



(12) 发明专利

(10) 授权公告号 CN 108285491 B

(45) 授权公告日 2021.08.10

(21) 申请号 201810089961.2

(22) 申请日 2013.02.28

(65) 同一申请的已公布的文献号  
申请公布号 CN 108285491 A

(43) 申请公布日 2018.07.17

(30) 优先权数据  
61/605,028 2012.02.29 US

(62) 分案原申请数据  
201380016072.9 2013.02.28

(73) 专利权人 桑格摩生物科学股份有限公司  
地址 美国加利福尼亚州

(72) 发明人 J·C·米勒 E·J·瑞巴  
H·S·张

(74) 专利代理机构 上海专利商标事务所有限公司 31100  
代理人 余颖

(51) Int.Cl.

C07K 19/00 (2006.01)

C12N 15/90 (2006.01)

A61K 38/17 (2006.01)

A61P 25/14 (2006.01)

(56) 对比文件

CN 101365801 A, 2009.02.11

CN 101668856 A, 2010.03.10

US 20110082093 A1, 2011.04.07

Mittelman D等.Zinc-finger directed double-strand breaks within CAG repeat tracts promote repeat instability in human cells.《Proc.natl.acad.sci.usa》.2009,第106卷(第24期),第9607-9612页.

审查员 夏颖

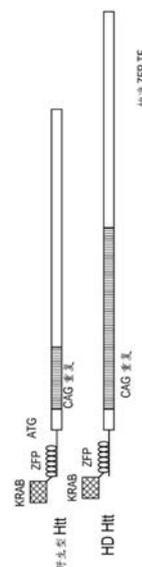
权利要求书3页 说明书46页  
序列表74页 附图41页

(54) 发明名称

治疗亨廷顿氏病的方法和组合物

(57) 摘要

本文公开了治疗或预防亨廷顿氏病的方法和组合物。



1. 优先结合含40个或更多个CAG重复的突变Htt基因的非天然存在的锌指蛋白,所述锌指蛋白包含以下编号为F1至F3、F1至F4、F1至F5和F1至F6的3、4、5或6个锌指结构域:

- (i) F1:RSDNLSE (SEQ ID NO:60) ,  
F2:KRCNLRC (SEQ ID NO:61) ,  
F3:QSGDLTR (SEQ ID NO:18) ,  
F4:QSGDLTR (SEQ ID NO:18) ,  
F5:RSDNLSE (SEQ ID NO:60) ,和  
F6:KRCNLRC (SEQ ID NO:61) ;或
- (ii) F1:QSGDLTR (SEQ ID NO:18) ,  
F2:QSGDLTR (SEQ ID NO:18) ,  
F3:RSDNLSE (SEQ ID NO:60) ,  
F4:KRCNLRC (SEQ ID NO:61) ,  
F5:QSGDLTR (SEQ ID NO:18) ,和  
F6:QSGDLTR (SEQ ID NO:18) ;或
- iii) F1:RSDNLSE (SEQ ID NO:60) ,  
F2:KRCNLRC (SEQ ID NO:61) ,  
F3:QSSDLR (SEQ ID NO:31) ,  
F4:QWSTRKR (SEQ ID NO:63) ,  
F5:QSSDLR (SEQ ID NO:31) ,和  
F6:QWSTRKR (SEQ ID NO:63) ;或
- (iv) F1:RSDNLSE (SEQ ID NO:60) ,  
F2:KRCNLRC (SEQ ID NO:61) ,  
F3:RSDNLSE (SEQ ID NO:60) ,  
F4:KRCNLRC (SEQ ID NO:61) ,  
F5:RSDNLSE (SEQ ID NO:60) ,和  
F6:KRCNLRC (SEQ ID NO:61) ;或
- (v) F1:RSDNLSE (SEQ ID NO:60) ,  
F2:KRCNLRC (SEQ ID NO:61) ,  
F3:QSSDLR (SEQ ID NO:31) ,  
F4:QWSTRKR (SEQ ID NO:63) ,和  
F5:QSGDLTR (SEQ ID NO:18) ;或
- (vi) F1:QSSDLR (SEQ ID NO:31) ,  
F2:QWSTRKR (SEQ ID NO:63) ,  
F3:QSSDLR (SEQ ID NO:31) ,  
F4:QWSTRKR (SEQ ID NO:63) ,和  
F5:QSGDLTR (SEQ ID NO:18) ;或
- (vii) F1:QSGDLTR (SEQ ID NO:18) ,  
F2:QSSDLR (SEQ ID NO:31) ,  
F3:QWSTRKR (SEQ ID NO:63) ,

- F4:QSSDLSR (SEQ ID NO:31), 和  
F5:QWSTRKR (SEQ ID NO:63); 或  
(viii) F1:RSDNLSE (SEQ ID NO:60),  
F2:KRCNLRC (SEQ ID NO:61),  
F3:QSGDLTR (SEQ ID NO:18),  
F4:QSSDLSR (SEQ ID NO:31), 和  
F5:QWSTRKR (SEQ ID NO:63); 或  
(ix) F1:QSGDLTR (SEQ ID NO:18),  
F2:QSSDLSR (SEQ ID NO:31), 和  
F3:QWSTRKR (SEQ ID NO:63); 或  
(x) F1:RSDNLRE (SEQ ID NO:65),  
F2:RSDNLSE (SEQ ID NO:60), 和  
F3:KRCNLRC (SEQ ID NO:61);  
所述锌指蛋白结合包含序列GCAGCAGCAGCAGCA的靶位点。
2. 如权利要求1所述的锌指蛋白, 所述锌指蛋白包含 (i)、(v) 或 (vi) 项所述的锌指结构域。
  3. 如权利要求1或2所述的锌指蛋白, 所述锌指蛋白结合Htt基因的CAG重复区内的序列。
  4. 如权利要求1所述的锌指蛋白, 还包含当与DNA结合时允许锌指蛋白多聚化的二聚化结构域。
  5. 一种融合蛋白, 包含权利要求1-4中任一项所述的锌指蛋白和功能结构域, 所述功能结构域选自: 转录激活结构域、转录阻遏结构域和核酸酶结构域。
  6. 一种多核苷酸, 编码一种或多种权利要求1-4所述的锌指蛋白或权利要求5所述的融合蛋白。
  7. 一种宿主细胞, 包含一种或多种权利要求1-4所述的锌指蛋白、权利要求5所述的融合蛋白或权利要求6所述的多核苷酸。
  8. 一种药物组合物, 包含一种或多种权利要求1-4所述的锌指蛋白、权利要求5所述的融合蛋白或权利要求6所述的多核苷酸。
  9. 一种修饰细胞内Htt基因表达的药物组合物, 所述药物组合物包含编码一种或多种权利要求5所述融合蛋白的一种或多种多核苷酸。
  10. 如权利要求9所述的药物组合物, 所述Htt基因包含至少一个突变等位基因。
  11. 如权利要求9所述的药物组合物, 所述Htt基因是野生型的。
  12. 如权利要求9-11中任一项所述的药物组合物, 所述融合蛋白包含核酸酶结构域或阻遏结构域, 且Htt基因表达被失活。
  13. 一种修饰细胞内Htt基因的药物组合物, 所述药物组合物包含编码一种或多种权利要求5所述融合蛋白的一种或多种多核苷酸, 所述一种或多种融合蛋白包含核酸酶结构域, 所述Htt基因的序列经修饰。
  14. 一种修饰细胞内Htt基因的方法, 所述方法包括向所述细胞施用编码一种或多种权利要求5所述融合蛋白的一种或多种多核苷酸, 所述融合蛋白包含核酸酶结构域, 所述Htt

基因的序列经修饰,所述方法不用于诊断和治疗目的。

15. 一种生成用于研究亨廷顿氏病 (HD) 的模型体系的方法,所述方法包括据权利要求14所述方法修饰Htt基因。

16. 如权利要求15所述方法,所述Htt基因经修饰为包含一个或多个突变等位基因。

17. 如权利要求16所述的方法,所述突变等位基因包含扩增的三核苷酸重复。

18. 如权利要求15所述的方法,所述细胞包括胚胎干细胞。

19. 一种细胞,由权利要求14所述方法制得。

20. 一种分离的细胞,包含权利要求9所述的药物组合物。

21. 一种分离的细胞,包含权利要求13所述的药物组合物。

22. 一种用于治疗亨廷顿氏病的药物组合物,包含一种或多种权利要求20或权利要求21所述的细胞。

23. 编码一种或多种权利要求5所述融合蛋白的一种或多种多核苷酸用于制造治疗HD的药物的用途。

## 治疗亨廷顿氏病的方法和组合物

[0001] 相关申请的交叉引用

[0002] 本申请要求2012年2月29日提交的美国临时申请第61/605,028号的权益,其内容据此通过引用整体并入。

### 技术领域

[0003] 本公开属于基因表达和基因组编辑的领域。

[0004] 发明背景

[0005] 亨廷顿氏病(Huntington's Disease, HD),也称为亨廷顿氏舞蹈症(Huntington's Chorea),是一种运动、认知和精神障碍的进行性病症。虽然约10%的病例中发作发生在21岁之前,但是这种疾病发作的平均年龄为35-44岁,并且在疾病诊断后的平均寿命为15-18年。在100,000个西方欧洲血统的人中患病率为3至7人。

[0006] 亨廷顿氏病是在上世纪90年代早期首次表征的三核苷酸重复扩增病症的实例(见Di Prospero和Fischbeck(2005) Nature Reviews Genetics 6:756-765)。这些病症牵涉三个核苷酸一组的不稳定重复的局部扩增并且可导致扩增重复所在基因的功能丧失、获得毒性功能或二者。三核苷酸重复可位于所述基因的任何部分,包括非编码和编码基因区。位于编码区内的重复通常包括重复谷氨酰胺编码三联体(CAG)或丙氨酸编码三联体(CGA)。非编码序列内的扩增重复区可导致基因表达异常,而编码区内的扩增重复(也称为密码子重复症)可引起错误折叠和蛋白质聚集。与异常蛋白质相关的确切病理生理学原因往往未知。通常,在经受三核苷酸扩增的野生型基因中,在正常人群中这些区域含有可变数量的重复序列,但是在患病人群中,重复的数量可翻升一倍至重复的数量增加一个对数级。在HD中,在大胞质蛋白亨廷顿蛋白(Htt)的N端编码区中插入重复。正常Htt等位基因含有15-20个CAG重复,而含35个或更多个重复的等位基因可视为可能引起HD的等位基因并且带来了发展所述疾病的风险。认为含36-39个重复的等位基因不完全渗透,并且携带有那些等位基因的那些个体可能或可能不发展所述疾病(或可能在以后的生活中发展症状),而认为含有40个重复或更多个重复的等位基因完全渗透。实际上,尚未报道含有具有许多这种重复的HD等位基因的无症状患者。常常发现有青少年型发作HD(<21岁)的那些个体具有60个或更多个CAG重复。除CAG重复增多外,还证实HD可涉及重复序列中+1和+2移码,以致所述区域将编码聚丝氨酸多肽(在为+1移码的情况下由AGC重复编码)簇而非聚谷氨酰胺(Davies和Rubinsztein(2006) Journal of Medical Genetics 43:893-896)。

[0007] 在HD中,突变Htt等位基因通常作为显性性状从父母一方遗传。如果父母另一方未患所述病症,则HD患者所生的任何孩子均有50%发展所述疾病的机会。在一些情况下,父母可能具有中间HD等位基因并且没有症状,而由于重复扩增,孩子表现出所述疾病。另外,HD等位基因也可表现出预期所知的现象,其中由于精子发生期间重复区的不稳定性质,经过几代观察到严重程度增加或发作的年龄减小。

[0008] 此外,Htt中的三核苷酸扩增导致纹状体中中型多棘 $\gamma$ -氨基丁酸(GABA)投射神经元的神经元丢失,其中在新皮质中也发生神经元丢失。含有脑啡肽并且投射到外苍白球的

中型多棘神经元比含有物质P并且投射到内苍白球的神经元更复杂。在患有亨廷顿氏病的人中受严重影响的其它脑部区域包括黑质、皮层3、5和6、海马CA1区、顶叶角回、小脑浦肯野细胞(Purkinje cell)、下丘脑外侧结节核和丘脑中央中核束旁核复合物(Walker (2007) Lancet 369:218-228)。

[0009] 正常Htt蛋白质的作用知之甚少,但是可能牵涉于神经发生、凋亡性细胞死亡和囊泡运输。另外,有证据表明野生型Htt刺激脑源性神经营养因子(BDNF)(纹状体神经元的促存活因子)的生成。已经证实在小鼠HD模型中HD的进展与BDNF表达减少相关联(Zuccato等(2005) Pharmacological Research 52(2):133-139),并且经由腺相关病毒(AAV)载体介导的基因递送的BDNF或胶质细胞系源性神经营养因子(GDNF)的递送可保护HD小鼠模型中的纹状体神经元(Kells等,(2004) Molecular Therapy 9(5):682-688)。

[0010] 对HD的治疗选择目前非常有限。设计用于预防与通过延伸的聚谷氨酰胺簇发生的蛋白质聚集相关的毒性的一些潜在方法例如过表达伴侣蛋白或用化合物格尔德霉素(geldanamycin)诱导热休克反应已经显示在体外模型中这些毒性降低。其它治疗靶向所述疾病的临床表现中细胞凋亡的作用。例如,在动物模型中,在双亲一方含有HD等位基因而双亲另一方具有半胱天冬酶1(caspase 1)的显性负等位基因的配对小鼠的后代中,已经通过阻断半胱天冬酶活性证实疾病症状减缓。另外,用半胱天冬酶裂解突变HD Htt可能对疾病的致病性起作用。与携带非半胱天冬酶抗性突变Htt等位基因的小鼠相比,发现携带半胱天冬酶-6抗性突变Htt的转基因小鼠保持了正常神经元功能并且不会发展纹状体神经退化(见Graham等(2006) Cell 125:1179-1191)。还已证实靶向凋亡途径成员的分子对症状具有减缓作用。例如,已经证实二者均抑制半胱天冬酶活性的化合物zVAD-fmk和二甲胺四环素(minocycline)减缓了小鼠中的疾病表现。药物瑞马酰胺(remacemide)也已经用于小HD人体试验,因为认为所述化合物防止突变Htt与NDMA受体结合以防止对神经细胞施加毒性作用。然而,在这些试验中未观察到神经元功能的统计上显著的改善。另外,亨廷顿研究小组(Huntington Study Group)使用辅酶Q进行了随机双盲研究。虽然在用辅酶Q10治疗的患者中观察到朝更缓慢疾病进展的趋势,但是在功能总量的下降速率上无显著变化。(Di Prospero和Fischbeck,如上)。美国专利公布2011/0082093公开了靶向Htt的特异性锌指蛋白。

