIII-FL-AA;SEQ ID NO:19)的氨基酸760/1582-1667(Oh等,Biotechnol.Prog.,17:1999(2001))。在一些实施方案中，Greengene型B-结构域接头相对于相应野生型序列具有一个氨基酸取代。在一些实施方案中，Greengene型B-结构域接头相对于相应野生型序列具有两个氨基酸取代。在一些实施方案中，将糖基化肽插入Greengene型B-结构域接头中。在一些实施方案中，糖基化肽是选自图13中所示的那些糖基化肽(按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75)。

[0281] 在一些实施方案中，本文所描述的编码的因子VIII多肽包括延长的SQ型B-结构域接头(SFSQNPPVLKRHQR;BDL-SQ-AA;SEQ ID NO:30)，所述延长的SQ型B-结构域接头包括野生型人因子VIII B-结构域(FVIII-FL-AA;SEQ ID NO:19)的氨基酸760-769/1657-1667(Thim等,Haemophilia,16:349(2010))。在一些实施方案中，延长的SQ型B-结构域接头相对于相应野生型序列具有一个氨基酸取代。在一些实施方案中，延长的SQ型B-结构域接头相对于相应野生型序列具有两个氨基酸取代。在一些实施方案中，将糖基化肽插入延长的SQ型B-结构域接头中。在一些实施方案中，糖基化肽是选自图13中所示的那些糖基化肽(按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75)。

[0282] 在一些实施方案中，本文所描述的编码的因子VIII多肽包括猪OBI-1型B-结构域接头，所述猪OBI-1型B-结构域接头包括来自野生型猪因子VIII B-结构域的氨基酸SFAQNSRPPSASAPKPPVLRRHQR(SEQ ID NO:31)(Toschi等,Curr.Opin.Mol.Ther.12:517(2010))。在一些实施方案中，猪OBI-1型B-结构域接头相对于相应野生型序列具有一个氨基酸取代。在一些实施方案中，猪OBI-1型B-结构域接头相对于相应野生型序列具有两个氨基酸取代。在一些实施方案中，将糖基化肽插入猪OBI-1型B-结构域接头中。在一些实施方案中，糖基化肽是选自图13中所示的那些糖基化肽(按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75)。

[0284] 在一些实施方案中，本文所描述的编码的因子VIII多肽包括人OBI-1型B-结构域接头，所述人OBI-1型B-结构域接头包括野生型人因子VIII B-结构域(FVIII-FL-AA;SEQ ID NO:19)的氨基酸760-772/1655-1667。在一些实施方案中，人OBI-1型B-结构域接头相对于相应野生型序列具有一个氨基酸取代。在一些实施方案中，人OBI-1型B-结构域接头相对于相应野生型序列具有两个氨基酸取代。在一些实施方案中，将糖基化肽插入人OBI-1型B-结构域接头中。在一些实施方案中，糖基化肽是选自图13中所示的那些糖基化肽(按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75)。

[0285] 在一些实施方案中，本文所描述的编码的因子VIII多肽包括O8型B-结构域接头，所述O8型B-结构域接头包括来自野生型猪因子VIII B-结构域的氨基酸SFSQNSRHQAYRYRRG(SEQ ID NO:32)(Toschi等,Curr.Opin.Mol.Ther.12:517(2010))。在一些实施方案中，猪OBI-1型B-结构域接头相对于相应野生型序列具有一个氨基酸取代。在一些实施方案中，猪OBI-1型B-结构域接头相对于相应野生型序列具有两个氨基酸取代。在一些实施方案中，将糖基化肽插入猪OBI-1型B-结构域接头中。在一些实施方案中，糖基化肽是选自图13中所示

的那些糖基化肽(按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75)。

[0286] 从因子VIII构建体去除B-结构域似乎不影响活化酶(例如FVIIIa)的活性,可能是因为在活化过程中去除了B-结构域。然而,因子VIII的B-结构域含有例如通过N-或O-连接的糖基化而翻译后修饰的若干残基。野生型因子VIII B-结构域的计算机模拟分析(Prediction of N-glycosylation sites in human proteins,R.Gupta,E.Jung和S.Brunak,in preparation(2004))预测这些位点中至少有四个在体内被糖基化(图14)。B-结构域内的这些修饰被认为有助于翻译后调控和/或因子VIII在体内的半衰期。

[0287] 虽然成熟因子VIIIa蛋白中不存在因子VIII B-结构域,但前体因子VIII分子的B-结构域内的糖基化可增加蛋白质在活化之前的循环半衰期。因此,在一些实施方案中,本文所描述的编码的因子VIII构建体的多肽接头包括一个或多个糖基化序列,以允许体内糖基化。在一些实施方案中,多肽接头包括至少一个共有糖基化序列(例如N-或O-连接的糖基化共有序列)。在一些实施方案中,多肽接头包括至少两个共有糖基化序列。在一些实施方案中,多肽接头包括至少三个共有糖基化序列。在一些实施方案中,多肽接头包括至少四个共有糖基化序列。在一些实施方案中,多肽接头包括至少五个共有糖基化序列。在一些实施方案中,多肽接头包括至少6、7、8、9、10或更多个共有糖基化序列。

[0288] 在一些实施方案中,多肽接头含有至少一个N-连接的糖基化序列N-X-S/T,其中X为不是P、S或T任何氨基酸。在一些实施方案中,多肽接头含有至少两个N-连接的糖基化序列N-X-S/T,其中X为不是P、S或T任何氨基酸。在一些实施方案中,多肽接头含有至少三个N-连接的糖基化序列N-X-S/T,其中X为不是P、S或T任何氨基酸。在一些实施方案中,多肽接头含有至少四个N-连接的糖基化序列N-X-S/T,其中X为不是P、S或T任何氨基酸。在一些实施方案中,多肽接头含有至少五个N-连接的糖基化序列N-X-S/T,其中X为不是P、S或T任何氨基酸。在一些实施方案中,多肽接头含有至少6、7、8、9、10或更多个N-连接的糖基化序列N-X-S/T,其中X为不是P、S或T任何氨基酸。

[0289] 在一些实施方案中,多肽接头包括与如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有高序列同一性的糖基化肽。在一些实施方案中,糖基化多肽与如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有至少92%的同一性。在一些实施方案中,糖基化肽相对于如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有不超过两个氨基酸取代。在一些实施方案中,糖基化肽相对于如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有不超过一个氨基酸取代。在一些实施方案中,糖基化肽具有选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种的氨基酸序列。

[0290] 在一些实施方案中,糖基化肽与如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有至少92%的同一性,并且由与选自如图13A-13B中所示分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74的相应核苷酸序列具有至少90%同一性的多核苷酸序列编码。在一些实施方案中,糖基化肽与如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、

67、69、71、73以及75中的任一种具有至少92%的同一性，并且由与选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74的相应核苷酸序列具有至少95%同一性的多核苷酸序列编码。在一些实施方案中，糖基化肽与如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有至少92%的同一性，并且由与选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74的相应核苷酸序列具有至少98%同一性的多核苷酸序列编码。

[0291] 在一些实施方案中，糖基化肽相对于如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有不超过两个氨基酸取代，并且由与选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74的相应核苷酸序列具有至少90%同一性的多核苷酸序列编码。在一些实施方案中，糖基化肽相对于如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有不超过两个氨基酸取代，并且由与选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74的相应核苷酸序列具有至少95%同一性的多核苷酸序列编码。在一些实施方案中，糖基化肽相对于如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有不超过两个氨基酸取代，并且由与选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74的相应核苷酸序列具有至少98%同一性的多核苷酸序列编码。

[0292] 在一些实施方案中，糖基化肽相对于如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有不超过一个氨基酸取代，并且由与选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74的相应核苷酸序列具有至少90%同一性的多核苷酸序列编码。在一些实施方案中，糖基化肽相对于如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有不超过一个氨基酸取代，并且由与选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74的相应核苷酸序列具有至少95%同一性的多核苷酸序列编码。在一些实施方案中，糖基化肽相对于如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有不超过一个氨基酸取代，并且由与选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74的相应核苷酸序列具有至少98%同一性的多核苷酸序列编码。

[0293] 在一些实施方案中，糖基化肽具有选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种的序列，并且由与选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74的相应核苷酸序列具有至少90%同一性的多核苷酸序列编码。在一些实施方案中，糖基化肽具有选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种的序列，并且由与选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74的相应核苷酸序列具有至少95%同一性的多核苷酸序列编码。在一些实施方案中，糖基化肽具有选自如图

13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种的序列，并且由与选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74的相应核苷酸序列具有至少98％同一性的多核苷酸序列编码。

[0294] 在一些实施方案中，由本文所描述的密码子改变的多核苷酸编码的因子VIII多肽具有B-结构域取代的接头，其中糖基化肽被插入SQ接头序列(CS04-FL-AA的氨基酸760-773;SEQ ID NO:2)中。在一个特定实施方案中，糖基化肽是选自：选自如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种的糖基化肽；与如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有至少92％同一性的糖基化肽；相对于如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有不超过两个氨基酸取代的糖基化肽；以及相对于如图13A-13B中所示按出现顺序分别为SEQ ID NO 51、53、55、57、59、61、63、65、67、69、71、73以及75中的任一种具有不超过一个氨基酸取代的糖基化肽。在一些实施方案中，将糖基化肽插入残基N768与P769(相对于CS04-FL-AA;SEQ ID NO:2)之间的SQ肽中。

[0295] 在一些实施方案中，因子VIII构建体的多肽接头是由与图6中所示的那些序列(按出现顺序分别为SEQ ID NO 5-7和36-48)中的任一种具有高序列同一性的第三核苷酸序列编码。在一些实施方案中，第三核苷酸序列与图13中所示的那些序列(按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74)中的任一种具有至少95％的同一性。在一些实施方案中，第三核苷酸序列与图13中所示的那些序列(按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74)中的任一种具有至少96％的同一性。在一些实施方案中，第三核苷酸序列与图13中所示的那些序列(按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74)中的任一种具有至少97％的同一性。在一些实施方案中，第三核苷酸序列与图13中所示的那些序列(按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74)中的任一种具有至少98％的同一性。在一些实施方案中，第三核苷酸序列与图13中所示的那些序列(按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74)中的任一种具有至少99％的同一性。在一些实施方案中，第三核苷酸序列与图13中所示的那些序列(按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74)中的任一种具有至少99.5％的同一性。在一些实施方案中，第三核苷酸序列与图13中所示的那些序列(按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74)中的任一种具有至少99.9％的同一性。在一些实施方案中，第三核苷酸序列与图13中所示的那些序列(按出现顺序分别为SEQ ID NO 50、52、54、56、58、60、62、64、66、68、70、72以及74)中的任一种相同。

[0296] C.具有可裂解接头的编码因子VIII变体的密码子改变的多核苷酸

[0297] CS04密码子改变的多核苷酸

[0298] 在一个实施方案中，本文所提供的密码子改变的多核苷酸包括编码具有在体内可裂解的接头的因子VIII变体多肽的核苷酸序列。因子VIII多肽包括因子VIII轻链、因子VIII重链以及将重链的C端连接至轻链的N端的多肽接头。因子VIII多肽的重链是由与CS04-HC-NA(SEQ ID NO:3)具有高序列同一性的第一核苷酸序列编码，CS04-HC-NA为编码

因子VIII重链的CS04-FL-NA(SEQ ID NO:1)的一部分。因子VIII多肽的轻链是由与CS04-LC-NA(SEQ ID NO:4)具有高序列同一性的第二核苷酸序列编码,CS04-LC-NA为编码因子VIII轻链的CS04-FL-NA(SEQ ID NO:1)的一部分。多肽接头包括弗林蛋白酶裂解位点,所述弗林蛋白酶裂解位点允许体内成熟(例如在前体多肽在体内表达或施用之后)。

[0299] 在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少95%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少96%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少97%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少98%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少99%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少99.5%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少99.9%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)相同。

[0300] 在一些实施方案中,因子VIII构建体的多肽接头是由与BDL004(SEQ ID NO:6)具有高序列同一性的第三核苷酸序列编码,BDL004编码对应于CS04-FL-AA(SEQ ID NO:2)的氨基酸760-773的14-氨基酸接头。在一些实施方案中,第三核苷酸序列与BDL004(SEQ ID NO:6)具有至少95%的同一性。在一些实施方案中,第三核苷酸序列与BDL004(SEQ ID NO:6)具有至少96%的同一性。在一些实施方案中,第三核苷酸序列与BDL004(SEQ ID NO:6)具有至少97%的同一性。在一些实施方案中,第三核苷酸序列与BDL004(SEQ ID NO:6)具有至少98%的同一性。在一些实施方案中,第三核苷酸序列与BDL004(SEQ ID NO:6)相同。

[0301] 在一些实施方案中,密码子改变的多核苷酸具有与CS04-FL-NA(SEQ ID NO:1)具有高序列同一性的核苷酸序列。在一些实施方案中,核苷酸序列与CS04-FL-NA(SEQ ID NO:1)具有至少95%的同一性。在一些实施方案中,核苷酸序列与CS04-FL-NA(SEQ ID NO:1)具有至少96%的同一性。在一些实施方案中,核苷酸序列与CS04-FL-NA(SEQ ID NO:1)具有至少97%的同一性。在一些实施方案中,核苷酸序列与CS04-FL-NA(SEQ ID NO:1)具有至少98%的同一性。在一些实施方案中,核苷酸序列与CS04-FL-NA(SEQ ID NO:1)具有至少99%的同一性。在一些实施方案中,核苷酸序列与CS04-FL-NA(SEQ ID NO:1)具有至少99.5%的同一性。在一些实施方案中,核苷酸序列与CS04-FL-NA(SEQ ID NO:1)具有至少99.9%的同一性。在一些实施方案中,核苷酸序列与CS04-FL-NA(SEQ ID NO:1)相同。

[0302] 在一些实施方案中,由密码子改变的多核苷酸编码的因子VIII变体具有与CS04-FL-AA(SEQ ID NO:2)具有高序列同一性的氨基酸序列。在一些实施方案中,氨基酸序列与CS04-FL-AA(SEQ ID NO:2)具有至少97%的同一性。在一些实施方案中,氨基酸序列与CS04-FL-AA(SEQ ID NO:2)具有至少98%的同一性。在一些实施方案中,氨基酸序列与CS04-FL-AA(SEQ ID NO:2)具有至少99%的同一性。在一些实施方案中,氨基酸序列与CS04-FL-AA(SEQ ID NO:2)具有至少99.5%的同一性。在一些实施方案中,氨基酸序列与

CS04-FL-AA(SEQ ID NO:2)具有至少99.9%的同一性。在一些实施方案中,氨基酸序列与CS04-FL-AA(SEQ ID NO:2)相同。

[0303] 在一些实施方案中,由CS04多核苷酸编码的与CS04-FL-AA具有高序列同源性(例如至少95%、96%、97%、98%、99%、99.5%或99.9%的同一性)的因子VIII变体包含一个或多个选自m1、m2、m3、m4以及m5的氨基酸取代。

[0304] 在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m1氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m2氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m3氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m4氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m5氨基酸取代。

[0305] 在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m12氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m13氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m23氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m24氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m25氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m34氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m35氨基酸取代。

[0306] 在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m123氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m234氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m125氨基酸取代。

[0307] CS01密码子改变的多核苷酸

[0308] 在一个实施方案中,本文所提供的密码子改变的多核苷酸包括编码具有在体内可裂解的接头的因子VIII变体多肽的核苷酸序列。因子VIII多肽包括因子VIII轻链、因子VIII重链以及将重链的C端连接至轻链的N端的多肽接头。因子VIII多肽的重链是由与CS01-HC-NA(SEQ ID NO:24)具有高序列同一性的第一核苷酸序列编码,CS01-HC-NA为编码因子VIII重链的CS01-FL-NA(SEQ ID NO:13)的一部分。因子VIII多肽的轻链是由与CS01-LC-NA(SEQ ID NO:25)具有高序列同一性的第二核苷酸序列编码,CS01-LC-NA为编码因子VIII轻链的CS01-FL-NA(SEQ ID NO:13)的一部分。多肽接头包括弗林蛋白酶裂解位点,所述弗林蛋白酶裂解位点允许体内成熟(例如在前体多肽在体内表达或施用之后)。

[0309] 在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少95%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少96%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少97%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少98%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少99%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少99.5%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸

序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少99.9%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)相同。

[0310] 在一些实施方案中,因子VIII构建体的多肽接头是由与BDL004(SEQ ID NO:6)具有高序列同一性的第三核苷酸序列编码,BDL004编码对应于CS01-FL-AA(SEQ ID NO:2)的氨基酸760-773的14-氨基酸接头。在一些实施方案中,第三核苷酸序列与BDL004(SEQ ID NO:6)具有至少95%的同一性。在一些实施方案中,第三核苷酸序列与BDL004(SEQ ID NO:6)具有至少96%的同一性。在一些实施方案中,第三核苷酸序列与BDL004(SEQ ID NO:6)具有至少97%的同一性。在一些实施方案中,第三核苷酸序列与BDL004(SEQ ID NO:6)具有至少98%的同一性。在一些实施方案中,第三核苷酸序列与BDL004(SEQ ID NO:6)相同。

[0311] 在一些实施方案中,密码子改变的多核苷酸具有与CS01-FL-NA(SEQ ID NO:13)具有高序列同一性的核苷酸序列。在一些实施方案中,核苷酸序列与CS01-FL-NA(SEQ ID NO:13)具有至少95%的同一性。在一些实施方案中,核苷酸序列与CS01-FL-NA(SEQ ID NO:13)具有至少96%的同一性。在一些实施方案中,核苷酸序列与CS01-FL-NA(SEQ ID NO:13)具有至少97%的同一性。在一些实施方案中,核苷酸序列与CS01-FL-NA(SEQ ID NO:13)具有至少98%的同一性。在一些实施方案中,核苷酸序列与CS01-FL-NA(SEQ ID NO:13)具有至少99%的同一性。在一些实施方案中,核苷酸序列与CS01-FL-NA(SEQ ID NO:13)具有至少99.5%的同一性。在一些实施方案中,核苷酸序列与CS01-FL-NA(SEQ ID NO:13)具有至少99.9%的同一性。在一些实施方案中,核苷酸序列与CS01-FL-NA(SEQ ID NO:13)相同。

[0312] 在一些实施方案中,由密码子改变的多核苷酸编码的因子VIII变体具有与CS01-FL-AA(SEQ ID NO:2)具有高序列同一性的氨基酸序列。在一些实施方案中,氨基酸序列与CS01-FL-AA(SEQ ID NO:2)具有至少97%的同一性。在一些实施方案中,氨基酸序列与CS01-FL-AA(SEQ ID NO:2)具有至少98%的同一性。在一些实施方案中,氨基酸序列与CS01-FL-AA(SEQ ID NO:2)具有至少99%的同一性。在一些实施方案中,氨基酸序列与CS01-FL-AA(SEQ ID NO:2)具有至少99.5%的同一性。在一些实施方案中,氨基酸序列与CS01-FL-AA(SEQ ID NO:2)具有至少99.9%的同一性。在一些实施方案中,氨基酸序列与CS01-FL-AA(SEQ ID NO:2)相同。

[0313] 在一些实施方案中,由CS01多核苷酸编码的与CS01-FL-AA具有高序列同源性(例如至少95%、96%、97%、98%、99%、99.5%或99.9%的同一性)的因子VIII变体包含一个或多个选自m1、m2、m3、m4以及m5的氨基酸取代。

[0314] 在一个实施方案中,由CS01多核苷酸编码的因子VIII变体包含m1氨基酸取代。在一个实施方案中,由CS01多核苷酸编码的因子VIII变体包含m2氨基酸取代。在一个实施方案中,由CS01多核苷酸编码的因子VIII变体包含m3氨基酸取代。在一个实施方案中,由CS01多核苷酸编码的因子VIII变体包含m4氨基酸取代。在一个实施方案中,由CS01多核苷酸编码的因子VIII变体包含m5氨基酸取代。

[0315] 在一个实施方案中,由CS01多核苷酸编码的因子VIII变体包含m12氨基酸取代。在一个实施方案中,由CS01多核苷酸编码的因子VIII变体包含m13氨基酸取代。在一个实施方案中,由CS01多核苷酸编码的因子VIII变体包含m23氨基酸取代。在一个实施方案中,由CS01多核苷酸编码的因子VIII变体包含m24氨基酸取代。在一个实施方案中,由CS01多核苷

酸编码的因子VIII变体包含m25氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m34氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m35氨基酸取代。

[0316] 在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m123氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m234氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m125氨基酸取代。

[0317] CS23密码子改变的多核苷酸

[0318] 在一个实施方案中，本文所提供的密码子改变的多核苷酸包括编码具有在体内可裂解的接头的因子VIII变体多肽的核苷酸序列。因子VIII多肽包括因子VIII轻链、因子VIII重链以及将重链的C端连接至轻链的N端的多肽接头。因子VIII多肽的重链是由与CS23-HC-NA(SEQ ID NO:22)具有高序列同一性的第一核苷酸序列编码，CS23-HC-NA为编码因子VIII重链的CS23-FL-NA(SEQ ID NO:20)的一部分。因子VIII多肽的轻链是由与CS23-LC-NA(SEQ ID NO:23)具有高序列同一性的第二核苷酸序列编码，CS23-LC-NA为编码因子VIII轻链的CS23-FL-NA(SEQ ID NO:20)的一部分。多肽接头包括弗林蛋白酶裂解位点，所述弗林蛋白酶裂解位点允许体内成熟(例如在前体多肽在体内表达或施用之后)。

[0319] 在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少95％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少96％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少97％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少98％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少99％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少99.5％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少99.9％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)相同。

[0320] 在一些实施方案中，因子VIII构建体的多肽接头是由与BDLO04(SEQ ID NO:6)具有高序列同一性的第三核苷酸序列编码，BDLO04编码对应于CS23-FL-AA(SEQ ID NO:21)的氨基酸760-773的14-氨基酸接头。在一些实施方案中，第三核苷酸序列与BDLO04(SEQ ID NO:6)具有至少95％的同一性。在一些实施方案中，第三核苷酸序列与BDLO04(SEQ ID NO:6)具有至少96％的同一性。在一些实施方案中，第三核苷酸序列与BDLO04(SEQ ID NO:6)具有至少97％的同一性。在一些实施方案中，第三核苷酸序列与BDLO04(SEQ ID NO:6)具有至少98％的同一性。在一些实施方案中，第三核苷酸序列与BDLO04(SEQ ID NO:6)相同。

[0321] 在一些实施方案中，密码子改变的多核苷酸具有与CS23-FL-NA(SEQ ID NO:20)具有高序列同一性的核苷酸序列。在一些实施方案中，核苷酸序列与CS23-FL-NA(SEQ ID NO:20)具有至少95％的同一性。在一些实施方案中，核苷酸序列与CS23-FL-NA(SEQ ID NO:20)具有至少96％的同一性。在一些实施方案中，核苷酸序列与CS23-FL-NA(SEQ ID NO:20)具

有至少97％的同一性。在一些实施方案中，核苷酸序列与CS23-FL-NA(SEQ ID NO:20)具有至少98％的同一性。在一些实施方案中，核苷酸序列与CS23-FL-NA(SEQ ID NO:20)具有至少99％的同一性。在一些实施方案中，核苷酸序列与CS23-FL-NA(SEQ ID NO:20)具有至少99.5％的同一性。在一些实施方案中，核苷酸序列与CS23-FL-NA(SEQ ID NO:20)具有至少99.9％的同一性。在一些实施方案中，核苷酸序列与CS23-FL-NA(SEQ ID NO:20)相同。

[0322] 在一些实施方案中，由密码子改变的多核苷酸编码的因子VIII变体具有与CS23-FL-AA(SEQ ID NO:21)具有高序列同一性的氨基酸序列。在一些实施方案中，氨基酸序列与CS23-FL-AA(SEQ ID NO:21)具有至少97％的同一性。在一些实施方案中，氨基酸序列与CS23-FL-AA(SEQ ID NO:21)具有至少98％的同一性。在一些实施方案中，氨基酸序列与CS23-FL-AA(SEQ ID NO:21)具有至少99％的同一性。在一些实施方案中，氨基酸序列与CS23-FL-AA(SEQ ID NO:21)具有至少99.5％的同一性。在一些实施方案中，氨基酸序列与CS23-FL-AA(SEQ ID NO:21)具有至少99.9％的同一性。在一些实施方案中，氨基酸序列与CS23-FL-AA(SEQ ID NO:21)相同。

[0323] 在一些实施方案中，由CS23多核苷酸编码的与CS23-FL-AA具有高序列同源性(例如至少95％、96％、97％、98％、99％、99.5％或99.9％的同一性)的因子VIII变体包含一个或多个选自m1、m2、m3、m4以及m5的氨基酸取代。

[0324] 在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m1氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m2氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m3氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m4氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m5氨基酸取代。

[0325] 在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m12氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m13氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m23氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m24氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m25氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m34氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m35氨基酸取代。

[0326] 在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m123氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m234氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m125氨基酸取代。

[0327] D.编码单链因子VIII蛋白的密码子改变的多核苷酸

[0328] 尽管不能发生因子VIII分子的正常成熟，但定位于B-结构域的C末端的弗林蛋白酶裂解位点被去除的因子VIII构建体保留如单链多肽的活性(Leyte等(1991))。类似地，具有减弱的弗林蛋白酶位点(含有R1664H氨基酸取代)的B-结构域缺失的因子VIII构建体比具有野生型弗林蛋白酶裂解位点的相应因子VIII构建体更具生物活性(Siner等(2013))。因此，在一些实施方案中，本文所提供的密码子改变的多核苷酸包括编码单链因子VIII变体多肽的核苷酸序列。单链因子VIII多肽包括因子VIII轻链、因子VIII重链以及将重链的C端连接至轻链的N端的多肽接头。多肽接头不包括弗林蛋白酶裂解位点。

[0329] 单链CS04密码子改变的多核苷酸

[0330] 在一个实施方案中,本文所提供的密码子改变的多核苷酸包括编码单链因子VIII变体多肽的核苷酸序列。因子VIII多肽包括因子VIII轻链、因子VIII重链以及任选的将重链的C端连接至轻链的N端的多肽接头。因子VIII多肽的重链是由与CS04-HC-NA(SEQ ID NO:3)具有高序列同一性的第一核苷酸序列编码,CS04-HC-NA为编码因子VIII重链的CS04-FL-NA(SEQ ID NO:1)的一部分。因子VIII多肽的轻链是由与CS04-LC-NA(SEQ ID NO:4)具有高序列同一性的第二核苷酸序列编码,CS04-LC-NA为编码因子VIII轻链的CS04-FL-NA(SEQ ID NO:1)的一部分。任选的多肽接头不包括弗林蛋白酶裂解位点。

[0331] 在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少95%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少96%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少97%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少98%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少99%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少99.5%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)具有至少99.9%的序列同一性。在一些实施方案中,第一核苷酸序列和第二核苷酸序列分别与CS04-HC-NA和CS04-LC-NA(SEQ ID NO 3和4)相同。

[0332] 在一些实施方案中,密码子改变的多核苷酸具有与CS04-SC1-NA(SEQ ID NO:9)具有高序列同一性的核苷酸序列。在一些实施方案中,核苷酸序列与CS04-SC1-NA(SEQ ID NO:9)具有至少95%的同一性。在一些实施方案中,核苷酸序列与CS04-SC1-NA(SEQ ID NO:9)具有至少96%的同一性。在一些实施方案中,核苷酸序列与CS04-SC1-NA(SEQ ID NO:9)具有至少97%的同一性。在一些实施方案中,核苷酸序列与CS04-SC1-NA(SEQ ID NO:9)具有至少98%的同一性。在一些实施方案中,核苷酸序列与CS04-SC1-NA(SEQ ID NO:9)具有至少99%的同一性。在一些实施方案中,核苷酸序列与CS04-SC1-NA(SEQ ID NO:9)具有至少99.5%的同一性。在一些实施方案中,核苷酸序列与CS04-SC1-NA(SEQ ID NO:9)具有至少99.9%的同一性。在一些实施方案中,核苷酸序列与CS04-SC1-NA(SEQ ID NO:9)相同。

[0333] 在一些实施方案中,密码子改变的多核苷酸具有与CS04-SC2-NA(SEQ ID NO:11)具有高序列同一性的核苷酸序列。在一些实施方案中,核苷酸序列与CS04-SC2-NA(SEQ ID NO:11)具有至少95%的同一性。在一些实施方案中,核苷酸序列与CS04-SC2-NA(SEQ ID NO:11)具有至少96%的同一性。在一些实施方案中,核苷酸序列与CS04-SC2-NA(SEQ ID NO:11)具有至少97%的同一性。在一些实施方案中,核苷酸序列与CS04-SC2-NA(SEQ ID NO:11)具有至少98%的同一性。在一些实施方案中,核苷酸序列与CS04-SC2-NA(SEQ ID NO:11)具有至少99%的同一性。在一些实施方案中,核苷酸序列与CS04-SC2-NA(SEQ ID NO:11)具有至少99.5%的同一性。在一些实施方案中,核苷酸序列与CS04-SC2-NA(SEQ ID NO:11)具有至少99.9%的同一性。在一些实施方案中,核苷酸序列与CS04-SC2-NA(SEQ ID

NO:11)相同。

[0334] 在一些实施方案中,由密码子改变的多核苷酸编码的单链因子VIII变体具有与CS04-SC1-AA(SEQ ID NO:10;人因子VIIIΔ(760-1667)(SPI;HsFVIIIΔ(741-1648),SPE))具有高序列同一性的氨基酸序列。在一些实施方案中,由密码子改变的多核苷酸编码的因子VIII变体具有与CS04-SC1-AA(SEQ ID NO:10)具有高序列同一性的氨基酸序列。在一些实施方案中,氨基酸序列与CS04-SC1-AA(SEQ ID NO:10)具有至少97%的同一性。在一些实施方案中,氨基酸序列与CS04-SC1-AA(SEQ ID NO:10)具有至少98%的同一性。在一些实施方案中,氨基酸序列与CS04-SC1-AA(SEQ ID NO:10)具有至少99%的同一性。在一些实施方案中,氨基酸序列与CS04-SC1-AA(SEQ ID NO:10)具有至少99.5%的同一性。在一些实施方案中,氨基酸序列与CS04-SC1-AA(SEQ ID NO:10)具有至少99.9%的同一性。在一些实施方案中,氨基酸序列与CS04-SC1-AA(SEQ ID NO:10)相同。

[0335] 在一些实施方案中,由CS04-SC1多核苷酸编码的与CS04-SC1-AA具有高序列同源性(例如至少95%、96%、97%、98%、99%、99.5%或99.9%的同一性)的因子VIII变体包含一个或多个选自m1、m2、m3、m4以及m5的氨基酸取代。

[0336] 在一些实施方案中,由密码子改变的多核苷酸编码的单链因子VIII变体具有与CS04-SC2-AA(SEQ ID NO:12;人因子VIIIΔ(772-1667)(SPI;HsFVIIIΔ(753-1648),SPE))具有高序列同一性的氨基酸序列。在一些实施方案中,由密码子改变的多核苷酸编码的因子VIII变体具有与CS04-SC2-AA(SEQ ID NO:12)具有高序列同一性的氨基酸序列。在一些实施方案中,氨基酸序列与CS04-SC2-AA(SEQ ID NO:12)具有至少97%的同一性。在一些实施方案中,氨基酸序列与CS04-SC2-AA(SEQ ID NO:12)具有至少98%的同一性。在一些实施方案中,氨基酸序列与CS04-SC2-AA(SEQ ID NO:12)具有至少99%的同一性。在一些实施方案中,氨基酸序列与CS04-SC2-AA(SEQ ID NO:12)具有至少99.5%的同一性。在一些实施方案中,氨基酸序列与CS04-SC2-AA(SEQ ID NO:12)具有至少99.9%的同一性。在一些实施方案中,氨基酸序列与CS04-SC2-AA(SEQ ID NO:12)相同。

[0337] 在一些实施方案中,由CS04-SC2多核苷酸编码的与CS04-SC2-AA具有高序列同源性(例如至少95%、96%、97%、98%、99%、99.5%或99.9%的同一性)的单链因子VIII变体包含一个或多个选自m1、m2、m3、m4以及m5的氨基酸取代。

[0338] 在一个实施方案中,由CS04多核苷酸编码的单链因子VIII变体包含m1氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m2氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m3氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m4氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m5氨基酸取代。

[0339] 在一个实施方案中,由CS04多核苷酸编码的单链因子VIII变体包含m12氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m13氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m23氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m24氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m25氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m34氨基酸取代。在一个实施方案中,由CS04多核苷酸编码的因子VIII变体包含m35氨基酸取代。

[0340] 在一个实施方案中，由CS04多核苷酸编码的单链因子VIII变体包含m123氨基酸取代。在一个实施方案中，由CS04多核苷酸编码的因子VIII变体包含m234氨基酸取代。在一个实施方案中，由CS04多核苷酸编码的因子VIII变体包含m125氨基酸取代。

[0341] 单链CS01密码子改变的多核苷酸

[0342] 在一个实施方案中，本文所提供的密码子改变的多核苷酸包括编码单链因子VIII变体多肽的核苷酸序列。因子VIII多肽包括因子VIII轻链、因子VIII重链以及任选的将重链的C端连接至轻链的N端的多肽接头。因子VIII多肽的重链是由与CS01-HC-NA(SEQ ID NO:24)具有高序列同一性的第一核苷酸序列编码，CS01-HC-NA为编码因子VIII重链的CS01-FL-NA(SEQ ID NO:13)的一部分。因子VIII多肽的轻链是由与CS01-LC-NA(SEQ ID NO:25)具有高序列同一性的第二核苷酸序列编码，CS01-LC-NA为编码因子VIII轻链的CS01-FL-NA(SEQ ID NO:13)的一部分。任选的多肽接头不包括弗林蛋白酶裂解位点。

[0343] 在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少95％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少96％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少97％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少98％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少99％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少99.5％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)具有至少99.9％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS01-HC-NA和CS01-LC-NA(SEQ ID NO 24和25)相同。

[0344] 在一些实施方案中，密码子改变的多核苷酸具有与CS01-SC1-NA(SEQ ID NO:26)具有高序列同一性的核苷酸序列。在一些实施方案中，核苷酸序列与CS01-SC1-NA(SEQ ID NO:26)具有至少95％的同一性。在一些实施方案中，核苷酸序列与CS01-SC1-NA(SEQ ID NO:26)具有至少96％的同一性。在一些实施方案中，核苷酸序列与CS01-SC1-NA(SEQ ID NO:26)具有至少97％的同一性。在一些实施方案中，核苷酸序列与CS01-SC1-NA(SEQ ID NO:26)具有至少98％的同一性。在一些实施方案中，核苷酸序列与CS01-SC1-NA(SEQ ID NO:26)具有至少99％的同一性。在一些实施方案中，核苷酸序列与CS01-SC1-NA(SEQ ID NO:26)具有至少99.5％的同一性。在一些实施方案中，核苷酸序列与CS01-SC1-NA(SEQ ID NO:26)具有至少99.9％的同一性。在一些实施方案中，核苷酸序列与CS01-SC1-NA(SEQ ID NO:26)相同。

[0345] 在一些实施方案中，密码子改变的多核苷酸具有与CS01-SC2-NA(SEQ ID NO:27)具有高序列同一性的核苷酸序列。在一些实施方案中，核苷酸序列与CS01-SC2-NA(SEQ ID NO:27)具有至少95％的同一性。在一些实施方案中，核苷酸序列与CS01-SC2-NA(SEQ ID NO:27)具有至少96％的同一性。在一些实施方案中，核苷酸序列与CS01-SC2-NA(SEQ ID NO:27)具有至少97％的同一性。在一些实施方案中，核苷酸序列与CS01-SC2-NA(SEQ ID

NO:27)具有至少98％的同一性。在一些实施方案中，核苷酸序列与CS01-SC2-NA(SEQ ID NO:27)具有至少99％的同一性。在一些实施方案中，核苷酸序列与CS01-SC2-NA(SEQ ID NO:27)具有至少99.5％的同一性。在一些实施方案中，核苷酸序列与CS01-SC2-NA(SEQ ID NO:27)具有至少99.9％的同一性。在一些实施方案中，核苷酸序列与CS01-SC2-NA(SEQ ID NO:27)相同。

[0346] 在一些实施方案中，由密码子改变的多核苷酸编码的单链因子VIII变体具有与CS01-SC1-AA(SEQ ID NO:10;人因子VIIIΔ(760-1667)(SPI;HsFVIIIΔ(741-1648),SPE))具有高序列同一性的氨基酸序列。在一些实施方案中，由密码子改变的多核苷酸编码的因子VIII变体具有与CS01-SC1-AA(SEQ ID NO:10)具有高序列同一性的氨基酸序列。在一些实施方案中，氨基酸序列与CS01-SC1-AA(SEQ ID NO:10)具有至少97％的同一性。在一些实施方案中，氨基酸序列与CS01-SC1-AA(SEQ ID NO:10)具有至少98％的同一性。在一些实施方案中，氨基酸序列与CS01-SC1-AA(SEQ ID NO:10)具有至少99％的同一性。在一些实施方案中，氨基酸序列与CS01-SC1-AA(SEQ ID NO:10)具有至少99.5％的同一性。在一些实施方案中，氨基酸序列与CS01-SC1-AA(SEQ ID NO:10)具有至少99.9％的同一性。在一些实施方案中，氨基酸序列与CS01-SC1-AA(SEQ ID NO:10)相同。

[0347] 在一些实施方案中，由CS01-SC1多核苷酸编码的与CS01-SC1-AA具有高序列同源性(例如至少95％、96％、97％、98％、99％、99.5％或99.9％的同一性)的因子VIII变体包含一个或多个选自m1、m2、m3、m4以及m5的氨基酸取代。

[0348] 在一些实施方案中，由密码子改变的多核苷酸编码的单链因子VIII变体具有与CS01-SC2-AA(SEQ ID NO:12;人因子VIIIΔ(772-1667)(SPI;HsFVIIIΔ(753-1648),SPE))具有高序列同一性的氨基酸序列。在一些实施方案中，由密码子改变的多核苷酸编码的因子VIII变体具有与CS01-SC2-AA(SEQ ID NO:12)具有高序列同一性的氨基酸序列。在一些实施方案中，氨基酸序列与CS01-SC2-AA(SEQ ID NO:12)具有至少97％的同一性。在一些实施方案中，氨基酸序列与CS01-SC2-AA(SEQ ID NO:12)具有至少98％的同一性。在一些实施方案中，氨基酸序列与CS01-SC2-AA(SEQ ID NO:12)具有至少99％的同一性。在一些实施方案中，氨基酸序列与CS01-SC2-AA(SEQ ID NO:12)具有至少99.5％的同一性。在一些实施方案中，氨基酸序列与CS01-SC2-AA(SEQ ID NO:12)具有至少99.9％的同一性。在一些实施方案中，氨基酸序列与CS01-SC2-AA(SEQ ID NO:12)相同。

[0349] 在一些实施方案中，由CS01-SC2多核苷酸编码的与CS01-SC2-AA具有高序列同源性(例如至少95％、96％、97％、98％、99％、99.5％或99.9％的同一性)的单链因子VIII变体包含一个或多个选自m1、m2、m3、m4以及m5的氨基酸取代。

[0350] 在一个实施方案中，由CS01多核苷酸编码的单链因子VIII变体包含m1氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m2氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m3氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m4氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m5氨基酸取代。

[0351] 在一个实施方案中，由CS01多核苷酸编码的单链因子VIII变体包含m12氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m13氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m23氨基酸取代。在一个实施方案

中，由CS01多核苷酸编码的因子VIII变体包含m24氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m25氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m34氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m35氨基酸取代。

[0352] 在一个实施方案中，由CS01多核苷酸编码的单链因子VIII变体包含m123氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m234氨基酸取代。在一个实施方案中，由CS01多核苷酸编码的因子VIII变体包含m125氨基酸取代。

[0353] 单链CS23密码子改变的多核苷酸

[0354] 在一个实施方案中，本文所提供的密码子改变的多核苷酸包括编码单链因子VIII变体多肽的核苷酸序列。因子VIII多肽包括因子VIII轻链、因子VIII重链以及任选的将重链的C端连接至轻链的N端的多肽接头。因子VIII多肽的重链是由与CS23-HC-NA(SEQ ID NO:22)具有高序列同一性的第一核苷酸序列编码，CS23-HC-NA为编码因子VIII重链的CS23-FL-NA(SEQ ID NO:20)的一部分。因子VIII多肽的轻链是由与CS23-LC-NA(SEQ ID NO:23)具有高序列同一性的第二核苷酸序列编码，CS23-LC-NA为编码因子VIII轻链的CS23-FL-NA(SEQ ID NO:20)的一部分。任选的多肽接头不包括弗林蛋白酶裂解位点。

[0355] 在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少95％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少96％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少97％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少98％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少99％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少99.5％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)具有至少99.9％的序列同一性。在一些实施方案中，第一核苷酸序列和第二核苷酸序列分别与CS23-HC-NA和CS23-LC-NA(SEQ ID NO 22和23)相同。

[0356] 在一些实施方案中，密码子改变的多核苷酸具有与CS23-SC1-NA(SEQ ID NO:28)具有高序列同一性的核苷酸序列。在一些实施方案中，核苷酸序列与CS23-SC1-NA(SEQ ID NO:28)具有至少95％的同一性。在一些实施方案中，核苷酸序列与CS23-SC1-NA(SEQ ID NO:28)具有至少96％的同一性。在一些实施方案中，核苷酸序列与CS23-SC1-NA(SEQ ID NO:28)具有至少97％的同一性。在一些实施方案中，核苷酸序列与CS23-SC1-NA(SEQ ID NO:28)具有至少98％的同一性。在一些实施方案中，核苷酸序列与CS23-SC1-NA(SEQ ID NO:28)具有至少99％的同一性。在一些实施方案中，核苷酸序列与CS23-SC1-NA(SEQ ID NO:28)具有至少99.5％的同一性。在一些实施方案中，核苷酸序列与CS23-SC1-NA(SEQ ID NO:28)具有至少99.9％的同一性。在一些实施方案中，核苷酸序列与CS23-SC1-NA(SEQ ID NO:28)相同。

[0357] 在一些实施方案中，密码子改变的多核苷酸具有与CS23-SC2-NA(SEQ ID NO:29)

具有高序列同一性的核苷酸序列。在一些实施方案中，核苷酸序列与CS23-SC2-NA(SEQ ID NO:29)具有至少95％的同一性。在一些实施方案中，核苷酸序列与CS23-SC2-NA(SEQ ID NO:29)具有至少96％的同一性。在一些实施方案中，核苷酸序列与CS23-SC2-NA(SEQ ID NO:29)具有至少97％的同一性。在一些实施方案中，核苷酸序列与CS23-SC2-NA(SEQ ID NO:29)具有至少98％的同一性。在一些实施方案中，核苷酸序列与CS23-SC2-NA(SEQ ID NO:29)具有至少99％的同一性。在一些实施方案中，核苷酸序列与CS23-SC2-NA(SEQ ID NO:29)具有至少99.5％的同一性。在一些实施方案中，核苷酸序列与CS23-SC2-NA(SEQ ID NO:29)具有至少99.9％的同一性。在一些实施方案中，核苷酸序列与CS23-SC2-NA(SEQ ID NO:29)相同。

[0358] 在一些实施方案中，由密码子改变的多核苷酸编码的单链因子VIII变体具有与CS23-SC1-AA(SEQ ID NO:10;人因子VIIIΔ(760-1667)(SPI;CS04Δ(741-1648),SPE))具有高序列同一性的氨基酸序列。在一些实施方案中，由密码子改变的多核苷酸编码的因子VIII变体具有与CS23-SC1-AA(SEQ ID NO:10)具有高序列同一性的氨基酸序列。在一些实施方案中，氨基酸序列与CS23-SC1-AA(SEQ ID NO:10)具有至少97％的同一性。在一些实施方案中，氨基酸序列与CS23-SC1-AA(SEQ ID NO:10)具有至少98％的同一性。在一些实施方案中，氨基酸序列与CS23-SC1-AA(SEQ ID NO:10)具有至少99％的同一性。在一些实施方案中，氨基酸序列与CS23-SC1-AA(SEQ ID NO:10)具有至少99.5％的同一性。在一些实施方案中，氨基酸序列与CS23-SC1-AA(SEQ ID NO:10)具有至少99.9％的同一性。在一些实施方案中，氨基酸序列与CS23-SC1-AA(SEQ ID NO:10)相同。

[0359] 在一些实施方案中，由CS23-SC1多核苷酸编码的与CS23-SC1-AA具有高序列同源性(例如至少95％、96％、97％、98％、99％、99.5％或99.9％的同一性)的因子VIII变体包含一个或多个选自m1、m2、m3、m4以及m5的氨基酸取代。

[0360] 在一些实施方案中，由密码子改变的多核苷酸编码的单链因子VIII变体具有与CS23-SC2-AA(SEQ ID NO:12;人因子VIIIΔ(772-1667)(SPI;HsFVIIIΔ(753-1648),SPE))具有高序列同一性的氨基酸序列。在一些实施方案中，由密码子改变的多核苷酸编码的因子VIII变体具有与CS23-SC2-AA(SEQ ID NO:12)具有高序列同一性的氨基酸序列。在一些实施方案中，氨基酸序列与CS23-SC2-AA(SEQ ID NO:12)具有至少97％的同一性。在一些实施方案中，氨基酸序列与CS23-SC2-AA(SEQ ID NO:12)具有至少98％的同一性。在一些实施方案中，氨基酸序列与CS23-SC2-AA(SEQ ID NO:12)具有至少99％的同一性。在一些实施方案中，氨基酸序列与CS23-SC2-AA(SEQ ID NO:12)具有至少99.5％的同一性。在一些实施方案中，氨基酸序列与CS23-SC2-AA(SEQ ID NO:12)具有至少99.9％的同一性。在一些实施方案中，氨基酸序列与CS23-SC2-AA(SEQ ID NO:12)相同。

[0361] 在一些实施方案中，由CS23-SC2多核苷酸编码的与CS23-SC2-AA具有高序列同源性(例如至少95％、96％、97％、98％、99％、99.5％或99.9％的同一性)的单链因子VIII变体包含一个或多个选自m1、m2、m3、m4以及m5的氨基酸取代。

[0362] 在一个实施方案中，由CS23多核苷酸编码的单链因子VIII变体包含m1氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m2氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m3氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m4氨基酸取代。在一个实施方案中，由CS23多核苷

酸编码的因子VIII变体包含m5氨基酸取代。

[0363] 在一个实施方案中，由CS23多核苷酸编码的单链因子VIII变体包含m12氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m13氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m23氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m24氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m25氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m34氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m35氨基酸取代。

[0364] 在一个实施方案中，由CS23多核苷酸编码的单链因子VIII变体包含m123氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m234氨基酸取代。在一个实施方案中，由CS23多核苷酸编码的因子VIII变体包含m125氨基酸取代。

[0365] E.因子VIII表达载体

[0366] 在一些实施方案中，将本文所描述的密码子改变的多核苷酸整合至表达载体中。表达载体的非限制性实例包括病毒载体(例如适合于基因疗法的载体)、质粒载体、噬菌体载体、粘粒、噬菌粒、人工染色体等。

[0367] 病毒载体的非限制性实例包括：逆转录病毒，例如莫洛尼鼠白血病病毒(Moloney murine leukemia virus，MMLV)、哈维鼠肉瘤病毒(Harvey murine sarcoma virus)、鼠乳腺肿瘤病毒以及劳斯肉瘤病毒(Rous sarcoma virus)；腺病毒、腺相关病毒；SV40型病毒；多瘤病毒；爱泼斯坦-巴尔病毒(Epstein-Barr viruse)；乳头瘤病毒；疱疹病毒；牛痘病毒；以及脊髓灰质炎病毒。

[0368] 在一些实施方案中，将本文所描述的密码子改变的多核苷酸整合至基因疗法载体中。在一些实施方案中，基因疗法载体为逆转录病毒，并且特别地为复制缺陷型逆转录病毒。用于制备复制缺陷型逆转录病毒的方案是本领域中已知的。关于综述，参见Kriegler，M.，Gene Transfer and Expression，A Laboratory Manual，W.H.Freeman Co.，New York(1990)以及Murry，E.J.，Methods in Molecular Biology，第7卷，Humana Press，Inc.，Cliffton，N.J.(1991)。

[0369] 在一个实施方案中，基因疗法载体为基于腺相关病毒(AAV)的基因疗法载体。AAV系统先前已有描述，并且总体上是本领域中熟知的(Kelleher和Vos，Biotechniques，17(6)：1110-17(1994)；Cotten等，Proc Natl Acad Sci USA，89(13)：6094-98(1992)；Curiel，Nat Immun，13(2-3)：141-64(1994)；Muzyczka，Curr Top Microbiol Immunol，158：97-129(1992)；以及Asokan A等，Mol.Ther.，20(4)：699-708(2012)，这些文献各自出于所有目的以全文引用的方式并入本文中)。例如美国专利号5，139，941和4，797，368中描述了关于rAAV载体的产生和使用的细节，这些专利各自出于所有目的以全文引用的方式并入本文中。在一个特定实施方案中，AAV载体为AAV-8载体。

[0370] 在一些实施方案中，将本文所描述的密码子改变的多核苷酸整合至逆转录病毒表达载体中。这些系统先前已有描述，并且总体上是本领域中熟知的(Mann等，Cell，33：153-159，1983；Nicolas和Rubinstein，Vectors：A survey of molecular cloning vectors and their uses，Rodriguez和Denhardt编，Stoneham：Butterworth，第494-513页，1988；Temin，Gene Transfer，Kucherlapati(编)，New York：Plenum Press，第149-188页，1986)。

在一个特定实施方案中,逆转录病毒载体为慢病毒载体(参见例如Naldini等,Science,272(5259):263-267,1996;Zufferey等,Nat Biotechnol,15(9):871-875,1997;Blomer等,J Virol.,71(9):6641-6649,1997;美国专利号6,013,516和5,994,136)。

[0371] 为使因子VIII多肽在细胞培养物中由密码子改变的多肽表达可使用多种载体,包括真核和原核表达载体。在某些实施方案中,考虑将质粒载体用于在细胞培养物中表达因子VIII多肽。一般来说,将含有从与宿主细胞相容的物种衍生的复制子和控制序列的质粒载体与这些宿主结合使用。载体可带有复制位点,以及能够在转化的细胞中提供表型选择的标记序列。质粒将包括可操作地连接至一个或多个控制序列(例如启动子)的编码因子VIII多肽的密码子改变的多核苷酸。

[0372] 用于原核表达的载体的非限制性实例包括质粒,诸如pRSET、pET、pBAD等,其中原核表达载体中所用的启动子包括lac、trc、trp、recA、araBAD等。用于真核表达的载体的实例包括:(i)对于在酵母中表达,载体为诸如pAO、pPIC、pYES、pMET,使用诸如AOX1、GAP、GAL1、AUG1等启动子;(ii)对于在昆虫细胞中表达,载体为诸如pMT、pAc5、pIB、pMIB、pBAC等,使用诸如PH、p10、MT、Ac5、OpIE2、gp64、polh等启动子,以及(iii)对于在哺乳动物细胞中表达,载体为诸如pSVL、pCMV、pRc/RSV、pcDNA3、pBPV等,以及来源于诸如牛痘病毒、腺相关病毒、疱疹病毒、逆转录病毒等病毒系统的载体,使用诸如CMV、SV40、EF-1、UbC、RSV、ADV、BPV以及β-肌动蛋白等启动子。

IV.实施例

[0373] 实施例1-密码子改变的因子VIII变体表达序列的构建

[0374] 为了形成对于血友病A的基因治疗有效的因子VIII编码序列,必须克服两个障碍。第一,由于常规基因疗法递送载体(例如AAV病毒体)的基因组尺寸限制,必须使编码的因子VIII多肽明显缩短。第二,必须改变编码序列以:(i)稳定递送载体内的包装相互作用,(ii)稳定mRNA中间形态,并且(iii)改善mRNA转录/翻译的稳健性。

[0375] 为了实现第一个目标,申请人从在本文中称为“FVIII-BDD-SQ”的B-结构域缺失的因子VIII变体构建体入手。在此构建体中,将B-结构域用被称为“SQ”序列的14氨基酸序列替换。重组FVIII-BDD-SQ是以商品名REFACTO®出售,并且已显示对于血友病A的管理是有效的。然而,包括因子VIII重链和轻链的人野生型核酸序列的FVIII-BDD-SQ的天然编码序列在基因疗法载体中无效地表达。

[0376] 为解决天然FVIII-BDD-SQ的不良表达,将如在Ward等(Blood,117:798(2011))和McIntosh等(Blood,121,3335-3344(2013))中所描述而改良的在Fath等(PLoS ONE,6:e17596(2011))中所描述的密码子优化算法应用于FVIII-BDD-SQ序列以形成第一中间编码序列CS04a。然而,申请人认识到可通过对序列进一步修饰来改进使用改良算法所形成的CS04a序列。因此,申请人重新引入了CpG二核苷酸,重新引入了用于精氨酸的CGC密码子,改变了亮氨酸和丝氨酸密码子分布,重新引入了高度保守的密码子对,并且去除了隐蔽的TATA盒、CCAAT盒以及剪接位点元件,同时避免CpG岛以及富含AT和富含GC的序列段(stretch)的局部过表达。

[0377] 首先,改良的算法用非CpG-二核苷酸密码子系统地替换含有CpG-二核苷酸的密码子(例如精氨酸密码子),并且消除/避免由相邻密码子形成的CpG-二核苷酸。通常这样严格

避免CpG二核苷酸以防止在肌肉内注射DNA疫苗后TLR诱导的免疫性。然而，这样做限制了密码子优化的可能性。举例来说，改良的算法排除使用完整的CGX精氨酸密码子集合。这在用于在人细胞中表达的基因的编码中特别具有破坏性，因为CGC是高度表达的人基因中最常使用的精氨酸密码子。此外，避免由相邻密码子形成CpG进一步限制优化可能性(例如限制可一起使用的密码子对的数目)。

[0378] 因为预期TLR诱导的免疫性不是与肝定向的基于AAV的基因疗法相关的问题，所以将包括CpG的密码子和形成CpG的相邻密码子优先在编码因子VIII轻链的序列中(例如在FVIII-BDD-SQ编码序列的3'末端)重新引入中间编码序列CS04a中。这允许更频繁地使用优选的人密码子，特别是那些用于精氨酸的密码子。然而，谨慎避免形成CpG岛，所述CpG岛为编码序列中具有高CpG位点频率的区域。这与Krinner等(Nucleic Acids Res.,42(6):3551-64(2014))的教示不同，所述文献表明转录起始位点下游的CpG结构域促进高水平的基因表达。

[0379] 其次，改良算法排他性地应用于某些密码子，诸如用于亮氨酸的CTG、用于缬氨酸的GTG以及用于谷氨酰胺的CAG。然而，这违反了例如如Haas等(Current Biology,6(3):315-24(1996))中所提出的平衡密码子使用的原则。为了说明改良算法对优选密码子的过度使用，在适用于密码子改变的其他规则(例如CpG频率和GC含量)允许的情况下，重新引入替代亮氨酸密码子。

[0380] 第三，当某些准则(例如存在CG-二核苷酸)得到满足时，改良算法替换密码子对而不考虑它们在自然界中有多保守。为了说明通过进化可能保守的有益特性，对由算法替换的最保守的密码子对和例如如Tats等(BMC Genomics 9:463(2008))中所描述的最保守的优选密码子对进行分析，并且在适用于密码子改变的其他规则(例如CpG频率和GC含量)允许的情况下进行调整。

[0381] 第四，还对中间编码序列中所用的丝氨酸密码子进行重新工程化。具体地说，将AGC、TCC以及TCT丝氨酸密码子以更高的频率引入修饰的编码序列中，以总体上更好地匹配人密码子使用(Haas等,同上)。

[0382] 第五，筛检TATA盒、CCAAT盒元件以及内含子/外显子剪接位点并且从修饰的编码序列中去除。当修饰编码序列时，应谨慎避免富含AT或富含GC的序列段的局部过表达。

[0383] 最后，除了优化编码序列内的密码子使用之外，当进一步精制中间编码序列CS04a时，还考虑基础AAV病毒体的结构要求。将AAV载体(例如AAV病毒体的核酸部分)以单链DNA分子形式包装到它们的衣壳中(关于综述，参见Daya和Berns,Clin.Microbiol Rev.,21(4):583-93(2008))。因此，载体的GC含量可能影响基因组的包装，并且因而影响制备期间的载体产量。像许多算法一样，在此所用的改良算法形成GC含量为至少60％的优化基因序列(参见Fath等,PLoS One,6(3):e17596(2011)(于PLoS One,(6)3(2011)中更正)。然而，AAV8衣壳蛋白是由具有约56％的较低GC含量的核苷酸序列编码。因此，为了更好地模拟天然AAV8衣壳蛋白编码序列，使中间编码序列CS04a的GC含量降至56％。

[0384] 图2中所示的所得CS04编码序列的总GC含量为56％。序列的CpG-二核苷酸含量为适度的。然而，CpG二核苷酸主要存在于编码序列的下游部分，例如编码因子VIII轻链的部分。CS04序列与野生型因子VIII中的相应编码序列具有79.77％的核苷酸序列同一性(Genbank登录号M14113)。

[0385] 出于比较的目的,制备了若干其他密码子优化的ReFacto构建体。如关于CS04所做的那样,通过应用如由Ward等改良的Fath等的密码子优化算法来构建CS01。然而,不同于CS04,CS01构建体不含任何CpG岛。如Radcliff P.M.等,Gene Therapy,15:289-97(2008)中所描述对CS08 ReFacto构建体进行密码子优化,该文献的内容在此出于所有目的以全文引用的方式并入本文中。从Eurofins Genomics(Ebersberg,Germany)获得CS10密码子优化的ReFacto构建体。从Integrated DNA Technologies,Inc.(Coralville,USA)获得CS11密码子优化的ReFacto构建体。从ThermoFischer Scientific的GeneArt服务中心(Regensburg,Germany)获得CH25密码子优化的ReFacto构建体。CS40 ReFacto构建体由野生型因子VIII编码序列组成。用于构建CS23的算法是基于JCAT工具(www.jcat.de),这是用于密码子优化的一种在线工具(Grote等,2005;Nucl.Acids Res.W526-31)。对序列进行进一步修饰以更好地反映白蛋白超家族的密码子使用(Mirsafian等2014:Sc.Word Journal 2014,ID 639682)。下表2中示出了各个ReFacto编码序列之间拥有的序列同一性。

[0386] 表2-密码子改变的因子VIII构建体的同一性百分比矩阵.

[0387]

	CS01	CS04	CS08	CS10	CS11	CS40	CH25	CS23
CS01	100%							
CS04	93.0%	100%						
CS08	80.7%	82.2.%	100%					
CS10	79.1%	79.4%	78.4%	100%				
CS11	78.3%	78.3%	78.1%	77.5%	100%			
CS40	79.6%	79.8%	76.7%	77.6%	75.4%	100%		
CH25	81.3%	85.1%	85.0%	79.9%	79.4%	75.8%	100%	
CS23	84.3%	89.2%	85.1%	80.3%	79.9	76.5%	93.2%	100%

[0388] 通过将不同的合成DNA片段克隆至相同的载体骨架质粒(pCh-BB01)中来构建各构建体的质粒。通过ThermoFischer Scientific(Regensburg,Germany)来进行具有侧接的AscI和NotI酶限制位点的Refacto型BDD-FVIII片段的DNA合成。载体骨架含有涵盖来源于肝特异性鼠转甲状腺素蛋白基因的启动子/增强子序列的两个侧接的AAV2衍生的反向末端重复序列(ITR)、用于插入相应Refacto型BDD-FVIII的AscI和NotI酶限制位点以及合成的polyA位点。在经由AscI和NotI位点连结所制备的载体骨架与插入物之后,以毫克规模对所得质粒进行扩增。通过直接测序(Microsynth,Balgach,Switzerland)来验证构建体的Refacto型BDD-FVIII序列。克隆产生命名为pCS40、pCS01、pCS04、pCS08、pCS10、pCS11以及pCh25的七种不同的质粒构建体(图23)。这些构建体具有相同的载体骨架并且编码相同的B-结构域缺失的FVIII蛋白(Refacto型BDD-FVIII),但其FVIII编码序列有所不同。

[0389] 如Grieger JC等(Virus Vectors Using Suspension HEK293 Cells and Continuous Harvest of Vector From the Culture Media for GMP FIX and FLT1 Clinical Vector,Mol Ther.,10月6日.(2015)doi:10.1038/mt.2015.187.[印刷版之前的电子版])中所描述,通过三种质粒转染方法制备基于AAV8的载体,该文献的内容在此出于所有目的以全文引用的方式明确并入本文中。使用HEK293悬浮液细胞使用相应FVIII载体质粒、辅助质粒pXX6-80(携带腺病毒辅助基因)以及包装质粒pGSK2/8(贡献rep2和cap8基因)进行质粒转染。为了分离AAV8构建体,如Grieger等(2015,同上)中所描述,使用碘克沙

醇(iodixanol)梯度对1升培养物的细胞沉淀进行加工。该程序产生称为vCS01、vCS04、vCS08、vCS10、vCS11以及vCH25的载体制剂。通过qPCR使用靶向AAV2反向末端重复序列的通用qPCR程序对载体进行定量(Aurnhammer,Human Gene Therapy Methods:B部分23:18-28(2012))。携带AAV2反向末端重复序列的对照载体质粒用于制备标准曲线。在图7A-7C中以SEQ ID NO:8形式呈现所得vCS04构建体。

[0390] 通过AAV琼脂糖凝胶电泳分析载体基因组的完整性。如Fagone等,Human Gene Therapy Methods 23:1-7(2012)中所描述进行电泳。简言之,将AAV载体制剂在0.5%SDS存在下在75℃下孵育10分钟,并且然后冷却至室温。在1%1xTAE琼脂糖凝胶上每个泳道加载约1.5E10个载体基因组(vg)并且在每厘米凝胶长度7V下电泳60min。然后将凝胶在2x GelRed(Biotium目录号41003)溶液中染色并通过ChemiDocTMMP(Biorad)成像。图24中所示的结果证实vCS01、vCS04以及vCS40病毒载体具有由5kb范围内的独特条带(图24,泳道2-4)指示的相同尺寸的基因组。尽管载体尺寸为约5.2kb,但基因组为均匀条带,证实尺寸有些过大的基因组(相对于4.7kb的AAV野生型基因组)的恰当包装。所有其他vCS载体制剂显示相同的基因组尺寸(数据未显示)。

[0391] 为了确认预期衣壳蛋白模式,用载体vCS01、vCS04以及vCS40进行SDS PAGE,随后进行银染色(图25)。如图中所示,下游纯化程序产生展示VP1、VP2以及VP3的预期蛋白模式的高度纯化的材料(图25,泳道2-4)。在所有其他病毒制剂(未显示)情况下观察到相同的模式。根据标准程序来完成AAV制剂的SDS-PAGE程序。各泳道含有1E10 vg的相应病毒构建体,并且根据制造商的说明书在4-12%Bis-Tris(NuPAGE®Novex,Life Technologies)凝胶上进行分离。根据制造商的说明书用SilverQuestTM试剂盒(Novex,Life Technologies)进行银染色。

[0392] 令人惊讶地,与vCS40野生型编码构建体和其他密码子优化的构建体相比,AAV载体vCS01和vCS04具有由在AAV病毒制备中的更高产量度量的更高程度的病毒体包装。如表3中所示,vCS01和vCS04载体基本上比vCS40更好地复制,从而在AAV效价中提供5-7倍的产量增加。

[0393] 表3-使用如从细胞沉淀纯化的AAV载体构建体vCS01、vCS04以及vCD40获得的每升细胞培养物的产量.

[0394]

构建体	载体浓度 [vg/ml] x10E12	产量 [vg/升] x10E12	相对于wt的增加倍数
vCS40	2.0	11.0	-
vCS01	9.2	51.4	4.7

[0395]

vCS04–样品1	17.6	79.2	7.2
vCS04–样品2	15.9	58.8	5.4

[0396] 实施例2-密码子改变的因子VIII变体表达序列的体内表达

[0397] 为了测试密码子改变的因子VIII变体序列的生物效能，将实施例1中所描述的ReFacto型FVIII构建体施用给缺乏因子VIII的小鼠。简言之，在C57Bl/6FVIII敲除(ko)小鼠中(每组6-8个动物)通过每千克小鼠体重尾部静脉注射4E12个载体基因组(vg)来进行分析。注射后14天时通过眶后穿刺取血，并且使用标准程序制备血浆并冷冻。选择第14天的表达水平，因为此时抑制性抗体的影响极小，所述抑制性抗体在更晚的时间在这种小鼠模型的一些动物中可被观察到。使用如制造商(Technoclone,Vienna,Austria)建议所进行的Technochrome FVIII分析来测定小鼠血浆中的FVIII活性，仅进行少量修改。对于该分析，将血浆样品适当稀释并与含有凝血酶、活化的因子IX(FIXa)、磷脂、因子X以及钙的分析试剂混合。在FVIII由凝血酶活化之后，形成与FIXa、磷脂以及钙的复合物。此复合物将FX活化为活化的FX(FXa)，活化的FX进而从显色底物上裂解对硝基苯胺(pNA)。在405nm下测量pNA形成的动力学。速率与样品中的FVIII浓度成正比。从参考曲线读取FVIII浓度，并且以IU FVIII/毫升的形式给出结果。

[0398] 下表4中所呈现的结果证实使用商业算法设计的密码子改变的序列(CS10、CS11以及CH25)与野生型BDD-因子VIII构建体(CS40)相比仅提供BDD-因子VIII的适度增加(3-4倍)。类似地，如Radcliffe等(CS08)中所描述制备的密码子改变的BDD-因子VIII构建体仅提供BDD-FVIII表达的3-4倍的增加。此结果与Radcliff等中报道的结果一致。令人惊讶地，CS01、CS04以及CS23构建体在体内生物效能分析中提供高得多的BDD-FVIII表达(分别增加18倍、74倍以及30倍)。

[0399] 表4-在FVIII-敲除小鼠的血浆中由不同的AAV载体构建体诱导的FVIII表达.

[0400]

构建体	密码子算法	第 14 天的平均 FVIII 表达 [IU/ml]	标准偏差	小鼠数目	相对于 wt 的增加倍数
vCS40	人野生型	0.03	0.03	12	-
vCS01	申请人的	0.55	0.28	22	18.3
vCS04	申请人的	2.21	1.20	55	73.7
vCS08	Radcliffe 等	0.11	0.01	6	3.6
vCS10	Eurofins	0.09	0.01	7	3.0
vCS11	IDT	0.08	0.02	8	2.7
vCH25	GeneArt	0.13	0.12	18	4.3
vCS23	申请人的	0.91	0.32	5	30.3

[0401] 实施例3-B-结构域取代的接头的糖基化肽的设计

[0402] 其他人已证实，将含有来自野生型因子VIII B-结构域的六个推定的N-连接的糖基化位点的小肽(“V3肽”)纳入到B-结构域缺失的基因疗法构建体中使小鼠血浆中的因子

VIII水平增加(McIntosh等,Blood 121(17):3335-44(2013))。然而,为了保持B-结构域取代的接头的小尺寸,将糖基化位点从野生型B-结构域的背景中去除。含有V3肽的接头的计算机模拟预测(Gupta等,同上)表明V3肽中的这些糖基化位点中只有两个将在体内被修饰(图15)。

[0403] 因此,申请人尝试鉴定将支持更高水平的体内糖基化的替代糖基化肽,这种更高水平的体内糖基化比V3肽更接近于野生型糖基化。申请人经由计算机模拟设计和测试了若干替代糖基化肽。图13A-13B中所出的这些肽中有若干被预测当被放置于SEQ ID NO:2中的B-结构域取代的接头的氨基酸N768与P769之间时,具有与V3肽相同或更高的体内糖基化。下表5中示出了计算机模拟预测的结果。表5还报道了针对若干构建体进行的表达实验的结果,所述若干构建体在CS01密码子优化的背景中编码具有合并至B-结构域取代的接头中的糖基化肽的ReFacto型因子VIII蛋白。

[0404] 表5-B-结构域取代的接头肽中的N-糖基化的预测以及AAV载体构建体在体内的表现.

[0405]

序列	预测的 N-糖基化位点的数目	第 28 天的表达[IU/ml]	SD	小鼠数目[n]	表达倍数
vCS01	0	0.74	0.52	5	21
vNG1/CS01	4	n.d.	-	-	-
vNG4/CS01	**3**	**1.93**	**0.57**	**6**	**55**
vNG5/CS01	2	n.d.	-	-	-
vNG6/CS01	1	0.80	0.67	5	23
vNG9/CS01	1	n.d.	-	-	-
vNG10/CS01	**2**	**2.66**	**0.52**	**6**	**76**
vNG16/CS01	**2**	**1.59**	**0.57**	**6**	**45**
vNG17/CS01	2	n.d.	-	-	-
vNG18/CS01	2	n.d.	-	-	-
vNG19/CS01	2	0.88	0.25	5	25
vNG20/CS01	2	n.d.	-	-	-
vNG21/CS01	2	n.d.	-	-	-
vCS40	0	0.035	0.030	12	1

[0406] 如实施例1中所描述构建含有NG变体的AAV载体并且如实施例2中所描述在FVIII敲除小鼠中进行测试。表5中所示的所有病毒载体(除了对照载体vCS40)是基于如vCS01中所用的算法。还制备了使用vCS04的算法的一组平行构建体(vNG/CS04系列),并且在小鼠模型中进行测试。将结果与用野生型vCS40构建体实现的表达水平相比较。在本实施例中选择

第28天的表达水平，因为大部分构建体的表达水平在此时达到最高水平。三种AAV载体实现大于40倍的FVIII表达水平，包括vNG4/CS01、vNG10/CS01以及vNG16/CS01(表5)。相应构建体vNG4/CS04、vNG10/CS04以及vNG16/CS04预期显示甚至更高的表达，因为它们是基于优良的vCS04算法。

[0407] 令人惊讶地，与vCS40野生型编码构建体相比，vNG/CS01系列的AAV载体具有由AAV病毒制备中的更高产量度量的更高程度的病毒体包装。如表6中所示，基于vNG/CS01的载体基本上比vCS40更好地复制，从而在AAV效价中提供约3倍的产量增加。

[0408] 表6-使用如从细胞沉淀纯化的AAV载体构建体获得的每升细胞培养物的产量.

[0409]

序列	载体浓度 [vg/ml] x10^{12}	产量[vg/升] x10^{12}	相较于野生型的增加倍数
vCS01	9.17	51.35	4.7
vNG1/CS01	2.13	17.04	1.5
vNG4/CS01	5.74	33.01	3.0
vNG5/CS01	6.91	27.29	2.5
vNG6/CS01	7.01	40.66	3.7
vNG9/CS01	6.39	29.39	2.7
vNG10/CS01	8.57	37.71	3.4
vNG16/CS01	5.3	28.36	2.6
vNG17/CS01	4.24	32.22	2.9
vNG18/CS01	6.11	37.88	3.4
vNG19/CS01	9.42	39.56	3.6
vNG20/CS01	4.09	30.27	2.8
vNG21/CS01	n.d	-	-
vCS40	2.03	11	1.0

[0410] 实施例4-突变体BDD-FVIII构建体的构建

[0411] 对在因子VIII重链和/或B-结构域取代的接头内携带氨基酸突变的许多不同的突变的Refacto型BDD-FVIII构建体进行克隆并筛选。如本文中被称为“vCS”系列载体的相应载体在CS01、CS04以及CS23密码子改变的背景中编码BDD-FVIII变体。实施例1中描述了用于构建CS01和CS04背景的方法。用于构建CS23的方法是基于JCAT工具(www.jcat.de)，这是用于密码子优化的一种在线工具(Grote等,2005;Nucl.Acids Res.W526-31)。对序列进行进一步修饰以更好地反映白蛋白超家族的密码子使用(Mirsafian等,Sc.Word Journal,ID 639682(2014))，该文献的内容在此出于所有目的以全文引用的方式明确并入本文中。

[0412] 在vCS系列构建体的FVIII序列中包括三种类型突变的组合。引入FVIII序列中的第一种氨基酸改变为X1突变(TTYVNRSL(SEQ ID NO:33);X.Xiao)，所述X1突变在B-结构域取代的接头附近引入额外的糖基化位点。X1突变在本文中也被称为“m3”突变。在FVIII序列中进行的第二种氨基酸改变包括F328S(SPI,F309S SPE)突变，这是一种已知改善FVIII分泌的氨基酸改变(Swaaroop,J.Biol.Chem.,272:24121-24(1997))。此突变在本文中也被称为“m1”突变。第三种改变为所谓的X5突变，所述X5突变为在重链的A1结构域中的五种氨基

酸变化的组合,改善BDD-FVIII的比活性和分泌(Cao等,2014;ASGCT摘要编号460;口头报告中公开了突变细节)。X5突变在本文中也被称为“m2”突变。接下来,实现X1和F328S(SPI,F309S SPE)的组合,随后为X1和X5的组合(也称为“X6”),并且实现X5和F328S(SPI,F309S SPE)的其他组合(表7)。

[0413] 载体质粒的基因合成和克隆.通过将不同的合成DNA片段克隆至相同的载体骨架质粒(pCh-BB01)中来构建质粒。通过ThermoFischer Scientific(Regensburg,Germany)来进行具有侧接的AscI和NotI酶限制位点的Refacto型BDD-FVIII片段的DNA合成。载体骨架含有涵盖来源于肝特异性鼠转甲状腺素蛋白基因的启动子/增强子序列的两个侧接的AAV2衍生的反向末端重复序列(ITR)、用于插入相应Refacto型BDD-FVIII的AscI和NotI酶限制位点以及合成的polyA位点。在经由AscI和NotI位点连结所制备的载体骨架与插入物之后,以毫克规模对所得质粒进行扩增。通过直接测序(Microsynth,Balgach,Switzerland)来验证构建体的Refacto型BDD-FVIII序列。如图44中所示,克隆产生不同的质粒构建体。

[0414] 小规模载体制备以及通过定量PCR(qPCR)进行定量.大体上如Grieger等(2015,Supra)中所描述通过三种质粒转染方法制备基于AAV8的载体。使用HEK293悬浮液细胞使用相应FVIII载体质粒、辅助质粒pXX6X80(携带腺病毒辅助基因)以及包装质粒pGSK2/8(贡献rep2和cap8基因)进行质粒转染。在下游工艺中,如上文所描述使用碘克沙醇梯度对一升培养物的细胞沉淀进行加工。该程序产生如表8中所概述的载体制剂。通过qPCR使用靶向AAV2反向末端重复序列的通用qPCR程序对载体进行定量(Aurnhammer,HUMAN GENE THERAPY METHODS:B部分23:18-28(2012))。精确定量的携带AAV2反向末端重复序列的载体质粒用于制备标准曲线。

[0415] AAV载体表征.通过AAV琼脂糖凝胶电泳分析载体基因组的完整性。与Fagone等(Human Gene Therapy Methods,23:1-7(2012))中所描述类似地进行电泳。将AAV载体制剂在0.5%SDS存在下在75℃下孵育10分钟,并且然后冷却至室温。在1%1xTAE琼脂糖凝胶上每个泳道加载约1.5E10个载体基因组(vg)并且在每厘米凝胶长度7V下电泳60min。然后将凝胶在2x GelRed(Biotium目录号41003)溶液中染色并通过ChemiDoc™MP(Biorad)成像。图45中示出了载体选择的结果。病毒载体vCS04(对照)、vCS17、vCS20、vCS24、vCS16以及vCS40(对照)全部以在5kb范围内的独特条带(图45,泳道2-7;右侧的箭头)的形式显示相同尺寸的基因组。尽管载体尺寸为约5.2kb,但基因组为均匀条带,证实尺寸有些过大的基因组(相对于4.7kb的AAV野生型基因组)的恰当包装。

[0416] 为了确认载体的纯度和预期衣壳蛋白模式,如图46中所示,用载体进行SDS PAGE,随后进行银染色。如图中所示,下游纯化程序产生展示VP1、VP2以及VP3的预期蛋白模式的高度纯化的材料(图46泳道2-9;右手侧的箭头)。根据标准程序来完成AAV制剂的SDS-PAGE程序。将每一泳道1E10 vg的量根据制造商的说明书在4-12% Bis-Tris(NuPAGE® Novex,Life Technologies)凝胶上进行分离。根据制造商的说明书用SilverQuest™试剂盒(Novex,Life Technologies)进行银染色。

[0417] 载体的体内生物效能筛选.在小鼠中对不同的Refacto型BDD-FVIII构建体进行筛选。在C57Bl/6FVIII敲除(ko)小鼠中(每组6-8个动物)通过每千克小鼠体重尾部静脉注射4E12个载体基因组(vg)来进行分析。注射后14天时通过眶后穿刺取血,并且使用标准程序制备血浆并冷冻。使用来自Technoclone的显色分析(Technochrome FVIII,Technoclone,

Vienna,Austria)进行少量修改来测定小鼠血浆中的FVIII活性。简言之,将血浆样品适当稀释并与含有凝血酶、活化的因子IX(FIXa)、磷脂、因子X以及钙的分析试剂混合。在FVIII由凝血酶活化之后,形成与FIXa、磷脂以及钙的复合物。此复合物将FX活化为活化的FX(FXa),活化的FX进而从显色底物上裂解对硝基苯胺(pNA)。在405nm下测量pNA形成的动力学。速率与样品中的FVIII浓度成正比。从参考曲线读取FVIII浓度,并且以IU FVIII/毫升的形式给出结果。

[0418] 小鼠生物效能分析的结果(在小鼠血浆中的第14天的FVIII表达数据(国际单位/毫升[IU/ml])和与野生型vCS40对照相比的表达倍数)在表7中示出。AAV载体vCS19、vCS26以及vCS32全部分别在CS01、CS04以及CS23密码子改变的背景中含有X1糖基化位点。如表7中可见,与野生型构建体vCS40(水平规定为1)相比,获得令人惊讶的高表达水平。举例来说,与野生型vCS40载体相比,vCS26表达高202倍的水平。X1系列载体的在Geneart密码子背景中含有X1突变的另一对照构建体vCH111显示更适度的表达增加(12倍)。

[0419] 载体vCS16、vCS28以及vCS34全部分别在CS01、CS04以及CS23密码子改变的背景中含有增强分泌的F328S(SPI,F309S SPE)突变。如表7中可见,在vCS16和vCS28情况下获得高表达水平(比wt vCS40对照高45-93倍)。

[0420] 载体vCS20、vCS24以及vCS33分别在CS01、CS04以及CS23密码子改变的背景中含有X5突变。X5系列中表现最佳的变体为vCS20,在第14天之后实现>3个单位/ml的水平并且相较于wt vCS40对照增加121倍。

[0421] 载体vCS17、vCS29以及vCS31分别在CS01、CS04以及CS23密码子改变的背景中含有X1和F328S(SPI,F309S SPE)突变的组合(表6)。vCS17和vCS29构建体在小鼠研究中实现非常高的表达水平(相较于vCS40对照增加115至246倍)。值得注意的是,在所用的FVIII KO小鼠模型中,用vCS17构建体处理的小鼠中大多数随时间推移不产生中和抗体,这由在更晚的时间点(例如第28天和第42天;数据未显示)不断增加的FVIII水平得到证实。这是一个意外的发现,因为在一些其他构建体中,由于形成中和抗体,表达水平随时间推移开始降低。CS01背景与增强分泌的突变F328S(SPI,F309S SPE)和X1组合产生低免疫原性诱导。

[0422] 载体vCS18、vCS27以及vCS35分别在CS01、CS04以及CS23密码子改变的背景中含有X1和X5突变的组合。这两种突变的组合也非常高效。例如在vCS18情况下可实现相较于vCS40对照增加145倍(表7)。

[0423] 载体vCS48和vCS49分别在CS01和CS04密码子改变的背景中含有X5和F328S(SPI,F309S SPE)突变的组合。这两种突变的组合也非常高效。所有突变体中增加极多的中的一种是在vCS49情况下可实现相较于vCS40对照增加239倍,这证实了包括F328S(SPI,F309S SPE)突变的组合的特殊价值。

[0424] 另一个令人惊讶的观察结果是突变体AAV载体基本上比具有野生型BDD-FVIII密码子的vCS40构建体生长得更好。序列优化使得载体制备时的产量增加若干倍。在一些表达极佳的构建体(例如vCS29、vCS17、vCS20以及vCS26)中,归因于密码子改变和/或突变序列的产量增加为与野生型载体相比高约3-5倍(表8)。

[0425] 表7中示出了在FVIII-敲除小鼠的血浆中由不同的AAV载体构建体诱导的BDD-FVIII表达。构建体具有相同的载体骨架,然而,编码不同类型的突变的FVIII,包括不同的密码子优化背景。选择第14天的表达水平,因为在此时间点抑制性抗体的影响极小,所述抑

制性抗体在更晚的时间在小鼠模型中在一些动物中可被观察到。n.d.,未检出。

[0426] 表7-vCS构建体的体内生物效能数据.

[0427]

编号	载体	算法，突变	第 14 天的表达 [IU/ml]	SD	小鼠数目 [n]	表达倍数
1	vCS19	CS01，X1	2.34	1.10	13	78
2	vCS26	CS04，X1	6.07	2.72	12	202
3	vCS32	CS23，X1	n.d.	-	-	-
4	vCS16	CS01，F328S	1.35	0.88	6	45
5	vCS28	CS04，F328S	2.78	0.92	7	93
6	vCS34	CS23，F328S	n.d.	-	-	-
7	vCS20	CS01，X5	3.62	1.96	21	121
8	vCS24	CS04，X5	0.79	0.89	18	26
9	vCS33	CS23，X5	n.d.	-	-	n.d.
10	vCS17	CS01，X1，F328S	3.44	1.92	20	115
11	vCS29	CS04，X1，F328S	7.39	2.64	9	246

[0428]

12	vCS31	CS23，X1，F328S	n.d.			n.d.
13	vCS18	CS01，X1+X5 (X6)	4.34	2.50	6	145
14	vCS27	CS04，X1+X5 (X6)	8.03	3.97-	6-	268-
15	vCS35	CS23，X1+X5 (X6)	n.d.	-	-	-
19	vCS48	CS01，X5，F328S	2.54	0.72	8	85
20	vCS49	CS04，X5，F328S	7.17	1.30	7	239
	对照					
16	vCS40	人野生型	0.03	0.03	12	1
17	vCh25	Geneart	0.13	0.12	18	4
18	vCh111	Geneart +X1	0.37	0.21	17	12

[0429] 表8-使用不同的AAV载体构建体获得的每升细胞培养物的产量(包装效率).将载体从细胞沉淀中纯化出来;n.d.,未检出。

[0430]

	构建体	**算法，突变**	**载体浓度 [vg/ml] x10^{12}**	**产量 [vg/升] x10^{12}**	**相对于wt的增加倍数**
1	vCS19	CS01，X1	9.71	36	3.22
2	**vCS26**	CS04，X1	5.93	32	2.87
3	vCS32	CS23，X1	n.d.	n.d.	n.d.
4	vCS16	CS01，F328S	6.51	29	2.56
5	vCS28	CS04，F328S	5.85	32	2.88
6	vCS34	CS23，F328S	n.d.	n.d.	n.d.
7	**vCS20**	CS01，X5	9.90	50	4.48
8	vCS24	CS04，X5	3.00	16	1.46
9	vCS33	CS23，X5	n.d.	n.d.	n.d.
10	**vCS17**	CS01，X1，F328S	8.94	37	3.34

[0431]

11	**vCS29**	CS04，X1，F328S	7.42	53	4.72
12	vCS31	CS23，X1，F328S	n.d.	n.d.	n.d.
13	vCS18	CS01，X1+X5(X6)	21.20	53	4.75
14	vCS27	CS04，X1+X5(X6)	4.15	19	1.67
15	vCS35	CS23，X1+X5(X6)	n.d.	n.d.	n.d.
16	vCS48	CS01，X5，F328S	7.14	42.1	3.77
17	vCS49	CS04，X5，F328S	8.27	37.2	3.33
18	vCS40	人野生型	2.03	11	1.00

[0432] 应当理解，本文所描述的实施例和实施方案仅是用于说明的目的，并且据其进行的各种修改或变化将对本领域技术人员作出提示，并且包括在本申请的精神和权限以及所附权利要求书的范围内。本文所引用的所有出版物、专利以及专利申请在此出于所有目的以全文引用的方式并入本文中。

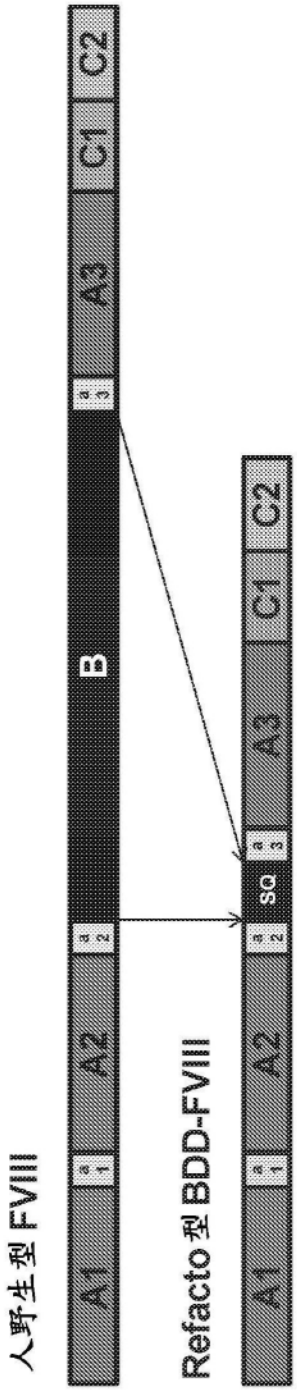

人野生型 FVIII
A1
a1
A2
a2
B
a3
A3
C1
C2
Refacto型 BDD-FVIII
A1
a1
A2
a2
sq
a3
A3
C1
C2

图1

CS04-FL-NA

atgcagattgagctgagcacctgcttcttcctgtgcctgctgaggttctgcttctctgccaccagga
gatactacctgggggctgtggagctttcttgggactacatgcagtctgacctgggggagctgcctgt
ggatgccaggttcccacccagagtgcccaaatccttcccattcaacacctctgtggtctacaagaag
accctctttgtggagttcactgaccacctgttcaacattgccaaacccaggccaccctggatgggac
tcctgggacccaccattcaggctgaggtgtatgacactgtggtcatcaccctcaagaacatggcctc
ccaccctgtgagcctgcatgctgtgggggtcagctactggaaggcctctgagggggctgagtatgat
gaccagacctcccagagggagaaggaggatgacaaagtgttccctgggggcagccacacctatgtgt
ggcaggtcctcaaggagaatggccccatggcctctgacccactctgcctgacctactcctacctttc
tcatgtggacctggtcaaggacctcaactctggactgattggggccctgctggtgtgcagggagggc
tccctggccaaagagaagacccagaccctgcacaagttcattctcctgtttgctgtctttgatgagg
gcaagagctggcactctgaaaccaagaactccctgatgcaggacagggatgctgcctctgccagggc
ctggcccaagatgcacactgtgaatggctatgtgaacaggagcctgcctggactcattggctgccac
aggaaatctgtctactggcatgtgattggcatggggacaacccctgaggtgcactccatttcctgg
agggccacaccttcctggtcaggaaccacagacaggccagcctggagatcagccccatcaccttcct
cactgcccagaccctgctgatggacctcggacagttcctgctgttctgccacatcagctcccaccag
catgatggcatggaggcctatgtcaaggtggacagctgccctgaggagccacagctcaggatgaaga
acaatgaggaggctgaggactatgatgatgacctgactgactctgagatggatgtggtccgctttga
tgatgacaacagcccatccttcattcagatcaggtctgtggccaagaaacaccccaagacctgggtg
cactacattgctgctgaggaggaggactgggactatgccccactggtcctggcccctgatgacagga
gctacaagagccagtacctcaacaatggcccacagaggattggacgcaagtacaagaaagtcaggtt
catggcctacactgatgaaaccttcaagaccagggaggccattcagcatgagtctggcatcctgggc
ccactcctgtatggggaggtgggggacaccctgctcatcatcttcaagaaccaggcctccaggccct
acaacatctacccacatggcatcactgatgtcaggcccctgtacagccgcaggctgccaaagggggt
gaaacacctcaaggacttccccattctgcctggggagatcttcaagtacaagtggactgtcactgtg
gaggatggaccaaccaaatctgaccccaggtgcctcaccagatactactccagctttgtgaacatgg
agagggacctggcctctggcctgattggcccactgctcatctgctacaaggagtctgtggaccagag
gggaaaccagatcatgtctgacaagaggaatgtgattctgttctctgtctttgatgagaacaggagc
tggtacctgactgagaacattcagcgcttcctgcccaaccctgctggggtgcagctggaggaccctg
agttccaggccagcaacatcatgcactccatcaatggctatgtgtttgacagcctccagctttctgt
ctgcctgcatgaggtggcctactggtacattctttctattggggcccagactgacttcctttctgtc
ttcttctctggctacaccttcaaacacaagatggtgtatgaggacaccctgaccctcttcccattct
ctggggagactgtgttcatgagcatggagaaccctggcctgtggattctgggatgccacaactctga
cttccgcaacaggggcatgactgccctgctcaaagtctcctcctgtgacaagaacactggggactac
tatgaggacagctatgaggacatctctgcctacctgctcagcaagaacaatgccattgagcccagga
gcttcagccagaatccacctgtcctgaaacgccaccagagggagatcaccaggaccaccctccagtc
tgaccaggaggagattgactatgatgacaccatttctgtggagatgaagaaagaggactttgacatc
tatgacgaggacgagaaccagagcccaaggagcttccagaagaagaccaggcactacttcattgctg
ctgtggagcgcctgtgggactatggcatgagctccagcccccatgtcctcaggaacagggcccagtc
tggctctgtgccacagttcaagaaagtggtcttccaagagttcactgatggcagcttcacccagccc
ctgtacagaggggagctgaatgagcacctgggactcctgggcccatacatcagggctgaggtggagg
acaacatcatggtgaccttccgcaaccaggcctccaggccctacagcttctacagctccctcatcag
ctatgagGaggaccagaggcaggggctgagccacgcaagaactttgtgaaacccaatgaaaccaag
acctacttctggaaagtccagcaccacatggccccccaccaaggatgagtttgactgcaaggcctggg

（续）

图2A

cctacttctctgatgtggacctggagaaggatgtgcactctggcctgattggcccactcctggtctg
ccacaccaacaccctgaaccctgcccatggaaggcaagtgactgtgcaggagtttgccctcttcttc
accatctttgatgaaaccaagagctggtacttcactgagaacatggagcgcaactgcagggccccat
gcaacattcagatggaggaccccaccttcaaagagaactaccgcttccatgccatcaatggctacat
catggacaccctgcctgggcttgtcatggcccaggaccagaggatcaggtggtacctgctttctatg
ggctccaatgagaacattcactccatccacttctctgggcatgtcttcactgtgcgcaagaaggagg
agtacaagatggccctgtacaacctctaccctggggtctttgagactgtggagatgctgccctccaa
agctggcatctggagggtggagtgcctcattggggagcacctgcatgctggcatgagcaccctgttc
ctggtctacagcaacaagtgccagacccccctgggaatggcctctggccacatcagggacttccaga
tcactgcctctggccagtatggccagtgggcccccaagctggccaggctccactactctggatccat
caatgcctggagcaccaaggagccattcagctggatcaaagtggacctgctggcccccatgatcatc
catggcatcaagacccaggggccaggcagaagttctccagcctgtacatcagccagttcatcatca
tgtacagcctggatggcaagaaatggcagacctacagaggcaactccactggaacactcatggtctt
ctttggcaatgtggacagctctggcatcaagcacaacatcttcaacccccccaatcatcgccagatac
atcaggctgcaccccacccactacagcatccgcagcaccctcaggatggagctgatgggctgtgacc
tgaactcctgcagcatgcccctgggcatggagagcaaggccatttctgatgcccagatcactgcctc
cagctacttcaccaacatgtttgccacctggagcccaagcaaggccaggctgcacctccagggaagg
agcaatgcctggaggccccaggtcaacaacccaaaggagtggctgcaggtggacttccagaagacca
tgaaggtcactggggtgaccacccaggggtcaagagcctgctcaccagcatgtatgtgaaggagtt
cctgatcagctccagccaggatggccaccagtggaccctcttcttccagaatggcaaggtcaaggtg
ttccagggcaaccaggacagcttcacCcctgtggtgaacagcctggacccccccctcctgaccagat
acctgaggattcacccccagagctgggtccaccagattgccctgaggatggaggtcctgggatgtga
ggcccaggacctgtactga (SEQ ID NO:1)