[0011] 因此,仍需要用于治疗 and 预防亨廷顿氏病的组合物和方法。

[0012] 发明概述

[0013] 本文公开了治疗亨廷顿氏病的方法和组合物。具体而言,本文提供了修饰HD Htt等位基因(例如,调节其表达)以便治疗亨廷顿氏病的方法和组合物。还提供了产生亨廷顿氏病的动物模型的方法和组合物。

[0014] 因此,一方面,提供了调节HD等位基因(例如,Htt)的表达的工程化DNA结合结构域(例如,锌指蛋白或TAL效应(TALE)蛋白)。工程化锌指蛋白或TALE是其DNA结合结构域(例如,识别螺旋或RVD)已经改变(例如,通过选择和/或合理设计)为与预选靶位点结合的非天然存在的锌指或TALE蛋白。本文描述的任何锌指蛋白可能包括1、2、3、4、5、6个或更多个锌指,每个锌指具有与所选序列(例如,基因)中的靶亚位点结合的识别螺旋。类似地,本文描述的任何TALE蛋白可能包括任何数量的TALE RVD。在一些实施方案中,至少一个RVD具有非特异性DNA结合。在一些实施方案中,至少一个识别螺旋(或RVD)是非天然存在的。在某些实

实施方案中,锌指蛋白具有表1A和1B中示出的识别螺旋。在其它实施方案中,锌指蛋白与表2A和2B中示出的靶序列结合。在一些实施方案中,锌指蛋白包含表2C中的识别螺旋。在某些实施方案中,将锌指蛋白配制成例如用于向受试者施用的药物组合物。

[0015] 一方面,提供了与完全或部分在Htt的CAG重复区外的序列结合的阻遏因子(ZFP-TF或TALE-TF)。另一方面,提供了与Htt的CAG重复区内的序列结合的ZFP或TALE阻遏因子(ZFP-TF或TALE-TF)。在一些实施方案中,相对于野生型长度的重复簇,这些ZFP-TF或TALE-TF优先与扩增三核苷酸簇结合,从而实现了对扩增等位基因的优先阻遏。在一些实施方案中,这些ZFP-TF或TALE-TF包括当与DNA结合时允许多聚化的蛋白质相互作用结构域(或“二聚化结构域”)。在一些实施方案中,这些ZFP-TF或TALE-TF实现与重复序列的协同DNA结合,以致比野生型等位基因数量更大的ZFP或TALE蛋白更有效地结合扩增等位基因,从而允许对突变等位基因的优先阻遏。这些协同结合ZFP-TF或TALE-TF可能或可能不进一步含有当与DNA结合时允许多聚化的蛋白质相互作用结构域。在一些实施方案中,ZFP-TF或TALE-TF形成指定尺寸的多聚体的稳定复合物,并且因此能够优先与超出某一最小尺寸的CAG簇相互作用,其中该最小尺寸大于野生型CAG簇的长度。

[0016] 在某些实施方案中,如本文所述的ZFP或TALE蛋白(例如,双手、多聚化等)优先修饰突变Htt等位基因的表达。在一些实施方案中,ZFP或TALE与其中扩增簇编码多聚谷氨酰胺的突变Htt等位基因特异性结合,而在其它实施方案中,ZFP或TALE与其中扩增簇编码多聚丝氨酸的突变Htt等位基因特异性结合。因此,在一些实施方案中,ZFP-TF或TALE-TF调节野生型和突变形式的Htt等位基因。在某些实施方案中,ZFP或TALE仅调节野生型Htt等位基因。在其它实施方案中,ZFP或TALE仅调节突变形式的Htt。

[0017] 在其它实施方案中,提供了优先和与扩增HD Htt等位基因相关的已知SNP结合的阻遏ZFP-TF或TALE-TF。用这种方式,ZFP-TF或TALE-TF对含有SNP的突变Htt等位基因有特异性,从而允许特异性阻遏突变Htt等位基因。另一方面,提供了通过和与野生型等位基因相关的SNP相互作用,特异性激活野生型Htt等位基因的ZFP-TF或TALE-TF。用这种方式,仅激活了野生型Htt等位基因。

[0018] 在某些实施方案中,如本文所述的锌指蛋白(ZFP)或TALE蛋白可置于和作为融合蛋白的一部分的调控结构域(或功能结构域)的操作性连接中。例如,功能结构域可为转录激活结构域、转录阻遏结构域和/或核酸酶(裂解)结构域。通过选择与ZFP或TALE融合的激活结构域或阻遏结构域,此类融合蛋白可用于激活或阻遏基因表达。在一些实施方案中,提供了包含靶向如本文所述的突变Htt,与可用于下调突变Htt表达的转录阻遏结构域融合的ZFP或TALE的融合蛋白。在一些实施方案中,提供了包含靶向野生型Htt等位基因,与可上调野生型Htt等位基因的转录激活结构域融合的ZFP或TALE的融合蛋白。在某些实施方案中,调控结构域的活性由外源性小分子或配体调节,以致在外源性配体缺乏时不会发生与细胞转录机构的相互作用。此类外部配体控制ZFP-TF或TALE-TF与转录机构的相互作用程度。调控结构域可与一个或多个ZFP或TALE的任何部分可操作地连接,包括介于一个或多个ZFP或TALE之间,在一个或多个ZFP或TALE外部及其任何组合。可将本文所述任何融合蛋白配制成药物组合物。

[0019] 在一些实施方案中,如本文所述的工程化DNA结合结构域可置于和作为融合蛋白的一部分的核酸酶(裂解)结构域的操作性连接中。在其它实施方案中,核酸酶体系例如

CRISPR/Cas体系可与特异性单导向RNA一起用于将核酸酶靶向DNA中的靶位置。在某些实施方案中,此类核酸酶和核酸酶融合物可用于靶向干细胞例如诱导性多能干细胞(iPSC)、人胚胎干细胞(hESC)、间充质干细胞(MSC)或神经元干细胞中的突变Htt等位基因,其中核酸酶融合物的活性将产生含野生型数量的CAG重复的Htt等位基因。在某些实施方案中,提供了包含经修饰干细胞的药物组合物。

[0020] 再一方面,提供了编码本文所述任何DNA结合蛋白的多核苷酸。另一方面,提供了编码CRISPR/Cas核酸酶和单导向RHA的多核苷酸。可将此类多核苷酸施用给需要治疗亨廷顿氏病的受试者。

[0021] 又一方面,本发明提供了产生用于研究亨廷顿氏病的特异性模型体系的方法和组合物。在某些实施方案中,本文提供了这样的模型,其中使用胚胎干细胞产生突变Htt等位基因以使用锌指核酸酶(ZFN)、TALE-核酸酶(TALEN)或CRISPR/Cas核酸酶驱动的靶向整合生成其中向野生型Htt等位基因中插入了特定长度的三核苷酸扩增簇(例如,50、80、109和180个CAG重复)的细胞和动物系。在某些实施方案中,模型体系包含体外细胞系,而在其它实施方案中,模型体系包含转基因动物。在本文所述任何动物模型中,例如动物可能为啮齿动物(例如,大鼠、小鼠)、灵长类动物(例如,非人灵长类动物)或兔。

[0022] 再一方面,提供了包含本文所述任何多核苷酸的基因递送载体。在某些实施方案中,载体为腺病毒载体(例如,Ad5/F35载体)、慢病毒载体(LV)(包括能整合型或整合缺陷型慢病毒载体)或腺病毒相关病毒载体(AAV)。因此,本文还提供了包含编码至少一种核酸酶(ZFN或TALEN)的序列和/或用于靶向整合到靶基因中的供体序列的腺病毒(Ad)载体、LV或腺病毒相关病毒载体(AAV)。在某些实施方案中,Ad载体为嵌合Ad载体,例如Ad5/F35载体。在某些实施方案中,慢病毒载体为整合酶缺陷型慢病毒载体(IDLV)或能整合型慢病毒载体。在某些实施方案中,载体为具有VSV-G包膜或具有其它包膜的假型。

[0023] 在一些实施方案中,提供模型体系用于亨廷顿氏病,其中靶等位基因(例如,突变Htt)经表达标志标记。在某些实施方案中,突变等位基因(例如,突变Htt)经标记。在一些实施方案中,野生型等位基因(例如,野生型Htt)经标记,并且在另外的实施方案中,野生型和突变等位基因经单独表达标志标记。在某些实施方案中,模型体系包含体外细胞系,而在其它实施方案中,模型体系包含转基因动物。

[0024] 另外,还提供了包含核酸和/或蛋白质(例如,ZFP或TALE或包含ZFP或TALE的融合蛋白)的药物组合物。例如,某些组合物包含核酸与药学上可接受的载体或稀释剂组合,所述核酸包含编码本文所述ZFP或TALE之一与调控序列可操作地连接的序列,其中所述调控序列允许所述核酸在细胞中表达。在某些实施方案中,编码的ZFP或TALE对HD Htt等位基因有特异性。在一些实施方案中,药物组合物包含调节HD Htt等位基因的ZFP或TALE和调节神经营养因子的ZFP或TALE。基于蛋白质的组合物包含如本文公开的多种ZFP或TALE之一和药学上可接受的载体或稀释剂。

[0025] 再一方面,还提供了包含如本文所述任何蛋白质、多核苷酸和/或组合物的分离细胞。

[0026] 另一方面,本文提供了使用本文所述方法和组合物治疗和/或预防亨廷顿氏病的方法。在一些实施方案中,所述方法涉及其中可使用病毒载体、非病毒载体(例如,质粒)和/或其组合递送多核苷酸和/或蛋白质的组合物。在一些实施方案中,所述方法涉及包含含

ZFP或TALE或经本发明的ZFN、TALEN或CRISPR/Cas核酸酶体系改变的干细胞群的组合物。

[0027] 总体上根据公开内容,这些和其它方面将对技术人员显而易见。

[0028] 附图简述

[0029] 图1,图A至E,为描绘了野生型和突变(亨廷顿氏病,HD)亨廷顿蛋白(Htt)等位基因和与那些等位基因结合的各种ZFP-TF的示意图。图1A示出了结合在CAG重复外,并且因此预测同样与野生型等位基因和突变(HD)等位基因结合的ZFP设计。“KRAB”指来自KOX1基因的KRAB阻遏结构域并且“ZFP”指锌指DNA结合蛋白。“标准ZFP TF”为ZFP转录因子融合蛋白,其中锌指DNA结合结构域与KRAB阻遏结构域连接。图1B示出了设计为在CAG区内结合的ZFP-TF。图1C描绘了“双手ZFP TF”,其为其中两组锌指结构域经刚性蛋白质序列分离的ZFP转录因子。在该图中描述了在一个ZFP外部的功能(阻遏)结构域,但是明显的是,功能结构域可能介于ZFP之间或在蛋白质任一末端的ZFP外部。图1D描绘了“多聚化ZFP TF”,其为能够通过多聚化结构域(如斑点框所示)多聚化的ZFP TF。图1E描绘了ZFP-ZFP-KRAB构型,其中两个锌指DNA结合结构域通过柔性接头连接并且还和KRAB结构域融合。对技术人员而言明显的是,在所有融合蛋白中,功能结构域可在DNA结合结构域的任一末端,并且DNA结合结构域可能包含大量锌指。图1中如带黑色菱形的方框还描绘了功能结构域(例如,激活、阻遏、裂解结构域)。对技术人员而言明显的是,图中呈现的示例性模型也可适用于TALE TF。

[0030] 图2,图A至E,描绘了使用不与CAG重复序列结合的ZFP TF图1A中所述ZFP对Htt两个等位基因的阻遏。在柱下方指示了如表1A和1B中所示的ZFP标识号。图2A描绘了使用靶向人基因中的5个基因座的ZFP,对HEK293细胞中的人Htt等位基因的阻遏。示出了人Htt基因的图表并且示出了ZFP结合位点的位置。对于每个ZFP组而言,每根柱表示独立转染。图2B描绘了显示经GFP对照或18856ZFP TF阻遏因子(包含KOX1的KRAB阻遏结构域)转染的HEK293细胞中的Htt蛋白质水平的蛋白质印迹,其中NF B p65水平(“p65”)用于确认等量蛋白质负载。蛋白质印迹确认Htt表达受ZFP-TF阻遏。图2C描绘了对于Neuro2A细胞中小鼠Htt特异性ZFP,与图2A相似的一组数据。如图2A所示,示出了小鼠Htt基因的图表并且指示了ZFP结合位点的位置。图2D和2E证明了永生化纹状体细胞中小鼠Htt基因表达(RNA)的阻遏,其中将不同剂量的ZFP-TF mRNA用于转染。在除图2B外的所有情况下,通过实时RT-PCR测量Htt mRNA水平并标准化为肌动蛋白mRNA的水平。

[0031] 图3,图A至G,描绘了使用如图1B所示在CAG重复区内结合的ZFP对突变Htt的选择性阻遏。这个模型说明突变等位基因中CAG重复区越长,允许CAG靶向ZFP阻遏分子的结合越多。图3A描绘了HEK293细胞中CAG靶向ZFP对内源性Htt基因(具有正常CAG重复长度)的不同阻遏因子活性。图3B示出了受含有不同长度的CAG重复,范围从10至47个CAG重复的Htt启动子/外显子1片段控制的荧光素酶报告基因的阻遏。CAG10(对于所示两种情况的每一种,最左侧柱)示出了具有10个CAG重复的结果;CAG17(对于所示两种情况的每一种,左侧第二根柱)示出了具有17个CAG重复的结果;CAG23(对于所示两种情况的每一种,右侧第二根柱)示出了具有23个CAG重复的结果;并且CAG47(对于所示两种情况的每一种,最右侧柱)示出了具有47个CAG重复的结果。图上方的示意图描绘了用于该体系中的Htt启动子、外显子1、CAG重复和荧光素酶报告基因的排列。数据证明增加CAG的数量导致CAG靶向的ZFP降低从Htt启动子的表达。此外,图3C证明,虽然相对较弱的CAG靶向ZFP不阻遏含有正常长度CAG重复以及强CAG阻遏因子的荧光素酶报告基因,但是其在试验的所有剂量下,驱动含有作为强CAG