图2B

CS04-FL-AA

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDAR'FPPRVPKSFPFNTSVVYK
KTLFVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEY
DDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCRE
GSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGC
HRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSH
QHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKTW
VHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGIL
GPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVT
VEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENR
SWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLS
VFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGD
YYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFD
IYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQ
PLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNET
KTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALF
FTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLS
MGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTL
FLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMI
IHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIAR
YIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQG
RSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVK
VFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY (SEQ ID NO:2)

图3

CS04-HC-NA

gcc
accaggagat actacctggg ggctgtggag ctttcttggg actacatgca gtctgacctg
ggggagctgc ctgtggatgc caggttccca cccagagtgc ccaaatcctt cccattcaac
acctctgtgg tctacaagaa gaccctcttt gtggagttca ctgaccacct gttcaacatt
gccaaaccca ggccaccctg gatgggactc ctgggaccca ccattcaggc tgaggtgtat
gacactgtgg tcatcaccct caagaacatg gcctcccacc ctgtgagcct gcatgctgtg
ggggtcagct actggaaggc ctctgagggg gctgagtatg atgaccagac ctcccagagg
gagaaggagg atgacaaagt gttccctggg ggcagccaca cctatgtgtg gcaggtcctc
aaggagaatg gccccatggc ctctgaccca ctctgcctga cctactccta cctttctcat
gtggacctgg tcaaggacct caactctgga ctgattgggg ccctgctggt gtgcagggag
ggctccctgg ccaaagagaa gacccagacc ctgcacaagt tcattctcct gtttgctgtc
tttgatgagg gcaagagctg gcactctgaa accaagaact ccctgatgca ggacagggat
gctgcctctg ccagggcctg gcccaagatg cacactgtga atggctatgt gaacaggagc
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtgat tggcatgggg
acaacccctg aggtgcactc cattttcctg gagggccaca ccttcctggt caggaaccac
agacaggcca gcctggagat cagccccatc accttcctca ctgcccagac cctgctgatg
gacctcggac agttcctgct gttctgccac atcagctccc accagcatga tggcatggag
gcctatgtca aggtggacag ctgccctgag gagccacagc tcaggatgaa gaacaatgag
gaggctgagg actatgatga tgacctgact gactctgaga tggatgtggt ccgctttgat
gatgacaaca gcccatcctt cattcagatc aggtctgtgg ccaagaaaca ccccaagacc
tgggtgcact acattgctgc tgaggaggag gactgggact atgccccact ggtcctggcc
cctgatgaca ggagctacaa gagccagtac ctcaacaatg gcccacagag gattggacgc
aagtacaaga aagtcaggtt catggcctac actgatgaaa ccttcaagac cagggaggcc
attcagcatg agtctggcat cctgggccca ctcctgtatg gggaggtggg ggacaccctg
ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctacccaca tggcatcact
gatgtcaggc ccctgtacag ccgcaggctg ccaaaggggg tgaaacacct caaggacttc
cccattctgc ctggggagat cttcaagtac aagtggactg tcactgtgga ggatggacca
accaaatctg accccaggtg cctcaccaga tactactcca gctttgtgaa catggagagg
gacctggcct ctggcctgat tggcccactg ctcatctgct acaaggagtc tgtggaccag
aggggaaacc agatcatgtc tgacaagagg aatgtgattc tgttctctgt ctttgatgag
aacaggagct ggtacctgac tgagaacatt cagcgcttcc tgcccaaccc tgctggggtg
cagctggagg accctgagtt ccaggccagc aacatcatgc actccatcaa tggctatgtg
tttgacagcc tccagctttc tgtctgcctg catgaggtgg cctactggta cattctttct
attggggccc agactgactt cctttctgtc ttcttctctg gctacacctt caaacacaag
atggtgtatg aggacaccct gaccctcttc ccattctctg gggagactgt gttcatgagc
atggagaacc ctggcctgtg gattctggga tgccacaact ctgacttccg caacaggggc
atgactgccc tgctcaaagt ctcctcctgt gacaagaaca ctggggacta ctatgaggac
agctatgagg acatctctgc ctacctgctc agcaagaaca atgccattga gcccagg
(SEQ ID NO:3)

图4

CS04-LC-NA

g agatcaccag gaccaccctc
cagtctgacc aggaggagat tgactatgat gacaccattt ctgtggagat gaagaaagag
gactttgaca tctatgacga ggacgagaac cagagcccaa ggagcttcca gaagaagacc
aggcactact tcattgctgc tgtggagcgc ctgtgggact atggcatgag ctccagcccc
catgtcctca ggaacagggc ccagtctggc tctgtgccac agttcaagaa agtggtcttc
caagagttca ctgatggcag cttcacccag cccctgtaca gaggggagct gaatgagcac
ctgggactcc tgggcccata catcagggct gaggtggagg acaacatcat ggtgaccttc
cgcaaccagg cctccaggcc ctacagcttc tacagctccc tcatcagcta tgaggaggac
cagaggcagg gggctgagcc acgcaagaac tttgtgaaac ccaatgaaac caagacctac
ttctggaaag tccagcacca catggccccc acccaaggatg agtttgactg caaggcctgg
gcctacttct ctgatgtgga cctggagaag gatgtgcact ctggcctgat tggcccactc
ctggtctgcc acaccaacac cctgaaccct gcccatggaa ggcaagtgac tgtgcaggag
tttgccctct tcttcaccat ctttgatgaa accaagagct ggtacttcac tgagaacatg
gagcgcaact gcagggcccc atgcaacatt cagatggagg accccacctt caaagagaac
taccgcttcc atgccatcaa tggctacatc atggacaccc tgcctgggct tgtcatggcc
caggaccaga ggatcaggtg gtacctgctt tctatgggct ccaatgagaa cattcactcc
atccacttct ctgggcatgt cttcactgtg cgcaagaagg aggagtacaa gatggccctg
tacaacctct accctggggt ctttgagact gtggagatgc tgccctccaa agctggcatc
tggagggtgg agtgcctcat tggggagcac ctgcatgctg gcatgagcac cctgttcctg
gtctacagca acaagtgcca gacccccctg ggaatggcct ctggccacat cagggacttc
cagatcactg cctctggcca gtatggccag tgggccccca agctggccag gctccactac
tctggatcca tcaatgcctg gagcaccaag gagccattca gctggatcaa agtggacctg
ctggccccca tgatcatcca tggcatcaag acccaggggg ccaggcagaa gttctccagc
ctgtacatca gccagttcat catcatgtac agcctggatg gcaagaaatg gcagacctac
agaggcaact ccactggaac actcatggtc ttctttggca atgtggacag ctctggcatc
aagcacaaca tcttcaaccc cccaatcatc gccagataca tcaggctgca ccccacccac
tacagcatcc gcagcaccct caggatggag ctgatgggct gtgacctgaa ctcctgcagc
atgcccctgg gcatggagag caaggccatt tctgatgccc agatcactgc ctccagctac
ttcaccaaca tgtttgccac ctggagccca agcaaggcca ggctgcacct ccagggaagg
agcaatgcct ggaggcccca ggtcaacaac ccaaaggagt ggctgcaggt ggacttccag
aagaccatga aggtcactgg ggtgaccacc caggggtca agagcctgct caccagcatg
tatgtgaagg agttcctgat cagctccagc caggatggcc accagtggac cctcttcttc
cagaatggca aggtcaaggt gttccagggc aaccaggaca gcttcacccc tgtggtgaac
agcctggacc ccccctcct gaccagatac ctgaggattc acccccagag ctgggtccac
cagattgccc tgaggatgga ggtcctggga tgtgaggccc aggacctgta c
(SEQ ID NO:4)

图5

```
BDLO01  - agc ttctctcaga atccacctgt cctgaagaga caccagaga (SEQ ID NO:5)
BDLO04  - agc ttcagccaga atccacctgt cctgaaacgc caccagagg (SEQ ID NO:6)
BDLO23  - agc ttcagccaga acccccccgt gctgaagagg caccagagg (SEQ ID NO:7)
BDLNG1  - agcttcagccagaatGTGAGCAACAATGTGAGCAACAATGCCACCAATAATGCTACCAACcacctgtcctgaaacgccaccagagg  (SEQ ID NO:36)
BDLNG4  - agcttcagccagaatGTGAGCAACAATGCCACCAACAATGTGAGCAACcacctgtcctgaaacgccaccagagg  (SEQ ID NO:37)
BDLNG5  - agcttcagccagaatGTGAGCAATAATGCCACCAACcacctgtcctgaaacgccaccagagg  (SEQ ID NO:38)
BDLNG6  - agcttcagccagaatGTGAGCAATAATccacctgtcctgaaacgccaccagagg  (SEQ ID NO:39)
BDLNG9  - agcttcagccagaatAGGAGCCTGccacctgtcctgaaacgccaccagagg  (SEQ ID NO:40)
BDLNG10 - agcttcagccagaatGCCACTAATGTGTCTAACAACTCTGCTACCTCTGCTGACTCTGCTGTGAGCcacctgtcctgaaacgccaccagagg
(SEQ ID NO:41)
BDLNG16 - agcttcagccagaatGCCACCAACTATGTGTCTAACAACTCTGCTACCTCTGCTGACTCTGCTGTGAGCcacctgtcctgaaacgccaccagagg
(SEQ ID NO:42)
BDLNG17 - agcttcagccagaatGCCACCAACTATGTGAACAGGAGCCTGTCTGCCACCTCTGCTGACTCTGCTGTGAGCCAGAATccacctgtcctgaaacgccaccag
agg(SEQ ID NO:43)
BDLNG18 - agcttcagccagaatGTGAGCAACAATGTGAGCAATGCTGTGTCTGCTGTGTCTGCTccacctgtcctgaaacgccaccagagg  (SEQ ID NO:44)
BDLNG19 - agcttcagccagaatATCACTGTGGCCTCTGCCACCTCTAACATCACTGTGGCCTCTGCTGACccacctgtcctgaaacgccaccagagg
(SEQ ID NO:45)
BDLNG20 - agcttcagccagaatATCACTGTGACCAACATCACTGTGACTGCCccacctgtcctgaaacgccaccagagg  (SEQ ID NO:46)
BDLNG21 - agcttcagccagaatCAGACTGTGACCAACATCACTGTGACTGCCccacctgtcctgaaacgccaccagagg  (SEQ ID NO:47)
BDLNGV  - agcttcagccagaatGCCACTAATGTGTCTAACAACAGCAACACCAGCAATGACAGCAATGTGTCTccacctgtcctgaaacgccaccagagg
(SEQ ID NO:48)
```

图6

CS04-AV-NA

```
1    tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca
61   cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg
121  ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc
181  accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc
241  attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat
301  tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt
361  tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cctcgagatt taaatgacgt
421  tggccactcc ctctctgcgc gctcgctcgc tcactgaggc cgggcgacca aaggtcgccc
481  gacgcccggg ctttgcccgg gcggcctcag tgagcgagcg agcgcgcaga gagggagtgg
541  ccaactccat cactaggggt tcctgagttt aaacttcgtc gacgattcga gcttgggctg
601  caggtcgagg gcactgggag gatgttgagt aagatggaaa actactgatg acccttgcag
661  agacagagta ttaggacatg tttgaacagg ggccgggcga tcagcaggta gctctagagg
721  atccccgtct gtctgcacat ttcgtagagc gagtgttccg atactctaat ctccctaggc
781  aaggttcata tttgtgtagg ttacttattc tccttttgtt gactaagtca ataatcagaa
841  tcagcaggtt tggagtcagc ttggcaggga tcagcagcct gggttggaag gaggggtat
901  aaaagccccc tcaccaggag aagccgtcac acagactagg cgcgccaccg ccaccatgca
961  gattgagctg agcacctgct tcttcctgtg cctgctgagg ttctgcttct ctgccaccag
1021 gagatactac ctgggggctg tggagctttc ttgggactac atgcagtctg acctgggggа
1081 gctgcctgtg gatgccaggt tcccacccag agtgcccaaa tccttcccat tcaacacctc
1141 tgtggtctac aagaagaccc tctttgtgga gttcactgac cacctgttca acattgccaa
1201 acccaggcca ccctggatgg gactcctggg acccaccatt caggctgagg tgtatgacac
1261 tgtggtcatc accctcaaga acatggcctc ccaccctgtg agcctgcatg ctgtgggggt
1321 cagctactgg aaggcctctg agggggctga gtatgatgac cagacctccc agagggagaa
1381 ggaggatgac aaagtgttcc ctgggggcag ccacacctat gtgtggcagg tcctcaagga
1441 gaatggcccc atggcctctg acccactctg cctgacctac tcctaccttt ctcatgtgga
1501 cctggtcaag gacctcaact ctggactgat tggggccctg ctggtgtgca gggagggctc
1561 cctggccaaa gagaagaccc agaccctgca caagttcatt ctcctgtttg ctgtctttga
1621 tgagggcaag agctggcact ctgaaaccaa gaactccctg atgcaggaca gggatgctgc
1681 ctctgccagg gcctggccca agatgcacac tgtgaatggc tatgtgaaca ggagcctgcc
1741 tggactcatt ggctgccaca ggaaatctgt ctactggcat gtgattggca tggggacaac
1801 ccctgaggtg cactccattt tcctggaggg ccacaccttc ctggtcagga accacagaca
1861 ggccagcctg gagatcagcc ccatcacctt cctcactgcc cagaccctgc tgatggacct
1921 cggacagttc ctgctgttct gccacatcag ctcccaccag catgatggca tggaggccta
1981 tgtcaaggtg gacagctgcc ctgaggagcc acagctcagg atgaagaaca atgaggaggc
2041 tgaggactat gatgatgacc tgactgactc tgagatggat gtggtccgct ttgatgatga
2101 caacagccca tccttcattc agatcaggtc tgtggccaag aaacacccca agacctgggt
2161 gcactacatt gctgctgagg aggaggactg ggactatgcc ccactggtcc tggcccctga
2221 tgacaggagc tacaagagcc agtacctcaa caatggccca cagaggattg gacgcaagta
2281 caagaaagtc aggttcatgg cctacactga tgaaaccttc aagaccaggg aggccattca
2341 gcatgagtct ggcatcctgg gcccactcct gtatggggag gtgggggaca ccctgctcat
2401 catcttcaag aaccaggcct ccaggcccta caacatctac ccacatggca tcactgatgt
2461 caggcccctg tacagccgca ggctgccaaa gggggtgaaa cacctcaagg acttccccat
```

图7A

```
2521 tctgcctggg gagatcttca agtacaagtg gactgtcact gtggaggatg gaccaaccaa
2581 atctgacccc aggtgcctca ccagatacta ctccagcttt gtgaacatgg agagggacct
2641 ggcctctggc ctgattggcc cactgctcat ctgctacaag gagtctgtgg accagagggg
2701 aaaccagatc atgtctgaca agaggaatgt gattctgttc tctgtctttg atgagaacag
2761 gagctggtac ctgactgaga acattcagcg cttcctgccc aaccctgctg gggtgcagct
2821 ggaggaccct gagttccagg ccagcaacat catgcactcc atcaatggct atgtgtttga
2881 cagcctccag ctttctgtct gcctgcatga ggtggcctac tggtacattc tttctattgg
2941 ggcccagact gacttccttt ctgtcttctt ctctggctac accttcaaac acaagatggt
3001 gtatgaggac accctgaccc tcttcccatt ctctggggag actgtgttca tgagcatgga
3061 gaaccctggc ctgtggattc tgggatgcca caactctgac ttccgcaaca ggggcatgac
3121 tgccctgctc aaagtctcct cctgtgacaa gaacactggg gactactatg aggacagcta
3181 tgaggacatc tctgcctacc tgctcagcaa gaacaatgcc attgagccca ggagcttcag
3241 ccagaatcca cctgtcctga aacgccacca gagggagatc accaggacca ccctccagtc
3301 tgaccaggag gagattgact atgatgacac catttctgtg gagatgaaga aagaggactt
3361 tgacatctat gacgaggacg agaaccagag cccaaggagc ttccagaaga agaccaggca
3421 ctacttcatt gctgctgtgg agcgcctgtg ggactatggc atgagctcca gccccatgt
3481 cctcaggaac agggcccagt ctggctctgt gccacagttc aagaaagtgg tcttccaaga
3541 gttcactgat ggcagcttca cccagccct gtacagaggg gagctgaatg agcacctggg
3601 actcctgggc ccatacatca gggctgaggt ggaggacaac atcatggtga ccttccgcaa
3661 ccaggcctcc aggccctaca gcttctacag ctccctcatc agctatgagg aggaccagag
3721 gcaggggct gagccacgca agaactttgt gaaacccaat gaaaccaaga cctacttctg
3781 gaaagtccag caccacatgg ccccaccaa ggatgagttt gactgcaagg cctgggccta
3841 cttctctgat gtggacctgg agaaggatgt gcactctggc ctgattggcc cactcctggt
3901 ctgccacacc aacaccctga accctgccca tggaaggcaa gtgactgtgc aggagtttgc
3961 cctcttcttc accatctttg atgaaaccaa gagctggtac ttcactgaga acatggagcg
4021 caactgcagg gccccatgca acattcagat ggaggacccc accttcaaag agaactaccg
4081 cttccatgcc atcaatggct acatcatgga caccctgcct gggcttgtca tggcccagga
4141 ccagaggatc aggtggtacc tgctttctat gggctccaat gagaacattc actccatcca
4201 cttctctggg catgtcttca ctgtgcgcaa gaaggaggag tacaagatgg ccctgtacaa
4261 cctctaccct ggggtctttg agactgtgga gatgctgccc tccaaagctg gcatctggag
4321 ggtggagtgc ctcattgggg agcacctgca tgctggcatg agcaccctgt tcctggtcta
4381 cagcaacaag tgccagaccc ccctgggaat ggcctctggc cacatcaggg acttccagat
4441 cactgcctct ggccagtatg gccagtgggc ccccaagctg gccaggctcc actactctgg
4501 atccatcaat gcctggagca ccaaggagcc attcagctgg atcaaagtgg acctgctggc
4561 ccccatgatc atccatggca tcaagaccca gggggccagg cagaagttct ccagcctgta
4621 catcagccag ttcatcatca tgtacagcct ggatggcaag aaatggcaga cctacagagg
4681 caactccact ggaacactca tggtcttctt tggcaatgtg gacagctctg gcatcaagca
4741 caacatcttc aacccccaa tcatcgccag atacatcagg ctgcacccca cccactacag
4801 catccgcagc accctcagga tggagctgat gggctgtgac ctgaactcct gcagcatgcc
4861 cctgggcatg gagagcaagg ccatttctga tgcccagatc actgcctcca gctacttcac
4921 caacatgttt gccacctgga gcccaagcaa ggccaggctg cacctccagg gaaggagcaa
4981 tgcctggagg ccccaggtca acaacccaaa ggagtggctg caggtggact tccagaagac
```

图7B

```
5041 catgaaggtc actggggtga ccacccaggg ggtcaagagc ctgctcacca gcatgtatgt
5101 gaaggagttc ctgatcagct ccagccagga tggccaccag tggaccctct tcttccagaa
5161 tggcaaggtc aaggtgttcc agggcaacca ggacagcttc acccctgtgg tgaacagcct
5221 ggaccccccc ctcctgacca gatacctgag gattcacccc cagagctggg tccaccagat
5281 tgccctgagg atggaggtcc tgggatgtga ggcccaggac ctgtactgat gacgagcggc
5341 cgctcttagt agcagtatcg ataataaaag atctttattt tcattagatc tgtgtgttgg
5401 ttttttgtgt gttaattaag ctcgcgaagg aacccctagt gatggagttg gccactccct
5461 ctctgcgcgc tcgctcgctc actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct
5521 ttgcccgggc ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aagacgattt
5581 aaatgacaag cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc
5641 tcacaattcc acacaacata cgagccggaa gcataaagtg taaagcctgg ggtgcctaat
5701 gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc
5761 tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg
5821 ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag
5881 cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag
5941 gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc
6001 tggcgttttt ccataggctc cgcccccctg acgagcatca caaaaatcga cgctcaagtc
6061 agaggtggcg aaacccgaca ggactataaa gataccaggc gtttcccccc ggaagctccc
6121 tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt
6181 cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg
6241 ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat
6301 ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag
6361 ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt
6421 ggtggcctaa ctacggctac actagaagaa cagtatttgg tatctgcgct ctgctgaagc
6481 cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta
6541 gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag
6601 atcctttgat cttttctacg gggtctgacg ctcagtggaa cgaaaactca cgttaaggga
6661 ttttggtcat gagattatca aaaaggatct tcacctagat cctttttaaat taaaaatgaa
6721 gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa
6781 tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc
6841 ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga
6901 taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa
6961 gggccgagcg cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt
7021 gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg
7081 ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc
7141 aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg
7201 gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag
7261 cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt
7321 actcaaccaa gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt
7381 caatacggga taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac
7441 gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac
7501 ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag
7561 caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa
7621 tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga
7681 gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc
7741 cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa
7801 ataggcgtat cacgaggccc tttcgtc  (SEQ ID NO:8)
```

图7C

CS01m1-FL-NA

ATGCAGATTGAGCTGTCCACCTGCTTCTTTCTGTGCCTGCTGAGATTCTGCTTCTCTGCCACCAGGA
GATACTACCTGGGGGCTGTGGAACTTTCTTGGGACTACATGCAGTCTGACCTGGGAGAGCTGCCTGT
GGATGCCAGGTTCCCACCCAGAGTGCCCAAGTCCTTCCCATTCAACACCTCTGTGGTCTACAAGAAG
ACACTCTTTGTGGAATTCACTGACCACCTGTTCAACATTGCAAAACCCAGACCACCCTGGATGGGAC
TCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCATCACCCTCAAGAACATGGCATC
CCACCCTGTGTCTCTGCATGCTGTGGGAGTCTCATACTGGAAAGCCTCTGAAGGGGCTGAGTATGAT
GACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGGGATCTCACACCTATGTGT
GGCAAGTCCTCAAGGAGAATGGACCCATGGCATCTGACCCACTCTGCCTGACATACTCCTACCTTTC
TCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCACTGCTGGTGTGCAGGGAAGGA
TCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCATTCTCCTGTTTGCTGTCTTTGATGAGG
GCAAGTCTTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGATGCTGCCTCTGCCAGGGC
ATGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCTGGACTCATTGGCTGCCAC
AGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAAGTGCACTCCATTTTCCTGG
AGGGACACACCTTCCTGGTCAGGAACCACAGACAAGCCTCTCTGGAGATCTCTCCCATCACCTTCCT
CACTGCACAGACACTGCTGATGGACCTTGGACAGTTCCTGCTGTCCTGCCACATCTCTTCCCACCAG
CATGATGGCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCACAGCTCAGGATGAAGA
ACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATGGATGTGGTCAGATTTGA
TGATGACAACTCTCCATCCTTCATTCAGATCAGGTCTGTGGCAAAGAAACACCCCAAGACATGGGTG
CACTACATTGCTGCTGAGGAAGAGGACTGGGACTATGCACCACTGGTCCTGGCCCCTGATGACAGGA
GCTACAAGTCTCAGTACCTCAACAATGGCCCACAAAGAATTGGAAGAAAGTACAAGAAAGTCAGATT
CATGGCCTACACTGATGAAACCTTCAAGACAAGAGAAGCCATTCAGCATGAGTCTGGCATTCTGGGA
CCACTCCTGTATGGGGAAGTGGGAGACACCCTGCTCATCATCTTCAAGAACCAGGCCTCCAGGCCCT
ACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTACAGCAGGAGACTGCCAAAAGGGGT
GAAACACCTCAAGGACTTCCCCATTCTGCCTGGAGAGATCTTCAAGTACAAGTGGACTGTCACTGTG
GAGGATGGACCAACAAAGTCTGACCCCAGGTGCCTCACCAGATACTACTCCTCTTTTGTGAACATGG
AGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATCTGCTACAAGGAGTCTGTGGACCAGAG
AGGCAACCAGATCATGTCTGACAAGAGAAATGTGATTCTGTTCTCTGTCTTTGATGAGAACAGATCA
TGGTACCTGACTGAGAACATTCAGAGATTCCTGCCCAACCCTGCTGGGGTGCAACTGGAAGACCCTG
AGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGCTATGTGTTTGACTCTCTCCAGCTTTCTGT
CTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCTATTGGGGCACAAACTGACTTCCTTTCTGTC
TTCTTCTCTGGATACACCTTCAAGCACAAGATGGTGTATGAGGACACCCTGACACTCTTCCCATTCT
CTGGGGAAACTGTGTTCATGAGCATGGAGAACCCTGGACTGTGGATTCTGGGATGCCACAACTCTGA
CTTCAGAAACAGGGGAATGACTGCACTGCTCAAAGTCTCCTCCTGTGACAAGAACACTGGGGACTAC
TATGAGGACTCTTATGAGGACATCTCTGCCTACCTGCTCAGCAAGAACAATGCCATTGAGCCCAGAA
GCTTCTCTCAGAATCCACCTGTCCTGAAGAGACACCAGAGAGAGATCACCAGGACAACCCTCCAGTC
TGACCAGGAAGAGATTGACTATGATGACACCATTTCTGTGGAGATGAAGAAGGAGGACTTTGACATC
TATGATGAGGACGAGAACCAGTCTCCAAGATCATTCCAGAAGAAGACAAGACACTACTTCATTGCTG
CTGTGGAAAGACTGTGGGACTATGGCATGTCTTCCTCTCCCCATGTCCTCAGGAACAGGGCACAGTC
TGGCTCTGTGCCACAGTTCAAGAAAGTGGTCTTCCAGGAGTTCACTGATGGCTCATTCACCCAGCCC
CTGTACAGAGGGGAACTGAATGAGCACCTGGGACTCCTGGGACCATACATCAGGGCTGAGGTGGAAG
ACAACATCATGGTGACATTCAGAAACCAGGCCTCCAGGCCCTACAGCTTCTACTCTTCCCTCATCAG
CTATGAGGAAGACCAGAGACAAGGGGCTGAGCCAAGAAAGAACTTTGTGAAACCCAATGAAACCAAG
ACCTACTTCTGGAAAGTCCAGCACCACATGGCACCCACCAAGGATGAGTTTGACTGCAAGGCCTGGG

（续）

图8A

CATACTTCTCTGATGTGGACCTGGAGAAAGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTG
CCACACCAACACCCTGAACCCTGCACATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTC
ACCATCTTTGATGAAACCAAGTCATGGTACTTCACTGAGAACATGGAGAGAAACTGCAGAGCACCAT
GCAACATTCAGATGGAAGACCCCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACAT
CATGGACACCCTGCCTGGGCTTGTCATGGCACAGGACCAGAGAATCAGATGGTACCTGCTTTCTATG
GGATCCAATGAGAACATTCACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGAGAAAGAAGGAGG
AATACAAGATGGCCCTGTACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAA
AGCTGGCATCTGGAGGGTGGAATGCCTCATTGGGGAGCACCTGCATGCTGGCATGTCAACCCTGTTC
CTGGTCTACAGCAACAAGTGCCAGACACCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGA
TCACTGCCTCTGGCCAGTATGGCCAGTGGGCACCCAAACTGGCCAGGCTCCACTACTCTGGCTCCAT
CAATGCATGGTCAACCAAGGAGCCATTCTCTTGGATCAAGGTGGACCTGCTGGCACCCATGATCATT
CATGGCATCAAGACACAGGGGGCAAGACAGAAATTCTCCTCTCTGTACATCTCACAGTTCATCATCA
TGTACTCTCTGGATGGCAAGAAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCTCATGGTCTT
CTTTGGCAATGTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCTCCCATCATTGCCAGATAC
ATCAGGCTGCACCCCACCCACTACTCAATCAGATCAACCCTCAGGATGGAACTGATGGGATGTGACC
TGAACTCCTGCTCAATGCCCCTGGGAATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCATC
CTCTTACTTCACCAACATGTTTGCCACCTGGTCACCATCAAAAGCCAGGCTGCACCTCCAGGGAAGA
AGCAATGCCTGGAGACCCCAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAA
TGAAAGTCACTGGGGTGACAACCCAGGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGAGTT
CCTGATCTCTTCCTCACAGGATGGCCACCAGTGGACACTCTTCTTCCAGAATGGCAAAGTCAAGGTG
TTCCAGGGCAACCAGGACTCTTTCACACCTGTGGTGAACTCACTGGACCCCCCCCTCCTGACAAGAT
ACCTGAGAATTCACCCCCAGTCTTGGGTCCACCAGATTGCCCTGAGAATGGAAGTCCTGGGATGTGA
GGCACAAGACCTGTACTGA (SEQ ID NO:49)

图8B

CS04Δ(760-1667) - CS04-SC1-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGGAGATAC
TACCTGGGGGCTGTGGAGCTTTCTTGGGACTACATGCAGTCTGACCTGGGGGAGCTGCCTGTGGATGCCAGG
TTCCCACCCAGAGTGCCCAAATCCTTCCCATTCAACACCTCTGTGGTCTACAAGAAGACCCTCTTTGTGGAG
TTCACTGACCACCTGTTCAACATTGCCAAACCCAGGCCACCCTGGATGGGACTCCTGGGACCCACCATTCAG
GCTGAGGTGTATGACACTGTGGTCATCACCCTCAAGAACATGGCCTCCCACCCTGTGAGCCTGCATGCTGTG
GGGGTCAGCTACTGGAAGGCCTCTGAGGGGGCTGAGTATGATGACCAGACCTCCCAGAGGGAGAAGGAGGAT
GACAAAGTGTTCCCTGGGGGCAGCCACACCTATGTGTGGCAGGTCCTCAAGGAGAATGGCCCCATGGCCTCT
GACCCACTCTGCCTGACCTACTCCTACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATT
GGGGCCCTGCTGGTGTGCAGGGAGGGCTCCCTGGCCAAAGAGAAGACCCAGACCCTGCACAAGTTCATTCTC
CTGTTTGCTGTCTTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACTCCCTGATGCAGGACAGGGAT
GCTGCCTCTGCCAGGGCCTGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAGGTGCACTCCATT
TTCCTGGAGGGCCACACCTTCCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGCCCCATCACCTTC
CTCACTGCCCAGACCCTGCTGATGGACCTCGGACAGTTCCTGCTGTTCTGCCACATCAGCTCCCACCAGCAT
GATGGCATGGAGGCCTATGTCAAGGTGGACAGCTGCCCTGAGGAGCCACAGCTCAGGATGAAGAACAATGAG
GAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATGGATGTGGTCCGCTTTGATGATGACAACAGC
CCATCCTTCATTCAGATCAGGTCTGTGGCCAAGAAACACCCCAAGACCTGGGTGCACTACATTGCTGCTGAG
GAGGAGGACTGGGACTATGCCCCACTGGTCCTGGCCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAAC
AATGGCCCACAGAGGATTGGACGCAAGTACAAGAAAGTCAGGTTCATGGCCTACACTGATGAAACCTTCAAG
ACCAGGGAGGCCATTCAGCATGAGTCTGGCATCCTGGGCCCACTCCTGTATGGGGAGGTGGGGGACACCCTG
CTCATCATCTTCAAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCC
CTGTACAGCCGCAGGCTGCCAAAGGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGGGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACCAAATCTGACCCCAGGTGCCTCACCAGATACTAC
TCCAGCTTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCCACTGCTCATCTGCTACAAGGAG
TCTGTGGACCAGAGGGGAAACCAGATCATGTCTGACAAGAGGAATGTGATTCTGTTCTCTGTCTTTGATGAG
AACAGGAGCTGGTACCTGACTGAGAACATTCAGCGCTTCCTGCCCAACCCTGCTGGGGTGCAGCTGGAGGAC
CCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGCTATGTGTTTGACAGCCTCCAGCTTTCTGTC
TGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCTATTGGGGCCCAGACTGACTTCCTTTCTGTCTTCTTC
TCTGGCTACACCTTCAAACACAAGATGGTGTATGAGGACACCCTGACCCTCTTCCCATTCTCTGGGGAGACT
GTGTTCATGAGCATGGAGAACCCTGGCCTGTGGATTCTGGGATGCCACAACTCTGACTTCCGCAACAGGGGC
ATGACTGCCCTGCTCAAAGTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACAGCTATGAGGAC
ATCTCTGCCTACCTGCTCAGCAAGAACAATGCCATTGAGCCCAGGGAGATCACCAGGACCACCCTCCAGTCT
GACCAGGAGGAGATTGACTATGATGACACCATTTCTGTGGAGATGAAGAAAGAGGACTTTGACATCTATGAC
GAGGACGAGAACCAGAGCCCAAGGAGCTTCCAGAAGAAGACCAGGCACTACTTCATTGCTGCTGTGGAGCGC
CTGTGGGACTATGGCATGAGCTCCAGCCCCCATGTCCTCAGGAACAGGGCCCAGTCTGGCTCTGTGCCACAG
TTCAAGAAAGTGGTCTTCCAAGAGTTCACTGATGGCAGCTTCACCCAGCCCCTGTACAGAGGGGAGCTGAAT
GAGCACCTGGGACTCCTGGGCCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACCTTCCGCAAC
CAGGCCTCCAGGCCCTACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAGGGGGCTGAG
CCACGCAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCACCACATGGCCCCC

（续）

图9A

ACCAAGGATGAGTTTGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGACCTGGAGAAGGATGTGCACTCT
GGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAACCCTGCCCATGGAAGGCAAGTGACTGTG
CAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACCAAGAGCTGGTACTTCACTGAGAACATGGAGCGC
AACTGCAGGGCCCCATGCAACATTCAGATGGAGGACCCCACCTTCAAAGAGAACTACCGCTTCCATGCCATC
AATGGCTACATCATGGACACCCTGCCTGGGCTTGTCATGGCCCAGGACCAGAGGATCAGGTGGTACCTGCTT
TCTATGGGCTCCAATGAGAACATTCACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGCGCAAGAAGGAG
GAGTACAAGATGGCCCTGTACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCT
GGCATCTGGAGGGTGGAGTGCCTCATTGGGGAGCACCTGCATGCTGGCATGAGCACCCTGTTCCTGGTCTAC
AGCAACAAGTGCCAGACCCCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGC
CAGTATGGCCAGTGGGCCCCCAAGCTGGCCAGGCTCCACTACTCTGGATCCATCAATGCCTGGAGCACCAAG
GAGCCATTCAGCTGGATCAAAGTGGACCTGCTGGCCCCCATGATCATCCATGGCATCAAGACCCAGGGGGCC
AGGCAGAAGTTCTCCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTGGATGGCAAGAAATGGCAG
ACCTACAGAGGCAACTCCACTGGAACACTCATGGTCTTCTTTGGCAATGTGGACAGCTCTGGCATCAAGCAC
AACATCTTCAACCCCCCAATCATCGCCAGATACATCAGGCTGCACCCCACCCACTACAGCATCCGCAGCACC
CTCAGGATGGAGCTGATGGGCTGTGACCTGAACTCCTGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATT
TCTGATGCCCAGATCACTGCCTCCAGCTACTTCACCAACATGTTTGCCACCTGGAGCCCAAGCAAGGCCAGG
CTGCACCTCCAGGGAAGGAGCAATGCCTGGAGGCCCCAGGTCAACAACCCAAAGGAGTGGCTGCAGGTGGAC
TTCCAGAAGACCATGAAGGTCACTGGGGTGACCACCCAGGGGGTCAAGAGCCTGCTCACCAGCATGTATGTG
AAGGAGTTCCTGATCAGCTCCAGCCAGGATGGCCACCAGTGGACCCTCTTCTTCCAGAATGGCAAGGTCAAG
GTGTTCCAGGGCAACCAGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCCCTCCTGACCAGATAC
CTGAGGATTCACCCCCAGAGCTGGGTCCACCAGATTGCCCTGAGGATGGAGGTCCTGGGATGTGAGGCCCAG
GACCTGTACTGA (SEQ ID NO:9)

图9B

CS04Δ(760-1667) - CS04-SC1-AA

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKK
TLFVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEYD
DQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREG
SLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCH
RKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQ
HDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKTWV
HYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILG
PLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTV
EDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRS
WYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSV
FFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDY
YEDSYEDISAYLLSKNNAIEPREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQK
KTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLG
PYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTK
DEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTEN
MERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGH
VFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMA
SGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSS
LYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTL
RMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEW
LQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNS
LDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY (SEQ ID NO:10)

图10

CS04Δ(772-1667) - CS04-SC2-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGCTGTGGAGCTTTCTTGGGACTACATGCAGTCTGACCTGGGGGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAATCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACCCTCTTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAAACCCAGGCCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCATCACCCTCAAGAACATG
GCCTCCCACCCTGTGAGCCTGCATGCTGTGGGGGTCAGCTACTGGAAGGCCTCTGAGGGGGCTGAG
TATGATGACCAGACCTCCCAGAGGGAGAAGGAGGATGACAAAGTGTTCCCTGGGGGCAGCCACACC
TATGTGTGGCAGGTCCTCAAGGAGAATGGCCCCATGGCCTCTGACCCACTCTGCCTGACCTACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCCCTGCTGGTGTGC
AGGGAGGGCTCCCTGGCCAAAGAGAAGACCCAGACCCTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACTCCCTGATGCAGGACAGGGATGCTGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAGGTGCAC
TCCATTTTCCTGGAGGGCCACACCTTCCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCCTCACTGCCCAGACCCTGCTGATGGACCTCGGACAGTTCCTGCTGTTCTGCCAC
ATCAGCTCCCACCAGCATGATGGCATGGAGGCCTATGTCAAGGTGGACAGCTGCCCTGAGGAGCCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCCGCTTTGATGATGACAACAGCCCATCCTTCATTCAGATCAGGTCTGTGGCCAAGAAA
CACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGGACTGGGACTATGCCCCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGGCCCACAGAGGATTGGACGC
AAGTACAAGAAAGTCAGGTTCATGGCCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATTCAG
CATGAGTCTGGCATCCTGGGCCCACTCCTGTATGGGGAGGTGGGGGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCCGCAGGCTGCCAAAGGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGGGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACCAAATCTGACCCCAGGTGCCTCACCAGA
TACTACTCCAGCTTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGGGGAAACCAGATCATGTCTGACAAGAGGAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAGCGCTTCCTGCCCAAC
CCTGCTGGGGTGCAGCTGGAGGACCCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACAGCCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCCCAGACTGACTTCCTTTCTGTCTTCTTCTCTGGCTACACCTTCAAACACAAGATGGTG
TATGAGGACACCCTGACCCTCTTCCCATTCTCTGGGGAGACTGTGTTCATGAGCATGGAGAACCCT
GGCCTGTGGATTCTGGGATGCCACAACTCTGACTTCCGCAACAGGGGCATGACTGCCCTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACAGCTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATGCCATTGAGCCCAGGAGCTTCAGCCAGAATTCCAGACACCCCAGCACC
AGGGAGATCACCAGGACCACCCTCCAGTCTGACCAGGAGGAGATTGACTATGATGACACCATTTCT
GTGGAGATGAAGAAAGAGGACTTTGACATCTATGACGAGGACGAGAACCAGAGCCCAAGGAGCTTC

（续）

图11A

CAGAAGAAGACCAGGCACTACTTCATTGCTGCTGTGGAGCGCCTGTGGGACTATGGCATGAGCTCC
AGCCCCCATGTCCTCAGGAACAGGGCCCAGTCTGGCTCTGTGCCACAGTTCAAGAAAGTGGTCTTC
CAAGAGTTCACTGATGGCAGCTTCACCCAGCCCCTGTACAGAGGGGAGCTGAATGAGCACCTGGGA
CTCCTGGGCCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACCTTCCGCAACCAGGCC
TCCAGGCCCTACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAGGGGGCTGAG
CCACGCAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCACCACATG
GCCCCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGACCTGGAGAAG
GATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAACCCTGCCCAT
GGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACCAAGAGCTGG
TACTTCACTGAGAACATGGAGCGCAACTGCAGGGCCCCATGCAACATTCAGATGGAGGACCCCACC
TTCAAAGAGAACTACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCTGGGCTTGTC
ATGGCCCAGGACCAGAGGATCAGGTGGTACCTGCTTTCTATGGGCTCCAATGAGAACATTCACTCC
ATCCACTTCTCTGGGCATGTCTTCACTGTGCGCAAGAAGGAGGAGTACAAGATGGCCCTGTACAAC
CTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGGGTGGAG
TGCCTCATTGGGGAGCACCTGCATGCTGGCATGAGCACCCTGTTCCTGGTCTACAGCAACAAGTGC
CAGACCCCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGCCAGTAT
GGCCAGTGGGCCCCCAAGCTGGCCAGGCTCCACTACTCTGGATCCATCAATGCCTGGAGCACCAAG
GAGCCATTCAGCTGGATCAAAGTGGACCTGCTGGCCCCCATGATCATCCATGGCATCAAGACCCAG
GGGGCCAGGCAGAAGTTCTCCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTGGATGGC
AAGAAATGGCAGACCTACAGAGGCAACTCCACTGGAACACTCATGGTCTTCTTTGGCAATGTGGAC
AGCTCTGGCATCAAGCACAACATCTTCAACCCCCCAATCATCGCCAGATACATCAGGCTGCACCCC
ACCCACTACAGCATCCGCAGCACCCTCAGGATGGAGCTGATGGGCTGTGACCTGAACTCCTGCAGC
ATGCCCCTGGGCATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCCTCCAGCTACTTCACC
AACATGTTTGCCACCTGGAGCCCAAGCAAGGCCAGGCTGCACCTCCAGGGAAGGAGCAATGCCTGG
AGGCCCCAGGTCAACAACCCAAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATGAAGGTCACT
GGGGTGACCACCCAGGGGGTCAAGAGCCTGCTCACCAGCATGTATGTGAAGGAGTTCCTGATCAGC
TCCAGCCAGGATGGCCACCAGTGGACCCTCTTCTTCCAGAATGGCAAGGTCAAGGTGTTCCAGGGC
AACCAGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCCCTCCTGACCAGATACCTGAGG
ATTCACCCCCAGAGCTGGGTCCACCAGATTGCCCTGAGGATGGAGGTCCTGGGATGTGAGGCCCAG
GACCTGTACTGA (SEQ ID NO:11)

图11B

CS04Δ(772-1667) - CS04-SC2-AA

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEF
TDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDK
VFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFA
VFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFLEG
HTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDY
DDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRI
GRKYKKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLP
KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQ
IMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWY
ILSIGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSC
DKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNSRHPSTREITRTTLQSDQEEIDYDDTISVEMKKEDFDIY
DEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELN
EHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPT
KDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNC
RAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYK
MALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQ
WAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGN
STGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQIT
ASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISS
SQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY
(SEQ ID NO:12)

图12

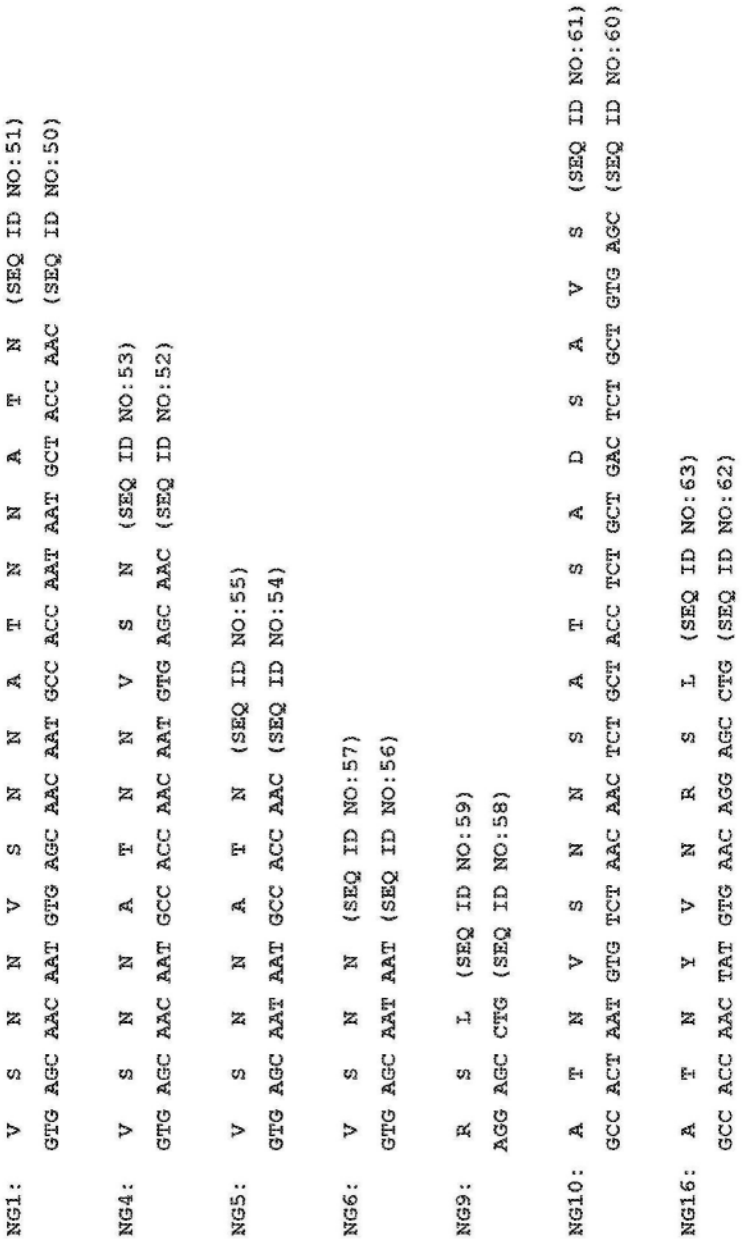

NG1: V S N N V S N N A T N N A T N (SEQ ID NO:51)
GTG AGC AAC AAT GTG AGC AAC AAT GCC ACC AAT AAT GCT ACC AAC (SEQ ID NO:50)
NG4: V S N N A T N N V S N (SEQ ID NO:53)
GTG AGC AAC AAT GCC ACC AAC AAT GTG AGC AAC (SEQ ID NO:52)
NG5: V S N N A T N (SEQ ID NO:55)
GTG AGC AAT AAT GCC ACC AAC (SEQ ID NO:54)
NG6: V S N N (SEQ ID NO:57)
GTG AGC AAT AAT (SEQ ID NO:56)
NG9: R S L (SEQ ID NO:59)
AGG AGC CTG (SEQ ID NO:58)
NG10: A T N V S N N S A T S A D S A V S (SEQ ID NO:61)
GCC ACT AAT GTG TCT AAC AAC TCT GCT ACC TCT GAC TCT GCT GTG AGC (SEQ ID NO:60)
NG16: A T N Y V N R S L (SEQ ID NO:63)
GCC ACC AAC TAT GTG AAC AGG AGC CTG (SEQ ID NO:62)

图13A

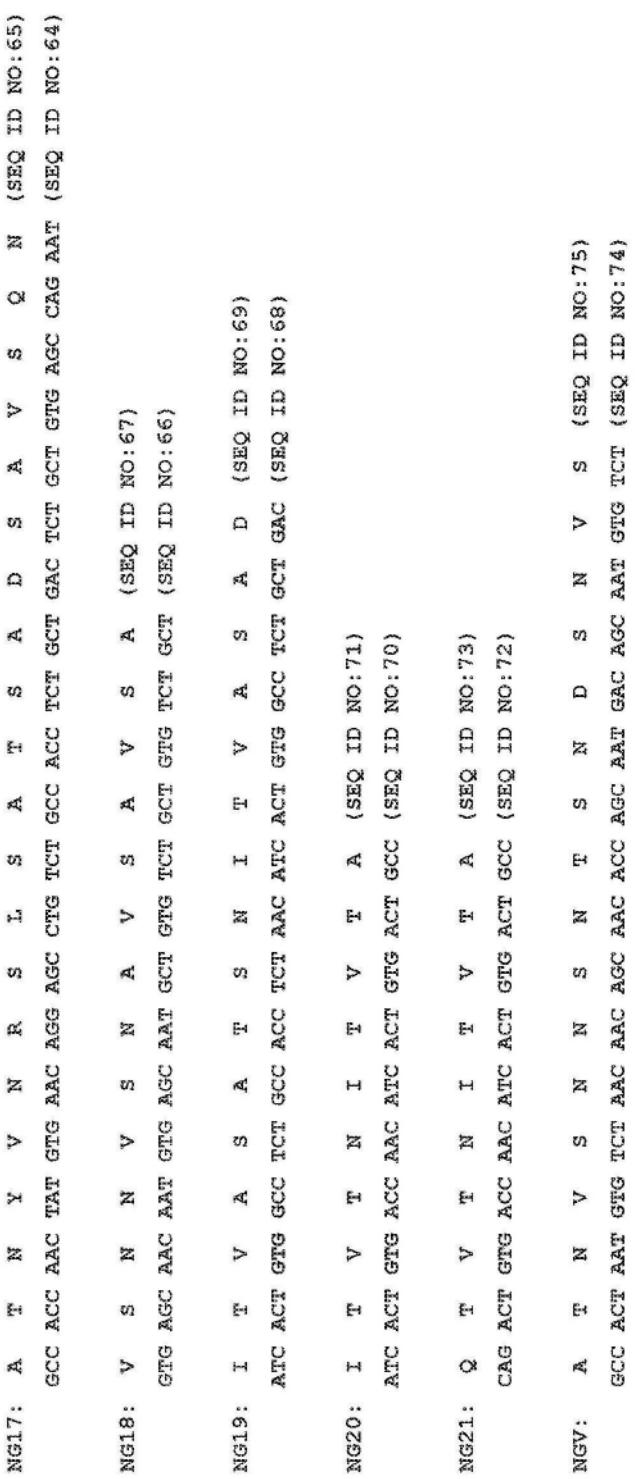

NG17: A T N Y V N R S L S A T S A D S A V S Q N (SEQ ID NO:65)
GCC ACC AAC TAT GTG AAC AGG AGC CTG TCT GCC ACC TCT GCT GAC TCT GCT GTG AGC CAG AAT (SEQ ID NO:64)
NG18: V S N N V S N A V S A V S A (SEQ ID NO:67)
GTG AGC AAC AAT GTG AGC AAT GCT GTG TCT GCT GTG TCT GCT (SEQ ID NO:66)
NG19: I T V A S A T S N I T V A S A D (SEQ ID NO:69)
ATC ACT GTG GCC TCT GCC ACC TCT AAC ATC ACT GTG GCC TCT GCT GAC (SEQ ID NO:68)
NG20: I T V T N I T V T A (SEQ ID NO:71)
ATC ACT GTG ACC AAC ATC ACT GTG ACT GCC (SEQ ID NO:70)
NG21: Q T V T N I T V T A (SEQ ID NO:73)
CAG ACT GTG ACC AAC ATC ACT GTG ACT GCC (SEQ ID NO:72)
NGV: A T N V S N N S N T S N D S N V S (SEQ ID NO:75)
GCC ACT AAT GTG TCT AAC AAC AGC AAC ACC AGC AAT GAC AGC AAT GTG TCT (SEQ ID NO:74)

图13B

图14

名称:序列 长度: 41

LSKNNAIEPRSFSQNATNVSNNSNTSNDSNVSPPVLKRHQR

. . . . . N..N. . . . . .

(阈值=0.5)

序列名称	位置	潜能	评审一致	N-Glyc 结果
序列	15 NATN	0.6224	(8/9)	+
序列	18 NVSN	0.6453	(9/9)	++
序列	21 NNSN	0.4158	(8/9)	-
序列	24 NTSN	0.4158	(8/9)	-
序列	27 NDSN	0.3619	(8/9)	-
序列	30 NVSP	0.1149	(9/9)	---

图15

CS01-FL-NA

ATGCAGATTGAGCTGTCCACCTGCTTCTTTCTGTGCCTGCTGAGATTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGCTGTGGAACTTTCTTGGGACTACATGCAGTCTGACCTGGGAGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAGTCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACACTCTTTGTGGAATTCACTGACCACCTGTTCAACATTGCAAAACCCAGACCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCATCACCCTCAAGAACATG
GCATCCCACCCTGTGTCTCTGCATGCTGTGGGAGTCTCATACTGGAAAGCCTCTGAAGGGGCTGAG
TATGATGACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGGGATCTCACACC
TATGTGTGGCAAGTCCTCAAGGAGAATGGACCCATGGCATCTGACCCACTCTGCCTGACATACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCACTGCTGGTGTGC
AGGGAAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGTCTTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGATGCTGCC
TCTGCCAGGGCATGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAAGTGCAC
TCCATTTTCCTGGAGGGACACACCTTCCTGGTCAGGAACCACAGACAAGCCTCTCTGGAGATCTCT
CCCATCACCTTCCTCACTGCACAGACACTGCTGATGGACCTTGGACAGTTCCTGCTGTTCTGCCAC
ATCTCTTCCCACCAGCATGATGGCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCAGATTTGATGATGACAACTCTCCATCCTTCATTCAGATCAGGTCTGTGGCAAAGAAA
CACCCCAAGACATGGGTGCACTACATTGCTGCTGAGGAAGAGGACTGGGACTATGCACCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAACAATGGCCCACAAAGAATTGGAAGA
AAGTACAAGAAAGTCAGATTCATGGCCTACACTGATGAAACCTTCAAGACAAGAGAAGCCATTCAG
CATGAGTCTGGCATTCTGGGACCACTCCTGTATGGGGAAGTGGGAGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCAGGAGACTGCCAAAAGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGAGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCCAGGTGCCTCACCAGA
TACTACTCCTCTTTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGAGGCAACCAGATCATGTCTGACAAGAGAAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTCAGAGATTCCTGCCCAAC
CCTGCTGGGGTGCAACTGGAAGACCCTGAGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACTCTCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCACAAACTGACTTCCTTTCTGTCTTCTTCTCTGGATACACCTTCAAGCACAAGATGGTG
TATGAGGACACCCTGACACTCTTCCCATTCTCTGGGGAAACTGTGTTCATGAGCATGGAGAACCCT
GGACTGTGGATTCTGGGATGCCACAACTCTGACTTCAGAAACAGGGGAATGACTGCACTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATGCCATTGAGCCCAGAAGCTTCTCTCAGAATCCACCTGTCCTGAAGAGA
CACCAGAGAGAGATCACCAGGACAACCCTCCAGTCTGACCAGGAAGAGATTGACTATGATGACACC
ATTTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTATGATGAGGACGAGAACCAGTCTCCAAGA
TCATTCCAGAAGAAGACAAGACACTACTTCATTGCTGCTGTGGAAAGACTGTGGGACTATGGCATG
TCTTCCTCTCCCCATGTCCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAAGAAAGTG
GTCTTCCAGGAGTTCACTGATGGCTCATTCACCCAGCCCCTGTACAGAGGGGAACTGAATGAGCAC

（续）

图16A

CTGGGACTCCTGGGACCATACATCAGGGCTGAGGTGGAAGACAACATCATGGTGACATTCAGAAAC
CAGGCCTCCAGGCCCTACAGCTTCTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAAGGG
GCTGAGCCAAGAAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCAC
CACATGGCACCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCATACTTCTCTGATGTGGACCTG
GAGAAAGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAACCCT
GCACATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACCAAG
TCATGGTACTTCACTGAGAACATGGAGAGAAACTGCAGAGCACCATGCAACATTCAGATGGAAGAC
CCCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCTGGG
CTTGTCATGGCACAGGACCAGAGAATCAGATGGTACCTGCTTTCTATGGGATCCAATGAGAACATT
CACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGAGAAAGAAGGAGGAATACAAGATGGCCCTG
TACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGG
GTGGAATGCCTCATTGGGGAGCACCTGCATGCTGGCATGTCAACCCTGTTCCTGGTCTACAGCAAC
AAGTGCCAGACACCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGC
CAGTATGGCCAGTGGGCACCCAAACTGGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGGTCA
ACCAAGGAGCCATTCTCTTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTCATGGCATCAAG
ACACAGGGGGCAAGACAGAAATTCTCCTCTCTGTACATCTCACAGTTCATCATCATGTACTCTCTG
GATGGCAAGAAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCTCATGGTCTTCTTTGGCAAT
GTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCTCCCATCATTGCCAGATACATCAGGCTG
CACCCCACCCACTACTCAATCAGATCAACCCTCAGGATGGAACTGATGGGATGTGACCTGAACTCC
TGCTCAATGCCCCTGGGAATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCATCCTCTTAC
TTCACCAACATGTTTGCCACCTGGTCACCATCAAAAGCCAGGCTGCACCTCCAGGGAAGAAGCAAT
GCCTGGAGACCCCAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAATGAAA
GTCACTGGGGTGACAACCCAGGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGAGTTCCTG
ATCTCTTCCTCACAGGATGGCCACCAGTGGACACTCTTCTTCCAGAATGGCAAAGTCAAGGTGTTC
CAGGGCAACCAGGACTCTTTCACACCTGTGGTGAACTCACTGGACCCCCCCCTCCTGACAAGATAC
CTGAGAATTCACCCCCAGTCTTGGGTCCACCAGATTGCCCTGAGAATGGAAGTCCTGGGATGTGAG
GCACAAGACCTGTACTGA (SEQ ID NO:13)

图16B

CS08-FL-NA

ATGCAGATCGAACTGAGCACTTGCTTCTTCCTGTGTCTCCTGCGCTTTTGCTTCTCCGCCACAAGG
AGATACTATCTCGGTGCCGTGGAGCTCAGCTGGGACTACATGCAGAGCGACTTGGGTGAACTGCCT
GTGGACGCCAGGTTTCCACCCCGCGTGCCCAAGAGTTTCCCGTTCAACACCAGTGTCGTGTACAAG
AAAACCCTCTTCGTGGAATTCACCGACCACCTGTTCAACATCGCCAAACCGCGCCCTCCCTGGATG
GGGCTGCTCGGCCCGACGATCCAGGCTGAGGTCTATGACACGGTGGTGATTACCCTCAAGAACATG
GCTAGCCACCCGGTGAGCCTGCACGCCGTGGGCGTGTCCTATTGGAAAGCGTCCGAGGGTGCGGAG
TACGATGACCAGACTTCACAGCGGGAGAAGGAAGACGACAAAGTGTTCCCCGGGGGTTCCCACACC
TATGTCTGGCAGGTCCTGAAGGAGAATGGTCCTATGGCCTCCGACCCATTGTGCCTCACCTACTCT
TACCTAAGCCATGTGGATCTCGTCAAGGACCTGAACTCGGGGCTGATCGGCGCCCTGCTCGTGTGC
CGGGAGGGCTCACTGGCCAAGGAGAAGACCCAAACTCTGCACAAGTTCATCCTGCTGTTCGCGGTA
TTCGACGAGGGGAAGTCCTGGCACTCCGAGACCAAGAACAGCCTGATGCAGGACCGCGACGCAGCC
TCGGCCCGTGCGTGGCCAAAGATGCACACCGTGAACGGCTACGTTAACAGGAGCCTACCCGGCCTG
ATCGGCTGCCACCGCAAATCGGTCTACTGGCATGTGATCGGAATGGGCACAACGCCCGAGGTCCAC
AGTATCTTCCTCGAGGGCCACACTTTCCTGGTCCGGAATCACCGCCAGGCCAGCCTGGAGATCAGC
CCCATAACCTTTCTGACGGCGCAGACCTTACTCATGGATCTCGGCCAGTTCCTCCTGTTCTGCCAC
ATTTCGTCCCACCAGCACGATGGGATGGAAGCATATGTGAAAGTGGACTCCTGCCCCGAGGAACCC
CAGCTTAGGATGAAGAACAATGAGGAGGCCGAGGACTACGACGATGACCTTACCGATTCAGAAATG
GACGTAGTACGCTTTGACGACGACAACTCTCCATCCTTCATACAGATTCGCTCCGTCGCCAAGAAG
CACCCTAAGACTTGGGTGCACTACATCGCGGCCGAGGAGGAGGACTGGGATTATGCTCCCCTGGTG
CTGGCCCCCGACGACCGCAGCTACAAGAGCCAGTACCTGAATAACGGGCCCCAGCGCATCGGCCGG
AAGTACAAGAAAGTGCGGTTCATGGCTTACACGGACGAGACCTTCAAGACCCGGGAGGCTATCCAG
CATGAGAGCGGCATCTTGGGGCCCCTCCTGTACGGCGAAGTTGGAGACACACTGCTGATCATCTTC
AAGAACCAGGCGAGCAGGCCCTACAACATCTACCCCCACGGCATTACCGATGTCCGGCCGTTGTAC
AGCCGACGGCTGCCCAAGGGCGTGAAGCACCTGAAGGACTTTCCGATCCTGCCGGGGCGAGATCTTC
AAGTACAAGTGGACTGTGACCGTGGAGGATGGGCCGACCAAGAGCGATCCGCGCTGCCTGACCCGT
TACTACTCCAGCTTTGTCAATATGGAGCGCGACCTCGCTAGCGGCTTGATTGGCCCTCTGCTGATC
TGCTACAAGGAGTCCGTGGACCAGAGGGGGAATCAGATCATGAGTGACAAGAGGAACGTGATCCTG
TTCTCCGTGTTCGACGAAAACCGCAGCTGGTATCTCACCGAGAATATCCAGCGCTTCCTGCCCAAC
CCGGCCGGTGTGCAGCTGGAGGACCCCGAGTTTCAGGCCAGCAACATCATGCATTCTATCAACGGA
TATGTGTTTGATTCCCTGCAGCTCTCAGTGTGTCTGCACGAGGTCGCCTACTGGTATATCCTCAGC
ATTGGGGCACAGACCGACTTCCTGAGCGTGTTCTTCTCCGGGTATACCTTCAAGCACAAGATGGTG
TACGAGGATACCCTGACCCTGTTCCCCTTTAGCGGCGAAACCGTGTTTATGTCTATGGAGAACCCC
GGGCTCTGGATCCTTGGCTGCCATAACTCCGACTTCCGCAACCGCGGAATGACCGCGCTCCTGAAA
GTGTCGAGTTGTGACAAGAACACCGGCGACTATTACGAGGACAGTTACGAGGACATCTCTGCGTAC
CTCCTTAGCAAGAATAACGCCATCGAGCCAAGATCCTTCAGCCAGAACCCCCCAGTGCTGAAGAGG
CATCAGCGGGAGATCACCCGCACGACCCTGCAGTCGGATCAGGAGGAGATTGATTACGACGACACG
ATCAGTGTGGAGATGAAGAAGGAGGACTTCGACATCTACGACGAAGATGAAAACCAGTCCCCTCGG
TCCTTCCAAAAGAAGACCCGGCACTACTTCATCGCCGCTGTGGAACGCCTGTGGGACTATGGAATG

（续）

图17A

TCTTCTAGCCCTCACGTTTTGAGGAACCGCGCCCAGTCGGGCAGCGTGCCCCAGTTCAAGAAAGTG
GTGTTCCAGGAGTTCACCGACGGCTCCTTCACCCAGCCACTTTACCGGGGCGAGCTCAATGAACAT
CTGGGCCTGCTGGGACCCTACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACATTCCGGAAT
CAGGCCAGCAGACCATACAGTTTCTACAGTTCACTCATCTCCTACGAGGAGGACCAGCGCCAGGGG
GCTGAACCCCGTAAGAACTTCGTGAAGCCAAACGAAACAAAGACCTACTTCTGGAAGGTCCAGCAC
CACATGGCACCTACCAAGGACGAGTTCGATTGCAAGGCCTGGGCCTACTTCTCCGACGTGGACCTG
GAGAAAGATGTGCACAGCGGCCTGATTGGCCCTCTGCTGGTGTGTCACACGAACACACTCAACCCT
GCACACGGGCGGCAGGTCACTGTGCAGGAATTCGCCCTGTTCTTTACCATCTTTGATGAGACGAAG
TCCTGGTATTTCACCGAAAACATGGAGAGGAACTGCCGCGCACCCTGCAACATCCAGATGGAAGAT
CCGACATTCAAGGAGAACTACCGGTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCTGGC
CTCGTGATGGCCCAAGACCAGCGTATCCGCTGGTATCTGCTGTCGATGGGCTCCAACGAGAACATC
CATAGTATCCACTTCAGCGGGCATGTCTTCACGGTGAGGAAAAAGGAGGAGTACAAGATGGCACTG
TACAACCTCTATCCCGGCGTGTTCGAGACCGTGGAGATGCTGCCCTCCAAGGCCGGCATCTGGAGA
GTGGAATGCCTGATCGGCGAGCACCTCCACGCTGGGATGTCCACGCTGTTCCTCGTTTACAGCAAT
AAGTGCCAGACCCCTCTGGGCATGGCGAGCGGCCACATCCGCGACTTCCAGATTACAGCCAGCGGC
CAGTACGGTCAGTGGGCTCCAAAGCTGGCCCGTCTGCACTACTCCGGATCCATCAACGCCTGGTCC
ACCAAGGAACCGTTCTCCTGGATCAAAGTAGACCTGCTAGCCCCCATGATCATTCACGGCATCAAG
ACACAAGGCGCCCGACAGAAGTTCTCGAGCCTCTATATCTCCCAGTTCATCATCATGTATAGCCTG
GACGGAAAGAAGTGGCAGACTTACCGCGGAAACTCGACAGGGACCCTGATGGTATTCTTCGGTAAC
GTGGACAGCTCCGGAATCAAGCACAACATCTTCAACCCACCCATTATCGCCCGCTACATCCGCCTG
CACCCCACTCACTATAGCATTAGGTCCACCCTGCGAATGGAGCTCATGGGCTGTGACCTGAACAGC
TGTAGCATGCCCCTCGGCATGGAGTCTAAGGCGATCTCCGACGCACAGATAACGGCATCATCCTAC
TTTACCAACATGTTCGCTACCTGGTCCCCCTCCAAGGCCCGACTCCACCTGCAAGGGAGATCCAAC
GCCTGGCGGCCACAGGTCAACAATCCCAAGGAGTGGCTGCAAGTGGACTTTCAGAAAACTATGAAA
GTCACCGGAGTGACCACACAGGGAGTGAAGTCTCTGCTGACCAGCATGTACGTGAAGGAGTTCCTC
ATCTCCAGTTCGCAGGATGGCCACCAGTGGACGTTGTTCTTCCAAAACGGTAAAGTCAAAGTCTTC
CAAGGGAACCAGGACAGCTTTACACCCGTCGTGAACTCCCTGGACCCCCCGCTTCTCACTAGATAC
CTCCGCATCCACCCTCAGAGCTGGGTGCACCAGATTGCCCTGCGCATGGAGGTTCTGGGGTGTGAA
GCCCAGGACCTGTACTAA (SEQ ID NO:14)

图17B

CS10-FL-NA

ATGCAGATTGAGCTCTCCACCTGCTTCTTTCTCTGCCTTCTTCGCTTCTGCTTTTCTGCCACACGC
AGGTACTATTTGGGAGCAGTGGAACTGAGCTGGGATTACATGCAGAGTGACCTTGGTGAACTTCCT
GTGGACGCTCGTTTTCCACCTAGAGTTCCCAAGTCCTTCCCCTTCAACACCTCAGTGGTCTACAAG
AAAACGCTGTTTGTGGAGTTCACTGACCACCTCTTCAACATTGCCAAACCAAGACCCCCTTGGATG
GGATTGCTGGGACCCACAATACAAGCAGAAGTCTACGACACGGTGGTGATTACCCTGAAGAACATG
GCGTCACACCCTGTTTCACTTCACGCTGTTGGGGTCAGTTATTGGAAAGCCTCAGAGGGTGCGGAA
TACGATGATCAAACCAGCCAGAGGGAGAAGGAAGATGACAAGGTCTTTCCTGGGGGTAGCCATACC
TATGTTTGGCAGGTGCTGAAAGAGAATGGGCCTATGGCCTCTGATCCCTTGTGCCTCACATACTCT
TACCTGAGTCACGTCGACCTGGTGAAAGACCTGAATAGCGGTCTGATTGGTGCACTGCTTGTTTGT
AGAGAGGGGAGTTTGGCCAAGGAGAAAACTCAGACTCTCCACAAGTTTATCCTCCTGTTTGCTGTG
TTCGACGAGGGCAAGTCTTGGCACTCTGAAACAAAGAACTCCCTGATGCAGGACAGAGATGCTGCA
TCTGCAAGGGCTTGGCCAAAAATGCACACAGTGAACGGCTATGTGAATCGATCACTGCCAGGACTG
ATAGGCTGTCATCGCAAGTCAGTGTATTGGCACGTTATCGGGATGGGAACAACTCCAGAAGTGCAC
AGCATCTTCCTTGAGGGCCACACTTTCCTGGTTCGGAATCATAGACAGGCCAGCCTTGAGATCAGC
CCAATCACCTTTCTGACTGCCCAAACCTTGCTGATGGATCTGGGACAGTTCCTCCTGTTTTGTCAC
ATCTCCTCCCACCAACATGACGGGATGGAGGCTTATGTGAAGGTCGATAGCTGTCCGGAGGAACCA
CAACTGAGGATGAAGAACAACGAAGAGGCAGAGGACTATGACGACGATCTGACTGACAGTGAAATG
GACGTGGTTCGGTTCGACGATGACAATTCTCCTTCATTTATCCAGATCCGTTCCGTGGCCAAGAAG
CACCCCAAGACTTGGGTTCATTACATCGCTGCTGAGGAGGAGGATTGGGACTACGCGCCCTTGGTG
TTGGCCCCAGACGATCGCTCATACAAGAGCCAGTACCTTAACAATGGTCCACAAAGGATCGGCCGG
AAGTACAAGAAGGTTAGATTTATGGCTTATACCGACGAGACTTTTAAAACTAGGGAAGCAATTCAG
CATGAAAGTGGCATTCTTGGACCCCTGCTGTATGGCGAGGTTGGCGACACCCTGCTGATTATCTTT
AAGAACCAGGCAAGCCGGCCCTACAACATCTACCCGCACGGCATAACCGATGTACGACCCCTGTAC
AGTCGCAGACTTCCTAAAGGGGTGAAACACCTGAAGGACTTCCCAATTCTGCCCGGGGAGATCTTC
AAGTATAAATGGACCGTGACGGTTGAGGATGGTCCCACAAAGTCCGATCCGAGATGCCTTACCCGA
TATTATTCCAGCTTCGTGAACATGGAAAGGGACCTGGCCAGCGGGCTGATTGGCCCACTGCTGATT
TGTTACAAGGAGTCTGTCGATCAAAGAGGAAACCAAATAATGAGCGACAAACGTAACGTCATCCTG
TTCAGCGTCTTTGATGAGAATAGAAGCTGGTACCTCACAGAAAATATTCAGCGGTTTCTGCCTAAC
CCCGCAGGCGTCCAGCTGGAAGATCCCGAGTTCCAAGCCTCAAACATCATGCATAGCATCAACGGA
TACGTATTCGATAGCCTGCAGCTGTCCGTCTGTCTCCATGAAGTGGCATATTGGTACATCCTGAGT
ATCGGGGCGCAGACCGACTTCCTGAGCGTGTTCTTTTCTGGATACACGTTCAAACACAAAATGGTC
TATGAAGATACCCTGACTCTGTTTCCATTCTCAGGAGAGACAGTCTTTATGAGTATGGAAAATCCT
GGACTGTGGATCCTGGGCTGTCACAATTCTGATTTTCGGAACAGAGGCATGACAGCCCTGCTTAAA
GTGAGCTCATGCGACAAGAACACCGGTGATTACTACGAAGATAGCTATGAGGACATCAGTGCGTAT
TTGCTCTCCAAGAACAACGCTATCGAGCCACGGTCTTTCAGTCAGAATCCTCCCGTTCTGAAGCGG
CATCAGCGCGAAATAACACGCACAACCCTTCAGTCAGACCAAGAGGAAATCGACTACGATGATACT
ATCTCTGTGGAGATGAAGAAGGAGGATTTCGACATTTACGACGAGGACGAGAATCAGTCCCCAAGG
AGCTTTCAGAAGAAAACAAGACACTATTTCATTGCCGCCGTGGAGCGACTGTGGGACTACGGCATG

（续）

图18A

TCTAGCTCTCCGCATGTACTTAGAAATAGGGCACAAAGCGGATCCGTGCCTCAGTTTAAGAAAGTT
GTCTTTCAGGAGTTTACAGATGGCTCCTTCACCCAGCCCTTGTATCGCGGGGAACTCAATGAACAC
CTGGGCCTCCTGGGTCCTTATATTAGGGCCGAAGTCGAGGACAATATCATGGTGACCTTTAGGAAC
CAGGCATCTAGACCTTACTCTTTCTACTCCTCCCTGATATCCTATGAGGAGGACCAGCGGCAAGGC
GCTGAGCCTCGGAAGAACTTTGTGAAGCCAAATGAAACCAAAACATACTTTTGGAAAGTTCAGCAC
CACATGGCTCCCACGAAGGACGAATTTGACTGTAAAGCCTGGGCCTACTTCTCAGATGTAGATCTC
GAGAAAGACGTGCACTCAGGGCTCATTGGTCCCCTCCTGGTCTGTCATACTAATACCCTCAATCCA
GCACACGGACGTCAGGTAACCGTCCAGGAATTTGCCCTGTTCTTTACCATTTTCGATGAGACTAAA
TCCTGGTACTTTACCGAAAACATGGAGAGGAATTGCAGAGCCCCATGCAACATCCAGATGGAGGAC
CCTACCTTCAAAGAGAACTATCGCTTCCATGCCATTAACGGTTACATTATGGATACTCTCCCAGGA
CTTGTGATGGCACAGGATCAGCGGATAAGATGGTATCTGTTGAGCATGGGCTCCAACGAGAATATT
CACAGCATCCATTTCTCCGGTCACGTGTTTACAGTGAGAAAGAAAGAAGAGTACAAGATGGCTCTG
TATAATCTCTATCCAGGCGTATTCGAAACGGTGGAGATGTTGCCTAGCAAGGCCGGCATTTGGCGA
GTAGAATGCCTTATCGGGGAACATCTGCATGCCGGAATGAGCACGCTCTTCCTGGTGTATAGTAAC
AAGTGCCAGACTCCGCTGGGCATGGCATCTGGCCATATACGGGACTTTCAGATTACGGCTAGCGGG
CAGTATGGGCAGTGGGCACCCAAACTTGCGCGACTGCACTATTCAGGCTCTATCAATGCATGGTCC
ACCAAGGAACCCTTCTCTTGGATTAAGGTGGACCTTTTGGCGCCCATGATAATCCATGGGATCAAA
ACCCAGGGCGCTCGTCAGAAATTCTCATCACTCTACATCTCTCAGTTCATAATAATGTATTCACTG
GATGGGAAGAAATGGCAGACTTACAGAGGAAACAGCACCGGGACGCTGATGGTGTTCTTTGGCAAC
GTGGACAGCAGCGGCATCAAACACAACATCTTCAATCCTCCCATTATTGCCCGTTATATTAGACTG
CATCCCACTCACTACTCTATACGCAGCACACTTAGGATGGAGCTCATGGGATGCGACCTGAACAGT
TGTAGTATGCCCTTGGGGATGGAGTCCAAAGCTATAAGCGACGCACAAATTACAGCTAGCTCTTAC
TTTACGAATATGTTCGCCACGTGGAGCCCAAGCAAAGCCCGGCTGCATTTGCAGGGTCGGAGTAAT
GCTTGGCGCCCACAGGTGAATAACCCTAAGGAATGGTTGCAAGTAGATTTCCAGAAAACTATGAAG
GTAACCGGCGTCACTACACAGGGAGTCAAGTCCCTCTTGACCTCTATGTACGTCAAGGAGTTCCTG
ATTAGCAGCAGTCAGGATGGGCACCAATGGACACTGTTCTTCCAGAATGGGAAAGTTAAAGTATTT
CAGGGTAACCAGGACTCCTTTACACCTGTGGTGAATAGCCTCGACCCACCCCTGCTGACACGATAC
CTCCGCATCCACCCTCAGTCTTGGGTGCATCAAATTGCCCTGCGAATGGAGGTGTTGGGATGCGAA
GCTCAGGACCTCTACTGA (SEQ ID NO:15)

图18B

CS11-FL-NA

ATGCAGATCGAACTCTCTACTTGCTTCTTCCTGTGCCTTCTGAGGTTCTGCTTCTCTGCCACTCGC
CGATATTACCTCGGGGCCGTGGAGTTGAGTTGGGACTACATGCAATCAGATCTGGGCGAACTCCCT
GTGGATGCCCGATTCCCACCGCGCGTGCCCAAGTCTTTCCCATTTAATACTTCTGTGGTGTACAAG
AAGACATTGTTTGTGGAGTTTACCGATCACCTGTTCAACATCGCCAAACCGCGGCCCCCATGGATG
GGTCTGCTTGGGCCCACCATTCAAGCGGAGGTCTATGATACAGTGGTGATAACGCTTAAGAACATG
GCGAGCCACCCAGTGTCTCTGCATGCCGTTGGTGTATCATATTGGAAGGCCAGCGAAGGAGCGGAG
TACGATGACCAGACCTCTCAGAGAGAGAAGGAAGACGATAAGGTTTTTCCTGGCGGAAGTCATACA
TATGTATGGCAGGTCCTGAAAGAGAATGGGCCGATGGCTTCTGACCCCCTTTGTCTTACCTATAGT
TATCTGAGCCACGTGGACCTGGTCAAGGACCTCAACAGTGGTCTGATTGGGGCTCTGCTTGTTTGT
AGAGAGGGTAGCTTGGCTAAGGAGAAAACCCAAACACTCCATAAGTTCATTTTGCTGTTCGCGGTG
TTCGACGAGGGAAAGAGTTGGCACAGCGAAACAAAGAATTCACTGATGCAAGACAGGGACGCCGCT
TCCGCAAGGGCTTGGCCTAAGATGCATACGGTGAATGGGTATGTGAACCGGAGCCTCCCGGGGCTG
ATCGGGTGCCATCGCAAGTCTGTTTACTGGCACGTCATTGGAATGGGGACAACGCCAGAGGTACAT
AGTATATTTCTTGAAGGCCACACGTTCCTCGTACGGAACCACCGACAGGCTTCCCTGGAGATAAGC
CCCATTACCTTTCTGACCGCTCAGACTCTGCTGATGGACCTTGGCCAGTTTCTCCTGTTCTGCCAT
ATTAGCAGCCACCAGCACGACGGTATGGAAGCATACGTGAAAGTCGATAGCTGTCCTGAGGAGCCT
CAGCTCAGAATGAAGAACAACGAGGAGGCCGAAGACTATGACGATGACCTTACAGATTCCGAGATG
GACGTGGTGCGCTTTGACGACGATAACAGTCCTAGTTTCATTCAAATCAGATCCGTAGCCAAAAAG
CATCCAAAGACATGGGTGCATTACATTGCAGCCGAAGAGGAGGATTGGGATTATGCGCCCCTTGTT
CTGGCTCCAGATGACAGGAGCTATAAGTCCCAGTACTTGAACAACGGGCCACAGCGAATCGGTAGA
AAATATAAGAAGGTAAGATTCATGGCCTACACTGACGAAACATTTAAAACCAGGGAAGCTATCCAA
CACGAATCTGGAATTCTCGGCCCTCTGCTCTACGGTGAGGTGGGGGACACCTTGCTGATCATTTTC
AAAAATCAGGCATCCAGGCCTTACAACATATACCCCCATGGCATCACCGATGTCCGCCCGCTGTAT
TCCAGAAGACTCCCCAAGGGAGTGAAACATCTGAAAGATTTTCCCATCCTGCCGGGCGAGATCTTT
AAATACAAATGGACTGTGACTGTAGAGGACGGGCCTACAAAATCAGACCCACGGTGCCTGACAAGG
TATTACAGTAGCTTCGTCAACATGGAACGCGACCTCGCCAGCGGACTCATTGGCCCACTGTTGATC
TGTTACAAAGAGTCAGTGGATCAGAGGGGAAATCAGATCATGAGCGATAAGAGAAACGTTATCCTG
TTTAGTGTCTTCGACGAGAACCGGTCTTGGTACCTTACTGAGAACATCCAGAGGTTCCTGCCGAAT
CCGGCTGGCGTTCAGCTCGAGGACCCAGAGTTCCAGGCCAGTAATATAATGCACTCAATCAACGGT
TATGTGTTCGATAGCCTGCAGCTGAGCGTCTGCCTCCACGAGGTAGCCTATTGGTACATATTGTCC
ATCGGGGCTCAGACCGATTTTCTGTCCGTGTTCTTTAGCGGGTATACCTTTAAACATAAAATGGTC
TATGAAGACACCCTGACCCTGTTCCCATTCTCCGGTGAGACTGTGTTCATGTCCATGGAGAACCCA
GGGCTGTGGATCCTGGGGTGTCACAATAGTGACTTTAGGAATCGGGGAATGACGGCACTGCTGAAG
GTGAGTTCTTGCGATAAAAATACAGGAGATTACTATGAGGATAGTTACGAGGATATCAGTGCCTAT
CTGCTTTCAAAAAACAACGCAATTGAGCCCCGGTCTTTCTCACAAAACCCCCCGGTGCTGAAGCGC
CACCAGCGCGAAATTACCCGGACAACCTTGCAGTCCGACCAGGAGGAAATCGATTATGACGATACT
ATCAGTGTAGAAATGAAAAAGGAGGATTTTGATATTTACGACGAAGACGAGAACCAGTCTCCGCGA

（续）

图19A

AGTTTTCAGAAGAAAACGCGACACTACTTTATAGCTGCCGTGGAACGACTCTGGGATTATGGCATG
TCCTCCAGCCCTCATGTCCTTAGGAATCGAGCGCAGAGTGGCTCTGTGCCTCAGTTCAAAAAGGTT
GTGTTCCAGGAATTCACCGACGGCTCATTTACCCAGCCGCTGTACAGAGGCGAACTCAACGAACAC
CTTGGGCTGCTTGGGCCATATATTCGAGCAGAGGTGGAAGATAATATCATGGTAACCTTTAGAAAC
CAGGCGTCAAGACCCTATTCCTTCTACAGTTCTCTGATCAGCTACGAGGAGGACCAAAGACAGGGA
GCTGAACCCAGGAAGAACTTTGTGAAACCTAATGAGACCAAGACCTACTTCTGGAAGGTCCAGCAC
CATATGGCCCCAACTAAAGATGAATTCGATTGCAAGGCCTGGGCTTATTTCAGCGACGTGGATCTC
GAAAAGGATGTGCACAGCGGGTTGATCGGACCGCTTTTGGTGTGCCACACAAATACCCTCAATCCT
GCCCACGGGCGGCAGGTCACAGTTCAAGAGTTTGCACTCTTCTTTACAATATTTGACGAGACAAAG
TCATGGTATTTTACAGAGAATATGGAGAGAAATTGTCGCGCACCTTGCAACATTCAGATGGAGGAC
CCCACATTTAAGGAGAATTACAGATTTCATGCTATCAATGGGTACATTATGGATACTCTGCCTGGT
CTGGTCATGGCCCAGGATCAGCGCATAAGGTGGTACTTGCTGAGCATGGGATCTAATGAGAATATA
CACAGCATTCACTTCAGTGGCCACGTTTTTACTGTTAGAAAGAAGGAGGAGTACAAAATGGCGCTC
TACAACCTTTACCCGGGTGTGTTTGAGACAGTGGAGATGCTGCCAAGCAAGGCAGGCATCTGGAGG
GTTGAGTGTCTTATTGGGGAGCATCTGCATGCTGGAATGTCCACCCTCTTTCTTGTGTACAGCAAT
AAGTGCCAGACACCGCTTGGCATGGCCAGCGGCCACATTAGGGACTTTCAGATAACTGCCAGTGGA
CAGTACGGCCAGTGGGCTCCCAAGCTTGCAAGACTCCACTACTCCGGAAGCATAAACGCATGGAGC
ACCAAGGAACCCTTCTCTTGGATTAAGGTGGACCTGCTGGCGCCAATGATCATTCACGGCATAAAA
ACCCAAGGGGCACGACAGAAATTTTCATCTTTGTATATTAGTCAGTTTATCATCATGTACAGCTTG
GATGGAAAGAAGTGGCAGACGTACAGGGGCAATTCTACAGGAACACTTATGGTGTTTTTTGGGAAT
GTCGATTCCAGCGGGATCAAACATAACATCTTCAATCCTCCTATTATCGCCCGATATATCCGCCTG
CACCCTACGCATTACTCCATCAGGTCCACATTGAGAATGGAACTGATGGGGTGCGACCTGAATAGT
TGTAGTATGCCACTGGGCATGGAGTCTAAAGCCATCAGCGATGCACAGATCACTGCCAGCTCTTAC
TTCACCAACATGTTTGCAACTTGGTCCCCCTCTAAAGCTCGCCTGCATCTGCAGGGACGCTCAAAT
GCATGGCGACCACAGGTGAACAATCCAAAAGAGTGGCTCCAGGTCGACTTTCAGAAGACAATGAAG
GTAACAGGAGTGACAACCCAGGGTGTAAAAAGCCTCCTTACGAGTATGTACGTTAAGGAGTTTCTG
ATTTCTAGCTCCCAGGACGGACACCAGTGGACTCTGTTCTTCCAGAACGGCAAAGTGAAGGTATTT
CAGGGAAACCAGGATTCTTTTACCCCGGTAGTGAATAGCCTGGATCCACCGTTGCTGACCCGCTAT
CTGAGAATTCATCCACAATCCTGGGTGCATCAGATTGCCCTCCGGATGGAAGTGCTCGGCTGTGAA
GCTCAGGATCTGTATTAG (SEQ ID NO:16)

图19B

CS40-FL-NA

ATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGCGATTCTGCTTTAGTGCCACCAGA
AGATACTACCTGGGTGCAGTGGAACTGTCATGGGACTATATGCAAAGTGATCTCGGTGAGCTGCCT
GTGGACGCAAGATTTCCTCCTAGAGTGCCAAAATCTTTTCCATTCAACACCTCAGTCGTGTACAAA
AAGACTCTGTTTGTAGAATTCACGGATCACCTTTTCAACATCGCTAAGCCAAGGCCACCCTGGATG
GGTCTGCTAGGTCCTACCATCCAGGCTGAGGTTTATGATACAGTGGTCATTACACTTAAGAACATG
GCTTCCCATCCTGTCAGTCTTCATGCTGTTGGTGTATCCTACTGGAAAGCTTCTGAGGGAGCTGAA
TATGATGATCAGACCAGTCAAAGGGAGAAAGAAGATGATAAAGTCTTCCCTGGTGGAAGCCATACA
TATGTCTGGCAGGTCCTGAAAGAGAATGGTCCAATGGCCTCTGACCCACTGTGCCTTACCTACTCA
TATCTTTCTCATGTGGACCTGGTAAAAGACTTGAATTCAGGCCTCATTGGAGCCCTACTAGTATGT
AGAGAAGGGAGTCTGGCCAAGGAAAAGACACAGACCTTGCACAAATTTATACTACTTTTTGCTGTA
TTTGATGAAGGGAAAAGTTGGCACTCAGAAACAAAGAACTCCTTGATGCAGGATAGGGATGCTGCA
TCTGCTCGGGCCTGGCCTAAAATGCACACAGTCAATGGTTATGTAAACAGGTCTCTGCCAGGTCTG
ATTGGATGCCACAGGAAATCAGTCTATTGGCATGTGATTGGAATGGGCACCACTCCTGAAGTGCAC
TCAATATTCCTCGAAGGTCACACATTTCTTGTGAGGAACCATCGCCAGGCGTCCTTGGAAATCTCG
CCAATAACTTTCCTTACTGCTCAAACACTCTTGATGGACCTTGGACAGTTTCTACTGTTTTGTCAT
ATCTCTTCCCACCAACATGATGGCATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCC
CAACTACGAATGAAAAATAATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTCTGAAATG
GATGTGGTCAGGTTTGATGATGACAACTCTCCTTCCTTTATCCAAATTCGCTCAGTTGCCAAGAAG
CATCCTAAAACTTGGGTACATTACATTGCTGCTGAAGAGGAGGACTGGGACTATGCTCCCTTAGTC
CTCGCCCCCGATGACAGAAGTTATAAAAGTCAATATTTGAACAATGGCCCTCAGCGGATTGGTAGG
AAGTACAAAAAAGTCCGATTTATGGCATACACAGATGAAACCTTTAAGACTCGTGAAGCTATTCAG
CATGAATCAGGAATCTTGGGACCTTTACTTTATGGGGAAGTTGGAGACACACTGTTGATTATATTT
AAGAATCAAGCAAGCAGACCATATAACATCTACCCTCACGGAATCACTGATGTCCGTCCTTTGTAT
TCAAGGAGATTACCAAAAGGTGTAAAACATTTGAAGGATTTTCCAATTCTGCCAGGAGAAATATTC
AAATATAAATGGACAGTGACTGTAGAAGATGGGCCAACTAAATCAGATCCTCGGTGCCTGACCCGC
TATTACTCTAGTTTCGTTAATATGGAGAGAGATCTAGCTTCAGGACTCATTGGCCCTCTCCTCATC
TGCTACAAAGAATCTGTAGATCAAAGAGGAAACCAGATAATGTCAGACAAGAGGAATGTCATCCTG
TTTTCTGTATTTGATGAGAACCGAAGCTGGTACCTCACAGAGAATATACAACGCTTTCTCCCCAAT
CCAGCTGGAGTGCAGCTTGAGGATCCAGAGTTCCAAGCCTCCAACATCATGCACAGCATCAATGGC
TATGTTTTTGATAGTTTGCAGTTGTCAGTTTGTTTGCATGAGGTGGCATACTGGTACATTCTAAGC
ATTGGAGCACAGACTGACTTCCTTTCTGTCTTCTTCTCTGGATATACCTTCAAACACAAAATGGTC
TATGAAGACACACTCACCCTATTCCCATTCTCAGGAGAAACTGTCTTCATGTCGATGGAAAACCCA
GGTCTATGGATTCTGGGGTGCCACAACTCAGACTTTCGGAACAGAGGCATGACCGCCTTACTGAAG
GTTTCTAGTTGTGACAAGAACACTGGTGATTATTACGAGGACAGTTATGAAGATATTTCAGCATAC
TTGCTGAGTAAAAACAATGCCATTGAACCAAGAAGCTTCTCCCAGAATCCACCAGTCTTGAAACGC
CATCAACGGGAAATAACTCGTACTACTCTTCAGTCAGATCAAGAGGAAATTGACTATGATGATACC
ATATCAGTTGAAATGAAGAAGGAAGATTTTGACATTTATGATGAGGATGAAAATCAGAGCCCCCGC
AGCTTTCAAAAGAAAACACGACACTATTTTATTGCTGCAGTGGAGAGGCTCTGGGATTATGGGATG

（续）

图20A

AGTAGCTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTT
GTTTTCCAGGAATTTACTGATGGCTCCTTTACTCAGCCCTTATACCGTGGAGAACTAAATGAACAT
TTGGGACTCCTGGGGCCATATATAAGAGCAGAAGTTGAAGATAATATCATGGTAACTTTCAGAAAT
CAGGCCTCTCGTCCCTATTCCTTCTATTCTAGCCTTATTTCTTATGAGGAAGATCAGAGGCAAGGA
GCAGAACCTAGAAAAAACTTTGTCAAGCCTAATGAAACCAAAACTTACTTTTGGAAAGTGCAACAT
CATATGGCACCCACTAAAGATGAGTTTGACTGCAAAGCCTGGGCTTATTTCTCTGATGTTGACCTG
GAAAAAGATGTGCACTCAGGCCTGATTGGACCCCTTCTGGTCTGCCACACTAACACACTGAACCCT
GCTCATGGGAGACAAGTGACAGTACAGGAATTTGCTCTGTTTTTCACCATCTTTGATGAGACCAAA
AGCTGGTACTTCACTGAAAATATGGAAAGAAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGAT
CCCACTTTTAAAGAGAATTATCGCTTCCATGCAATCAATGGCTACATAATGGATACACTACCTGGC
TTAGTAATGGCTCAGGATCAAAGGATTCGATGGTATCTGCTCAGCATGGGCAGCAATGAAAACATC
CATTCTATTCATTTCAGTGGACATGTGTTCACTGTACGAAAAAAAGAGGAGTATAAAATGGCACTG
TACAATCTCTATCCAGGTGTTTTTGAGACAGTGGAAATGTTACCATCCAAAGCTGGAATTTGGCGG
GTGGAATGCCTTATTGGCGAGCATCTACATGCTGGGATGAGCACACTTTTTCTGGTGTACAGCAAT
AAGTGTCAGACTCCCCTGGGAATGGCTTCTGGACACATTAGAGATTTTCAGATTACAGCTTCAGGA
CAATATGGACAGTGGGCCCCAAAGCTGGCCAGACTTCATTATTCCGGATCAATCAATGCCTGGAGC
ACCAAGGAGCCCTTTTCTTGGATCAAGGTGGATCTGTTGGCACCAATGATTATTCACGGCATCAAG
ACCCAGGGTGCCCGTCAGAAGTTCTCCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTT
GATGGGAAGAAGTGGCAGACTTATCGAGGAAATTCCACTGGAACCTTAATGGTCTTCTTTGGCAAT
GTGGATTCATCTGGGATAAAACACAATATTTTTAACCCTCCAATTATTGCTCGATACATCCGTTTG
CACCCAACTCATTATAGCATTCGCAGCACTCTTCGCATGGAGTTGATGGGCTGTGATTTAAATAGT
TGCAGCATGCCATTGGGAATGGAGAGTAAAGCAATATCAGATGCACAGATTACTGCTTCATCCTAC
TTTACCAATATGTTTGCCACCTGGTCTCCTTCAAAAGCTCGACTTCACCTCCAAGGGAGGAGTAAT
GCCTGGAGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGGACTTCCAGAAGACAATGAAA
GTCACAGGAGTAACTACTCAGGGAGTAAAATCTCTGCTTACCAGCATGTATGTGAAGGAGTTCCTC
ATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTCTTTTTTCAGAATGGCAAAGTAAAGGTTTTT
CAGGGAAATCAAGACTCCTTCACACCTGTGGTGAACTCTCTAGACCCACCGTTACTGACTCGCTAC
CTTCGAATTCACCCCCAGAGTTGGGTGCACCAGATTGCCCTGAGGATGGAGGTTCTGGGCTGCGAG
GCACAGGACCTCTACTGA (SEQ ID NO:17)