靶向ZFP的扩增CAG重复的荧光素酶报告基因的相似阻遏。“pRL-Htt-CAG23-内含子1”（每一对的左侧柱）对应于从野生型等位基因的表达，而“pRL-HttCAG47-内含子1”（每一对的右侧柱）与从突变扩增Htt等位基因（含47个CAG重复）的表达相关联。图3D为描绘了源自HdH（Q111/Q7）敲入小鼠的永生化小鼠纹状体细胞中CAG靶向ZFP对突变Htt（111CAG）的阻遏的图表。在每一对的左侧柱中示出了野生型表达而在每一对的右侧柱中示出了敲入表达。在转染中使用3个不同浓度的ZFP mRNA试验了包含与KRAB阻遏结构域融合的指定ZFP的ZFP-TF。图3E描绘了通过HD患者来源的成纤维细胞系（CAG15/70）中的CAG靶向ZFP的突变Htt阻遏。在这种成纤维细胞系中，野生型Htt等位基因包含15个CAG重复（“099T（CAG15）”，每种指示情况的中间柱）并且突变扩增Htt等位基因包含70个CAG重复（“099C（CAG70）”，每种指示情况的右侧柱）。图3F示出了在4种不同HD患者来源的成纤维细胞系中，对突变Htt表达的选择性阻遏。每个分组上方的数字表示野生型Htt等位基因（例如15或18）和突变等位基因（例如70、67、45和44）的CAG重复的数量；其中试验了两个不同剂量的ZFP mRNA。每一对的左侧柱示出了野生型Htt表达而每一对的右侧柱示出了突变Htt的表达。图3G描绘了在ZFP-TF 30640、32528和30657的存在下，通过蛋白质印迹分析测定的HD患者来源的成纤维细胞中的Htt表达。移动较慢的蛋白质条带是扩增突变Htt等位基因产生的蛋白质条带。32528与Htt的转录起始位点（TSS）结合并且因此抑制从两个等位基因的表达，而30640和30657与CAG重复（CAG）结合。

[0032] 图4，图A和B，描绘了HD患者来源的成纤维细胞系中突变Htt受一组靶向CAG重复的ZFP的阻遏。使用一系列RNA浓度，从0.1ng至3 $\mu$ g。在图4A和4B中，每种指示情况的左侧柱示出了总Htt表达，中间柱示出了成纤维细胞中Htt的表达，其中Htt等位基因包含18个CAG重复（“099T（CAG18）”）而突变扩增Htt等位基因包含45个CAG重复099T（CAG45）

[0033] 图5示出了CAG靶向ZFP阻遏因子对HD患者来源的成纤维细胞中Htt和其它含CAG的基因的表达的影响。每种指示情况下的左侧柱示出了30640的结果；每种指示情况下的中间柱示出了30675的结果；并且右侧柱示出了模拟转染。

[0034] 图6，图A和B，描绘了检查3种CAG靶向ZFP的全基因组特异性的实验。图6A描绘了使用30640、30645或33074对HD成纤维细胞（CAG18（中间柱）/CAG45（右侧柱））的6种生物复制物（6次单独转染）进行的Htt阻遏qPCR分析。然后通过qPCR选择4种最相似的复制物进行微阵列分析，并将数据呈现于图6B中。

[0035] 图7描绘了CAG17/69神经元干细胞（NSC）中的Htt阻遏。细胞经指示剂量的ZFP mRNA转染。每个指示剂量下的左侧柱示出了在CAG17细胞中的结果，中间柱示出了在野生型细胞中的结果，右侧柱示出了在CAG69细胞中的结果。

[0036] 图8描绘了由经ZFP TF处理的HD胚胎干细胞（ESC）（CAG 17/48）分化的神经元中的Htt表达。细胞经指示剂量的ZFP mRNA转染。

[0037] 图9描绘了用ZFP TF 30640处理后，在R6/2小鼠中突变Htt转基因表达的阻遏。

[0038] 图10，图A至D，描绘了如图1D所示，具有特异性靶向扩增CAG重复的多聚化结构域的ZFP。图10A示出了具有4个组成部分的单个ZFP：(i) KOX阻遏因子结构域（椭圆形标记的“阻遏因子”）；(ii) 与(CAG)<sub>N</sub>或这个序列的置换结合的2-6指阵列（示出两指，小椭圆形标记的“Z”）；和(iii) 呈反平行构型相互作用的两个二聚化结构域（矩形标记的“d1”和“d2”）。这些结构域允许ZFP在CAG簇的大沟中聚合。图10B示出了与3个ZFP的多聚体的结合事件的草

图。将明显的是,可使用任何数量的多聚体并且功能结构域可位于一个或多个单独ZFP的任何地方,并且这些图也适于TALE-TF。图10C示出了设计为通过二聚化锌指(DZ)之间的相互作用多聚化的4个ZFP单体支架的蛋白质序列。支架命名为DZ1(SEQ ID NO:180)、DZ2(SEQ ID NO:181)、DZ3(SEQ ID NO:182)和DZ4(SEQ ID NO:183)。二聚化锌指结构域加有下划线,而阻遏结构域和核定位序列用粗体下划线和斜体字(分别)示出。图10D示出了设计为通过卷曲螺旋(CC)之间的相互作用多聚化的7个ZFP单体支架的蛋白质序列。支架命名为CC1(SEQ ID NO:184)、CC2(SEQ ID NO:185)、CC3(SEQ ID NO:186)、CC4(SEQ ID NO:187)、CC5(SEQ ID NO:188)、CC6(SEQ ID NO:189)和CC7(SEQ ID NO:190)。卷曲螺旋序列加有下划线,而阻遏结构域和核定位序列用粗体下划线和斜体字(分别)示出。用“[ZFP]”指示将随设计改变的每个支架的ZFP区位置。用“[ZFP]”指示将随设计改变的每个支架的(DNA结合)ZFP区的位置。

[0039] 图11,图A和B,描绘了具有二聚化结构域的ZFP-TF的活性。在图11A中,用荧光素酶报告基因试验了具有“卷曲螺旋”(CC)结构域的ZFP-TF。pRL-Htt CAG17(每一对的左侧柱)代表受具有17个CAG的人Htt启动子/外显子1片段控制的海肾荧光素酶报告基因;pGL3-Htt-CAG47(每一对的右侧柱)代表受具有47个CAG重复的人Htt启动子/外显子1片段控制的萤火虫荧光素酶报告基因。对各二聚化结构域的描述,见实施例10中的文本。在图11B中,用相同荧光素酶报告基因试验了具有二聚化锌指“DZ”结构域的ZFP,并且证明用一些ZFP-TF二聚化结构域,阻遏增加。每一对中的左侧柱指示由17CAG重复Htt等位基因的表达,而右侧柱指示由47CAG重复Htt等位基因的表达。

[0040] 图12,图A和B,描绘了ZFP-ZFP-KOX蛋白质对Htt的阻遏。图12A描绘了在野生型(左侧柱)、CAG18(中间柱)和CAG45(右侧柱)(图12A)HD成纤维细胞中受单个33088和33084ZFP-TF的Htt阻遏和受33088-33088和33088-33084ZFP-ZFP-KOX蛋白质的阻遏;图12B描绘了在野生型(左侧柱)、CAG20(中间柱)和CAG41(右侧柱)HD成纤维细胞中受33088-33088和33088-33084ZFP-ZFP-KOX的Htt阻遏。

[0041] 图13,图A至E,描绘了小鼠Htt的激活。图13A证明了使用与p65激活结构域融合的ZFP,Neuro2A细胞中在RNA水平上,ZFP-TF驱动的小鼠Htt基因上调。双柱指示重复转染。图13B描绘了证明受ZFP驱动Htt蛋白质生成增加的蛋白质印迹。图13C描绘了野生型小鼠Htt等位基因和“敲入”Htt等位基因,其中小鼠序列(大部分外显子1和部分内含子1,野生型等位基因示意图上面的线)已经被经CAG扩增的相应人序列(敲入等位基因示意图上面的线)置换。图13D描绘了小鼠序列(SEQ ID NO:191)之间的比对,其经相应人序列(SEQ ID NO:192)置换,以致敲入等位基因具有足以允许将ZFP(示于A和B中)设计为特异性结合小鼠序列的序列偏差。图13E描绘了在源自HdhQ111/Q7敲入小鼠的永生化纹状体细胞中野生型小鼠Htt等位基因的特异性激活。左侧柱示出了野生型细胞的结果而右侧柱示出了敲入突变等位基因细胞的结果。

[0042] 图14,图A和B,描绘了用Htt特异性ZFN对处理K562细胞后,Cel-I错配测定(Surveyor™,Transgenomics)的结果。在相应条带底部示出了活性ZFN的NHEJ活性百分比(插入-缺失(in-del))。“GFP”指示已经用GFP编码质粒转染的细胞。图14A描绘了来自早期裂解Htt外显子的ZFN的结果而图14B描绘了来自在终止密码子附近裂解的ZFN的结果。也观察到失活ZFN对(未标注插入-缺失百分比的条带)。

[0043] 图15描绘了几种候选TALE-TF蛋白的Htt阻遏结果的图表。在HD患者来源的成纤维细胞(CAG 20/41)中试验了TALE-TF。结果证明一些TALE TF在阻遏总体Htt表达中有活性,而其它表现出突变Htt优先阻遏。

[0044] 发明详述

[0045] 本文公开了治疗亨廷顿氏病(HD)的组合物和方法。具体而言,提供了包含锌指蛋白(ZFP TF)或TALE(TALE-TF)的Htt调节转录因子和利用此类蛋白的方法用于治疗或预防亨廷顿氏病。例如,提供了阻遏突变Htt等位基因表达或激活野生型Htt等位基因表达的ZFP-TF或TALE-TF。另外,提供了修饰HD相关基因的基因组结构的锌指核酸酶(ZFN)、TALE核酸酶(TALEN)或CRISPR/Cas核酸酶体系。例如,提供了能够特异性改变部分突变形式的Htt的ZFN、TALEN或CRISPR/Cas核酸酶体系。这些包括使用工程化锌指蛋白或工程化TALE蛋白,即与预定核酸靶序列结合的非天然存在的蛋白质的组合物和方法。

[0046] 因此,本文所述方法和组合物提供了治疗和预防亨廷顿氏病的方法,并且这些方法和组合物可包含能够调节靶基因的锌指转录因子或TALE转录因子以及能够修饰或编辑Htt的工程化锌指和TALE核酸酶和CRISPR/Cas核酸酶体系。

[0047] 概述

[0048] 除非另外指出,所述方法的实践以及本文公开的组合物的制备和使用采用分子生物学、生物化学、染色质结构和分析、计算化学、细胞培养、重组DNA和属于本领域技术范围内的相关领域的常规技术。在文献中全面解释了这些技术。见,例如,Sambrook等,MOLECULAR CLONING:A LABORATORY MANUAL,第2版,Cold Spring Harbor Laboratory Press,1989和第3版,2001;Ausubel等,CURRENT PROTOCOLS IN MOLECULAR BIOLOGY,John Wiley&Sons,New York,1987和定期更新;METHODS IN ENZYMOLOGY系列,Academic Press, San Diego;Wolffe,CHROMATIN STRUCTURE AND FUNCTION,第3版,Academic Press, San Diego,1998;METHODS IN ENZYMOLOGY,第304卷,“Chromatin”(P.M.Wasserman和A.P.Wolffe编辑),Academic Press, San Diego,1999;和METHODS IN MOLECULAR BIOLOGY,第119卷,“Chromatin Protocols”(P.B.Becker编辑)Humana Press,Totowa, 1999。

[0049] 定义

[0050] 术语“核酸”、“多核苷酸”和“寡核苷酸”可交换使用并且指呈线性或环状构象并且呈单链或双链形式的脱氧核糖核苷酸或核糖核苷酸聚合物。为了本公开的目的,这些术语不得视为关于聚合物长度的限制。术语可涵盖天然核苷酸的已知类似物以及在碱基、糖和/或磷酸部分(硫代磷酸骨架)中经修饰的核苷酸。一般而言,特定核苷酸的类似物具有相同碱基配对特异性;即,A的类似物将与T碱基配对。

[0051] 术语“多肽”、“肽”和“蛋白质”可交换使用并且指氨基酸残基的聚合物。术语还适用于其中一个或多个氨基酸为相应天然存在的氨基酸的化学类似物或经修饰衍生物的氨基酸聚合物。

[0052] “结合”指大分子之间(例如蛋白质和核酸之间)的序列特异性、非共价相互作用。并非结合相互作用的所有组成部分均需有序列特异性(例如,与DNA骨架中的磷酸碱基接触),只要总体上相互作用有序列特异性。此类相互作用通常特征在于解离常数( $K_d$ )为 $10^{-6}$ M<sup>-1</sup>或更低。“亲和力”指结合强度:结合亲和力增加与较低的 $K_d$ 相关联。

[0053] “结合蛋白”是能够与另一分子非共价结合的蛋白质。例如,结合蛋白可与DNA分子(DNA结合蛋白)、RNA分子(RNA结合蛋白)和/或蛋白质分子(蛋白质结合蛋白)结合。在为蛋白质结合蛋白的情况下,其可与自身结合(形成同型二聚体、同型三聚体等)和/或其可与不同蛋白质的一个或多个分子结合。结合蛋白可具有一种以上类型的结合活性。例如,锌指蛋白具有DNA结合、RNA结合和蛋白质结合活性。

[0054] “锌指DNA结合蛋白”(或结合结构域)是以序列特异性方式,通过一个或多个锌指结合DNA的蛋白质或较大蛋白质中的结构域,锌指是其结构通过锌离子配位稳定的结合结构域中的氨基酸序列区域。术语锌指DNA结合蛋白常常缩写为锌指蛋白或ZFP。

[0055] “TALE DNA结合结构域”或“TALE”是包含一个或多个TALE重复结构域/单元的多肽。重复结构域牵涉于TALE与其同源靶DNA序列的结合中。单个“重复单元”(也称为“重复”)通常长度为33-35个氨基酸并且与天然存在的TALE蛋白中的其它TALE重复序列至少表现出一定的序列同源性。

[0056] 例如,可通过工程化天然存在的锌指蛋白的识别螺旋区(或改变一个或多个氨基酸)或通过工程化TALE蛋白的RVD将锌指结合结构域或TALE DNA结合结构域“工程化”为与预定核苷酸序列结合。因此,工程化锌指蛋白或TALE是非天然存在的蛋白质。工程化锌指蛋白或TALE的方法的非限制性实例为设计和选择。设计的锌指蛋白或TALE是在自然界中不存在的蛋白质,其设计/组成主要由合理标准产生。设计的合理标准包括应用取代规则和计算机算法处理保存现有ZFP设计和结合数据信息的数据库中的信息。见,例如,美国专利6,140,081、6,453,242和6,534,261;还见WO 98/53058、WO 98/53059、WO 98/53060、WO 02/016536和WO 03/016496和美国公布第20110301073号。

[0057] “选定的”锌指蛋白或TALE是在自然界中未发现的蛋白质,其生成主要由经验过程引起,例如噬菌体展示、相互作用陷阱或杂交选择。见例如US 5,789,538、US 5,925,523、US 6,007,988、US 6,013,453、US 6,200,759、WO 95/19431、WO 96/06166、WO 98/53057、WO 98/54311、WO 00/27878、WO 01/60970、WO 01/88197、WO 02/099084和WO 2011/146121(美国专利公布第20110301073号)。

[0058] “重组”指两个多核苷酸之间交换遗传信息的过程。为了本公开的目的,“同源重组(HR)”指例如在细胞中通过同源介导修复机制修复双链断裂期间发生的此类交换的特化形式。这个过程需要核苷酸序列同源性,将“供体”分子用于“靶”分子(即,经历双链断裂的分子)的模板修复,并且因为其导致遗传信息从供体转移到靶标,所以不同地称为“非交换型基因转化(non-crossover gene conversion)”或“短序列基因转化(short tract gene conversion)”。不希望受任何特定理论约束,此类转移可涉及在断裂靶标和供体之间形成的异源双链体DNA的错配校正和/或其中供体用于重新合成将成为靶标的一部分的遗传信息的“合成依赖式链退火”和/或相关过程。此类特化HR常常导致靶分子的序列改变,以致供体多核苷酸的部分或全部序列并入靶多核苷酸中。

[0059] 在本公开的方法中,如本文所述的一种或多种靶向核酸酶在靶序列(例如,细胞染色质)的预定位点产生双链断裂,并且可将与断裂区内的核苷酸序列具有同源性的“供体”多核苷酸引入细胞中。已经证实双链断裂的存在利于供体序列的整合。供体序列可物理整合,或可选地,将供体多核苷酸用作通过同源重组修复断裂的模板,导致与供体中一样,所有或部分核苷酸序列引入细胞染色质中。因此,细胞染色质中的第一序列可改变并且,在某

些实施方案中,可转化为供体多核苷酸中存在的序列。因此,术语“置换(replace)”或“置换(replacement)”的使用可理解为表示一个核苷酸序列经另一个置换,(即,在信息意义上置换序列),并不一定需要一个多核苷酸经另一个物理或化学置换。

[0060] 在本文所述的任何方法中,附加对的锌指或TALE蛋白可用于细胞中附加靶位点的附加双链裂解。

[0061] 在靶向重组和/或置换和/或改变细胞染色质中目标区域内的序列的方法的某些实施方案中,通过与外源“供体”核苷酸序列的同源重组改变染色体序列。如果存在与断裂区域同源的序列,则细胞染色质中双链断裂的存在刺激此类同源重组。

[0062] 在本文所述任何方法中,第一核苷酸序列(“供体序列”)可含有与目标区域内的基因组序列同源,但是不相同的序列,从而刺激同源重组以将不同序列插入目标区域。因此,在某些实施方案中,与目标区域内的序列同源的供体序列的部分表现出与被置换基因组序列约80-99%(或之间的任何整数)序列同一性。在其它实施方案中,例如如果在100个以上连续碱基对的供体和基因组序列之间仅1个核苷酸不同,则供体和基因组序列之间的同源性高于99%。在某些情况下,供体序列的非同源部分可含有在目标区域内不存在的序列,以致将新序列引入目标区域。在这些情况下,非同源序列两侧通常为与目标区域内的序列同源或相同的50-1,000个碱基对(或之间的任何整数值)或大于1,000的任何数量的碱基对的序列。在其它实施方案中,供体序列与第一序列非同源,并且通过非同源重组机制插入基因组中。

[0063] 本文所述任何方法均可用于通过靶向整合破坏目标基因表达的供体序列,使细胞中的一个或多个靶序列部分或完全失活。还提供了具有部分或完全失活基因的细胞系。

[0064] 此外,如本文所述靶向整合的方法也可用于整合一个或多个外源序列。例如,外源性核酸序列可包含一个或多个基因或cDNA分子或任何类型的编码或非编码序列以及一个或多个控制元件(例如,启动子)。另外,外源性核酸序列可产生一种或多种RNA分子(例如,小发夹RNA(shRNA)、抑制RNA(RNAi)、微RNA(miRNA)等)。

[0065] “裂解”指DNA分子的共价骨架的断裂。可通过多种方法,包括但不限于酶或化学水解磷酸二酯键引发裂解。单链裂解和双链裂解均有可能,并且双链裂解可由于两个不同单链裂解事件而发生。DNA裂解可导致产生平末端或交错末端。在某些实施方案中,融合多肽用于靶向双链DNA裂解。

[0066] “裂解半结构域”是连同第二多肽(相同或不同)一起形成具有裂解活性(优选为双链裂解活性)的复合物的多肽序列。术语“第一和第二裂解半结构域”、“+和-裂解半结构域”和“右侧和左侧裂解半结构域”可交换用于指二聚化的成对裂解半结构域。

[0067] “工程化裂解半结构域”是已经修饰以便与另一裂解半结构域(例如,另一工程化裂解半结构域)一起形成专性杂二聚体的裂解半结构域。同样见通过引用整体并入本文的美国专利公布第2005/0064474、20070218528、2008/0131962和2011/0201055号。

[0068] 术语“序列”指可为DNA或RNA,可为直链、环状或支链并且可为单链或双链的任何长度的核苷酸序列。术语“供体序列”指插入基因组中的核苷酸序列。供体序列可具有任何长度,例如长度介于2和10,000个核苷酸之间(或之间或以上的任何整数值),优选长度介于约100和1,000个核苷酸之间(或之间的任何整数值),更优选长度介于约200和500个核苷酸之间。

[0069] “染色质”是包含细胞基因组的核蛋白结构。细胞染色质包含核酸(主要为DNA)和蛋白质,包括组蛋白和非组蛋白染色体蛋白。大部分真核细胞染色质呈核小体形式存在,其中核小体核包含与包含组蛋白H2A、H2B、H3和H4各两个的八聚物缔合的DNA的约150个碱基对;并且接头DNA(根据生物而言长度可变)在核小体核之间延伸。组蛋白H1的分子通常与接头DNA缔合。为了本公开的目的,术语“染色质”意在涵盖所有类型的原核和真核细胞核蛋白。细胞染色质包括染色体和附加体染色质。

[0070] “染色体”是包含细胞的全部或部分基因组的染色质复合物。细胞基因组常常特征在于其核型,这是包含细胞基因组的所有染色体的集合。细胞基因组可包含一条或多条染色体。

[0071] “附加体”为复制核酸、核蛋白复合物或包含并非细胞染色体核型的一部分的核酸的其它结构。附加体的实例包括质粒和某些病毒基因组。

[0072] “靶位点”或“靶序列”是限定结合分子将与之结合的核酸部分的核酸序列,条件是存在供结合的足够条件。

[0073] “外源”分子是细胞中通常不存在,但是可通过一种或多种遗传、生物化学或其它方法引入细胞中的分子。相对于细胞的特定发育阶段和环境条件测定“细胞中正常存在”。因此,例如,仅在肌肉胚胎发育期存在的分子相对于成人肌肉细胞为外源分子。类似地,由热休克诱导的分子相对于非热休克细胞为外源分子。例如,外源分子可包含功能形式的功能障碍内源分子或功能障碍形式的正常功能内源分子。

[0074] 其中,外源分子可为小分子,例如通过组合化学过程生成的小分子,或大分子例如蛋白质、核酸、碳水化合物、脂质、糖蛋白、脂蛋白、多糖、以上分子的任何经修饰衍生物或包含以上一种或多种分子的任何复合物。核酸包括DNA和RNA,可为单链或双链;可为直链、支链或环状;并且可具有任何长度。核酸包括能够形成双链体的核酸以及形成三链体的核酸。见,例如,美国专利第5,176,996和5,422,251号。蛋白质包括但不限于DNA结合蛋白、转录因子、染色质重构因子、甲基化DNA结合蛋白、聚合酶、甲基化酶、脱甲基酶、乙酰基转移酶、脱乙酰基酶、激酶、磷酸酶、整合酶、重组酶、连接酶、拓扑异构酶、促旋酶和解螺旋酶。

[0075] 外源分子可为与内源分子相同类型的分子,例如外源蛋白质或核酸。例如,外源性核酸可包括感染病毒基因组、引入细胞的质粒或附加体或细胞中通常不存在的染色体。将外源分子引入细胞的方法为本领域的技术人员已知并且包括但不限于脂质介导的转移(即,脂质体,包括中性和阳离子脂质)、电穿孔、直接注入、细胞融合、粒子轰击、磷酸钙共沉淀法、DEAE-葡聚糖介导的转移和病毒载体介导的转移。外源分子也可为与内源分子相同类型的分子,但是源自与所述细胞来源不同的物种。例如,人核酸序列可引入最初源自小鼠或仓鼠的细胞系中。

[0076] 相反,“内源”分子是特定细胞中,在特定环境条件下的特定发育阶段通常存在的分子。例如,内源性核酸可包括染色体、线粒体、叶绿体或其它细胞器的基因组或天然存在的附加体核酸。另外的内源分子可包括蛋白质,例如转录因子和酶。

[0077] “融合”分子是其中两个或更多个亚基分子优选共价连接的分子。亚基分子可为相同化学类型的分子,或可为不同化学类型的分子。第一种类型的融合分子的实例包括但不限于融合蛋白(例如,ZFP或TALE DNA结合结构域和一个或多个激活结构域之间的融合)和融合核酸(例如,编码如上所述融合蛋白的核酸)。第二种类型的融合分子的实例包括但不

限于形成三链体的核酸与多肽之间的融合和小沟结合物与核酸之间的融合。

[0078] 细胞中融合蛋白的表达可由向细胞递送融合蛋白引起或通过向细胞递送编码融合蛋白的多核苷酸引起,其中转录多核苷酸并且翻译转录物以生成融合蛋白。在细胞中蛋白质的表达中还可牵涉反式剪接、多肽裂解和多肽连接。在本公开其它地方提出了向细胞递送多核苷酸和多肽的方法。

[0079] “多聚化结构域”(也称为“二聚化结构域”或“蛋白质相互作用结构域”)是在ZFP TF或TALE TF的氨基、羧基或氨基和羧基末端区域并入的结构域。这些结构域允许多个ZFP TF或TALE TF单元多聚化,以致相对于具有野生型数量长度的较短簇,较大簇的三核苷酸重复结构域变得优先由多聚化ZFP TF或TALE TF结合。多聚化结构域的实例包括亮氨酸拉链。多聚化结构域也可能受小分子调控,其中多聚化结构域呈现适当构象以允许仅在小分子或外部配体存在下,与另一多聚化结构域相互作用。这样,外源性配体可用于调控这些结构域的活性。

[0080] 为了本公开的目的,“基因”包括编码基因产物的DNA区域(见下文)以及调控基因产物的生成的所有DNA区域,不论此类调控序列是否与编码和/或转录序列相邻。相应地,基因包括但不一定限于启动子序列、终止子、翻译调控序列(例如核糖体结合位点和内部核糖体进入位点)、增强子、沉默子、绝缘子、边界元件、复制起点、基质附着位点和基因座控制区。

[0081] “基因表达”指将基因内所含信息转化为基因产物。基因产物可为基因的直接转录产物(例如,mRNA、tRNA、rRNA、反义RNA、核酶、结构RNA或任何其它类型的RNA)或通过翻译mRNA生成的蛋白质。基因产物还包括通过例如加帽、聚腺苷酸化、甲基化和编辑的过程修饰的RNA和(例如)通过甲基化、乙酰化、磷酸化、泛素化、ADP-核糖基化、豆蔻酰化和糖基化修饰的蛋白质。

[0082] 基因表达的“调节”指基因活性的变化。表达的调节可包括但不限于基因激活和基因阻遏。基因组编辑(例如,裂解、改变、失活、随机突变)可用于调节表达。基因失活指与不包括如本文所述的ZFP或TALE蛋白质的细胞相比,基因表达的任何减少。因此,基因失活可为部分或全部。

[0083] “目标区域”为细胞染色质的任何区域,例如需要结合外源分子的基因或基因内或相邻的非编码序列。结合可能是为了靶向DNA裂解和/或靶向重组的目的。例如,目标区域可存在于染色体、附加体、细胞器基因组(例如,线粒体、叶绿体)或感染病毒基因组。目标区域可在基因的编码区内、经转录的非编码区例如前导序列、尾随序列或内含子内,或在编码区上游或下游的非转录区内。目标区域长度可小到单个核苷酸对或多达2,000个核苷酸对或任何整数值的核苷酸对。

[0084] “真核”细胞包括但不限于真菌细胞(例如酵母)、植物细胞、动物细胞、哺乳动物细胞和人细胞(例如,T细胞)。

[0085] 术语“操作性连接”和“操作性地连接”(或“可操作地连接”)对于两个或更多个组成部分(例如序列元件)的并置可交换使用,其中组成部分排列成以致两个组成部分正常工作并且允许至少一个组成部分可介导对其它组成部分中的至少一个发挥的作用的可能性。举例而言,如果转录调控序列控制编码序列响应于一个或多个转录调控因子的存在或缺乏而转录的水平,则转录调控序列例如启动子与编码序列操作性地连接。转录调控序列通常