图20B

CH25-FL-NA

ATGCAGATCGAGCTGTCCACATGCTTTTTTCTGTGCCTGCTGCGGTTCTGCTTCAGCGCCACCCGG
CGGTACTACCTGGGCGCCGTGGAGCTGTCCTGGGACTACATGCAGAGCGACCTGGGCGAGCTGCCC
GTGGACGCCCGGTTCCCCCCCAGAGTGCCCAAGAGCTTCCCCTTCAACACCAGCGTGGTGTACAAG
AAAACCCTGTTCGTGGAGTTCACCGACCACCTGTTCAACATCGCCAAGCCCAGGCCCCCCTGGATG
GGCCTGCTGGGCCCCACCATCCAGGCCGAGGTGTACGACACCGTGGTGATCACCCTGAAGAACATG
GCCAGCCACCCCGTGAGCCTGCACGCCGTGGGCGTGAGCTACTGGAAGGCCTCCGAGGGCGCCGAG
TACGACGACCAGACCAGCCAGCGGGAGAAAGAGGACGACAAAGTCTTTCCTGGCGGCAGCCACACC
TACGTGTGGCAGGTCCTGAAAGAAAACGGCCCCATGGCCTCCGACCCCCTGTGCCTGACCTACAGC
TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACAGCGGGCTGATTGGGGCCCTGCTGGTCTGC
CGGGAGGGCAGCCTGGCCAAAGAGAAAACCCAGACCCTGCACAAGTTCATCCTGCTGTTCGCCGTG
TTCGACGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAGGACCGGGACGCCGCC
TCTGCCAGAGCCTGGCCCAAGATGCACACCGTGAACGGCTACGTGAACAGAAGCCTGCCCGGCCTG
ATTGGCTGCCACCGGAAGAGCGTGTACTGGCACGTGATCGGCATGGGCACCACACCCGAGGTGCAC
AGCATCTTTCTGGAAGGGCACACCTTTCTGGTGCGGAACCACCGGCAGGCCAGCCTGGAAATCAGC
CCTATCACCTTCCTGACCGCCCAGACACTGCTGATGGACCTGGGCCAGTTCCTGCTGTTTTGCCAC
ATCAGCTCTCACCAGCACGACGGCATGGAAGCCTACGTGAAGGTGGACTCCTGCCCCGAGGAACCC
CAGCTGCGGATGAAGAACAACGAGGAAGCCGAGGACTACGACGACGACCTGACCGACAGCGAGATG
GACGTGGTGCGGTTCGACGACGACAACAGCCCCAGCTTCATCCAGATCAGAAGCGTGGCCAAGAAG
CACCCCAAGACCTGGGTGCACTACATCGCCGCCGAGGAAGAGGACTGGGACTACGCCCCCTGGTG
CTGGCCCCCGACGACAGAAGCTACAAGAGCCAGTACCTGAACAATGGCCCCCAGCGGATCGGCCGG
AAGTACAAGAAAGTGCGGTTCATGGCCTACACCGACGAGACCTTCAAGACCCGGGAGGCCATCCAG
CACGAGAGCGGCATCCTGGGCCCCCTGCTGTACGGCGAAGTGGGCGACACACTGCTGATCATCTTC
AAGAACCAGGCCAGCCGGCCCTACAACATCTACCCCCACGGCATCACCGACGTGCGGCCCCTGTAC
AGCAGGCGGCTGCCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCCTGCCCGGCGAGATCTTC
AAGTACAAGTGGACCGTGACCGTGGAGGACGGCCCCACCAAGAGCGACCCCAGATGCCTGACCCGG
TACTACAGCAGCTTCGTGAACATGGAACGGGACCTGGCCTCCGGGCTGATCGGACCTCTGCTGATC
TGCTACAAAGAAAGCGTGGACCAGCGGGGCAACCAGATCATGAGCGACAAGCGGAACGTGATCCTG
TTCAGCGTGTTCGATGAGAACCGGTCCTGGTATCTGACCGAGAACATCCAGCGGTTTCTGCCCAAC
CCTGCCGGGGTGCAGCTGGAAGATCCCGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TACGTGTTCGACAGCCTGCAGCTGTCCGTGTGTCTGCACGAGGTGGCCTACTGGTACATCCTGAGC
ATCGGCGCCCAGACCGACTTCCTGAGCGTGTTCTTCAGCGGCTACACCTTCAAGCACAAGATGGTG
TACGAGGACACCCTGACCCTGTTCCCTTTCAGCGGCGAGACCGTGTTCATGAGCATGGAAAACCCC
GGCCTGTGGATCCTGGGCTGCCACAACAGCGACTTCCGGAACCGGGGCATGACCGCCCTGCTGAAG
GTGTCCAGCTGCGACAAGAACACCGGCGACTACTACGAGGACAGCTACGAGGATATCAGCGCCTAC
CTGCTGTCCAAGAACAACGCCATCGAGCCCAGAAGCTTCAGCCAGAACCCCCCTGTGCTGAAGCGG
CACCAGAGAGAGATCACCCGGACCACCCTGCAGTCCGACCAGGAAGAGATCGATTACGACGACACC

（续）

图21A

ATCAGCGTGGAGATGAAAAAAGAAGATTTCGACATCTACGACGAGGACGAGAACCAGAGCCCCCGG
TCCTTCCAGAAGAAAACCCGGCACTACTTTATCGCCGCCGTGGAGCGGCTGTGGGACTACGGCATG
AGCAGCAGCCCCCACGTGCTGCGGAACCGGGCCCAGAGCGGCAGCGTGCCCCAGTTCAAGAAAGTG
GTGTTCCAGGAATTCACCGACGGCAGCTTCACCCAGCCCCTGTACCGGGGCGAGCTGAACGAGCAC
CTGGGGCTGCTGGGGCCCTACATCAGGGCCGAAGTGGAGGACAACATCATGGTGACCTTCCGGAAT
CAGGCCAGCAGACCCTACTCCTTCTACAGCAGCCTGATCAGCTACGAAGAGGACCAGCGGCAGGGC
GCTGAACCCCGGAAGAACTTCGTGAAGCCCAATGAGACCAAGACCTACTTCTGGAAAGTGCAGCAC
CACATGGCCCCCACCAAGGACGAGTTCGACTGCAAGGCCTGGGCCTACTTCAGCGACGTGGATCTG
GAAAAGGACGTGCACTCTGGACTGATTGGCCCTCTGCTGGTGTGCCACACCAACACCCTGAACCCC
GCCCACGGCCGGCAGGTGACCGTGCAGGAATTCGCCCTGTTCTTCACCATCTTCGACGAGACCAAG
TCCTGGTACTTCACCGAGAATATGGAACGGAACTGCAGAGCCCCCTGCAACATCCAGATGGAAGAT
CCTACCTTCAAAGAGAACTACCGGTTCCACGCCATCAACGGCTACATCATGGACACCCTGCCTGGC
CTGGTGATGGCCCAGGACCAGAGGATCCGGTGGTATCTGCTGTCCATGGGCAGCAACGAGAATATC
CACAGCATCCACTTCAGCGGCCACGTGTTCACCGTGAGGAAGAAAGAAGAGTACAAGATGGCCCTG
TACAACCTGTACCCCGGCGTGTTCGAGACCGTGGAGATGCTGCCCAGCAAGGCCGGCATCTGGCGG
GTGGAGTGTCTGATCGGCGAGCACCTGCATGCCGGGATGAGCACCCTGTTTCTGGTGTACAGCAAC
AAGTGCCAGACCCCCCTGGGCATGGCCAGCGGCCACATCCGGGACTTCCAGATCACCGCCTCCGGC
CAGTACGGCCAGTGGGCCCCCAAGCTGGCCCGGCTGCACTACAGCGGCAGCATCAACGCCTGGTCC
ACCAAAGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCTATGATCATCCACGGCATTAAG
ACCCAGGGCGCCAGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTG
GACGGCAAGAAGTGGCAGACCTACCGGGGCAACAGCACCGGCACCCTGATGGTGTTCTTCGGCAAC
GTGGACAGCAGCGGCATCAAGCACAACATCTTCAACCCCCCCATCATCGCCCGGTACATCCGGCTG
CACCCCACCCACTACAGCATCAGATCCACCCTGCGGATGGAACTGATGGGCTGCGACCTGAACTCC
TGCAGCATGCCTCTGGGCATGGAAAGCAAGGCCATCAGCGACGCCCAGATCACAGCCAGCAGCTAC
TTCACCAACATGTTCGCCACCTGGTCCCCCTCCAAGGCCAGGCTGCACCTGCAGGGCCGGTCCAAC
GCCTGGCGGCCTCAGGTGAACAACCCCAAAGAATGGCTGCAGGTGGACTTTCAGAAAACCATGAAG
GTGACCGGCGTGACCACCCAGGGCGTGAAAAGCCTGCTGACCAGCATGTACGTGAAAGAGTTTCTG
ATCAGCAGCAGCCAGGACGGCCACCAGTGGACCCTGTTCTTTCAGAACGGCAAGGTGAAAGTGTTC
CAGGGCAACCAGGACTCCTTCACCCCCGTGGTGAACTCCCTGGACCCCCCCCCTGCTGACCCGCTAC
CTGCGGATCCACCCCCAGTCTTGGGTGCACCAGATCGCCCTGAGGATGGAAGTGCTGGGATGTGAG
GCCCAGGATCTGTACTGA (SEQ ID NO:18)

图21B

FVIII-FL-AA

mqielstcff lcllrfcfsa trryylgave lswdymqsdl gelpvdarfp prvpksfpfn
tsvvykktlf veftdhlfni akprppwmgl lgptiqaevy dtvvitlknm ashpvslhav
gvsywkaseg aeyddqtsqr ekeddkvfpg gshtyvwqvl kengpmasdp lcltysylsh
vdlvkdlnsg ligallvcre gslakektqt lhkfillfav fdegkswhse tknslmqdrd
aasarawpkm htvngyvnrs lpgligchrk svywhvigmg ttpevhsifl eghtflvrnh
rqasleispi tfltaqtllm dlgqfllfch isshqhdgme ayvkvdscpe epqlrmknne
eaedydddlt dsemdvvrfd ddnspsfiqi rsvakkhpkt wvhyiaaeee dwdyaplvla
pddrsyksqy lnngpqrigr kykkvrfmay tdetfktrea iqhesgilgp llygevgdtl
liifknqasr pyniyphgit dvrplysrrl pkgvkhlkdf pilpgeifky kwtvtvedgp
tksdprcltr yyssfvnmer dlasgligpl licykesvdq rgnqimsdkr nvilfsvfde
nrswylteni qrflpnpagv qledpefqas nimhsingyv fdslqlsvcl hevaywyils
igaqtdflsv ffsgytfkhk mvyedtltlf pfsgetvfms menpglwilg chnsdfrnrg
mtallkvssc dkntgdyyed syedisayll sknnaieprs fsqnsrhpst rqkqfnatti
pendiektdp wfahrtpmpk iqnvsssdll mllrqsptph glslsdlqea kyetfsddps
pgaidsnnsl semthfrpql hhsgdmvftp esglqlrlne klgttaatel kkldfkvsst
snnlistips dnlaaqtdnt sslgppsmpv hydsqldttl fgkkssplte sggplslsee
nndskllesg lmnsqesswg knvsstesgr lfkgkrahgp alltkdnalf kvsislliktn
ktsnnsatnr kthidgpsll ienspsvwqn ilesdtefkk vtplihdrml mdknatalrl
nhmsnkttss knmemvqqkk egpippdaqn pdmsffkmlf lpesarwiqr thgknslnsg
qgpspkqlvs lgpeksvegq nflseknkvv vgkgeftkdv glkemvfpss rnlfltnldn
lhennthnqe kkiqeeiekk etliqenvvl pqihtvtgtk nfmknlflls trqnvegsyd
gayapvlqdf rslndstnrt kkhtahfskk geeenleglg nqtkqiveky acttrispnt
sqqnfvtqrs kralkqfrlp leetelekri ivddtstqws knmkhltpst ltqidyneke
kgaitqspls dcltrshsip qanrsplpia kvssfpsirp iyltrvlfqd nsshlpaasy
rkkdsgvqes shflqgakkn nlslailtle mtgdqrevgs lgtsatnsvt ykkventvlp
kpdlpktsgk vellpkvhiy qkdlfptets ngspghldlv egsllqgteg aikwneanrp
gkvpflrvat essaktpskl ldplawdnhy gtqipkeewk sqekspekta fkkkdtilsl
nacesnhaia ainegqnkpe ievtwakqgr terlcsqnpp vlkrhqreit rttlqsdqee
idyddtisve mkkedfdiyd edenqsprsf qkktrhyfia averlwdygm sssphvlrnr
aqsgsvpqfk kvvfqeftdg sftqplyrge lnehlgllgp yiraevedni mvtfrnqasr
pysfysslis yeedqrqgae prknfvkpne tktyfwkvqh hmaptkdefd ckawayfsdv
dlekdvhsgl igpllvchtn tlnpahgrqv tvqefalfft ifdetkswyf tenmerncra
pcniqmedpt fkenyrfhai ngyimdtlpg lvmaqdqrir wyllsmgsne nihsihfsgh
vftvrkkeey kmalynlypg vfetvemlps kagiwrvecl igehlhagms tlflvysnkc
qtplgmasgh irdfqitasg qygqwapkla rlhysgsina wstkepfswi kvdllapmii
hgiktqgarq kfsslyisqf iimysldgkk wqtyrgnstg tlmvffgnvd ssgikhnifn
ppiiaryirl hpthysirst lrmelmgcdl nscsmplgme skaisdaqit assyftnmfa
twspskarlh lqgrsnawrp qvnnpkewlq vdfqktmkvt gvttqgvksl ltsmyvkefl
isssqdghqw tlffqngkvk vfqgnqdsft pvvnsldppl ltrylrihpq swvhqialrm
evlgceaqdl y (SEQ ID NO:19)

图22

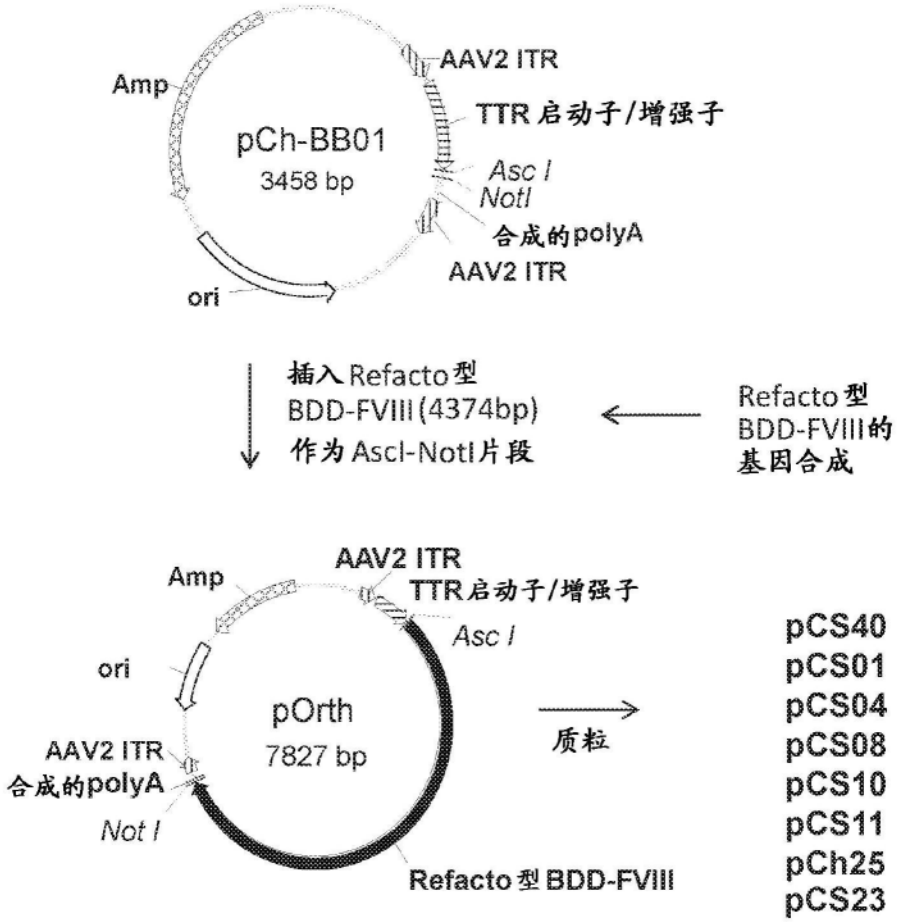

Amp
AAV2 ITR
TTR 启动子/增强子
pCh-BB01
3458 bp
Asc I
NotI
合成的polyA
AAV2 ITR
ori
插入Refacto型
BDD-FVIII(4374bp)
作为AscI-NotI片段
Refacto型
BDD-FVIII的
基因合成
Amp
AAV2 ITR
TTR启动子/增强子
Asc I
ori
pOrth
7827 bp
AAV2 ITR
合成的polyA
Not I
质粒
Refacto型BDD-FVIII
pCS40
pCS01
pCS04
pCS08
pCS10
pCS11
pCh25
pCS23

图23

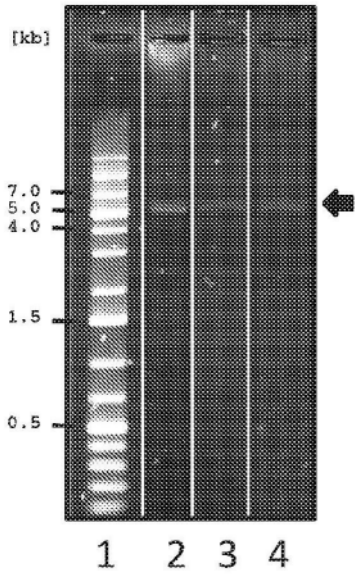

[kb]
7.0
5.0
4.0
1.5
0.5
1
2
3
4

图24

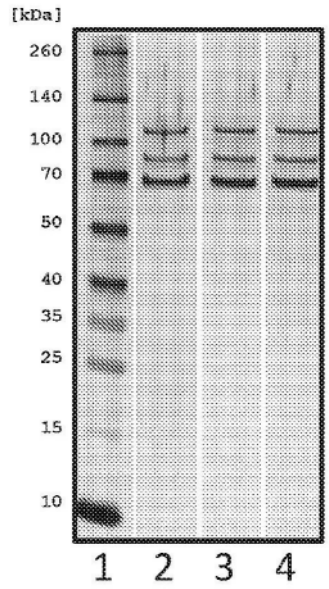


图25

CS23-FL-NA

atgcagattgagctgagcacctgcttcttcctgtgcctgctgaggttctgcttctctgccaccagg
agatactacctgggcgccgtggagctgagctgggactacatgcagtctgacctgggcgagctgcct
gtggacgccaggttcccccccagagtgcccaagagcttcccttcaacacctcagtggtgtacaag
aagaccctgttcgtggagttcaccgaccacctgttcaacatcgccaagcccaggcccccctggatg
ggcctgctgggccccaccatccaggccgaggtgtacgacaccgtggtgatcaccctgaagaacatg
gccagccaccccgtgagcctgcacgccgtgggcgtgagctactggaaggcctctgagggcgccgag
tatgacgaccagaccagccagagggagaaggaggacgacaaggtgttccccggcggcagccacacc
tacgtgtggcaggtgctgaaggagaacggccccatggccagcgacccccctgtgcctgacctacagc
tacctgagccacgtggacctggtgaaggacctgaactctggcctgatcggcgccctgctggtgtgc
agggagggcagcctggccaaggagaagacccagaccctgcacaagttcatcctgctgttcgccgtg
ttcgatgagggcaagagctggcacagcgagaccaagaacagcctgatgcaggacagggatgccgcc
tctgccagggcctggcccaagatgcacaccgtgaacggctacgtgaacaggagcctgcccggcctg
atcggctgccacaggaagtctgtgtactggcacgtgatcggcatgggcaccacccccgaggtgcac
agcatcttcctggagggccacaccttcctggtgaggaaccacaggcaggccagcctggagatcagc
cccatcaccttcctgaccgcccagaccctgctgatggacctgggccagttcctgctgttctgccac
atcagcagccaccagcacgacggcatggaggcctacgtgaaggtggacagctgccccgaggagccc
cagctgaggatgaagaacaacgaggaggccgaggactatgatgatgacctgaccgactctgagatg
gacgtggtgaggtttgatgatgacaacagccccagcttcatccagatcaggtctgtggccaagaag
caccccaagacctgggtgcactacatcgccgccgaggaggaggactgggactacgccccctggtg
ctggcccccgacgacaggagctacaagagccagtacctgaacaacggccccccagaggatcggcagg
aagtacaagaaggtcagattcatggcctacaccgacgagaccttcaagaccagggaggccatccag
cacgagtctggcatcctgggcccccctgctgtacggcgaggtgggcgacaccctgctgatcatcttc
aagaaccaggccagcaggccctacaacatctacccccacggcatcaccgatgtgaggcccctgtac
agcaggaggctgcccaagggcgtgaagcacctgaaggacttccccatcctgcccggcgagatcttc
aagtacaagtggaccgtgaccgtggaggatggccccaccaagtctgaccccaggtgcctgaccagg
tactacagcagcttcgtgaacatggagagggacctggcctctggcctgatcggcccccctgctgatc
tgctacaaggagagcgtggaccagaggggcaaccagatcatgtctgacaagaggaacgtgatcctg
ttctctgtgttcgatgagaacaggagctggtatctgaccgagaacatccagaggttcctgcccaac
cccgccggcgtgcagctggaggaccccgagttccaggccagcaacatcatgcacagcatcaacggc
tacgtgttcgacagcctgcagctgtctgtgtgcctgcacgaggtggcctactggtacatcctgagc
atcggcgcccagaccgacttcctgtctgtgttcttctctggctacaccttcaagcacaagatggtg
tacgaggacaccctgaccctgttcccccttcagcggcgagaccgtgttcatgagcatggagaacccc
ggcctgtggatcctgggctgccacaacagcgacttcaggaacaggggcatgaccgccctgctgaaa
gtcagcagctgcgacaagaacaccggcgactactacgaggacagctacgaggacatcagcgcctac
ctgctgagcaagaacaacgccatcgagcccaggagcttcagccagaacccccccgtgctgaagagg
caccagagggagatcaccaggaccaccctgcagagcgaccaggaggagatcgactatgatgacacc

（续）

图26A

atcagcgtggagatgaagaaggaggacttcgacatctacgacgaggacgagaaccagagccccagg
agcttccagaagaagaccaggcactacttcatcgccgccgtggagaggctgtgggactatggcatg
agcagcagccccacgtgctgaggaacagggcccagagcggcagcgtgccccagttcaagaaggtg
gtgttccaggagttcaccgacggcagcttcacccagcccctgtacagaggcgagctgaacgagcac
ctgggcctgctgggcccctacatcagggccgaggtggaggacaacatcatggtgaccttcaggaac
caggccagcaggccctacagcttctacagcagcctgatcagctacgaggaggaccagaggcagggc
gccgagcccaggaagaacttcgtgaagcccaacgagaccaagacctacttctggaaggtgcagcac
cacatggcccccaccaaggacgagttcgactgcaaggcctgggcctacttctctgatgtggacctg
gagaaggacgtgcacagcggcctgatcggcccctgctggtgtgccacaccaacaccctgaacccc
gcccacggcaggcaggtgaccgtgcaggagttcgccctgttcttcaccatcttcgacgagaccaag
agctggtacttcaccgagaacatggagaggaactgcagggcccctgcaacatccagatggaggac
cccaccttcaaggagaactacaggttccacgccatcaacggctacatcatggacaccctgcccggc
ctggtgatggcccaggaccagaggatcaggtggtatctgctgagcatgggcagcaacgagaacatc
cacagcatccacttcagcggccacgtgttcaccgtgaggaagaaggaggagtacaagatggccctg
tacaacctgtaccccggcgtgttcgagaccgtggagatgctgcccagcaaggccggcatctggagg
gtggagtgcctgatcggcgagcacctgcacgccggcatgagcaccctgttcctggtgtacagcaac
aagtgccagacccccctgggcatggccagcggccacatcagggacttccagatcaccgcctctggc
cagtacggccagtgggcccccaagctggccaggctgcactacagcggcagcatcaacgcctggagc
accaaggagcccttcagctggatcaaggtggacctgctggcccccatgatcatccacggcatcaag
acccagggcgccaggcagaagttcagcagcctgtacatcagccagttcatcatcatgtacagcctg
gacggcaagaagtggcagacctacaggggcaacagcaccggcaccctgatggtgttcttcggcaac
gtggacagcagcggcatcaagcacaacatcttcaacccccccatcatcgccaggtacatcaggctg
caccccacccactacagcatcaggagcaccctgcggatggaactgatgggctgcgacctgaacagc
tgcagcatgcccctgggcatggagagcaaggccatctctgacgcccagatcaccgccagcagctac
ttcaccaacatgttcgccacctggagccccagcaaggccaggctgcacctgcagggcaggagcaac
gcctggaggccccaggtgaacaaccccaaggagtggctgcaggtggacttccagaagaccatgaag
gtgaccggcgtgaccacccagggcgtgaagagcctgctgaccagcatgtacgtgaaggagttcctg
atcagcagcagccaggacggccaccagtggaccctgttcttccagaacggcaaagtgaaggtgttc
cagggcaaccaggacagcttcacCcccgtggtgaacagcctggacccccccctgctgaccaggtat
ctgaggatccacccccagagctgggtgcaccagatcgccctgagaatggaagtgctgggatgcgag
gcccaggacctgtactga (SEQ ID NO:20)

图26B

CS23-FL-AA

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDAR'FPPRVPKSFPFNTSVVYK
KTLFVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEY
DDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCRE
GSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGC
HRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSH
QHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKTW
VHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGIL
GPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVT
VEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENR
SWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLS
VFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGD
YYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFD
IYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQ
PLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNET
KTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALF
FTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLS
MGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTL
FLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMI
IHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIAR
YIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQG
RSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVK
VFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY (SEQ ID NO:21)

图27

CS23-HC-NA

gcc
accaggagat actacctggg cgccgtggag ctgagctggg actacatgca gtctgacctg
ggcgagctgc ctgtggacgc caggttcccc cccagagtgc ccaagagctt ccccttcaac
acctcagtgg tgtacaagaa gaccctgttc gtggagttca ccgaccacct gttcaacatc
gccaagccca ggcccccctg gatgggcctg ctgggcccca ccatccaggc cgaggtgtac
gacaccgtgg tgatcaccct gaagaacatg gccagccacc ccgtgagcct gcacgccgtg
ggcgtgagct actggaaggc ctctgagggc gccgagtatg acgaccagac cagccagagg
gagaaggagg acgacaaggt gttccccggc ggcagccaca cctacgtgtg gcaggtgctg
aaggagaacg gccccatggc cagcgacccc ctgtgcctga cctacagcta cctgagccac
gtggacctgg tgaaggacct gaactctggc ctgatcggcg ccctgctggt gtgcagggag
ggcagcctgg ccaaggagaa gacccagacc ctgcacaagt tcatcctgct gttcgccgtg
ttcgatgagg gcaagagctg gcacagcgag accaagaaca gcctgatgca ggacagggat
gccgcctctg ccagggcctg gcccaagatg cacaccgtga acggctacgt gaacaggagc
ctgcccggcc tgatcggctg ccacaggaag tctgtgtact ggcacgtgat cggcatgggc
accacccccg aggtgcacag catcttcctg gagggccaca ccttcctggt gaggaaccac
aggcaggcca gcctggagat cagccccatc accttcctga ccgcccagac cctgctgatg
gacctgggcc agttcctgct gttctgccac atcagcagcc accagcacga cggcatggag
gcctacgtga aggtggacag ctgccccgag gagccccagc tgaggatgaa gaacaacgag
gaggccgagg actatgatga tgacctgacc gactctgaga tggacgtggt gaggtttgat
gatgacaaca gccccagctt catccagatc aggtctgtgg ccaagaagca ccccaagacc
tgggtgcact acatcgccgc cgaggaggag gactgggact acgccccccc ggtgctggcc
cccgacgaca ggagctacaa gagccagtac ctgaacaacg gcccccagag gatcggcagg
aagtacaaga aggtcagatt catggcctac accgacgaga ccttcaagac cagggaggcc
atccagcacg agtctggcat cctgggcccc ctgctgtacg gcgaggtggg cgacaccctg
ctgatcatct tcaagaacca ggccagcagg ccctacaaca tctaccccca cggcatcacc
gatgtgaggc ccctgtacag caggaggctg cccaagggcg tgaagcacct gaaggacttc
cccatcctgc ccggcgagat cttcaagtac aagtggaccg tgaccgtgga ggatggcccc
accaagtctg accccaggtg cctgaccagg tactacagca gcttcgtgaa catggagagg
gacctggcct ctggcctgat cggcccccctg ctgatctgct acaaggagag cgtggaccag
aggggcaacc agatcatgtc tgacaagagg aacgtgatcc tgttctctgt gttcgatgag
aacaggagct ggtatctgac cgagaacatc cagaggttcc tgcccaaccc cgccggcgtg
cagctggagg accccgagtt ccaggccagc aacatcatgc acagcatcaa cggctacgtg
ttcgacagcc tgcagctgtc tgtgtgcctg cacgaggtgg cctactggta catcctgagc
atcggcgccc agaccgactt cctgtctgtg ttcttctctg gctacacctt caagcacaag
atggtgtacg aggacaccct gaccctgttc cccttcagcg gcgagaccgt gttcatgagc
atggagaacc ccggcctgtg gatcctgggc tgccacaaca gcgacttcag gaacaggggc
atgaccgccc tgctgaaagt cagcagctgc gacaagaaca ccggcgacta ctacgaggac
agctacgagg acatcagcgc ctacctgctg agcaagaaca acgccatcga gcccagg
(SEQ ID NO:22)

图28

CS23-LC-NA

g agatcaccag gaccaccctg
cagagcgacc aggaggagat cgactatgat gacaccatca gcgtggagat gaagaaggag
gacttcgaca tctacgacga ggacgagaac cagagcccca ggagcttcca gaagaagacc
aggcactact tcatcgccgc cgtggagagg ctgtgggact atggcatgag cagcagcccc
cacgtgctga ggaacagggc ccagagcggc agcgtgcccc agttcaagaa ggtggtgttc
caggagttca ccgacggcag cttcacccag cccctgtaca gaggcgagct gaacgagcac
ctgggcctgc tgggccccta catcagggcc gaggtggagg acaacatcat ggtgaccttc
aggaaccagg ccagcaggcc ctacagcttc tacagcagcc tgatcagcta cgaggaggac
cagaggcagg gcgccgagcc caggaagaac ttcgtgaagc ccaacgagac caagacctac
ttctggaagg tgcagcacca catggccccc accaaggacg agttcgactg caaggcctgg
gcctacttct ctgatgtgga cctggagaag gacgtgcaca gcggcctgat cggccccctg
ctggtgtgcc acaccaacac cctgaacccc gcccacggca ggcaggtgac cgtgcaggag
ttcgccctgt tcttcaccat cttcgacgag accaagagct ggtacttcac cgagaacatg
gagaggaact gcagggcccc ctgcaacatc cagatggagg accccacctt caaggagaac
tacaggttcc acgccatcaa cggctacatc atggacaccc tgcccggcct ggtgatggcc
caggaccaga ggatcaggtg gtatctgctg agcatgggca gcaacgagaa catccacagc
atccacttca gcggccacgt gttcaccgtg aggaagaagg aggagtacaa gatggccctg
tacaacctgt accccggcgt gttcgagacc gtggagatgc tgcccagcaa ggccggcatc
tggagggtgg agtgcctgat cggcgagcac ctgcacgccg gcatgagcac cctgttcctg
gtgtacagca acaagtgcca gacccccctg ggcatggcca gcggccacat cagggacttc
cagatcaccg cctctggcca gtacggccag tgggcccccа agctggccag gctgcactac
agcggcagca tcaacgcctg gagcaccaag gagcccttca gctggatcaa ggtggacctg
ctggcccccа tgatcatcca cggcatcaag acccagggcg ccaggcagaa gttcagcagc
ctgtacatca gccagttcat catcatgtac agcctggacg gcaagaagtg gcagacctac
aggggcaaca gcaccggcac cctgatggtg ttcttcggca acgtggacag cagcggcatc
aagcacaaca tcttcaaccc ccccatcatc gccaggtaca tcaggctgca ccccacccac
tacagcatca ggagcaccct gcggatggaa ctgatgggct gcgacctgaa cagctgcagc
atgcccctgg gcatggagag caaggccatc tctgacgccc agatcaccgc cagcagctac
ttcaccaaca tgttcgccac ctggagcccc agcaaggcca ggctgcacct gcagggcagg
agcaacgcct ggaggcccca ggtgaacaac cccaaggagt ggctgcaggt ggacttccag
aagaccatga aggtgaccgg cgtgaccacc cagggcgtga agagcctgct gaccagcatg
tacgtgaagg agttcctgat cagcagcagc caggacggcc accagtggac cctgttcttc
cagaacggca aagtgaaggt gttccagggc aaccaggaca gcttcacccc cgtggtgaac
agcctggacc ccccccctgct gaccaggtat ctgaggatcc acccccagag ctgggtgcac
cagatcgccc tgagaatgga agtgctggga tgcgaggccc aggacctgta c
(SEQ ID NO:23)

图29

CS01m13-FL-NA

ATGCAGATTGAGCTGTCCACCTGCTTCTTTCTGTGCCTGCTGAGATTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGCTGTGGAACTTTCTTGGGACTACATGCAGTCTGACCTGGGAGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAGTCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACACTCTTTGTGGAATTCACTGACCACCTGTTCAACATTGCAAAACCCAGACCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCATCACCCTCAAGAACATG
GCATCCCACCCTGTGTCTCTGCATGCTGTGGGAGTCTCATACTGGAAAGCCTCTGAAGGGGCTGAG
TATGATGACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGGGATCTCACACC
TATGTGTGGCAAGTCCTCAAGGAGAATGGACCCATGGCATCTGACCCACTCTGCCTGACATACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCACTGCTGGTGTGC
AGGGAAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGTCTTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGATGCTGCC
TCTGCCAGGGCATGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAAGTGCAC
TCCATTTTCCTGGAGGGACACACCTTCCTGGTCAGGAACCACAGACAAGCCTCTCTGGAGATCTCT
CCCATCACCTTCCTCACTGCACAGACACTGCTGATGGACCTTGGACAGTTCCTGCTGTCCTGCCAC
ATCTCTTCCCACCAGCATGATGGCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCAGATTTGATGATGACAACTCTCCATCCTTCATTCAGATCAGGTCTGTGGCAAAGAAA
CACCCCAAGACATGGGTGCACTACATTGCTGCTGAGGAAGAGGACTGGGACTATGCACCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAACAATGGCCCACAAAGAATTGGAAGA
AAGTACAAGAAAGTCAGATTCATGGCCTACACTGATGAAACCTTCAAGACAAGAGAAGCCATTCAG
CATGAGTCTGGCATTCTGGGACCACTCCTGTATGGGGAAGTGGGAGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCAGGAGACTGCCAAAAGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGAGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCCAGGTGCCTCACCAGA
TACTACTCCTCTTTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGAGGCAACCAGATCATGTCTGACAAGAGAAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTCAGAGATTCCTGCCCAAC
CCTGCTGGGGTGCAACTGGAAGACCCTGAGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACTCTCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCACAAACTGACTTCCTTTCTGTCTTCTTCTCTGGATACACCTTCAAGCACAAGATGGTG
TATGAGGACACCCTGACACTCTTCCCATTCTCTGGGGAAACTGTGTTCATGAGCATGGAGAACCCT
GGACTGTGGATTCTGGGATGCCACAACTCTGACTTCAGAAACAGGGGAATGACTGCACTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATACCACCTACGTGAACCGCTCCCTGTCTCAGAATCCACCTGTCCTGAAG
AGACACCAGAGAGAGATCACCAGGACAACCCTCCAGTCTGACCAGGAAGAGATTGACTATGATGAC
ACCATTTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTATGATGAGGACGAGAACCAGTCTCCA
AGATCATTCCAGAAGAAGACAAGACACTACTTCATTGCTGCTGTGGAAAGACTGTGGGACTATGGC
ATGTCTTCCTCTCCCCATGTCCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAAGAAA

（续）

图30A

GTGGTCTTCCAGGAGTTCACTGATGGCTCATTCACCCAGCCCCTGTACAGAGGGGAACTGAATGAG
CACCTGGGACTCCTGGGACCATACATCAGGGCTGAGGTGGAAGACAACATCATGGTGACATTCAGA
AACCAGGCCTCCAGGCCCTACAGCTTCTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAA
GGGGCTGAGCCAAGAAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAG
CACCACATGGCACCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCATACTTCTCTGATGTGGAC
CTGGAGAAAGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAAC
CCTGCACATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACC
AAGTCATGGTACTTCACTGAGAACATGGAGAGAAACTGCAGAGCACCATGCAACATTCAGATGGAA
GACCCCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCT
GGGCTTGTCATGGCACAGGACCAGAGAATCAGATGGTACCTGCTTTCTATGGGATCCAATGAGAAC
ATTCACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGAGAAAGAAGGAGGAATACAAGATGGCC
CTGTACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGG
AGGGTGGAATGCCTCATTGGGGAGCACCTGCATGCTGGCATGTCAACCCTGTTCCTGGTCTACAGC
AACAAGTGCCAGACACCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCT
GGCCAGTATGGCCAGTGGGCACCCAAACTGGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGG
TCAACCAAGGAGCCATTCTCTTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTCATGGCATC
AAGACACAGGGGGCAAGACAGAAATTCTCCTCTCTGTACATCTCACAGTTCATCATCATGTACTCT
CTGGATGGCAAGAAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCTCATGGTCTTCTTTGGC
AATGTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCTCCCATCATTGCCAGATACATCAGG
CTGCACCCCACCCACTACTCAATCAGATCAACCCTCAGGATGGAACTGATGGGATGTGACCTGAAC
TCCTGCTCAATGCCCCTGGGAATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCATCCTCT
TACTTCACCAACATGTTTGCCACCTGGTCACCATCAAAAGCCAGGCTGCACCTCCAGGGAAGAAGC
AATGCCTGGAGACCCCAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAATG
AAAGTCACTGGGGTGACAACCCAGGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGAGTTC
CTGATCTCTTCCTCACAGGATGGCCACCAGTGGACACTCTTCTTCCAGAATGGCAAAGTCAAGGTG
TTCCAGGGCAACCAGGACTCTTTCACACCTGTGGTGAACTCACTGGACCCCCCCCTCCTGACAAGA
TACCTGAGAATTCACCCCCAGTCTTGGGTCCACCAGATTGCCCTGAGAATGGAAGTCCTGGGATGT
GAGGCACAAGACCTGTACTGA (SEQ ID NO:90)

图30B

CS01m23-FL-NA

ATGCAGATTGAGCTGTCCACCTGCTTCTTTCTGTGCCTGCTGAGATTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGCTGTGGAACTTTCTTGGGACTACATGCAGTCTGACCTGGGAGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAGTCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACACTCTTTGTGGAATTCACTGACCACCTGTTCAACATTGCAAAACCCAGACCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCGTCACCCTCAAGAACATG
GCATCCCACCCTGTGTCTCTGCATGCTGTGGGAGTCTCATACTGGAAATCCTCTGAAGGGGCTGAG
TATGATGACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGAAGTCTCACACC
TATGTGTGGCAAGTCCTCAAGGAGAATGGACCCACTGCATCTGACCCACCCTGCCTGACATACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCACTGCTGGTGTGC
AGGGAAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGTCTTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGATGCTGCC
TCTGCCAGGGCATGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAAGTGCAC
TCCATTTTCCTGGAGGGACACACCTTCCTGGTCAGGAACCACAGACAAGCCTCTCTGGAGATCTCT
CCCATCACCTTCCTCACTGCACAGACACTGCTGATGGACCTTGGACAGTTCCTGCTGTTCTGCCAC
ATCTCTTCCCACCAGCATGATGGCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCAGATTTGATGATGACAACTCTCCATCCTTCATTCAGATCAGGTCTGTGGCAAAGAAA
CACCCCAAGACATGGGTGCACTACATTGCTGCTGAGGAAGAGGACTGGGACTATGCACCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAACAATGGCCCACAAAGAATTGGAAGA
AAGTACAAGAAAGTCAGATTCATGGCCTACACTGATGAAACCTTCAAGACAAGAGAAGCCATTCAG
CATGAGTCTGGCATTCTGGGACCACTCCTGTATGGGGAAGTGGGAGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCAGGAGACTGCCAAAAGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGAGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCCAGGTGCCTCACCAGA
TACTACTCCTCTTTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGAGGCAACCAGATCATGTCTGACAAGAGAAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTCAGAGATTCCTGCCCAAC
CCTGCTGGGGTGCAACTGGAAGACCCTGAGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACTCTCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCACAAACTGACTTCCTTTCTGTCTTCTTCTCTGGATACACCTTCAAGCACAAGATGGTG
TATGAGGACACCCTGACACTCTTCCCATTCTCTGGGGAAACTGTGTTCATGAGCATGGAGAACCCT
GGACTGTGGATTCTGGGATGCCACAACTCTGACTTCAGAAACAGGGGAATGACTGCACTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATACCACCTACGTGAACCGCTCCCTGTCTCAGAATCCACCTGTCCTGAAG
AGACACCAGAGAGAGATCACCAGGACAACCCTCCAGTCTGACCAGGAAGAGATTGACTATGATGAC
ACCATTTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTATGATGAGGACGAGAACCAGTCTCCA
AGATCATTCCAGAAGAAGACAAGACACTACTTCATTGCTGCTGTGGAAAGACTGTGGGACTATGGC
ATGTCTTCCTCTCCCCATGTCCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAAGAAA

（续）

图31A

GTGGTCTTCCAGGAGTTCACTGATGGCTCATTCACCCAGCCCCTGTACAGAGGGGAACTGAATGAG
CACCTGGGACTCCTGGGACCATACATCAGGGCTGAGGTGGAAGACAACATCATGGTGACATTCAGA
AACCAGGCCTCCAGGCCCTACAGCTTCTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAA
GGGGCTGAGCCAAGAAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAG
CACCACATGGCACCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCATACTTCTCTGATGTGGAC
CTGGAGAAAGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAAC
CCTGCACATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACC
AAGTCATGGTACTTCACTGAGAACATGGAGAGAAACTGCAGAGCACCATGCAACATTCAGATGGAA
GACCCCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCT
GGGCTTGTCATGGCACAGGACCAGAGAATCAGATGGTACCTGCTTTCTATGGGATCCAATGAGAAC
ATTCACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGAGAAAGAAGGAGGAATACAAGATGGCC
CTGTACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGG
AGGGTGGAATGCCTCATTGGGGAGCACCTGCATGCTGGCATGTCAACCCTGTTCCTGGTCTACAGC
AACAAGTGCCAGACACCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCT
GGCCAGTATGGCCAGTGGGCACCCAAACTGGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGG
TCAACCAAGGAGCCATTCTCTTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTCATGGCATC
AAGACACAGGGGGCAAGACAGAAATTCTCCTCTCTGTACATCTCACAGTTCATCATCATGTACTCT
CTGGATGGCAAGAAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCTCATGGTCTTCTTTGGC
AATGTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCTCCCATCATTGCCAGATACATCAGG
CTGCACCCCACCCACTACTCAATCAGATCAACCCTCAGGATGGAACTGATGGGATGTGACCTGAAC
TCCTGCTCAATGCCCCTGGGAATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCATCCTCT
TACTTCACCAACATGTTTGCCACCTGGTCACCATCAAAAGCCAGGCTGCACCTCCAGGGAAGAAGC
AATGCCTGGAGACCCCAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAATG
AAAGTCACTGGGGTGACAACCCAGGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGAGTTC
CTGATCTCTTCCTCACAGGATGGCCACCAGTGGACACTCTTCTTCCAGAATGGCAAAGTCAAGGTG
TTCCAGGGCAACCAGGACTCTTTCACACCTGTGGTGAACTCACTGGACCCCCCCCTCCTGACAAGA
TACCTGAGAATTCACCCCCAGTCTTGGGTCCACCAGATTGCCCTGAGAATGGAAGTCCTGGGATGT
GAGGCACAAGACCTGTACTGA (SEQ ID NO:91)