与编码序列顺式操作性地连接,但不需要与之直接相邻。例如,增强子是与编码序列操作性地连接的转录调控序列,即使它们不连续。

[0086] 对于融合多肽,术语“操作性地连接”可指每个组成部分在与其它组成部分的连接中,执行与如果不这样连接时相同的功能。例如,对于其中ZFP或TALE DNA结合结构域与激活结构域融合的融合多肽,如果在融合多肽中,ZFP或TALE DNA结合结构域部分能够结合其靶位点和/或其结合位点,而激活结构域能够上调基因表达,则ZFP或TALE DNA结合结构域和激活结构域处于操作性连接。与能够调节基因表达的结构域融合的ZFP统称为“ZFP-TF”或“锌指转录因子”,而与能够调节基因表达的结构域融合的TALE统称为“TALE-TF”或“TALE转录因子”。对于ZFP DNA结合结构域与裂解结构域(“ZFN”或“锌指核酸酶”)融合的融合多肽,如果在融合多肽中,ZFP DNA结合结构域部分能够结合其靶位点和/或其结合位点,而裂解结构域能够在靶位点附近裂解DNA,则ZFP DNA结合结构域和裂解结构域处于操作性连接。对于TALE DNA结合结构域与裂解结构域(“TALEN”或“TALE核酸酶”)融合的融合多肽,如果在融合多肽中,TALE DNA结合结构域部分能够结合其靶位点和/或其结合位点,而裂解结构域能够在靶位点附近裂解DNA,则TALE DNA结合结构域和裂解结构域处于操作性连接。

[0087] 蛋白质、多肽或核酸的“功能片段”是其序列与全长蛋白质、多肽或核酸不同,但是保持了与全长蛋白质、多肽或核酸相同的功能的蛋白质、多肽或核酸。功能片段可具有比相应天然分子更多、更少或相同数量的残基,和/或可含有一个或多个氨基酸或核苷酸取代。测定核酸功能(例如,编码功能、与另一核酸杂交的能力)的方法在本领域中众所周知。类似地,测定蛋白质功能的方法众所周知。例如,可通过过滤结合、电泳迁移或免疫沉淀测定法测定多肽的DNA结合功能。可通过凝胶电泳测定DNA裂解。见Ausubel等,同上。例如,可通过遗传和生化免疫共沉淀、双杂交测定法或互补作用测定蛋白质与另一种蛋白质相互作用的能力。见,例如,Fields等(1989)Nature 340:245-246;美国专利第5,585,245号和PCT WO 98/44350。

[0088] “载体”能够将基因序列转移到靶细胞中。通常,“载体构建体”、“表达载体”和“基因转移载体”意指能够指导目标基因的表达并且可将基因序列转移到靶细胞中的任何核酸构建体。因此,术语包括克隆和表达媒介物以及整合载体。

[0089] “报告基因”或“报告序列”指生成优选不一定在常规测定中易于测量的蛋白质产物的任何序列。适合的报告基因包括但不限于编码介导抗生素抗性(例如,氨苄青霉素(ampicillin)抗性、新霉素(neomycin)抗性、G418抗性、嘌呤霉素(puromycin)抗性)的蛋白质的序列、编码有色或荧光或发光蛋白质(例如,绿色荧光蛋白、增强型绿色荧光蛋白、红色荧光蛋白、荧光素酶)和介导增强细胞生长和/或基因扩增的蛋白质(例如,二氢叶酸还原酶)的序列。表位标签包括(例如)一个或多个拷贝的FLAG、His、myc、Tap、HA或任何可检测的氨基酸序列。“表达标签”包括编码可能与所需基因序列可操作地连接以便监测目标基因的表达的报告基因的序列。

[0090] DNA结合结构域

[0091] 本文描述了包含与包含三核苷酸重复的任何基因,包括但不限于Htt中的靶序列特异性结合的DNA结合结构域的组合物。任何DNA结合结构域均可用于本文公开的组合物和方法中。

[0092] 在某些实施方案中,DNA结合结构域包含锌指蛋白。优选地,锌指蛋白为非天然存

在, 因为其经工程化为与所选靶位点结合。见, 例如, Beerli等(2002) *Nature Biotechnol.* 20:135-141; Pabo等(2001) *Ann. Rev. Biochem.* 70:313-340; Isalan等(2001) *Nature Biotechnol.* 19:656-660; Segal等(2001) *Curr. Opin. Biotechnol.* 12:632-637; Choo等(2000) *Curr. Opin. Struct. Biol.* 10:411-416; 美国专利第6,453,242、6,534,261、6,599,692、6,503,717、6,689,558、7,030,215、6,794,136、7,067,317、7,262,054、7,070,934、7,361,635、7,253,273号; 和美国专利公布第2005/0064474、2007/0218528、2005/0267061号, 全部通过引用整体并入本文。

[0093] 与天然存在的锌指蛋白相比, 工程化锌指结合结构域可具有新型的结合特异性。工程化方法包括但不限于合理的设计和各種类型的选择。合理的设计包括(例如) 使用包含三联体(或四联体) 核苷酸序列和单独锌指氨基酸序列的数据库, 其中每个三联体或四联体核苷酸序列与锌指的结合特定三联体或四联体序列的一个或多个氨基酸序列缔合。见, 例如, 共同拥有的美国专利6,453,242和6,534,261, 其通过引用整体并入本文。

[0094] 在美国专利5,789,538、5,925,523、6,007,988、6,013,453、6,410,248、6,140,466、6,200,759和6,242,568以及WO 98/37186、WO 98/53057、WO 00/27878、WO 01/88197和GB 2,338,237中公开了示例性选择方法, 包括噬菌体展示和双杂交体系。另外, 例如在共同拥有的WO 02/077227中已描述了锌指结合结构域结合特异性的增强。

[0095] 另外, 如这些和其它参考文献中所公开那样, 锌指结构域和/或多指锌指蛋白可使用任何适合的接头序列, 包括例如长度为5个或更多个氨基酸的接头连接在一起。对于长度为6个或更多个氨基酸的示例性接头, 同样见美国专利第6,479,626、6,903,185和7,153,949号。本文所述蛋白质可包括在蛋白质的单独锌指之间适合接头的任何组合。另外, 例如在共同拥有的WO 02/077227中描述了对于锌指结合结构域的结合特异性的增强。

[0096] 靶位点的选择、ZFP和设计和构建融合蛋白(和编码融合蛋白的多核苷酸)的方法为本领域的技术人员已知并且在美国专利第6,140,081、5,789,538、6,453,242、6,534,261、5,925,523、6,007,988、6,013,453、6,200,759号、WO 95/19431、WO 96/06166、WO 98/53057、WO 98/54311、WO 00/27878、WO 01/60970、WO 01/88197、WO 02/099084、WO 98/53058、WO 98/53059、WO 98/53060、WO 02/016536和WO 03/016496中有详细描述。

[0097] 另外, 如这些和其它参考文献中所公开那样, 锌指结构域和/或多指锌指蛋白可使用任何适合接头序列, 包括例如长度为5个或更多个氨基酸的接头连接在一起。对于长度为6个或更多个氨基酸的示例性接头, 同样见美国专利第6,479,626、6,903,185和7,153,949号。本文所述蛋白质可包括在蛋白质的单独锌指之间适合接头的任何组合。

[0098] 在某些实施方案中, DNA结合结构域为与Htt基因中的靶位点(以序列特异性方式)结合并调节Htt表达的工程化锌指蛋白。ZFP可与突变Htt等位基因或野生型Htt序列选择性结合。Htt靶位点通常包括至少一个锌指, 但是可包括多个锌指(例如, 2、3、4、5、6指或更多指)。通常, ZFP包括至少3指。某些ZFP包括4、5或6指, 而一些ZFP包括8、9、10、11或12指。包括3指的ZFP通常识别包括9或10个核苷酸的靶位点; 包括4指的ZFP通常识别包括12至14个核苷酸的靶位点; 而具有6指的ZFP可识别包括18至21个核苷酸的靶位点。ZFP也可为包括一个或多个调控结构域的融合蛋白, 所述结构域可为转录激活或阻遏结构域。在一些实施方案中, 融合蛋白包含两个连接在一起的ZFP DNA结合结构域。因此这些锌指蛋白可包含8、9、10、11、12指或更多指。在一些实施方案中, 两个DNA结合结构域经由可延伸柔性接头连接,

以致一个DNA结合结构域包含4、5或6个锌指而第二DNA结合结构域包含另外4、5或5个锌指。在一些实施方案中,接头为标准指间接头,以致指阵列包含一个含8、9、10、11、12指或更多指的DNA结合结构域。在其它实施方案中,接头为非典型接头例如柔性接头。DNA结合结构域与至少一个调控结构域融合并且可看作‘ZFP-ZFP-TF’构造。这些实施方案的特定实例可称为包含经柔性接头连接并且与KOX阻遏因子融合的两个DNA结合结构域的“ZFP-ZFP-KOX”和其中两个ZFP-KOX融合蛋白经由接头融合在一起的“ZFP-KOX-ZFP-KOX”。

[0099] 可选地,DNA结合结构域可源自核酸酶。例如,已知归巢内切核酸酶和大范围核酸酶的识别序列,例如I-SceI、I-CeuI、PI-PspI、PI-Sce、I-SceIV、I-CsmI、I-PanI、I-SceII、I-PpoI、I-SceIII、I-CreI、I-TevI、I-TevII和I-TevIII。还见美国专利第5,420,032号、美国专利第6,833,252号;Belfort等(1997) *Nucleic Acids Res.* 25:3379-3388;Dujon等(1989) *Gene* 82:115-118;Perler等(1994) *Nucleic Acids Res.* 22,1125-1127;Jasin(1996) *Trends Genet.* 12:224-228;Gimble等(1996) *J. Mol. Biol.* 263:163-180;Argast等(1998) *J. Mol. Biol.* 280:345-353和New England Biolabs目录。另外,归巢内切核酸酶和大范围核酸酶的DNA结合特异性可经工程化为结合非天然靶位点。见,例如,Chevalier等(2002) *Molec. Cell* 10:895-905;Epinat等(2003) *Nucleic Acids Res.* 31:2952-2962;Ashworth等(2006) *Nature* 441:656-659;Paques等(2007) *Current Gene Therapy* 7:49-66;美国专利公布第20070117128号。

[0100] “双手”锌指蛋白是其中通过插入氨基酸分隔开两簇锌指DNA结合结构域,以致所述两个锌指结构域与两个不连续的靶位点结合的那些蛋白质。双手型锌指结合蛋白的实例为SIP1,其中一簇4个锌指位于蛋白质的氨基末端并且一簇3指位于羧基末端(见Remacle等,(1999) *EMBO Journal* 18(18):5073-5084)。这些蛋白质中的每一簇锌指均能够与唯一靶序列结合并且两个靶序列之间的间隔可包含许多核苷酸。双手ZFP可能包括(例如)与一个或两个ZFP融合的功能结构域。因此,将明显的是,功能结构域可能连接至一个或两个ZFP外部(见,图1C)或可能位于ZFP之间(连接至两个ZFP)(见,图4)。

[0101] 在表1A和1B中公开了Htt靶向的ZFP的具体实例。这个表中第一列为ZFP的内部引用名称(编号)并且与表2A和2B第1列中的相同名称相对应。“F”指的是指并且“F”后面的数字指哪个锌指(例如,“F1”指的是指1)。

[0102] 表1A:Htt靶向的锌指蛋白

[0103]

SBS #	设计					
	F1	F2	F3	F4	F5	F6
18856	RSDDLRS (SEQ ID NO: 1)	RNDNRK (SEQ ID NO:2)	RSDDLTR (SEQ ID NO:3)	RSDDRKT (SEQ ID NO:4)	RSADLTR (SEQ ID NO:5)	QSSDLRR (SEQ ID NO:6)
25920	RSAALSR (SEQ ID NO: 58)	RSDALAR (SEQ ID NO: 59)	RSDNLSE (SEQ ID NO: 60)	KRCNLRC (SEQ ID NO: 61)	QSSDLRR (SEQ ID NO: 6)	NA
25921	WRSCRSA (SEQ ID NO: 62)	DRSNLSR (SEQ ID NO: 9)	QRHLTQ (SEQ ID NO: 53)	RSAHLRS (SEQ ID NO: 46)	TSGHLRS (SEQ ID NO: 43)	NA
25923	RSDDLRS (SEQ ID NO: 1)	RNDNRK (SEQ ID NO: 2)	WRSCRSA (SEQ ID NO: 62)	RSDNLAR (SEQ ID NO: 7)	QSGHLRS (SEQ ID NO: 41)	NA
25922	RSAALSR (SEQ ID NO: 58)	RSDALAR (SEQ ID NO: 59)	RSDNLSE (SEQ ID NO: 60)	KRCNLRC (SEQ ID NO: 61)	QSSDLRS (SEQ ID NO: 31)	DRSHLAR (SEQ ID NO: 13)

[0104] 表1B:人和小鼠Htt靶向的锌指蛋白

SBS #	设计					
	F1	F2	F3	F4	F5	F6
32468	RSDNLAR (SEQ ID NO: 7)	WRGDRVK (SEQ ID NO: 8)	DRSNLSR (SEQ ID NO: 9)	TSGSLTR (SEQ ID NO: 10)	ERGTLAR (SEQ ID NO: 11)	RSDDRKT (SEQ ID NO: 4)
32501	RSDALSR (SEQ ID NO: 12)	DRSHLAR (SEQ ID NO: 13)	RSDHLSR (SEQ ID NO: 14)	QSSDLTR (SEQ ID NO: 15)	TSGNLTR (SEQ ID NO: 16)	DRSHLAR (SEQ ID NO: 13)
31809	RSDDLRS (SEQ ID NO: 1)	RNDNRK (SEQ ID NO: 2)	RSDDLTR (SEQ ID NO: 3)	RSDDRKT (SEQ ID NO: 4)	RSDDLTR (SEQ ID NO: 3)	QSSDLRS (SEQ ID NO: 6)
32528	QSGHLQR (SEQ ID NO: 17)	TSGNLTR (SEQ ID NO: 16)	QSGDLTR (SEQ ID NO: 18)	DRSHLAR (SEQ ID NO: 13)	RSDDLST (SEQ ID NO: 19)	VRSRLRR (SEQ ID NO: 20)
30580	RSDNLAR (SEQ ID NO: 7)	WRGDRVK (SEQ ID NO: 8)	DRSDLSR (SEQ ID NO: 22)	RSDALAR (SEQ ID NO: 59)	ERGTLAR (SEQ ID NO: 11)	RSDDRKT (SEQ ID NO: 4)
30929	DRSTLRQ (SEQ ID NO: 21)	DRSDLSR (SEQ ID NO: 22)	QSSTRAR (SEQ ID NO: 23)	RSDDLSE (SEQ ID NO: 24)	HRSRWG (SEQ ID NO: 25)	NA
32538	DRSDLSR (SEQ ID NO: 22)	RRDTLRS (SEQ ID NO: 26)	RSDHLST (SEQ ID NO: 27)	QSAHRIT (SEQ ID NO: 28)	QSGDLTR (SEQ ID NO: 18)	DRSHLAR (SEQ ID NO: 13)
32567	RSDHLSE (SEQ ID NO: 29)	QNAHRKT (SEQ ID NO: 30)	QSSDLRS (SEQ ID NO: 31)	HRSTRNR (SEQ ID NO: 32)	QSSDLRS (SEQ ID NO: 31)	HRSTRNR (SEQ ID NO: 32)
29627	DRSNLSR (SEQ ID NO: 9)	LRQDLKR (SEQ ID NO: 33)	DRSHLTR (SEQ ID NO: 34)	DRSNLTR (SEQ ID NO: 35)	RSDDLST (SEQ ID NO: 27)	QSAHRIT (SEQ ID NO: 28)
29628	TSGNLTR (SEQ ID NO: 16)	LKQMLAV (SEQ ID NO: 36)	RSDLSA (SEQ ID NO: 37)	DRSDLSR (SEQ ID NO: 22)	RSDALST (SEQ ID NO: 38)	DRSTRTK (SEQ ID NO: 39)
29631	QSSDLRS (SEQ ID NO: 31)	DRSALAR (SEQ ID NO: 40)	QSSDLRS (SEQ ID NO: 31)	QSGHLSR (SEQ ID NO: 41)	RSDDLSE (SEQ ID NO: 42)	TSGHLSR (SEQ ID NO: 43)
29632	RSDDLSE (SEQ ID NO: 24)	KLCNRKC (SEQ ID NO: 44)	TSGNLTR (SEQ ID NO: 16)	HRTSLTD (SEQ ID NO: 45)	RSALSR (SEQ ID NO: 46)	QSGNLAR (SEQ ID NO: 47)
29637	DRSNLSR (SEQ ID NO: 9)	QSGNLAR (SEQ ID NO: 47)	DRSNLSR (SEQ ID NO: 9)	LKHHLTD (SEQ ID NO: 48)	QSGDLTR (SEQ ID NO: 18)	YRWLRNN (SEQ ID NO: 49)
29638	RSDHLSQ	RSVRKN	QSSDLRS	QSGDLTR	WSTSLRA	NA

[0105]

[0106]

	(SEQ ID NO: 50)	(SEQ ID NO: 51)	(SEQ ID NO: 31)	(SEQ ID NO: 18)	(SEQ ID NO: 52)	
25917	DRSNLSR (SEQ ID NO: 9)	QRTHLTQ (SEQ ID NO: 53)	RSSHLSR (SEQ ID NO: 54)	TSGSLSR (SEQ ID NO: 55)	TRQNRDT (SEQ ID NO: 56)	NA
25916	DQSTLRN (SEQ ID NO: 57)	RSAALSR (SEQ ID NO: 58)	RSDALAR (SEQ ID NO: 59)	RSDNLSE (SEQ ID NO: 60)	KRCNLRC (SEQ ID NO: 61)	NA
33074	RSDNLSE (SEQ ID NO: 60)	KRCNLRC (SEQ ID NO: 61)	QSGDLTR (SEQ ID NO: 18)	QSGDLTR (SEQ ID NO: 18)	RSDNLSE (SEQ ID NO: 60)	KRCNLRC (SEQ ID NO: 61)
33080	QSGDLTR (SEQ ID NO: 18)	QSGDLTR (SEQ ID NO: 18)	RSDNLSE (SEQ ID NO: 60)	KRCNLRC (SEQ ID NO: 61)	QSGDLTR (SEQ ID NO: 18)	QSGDLTR (SEQ ID NO: 18)
33084	QSSDLR (SEQ ID NO: 31)	HRSTRNR (SEQ ID NO: 32)	RSDLTSE (SEQ ID NO: 24)	RRWTLVG (SEQ ID NO: 64)	NA	NA
33088	QSSDLR (SEQ ID NO: 31)	HRSTRNR (SEQ ID NO: 32)	RAVLSE (SEQ ID NO: 148)	QSSDLR (SEQ ID NO: 31)	HRSTRNR (SEQ ID NO: 32)	NA
30643	RSDNLSE (SEQ ID NO: 60)	KRCNLRC (SEQ ID NO: 61)	QSSDLR (SEQ ID NO: 31)	QWSTRKR (SEQ ID NO: 63)	QSSDLR (SEQ ID NO: 31)	QWSTRKR (SEQ ID NO: 63)
30648	RSDNLSE (SEQ ID NO: 60)	KRCNLRC (SEQ ID NO: 61)	RSDNLSE (SEQ ID NO: 60)	KRCNLRC (SEQ ID NO: 61)	RSDNLSE (SEQ ID NO: 60)	KRCNLRC (SEQ ID NO: 61)
30645	RSDNLSE (SEQ ID NO: 60)	KRCNLRC (SEQ ID NO: 61)	QSSDLR (SEQ ID NO: 31)	QWSTRKR (SEQ ID NO: 63)	QSGDLTR (SEQ ID NO: 18)	NA
30640	QSSDLR (SEQ ID NO: 31)	QWSTRKR (SEQ ID NO: 63)	QSSDLR (SEQ ID NO: 31)	QWSTRKR (SEQ ID NO: 63)	QSGDLTR (SEQ ID NO: 18)	NA
30657	RSDLTSE (SEQ ID NO: 24)	RRWTLVG (SEQ ID NO: 64)	QSSDLR (SEQ ID NO: 31)	HRSTRNR (SEQ ID NO: 32)	QSSDLR (SEQ ID NO: 31)	HRSTRNR (SEQ ID NO: 32)
30642	QSGDLTR (SEQ ID NO: 18)	QSSDLR (SEQ ID NO: 31)	QWSTRKR (SEQ ID NO: 63)	QSSDLR (SEQ ID NO: 31)	QWSTRKR (SEQ ID NO: 63)	NA
30646	RSDNLSE (SEQ ID NO: 60)	KRCNLRC (SEQ ID NO: 61)	QSGDLTR (SEQ ID NO: 18)	QSSDLR (SEQ ID NO: 31)	QWSTRKR (SEQ ID NO: 63)	NA
32220	RSDVLSE (SEQ ID NO: 42)	QSSDLR (SEQ ID NO: 31)	HRSTRNR (SEQ ID NO: 32)	NA	NA	NA

[0107]

32210	QSGDLTR (SEQ ID NO: 18)	QSSDLSR (SEQ ID NO: 31)	QWSTRKR (SEQ ID NO: 63)	NA	NA	NA
32215	RSDNLRE (SEQ ID NO: 65)	RSDNLSE (SEQ ID NO: 60)	KRCNLRC (SEQ ID NO: 61)	NA	NA	NA
30658	QSSDLSR (SEQ ID NO:31)	HRSTRNR (SEQ ID NO:32)	QSSDLSR (SEQ ID NO:31)	HRSTRNR (SEQ ID NO:32)	QSSDLSR (SEQ ID NO:31)	NA
32218	QSSDLSR (SEQ ID NO: 31)	QSSDLSR (SEQ ID NO: 31)	NA	NA	NA	NA
32427	ERGTLAR (SEQ ID NO:11)	TSGSLTR (SEQ ID NO:10)	RSDNLAR (SEQ ID NO:7)	DPSNRVG (SEQ ID NO:78)	RSDDLK (SEQ ID NO:149)	DNSNRIK (SEQ ID NO:150)
32653	RSDHLSE (SEQ ID NO: 29)	QSGHLR (SEQ ID NO: 41)	RSDDLTR (SEQ ID NO: 3)	YRWLLRS (SEQ ID NO: 66)	QSSDLSR (SEQ ID NO: 31)	RKDALVA (SEQ ID NO: 67)
32677	QSGDLTR (SEQ ID NO: 18)	RRADLSR (SEQ ID NO: 68)	DRSHLTR (SEQ ID NO: 34)	DRSHLAR (SEQ ID NO: 13)	DRSNLSR (SEQ ID NO: 9)	LAQPRNK (SEQ ID NO: 69)
33560	ERGTLAR (SEQ ID NO:11)	QSGSLTR (SEQ ID NO:84)	RSDNLAR (SEQ ID NO:7)	DDSHRKD (SEQ ID NO:151)	RSDDLK (SEQ ID NO:149)	DNSNRIK (SEQ ID NO:150)
33583	DRSNLSR (SEQ ID NO:9)	HKQHRDA (SEQ ID NO:76)	DRSDLR (SEQ ID NO:22)	RRTDLRR (SEQ ID NO:77)	RSANLAR (SEQ ID NO:73)	DRSHLAR (SEQ ID NO:13)
32685	RSDHLA (SEQ ID NO: 70)	RSADRTR (SEQ ID NO: 71)	RSDVLSE (SEQ ID NO: 42)	TSGHLR (SEQ ID NO: 43)	RSDDLTR (SEQ ID NO: 3)	TSSDRKK (SEQ ID NO: 72)
32422	RSANLAR (SEQ ID NO: 73)	RSDDLTR (SEQ ID NO: 3)	RSDTLSE (SEQ ID NO: 24)	HHSARRC (SEQ ID NO: 74)	ERGTLAR (SEQ ID NO: 11)	DRSNLTR (SEQ ID NO: 35)
32428	RSDVLST (SEQ ID NO: 19)	DNSSRTR (SEQ ID NO: 75)	DRSNLSR (SEQ ID NO: 9)	HKQHRDA (SEQ ID NO: 76)	DRSDLR (SEQ ID NO: 22)	RRTDLRR (SEQ ID NO: 77)
32430	RSDVLST (SEQ ID NO: 19)	VRSRLRR (SEQ ID NO: 20)	ERGTLAR (SEQ ID NO: 11)	TSGSLTR (SEQ ID NO: 10)	RSDNLAR (SEQ ID NO: 7)	DPSNRVG (SEQ ID NO: 78)
32432	RSDVLST (SEQ ID NO: 19)	VRSRLRR (SEQ ID NO: 20)	ERGTLAR (SEQ ID NO: 11)	TSGSLTR (SEQ ID NO: 10)	RSDHLA (SEQ ID NO: 70)	RSADLSR (SEQ ID NO: 79)
32714	RSDVLST (SEQ ID NO: 19)	DNSSRTR (SEQ ID NO: 75)	ERGTLAR (SEQ ID NO: 11)	QSGNLAR (SEQ ID NO: 47)	DRSHLTR (SEQ ID NO: 34)	RNDDRKK (SEQ ID NO: 80)
32733	DRSNLSR (SEQ ID NO: 9)	QKVTLAA (SEQ ID NO: 81)	RSAHLR (SEQ ID NO: 46)	TSGNLTR (SEQ ID NO: 16)	DRSDLR (SEQ ID NO: 22)	RRSTLRS (SEQ ID NO: 82)

[0108]

30901	DRSALSR (SEQ ID NO: 83)	QSGSLTR (SEQ ID NO: 84)	QSSDLR (SEQ ID NO: 31)	LKWNLRT (SEQ ID NO: 85)	RSDNLAR (SEQ ID NO: 7)	LKWDRQT (SEQ ID NO: 86)
31952	QSGALAR (SEQ ID NO:147)	RSDDLTR (SEQ ID NO:3)	DRSALSR (SEQ ID NO:83)	RSDHLTQ (SEQ ID NO:152)	QSGDLTR (SEQ ID NO:18)	WSTSLRA (SEQ ID NO:52)
31921	RSDSLR (SEQ ID NO:153)	RSDDLTR (SEQ ID NO:3)	QSGDLTR (SEQ ID NO:18)	RRDWLPQ (SEQ ID NO:154)	DRSNLSR (SEQ ID NO:9)	RSDDRKT (SEQ ID NO:4)
30906	DRSHLSR (SEQ ID NO: 87)	TSGNLTR (SEQ ID NO: 16)	QSGDLTR (SEQ ID NO: 18)	DRSHLAR (SEQ ID NO: 13)	RSDVLST (SEQ ID NO: 19)	VRSLRR (SEQ ID NO: 20)

[0109] 表2A和2B中公开了这些蛋白质靶位点的序列和位置。表2A和2B示出了指示的锌指蛋白的靶序列。用大写字母指示靶位点中与ZFP识别螺旋接触的核苷酸；用小写字母指示非接触核苷酸。

[0110] 表2A:人和小鼠Htt上的靶位点

[0111]

SBS#	靶位点
18856	AcGCTGCGCCGGCGGAGGCGgggcccgcg (SEQ ID NO:88)
25920	gcGCTCAGCAGGTGGTGaccttgtggac_ (SEQ ID NO:103)
25921	atGGTGGGAGAGACTGTgaggcggcagc_ (SEQ ID NO:104)
25923	tgGGAGAGacTGTGAGGCGgcagctggg (SEQ ID NO:105)
25922	atGGCGCTCAGCAGGTGGTGaccttgtg_ (SEQ ID NO:106)

[0112] 表2B:人和小鼠Htt上的靶位点

[0113]

SBS #	靶位点
32468	agCCGGCCGTGGACTCTGAGccgaggtg_ (SEQ ID NO:89)
32427	cgCACTCGcCGCGAGgGTTGCCgggacg_ (SEQ ID NO:155)
32501	gtGGCGATGCGGGGGCGTGgtgaggta_ (SEQ ID NO:90)
31809	acGCTGCGCCGGCGGAGGCGgggcccgcg_ (SEQ ID NO:88)
32528	ccGGGACGGGTCCAaGATGGAcggcccgc_ (SEQ ID NO:91)
30580	agCCGGCCGTGGACTCTGAGccgaggtg_ (SEQ ID NO:89)
30929	ccGTCCCGGCAGCCCCacgggcgcttg_ (SEQ ID NO:92)
30658	ctGCTGCTGCTGCTGCTgctggaaggac_ (SEQ ID NO:108)
32538	cgGGTCCAAGATGGACGGCCgctcaggt_ (SEQ ID NO:93)
32567	ctGCTGCTGCTGCTGGAAGGacttgagg_ (SEQ ID NO:94)
29627	tcAGATGGGACGGCGCTGACctggctgg_ (SEQ ID NO:95)
29628	ctGCCATGGACCTGAATGATgggaccca_ (SEQ ID NO:96)
29631	gtGGTCTGGGAGCTGTCGCTgatgggcg_ (SEQ ID NO:97)
29632	ccGAAGGCCTGATtCAGCTGttacccc_ (SEQ ID NO:98)
29637	aaCTTGCAAGTAACaGAAGACTcatcct_ (SEQ ID NO:99)

[0114]