图31B

CS01m3-FL-NA

ATGCAGATTGAGCTGTCCACCTGCTTCTTTCTGTGCCTGCTGAGATTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGCTGTGGAACTTTCTTGGGACTACATGCAGTCTGACCTGGGAGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAGTCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACACTCTTTGTGGAATTCACTGACCACCTGTTCAACATTGCAAAACCCAGACCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCATCACCCTCAAGAACATG
GCATCCCACCCTGTGTCTCTGCATGCTGTGGGAGTCTCATACTGGAAAGCCTCTGAAGGGGCTGAG
TATGATGACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGGGATCTCACACC
TATGTGTGGCAAGTCCTCAAGGAGAATGGACCCATGGCATCTGACCCACTCTGCCTGACATACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCACTGCTGGTGTGC
AGGGAAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGTCTTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGATGCTGCC
TCTGCCAGGGCATGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAAGTGCAC
TCCATTTTCCTGGAGGGACACACCTTCCTGGTCAGGAACCACAGACAAGCCTCTCTGGAGATCTCT
CCCATCACCTTCCTCACTGCACAGACACTGCTGATGGACCTTGGACAGTTCCTGCTGTTCTGCCAC
ATCTCTTCCCACCAGCATGATGGCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCAGATTTGATGATGACAACTCTCCATCCTTCATTCAGATCAGGTCTGTGGCAAAGAAA
CACCCCAAGACATGGGTGCACTACATTGCTGCTGAGGAAGAGGACTGGGACTATGCACCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAACAATGGCCCACAAAGAATTGGAAGA
AAGTACAAGAAAGTCAGATTCATGGCCTACACTGATGAAACCTTCAAGACAAGAGAAGCCATTCAG
CATGAGTCTGGCATTCTGGGACCACTCCTGTATGGGGAAGTGGGAGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCAGGAGACTGCCAAAAGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGAGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCCAGGTGCCTCACCAGA
TACTACTCCTCTTTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGAGGCAACCAGATCATGTCTGACAAGAGAAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTCAGAGATTCCTGCCCAAC
CCTGCTGGGGTGCAACTGGAAGACCCTGAGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACTCTCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCACAAACTGACTTCCTTTCTGTCTTCTTCTCTGGATACACCTTCAAGCACAAGATGGTG
TATGAGGACACCCTGACACTCTTCCCATTCTCTGGGGAAACTGTGTTCATGAGCATGGAGAACCCT
GGACTGTGGATTCTGGGATGCCACAACTCTGACTTCAGAAACAGGGGAATGACTGCACTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATACCACCTACGTGAACCGCTCCCTGTCTCAGAATCCACCTGTCCTGAAG
AGACACCAGAGAGAGATCACCAGGACAACCCTCCAGTCTGACCAGGAAGAGATTGACTATGATGAC
ACCATTTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTATGATGAGGACGAGAACCAGTCTCCA
AGATCATTCCAGAAGAAGACAAGACACTACTTCATTGCTGCTGTGGAAAGACTGTGGGACTATGGC
ATGTCTTCCTCTCCCCATGTCCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAAGAAA

（续）

图32A

GTGGTCTTCCAGGAGTTCACTGATGGCTCATTCACCCAGCCCCTGTACAGAGGGGAACTGAATGAG
CACCTGGGACTCCTGGGACCATACATCAGGGCTGAGGTGGAAGACAACATCATGGTGACATTCAGA
AACCAGGCCTCCAGGCCCTACAGCTTCTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAA
GGGGCTGAGCCAAGAAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAG
CACCACATGGCACCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCATACTTCTCTGATGTGGAC
CTGGAGAAAGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAAC
CCTGCACATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACC
AAGTCATGGTACTTCACTGAGAACATGGAGAGAAACTGCAGAGCACCATGCAACATTCAGATGGAA
GACCCCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCT
GGGCTTGTCATGGCACAGGACCAGAGAATCAGATGGTACCTGCTTTCTATGGGATCCAATGAGAAC
ATTCACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGAGAAAGAAGGAGGAATACAAGATGGCC
CTGTACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGG
AGGGTGGAATGCCTCATTGGGGAGCACCTGCATGCTGGCATGTCAACCCTGTTCCTGGTCTACAGC
AACAAGTGCCAGACACCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCT
GGCCAGTATGGCCAGTGGGCACCCAAACTGGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGG
TCAACCAAGGAGCCATTCTCTTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTCATGGCATC
AAGACACAGGGGGCAAGACAGAAATTCTCCTCTCTGTACATCTCACAGTTCATCATCATGTACTCT
CTGGATGGCAAGAAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCTCATGGTCTTCTTTGGC
AATGTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCTCCCATCATTGCCAGATACATCAGG
CTGCACCCCACCCACTACTCAATCAGATCAACCCTCAGGATGGAACTGATGGGATGTGACCTGAAC
TCCTGCTCAATGCCCCTGGGAATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCATCCTCT
TACTTCACCAACATGTTTGCCACCTGGTCACCATCAAAAGCCAGGCTGCACCTCCAGGGAAGAAGC
AATGCCTGGAGACCCCAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAATG
AAAGTCACTGGGGTGACAACCCAGGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGAGTTC
CTGATCTCTTCCTCACAGGATGGCCACCAGTGGACACTCTTCTTCCAGAATGGCAAAGTCAAGGTG
TTCCAGGGCAACCAGGACTCTTTCACACCTGTGGTGAACTCACTGGACCCCCCCCTCCTGACAAGA
TACCTGAGAATTCACCCCCAGTCTTGGGTCCACCAGATTGCCCTGAGAATGGAAGTCCTGGGATGT
GAGGCACAAGACCTGTACTGA (SEQ ID NO:92)

图32B

CS01m2-FL-NA

ATGCAGATTGAGCTGTCCACCTGCTTCTTTCTGTGCCTGCTGAGATTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGCTGTGGAACTTTCTTGGGACTACATGCAGTCTGACCTGGGAGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAGTCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACACTCTTTGTGGAATTCACTGACCACCTGTTCAACATTGCAAAACCCAGACCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCGTCACCCTCAAGAACATG
GCATCCCACCCTGTGTCTCTGCATGCTGTGGGAGTCTCATACTGGAAATCCTCTGAAGGGGCTGAG
TATGATGACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGAAGTCTCACACC
TATGTGTGGCAAGTCCTCAAGGAGAATGGACCCACTGCATCTGACCCACCCTGCCTGACATACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCACTGCTGGTGTGC
AGGGAAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGTCTTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGATGCTGCC
TCTGCCAGGGCATGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAAGTGCAC
TCCATTTTCCTGGAGGGACACACCTTCCTGGTCAGGAACCACAGACAAGCCTCTCTGGAGATCTCT
CCCATCACCTTCCTCACTGCACAGACACTGCTGATGGACCTTGGACAGTTCCTGCTGTTCTGCCAC
ATCTCTTCCCACCAGCATGATGGCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCAGATTTGATGATGACAACTCTCCATCCTTCATTCAGATCAGGTCTGTGGCAAAGAAA
CACCCCAAGACATGGGTGCACTACATTGCTGCTGAGGAAGAGGACTGGGACTATGCACCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAACAATGGCCCACAAAGAATTGGAAGA
AAGTACAAGAAAGTCAGATTCATGGCCTACACTGATGAAACCTTCAAGACAAGAGAAGCCATTCAG
CATGAGTCTGGCATTCTGGGACCACTCCTGTATGGGGAAGTGGGAGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCAGGAGACTGCCAAAAGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGAGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCCAGGTGCCTCACCAGA
TACTACTCCTCTTTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGAGGCAACCAGATCATGTCTGACAAGAGAAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTCAGAGATTCCTGCCCAAC
CCTGCTGGGGTGCAACTGGAAGACCCTGAGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACTCTCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCACAAACTGACTTCCTTTCTGTCTTCTTCTCTGGATACACCTTCAAGCACAAGATGGTG
TATGAGGACACCCTGACACTCTTCCCATTCTCTGGGGAAACTGTGTTCATGAGCATGGAGAACCCT
GGACTGTGGATTCTGGGATGCCACAACTCTGACTTCAGAAACAGGGGAATGACTGCACTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATGCCATTGAGCCCAGAAGCTTCTCTCAGAATCCACCTGTCCTGAAGAGA
CACCAGAGAGAGATCACCAGGACAACCCTCCAGTCTGACCAGGAAGAGATTGACTATGATGACACC
ATTTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTATGATGAGGACGAGAACCAGTCTCCAAGA
TCATTCCAGAAGAAGACAAGACACTACTTCATTGCTGCTGTGGAAAGACTGTGGGACTATGGCATG
TCTTCCTCTCCCCATGTCCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAAGAAAGTG

（续）

图33A

GTCTTCCAGGAGTTCACTGATGGCTCATTCACCCAGCCCCTGTACAGAGGGGAACTGAATGAGCAC
CTGGGACTCCTGGGACCATACATCAGGGCTGAGGTGGAAGACAACATCATGGTGACATTCAGAAAC
CAGGCCTCCAGGCCCTACAGCTTCTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAAGGG
GCTGAGCCAAGAAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCAC
CACATGGCACCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCATACTTCTCTGATGTGGACCTG
GAGAAAGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAACCCT
GCACATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACCAAG
TCATGGTACTTCACTGAGAACATGGAGAGAAACTGCAGAGCACCATGCAACATTCAGATGGAAGAC
CCCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCTGGG
CTTGTCATGGCACAGGACCAGAGAATCAGATGGTACCTGCTTTCTATGGGATCCAATGAGAACATT
CACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGAGAAAGAAGGAGGAATACAAGATGGCCCTG
TACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGG
GTGGAATGCCTCATTGGGGAGCACCTGCATGCTGGCATGTCAACCCTGTTCCTGGTCTACAGCAAC
AAGTGCCAGACACCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGC
CAGTATGGCCAGTGGGCACCCAAACTGGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGGTCA
ACCAAGGAGCCATTCTCTTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTCATGGCATCAAG
ACACAGGGGGCAAGACAGAAATTCTCCTCTCTGTACATCTCACAGTTCATCATCATGTACTCTCTG
GATGGCAAGAAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCTCATGGTCTTCTTTGGCAAT
GTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCTCCCATCATTGCCAGATACATCAGGCTG
CACCCCACCCACTACTCAATCAGATCAACCCTCAGGATGGAACTGATGGGATGTGACCTGAACTCC
TGCTCAATGCCCCTGGGAATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCATCCTCTTAC
TTCACCAACATGTTTGCCACCTGGTCACCATCAAAAGCCAGGCTGCACCTCCAGGGAAGAAGCAAT
GCCTGGAGACCCCAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAATGAAA
GTCACTGGGGTGACAACCCAGGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGAGTTCCTG
ATCTCTTCCTCACAGGATGGCCACCAGTGGACACTCTTCTTCCAGAATGGCAAAGTCAAGGTGTTC
CAGGGCAACCAGGACTCTTTCACACCTGTGGTGAACTCACTGGACCCCCCCCTCCTGACAAGATAC
CTGAGAATTCACCCCCAGTCTTGGGTCCACCAGATTGCCCTGAGAATGGAAGTCCTGGGATGTGAG
GCACAAGACCTGTACTGA (SEQ ID NO:93)

图33B

CS04m2-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGGCTGTGGAGCTTTCTTGGGACTACATGCAGTCTGACCTGGGGGGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAATCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACCCTCTTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAAACCCAGGCCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCGTCACCCTCAAGAACATG
GCCTCCCACCCTGTGAGCCTGCATGCTGTGGGGGTCAGCTACTGGAAGTCCTCTGAGGGGGCTGAG
TATGATGACCAGACCTCCCAGAGGGAGAAGGAGGATGACAAAGTGTTCCCTGGGAAGAGCCACACC
TATGTGTGGCAGGTCCTCAAGGAGAATGGCCCCACTGCCTCTGACCCACCCTGCCTGACCTACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCCCTGCTGGTGTGC
AGGGAGGGCTCCCTGGCCAAAGAGAAGACCCAGACCCTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACTCCCTGATGCAGGACAGGGATGCTGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAGGTGCAC
TCCATTTTCCTGGAGGGCCACACCTTCCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCCTCACTGCCCAGACCCTGCTGATGGACCTCGGACAGTTCCTGCTGTTCTGCCAC
ATCAGCTCCCACCAGCATGATGGCATGGAGGCCTATGTCAAGGTGGACAGCTGCCCTGAGGAGCCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCCGCTTTGATGATGACAACAGCCCATCCTTCATTCAGATCAGGTCTGTGGCCAAGAAA
CACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGGACTGGGACTATGCCCCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGGCCCACAGAGGATTGGACGC
AAGTACAAGAAAGTCAGGTTCATGGCCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATTCAG
CATGAGTCTGGCATCCTGGGCCCACTCCTGTATGGGGAGGTGGGGGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCCGCAGGCTGCCAAAGGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGGGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACCAAATCTGACCCCAGGTGCCTCACCAGA
TACTACTCCAGCTTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGGGGAAACCAGATCATGTCTGACAAGAGGAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAGCGCTTCCTGCCCAAC
CCTGCTGGGGTGCAGCTGGAGGACCCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACAGCCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCCCAGACTGACTTCCTTTCTGTCTTCTTCTCTGGCTACACCTTCAAACACAAGATGGTG
TATGAGGACACCCTGACCCTCTTCCCATTCTCTGGGGAGACTGTGTTCATGAGCATGGAGAACCCT
GGCCTGTGGATTCTGGGATGCCACAACTCTGACTTCCGCAACAGGGGCATGACTGCCCTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACAGCTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATGCCATTGAGCCCAGGAGCTTCAGCCAGAATCCACCTGTCCTGAAACGC
CACCAGAGGGAGATCACCAGGACCACCCTCCAGTCTGACCAGGAGGAGATTGACTATGATGACACC
ATTTCTGTGGAGATGAAGAAAGAGGACTTTGACATCTATGACGAGGACGAGAACCAGAGCCCAAGG
AGCTTCCAGAAGAAGACCAGGCACTACTTCATTGCTGCTGTGGAGCGCCTGTGGGACTATGGCATG
AGCTCCAGCCCCCATGTCCTCAGGAACAGGGCCCAGTCTGGCTCTGTGCCACAGTTCAAGAAAGTG

（续）

图34A

GTCTTCCAAGAGTTCACTGATGGCAGCTTCACCCAGCCCCTGTACAGAGGGGAGCTGAATGAGCAC
CTGGGACTCCTGGGCCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACCTTCCGCAAC
CAGGCCTCCAGGCCCTACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAGGGG
GCTGAGCCACGCAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCAC
CACATGGCCCCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGACCTG
GAGAAGGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAACCCT
GCCCATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACCAAG
AGCTGGTACTTCACTGAGAACATGGAGCGCAACTGCAGGGCCCCATGCAACATTCAGATGGAGGAC
CCCACCTTCAAAGAGAACTACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCTGGG
CTTGTCATGGCCCAGGACCAGAGGATCAGGTGGTACCTGCTTTCTATGGGCTCCAATGAGAACATT
CACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGCGCAAGAAGGAGGAGTACAAGATGGCCCTG
TACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGG
GTGGAGTGCCTCATTGGGGAGCACCTGCATGCTGGCATGAGCACCCTGTTCCTGGTCTACAGCAAC
AAGTGCCAGACCCCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGC
CAGTATGGCCAGTGGGCCCCCAAGCTGGCCAGGCTCCACTACTCTGGATCCATCAATGCCTGGAGC
ACCAAGGAGCCATTCAGCTGGATCAAAGTGGACCTGCTGGCCCCCATGATCATCCATGGCATCAAG
ACCCAGGGGGCCAGGCAGAAGTTCTCCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTG
GATGGCAAGAAATGGCAGACCTACAGAGGCAACTCCACTGGAACACTCATGGTCTTCTTTGGCAAT
GTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCCCCAATCATCGCCAGATACATCAGGCTG
CACCCCACCCACTACAGCATCCGCAGCACCCTCAGGATGGAGCTGATGGGCTGTGACCTGAACTCC
TGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCCTCCAGCTAC
TTCACCAACATGTTTGCCACCTGGAGCCCAAGCAAGGCCAGGCTGCACCTCCAGGGAAGGAGCAAT
GCCTGGAGGCCCCAGGTCAACAACCCAAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATGAAG
GTCACTGGGGTGACCACCCAGGGGGTCAAGAGCCTGCTCACCAGCATGTATGTGAAGGAGTTCCTG
ATCAGCTCCAGCCAGGATGGCCACCAGTGGACCCTCTTCTTCCAGAATGGCAAGGTCAAGGTGTTC
CAGGGCAACCAGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCCCTCCTGACCAGATAC
CTGAGGATTCACCCCCAGAGCTGGGTCCACCAGATTGCCCTGAGGATGGAGGTCCTGGGATGTGAG
GCCCAGGACCTGTACTGA (SEQ ID NO:94)

图34B

CS04m3-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGCTGTGGAGCTTTCTTGGGACTACATGCAGTCTGACCTGGGGGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAATCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACCCTCTTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAAACCCAGGCCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCATCACCCTCAAGAACATG
GCCTCCCACCCTGTGAGCCTGCATGCTGTGGGGGTCAGCTACTGGAAGGCCTCTGAGGGGGCTGAG
TATGATGACCAGACCTCCCAGAGGGAGAAGGAGGATGACAAAGTGTTCCCTGGGGGCAGCCACACC
TATGTGTGGCAGGTCCTCAAGGAGAATGGCCCCATGGCCTCTGACCCACTCTGCCTGACCTACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCCCTGCTGGTGTGC
AGGGAGGGCTCCCTGGCCAAAGAGAAGACCCAGACCCTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACTCCCTGATGCAGGACAGGGATGCTGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAGGTGCAC
TCCATTTTCCTGGAGGGCCACACCTTCCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCCTCACTGCCCAGACCCTGCTGATGGACCTCGGACAGTTCCTGCTGTTCTGCCAC
ATCAGCTCCCACCAGCATGATGGCATGGAGGCCTATGTCAAGGTGGACAGCTGCCCTGAGGAGCCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCCGCTTTGATGATGACAACAGCCCATCCTTCATTCAGATCAGGTCTGTGGCCAAGAAA
CACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGGACTGGGACTATGCCCCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGGCCCACAGAGGATTGGACGC
AAGTACAAGAAAGTCAGGTTCATGGCCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATTCAG
CATGAGTCTGGCATCCTGGGCCCACTCCTGTATGGGGAGGTGGGGGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCCGCAGGCTGCCAAAGGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGGGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACCAAATCTGACCCCAGGTGCCTCACCAGA
TACTACTCCAGCTTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGGGGAAACCAGATCATGTCTGACAAGAGGAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAGCGCTTCCTGCCCAAC
CCTGCTGGGGTGCAGCTGGAGGACCCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACAGCCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCCCAGACTGACTTCCTTTCTGTCTTCTTCTCTGGCTACACCTTCAAACACAAGATGGTG
TATGAGGACACCCTGACCCTCTTCCCATTCTCTGGGGAGACTGTGTTCATGAGCATGGAGAACCCT
GGCCTGTGGATTCTGGGATGCCACAACTCTGACTTCCGCAACAGGGGCATGACTGCCCTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACAGCTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATACCACCTACGTGAACCGCTCCCTGAGCCAGAATCCACCTGTCCTGAAA
CGCCACCAGAGGGAGATCACCAGGACCACCCTCCAGTCTGACCAGGAGGAGATTGACTATGATGAC
ACCATTTCTGTGGAGATGAAGAAAGAGGACTTTGACATCTATGACGAGGACGAGAACCAGAGCCCA
AGGAGCTTCCAGAAGAAGACCAGGCACTACTTCATTGCTGCTGTGGAGCGCCTGTGGGACTATGGC
ATGAGCTCCAGCCCCCATGTCCTCAGGAACAGGGCCCAGTCTGGCTCTGTGCCACAGTTCAAGAAA

（续）

图35A

GTGGTCTTCCAAGAGTTCACTGATGGCAGCTTCACCCAGCCCCTGTACAGAGGGGAGCTGAATGAG
CACCTGGGACTCCTGGGCCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACCTTCCGC
AACCAGGCCTCCAGGCCCTACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAG
GGGGCTGAGCCACGCAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAG
CACCACATGGCCCCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGAC
CTGGAGAAGGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAAC
CCTGCCCATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACC
AAGAGCTGGTACTTCACTGAGAACATGGAGCGCAACTGCAGGGCCCCATGCAACATTCAGATGGAG
GACCCCACCTTCAAAGAGAACTACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCT
GGGCTTGTCATGGCCCAGGACCAGAGGATCAGGTGGTACCTGCTTTCTATGGGCTCCAATGAGAAC
ATTCACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGCGCAAGAAGGAGGAGTACAAGATGGCC
CTGTACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGG
AGGGTGGAGTGCCTCATTGGGGAGCACCTGCATGCTGGCATGAGCACCCTGTTCCTGGTCTACAGC
AACAAGTGCCAGACCCCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCT
GGCCAGTATGGCCAGTGGGCCCCCAAGCTGGCCAGGCTCCACTACTCTGGATCCATCAATGCCTGG
AGCACCAAGGAGCCATTCAGCTGGATCAAAGTGGACCTGCTGGCCCCCATGATCATCCATGGCATC
AAGACCCAGGGGGCCAGGCAGAAGTTCTCCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGC
CTGGATGGCAAGAAATGGCAGACCTACAGAGGCAACTCCACTGGAACACTCATGGTCTTCTTTGGC
AATGTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCCCCAATCATCGCCAGATACATCAGG
CTGCACCCCACCCACTACAGCATCCGCAGCACCCTCAGGATGGAGCTGATGGGCTGTGACCTGAAC
TCCTGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCCTCCAGC
TACTTCACCAACATGTTTGCCACCTGGAGCCCAAGCAAGGCCAGGCTGCACCTCCAGGGAAGGAGC
AATGCCTGGAGGCCCCAGGTCAACAACCCAAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATG
AAGGTCACTGGGGTGACCACCCAGGGGGTCAAGAGCCTGCTCACCAGCATGTATGTGAAGGAGTTC
CTGATCAGCTCCAGCCAGGATGGCCACCAGTGGACCCTCTTCTTCCAGAATGGCAAGGTCAAGGTG
TTCCAGGGCAACCAGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCCCCTCCTGACCAGA
TACCTGAGGATTCACCCCCAGAGCTGGGTCCACCAGATTGCCCTGAGGATGGAGGTCCTGGGATGT
GAGGCCCAGGACCTGTACTGA (SEQ ID NO:95)

图35B

CS04m23-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGCTGTGGAGCTTTCTTGGGACTACATGCAGTCTGACCTGGGGGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAATCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACCCTCTTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAAACCCAGGCCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCGTCACCCTCAAGAACATG
GCCTCCCACCCTGTGAGCCTGCATGCTGTGGGGGTCAGCTACTGGAAGTCCTCTGAGGGGGCTGAG
TATGATGACCAGACCTCCCAGAGGGAGAAGGAGGATGACAAAGTGTTCCCTGGGAAGAGCCACACC
TATGTGTGGCAGGTCCTCAAGGAGAATGGCCCCACTGCCTCTGACCCACCCTGCCTGACCTACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCCCTGCTGGTGTGC
AGGGAGGGCTCCCTGGCCAAAGAGAAGACCCAGACCCTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACTCCCTGATGCAGGACAGGGATGCTGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAGGTGCAC
TCCATTTTCCTGGAGGGCCACACCTTCCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCCTCACTGCCCAGACCCTGCTGATGGACCTCGGACAGTTCCTGCTGTTCTGCCAC
ATCAGCTCCCACCAGCATGATGGCATGGAGGCCTATGTCAAGGTGGACAGCTGCCCTGAGGAGCCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCCGCTTTGATGATGACAACAGCCCATCCTTCATTCAGATCAGGTCTGTGGCCAAGAAA
CACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGGACTGGGACTATGCCCCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGGCCCACAGAGGATTGGACGC
AAGTACAAGAAAGTCAGGTTCATGGCCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATTCAG
CATGAGTCTGGCATCCTGGGCCCACTCCTGTATGGGGAGGTGGGGGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCCGCAGGCTGCCAAAGGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGGGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACCAAATCTGACCCCAGGTGCCTCACCAGA
TACTACTCCAGCTTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGGGGAAACCAGATCATGTCTGACAAGAGGAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAGCGCTTCCTGCCCAAC
CCTGCTGGGGTGCAGCTGGAGGACCCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACAGCCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCCCAGACTGACTTCCTTTCTGTCTTCTTCTCTGGCTACACCTTCAAACACAAGATGGTG
TATGAGGACACCCTGACCCTCTTCCCATTCTCTGGGGAGACTGTGTTCATGAGCATGGAGAACCCT
GGCCTGTGGATTCTGGGATGCCACAACTCTGACTTCCGCAACAGGGGCATGACTGCCCTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACAGCTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATACCACCTACGTGAACCGCTCCCTGAGCCAGAATCCACCTGTCCTGAAA
CGCCACCAGAGGGAGATCACCAGGACCACCCTCCAGTCTGACCAGGAGGAGATTGACTATGATGAC
ACCATTTCTGTGGAGATGAAGAAAGAGGACTTTGACATCTATGACGAGGACGAGAACCAGAGCCCA
AGGAGCTTCCAGAAGAAGACCAGGCACTACTTCATTGCTGCTGTGGAGCGCCTGTGGGACTATGGC
ATGAGCTCCAGCCCCCATGTCCTCAGGAACAGGGCCCAGTCTGGCTCTGTGCCACAGTTCAAGAAA

（续）

图36A

GTGGTCTTCCAAGAGTTCACTGATGGCAGCTTCACCCAGCCCCTGTACAGAGGGGAGCTGAATGAG
CACCTGGGACTCCTGGGCCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACCTTCCGC
AACCAGGCCTCCAGGCCCTACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAG
GGGGCTGAGCCACGCAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAG
CACCACATGGCCCCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGAC
CTGGAGAAGGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAAC
CCTGCCCATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACC
AAGAGCTGGTACTTCACTGAGAACATGGAGCGCAACTGCAGGGCCCCATGCAACATTCAGATGGAG
GACCCCACCTTCAAAGAGAACTACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCT
GGGCTTGTCATGGCCCAGGACCAGAGGATCAGGTGGTACCTGCTTTCTATGGGCTCCAATGAGAAC
ATTCACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGCGCAAGAAGGAGGAGTACAAGATGGCC
CTGTACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGG
AGGGTGGAGTGCCTCATTGGGGAGCACCTGCATGCTGGCATGAGCACCCTGTTCCTGGTCTACAGC
AACAAGTGCCAGACCCCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCT
GGCCAGTATGGCCAGTGGGCCCCCAAGCTGGCCAGGCTCCACTACTCTGGATCCATCAATGCCTGG
AGCACCAAGGAGCCATTCAGCTGGATCAAAGTGGACCTGCTGGCCCCCATGATCATCCATGGCATC
AAGACCCAGGGGGCCAGGCAGAAGTTCTCCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGC
CTGGATGGCAAGAAATGGCAGACCTACAGAGGCAACTCCACTGGAACACTCATGGTCTTCTTTGGC
AATGTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCCCCAATCATCGCCAGATACATCAGG
CTGCACCCCACCCACTACAGCATCCGCAGCACCCTCAGGATGGAGCTGATGGGCTGTGACCTGAAC
TCCTGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCCTCCAGC
TACTTCACCAACATGTTTGCCACCTGGAGCCCAAGCAAGGCCAGGCTGCACCTCCAGGGAAGGAGC
AATGCCTGGAGGCCCCAGGTCAACAACCCAAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATG
AAGGTCACTGGGGTGACCACCCAGGGGGTCAAGAGCCTGCTCACCAGCATGTATGTGAAGGAGTTC
CTGATCAGCTCCAGCCAGGATGGCCACCAGTGGACCCTCTTCTTCCAGAATGGCAAGGTCAAGGTG
TTCCAGGGCAACCAGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCCCCTCCTGACCAGA
TACCTGAGGATTCACCCCCAGAGCTGGGTCCACCAGATTGCCCTGAGGATGGAGGTCCTGGGATGT
GAGGCCCAGGACCTGTACTGA (SEQ ID NO:96)

图36B

CS04m1-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGCTGTGGAGCTTTCTTGGGACTACATGCAGTCTGACCTGGGGGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAATCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACCCTCTTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAAACCCAGGCCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCATCACCCTCAAGAACATG
GCCTCCCACCCTGTGAGCCTGCATGCTGTGGGGGTCAGCTACTGGAAGGCCTCTGAGGGGGCTGAG
TATGATGACCAGACCTCCCAGAGGGAGAAGGAGGATGACAAAGTGTTCCCTGGGGGCAGCCACACC
TATGTGTGGCAGGTCCTCAAGGAGAATGGCCCCATGGCCTCTGACCCACTCTGCCTGACCTACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCCCTGCTGGTGTGC
AGGGAGGGCTCCCTGGCCAAAGAGAAGACCCAGACCCTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACTCCCTGATGCAGGACAGGGATGCTGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAGGTGCAC
TCCATTTTCCTGGAGGGCCACACCTTCCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCCTCACTGCCCAGACCCTGCTGATGGACCTCGGACAGTTCCTGCTGTCCTGCCAC
ATCAGCTCCCACCAGCATGATGGCATGGAGGCCTATGTCAAGGTGGACAGCTGCCCTGAGGAGCCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCCGCTTTGATGATGACAACAGCCCATCCTTCATTCAGATCAGGTCTGTGGCCAAGAAA
CACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGGACTGGGACTATGCCCCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGGCCCACAGAGGATTGGACGC
AAGTACAAGAAAGTCAGGTTCATGGCCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATTCAG
CATGAGTCTGGCATCCTGGGCCCACTCCTGTATGGGGAGGTGGGGGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCCGCAGGCTGCCAAAGGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGGGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACCAAATCTGACCCCAGGTGCCTCACCAGA
TACTACTCCAGCTTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGGGGAAACCAGATCATGTCTGACAAGAGGAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAGCGCTTCCTGCCCAAC
CCTGCTGGGGTGCAGCTGGAGGACCCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACAGCCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCCCAGACTGACTTCCTTTCTGTCTTCTTCTCTGGCTACACCTTCAAACACAAGATGGTG
TATGAGGACACCCTGACCCTCTTCCCATTCTCTGGGGAGACTGTGTTCATGAGCATGGAGAACCCT
GGCCTGTGGATTCTGGGATGCCACAACTCTGACTTCCGCAACAGGGGCATGACTGCCCTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACAGCTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATGCCATTGAGCCCAGGAGCTTCAGCCAGAATCCACCTGTCCTGAAACGC
CACCAGAGGGAGATCACCAGGACCACCCTCCAGTCTGACCAGGAGGAGATTGACTATGATGACACC
ATTTCTGTGGAGATGAAGAAAGAGGACTTTGACATCTATGACGAGGACGAGAACCAGAGCCCAAGG
AGCTTCCAGAAGAAGACCAGGCACTACTTCATTGCTGCTGTGGAGCGCCTGTGGGACTATGGCATG
AGCTCCAGCCCCCATGTCCTCAGGAACAGGGCCCAGTCTGGCTCTGTGCCACAGTTCAAGAAAGTG

（续）

图37A

GTCTTCCAAGAGTTCACTGATGGCAGCTTCACCCAGCCCCTGTACAGAGGGGAGCTGAATGAGCAC
CTGGGACTCCTGGGCCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACCTTCCGCAAC
CAGGCCTCCAGGCCCTACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAGGGG
GCTGAGCCACGCAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCAC
CACATGGCCCCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGACCTG
GAGAAGGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAACCCT
GCCCATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACCAAG
AGCTGGTACTTCACTGAGAACATGGAGCGCAACTGCAGGGCCCCATGCAACATTCAGATGGAGGAC
CCCACCTTCAAAGAGAACTACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCTGGG
CTTGTCATGGCCCAGGACCAGAGGATCAGGTGGTACCTGCTTTCTATGGGCTCCAATGAGAACATT
CACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGCGCAAGAAGGAGGAGTACAAGATGGCCCTG
TACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGG
GTGGAGTGCCTCATTGGGGAGCACCTGCATGCTGGCATGAGCACCCTGTTCCTGGTCTACAGCAAC
AAGTGCCAGACCCCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGC
CAGTATGGCCAGTGGGCCCCCAAGCTGGCCAGGCTCCACTACTCTGGATCCATCAATGCCTGGAGC
ACCAAGGAGCCATTCAGCTGGATCAAAGTGGACCTGCTGGCCCCCATGATCATCCATGGCATCAAG
ACCCAGGGGGCCAGGCAGAAGTTCTCCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTG
GATGGCAAGAAATGGCAGACCTACAGAGGCAACTCCACTGGAACACTCATGGTCTTCTTTGGCAAT
GTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCCCCAATCATCGCCAGATACATCAGGCTG
CACCCCACCCACTACAGCATCCGCAGCACCCTCAGGATGGAGCTGATGGGCTGTGACCTGAACTCC
TGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCCTCCAGCTAC
TTCACCAACATGTTTGCCACCTGGAGCCCAAGCAAGGCCAGGCTGCACCTCCAGGGAAGGAGCAAT
GCCTGGAGGCCCCAGGTCAACAACCCAAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATGAAG
GTCACTGGGGTGACCACCCAGGGGGTCAAGAGCCTGCTCACCAGCATGTATGTGAAGGAGTTCCTG
ATCAGCTCCAGCCAGGATGGCCACCAGTGGACCCTCTTCTTCCAGAATGGCAAGGTCAAGGTGTTC
CAGGGCAACCAGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCCCTCCTGACCAGATAC
CTGAGGATTCACCCCCAGAGCTGGGTCCACCAGATTGCCCTGAGGATGGAGGTCCTGGGATGTGAG
GCCCAGGACCTGTACTGA (SEQ ID NO:97)

图37B

CS04m13-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGCTGTGGAGCTTTCTTGGGACTACATGCAGTCTGACCTGGGGGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAATCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACCCTCTTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAAACCCAGGCCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCATCACCCTCAAGAACATG
GCCTCCCACCCTGTGAGCCTGCATGCTGTGGGGGTCAGCTACTGGAAGGCCTCTGAGGGGGCTGAG
TATGATGACCAGACCTCCCAGAGGGAGAAGGAGGATGACAAAGTGTTCCCTGGGGGCAGCCACACC
TATGTGTGGCAGGTCCTCAAGGAGAATGGCCCCATGGCCTCTGACCCACTCTGCCTGACCTACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCCCTGCTGGTGTGC
AGGGAGGGCTCCCTGGCCAAAGAGAAGACCCAGACCCTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACTCCCTGATGCAGGACAGGGATGCTGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAGGTGCAC
TCCATTTTCCTGGAGGGCCACACCTTCCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCCTCACTGCCCAGACCCTGCTGATGGACCTCGGACAGTTCCTGCTGTCCTGCCAC
ATCAGCTCCCACCAGCATGATGGCATGGAGGCCTATGTCAAGGTGGACAGCTGCCCTGAGGAGCCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCCGCTTTGATGATGACAACAGCCCATCCTTCATTCAGATCAGGTCTGTGGCCAAGAAA
CACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGGACTGGGACTATGCCCCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGGCCCACAGAGGATTGGACGC
AAGTACAAGAAAGTCAGGTTCATGGCCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATTCAG
CATGAGTCTGGCATCCTGGGCCCACTCCTGTATGGGGAGGTGGGGGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCCGCAGGCTGCCAAAGGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGGGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACCAAATCTGACCCCAGGTGCCTCACCAGA
TACTACTCCAGCTTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGGGGAAACCAGATCATGTCTGACAAGAGGAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAGCGCTTCCTGCCCAAC
CCTGCTGGGGTGCAGCTGGAGGACCCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACAGCCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCCCAGACTGACTTCCTTTCTGTCTTCTTCTCTGGCTACACCTTCAAACACAAGATGGTG
TATGAGGACACCCTGACCCTCTTCCCATTCTCTGGGGAGACTGTGTTCATGAGCATGGAGAACCCT
GGCCTGTGGATTCTGGGATGCCACAACTCTGACTTCCGCAACAGGGGCATGACTGCCCTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACAGCTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATACCACCTACGTGAACCGCTCCCTGAGCCAGAATCCACCTGTCCTGAAA
CGCCACCAGAGGGAGATCACCAGGACCACCCTCCAGTCTGACCAGGAGGAGATTGACTATGATGAC
ACCATTTCTGTGGAGATGAAGAAAGAGGACTTTGACATCTATGACGAGGACGAGAACCAGAGCCCA
AGGAGCTTCCAGAAGAAGACCAGGCACTACTTCATTGCTGCTGTGGAGCGCCTGTGGGACTATGGC
ATGAGCTCCAGCCCCCATGTCCTCAGGAACAGGGCCCAGTCTGGCTCTGTGCCACAGTTCAAGAAA

〔续〕

图38A

GTGGTCTTCCAAGAGTTCACTGATGGCAGCTTCACCCAGCCCCTGTACAGAGGGGAGCTGAATGAG
CACCTGGGACTCCTGGGCCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACCTTCCGC
AACCAGGCCTCCAGGCCCTACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAG
GGGGCTGAGCCACGCAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAG
CACCACATGGCCCCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGAC
CTGGAGAAGGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAAC
CCTGCCCATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACC
AAGAGCTGGTACTTCACTGAGAACATGGAGCGCAACTGCAGGGCCCCATGCAACATTCAGATGGAG
GACCCCACCTTCAAAGAGAACTACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCT
GGGCTTGTCATGGCCCAGGACCAGAGGATCAGGTGGTACCTGCTTTCTATGGGCTCCAATGAGAAC
ATTCACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGCGCAAGAAGGAGGAGTACAAGATGGCC
CTGTACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGG
AGGGTGGAGTGCCTCATTGGGGAGCACCTGCATGCTGGCATGAGCACCCTGTTCCTGGTCTACAGC
AACAAGTGCCAGACCCCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCT
GGCCAGTATGGCCAGTGGGCCCCCAAGCTGGCCAGGCTCCACTACTCTGGATCCATCAATGCCTGG
AGCACCAAGGAGCCATTCAGCTGGATCAAAGTGGACCTGCTGGCCCCCATGATCATCCATGGCATC
AAGACCCAGGGGGCCAGGCAGAAGTTCTCCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGC
CTGGATGGCAAGAAATGGCAGACCTACAGAGGCAACTCCACTGGAACACTCATGGTCTTCTTTGGC
AATGTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCCCCAATCATCGCCAGATACATCAGG
CTGCACCCCACCCACTACAGCATCCGCAGCACCCTCAGGATGGAGCTGATGGGCTGTGACCTGAAC
TCCTGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCCTCCAGC
TACTTCACCAACATGTTTGCCACCTGGAGCCCAAGCAAGGCCAGGCTGCACCTCCAGGGAAGGAGC
AATGCCTGGAGGCCCCAGGTCAACAACCCAAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATG
AAGGTCACTGGGGTGACCACCCAGGGGGTCAAGAGCCTGCTCACCAGCATGTATGTGAAGGAGTTC
CTGATCAGCTCCAGCCAGGATGGCCACCAGTGGACCCTCTTCTTCCAGAATGGCAAGGTCAAGGTG
TTCCAGGGCAACCAGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCCCTCCTGACCAGA
TACCTGAGGATTCACCCCCAGAGCTGGGTCCACCAGATTGCCCTGAGGATGGAGGTCCTGGGATGT
GAGGCCCAGGACCTGTACTGA (SEQ ID NO:98)

图38B

CS23m13-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGCGCCGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGCGAGCTGCCT
GTGGACGCCAGGTTCCCCCCCAGAGTGCCCAAGAGCTTCCCCTTCAACACCTCAGTGGTGTACAAG
AAGACCCTGTTCGTGGAGTTCACCGACCACCTGTTCAACATCGCCAAGCCCAGGCCCCCCTGGATG
GGCCTGCTGGGCCCCACCATCCAGGCCGAGGTGTACGACACCGTGGTGATCACCCTGAAGAACATG
GCCAGCCACCCCGTGAGCCTGCACGCCGTGGGCGTGAGCTACTGGAAGGCCTCTGAGGGCGCCGAG
TATGACGACCAGACCAGCCAGAGGGAGAAGGAGGACGACAAGGTGTTCCCCGGCGGCAGCCACACC
TACGTGTGGCAGGTGCTGAAGGAGAACGGCCCCATGGCCAGCGACCCCCTGTGCCTGACCTACAGC
TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACTCTGGCCTGATCGGCGCCCTGCTGGTGTGC
AGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCCTGCACAAGTTCATCCTGCTGTTCGCCGTG
TTCGATGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAGGACAGGGATGCCGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACCGTGAACGGCTACGTGAACAGGAGCCTGCCCGGCCTG
ATCGGCTGCCACAGGAAGTCTGTGTACTGGCACGTGATCGGCATGGGCACCACCCCCGAGGTGCAC
AGCATCTTCCTGGAGGGCCACACCTTCCTGGTGAGGAACCACAGGCAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCCTGACCGCCCAGACCCTGCTGATGGACCTGGGCCAGTTCCTGCTGTCCTGCCAC
ATCAGCAGCCACCAGCACGACGGCATGGAGGCCTACGTGAAGGTGGACAGCTGCCCCGAGGAGCCC
CAGCTGAGGATGAAGAACAACGAGGAGGCCGAGGACTATGATGATGACCTGACCGACTCTGAGATG
GACGTGGTGAGGTTTGATGATGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCCAAGAAG
CACCCCAAGACCTGGGTGCACTACATCGCCGCCGAGGAGGAGGACTGGGACTACGCCCCCCTGGTG
CTGGCCCCCGACGACAGGAGCTACAAGAGCCAGTACCTGAACAACGGCCCCCAGAGGATCGGCAGG
AAGTACAAGAAGGTCAGATTCATGGCCTACACCGACGAGACCTTCAAGACCAGGGAGGCCATCCAG
CACGAGTCTGGCATCCTGGGCCCCCTGCTGTACGGCGAGGTGGGCGACACCCTGCTGATCATCTTC
AAGAACCAGGCCAGCAGGCCCTACAACATCTACCCCCACGGCATCACCGATGTGAGGCCCCTGTAC
AGCAGGAGGCTGCCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCCTGCCCGGCGAGATCTTC
AAGTACAAGTGGACCGTGACCGTGGAGGATGGCCCCACCAAGTCTGACCCCAGGTGCCTGACCAGG
TACTACAGCAGCTTCGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATCGGCCCCCTGCTGATC
TGCTACAAGGAGAGCGTGGACCAGAGGGGCAACCAGATCATGTCTGACAAGAGGAACGTGATCCTG
TTCTCTGTGTTCGATGAGAACAGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCCTGCCCAAC
CCCGCCGGCGTGCAGCTGGAGGACCCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCAACGGC
TACGTGTTCGACAGCCTGCAGCTGTCTGTGTGCCTGCACGAGGTGGCCTACTGGTACATCCTGAGC
ATCGGCGCCCAGACCGACTTCCTGTCTGTGTTCTTCTCTGGCTACACCTTCAAGCACAAGATGGTG
TACGAGGACACCCTGACCCTGTTCCCCTTCAGCGGCGAGACCGTGTTCATGAGCATGGAGAACCCC
GGCCTGTGGATCCTGGGCTGCCACAACAGCGACTTCAGGAACAGGGGCATGACCGCCCTGCTGAAA
GTCAGCAGCTGCGACAAGAACACCGGCGACTACTACGAGGACAGCTACGAGGACATCAGCGCCTAC
CTGCTGAGCAAGAACAACACCACCTACGTGAACCGCTCCCTGAGCCAGAACCCCCCCGTGCTGAAG
AGGCACCAGAGGGAGATCACCAGGACCACCCTGCAGAGCGACCAGGAGGAGATCGACTATGATGAC
ACCATCAGCGTGGAGATGAAGAAGGAGGACTTCGACATCTACGACGAGGACGAGAACCAGAGCCCC
AGGAGCTTCCAGAAGAAGACCAGGCACTACTTCATCGCCGCCGTGGAGAGGCTGTGGGACTATGGC
ATGAGCAGCAGCCCCCACGTGCTGAGGAACAGGGCCCAGAGCGGCAGCGTGCCCCAGTTCAAGAAG

（续）

图39A

GTGGTGTTCCAGGAGTTCACCGACGGCAGCTTCACCCAGCCCCTGTACAGAGGCGAGCTGAACGAG
CACCTGGGCCTGCTGGGCCCCTACATCAGGGCCGAGGTGGAGGACAACATCATGGTGACCTTCAGG
AACCAGGCCAGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCTACGAGGAGGACCAGAGGCAG
GGCGCCGAGCCCAGGAAGAACTTCGTGAAGCCCAACGAGACCAAGACCTACTTCTGGAAGGTGCAG
CACCACATGGCCCCCACCAAGGACGAGTTCGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGAC
CTGGAGAAGGACGTGCACAGCGGCCTGATCGGCCCCCTGCTGGTGTGCCACACCAACACCCTGAAC
CCCGCCCACGGCAGGCAGGTGACCGTGCAGGAGTTCGCCCTGTTCTTCACCATCTTCGACGAGACC
AAGAGCTGGTACTTCACCGAGAACATGGAGAGGAACTGCAGGGCCCCCTGCAACATCCAGATGGAG
GACCCCACCTTCAAGGAGAACTACAGGTTCCACGCCATCAACGGCTACATCATGGACACCCTGCCC
GGCCTGGTGATGGCCCAGGACCAGAGGATCAGGTGGTATCTGCTGAGCATGGGCAGCAACGAGAAC
ATCCACAGCATCCACTTCAGCGGCCACGTGTTCACCGTGAGGAAGAAGGAGGAGTACAAGATGGCC
CTGTACAACCTGTACCCCGGCGTGTTCGAGACCGTGGAGATGCTGCCCAGCAAGGCCGGCATCTGG
AGGGTGGAGTGCCTGATCGGCGAGCACCTGCACGCCGGCATGAGCACCCTGTTCCTGGTGTACAGC
AACAAGTGCCAGACCCCCCTGGGCATGGCCAGCGGCCACATCAGGGACTTCCAGATCACCGCCTCT
GGCCAGTACGGCCAGTGGGCCCCCAAGCTGGCCAGGCTGCACTACAGCGGCAGCATCAACGCCTGG
AGCACCAAGGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCCATGATCATCCACGGCATC
AAGACCCAGGGCGCCAGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGC
CTGGACGGCAAGAAGTGGCAGACCTACAGGGGCAACAGCACCGGCACCCTGATGGTGTTCTTCGGC
AACGTGGACAGCAGCGGCATCAAGCACAACATCTTCAACCCCCCCATCATCGCCAGGTACATCAGG
CTGCACCCCACCCACTACAGCATCAGGAGCACCCTGCGGATGGAACTGATGGGCTGCGACCTGAAC
AGCTGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATCTCTGACGCCCAGATCACCGCCAGCAGC
TACTTCACCAACATGTTCGCCACCTGGAGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGC
AACGCCTGGAGGCCCCAGGTGAACAACCCCAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATG
AAGGTGACCGGCGTGACCACCCAGGGCGTGAAGAGCCTGCTGACCAGCATGTACGTGAAGGAGTTC
CTGATCAGCAGCAGCCAGGACGGCCACCAGTGGACCCTGTTCTTCCAGAACGGCAAAGTGAAGGTG
TTCCAGGGCAACCAGGACAGCTTCACCCCCGTGGTGAACAGCCTGGACCCCCCCCCTGCTGACCAGG
TATCTGAGGATCCACCCCCAGAGCTGGGTGCACCAGATCGCCCTGAGAATGGAAGTGCTGGGATGC
GAGGCCCAGGACCTGTACTGA (SEQ ID NO:99)