29638	ctTGTACAGCTGTGAGGgtgagcataat_(SEQ ID NO:100)
25917	gcCATGGTGGGAGAGACTgtgagcggc_(SEQ ID NO:101)
25916	ctCAGCAGGTGGTGCCTtgtggacatt_(SEQ ID NO:102)
33074	agCAGCAGcaGCAGCAgCAGCAGcagca_(SEQ ID NO:157)
33080	caGCAGCAgCAGCAGcaGCAGCAgCAGcagc_(SEQ ID NO:107)
33084	tgCTGCTGctGCTGCTgctgctggaagg_(SEQ ID NO:109)
33088	ctGCTGCTgCTGctGCTGCTgctggaag_(SEQ ID NO:158)
30643	caGCAGCAGCAGCAgCAGCAGcagcagc_(SEQ ID NO:107)
30648	agCAGCAGCAGCAGCAGCAGcagcagca_(SEQ ID NO:157)
30645	caGCAGCAGCAgCAGCAGcagcagcagc_(SEQ ID NO:107)
30640	caGCAGCAGCAGCAGCAgCAGcagcagcagc_(SEQ ID NO:107)
30657	ctGCTGCTGCTGCTgCTGCTGgaaggac_(SEQ ID NO:108)
30642	caGCAGCAGCAGCAGCAgCAGcagcagcagc_(SEQ ID NO:107)
30646	caGCAGCAGCAgCAGCAGcagcagcagc_(SEQ ID NO:107)
32220	ctGCTGCTgCTGctgctgctgctggaagg_(SEQ ID NO:109)
32210	caGCAGCAGCAgCAGcagcagcagcagcagc_(SEQ ID NO:107)
32215	agCAGCAGCAGcagcagcagcagcagca_(SEQ ID NO:110)
32218	tGCTGCTgctgctgctgctgctggaagg_(SEQ ID NO:111)
32653	ggCTGGCTTTTGGCGGAAGGggcggggc_(SEQ ID NO:112)
32677	gaATTGACaGGCGGAtGCGTCGtctct_(SEQ ID NO:113)
33560	cgCACTCGcCGCGAGgGTTGCCgggacg_(SEQ ID NO:155)
33583	gcGGCGAGtGCGTCCCGTGACgtcatgc_(SEQ ID NO:158)
32685	atTCTGCGGGTCTGGCGTGGcctcgtct_(SEQ ID NO:114)
32422	gtGACGTCATGCCGGCGGAGacgaggcc_(SEQ ID NO:115)
32428	gtGCGTCCCGTGACGTCATGccggcgga_(SEQ ID NO:116)
32430	gcCGCGAGgGTTGCCGGGACGggcccaa_(SEQ ID NO:117)
32432	ccGCGAGGGTTGCCGGGACGggcccaaag_(SEQ ID NO:118)
32714	caTCGGGCagGAAGCCGTCATGgcaacc_(SEQ ID NO:119)
32733	tcCTGCCCGATGGGACAGACcctgaaga_(SEQ ID NO:120)
30901	gtACTGAGcAATGCTGTAGTCagcaatc_(SEQ ID NO:121)
31952	ccTGTCCAgAGGGTCCGGTAcctcct_(SEQ ID NO:159)
31921	tgCCGACCTGGCAGCGCGGtgggtggc_(SEQ ID NO:160)
30906	ccGGGACGGGTCCAaGATGGAcggcgc_(SEQ ID NO:91)

[0115] 在某些实施方案中，DNA结合结构域包含天然存在或经工程化(非天然存在)的TAL效应子(TALE)DNA结合结构域。见，例如，美国专利公布第20110301073号，通过引用整体并入本文。已知黄单孢菌属(*Xanthomonas*)的植物病原菌在重要作物中引起许多疾病。黄单孢菌属的病原性取决于向植物细胞注入25种以上不同效应子蛋白的保守型III分泌(T3S)系统。其中注入的蛋白质为模拟植物转录激活因子并且操纵植物转录组的转录激活因子样效应子(TALE)(见Kay等(2007) *Science* 318:648-651)。这些蛋白质含有DNA结合结构域和转录激活结构域。表征最清楚的TALE之一为来自野油菜黄单孢菌疱病致病变种(*Xanthomonas campestris* pv. *Vesicatoria*)的AvrBs3(见Bonas等(1989) *Mol Gen Genet* 218:127-136和W02010079430)。TALE含有串联重复的集中式结构域，每个重复含有对这些蛋白质的DNA结合特异性关键的大约34个氨基酸。另外，其含有核定位序列和酸性转录激活结构域(对于综述，见Schornack S等(2006) *J Plant Physiol* 163(3):256-272)。另外，在植物病原细菌青枯病菌(*Ralstonia solanacearum*)中，已经发现在青枯病菌(*R. solanacearum*)生物变种1菌株GMI1000和生物变种4菌株RS1000中称为brg11和hpx17的两个基因与黄单孢菌属的

AvrBs3家族同源(见Heuer等(2007) *Appl and Envir Micro* 73(13):4379-4384)。这些基因在核苷酸序列上98.9%彼此相同,但是不同之处在于hpx17的重复结构域中缺失1,575bp。然而,两种基因产物与黄单孢菌属的AvrBs3家族蛋白具有低于40%的序列同一性。

[0116] 这些TALE的特异性取决于串联重复中存在的序列。重复序列包含大约102bp并且重复通常相互91-100%同源(Bonas等,同上)。重复的多态性通常位于位置12和13并且在位置12和13的高变双残基的同一性与TALE靶序列中连续核苷酸的同一性之间似乎存在一一对应(见Moscou和Bogdanove,(2009) *Science* 326:1501和Boch等(2009) *Science* 326:1509-1512)。实验上,已经确定了供这些TALE的DNA识别的密码,以致位置12和13的HD序列导致与胞嘧啶(C)结合,NG与T结合,NI与A、C、G或T结合,NN与A或G结合,并且IG与T结合。这些DNA结合的重复序列已经组装成具有新的组合和重复数量的蛋白质,以产生能够与新序列相互作用并且激活植物细胞中非内源报告基因的表达的人工转录因子(Boch等,同上)。工程化TAL蛋白质已经与FokI裂解半结构域连接以产生在酵母报告基因测定(基于质粒的靶标)中表现出活性的TAL效应子结构域核酸酶融合(TALEN)。Christian等((2010) *Genetics* epub 10.1534/genetics.110.120717)。同样,见,美国专利公布第20110301073号,其通过引用整体并入。

[0117] 表3中列出了设计的与ZFP或TALE蛋白质一起使用的二聚化结构域的具体实例。列出了两类结构域,卷曲螺旋(CC)和二聚化锌指(DZ)的氨基酸序列。

[0118] 表3:设计的二聚化结构域

[0119]

设计名称	氨基酸序列
DZ1	TKCVHCGIVFLDEVMYALHMSCHGFRDPFECNICGYHSQDRYEFSSHIVRG EH (SEQ ID NO:122)
	TKCVHCGIVFLDEVMYALHMSCHGFRDPFECNICGYHSQDRYEFSSHIVRG EH(SEQ ID NO:122)
DZ2	FKCEHCRILFLDHVMFTIHMGCCHGFRDPFKCNMCGEKCDGPVGLFVHMARN AH(SEQ ID NO:123)
	TKCVHCGIVFLDEVMYALHMSCHGFRDPFECNICGYHSQDRYEFSSHIVRG EH(SEQ ID NO:122)
DZ3	FKCEHCRILFLDHVMFTIHMGCCHGFRDPFKCNMCGEKCDGPVGLFVHMARN AH(SEQ ID NO:123)
	HHCQHCDMYFADNILYTIHMGCCHGYENPFECNICGYHSQDRYEFSSHIVRG EH(SEQ ID NO:124)
DZ4	HHCQHCDMYFADNILYTIHMGCCHSCDDVFKCNMCGEKCDGPVGLFVHMARN AHGEKPTKCVHCGIVFLDEVMYALHMSCHGFRDPFECNICGYHSQDRYEFSS SHIVRGEH(SEQ ID NO:125)
	FKCEHCRILFLDHVMFTIHMGCCHGFRDPFKCNMCGEKCDGPVGLFVHMARN AHGEKPFYCEHCEITFRDVVMYSLHKGYHGFRDPFECNICGYHSQDRYEFSS SHIVRGEH(SEQ ID NO:126)
CC1	AQLEKELQALEKKLAQLEWENQALEKELAQ (SEQ ID NO:127)
	AQLKKKLQANKKELAQLKWKLQALKKKLAQ (SEQ ID NO:128)
CC2	EQLEKKLQALEKKLAQLEWKNQALEKKLAQ (SEQ ID NO:129)
	ALKKELQANKKELAQLKWELQALKKELAQ (SEQ ID NO:130)
CC3	EQLEKKLQALEKKLAQLEWKNQALEK (SEQ ID NO:131)
	ELQANKKELAQLKWELQALKKELAQ (SEQ ID NO:132)
CC4	EQLEKKLQALEKKLAQLEWKNQA (SEQ ID NO:133)
	QANKKELAQLKWELQALKKELAQ (SEQ ID NO:134)
CC5	EQLEKKLQALEKKLAQLEWKNQALEKKLAQ (SEQ ID NO:129)

[0120]		<u>ALKKELQANKKELAQLKWELQALKKELAQ (SEQ ID NO:130)</u>
	<b>CC6</b>	<u>EQLEKKLQALEKKLAQLEWKNQALEKKLAQ (SEQ ID NO:129)</u>
		<u>ALKKELQANKKELAQLKWELQALKKELAQ (SEQ ID NO:130)</u>
	<b>CC7</b>	<u>EQLEKKLQALEKKLAQLEWKNQALEKKLAQ (SEQ ID NO:129)</u>
		<u>ALKKELQANKKELAQLKWELQALKKELAQ (SEQ ID NO:130)</u>

[0121] 融合蛋白

[0122] 还提供了包含如本文所述的DNA结合蛋白(例如,ZFP或TALE)和异源调控(功能)结构域(或其功能片段)的融合蛋白。常见结构域包括(例如)转录因子结构域(激活因子、阻遏因子、辅助激活因子、辅助阻遏因子)、沉默子、致癌基因(例如,myc、jun、fos、myb、max、mad、rel、ets、bcl、myb、mos家族成员等);DNA修复酶及其相关因子和修饰因子;DNA重排酶及其相关因子和修饰因子;染色质相关蛋白及其修饰因子(例如,激酶、乙酰基转移酶和脱乙酰基酶);和DNA修饰酶(例如,甲基转移酶、拓扑异构酶、解螺旋酶、连接酶、激酶、磷酸酶、聚合酶、内切核酸酶)及其相关因子和修饰因子。关于DNA结合结构域和核酸酶裂解结构域融合的详情,见美国专利申请公布第20050064474、20060188987和2007/0218528号,其通过引用整体并入本文。

[0123] 用于实现激活的适合结构域包括HSV VP16激活结构域(见,例如Hagmann等, *J. Virol.* 71, 5952-5962 (1997))核激素受体(见,例如Torchia等, *Curr. Opin. Cell. Biol.* 10:373-383 (1998));核因子κB的p65亚基(Bitko和Barik, *J. Virol.* 72:5610-5618 (1998)及Doyle和Hunt, *Neuroreport* 8:2937-2942 (1997));Liu等, *Cancer Gene Ther.* 5:3-28 (1998)或人工嵌合功能结构域例如VP64(Beerli等, (1998) *Proc. Natl. Acad. Sci. USA* 95:14623-33)和降解决定子(Molinari等, (1999) *EMBO J.* 18, 6439-6447)。另外的示例性激活结构域包括Oct 1、Oct-2A、Sp1、AP-2和CTF1(Seipel等, *EMBO J.* 11, 4961-4968 (1992)以及p300、CBP、PCAF、SRC1PvALF、AtHD2A和ERF-2。见,例如, Robyr等(2000) *Mol. Endocrinol.* 14:329-347;Collingwood等(1999) *J. Mol. Endocrinol.* 23:255-275;Leo等(2000) *Gene* 245:1-11;Manteuffel-Cymborowska (1999) *Acta Biochim. Pol.* 46:77-89;McKenna等(1999) *J. Steroid Biochem. Mol. Biol.* 69:3-12;Malik等(2000) *Trends Biochem. Sci.* 25:277-283;和Lemon等(1999) *Curr. Opin. Genet. Dev.* 9:499-504。另外的示例性激活结构域包括但不限于OsGAI、HALF-1、C1、AP1、ARF-5、-6、-7、和-8、CPRF1、CPRF4、MYC-RP/GP和TRAB1。见,例如,Ogawa等(2000) *Gene* 245:21-29;Okanami等(1996) *Genes Cells* 1:87-99;Goff等(1991) *Genes Dev.* 5:298-309;Cho等(1999) *Plant Mol. Biol.* 40:419-429;Ulmason等(1999) *Proc. Natl. Acad. Sci. USA* 96:5844-5849;Sprenger-Haussels等(2000) *Plant J.* 22:1-8;Gong等(1999) *Plant Mol. Biol.* 41:33-44;和Hobo等(1999) *Proc. Natl. Acad. Sci. USA* 96:15,348-15,353。

[0124] 对于本领域的技术人员而言,明确的是在DNA结合结构域和功能结构域之间的融合蛋白(或编码融合蛋白的核酸)的形成中,激活结构域或与激活结构域相互作用的分子适合作为功能结构域。基本上能够为靶基因募集激活复合物和/或激活活性(例如,组蛋白乙酰化)的任何分子均可用作融合蛋白的激活结构域。例如在共同拥有的美国专利申请2002/

0115215和2003/0082552及共同拥有的WO 02/44376中描述了适合用作融合分子中的功能结构域的绝缘子结构域、定位结构域和染色质重构蛋白例如含ISWI的结构域和/或甲基结合结构域蛋白。

[0125] 示例性阻遏结构域包括但不限于KRAB A/B、KOX、TGF- $\beta$ -诱导型早期基因(TIEG)、v-erbA、SID、MBD2、MBD3、DNMT家族成员(例如, DNMT1、DNMT3A、DNMT3B)、Rb和MeCP2。见, 例如, Bird等(1999) Cell 99:451-454; Tyler等(1999) Cell 99:443-446; Knoepfler等(1999) Cell 99:447-450; 和Robertson等(2000) Nature Genet. 25:338-342。另外的示例性阻遏结构域包括但不限于ROM2和AtHD2A。见, 例如, Chem等(1996) Plant Cell 8:305-321; 和Wu等(2000) Plant J. 22:19-27。