图39B

CS23m3-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGCGCCGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGCGAGCTGCCT
GTGGACGCCAGGTTCCCCCCCAGAGTGCCCAAGAGCTTCCCCTTCAACACCTCAGTGGTGTACAAG
AAGACCCTGTTCGTGGAGTTCACCGACCACCTGTTCAACATCGCCAAGCCCAGGCCCCCCTGGATG
GGCCTGCTGGGCCCCACCATCCAGGCCGAGGTGTACGACACCGTGGTGATCACCCTGAAGAACATG
GCCAGCCACCCCGTGAGCCTGCACGCCGTGGGCGTGAGCTACTGGAAGGCCTCTGAGGGCGCCGAG
TATGACGACCAGACCAGCCAGAGGGAGAAGGAGGACGACAAGGTGTTCCCCGGCGGCAGCCACACC
TACGTGTGGCAGGTGCTGAAGGAGAACGGCCCCATGGCCAGCGACCCCCTGTGCCTGACCTACAGC
TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACTCTGGCCTGATCGGCGCCCTGCTGGTGTGC
AGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCCTGCACAAGTTCATCCTGCTGTTCGCCGTG
TTCGATGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAGGACAGGGATGCCGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACCGTGAACGGCTACGTGAACAGGAGCCTGCCCGGCCTG
ATCGGCTGCCACAGGAAGTCTGTGTACTGGCACGTGATCGGCATGGGCACCACCCCCGAGGTGCAC
AGCATCTTCCTGGAGGGCCACACCTTCCTGGTGAGGAACCACAGGCAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCCTGACCGCCCAGACCCTGCTGATGGACCTGGGCCAGTTCCTGCTGTTCTGCCAC
ATCAGCAGCCACCAGCACGACGGCATGGAGGCCTACGTGAAGGTGGACAGCTGCCCCGAGGAGCCC
CAGCTGAGGATGAAGAACAACGAGGAGGCCGAGGACTATGATGATGACCTGACCGACTCTGAGATG
GACGTGGTGAGGTTTGATGATGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCCAAGAAG
CACCCCAAGACCTGGGTGCACTACATCGCCGCCGAGGAGGAGGACTGGGACTACGCCCCCCTGGTG
CTGGCCCCCGACGACAGGAGCTACAAGAGCCAGTACCTGAACAACGGCCCCCAGAGGATCGGCAGG
AAGTACAAGAAGGTCAGATTCATGGCCTACACCGACGAGACCTTCAAGACCAGGGAGGCCATCCAG
CACGAGTCTGGCATCCTGGGCCCCCTGCTGTACGGCGAGGTGGGCGACACCCTGCTGATCATCTTC
AAGAACCAGGCCAGCAGGCCCTACAACATCTACCCCCACGGCATCACCGATGTGAGGCCCCTGTAC
AGCAGGAGGCTGCCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCCTGCCCGGCGAGATCTTC
AAGTACAAGTGGACCGTGACCGTGGAGGATGGCCCCACCAAGTCTGACCCCAGGTGCCTGACCAGG
TACTACAGCAGCTTCGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATCGGCCCCCTGCTGATC
TGCTACAAGGAGAGCGTGGACCAGAGGGGCAACCAGATCATGTCTGACAAGAGGAACGTGATCCTG
TTCTCTGTGTTCGATGAGAACAGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCCTGCCCAAC
CCCGCCGGCGTGCAGCTGGAGGACCCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCAACGGC
TACGTGTTCGACAGCCTGCAGCTGTCTGTGTGCCTGCACGAGGTGGCCTACTGGTACATCCTGAGC
ATCGGCGCCCAGACCGACTTCCTGTCTGTGTTCTTCTCTGGCTACACCTTCAAGCACAAGATGGTG
TACGAGGACACCCTGACCCTGTTCCCCTTCAGCGGCGAGACCGTGTTCATGAGCATGGAGAACCCC
GGCCTGTGGATCCTGGGCTGCCACAACAGCGACTTCAGGAACAGGGGCATGACCGCCCTGCTGAAA
GTCAGCAGCTGCGACAAGAACACCGGCGACTACTACGAGGACAGCTACGAGGACATCAGCGCCTAC
CTGCTGAGCAAGAACAACACCACCTACGTGAACCGCTCCCTGAGCCAGAACCCCCCCGTGCTGAAG
AGGCACCAGAGGGAGATCACCAGGACCACCCTGCAGAGCGACCAGGAGGAGATCGACTATGATGAC
ACCATCAGCGTGGAGATGAAGAAGGAGGACTTCGACATCTACGACGAGGACGAGAACCAGAGCCCC
AGGAGCTTCCAGAAGAAGACCAGGCACTACTTCATCGCCGCCGTGGAGAGGCTGTGGGACTATGGC
ATGAGCAGCAGCCCCCACGTGCTGAGGAACAGGGCCCAGAGCGGCAGCGTGCCCCAGTTCAAGAAG

（续）

图40A

GTGGTGTTCCAGGAGTTCACCGACGGCAGCTTCACCCAGCCCCTGTACAGAGGCGAGCTGAACGAG
CACCTGGGCCTGCTGGGCCCCTACATCAGGGCCGAGGTGGAGGACAACATCATGGTGACCTTCAGG
AACCAGGCCAGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCTACGAGGAGGACCAGAGGCAG
GGCGCCGAGCCCAGGAAGAACTTCGTGAAGCCCAACGAGACCAAGACCTACTTCTGGAAGGTGCAG
CACCACATGGCCCCCACCAAGGACGAGTTCGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGAC
CTGGAGAAGGACGTGCACAGCGGCCTGATCGGCCCCCTGCTGGTGTGCCACACCAACACCCTGAAC
CCCGCCCACGGCAGGCAGGTGACCGTGCAGGAGTTCGCCCTGTTCTTCACCATCTTCGACGAGACC
AAGAGCTGGTACTTCACCGAGAACATGGAGAGGAACTGCAGGGCCCCCTGCAACATCCAGATGGAG
GACCCCACCTTCAAGGAGAACTACAGGTTCCACGCCATCAACGGCTACATCATGGACACCCTGCCC
GGCCTGGTGATGGCCCAGGACCAGAGGATCAGGTGGTATCTGCTGAGCATGGGCAGCAACGAGAAC
ATCCACAGCATCCACTTCAGCGGCCACGTGTTCACCGTGAGGAAGAAGGAGGAGTACAAGATGGCC
CTGTACAACCTGTACCCCGGCGTGTTCGAGACCGTGGAGATGCTGCCCAGCAAGGCCGGCATCTGG
AGGGTGGAGTGCCTGATCGGCGAGCACCTGCACGCCGGCATGAGCACCCTGTTCCTGGTGTACAGC
AACAAGTGCCAGACCCCCCTGGGCATGGCCAGCGGCCACATCAGGGACTTCCAGATCACCGCCTCT
GGCCAGTACGGCCAGTGGGCCCCCAAGCTGGCCAGGCTGCACTACAGCGGCAGCATCAACGCCTGG
AGCACCAAGGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCCATGATCATCCACGGCATC
AAGACCCAGGGCGCCAGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGC
CTGGACGGCAAGAAGTGGCAGACCTACAGGGGCAACAGCACCGGCACCCTGATGGTGTTCTTCGGC
AACGTGGACAGCAGCGGCATCAAGCACAACATCTTCAACCCCCCCATCATCGCCAGGTACATCAGG
CTGCACCCCACCCACTACAGCATCAGGAGCACCCTGCGGATGGAACTGATGGGCTGCGACCTGAAC
AGCTGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATCTCTGACGCCCAGATCACCGCCAGCAGC
TACTTCACCAACATGTTCGCCACCTGGAGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGC
AACGCCTGGAGGCCCCAGGTGAACAACCCCAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATG
AAGGTGACCGGCGTGACCACCCAGGGCGTGAAGAGCCTGCTGACCAGCATGTACGTGAAGGAGTTC
CTGATCAGCAGCAGCCAGGACGGCCACCAGTGGACCCTGTTCTTCCAGAACGGCAAAGTGAAGGTG
TTCCAGGGCAACCAGGACAGCTTCACCCCCGTGGTGAACAGCCTGGACCCCCCCCCTGCTGACCAGG
TATCTGAGGATCCACCCCCAGAGCTGGGTGCACCAGATCGCCCTGAGAATGGAAGTGCTGGGATGC
GAGGCCCAGGACCTGTACTGA (SEQ ID NO:100)

图40B

CS23m2-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGCGCCGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGCGAGCTGCCT
GTGGACGCCAGGTTCCCCCCCAGAGTGCCCAAGAGCTTCCCCTTCAACACCTCAGTGGTGTACAAG
AAGACCCTGTTCGTGGAGTTCACCGACCACCTGTTCAACATCGCCAAGCCCAGGCCCCCCTGGATG
GGCCTGCTGGGCCCCACCATCCAGGCCGAGGTGTACGACACCGTGGTGGTCACCCTGAAGAACATG
GCCAGCCACCCCGTGAGCCTGCACGCCGTGGGCGTGAGCTACTGGAAGTCCTCTGAGGGCGCCGAG
TATGACGACCAGACCAGCCAGAGGGAGAAGGAGGACGACAAGGTGTTCCCCGGCAAGAGCCACACC
TACGTGTGGCAGGTGCTGAAGGAGAACGGCCCCACTGCCAGCGACCCCCCCTGCCTGACCTACAGC
TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACTCTGGCCTGATCGGCGCCCTGCTGGTGTGC
AGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCCTGCACAAGTTCATCCTGCTGTTCGCCGTG
TTCGATGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAGGACAGGGATGCCGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACCGTGAACGGCTACGTGAACAGGAGCCTGCCCGGCCTG
ATCGGCTGCCACAGGAAGTCTGTGTACTGGCACGTGATCGGCATGGGCACCACCCCCGAGGTGCAC
AGCATCTTCCTGGAGGGCCACACCTTCCTGGTGAGGAACCACAGGCAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCCTGACCGCCCAGACCCTGCTGATGGACCTGGGCCAGTTCCTGCTGTTCTGCCAC
ATCAGCAGCCACCAGCACGACGGCATGGAGGCCTACGTGAAGGTGGACAGCTGCCCCGAGGAGCCC
CAGCTGAGGATGAAGAACAACGAGGAGGCCGAGGACTATGATGATGACCTGACCGACTCTGAGATG
GACGTGGTGAGGTTTGATGATGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCCAAGAAG
CACCCCAAGACCTGGGTGCACTACATCGCCGCCGAGGAGGAGGACTGGGACTACGCCCCCCTGGTG
CTGGCCCCCGACGACAGGAGCTACAAGAGCCAGTACCTGAACAACGGCCCCCAGAGGATCGGCAGG
AAGTACAAGAAGGTCAGATTCATGGCCTACACCGACGAGACCTTCAAGACCAGGGAGGCCATCCAG
CACGAGTCTGGCATCCTGGGCCCCCTGCTGTACGGCGAGGTGGGCGACACCCTGCTGATCATCTTC
AAGAACCAGGCCAGCAGGCCCTACAACATCTACCCCCACGGCATCACCGATGTGAGGCCCCTGTAC
AGCAGGAGGCTGCCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCCTGCCCGGCGAGATCTTC
AAGTACAAGTGGACCGTGACCGTGGAGGATGGCCCCACCAAGTCTGACCCCAGGTGCCTGACCAGG
TACTACAGCAGCTTCGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATCGGCCCCCTGCTGATC
TGCTACAAGGAGAGCGTGGACCAGAGGGGCAACCAGATCATGTCTGACAAGAGGAACGTGATCCTG
TTCTCTGTGTTCGATGAGAACAGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCCTGCCCAAC
CCCGCCGGCGTGCAGCTGGAGGACCCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCAACGGC
TACGTGTTCGACAGCCTGCAGCTGTCTGTGTGCCTGCACGAGGTGGCCTACTGGTACATCCTGAGC
ATCGGCGCCCAGACCGACTTCCTGTCTGTGTTCTTCTCTGGCTACACCTTCAAGCACAAGATGGTG
TACGAGGACACCCTGACCCTGTTCCCCTTCAGCGGCGAGACCGTGTTCATGAGCATGGAGAACCCC
GGCCTGTGGATCCTGGGCTGCCACAACAGCGACTTCAGGAACAGGGGCATGACCGCCCTGCTGAAA
GTCAGCAGCTGCGACAAGAACACCGGCGACTACTACGAGGACAGCTACGAGGACATCAGCGCCTAC
CTGCTGAGCAAGAACAACGCCATCGAGCCCAGGAGCTTCAGCCAGAACCCCCCCGTGCTGAAGAGG
CACCAGAGGGAGATCACCAGGACCACCCTGCAGAGCGACCAGGAGGAGATCGACTATGATGACACC
ATCAGCGTGGAGATGAAGAAGGAGGACTTCGACATCTACGACGAGGACGAGAACCAGAGCCCCAGG
AGCTTCCAGAAGAAGACCAGGCACTACTTCATCGCCGCCGTGGAGAGGCTGTGGGACTATGGCATG
AGCAGCAGCCCCCACGTGCTGAGGAACAGGGCCCAGAGCGGCAGCGTGCCCCAGTTCAAGAAGGTG

（续）

图41A

GTGTTCCAGGAGTTCACCGACGGCAGCTTCACCCAGCCCCTGTACAGAGGCGAGCTGAACGAGCAC
CTGGGCCTGCTGGGCCCCTACATCAGGGCCGAGGTGGAGGACAACATCATGGTGACCTTCAGGAAC
CAGGCCAGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCTACGAGGAGGACCAGAGGCAGGGC
GCCGAGCCCAGGAAGAACTTCGTGAAGCCCAACGAGACCAAGACCTACTTCTGGAAGGTGCAGCAC
CACATGGCCCCCACCAAGGACGAGTTCGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGACCTG
GAGAAGGACGTGCACAGCGGCCTGATCGGCCCCCTGCTGGTGTGCCACACCAACACCCTGAACCCC
GCCCACGGCAGGCAGGTGACCGTGCAGGAGTTCGCCCTGTTCTTCACCATCTTCGACGAGACCAAG
AGCTGGTACTTCACCGAGAACATGGAGAGGAACTGCAGGGCCCCCTGCAACATCCAGATGGAGGAC
CCCACCTTCAAGGAGAACTACAGGTTCCACGCCATCAACGGCTACATCATGGACACCCTGCCCGGC
CTGGTGATGGCCCAGGACCAGAGGATCAGGTGGTATCTGCTGAGCATGGGCAGCAACGAGAACATC
CACAGCATCCACTTCAGCGGCCACGTGTTCACCGTGAGGAAGAAGGAGGAGTACAAGATGGCCCTG
TACAACCTGTACCCCGGCGTGTTCGAGACCGTGGAGATGCTGCCCAGCAAGGCCGGCATCTGGAGG
GTGGAGTGCCTGATCGGCGAGCACCTGCACGCCGGCATGAGCACCCTGTTCCTGGTGTACAGCAAC
AAGTGCCAGACCCCCCTGGGCATGGCCAGCGGCCACATCAGGGACTTCCAGATCACCGCCTCTGGC
CAGTACGGCCAGTGGGCCCCCAAGCTGGCCAGGCTGCACTACAGCGGCAGCATCAACGCCTGGAGC
ACCAAGGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCCATGATCATCCACGGCATCAAG
ACCCAGGGCGCCAGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTG
GACGGCAAGAAGTGGCAGACCTACAGGGGCAACAGCACCGGCACCCTGATGGTGTTCTTCGGCAAC
GTGGACAGCAGCGGCATCAAGCACAACATCTTCAACCCCCCCATCATCGCCAGGTACATCAGGCTG
CACCCCACCCACTACAGCATCAGGAGCACCCTGCGGATGGAACTGATGGGCTGCGACCTGAACAGC
TGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATCTCTGACGCCCAGATCACCGCCAGCAGCTAC
TTCACCAACATGTTCGCCACCTGGAGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGCAAC
GCCTGGAGGCCCCAGGTGAACAACCCCAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATGAAG
GTGACCGGCGTGACCACCCAGGGCGTGAAGAGCCTGCTGACCAGCATGTACGTGAAGGAGTTCCTG
ATCAGCAGCAGCCAGGACGGCCACCAGTGGACCCTGTTCTTCCAGAACGGCAAAGTGAAGGTGTTC
CAGGGCAACCAGGACAGCTTCACCCCCGTGGTGAACAGCCTGGACCCCCCCCCTGCTGACCAGGTAT
CTGAGGATCCACCCCCAGAGCTGGGTGCACCAGATCGCCCTGAGAATGGAAGTGCTGGGATGCGAG
GCCCAGGACCTGTACTGA (SEQ ID NO:101)

图41B

CS23m1-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGCGCCGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGCGAGCTGCCT
GTGGACGCCAGGTTCCCCCCCAGAGTGCCCAAGAGCTTCCCCTTCAACACCTCAGTGGTGTACAAG
AAGACCCTGTTCGTGGAGTTCACCGACCACCTGTTCAACATCGCCAAGCCCAGGCCCCCCTGGATG
GGCCTGCTGGGCCCCACCATCCAGGCCGAGGTGTACGACACCGTGGTGATCACCCTGAAGAACATG
GCCAGCCACCCCGTGAGCCTGCACGCCGTGGGCGTGAGCTACTGGAAGGCCTCTGAGGGCGCCGAG
TATGACGACCAGACCAGCCAGAGGGAGAAGGAGGACGACAAGGTGTTCCCCGGCGGCAGCCACACC
TACGTGTGGCAGGTGCTGAAGGAGAACGGCCCCATGGCCAGCGACCCCCTGTGCCTGACCTACAGC
TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACTCTGGCCTGATCGGCGCCCTGCTGGTGTGC
AGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCCTGCACAAGTTCATCCTGCTGTTCGCCGTG
TTCGATGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAGGACAGGGATGCCGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACCGTGAACGGCTACGTGAACAGGAGCCTGCCCGGCCTG
ATCGGCTGCCACAGGAAGTCTGTGTACTGGCACGTGATCGGCATGGGCACCACCCCCGAGGTGCAC
AGCATCTTCCTGGAGGGCCACACCTTCCTGGTGAGGAACCACAGGCAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCCTGACCGCCCAGACCCTGCTGATGGACCTGGGCCAGTTCCTGCTGTCCTGCCAC
ATCAGCAGCCACCAGCACGACGGCATGGAGGCCTACGTGAAGGTGGACAGCTGCCCCGAGGAGCCC
CAGCTGAGGATGAAGAACAACGAGGAGGCCGAGGACTATGATGATGACCTGACCGACTCTGAGATG
GACGTGGTGAGGTTTGATGATGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCCAAGAAG
CACCCCAAGACCTGGGTGCACTACATCGCCGCCGAGGAGGAGGACTGGGACTACGCCCCCCTGGTG
CTGGCCCCCGACGACAGGAGCTACAAGAGCCAGTACCTGAACAACGGCCCCCAGAGGATCGGCAGG
AAGTACAAGAAGGTCAGATTCATGGCCTACACCGACGAGACCTTCAAGACCAGGGAGGCCATCCAG
CACGAGTCTGGCATCCTGGGCCCCCTGCTGTACGGCGAGGTGGGCGACACCCTGCTGATCATCTTC
AAGAACCAGGCCAGCAGGCCCTACAACATCTACCCCCACGGCATCACCGATGTGAGGCCCCTGTAC
AGCAGGAGGCTGCCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCCTGCCCGGCGAGATCTTC
AAGTACAAGTGGACCGTGACCGTGGAGGATGGCCCCACCAAGTCTGACCCCAGGTGCCTGACCAGG
TACTACAGCAGCTTCGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATCGGCCCCCTGCTGATC
TGCTACAAGGAGAGCGTGGACCAGAGGGGCAACCAGATCATGTCTGACAAGAGGAACGTGATCCTG
TTCTCTGTGTTCGATGAGAACAGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCCTGCCCAAC
CCCGCCGGCGTGCAGCTGGAGGACCCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCAACGGC
TACGTGTTCGACAGCCTGCAGCTGTCTGTGTGCCTGCACGAGGTGGCCTACTGGTACATCCTGAGC
ATCGGCGCCCAGACCGACTTCCTGTCTGTGTTCTTCTCTGGCTACACCTTCAAGCACAAGATGGTG
TACGAGGACACCCTGACCCTGTTCCCCTTCAGCGGCGAGACCGTGTTCATGAGCATGGAGAACCCC
GGCCTGTGGATCCTGGGCTGCCACAACAGCGACTTCAGGAACAGGGGCATGACCGCCCTGCTGAAA
GTCAGCAGCTGCGACAAGAACACCGGCGACTACTACGAGGACAGCTACGAGGACATCAGCGCCTAC
CTGCTGAGCAAGAACAACGCCATCGAGCCCAGGAGCTTCAGCCAGAACCCCCCCGTGCTGAAGAGG
CACCAGAGGGAGATCACCAGGACCACCCTGCAGAGCGACCAGGAGGAGATCGACTATGATGACACC
ATCAGCGTGGAGATGAAGAAGGAGGACTTCGACATCTACGACGAGGACGAGAACCAGAGCCCCAGG
AGCTTCCAGAAGAAGACCAGGCACTACTTCATCGCCGCCGTGGAGAGGCTGTGGGACTATGGCATG
AGCAGCAGCCCCCACGTGCTGAGGAACAGGGCCCAGAGCGGCAGCGTGCCCCAGTTCAAGAAGGTG

（续）

图42A

GTGTTCCAGGAGTTCACCGACGGCAGCTTCACCCAGCCCCTGTACAGAGGCGAGCTGAACGAGCAC
CTGGGCCTGCTGGGCCCCTACATCAGGGCCGAGGTGGAGGACAACATCATGGTGACCTTCAGGAAC
CAGGCCAGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCTACGAGGAGGACCAGAGGCAGGGC
GCCGAGCCCAGGAAGAACTTCGTGAAGCCCAACGAGACCAAGACCTACTTCTGGAAGGTGCAGCAC
CACATGGCCCCCACCAAGGACGAGTTCGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGACCTG
GAGAAGGACGTGCACAGCGGCCTGATCGGCCCCCTGCTGGTGTGCCACACCAACACCCTGAACCCC
GCCCACGGCAGGCAGGTGACCGTGCAGGAGTTCGCCCTGTTCTTCACCATCTTCGACGAGACCAAG
AGCTGGTACTTCACCGAGAACATGGAGAGGAACTGCAGGGCCCCCTGCAACATCCAGATGGAGGAC
CCCACCTTCAAGGAGAACTACAGGTTCCACGCCATCAACGGCTACATCATGGACACCCTGCCCGGC
CTGGTGATGGCCCAGGACCAGAGGATCAGGTGGTATCTGCTGAGCATGGGCAGCAACGAGAACATC
CACAGCATCCACTTCAGCGGCCACGTGTTCACCGTGAGGAAGAAGGAGGAGTACAAGATGGCCCTG
TACAACCTGTACCCCGGCGTGTTCGAGACCGTGGAGATGCTGCCCAGCAAGGCCGGCATCTGGAGG
GTGGAGTGCCTGATCGGCGAGCACCTGCACGCCGGCATGAGCACCCTGTTCCTGGTGTACAGCAAC
AAGTGCCAGACCCCCCTGGGCATGGCCAGCGGCCACATCAGGGACTTCCAGATCACCGCCTCTGGC
CAGTACGGCCAGTGGGCCCCCAAGCTGGCCAGGCTGCACTACAGCGGCAGCATCAACGCCTGGAGC
ACCAAGGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCCATGATCATCCACGGCATCAAG
ACCCAGGGCGCCAGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTG
GACGGCAAGAAGTGGCAGACCTACAGGGGCAACAGCACCGGCACCCTGATGGTGTTCTTCGGCAAC
GTGGACAGCAGCGGCATCAAGCACAACATCTTCAACCCCCCCATCATCGCCAGGTACATCAGGCTG
CACCCCACCCACTACAGCATCAGGAGCACCCTGCGGATGGAACTGATGGGCTGCGACCTGAACAGC
TGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATCTCTGACGCCCAGATCACCGCCAGCAGCTAC
TTCACCAACATGTTCGCCACCTGGAGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGCAAC
GCCTGGAGGCCCCAGGTGAACAACCCCAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATGAAG
GTGACCGGCGTGACCACCCAGGGCGTGAAGAGCCTGCTGACCAGCATGTACGTGAAGGAGTTCCTG
ATCAGCAGCAGCCAGGACGGCCACCAGTGGACCCTGTTCTTCCAGAACGGCAAAGTGAAGGTGTTC
CAGGGCAACCAGGACAGCTTCACCCCCGTGGTGAACAGCCTGGACCCCCCCCCTGCTGACCAGGTAT
CTGAGGATCCACCCCCAGAGCTGGGTGCACCAGATCGCCCTGAGAATGGAAGTGCTGGGATGCGAG
GCCCAGGACCTGTACTGA (SEQ ID NO:102)

图42B

CS23m23-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGCGCCGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGCGAGCTGCCT
GTGGACGCCAGGTTCCCCCCCAGAGTGCCCAAGAGCTTCCCCTTCAACACCTCAGTGGTGTACAAG
AAGACCCTGTTCGTGGAGTTCACCGACCACCTGTTCAACATCGCCAAGCCCAGGCCCCCCTGGATG
GGCCTGCTGGGCCCCACCATCCAGGCCGAGGTGTACGACACCGTGGTGGTCACCCTGAAGAACATG
GCCAGCCACCCCGTGAGCCTGCACGCCGTGGGCGTGAGCTACTGGAAGTCCTCTGAGGGCGCCGAG
TATGACGACCAGACCAGCCAGAGGGAGAAGGAGGACGACAAGGTGTTCCCCGGCAAGAGCCACACC
TACGTGTGGCAGGTGCTGAAGGAGAACGGCCCCACTGCCAGCGACCCCCCCTGCCTGACCTACAGC
TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACTCTGGCCTGATCGGCGCCCTGCTGGTGTGC
AGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCCTGCACAAGTTCATCCTGCTGTTCGCCGTG
TTCGATGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAGGACAGGGATGCCGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACCGTGAACGGCTACGTGAACAGGAGCCTGCCCGGCCTG
ATCGGCTGCCACAGGAAGTCTGTGTACTGGCACGTGATCGGCATGGGCACCACCCCCGAGGTGCAC
AGCATCTTCCTGGAGGGCCACACCTTCCTGGTGAGGAACCACAGGCAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCCTGACCGCCCAGACCCTGCTGATGGACCTGGGCCAGTTCCTGCTGTTCTGCCAC
ATCAGCAGCCACCAGCACGACGGCATGGAGGCCTACGTGAAGGTGGACAGCTGCCCCGAGGAGCCC
CAGCTGAGGATGAAGAACAACGAGGAGGCCGAGGACTATGATGATGACCTGACCGACTCTGAGATG
GACGTGGTGAGGTTTGATGATGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCCAAGAAG
CACCCCAAGACCTGGGTGCACTACATCGCCGCCGAGGAGGAGGACTGGGACTACGCCCCCCTGGTG
CTGGCCCCCGACGACAGGAGCTACAAGAGCCAGTACCTGAACAACGGCCCCCAGAGGATCGGCAGG
AAGTACAAGAAGGTCAGATTCATGGCCTACACCGACGAGACCTTCAAGACCAGGGAGGCCATCCAG
CACGAGTCTGGCATCCTGGGCCCCCTGCTGTACGGCGAGGTGGGCGACACCCTGCTGATCATCTTC
AAGAACCAGGCCAGCAGGCCCTACAACATCTACCCCCACGGCATCACCGATGTGAGGCCCCTGTAC
AGCAGGAGGCTGCCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCCTGCCCGGCGAGATCTTC
AAGTACAAGTGGACCGTGACCGTGGAGGATGGCCCCACCAAGTCTGACCCCAGGTGCCTGACCAGG
TACTACAGCAGCTTCGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATCGGCCCCCTGCTGATC
TGCTACAAGGAGAGCGTGGACCAGAGGGGCAACCAGATCATGTCTGACAAGAGGAACGTGATCCTG
TTCTCTGTGTTCGATGAGAACAGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCCTGCCCAAC
CCCGCCGGCGTGCAGCTGGAGGACCCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCAACGGC
TACGTGTTCGACAGCCTGCAGCTGTCTGTGTGCCTGCACGAGGTGGCCTACTGGTACATCCTGAGC
ATCGGCGCCCAGACCGACTTCCTGTCTGTGTTCTTCTCTGGCTACACCTTCAAGCACAAGATGGTG
TACGAGGACACCCTGACCCTGTTCCCCTTCAGCGGCGAGACCGTGTTCATGAGCATGGAGAACCCC
GGCCTGTGGATCCTGGGCTGCCACAACAGCGACTTCAGGAACAGGGGCATGACCGCCCTGCTGAAA
GTCAGCAGCTGCGACAAGAACACCGGCGACTACTACGAGGACAGCTACGAGGACATCAGCGCCTAC
CTGCTGAGCAAGAACAACACCACCTACGTGAACCGCTCCCTGAGCCAGAACCCCCCCGTGCTGAAG
AGGCACCAGAGGGAGATCACCAGGACCACCCTGCAGAGCGACCAGGAGGAGATCGACTATGATGAC
ACCATCAGCGTGGAGATGAAGAAGGAGGACTTCGACATCTACGACGAGGACGAGAACCAGAGCCCC
AGGAGCTTCCAGAAGAAGACCAGGCACTACTTCATCGCCGCCGTGGAGAGGCTGTGGGACTATGGC
ATGAGCAGCAGCCCCCACGTGCTGAGGAACAGGGCCCAGAGCGGCAGCGTGCCCCAGTTCAAGAAG

（续）

图43A

GTGGTGTTCCAGGAGTTCACCGACGGCAGCTTCACCCAGCCCCTGTACAGAGGCGAGCTGAACGAG
CACCTGGGCCTGCTGGGCCCCTACATCAGGGCCGAGGTGGAGGACAACATCATGGTGACCTTCAGG
AACCAGGCCAGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCTACGAGGAGGACCAGAGGCAG
GGCGCCGAGCCCAGGAAGAACTTCGTGAAGCCCAACGAGACCAAGACCTACTTCTGGAAGGTGCAG
CACCACATGGCCCCCACCAAGGACGAGTTCGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGAC
CTGGAGAAGGACGTGCACAGCGGCCTGATCGGCCCCCTGCTGGTGTGCCACACCAACACCCTGAAC
CCCGCCCACGGCAGGCAGGTGACCGTGCAGGAGTTCGCCCTGTTCTTCACCATCTTCGACGAGACC
AAGAGCTGGTACTTCACCGAGAACATGGAGAGGAACTGCAGGGCCCCCTGCAACATCCAGATGGAG
GACCCCACCTTCAAGGAGAACTACAGGTTCCACGCCATCAACGGCTACATCATGGACACCCTGCCC
GGCCTGGTGATGGCCCAGGACCAGAGGATCAGGTGGTATCTGCTGAGCATGGGCAGCAACGAGAAC
ATCCACAGCATCCACTTCAGCGGCCACGTGTTCACCGTGAGGAAGAAGGAGGAGTACAAGATGGCC
CTGTACAACCTGTACCCCGGCGTGTTCGAGACCGTGGAGATGCTGCCCAGCAAGGCCGGCATCTGG
AGGGTGGAGTGCCTGATCGGCGAGCACCTGCACGCCGGCATGAGCACCCTGTTCCTGGTGTACAGC
AACAAGTGCCAGACCCCCCTGGGCATGGCCAGCGGCCACATCAGGGACTTCCAGATCACCGCCTCT
GGCCAGTACGGCCAGTGGGCCCCCAAGCTGGCCAGGCTGCACTACAGCGGCAGCATCAACGCCTGG
AGCACCAAGGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCCATGATCATCCACGGCATC
AAGACCCAGGGCGCCAGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGC
CTGGACGGCAAGAAGTGGCAGACCTACAGGGGCAACAGCACCGGCACCCTGATGGTGTTCTTCGGC
AACGTGGACAGCAGCGGCATCAAGCACAACATCTTCAACCCCCCCATCATCGCCAGGTACATCAGG
CTGCACCCCACCCACTACAGCATCAGGAGCACCCTGCGGATGGAACTGATGGGCTGCGACCTGAAC
AGCTGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATCTCTGACGCCCAGATCACCGCCAGCAGC
TACTTCACCAACATGTTCGCCACCTGGAGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGC
AACGCCTGGAGGCCCCAGGTGAACAACCCCAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATG
AAGGTGACCGGCGTGACCACCCAGGGCGTGAAGAGCCTGCTGACCAGCATGTACGTGAAGGAGTTC
CTGATCAGCAGCAGCCAGGACGGCCACCAGTGGACCCTGTTCTTCCAGAACGGCAAAGTGAAGGTG
TTCCAGGGCAACCAGGACAGCTTCACCCCCGTGGTGAACAGCCTGGACCCCCCCCCTGCTGACCAGG
TATCTGAGGATCCACCCCCAGAGCTGGGTGCACCAGATCGCCCTGAGAATGGAAGTGCTGGGATGC
GAGGCCCAGGACCTGTACTGA (SEQ ID NO:103)

图43B

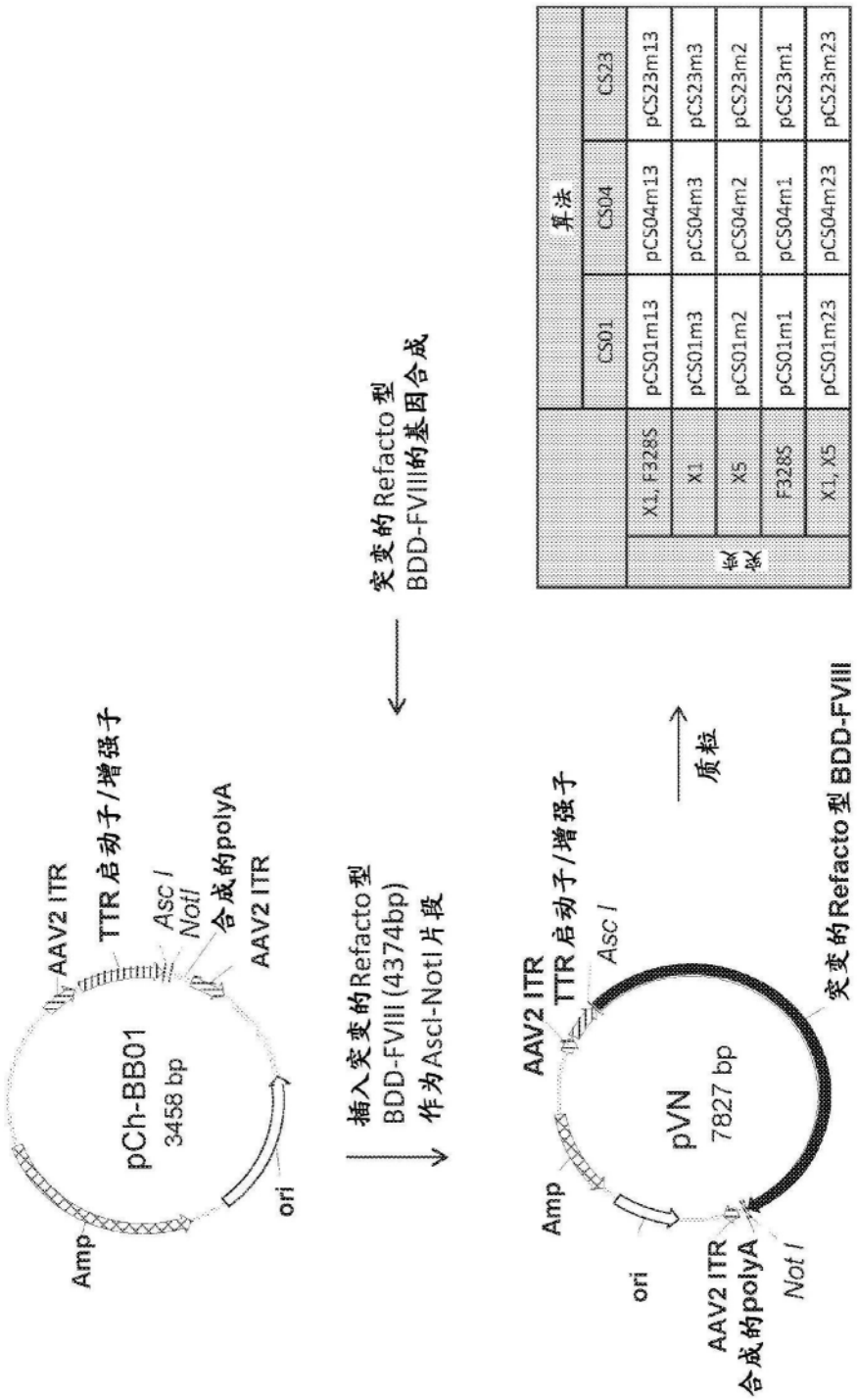

AAV2 ITR
TTR 启动子/增强子
Asc I
NotI
合成的polyA
AAV2 ITR
pCh-BB01
3458 bp
Amp
ori
插入突变的Refacto型
BDD-FVIII(4374bp)
作为AscI-NotI片段
突变的 Refacto 型
BDD-FVIII的基因合成
AAV2 ITR
TTR 启动子/增强子
Asc I
pVN
7827 bp
Amp
ori
AAV2 ITR
合成的polyA
Not I
突变的Refacto型 BDD-FVIII
质粒
算法
CS01
CS04
CS23
X1, F328S
X1
X5
F328S
X1, X5
pCS01m13
pCS01m3
pCS01m2
pCS01m1
pCS01m23
pCS04m13
pCS04m3
pCS04m2
pCS04m1
pCS04m23
pCS23m13
pCS23m3
pCS23m2
pCS23m1
pCS23m23

图44

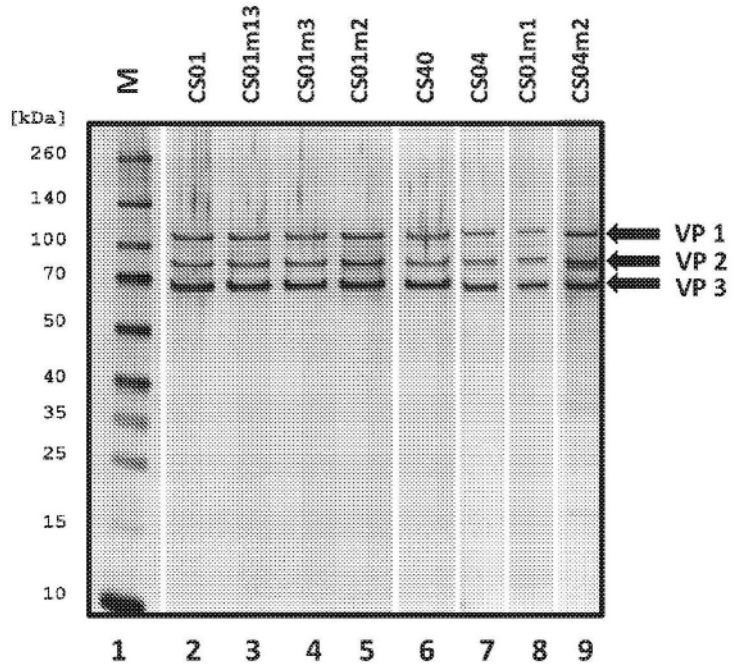

M
CS01
CS01m13
CS01m3
CS01m2
CS40
CS04
CS01m1
CS04m2
[kDa]
260
140
100
70
50
40
35
25
15
10
VP 1
VP 2
VP 3
1 2 3 4 5 6 7 8 9

图45

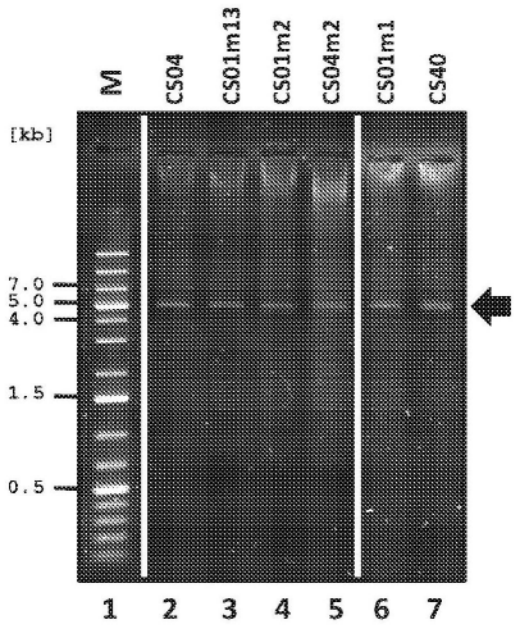

M
CS04
CS01m13
CS01m2
CS04m2
CS01m1
CS40
[kb]
7.0
5.0
4.0
1.5
0.5
1 2 3 4 5 6 7

图46

CS01-HC-NA

gcc
accaggagat actacctggg ggctgtggaa ctttcttggg actacatgca gtctgacctg
ggagagctgc ctgtggatgc caggttccca cccagagtgc ccaagtcctt cccattcaac
acctctgtgg tctacaagaa gacactcttt gtggaattca ctgaccacct gttcaacatt
gcaaaaccca gaccaccctg gatgggactc ctgggaccca ccattcaggc tgaggtgtat
gacactgtgg tcatcaccct caagaacatg gcatcccacc ctgtgtctct gcatgctgtg
ggagtctcat actggaaagc ctctgaaggg gctgagtatg atgaccagac atcccagaga
gagaaagagg atgacaaggt gttccctggg ggatctcaca cctatgtgtg gcaagtcctc
aaggagaatg gacccatggc atctgaccca ctctgcctga catactccta cctttctcat
gtggacctgg tcaaggacct caactctgga ctgattgggg cactgctggt gtgcagggaa
ggatccctgg ccaaggagaa aacccagaca ctgcacaagt tcattctcct gtttgctgtc
tttgatgagg gcaagtcttg gcactctgaa acaaagaact ccctgatgca agacagggat
gctgcctctg ccagggcatg gcccaagatg cacactgtga atggctatgt gaacagatca
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtgat tggcatgggg
acaacccctg aagtgcactc cattttcctg gagggacaca ccttcctggt caggaaccac
agacaagcct ctctggagat ctctcccatc accttcctca ctgcacagac actgctgatg
gaccttggac agttcctgct gttctgccac atctcttccc accagcatga tggcatggaa
gcctatgtca aggtggactc atgccctgag gaaccacagc tcaggatgaa gaacaatgag
gaggctgagg actatgatga tgacctgact gactctgaga tggatgtggt cagatttgat
gatgacaact ctccatcctt cattcagatc aggtctgtgg caaagaaaca ccccaagaca
tgggtgcact acattgctgc tgaggaagag gactgggact atgcaccact ggtcctggcc
cctgatgaca ggagctacaa gtctcagtac ctcaacaatg gcccacaaag aattggaaga
aagtacaaga aagtcagatt catggcctac actgatgaaa ccttcaagac aagagaagcc
attcagcatg agtctggcat tctgggacca ctcctgtatg gggaagtggg agacaccctg
ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctacccaca tggcatcact
gatgtcaggc ccctgtacag caggagactg ccaaaagggg tgaaacacct caaggacttc
cccattctgc ctggagagat cttcaagtac aagtggactg tcactgtgga ggatggacca
acaaagtctg accccaggtg cctcaccaga tactactcct cttttgtgaa catggagaga
gacctggcat ctggactgat tggaccactg ctcatctgct acaaggagtc tgtggaccag
agaggcaacc agatcatgtc tgacaagaga aatgtgattc tgttctctgt ctttgatgag
aacagatcat ggtacctgac tgagaacatt cagagattcc tgcccaaccc tgctggggtg
caactggaag accctgagtt ccaggcaagc aacatcatgc actccatcaa tggctatgtg
tttgactctc tccagctttc tgtctgcctg catgaggtgg cctactggta cattctttct
attggggcac aaactgactt cctttctgtc ttcttctctg gatacacctt caagcacaag
atggtgtatg aggacaccct gacactcttc ccattctctg gggaaactgt gttcatgagc
atggagaacc ctggactgtg gattctggga tgccacaact ctgacttcag aaacagggga
atgactgcac tgctcaaagt ctcctcctgt gacaagaaca ctggggacta ctatgaggac
tcttatgagg acatctctgc ctacctgctc agcaagaaca atgccattga gcccaga
(SEQ ID NO:24X)

图47

CS01-LC-NA

g agatcaccag gacaaccctc
cagtctgacc aggaagagat tgactatgat gacaccattt ctgtggagat gaagaaggag
gactttgaca tctatgatga ggacgagaac cagtctccaa gatcattcca gaagaagaca
agacactact tcattgctgc tgtggaaaga ctgtgggact atggcatgtc ttcctctccc
catgtcctca ggaacagggc acagtctggc tctgtgccac agttcaagaa agtggtcttc
caggagttca ctgatggctc attcacccag cccctgtaca gaggggaact gaatgagcac
ctgggactcc tgggaccata catcagggct gaggtggaag acaacatcat ggtgacattc
agaaaccagg cctccaggcc ctacagcttc tactcttccc tcatcagcta tgaggaagac
cagagacaag gggctgagcc aagaaagaac tttgtgaaac ccaatgaaac caagacctac
ttctggaaag tccagcacca catggcaccc accaaggatg agtttgactg caaggcctgg
gcatacttct ctgatgtgga cctggagaaa gatgtgcact ctggcctgat tggcccactc
ctggtctgcc acaccaacac cctgaaccct gcacatggaa ggcaagtgac tgtgcaggag
tttgccctct tcttcaccat ctttgatgaa accaagtcat ggtacttcac tgagaacatg
gagagaaact gcagagcacc atgcaacatt cagatggaag accccacctt caaggagaac
tacaggttcc atgccatcaa tggctacatc atggacaccc tgcctgggct tgtcatggca
caggaccaga gaatcagatg gtacctgctt tctatgggat ccaatgagaa cattcactcc
atccacttct ctgggcatgt cttcactgtg agaaagaagg aggaatacaa gatggccctg
tacaacctct accctggggt ctttgagact gtggagatgc tgccctccaa agctggcatc
tggagggtgg aatgcctcat tggggagcac ctgcatgctg gcatgtcaac cctgttcctg
gtctacagca acaagtgcca gacacccctg ggaatggcct ctggccacat cagggacttc
cagatcactg cctctggcca gtatggccag tgggcaccca aactggccag gctccactac
tctggctcca tcaatgcatg gtcaaccaag gagccattct cttggatcaa ggtggacctg
ctggcaccca tgatcattca tggcatcaag acacaggggg caagacagaa attctcctct
ctgtacatct cacagttcat catcatgtac tctctggatg gcaagaagtg gcagacatac
agaggcaact ccactggcac cctcatggtc ttctttggca atgtggacag ctctggcatc
aagcacaaca tcttcaaccc tcccatcatt gccagataca tcaggctgca ccccacccac
tactcaatca gatcaaccct caggatggaa ctgatgggat gtgacctgaa ctcctgctca
atgcccctgg gaatggagag caaggccatt tctgatgccc agatcactgc atcctcttac
ttcaccaaca tgtttgccac ctggtcacca tcaaaagcca ggctgcacct ccagggaaga
agcaatgcct ggagacccca ggtcaacaac ccaaaggaat ggctgcaagt ggacttccag
aagacaatga aagtcactgg ggtgacaacc caggggggtca agtctctgct cacctcaatg
tatgtgaagg agttcctgat ctcttcctca caggatggcc accagtggac actcttcttc
cagaatggca aagtcaaggt gttccagggc aaccaggact ctttcacacc tgtggtgaac
tcactggacc ccccctcct gacaagatac ctgagaattc acccccagtc ttgggtccac
cagattgccc tgagaatgga agtcctggga tgtgaggcac aagacctgta c
(SEQ ID NO:25)

图48

CS01Δ(760-1667) - CS01-SC1-NA

ATGCAGATTGAGCTGTCCACCTGCTTCTTTCTGTGCCTGCTGAGATTCTGCTTCTCTGCCACCAGGAGATAC
TACCTGGGGGCTGTGGAACTTTCTTGGGACTACATGCAGTCTGACCTGGGAGAGCTGCCTGTGGATGCCAGG
TTCCCACCCAGAGTGCCCAAGTCCTTCCCATTCAACACCTCTGTGGTCTACAAGAAGACACTCTTTGTGGAA
TTCACTGACCACCTGTTCAACATTGCAAAACCCAGACCACCCTGGATGGGACTCCTGGGACCCACCATTCAG
GCTGAGGTGTATGACACTGTGGTCATCACCCTCAAGAACATGGCATCCCACCCTGTGTCTCTGCATGCTGTG
GGAGTCTCATACTGGAAAGCCTCTGAAGGGGCTGAGTATGATGACCAGACATCCCAGAGAGAGAAAGAGGAT
GACAAGGTGTTCCCTGGGGGATCTCACACCTATGTGTGGCAAGTCCTCAAGGAGAATGGACCCATGGCATCT
GACCCACTCTGCCTGACATACTCCTACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATT
GGGGCACTGCTGGTGTGCAGGGAAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCATTCTC
CTGTTTGCTGTCTTTGATGAGGGCAAGTCTTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGAT
GCTGCCTCTGCCAGGGCATGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAAGTGCACTCCATT
TTCCTGGAGGGACACACCTTCCTGGTCAGGAACCACAGACAAGCCTCTCTGGAGATCTCTCCCATCACCTTC
CTCACTGCACAGACACTGCTGATGGACCTTGGACAGTTCCTGCTGTTCTGCCACATCTCTTCCCACCAGCAT
GATGGCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCACAGCTCAGGATGAAGAACAATGAG
GAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATGGATGTGGTCAGATTTGATGATGACAACTCT
CCATCCTTCATTCAGATCAGGTCTGTGGCAAAGAAACACCCCAAGACATGGGTGCACTACATTGCTGCTGAG
GAAGAGGACTGGGACTATGCACCACTGGTCCTGGCCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAAC
AATGGCCCACAAAGAATTGGAAGAAAGTACAAGAAAGTCAGATTCATGGCCTACACTGATGAAACCTTCAAG
ACAAGAGAAGCCATTCAGCATGAGTCTGGCATTCTGGGACCACTCCTGTATGGGGAAGTGGGAGACACCCTG
CTCATCATCTTCAAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCC
CTGTACAGCAGGAGACTGCCAAAAGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGAGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCCAGGTGCCTCACCAGATACTAC
TCCTCTTTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATCTGCTACAAGGAG
TCTGTGGACCAGAGAGGCAACCAGATCATGTCTGACAAGAGAAATGTGATTCTGTTCTCTGTCTTTGATGAG
AACAGATCATGGTACCTGACTGAGAACATTCAGAGATTCCTGCCCAACCCTGCTGGGGTGCAACTGGAAGAC
CCTGAGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGCTATGTGTTTGACTCTCTCCAGCTTTCTGTC
TGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCTATTGGGGCACAAACTGACTTCCTTTCTGTCTTCTTC
TCTGGATACACCTTCAAGCACAAGATGGTGTATGAGGACACCCTGACACTCTTCCCATTCTCTGGGGAAACT
GTGTTCATGAGCATGGAGAACCCTGGACTGTGGATTCTGGGATGCCACAACTCTGACTTCAGAAACAGGGGA
ATGACTGCACTGCTCAAAGTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACTCTTATGAGGAC
ATCTCTGCCTACCTGCTCAGCAAGAACAATGCCATTGAGCCCAGAGAGATCACCAGGACAACCCTCCAGTCT
GACCAGGAAGAGATTGACTATGATGACACCATTTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTATGAT
GAGGACGAGAACCAGTCTCCAAGATCATTCCAGAAGAAGACAAGACACTACTTCATTGCTGCTGTGGAAAGA
CTGTGGGACTATGGCATGTCTTCCTCTCCCCATGTCCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAG
TTCAAGAAAGTGGTCTTCCAGGAGTTCACTGATGGCTCATTCACCCAGCCCCTGTACAGAGGGGAACTGAAT
GAGCACCTGGGACTCCTGGGACCATACATCAGGGCTGAGGTGGAAGACAACATCATGGTGACATTCAGAAAC
CAGGCCTCCAGGCCCTACAGCTTCTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAAGGGGCTGAG
CCAAGAAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCACCACATGGCACCC

（续）

图49A

ACCAAGGATGAGTTTGACTGCAAGGCCTGGGCATACTTCTCTGATGTGGACCTGGAGAAAGATGTGCACTCT
GGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAACCCTGCACATGGAAGGCAAGTGACTGTG
CAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACCAAGTCATGGTACTTCACTGAGAACATGGAGAGA
AACTGCAGAGCACCATGCAACATTCAGATGGAAGACCCCACCTTCAAGGAGAACTACAGGTTCCATGCCATC
AATGGCTACATCATGGACACCCTGCCTGGGCTTGTCATGGCACAGGACCAGAGAATCAGATGGTACCTGCTT
TCTATGGGATCCAATGAGAACATTCACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGAGAAAGAAGGAG
GAATACAAGATGGCCCTGTACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCT
GGCATCTGGAGGGTGGAATGCCTCATTGGGGAGCACCTGCATGCTGGCATGTCAACCCTGTTCCTGGTCTAC
AGCAACAAGTGCCAGACACCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGC
CAGTATGGCCAGTGGGCACCCAAACTGGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGGTCAACCAAG
GAGCCATTCTCTTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTCATGGCATCAAGACACAGGGGGCA
AGACAGAAATTCTCCTCTCTGTACATCTCACAGTTCATCATCATGTACTCTCTGGATGGCAAGAAGTGGCAG
ACATACAGAGGCAACTCCACTGGCACCCTCATGGTCTTCTTTGGCAATGTGGACAGCTCTGGCATCAAGCAC
AACATCTTCAACCCTCCCATCATTGCCAGATACATCAGGCTGCACCCCACCCACTACTCAATCAGATCAACC
CTCAGGATGGAACTGATGGGATGTGACCTGAACTCCTGCTCAATGCCCCTGGGAATGGAGAGCAAGGCCATT
TCTGATGCCCAGATCACTGCATCCTCTTACTTCACCAACATGTTTGCCACCTGGTCACCATCAAAAGCCAGG
CTGCACCTCCAGGGAAGAAGCAATGCCTGGAGACCCCAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGAC
TTCCAGAAGACAATGAAAGTCACTGGGGTGACAACCCAGGGGGTCAAGTCTCTGCTCACCTCAATGTATGTG
AAGGAGTTCCTGATCTCTTCCTCACAGGATGGCCACCAGTGGACACTCTTCTTCCAGAATGGCAAAGTCAAG
GTGTTCCAGGGCAACCAGGACTCTTTCACACCTGTGGTGAACTCACTGGACCCCCCCCTCCTGACAAGATAC
CTGAGAATTCACCCCCAGTCTTGGGTCCACCAGATTGCCCTGAGAATGGAAGTCCTGGGATGTGAGGCACAA
GACCTGTACTGA (SEQ ID NO:26)

图49B

CS01Δ(772-1667) - CS01-SC2-NA

ATGCAGATTGAGCTGTCCACCTGCTTCTTTCTGTGCCTGCTGAGATTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGCTGTGGAACTTTCTTGGGACTACATGCAGTCTGACCTGGGAGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAGTCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACACTCTTTGTGGAATTCACTGACCACCTGTTCAACATTGCAAAACCCAGACCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCATCACCCTCAAGAACATG
GCATCCCACCCTGTGTCTCTGCATGCTGTGGGAGTCTCATACTGGAAAGCCTCTGAAGGGGCTGAG
TATGATGACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGGGATCTCACACC
TATGTGTGGCAAGTCCTCAAGGAGAATGGACCCATGGCATCTGACCCACTCTGCCTGACATACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCACTGCTGGTGTGC
AGGGAAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGTCTTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGATGCTGCC
TCTGCCAGGGCATGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAAGTGCAC
TCCATTTTCCTGGAGGGACACACCTTCCTGGTCAGGAACCACAGACAAGCCTCTCTGGAGATCTCT
CCCATCACCTTCCTCACTGCACAGACACTGCTGATGGACCTTGGACAGTTCCTGCTGTTCTGCCAC
ATCTCTTCCCACCAGCATGATGGCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCAGATTTGATGATGACAACTCTCCATCCTTCATTCAGATCAGGTCTGTGGCAAAGAAA
CACCCCAAGACATGGGTGCACTACATTGCTGCTGAGGAAGAGGACTGGGACTATGCACCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAACAATGGCCCACAAAGAATTGGAAGA
AAGTACAAGAAAGTCAGATTCATGGCCTACACTGATGAAACCTTCAAGACAAGAGAAGCCATTCAG
CATGAGTCTGGCATTCTGGGACCACTCCTGTATGGGGAAGTGGGAGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCAGGAGACTGCCAAAAGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGAGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCCAGGTGCCTCACCAGA
TACTACTCCTCTTTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGAGGCAACCAGATCATGTCTGACAAGAGAAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTCAGAGATTCCTGCCCAAC
CCTGCTGGGGTGCAACTGGAAGACCCTGAGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACTCTCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCACAAACTGACTTCCTTTCTGTCTTCTTCTCTGGATACACCTTCAAGCACAAGATGGTG
TATGAGGACACCCTGACACTCTTCCCATTCTCTGGGGAAACTGTGTTCATGAGCATGGAGAACCCT
GGACTGTGGATTCTGGGATGCCACAACTCTGACTTCAGAAACAGGGGAATGACTGCACTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATGCCATTGAGCCCAGAAGCTTCTCTCAGAATTCCAGACACCCCAGCACC
AGGGAGATCACCAGGACAACCCTCCAGTCTGACCAGGAAGAGATTGACTATGATGACACCATTTCT
GTGGAGATGAAGAAGGAGGACTTTGACATCTATGATGAGGACGAGAACCAGTCTCCAAGATCATTC

（续）

图50A

CAGAAGAAGACAAGACACTACTTCATTGCTGCTGTGGAAAGACTGTGGGACTATGGCATGTCTTCC
TCTCCCCATGTCCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAAGAAAGTGGTCTTC
CAGGAGTTCACTGATGGCTCATTCACCCAGCCCCTGTACAGAGGGGAACTGAATGAGCACCTGGGA
CTCCTGGGACCATACATCAGGGCTGAGGTGGAAGACAACATCATGGTGACATTCAGAAACCAGGCC
TCCAGGCCCTACAGCTTCTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAAGGGGCTGAG
CCAAGAAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCACCACATG
GCACCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCATACTTCTCTGATGTGGACCTGGAGAAA
GATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAACCCTGCACAT
GGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACCAAGTCATGG
TACTTCACTGAGAACATGGAGAGAAACTGCAGAGCACCATGCAACATTCAGATGGAAGACCCCACC
TTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCTGGGCTTGTC
ATGGCACAGGACCAGAGAATCAGATGGTACCTGCTTTCTATGGGATCCAATGAGAACATTCACTCC
ATCCACTTCTCTGGGCATGTCTTCACTGTGAGAAAGAAGGAGGAATACAAGATGGCCCTGTACAAC
CTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGGGTGGAA
TGCCTCATTGGGGAGCACCTGCATGCTGGCATGTCAACCCTGTTCCTGGTCTACAGCAACAAGTGC
CAGACACCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGCCAGTAT
GGCCAGTGGGCACCCAAACTGGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGGTCAACCAAG
GAGCCATTCTCTTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTCATGGCATCAAGACACAG
GGGGCAAGACAGAAATTCTCCTCTCTGTACATCTCACAGTTCATCATCATGTACTCTCTGGATGGC
AAGAAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCTCATGGTCTTCTTTGGCAATGTGGAC
AGCTCTGGCATCAAGCACAACATCTTCAACCCTCCCATCATTGCCAGATACATCAGGCTGCACCCC
ACCCACTACTCAATCAGATCAACCCTCAGGATGGAACTGATGGGATGTGACCTGAACTCCTGCTCA
ATGCCCCTGGGAATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCATCCTCTTACTTCACC
AACATGTTTGCCACCTGGTCACCATCAAAAGCCAGGCTGCACCTCCAGGGAAGAAGCAATGCCTGG
AGACCCCAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAATGAAAGTCACT
GGGGTGACAACCCAGGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGAGTTCCTGATCTCT
TCCTCACAGGATGGCCACCAGTGGACACTCTTCTTCCAGAATGGCAAAGTCAAGGTGTTCCAGGGC
AACCAGGACTCTTTCACACCTGTGGTGAACTCACTGGACCCCCCCCTCCTGACAAGATACCTGAGA
ATTCACCCCCAGTCTTGGGTCCACCAGATTGCCCTGAGAATGGAAGTCCTGGGATGTGAGGCACAA
GACCTGTACTGA (SEQ ID NO:27)

图50B

CS23Δ(760-1667) - CS23-SC1-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGGAGATAC
TACCTGGGCGCCGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGCGAGCTGCCTGTGGACGCCAGG
TTCCCCCCCAGAGTGCCCAAGAGCTTCCCCTTCAACACCTCAGTGGTGTACAAGAAGACCCTGTTCGTGGAG
TTCACCGACCACCTGTTCAACATCGCCAAGCCCAGGCCCCCCTGGATGGGCCTGCTGGGCCCCACCATCCAG
GCCGAGGTGTACGACACCGTGGTGATCACCCTGAAGAACATGGCCAGCCACCCCGTGAGCCTGCACGCCGTG
GGCGTGAGCTACTGGAAGGCCTCTGAGGGCGCCGAGTATGACGACCAGACCAGCCAGAGGGAGAAGGAGGAC
GACAAGGTGTTCCCCGGCGGCAGCCACACCTACGTGTGGCAGGTGCTGAAGGAGAACGGCCCCATGGCCAGC
GACCCCCTGTGCCTGACCTACAGCTACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACTCTGGCCTGATC
GGCGCCCTGCTGGTGTGCAGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCCTGCACAAGTTCATCCTG
CTGTTCGCCGTGTTCGATGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAGGACAGGGAT
GCCGCCTCTGCCAGGGCCTGGCCCAAGATGCACACCGTGAACGGCTACGTGAACAGGAGCCTGCCCGGCCTG
ATCGGCTGCCACAGGAAGTCTGTGTACTGGCACGTGATCGGCATGGGCACCACCCCCGAGGTGCACAGCATC
TTCCTGGAGGGCCACACCTTCCTGGTGAGGAACCACAGGCAGGCCAGCCTGGAGATCAGCCCCATCACCTTC
CTGACCGCCCAGACCCTGCTGATGGACCTGGGCCAGTTCCTGCTGTTCTGCCACATCAGCAGCCACCAGCAC
GACGGCATGGAGGCCTACGTGAAGGTGGACAGCTGCCCCGAGGAGCCCCAGCTGAGGATGAAGAACAACGAG
GAGGCCGAGGACTATGATGATGACCTGACCGACTCTGAGATGGACGTGGTGAGGTTTGATGATGACAACAGC
CCCAGCTTCATCCAGATCAGGTCTGTGGCCAAGAAGCACCCCAAGACCTGGGTGCACTACATCGCCGCCGAG
GAGGAGGACTGGGACTACGCCCCCCTGGTGCTGGCCCCCGACGACAGGAGCTACAAGAGCCAGTACCTGAAC
AACGGCCCCCAGAGGATCGGCAGGAAGTACAAGAAGGTCAGATTCATGGCCTACACCGACGAGACCTTCAAG
ACCAGGGAGGCCATCCAGCACGAGTCTGGCATCCTGGGCCCCCTGCTGTACGGCGAGGTGGGCGACACCCTG
CTGATCATCTTCAAGAACCAGGCCAGCAGGCCCTACAACATCTACCCCCACGGCATCACCGATGTGAGGCCC
CTGTACAGCAGGAGGCTGCCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCCTGCCCGGCGAGATCTTC
AAGTACAAGTGGACCGTGACCGTGGAGGATGGCCCCACCAAGTCTGACCCCAGGTGCCTGACCAGGTACTAC
AGCAGCTTCGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATCGGCCCCCTGCTGATCTGCTACAAGGAG
AGCGTGGACCAGAGGGGCAACCAGATCATGTCTGACAAGAGGAACGTGATCCTGTTCTCTGTGTTCGATGAG
AACAGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCCTGCCCAACCCCGCCGGCGTGCAGCTGGAGGAC
CCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCAACGGCTACGTGTTCGACAGCCTGCAGCTGTCTGTG
TGCCTGCACGAGGTGGCCTACTGGTACATCCTGAGCATCGGCGCCCAGACCGACTTCCTGTCTGTGTTCTTC
TCTGGCTACACCTTCAAGCACAAGATGGTGTACGAGGACACCCTGACCCTGTTCCCCTTCAGCGGCGAGACC
GTGTTCATGAGCATGGAGAACCCCGGCCTGTGGATCCTGGGCTGCCACAACAGCGACTTCAGGAACAGGGGC
ATGACCGCCCTGCTGAAAGTCAGCAGCTGCGACAAGAACACCGGCGACTACTACGAGGACAGCTACGAGGAC
ATCAGCGCCTACCTGCTGAGCAAGAACAACGCCATCGAGCCCAGGGAGATCACCAGGACCACCCTGCAGAGC
GACCAGGAGGAGATCGACTATGATGACACCATCAGCGTGGAGATGAAGAAGGAGGACTTCGACATCTACGAC
GAGGACGAGAACCAGAGCCCCAGGAGCTTCCAGAAGAAGACCAGGCACTACTTCATCGCCGCCGTGGAGAGG
CTGTGGGACTATGGCATGAGCAGCAGCCCCCACGTGCTGAGGAACAGGGCCCAGAGCGGCAGCGTGCCCCAG
TTCAAGAAGGTGGTGTTCCAGGAGTTCACCGACGGCAGCTTCACCCAGCCCCTGTACAGAGGCGAGCTGAAC
GAGCACCTGGGCCTGCTGGGCCCCTACATCAGGGCCGAGGTGGAGGACAACATCATGGTGACCTTCAGGAAC
CAGGCCAGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCTACGAGGAGGACCAGAGGCAGGGCGCCGAG

（续）

图51A

CCCAGGAAGAACTTCGTGAAGCCCAACGAGACCAAGACCTACTTCTGGAAGGTGCAGCACCACATGGCCCCC
ACCAAGGACGAGTTCGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGACCTGGAGAAGGACGTGCACAGC
GGCCTGATCGGCCCCCTGCTGGTGTGCCACACCAACACCCTGAACCCCGCCCACGGCAGGCAGGTGACCGTG
CAGGAGTTCGCCCTGTTCTTCACCATCTTCGACGAGACCAAGAGCTGGTACTTCACCGAGAACATGGAGAGG
AACTGCAGGGCCCCCTGCAACATCCAGATGGAGGACCCCACCTTCAAGGAGAACTACAGGTTCCACGCCATC
AACGGCTACATCATGGACACCCTGCCCGGCCTGGTGATGGCCCAGGACCAGAGGATCAGGTGGTATCTGCTG
AGCATGGGCAGCAACGAGAACATCCACAGCATCCACTTCAGCGGCCACGTGTTCACCGTGAGGAAGAAGGAG
GAGTACAAGATGGCCCTGTACAACCTGTACCCCGGCGTGTTCGAGACCGTGGAGATGCTGCCCAGCAAGGCC
GGCATCTGGAGGGTGGAGTGCCTGATCGGCGAGCACCTGCACGCCGGCATGAGCACCCTGTTCCTGGTGTAC
AGCAACAAGTGCCAGACCCCCCTGGGCATGGCCAGCGGCCACATCAGGGACTTCCAGATCACCGCCTCTGGC
CAGTACGGCCAGTGGGCCCCCAAGCTGGCCAGGCTGCACTACAGCGGCAGCATCAACGCCTGGAGCACCAAG
GAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCCATGATCATCCACGGCATCAAGACCCAGGGCGCC
AGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTGGACGGCAAGAAGTGGCAG
ACCTACAGGGGCAACAGCACCGGCACCCTGATGGTGTTCTTCGGCAACGTGGACAGCAGCGGCATCAAGCAC
AACATCTTCAACCCCCCCATCATCGCCAGGTACATCAGGCTGCACCCCACCCACTACAGCATCAGGAGCACC
CTGCGGATGGAACTGATGGGCTGCGACCTGAACAGCTGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATC
TCTGACGCCCAGATCACCGCCAGCAGCTACTTCACCAACATGTTCGCCACCTGGAGCCCCAGCAAGGCCAGG
CTGCACCTGCAGGGCAGGAGCAACGCCTGGAGGCCCCAGGTGAACAACCCCAAGGAGTGGCTGCAGGTGGAC
TTCCAGAAGACCATGAAGGTGACCGGCGTGACCACCCAGGGCGTGAAGAGCCTGCTGACCAGCATGTACGTG
AAGGAGTTCCTGATCAGCAGCAGCCAGGACGGCCACCAGTGGACCCTGTTCTTCCAGAACGGCAAAGTGAAG
GTGTTCCAGGGCAACCAGGACAGCTTCACCCCCGTGGTGAACAGCCTGGACCCCCCCCTGCTGACCAGGTAT
CTGAGGATCCACCCCCAGAGCTGGGTGCACCAGATCGCCCTGAGAATGGAAGTGCTGGGATGCGAGGCCCAG
GACCTGTACTGA (SEQ ID NO:28)

图51B

CS23Δ(772-1667) - CS23-SC2-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGCGCCGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGCGAGCTGCCT
GTGGACGCCAGGTTCCCCCCCAGAGTGCCCAAGAGCTTCCCCTTCAACACCTCAGTGGTGTACAAG
AAGACCCTGTTCGTGGAGTTCACCGACCACCTGTTCAACATCGCCAAGCCCAGGCCCCCCTGGATG
GGCCTGCTGGGCCCCACCATCCAGGCCGAGGTGTACGACACCGTGGTGATCACCCTGAAGAACATG
GCCAGCCACCCCGTGAGCCTGCACGCCGTGGGCGTGAGCTACTGGAAGGCCTCTGAGGGCGCCGAG
TATGACGACCAGACCAGCCAGAGGGAGAAGGAGGACGACAAGGTGTTCCCCGGCGGCAGCCACACC
TACGTGTGGCAGGTGCTGAAGGAGAACGGCCCCATGGCCAGCGACCCCCTGTGCCTGACCTACAGC
TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACTCTGGCCTGATCGGCGCCCTGCTGGTGTGC
AGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCCTGCACAAGTTCATCCTGCTGTTCGCCGTG
TTCGATGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAGGACAGGGATGCCGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACCGTGAACGGCTACGTGAACAGGAGCCTGCCCGGCCTG
ATCGGCTGCCACAGGAAGTCTGTGTACTGGCACGTGATCGGCATGGGCACCACCCCCGAGGTGCAC
AGCATCTTCCTGGAGGGCCACACCTTCCTGGTGAGGAACCACAGGCAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCCTGACCGCCCAGACCCTGCTGATGGACCTGGGCCAGTTCCTGCTGTTCTGCCAC
ATCAGCAGCCACCAGCACGACGGCATGGAGGCCTACGTGAAGGTGGACAGCTGCCCCGAGGAGCCC
CAGCTGAGGATGAAGAACAACGAGGAGGCCGAGGACTATGATGATGACCTGACCGACTCTGAGATG
GACGTGGTGAGGTTTGATGATGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCCAAGAAG
CACCCCAAGACCTGGGTGCACTACATCGCCGCCGAGGAGGAGGACTGGGACTACGCCCCCCTGGTG
CTGGCCCCCGACGACAGGAGCTACAAGAGCCAGTACCTGAACAACGGCCCCCAGAGGATCGGCAGG
AAGTACAAGAAGGTCAGATTCATGGCCTACACCGACGAGACCTTCAAGACCAGGGAGGCCATCCAG
CACGAGTCTGGCATCCTGGGCCCCCTGCTGTACGGCGAGGTGGGCGACACCCTGCTGATCATCTTC
AAGAACCAGGCCAGCAGGCCCTACAACATCTACCCCCACGGCATCACCGATGTGAGGCCCCTGTAC
AGCAGGAGGCTGCCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCCTGCCCGGCGAGATCTTC
AAGTACAAGTGGACCGTGACCGTGGAGGATGGCCCCACCAAGTCTGACCCCAGGTGCCTGACCAGG
TACTACAGCAGCTTCGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATCGGCCCCCTGCTGATC
TGCTACAAGGAGAGCGTGGACCAGAGGGGCAACCAGATCATGTCTGACAAGAGGAACGTGATCCTG
TTCTCTGTGTTCGATGAGAACAGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCCTGCCCAAC
CCCGCCGGCGTGCAGCTGGAGGACCCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCAACGGC
TACGTGTTCGACAGCCTGCAGCTGTCTGTGTGCCTGCACGAGGTGGCCTACTGGTACATCCTGAGC
ATCGGCGCCCAGACCGACTTCCTGTCTGTGTTCTTCTCTGGCTACACCTTCAAGCACAAGATGGTG
TACGAGGACACCCTGACCCTGTTCCCCTTCAGCGGCGAGACCGTGTTCATGAGCATGGAGAACCCC
GGCCTGTGGATCCTGGGCTGCCACAACAGCGACTTCAGGAACAGGGGCATGACCGCCCTGCTGAAA
GTCAGCAGCTGCGACAAGAACACCGGCGACTACTACGAGGACAGCTACGAGGACATCAGCGCCTAC
CTGCTGAGCAAGAACAACGCCATCGAGCCCAGGAGCTTCAGCCAGAACTCCAGACACCCCAGCACC

（续）

图52A

AGGGAGATCACCAGGACCACCCTGCAGAGCGACCAGGAGGAGATCGACTATGATGACACCATCAGC
GTGGAGATGAAGAAGGAGGACTTCGACATCTACGACGAGGACGAGAACCAGAGCCCCAGGAGCTTC
CAGAAGAAGACCAGGCACTACTTCATCGCCGCCGTGGAGAGGCTGTGGGACTATGGCATGAGCAGC
AGCCCCCACGTGCTGAGGAACAGGGCCCAGAGCGGCAGCGTGCCCCAGTTCAAGAAGGTGGTGTTC
CAGGAGTTCACCGACGGCAGCTTCACCCAGCCCCTGTACAGAGGCGAGCTGAACGAGCACCTGGGC
CTGCTGGGCCCCTACATCAGGGCCGAGGTGGAGGACAACATCATGGTGACCTTCAGGAACCAGGCC
AGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCTACGAGGAGGACCAGAGGCAGGGCGCCGAG
CCCAGGAAGAACTTCGTGAAGCCCAACGAGACCAAGACCTACTTCTGGAAGGTGCAGCACCACATG
GCCCCCACCAAGGACGAGTTCGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGACCTGGAGAAG
GACGTGCACAGCGGCCTGATCGGCCCCCTGCTGGTGTGCCACACCAACACCCTGAACCCCGCCCAC
GGCAGGCAGGTGACCGTGCAGGAGTTCGCCCTGTTCTTCACCATCTTCGACGAGACCAAGAGCTGG
TACTTCACCGAGAACATGGAGAGGAACTGCAGGGCCCCCTGCAACATCCAGATGGAGGACCCCACC
TTCAAGGAGAACTACAGGTTCCACGCCATCAACGGCTACATCATGGACACCCTGCCCGGCCTGGTG
ATGGCCCAGGACCAGAGGATCAGGTGGTATCTGCTGAGCATGGGCAGCAACGAGAACATCCACAGC
ATCCACTTCAGCGGCCACGTGTTCACCGTGAGGAAGAAGGAGGAGTACAAGATGGCCCTGTACAAC
CTGTACCCCGGCGTGTTCGAGACCGTGGAGATGCTGCCCAGCAAGGCCGGCATCTGGAGGGTGGAG
TGCCTGATCGGCGAGCACCTGCACGCCGGCATGAGCACCCTGTTCCTGGTGTACAGCAACAAGTGC
CAGACCCCCCTGGGCATGGCCAGCGGCCACATCAGGGACTTCCAGATCACCGCCTCTGGCCAGTAC
GGCCAGTGGGCCCCCAAGCTGGCCAGGCTGCACTACAGCGGCAGCATCAACGCCTGGAGCACCAAG
GAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCCATGATCATCCACGGCATCAAGACCCAG
GGCGCCAGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTGGACGGC
AAGAAGTGGCAGACCTACAGGGGCAACAGCACCGGCACCCTGATGGTGTTCTTCGGCAACGTGGAC
AGCAGCGGCATCAAGCACAACATCTTCAACCCCCCCATCATCGCCAGGTACATCAGGCTGCACCCC
ACCCACTACAGCATCAGGAGCACCCTGCGGATGGAACTGATGGGCTGCGACCTGAACAGCTGCAGC
ATGCCCCTGGGCATGGAGAGCAAGGCCATCTCTGACGCCCAGATCACCGCCAGCAGCTACTTCACC
AACATGTTCGCCACCTGGAGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGCAACGCCTGG
AGGCCCCAGGTGAACAACCCCAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATGAAGGTGACC
GGCGTGACCACCCAGGGCGTGAAGAGCCTGCTGACCAGCATGTACGTGAAGGAGTTCCTGATCAGC
AGCAGCCAGGACGGCCACCAGTGGACCCTGTTCTTCCAGAACGGCAAAGTGAAGGTGTTCCAGGGC
AACCAGGACAGCTTCACCCCCGTGGTGAACAGCCTGGACCCCCCCCCTGCTGACCAGGTATCTGAGG
ATCCACCCCCAGAGCTGGGTGCACCAGATCGCCCTGAGAATGGAAGTGCTGGGATGCGAGGCCCAG
GACCTGTACTGA (SEQ ID NO:29)

图52B

CS01m23-FL-AA (SEQ ID NO: 104)

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFN
TSVVYKKTLFVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVVTLKNMASHPVSLHAV
GVSYWKSSEGAEYDDQTSQREKEDDKVFPGKSHTYVWQVLKENGPTASDPPCLTYSYLSH
VDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRD
AASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNH
RQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNE
EAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLA
PDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGDTL
LIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGP
TKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDE
NRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILS
IGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRG
MTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNTTYVNRSLSQNPPVLKRHQREITRTT
LQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSS
PHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVT
FRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKA
WAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTEN
MERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIH
SIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLF
LVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVD
LLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSG
IKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASS
YFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTS
MYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWV
HQIALRMEVLGCEAQDLY

图53

CS04m3-FL-AA (SEQ ID NO: 105)

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFN
TSVVYKKTLFVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAV
GVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSH
VDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRD
AASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNH
RQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNE
EAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLA
PDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGDTL
LIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGP
TKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDE
NRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILS
IGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRG
MTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNTTYVNRSLSQNPPVLKRHQREITRTT
LQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSS
PHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVT
FRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKA
WAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTEN
MERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIH
SIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLF
LVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVD
LLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSG
IKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASS
YFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTS
MYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWV
HQIALRMEVLGCEAQDLY

图54

CS01-FL-AAm12 (SEQ ID NO: 106)

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYK
KTLFVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVVTLKNMASHPVSLHAVGVSYWKSSEGAE
YDDQTSQREKEDDKVFPGKSHTYVWQVLKENGPTASDPPCLTYSYLSHVDLVKDLNSGLIGALLVC
REGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGL
IGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLSCH
ISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKK
HPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQ
HESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIF
KYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVIL
FSVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILS
IGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLK
VSSCDKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQREITRTTLQSDQEEIDYDDT
ISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKV
VFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQG
AEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNP
AHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPG
LVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWR
VECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWS
TKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGN
VDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSY
FTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFL
ISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCE
AQDLY

图55

CS04-FL-AAm12 (SEQ ID NO: 107)

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYK
KTLFVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVVTLKNMASHPVSLHAVGVSYWKSSEGAE
YDDQTSQREKEDDKVFPGKSHTYVWQVLKENGPTASDPPCLTYSYLSHVDLVKDLNSGLIGALLVC
REGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGL
IGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLSCH
ISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKK
HPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQ
HESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIF
KYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVIL
FSVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILS
IGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLK
VSSCDKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQREITRTTLQSDQEEIDYDDT
ISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKV
VFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQG
AEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNP
AHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPG
LVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWR
VECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWS
TKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGN
VDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSY
FTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFL
ISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCE
AQDLY

图56

CS01-FL-NAm12 (SEQ ID NO: 108)

ATGCAGATTGAGCTGTCCACCTGCTTCTTTCTGTGCCTGCTGAGATTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGCTGTGGAACTTTCTTGGGACTACATGCAGTCTGACCTGGGAGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAGTCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACACTCTTTGTGGAATTCACTGACCACCTGTTCAACATTGCAAAACCCAGACCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCGTCACCCTCAAGAACATG
GCATCCCACCCTGTGTCTCTGCATGCTGTGGGAGTCTCATACTGGAAATCCTCTGAAGGGGCTGAG
TATGATGACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGAAGTCTCACACC
TATGTGTGGCAAGTCCTCAAGGAGAATGGACCCACTGCATCTGACCCACCCTGCCTGACATACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCACTGCTGGTGTGC
AGGGAAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGTCTTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGATGCTGCC
TCTGCCAGGGCATGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAAGTGCAC
TCCATTTTCCTGGAGGGACACACCTTCCTGGTCAGGAACCACAGACAAGCCTCTCTGGAGATCTCT
CCCATCACCTTCCTCACTGCACAGACACTGCTGATGGACCTTGGACAGTTCCTGCTGTCCTGCCAC
ATCTCTTCCCACCAGCATGATGGCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCAGATTTGATGATGACAACTCTCCATCCTTCATTCAGATCAGGTCTGTGGCAAAGAAA
CACCCCAAGACATGGGTGCACTACATTGCTGCTGAGGAAGAGGACTGGGACTATGCACCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAACAATGGCCCACAAAGAATTGGAAGA
AAGTACAAGAAAGTCAGATTCATGGCCTACACTGATGAAACCTTCAAGACAAGAGAAGCCATTCAG
CATGAGTCTGGCATTCTGGGACCACTCCTGTATGGGGAAGTGGGAGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCAGGAGACTGCCAAAAGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGAGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCCAGGTGCCTCACCAGA
TACTACTCCTCTTTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGAGGCAACCAGATCATGTCTGACAAGAGAAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTCAGAGATTCCTGCCCAAC
CCTGCTGGGGTGCAACTGGAAGACCCTGAGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACTCTCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCACAAACTGACTTCCTTTCTGTCTTCTTCTCTGGATACACCTTCAAGCACAAGATGGTG
TATGAGGACACCCTGACACTCTTCCCATTCTCTGGGGAAACTGTGTTCATGAGCATGGAGAACCCT
GGACTGTGGATTCTGGGATGCCACAACTCTGACTTCAGAAACAGGGGAATGACTGCACTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATGCCATTGAGCCCAGAAGCTTCTCTCAGAATCCACCTGTCCTGAAGAGA
CACCAGAGAGAGATCACCAGGACAACCCTCCAGTCTGACCAGGAAGAGATTGACTATGATGACACC
ATTTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTATGATGAGGACGAGAACCAGTCTCCAAGA
TCATTCCAGAAGAAGACAAGACACTACTTCATTGCTGCTGTGGAAAGACTGTGGGACTATGGCATG
TCTTCCTCTCCCCATGTCCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAAGAAAGTG

（续）

图57A

GTCTTCCAGGAGTTCACTGATGGCTCATTCACCCAGCCCCTGTACAGAGGGGAACTGAATGAGCAC
CTGGGACTCCTGGGACCATACATCAGGGCTGAGGTGGAAGACAACATCATGGTGACATTCAGAAAC
CAGGCCTCCAGGCCCTACAGCTTCTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAAGGG
GCTGAGCCAAGAAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCAC
CACATGGCACCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCATACTTCTCTGATGTGGACCTG
GAGAAAGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAACCCT
GCACATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACCAAG
TCATGGTACTTCACTGAGAACATGGAGAGAAACTGCAGAGCACCATGCAACATTCAGATGGAAGAC
CCCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCTGGG
CTTGTCATGGCACAGGACCAGAGAATCAGATGGTACCTGCTTTCTATGGGATCCAATGAGAACATT
CACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGAGAAAGAAGGAGGAATACAAGATGGCCCTG
TACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGG
GTGGAATGCCTCATTGGGGAGCACCTGCATGCTGGCATGTCAACCCTGTTCCTGGTCTACAGCAAC
AAGTGCCAGACACCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGC
CAGTATGGCCAGTGGGCACCCAAACTGGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGGTCA
ACCAAGGAGCCATTCTCTTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTCATGGCATCAAG
ACACAGGGGGCAAGACAGAAATTCTCCTCTCTGTACATCTCACAGTTCATCATCATGTACTCTCTG
GATGGCAAGAAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCTCATGGTCTTCTTTGGCAAT
GTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCTCCCATCATTGCCAGATACATCAGGCTG
CACCCCACCCACTACTCAATCAGATCAACCCTCAGGATGGAACTGATGGGATGTGACCTGAACTCC
TGCTCAATGCCCCTGGGAATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCATCCTCTTAC
TTCACCAACATGTTTGCCACCTGGTCACCATCAAAAGCCAGGCTGCACCTCCAGGGAAGAAGCAAT
GCCTGGAGACCCCAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAATGAAA
GTCACTGGGGTGACAACCCAGGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGAGTTCCTG
ATCTCTTCCTCACAGGATGGCCACCAGTGGACACTCTTCTTCCAGAATGGCAAAGTCAAGGTGTTC
CAGGGCAACCAGGACTCTTTCACACCTGTGGTGAACTCACTGGACCCCCCCCTCCTGACAAGATAC
CTGAGAATTCACCCCCAGTCTTGGGTCCACCAGATTGCCCTGAGAATGGAAGTCCTGGGATGTGAG
GCACAAGACCTGTACTGA

图57B

CS04-FL-NAm12 (SEQ ID NO: 109)

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGGGCTGTGGAGCTTTCTTGGGACTACATGCAGTCTGACCTGGGGGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAATCCTTCCCATTCAACACCTCTGTGGTCTACAAG
AAGACCCTCTTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAAACCCAGGCCACCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCGTCACCCTCAAGAACATG
GCCTCCCACCCTGTGAGCCTGCATGCTGTGGGGGTCAGCTACTGGAAGTCCTCTGAGGGGGCTGAG
TATGATGACCAGACCTCCCAGAGGGAGAAGGAGGATGACAAAGTGTTCCCTGGGAAGAGCCACACC
TATGTGTGGCAGGTCCTCAAGGAGAATGGCCCCACTGCCTCTGACCCACCCTGCCTGACCTACTCC
TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCCCTGCTGGTGTGC
AGGGAGGGCTCCCTGGCCAAAGAGAAGACCCAGACCCTGCACAAGTTCATTCTCCTGTTTGCTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACTCCCTGATGCAGGACAGGGATGCTGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAGGTGCAC
TCCATTTTCCTGGAGGGCCACACCTTCCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCCTCACTGCCCAGACCCTGCTGATGGACCTCGGACAGTTCCTGCTGTCCTGCCAC
ATCAGCTCCCACCAGCATGATGGCATGGAGGCCTATGTCAAGGTGGACAGCTGCCCTGAGGAGCCA
CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCCGCTTTGATGATGACAACAGCCCATCCTTCATTCAGATCAGGTCTGTGGCCAAGAAA
CACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGGACTGGGACTATGCCCCACTGGTC
CTGGCCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGGCCCACAGAGGATTGGACGC
AAGTACAAGAAAGTCAGGTTCATGGCCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATTCAG
CATGAGTCTGGCATCCTGGGCCCACTCCTGTATGGGGAGGTGGGGGACACCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCCTGTAC
AGCCGCAGGCTGCCAAAGGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGGGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACCAAATCTGACCCCAGGTGCCTCACCAGA
TACTACTCCAGCTTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGGGGAAACCAGATCATGTCTGACAAGAGGAATGTGATTCTG
TTCTCTGTCTTTGATGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAGCGCTTCCTGCCCAAC
CCTGCTGGGGTGCAGCTGGAGGACCCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACAGCCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
ATTGGGGCCCAGACTGACTTCCTTTCTGTCTTCTTCTCTGGCTACACCTTCAAACACAAGATGGTG
TATGAGGACACCCTGACCCTCTTCCCATTCTCTGGGGAGACTGTGTTCATGAGCATGGAGAACCCT
GGCCTGTGGATTCTGGGATGCCACAACTCTGACTTCCGCAACAGGGGCATGACTGCCCTGCTCAAA
GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACAGCTATGAGGACATCTCTGCCTAC
CTGCTCAGCAAGAACAATGCCATTGAGCCCAGGAGCTTCAGCCAGAATCCACCTGTCCTGAAACGC
CACCAGAGGGAGATCACCAGGACCACCCTCCAGTCTGACCAGGAGGAGATTGACTATGATGACACC
ATTTCTGTGGAGATGAAGAAAGAGGACTTTGACATCTATGACGAGGACGAGAACCAGAGCCCAAGG
AGCTTCCAGAAGAAGACCAGGCACTACTTCATTGCTGCTGTGGAGCGCCTGTGGGACTATGGCATG
AGCTCCAGCCCCCATGTCCTCAGGAACAGGGCCCAGTCTGGCTCTGTGCCACAGTTCAAGAAAGTG
GTCTTCCAAGAGTTCACTGATGGCAGCTTCACCCAGCCCCTGTACAGAGGGGAGCTGAATGAGCAC
CTGGGACTCCTGGGCCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACCTTCCGCAAC

（续）

图58A

CAGGCCTCCAGGCCCTACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAGGGG
GCTGAGCCACGCAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCAC
CACATGGCCCCCACCAAGGATGAGTTTGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGACCTG
GAGAAGGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAACCCT
GCCCATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACCAAG
AGCTGGTACTTCACTGAGAACATGGAGCGCAACTGCAGGGCCCCATGCAACATTCAGATGGAGGAC
CCCACCTTCAAAGAGAACTACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCTGGG
CTTGTCATGGCCCAGGACCAGAGGATCAGGTGGTACCTGCTTTCTATGGGCTCCAATGAGAACATT
CACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGCGCAAGAAGGAGGAGTACAAGATGGCCCTG
TACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGG
GTGGAGTGCCTCATTGGGGAGCACCTGCATGCTGGCATGAGCACCCTGTTCCTGGTCTACAGCAAC
AAGTGCCAGACCCCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGC
CAGTATGGCCAGTGGGCCCCCAAGCTGGCCAGGCTCCACTACTCTGGATCCATCAATGCCTGGAGC
ACCAAGGAGCCATTCAGCTGGATCAAAGTGGACCTGCTGGCCCCCATGATCATCCATGGCATCAAG
ACCCAGGGGGCCAGGCAGAAGTTCTCCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTG
GATGGCAAGAAATGGCAGACCTACAGAGGCAACTCCACTGGAACACTCATGGTCTTCTTTGGCAAT
GTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCCCCAATCATCGCCAGATACATCAGGCTG
CACCCCACCCACTACAGCATCCGCAGCACCCTCAGGATGGAGCTGATGGGCTGTGACCTGAACTCC
TGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATTTCTGATGCCCAGATCACTGCCTCCAGCTAC
TTCACCAACATGTTTGCCACCTGGAGCCCAAGCAAGGCCAGGCTGCACCTCCAGGGAAGGAGCAAT
GCCTGGAGGCCCCAGGTCAACAACCCAAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATGAAG
GTCACTGGGGTGACCACCCAGGGGGTCAAGAGCCTGCTCACCAGCATGTATGTGAAGGAGTTCCTG
ATCAGCTCCAGCCAGGATGGCCACCAGTGGACCCTCTTCTTCCAGAATGGCAAGGTCAAGGTGTTC
CAGGGCAACCAGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCCCTCCTGACCAGATAC
CTGAGGATTCACCCCCAGAGCTGGGTCCACCAGATTGCCCTGAGGATGGAGGTCCTGGGATGTGAG
GCCCAGGACCTGTACTGA

图58B