[0126] 通过本领域技术人员众所周知的克隆和生物化学偶联方法构建融合分子。融合分子包含DNA结合结构域和功能结构域(例如, 转录激活或阻遏结构域)。融合分子还任选包含核定位信号(例如, 来自SV40基质T抗原)和表位标签(例如, FLAG和血细胞凝集素)。设计融合蛋白(和编码融合蛋白的核酸), 以便将翻译阅读框保存在融合组成部分中。

[0127] 通过本领域技术人员众所周知的生物化学偶联方法构建一只手上功能结构域(或其功能片段)的多肽组成部分与另一只手上非蛋白DNA结合结构域(例如, 抗生素、嵌入剂、小沟结合剂、核酸)之间的融合。见, 例如, Pierce Chemical Company (Rockford, IL) Catalogue。已经描述了用于在小沟结合剂与多肽之间产生融合的方法和组合物。Mapp等(2000) Proc. Natl. Acad. Sci. USA 97:3930-3935。

[0128] 在某些实施方案中, 由DNA结合结构域结合的靶位点存在于细胞染色质的可达区域。例如, 可如共同拥有的国际公布WO 01/83732中所述测定可达区域。如果靶位点不存在于细胞染色质的可达区域, 可如共同拥有的WO 01/83793中所述生成一个或多个可达区域。在另外的实施方案中, 融合分子的DNA结合结构域能够与细胞染色质结合, 不管其靶位点是否在可达区域内。例如, 此类DNA结合结构域能够与接头DNA和/或核小体DNA结合。在某些类固醇受体和肝细胞核因子3(HNF3)中发现这类“先锋”DNA结合结构域的实例。Cordingley等(1987) Cell 48:261-270; Pina等(1990) Cell 60:719-731; 和Cirillo等(1998) EMBO J. 17:244-254。

[0129] 如本领域的技术人员所知, 融合分子可与药学上可接受的载体配制。见, 例如, Remington's Pharmaceutical Sciences, 第17版, 1985; 和共同拥有的WO 00/42219。

[0130] 融合分子的功能组成部分/结构域可选自一旦融合分子经由其DNA结合结构域与靶序列结合, 就能够影响基因转录的多种不同组成部分。因此, 功能组成部分可包括但不限于各种转录因子结构域, 例如激活因子、阻遏因子、辅助激活因子、辅助阻遏因子和沉默子。

[0131] 例如, 在共同拥有的美国专利第6,534,261号和美国专利申请公布第2002/0160940号中公开了另外的示例性功能结构域。

[0132] 也可选择受外源小分子或配体调控的功能结构域。例如, 可采用RheoSwitch®技术, 其中功能结构域只有在外部RheoChem™配体的存在下呈现其活性构象(见例如US 20090136465)。因此, ZFP或TALE可能与可调控的功能结构域可操作地连接, 其中所产生的ZFP-TF或TALE-TF活性受外部配体控制。

[0133] 核酸酶

[0134] 在某些实施方案中, 融合蛋白包含DNA结合结构域和裂解(核酸酶)结构域。同样,

可使用核酸酶,例如工程化核酸酶实现基因修饰。工程化核酸酶技术基于天然存在的DNA结合蛋白的工程化。例如,已经描述了具有定制DNA结合特异性的归巢内勤核酸酶的工程化。(见,Chames等(2005) *Nucleic Acids Res* 33(20):e178;Arnould等(2006) *J.Mol.Biol.*355:443-458)。另外,还已经描述了ZFP的工程化。见,例如,美国专利第6,534,261、6,607,882、6,824,978、6,979,539、6,933,113、7,163,824和7,013,219号。

[0135] 另外,ZFP和TALE已经与核酸酶结构域融合以产生能够通过其经工程化的(ZFP或TALE)DNA结合结构域识别其预期核酸靶标并经由核酸酶活性引起在ZFP或TALE DNA结合位点附近切割DNA的ZFN和TALEN-功能实体。见,例如,Kim等(1996) *Proc Natl Acad Sci USA* 93(3):1156-1160。最近,ZFN已经用于多种生物中的基因组修饰。见,例如,美国专利公布20030232410、20050208489、20050026157、20050064474、20060188987、20060063231和国际公布WO 07/014275。

[0136] 因此,本文所述的方法和组合物广泛适用并且可能牵涉任何目标核酸酶。核酸酶的非限制性实例包括大范围核酸酶、TALEN和锌指核酸酶。核酸酶可能包含异源DNA结合和裂解结构域(例如,锌指核酸酶;TALEN;大范围核酸酶DNA结合结构域与异源裂解结构域),或可选地,天然存在的核酸酶的DNA结合结构域可改变为与所选靶位点结合(例如,已经工程化为与不同于同源结合位点的位点结合的大范围核酸酶)。

[0137] 在某些实施方案中,核酸酶为大范围核酸酶(归巢内切核酸酶)。天然存在的大范围核酸酶识别15-40个碱基对裂解位点并且常分为4个家族:LAGLIDADG家族、GIY-YIG家族、His-Cyst盒家族和HNH家族。示例性归巢内切核酸酶包括I-SceI、I-CeuI、PI-PspI、PI-Sce、I-SceIV、I-CsmI、I-PanI、I-SceII、I-PpoI、I-SceIII、I-CreI、I-TevI、I-TevII和I-TevIII。其识别序列已知。同样见美国专利第5,420,032号;美国专利第6,833,252号;Belfort等(1997) *Nucleic Acids Res.*25:3379-3388;Dujon等(1989) *Gene* 82:115-118;Perler等(1994) *Nucleic Acids Res.*22,1125-1127;Jasin(1996) *Trends Genet.*12:224-228;Gimble等(1996) *J.Mol.Biol.*263:163-180;Argast等(1998) *J.Mol.Biol.*280:345-353和New England Biolabs目录。

[0138] 来自主要来自LAGLIDADG家族的天然存在的大范围核酸酶的DNA结合结构域已经用于促进植物、酵母、果蝇(*Drosophila*)、哺乳动物细胞和小鼠中的位点特异性基因组修饰,但是这种方法已限于保存大范围核酸酶识别序列的同源基因的修饰(Monet等(1999), *Biochem.Biophysics.Res.Common.*255:88-93)或已经向其中引入了识别序列的预工程化基因组(Route等(1994), *Mol.Cell.Biol.*14:8096-106;Chilton等(2003) *Plant Physiology* 133:956-65;Puchta等(1996), *Proc.Natl.Acad.Sci.USA* 93:5055-60;Rong等(2002), *Genes Dev.*16:1568-81;Gouble等(2006), *J.Gene Med.*8(5):616-622)。相应地,已经尝试在医学或生物技术相关位点工程化大范围核酸酶以表现出新型结合特异性(Porteus等(2005), *Nat.Biotechnol.*23:967-73;Sussman等(2004), *J.Mol.Biol.*342:31-41;Epinat等(2003), *Nucleic Acids Res.*31:2952-62;Chevalier等(2002) *Molec.Cell* 10:895-905;Epinat等(2003) *Nucleic Acids Res.*31:2952-2962;Ashworth等(2006) *Nature* 441:656-659;Paques等(2007) *Current Gene Therapy* 7:49-66;美国专利公布第20070117128、20060206949、20060153826、20060078552和20040002092号)。另外,来自大范围核酸酶的非天然存在或经工程化的DNA结合结构域也已经与来自异源核酸酶(例如,

FokI)的裂解结构域可操作地连接。

[0139] 在其它实施方案中,核酸酶为锌指核酸酶(ZFN)。ZFN包含已经工程化为与所选基因中的靶位点和裂解结构域或裂解半结构域结合的锌指蛋白。

[0140] 如以上所详细描述,锌指结合结构域可经工程化为与所选序列结合。见,例如,Beerli等(2002) *Nature Biotechnol.* 20:135-141;Pabo等(2001) *Ann.Rev.Biochem.* 70:313-340;Isalan等(2001) *Nature Biotechnol.* 19:656-660;Segal等(2001) *Curr.Opin.Biotechnol.* 12:632-637;Choo等(2000) *Curr.Opin.Struct.Biol.* 10:411-416。与天然存在的锌指蛋白相比,工程化锌指结合结构域可具有新型结合特异性。工程化方法包括但不限于合理的设计和类型的选择。合理的设计包括(例如)使用包含三联体(或四联体)核苷酸序列和单独锌指氨基酸序列的数据库,其中每个三联体或四联体核苷酸序列与锌指的结合特定三联体或四联体序列的一个或多个氨基酸序列缔合。见,例如,共同拥有的美国专利6,453,242和6,534,261,其通过引用整体并入本文。

[0141] 在美国专利5,789,538、5,925,523、6,007,988、6,013,453、6,410,248、6,140,466、6,200,759和6,242,568以及WO 98/37186、WO 98/53057、WO 00/27878、WO 01/88197和GB 2,338,237中公开了示例性选择方法,包括噬菌体展示和双杂交体系。另外,例如在共同拥有的WO 02/077227中描述了锌指结合结构域结合特异性的增强。

[0142] 另外,如这些和其它参考文献中所公开那样,锌指结构域和/或多指锌指蛋白可使用任何适合的接头序列,包括例如长度为5个或更多个氨基酸的接头(例如,TGEKP (SEQ ID NO:135)、TGGQRP (SEQ ID NO:136)、TGQKP (SEQ ID NO:137)和/或TGSQKP (SEQ ID NO:138))连接在一起。对于长度为6个或更多个氨基酸的示例性接头,见例如美国专利第6,479,626、6,903,185和7,153,949号。本文所述蛋白质可包括在蛋白质的单独锌指之间适合接头的任何组合。同样见,美国临时专利公布第20110287512号。

[0143] CRISPR(规律成簇的间隔短回文重复)/Cas(CRISPR相关)核酸酶体系是基于可用于基因组工程化的细菌体系的最近工程化的核酸酶体系。其基于许多细菌和古细菌的适应性免疫反应的一部分。当病毒或质粒侵入细菌时,侵入菌的DNA片段通过‘免疫’反应转化为CRISPR RNA(crRNA)。然后这种crRNA通过部分互补区域与另一类称为tracrRNA的RNA缔合以将Cas9核酸酶引导至称为“原间隔区(protospacer)”的靶DNA中与crRNA同源的区域。Cas9在由crRNA转录产物中所含的20-核苷酸引导序列指定的位点裂解DNA以在DSB处产生平末端。Cas9需要crRNA和tracrRNA进行位点特异性DNA识别和裂解。这种体系现已经工程化,以致crRNA和tracrRNA可结合成一个分子(“单个引导RNA”),单个引导RNA的crRNA等效部分可经工程化以引导Cas9核酸酶以靶向任何所需序列(见Jinek等(2012) *Science* 337,第816-821页,Jinek等(2013) *eLife* 2:e00471,和David Segal,(2013) *eLife* 2:e00563)。因此,CRISPR/Cas体系可经工程化以在基因组的所需靶标处产生DSB,并且可通过使用修复抑制剂影响DSB的修复以引起易错修复增加。

[0144] 核酸酶例如ZFN、TALEN和/或大范围核酸酶还包含核酸酶(裂解结构域、裂解半结构域)。如上所述,裂解结构域可能与DNA结合结构域异源,例如锌指DNA结合结构域和来自核酸酶的裂解结构域,或大范围核酸酶DNA结合结构域和来自不同核酸酶的裂解结构域。异源裂解结构域可从内切核酸酶或外切核酸酶获得。可从中得到裂解结构域的示例性内切核酸酶包括但不限于限制性内切核酸酶和归巢内切核酸酶。见,例如,2002-2003 Catalogue,

New England Biolabs, Beverly, MA; 和 Belfort 等 (1997) *Nucleic Acids Res.* 25:3379-3388。已知另外的裂解 DNA 的酶 (例如, S1 核酸酶; 绿豆核酸酶; 胰腺 DNA 酶 I; 微球菌核酸酶; 酵母 H0 内切核酸酶; 同样见 Linn 等 (编辑) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993)。这些酶中的一种或多种 (或其功能片段) 可用作裂解结构域和裂解半结构域的来源。

[0145] 类似地, 裂解半结构域可源自正如上面提出的对于裂解活性需要二聚化的任何核酸酶或其部分。一般而言, 如果融合蛋白包含裂解半结构域, 则裂解需要两个融合蛋白。可选地, 可使用包含两个裂解半结构域的统一蛋白质。所述两个裂解半结构域可源自相同内切核酸酶 (或其功能片段), 或每个裂解半结构域可源自不同内切核酸酶 (或其功能片段)。另外, 两个融合蛋白的靶位点优选相对于彼此布置, 以致两个融合蛋白与其各自靶位点的结合使裂解半结构域处于允许 (例如) 通过二聚化使裂解半结构域形成功能裂解结构域的相互空间定位。因此, 在某些实施方案中, 靶位点的近边间隔 5-8 个核苷酸或 15-18 个核苷酸。然而, 任何整数数量的核苷酸或核苷酸对可插入两个靶位点之间 (例如, 2 至 50 个核苷酸对或更多)。一般而言, 裂解位点位于靶位点之间。

[0146] 限制性内切核酸酶 (限制酶) 存在于许多物种中并且能够与 DNA 进行序列特异性结合 (在识别位点), 并在结合位点或附近裂解 DNA。某些限制酶 (例如 IIS 型) 在从识别位点移出的位点处裂解 DNA 并且具有可分离的结合和裂解结构域。例如, IIS 型酶 Fok I 在从一条链上的识别位点开始 9 个核苷酸处和从另一条链上的识别位点开始 13 个核苷酸处催化 DNA 的双链裂解。见, 例如, 美国专利 5,356,802、5,436,150 和 5,487,994; 以及 Li 等 (1992) *Proc. Natl. Acad. Sci. USA* 89:4275-4279; Li 等 (1993) *Proc. Natl. Acad. Sci. USA* 90:2764-2768; Kim 等 (1994a) *Proc. Natl. Acad. Sci. USA* 91:883-887; Kim 等 (1994b) *J. Biol. Chem.* 269:31,978-31,982。因此, 在一个实施方案中, 融合蛋白包含来自至少一种 IIS 型限制酶的裂解结构域 (或裂解半结构域) 和可能或可能未经工程化的一个或多个锌指结合结构域。

[0147] 裂解结构域可与结合结构域分离的示例性 IIS 型限制酶为 Fok I。这种特殊的酶呈二聚体时有活性。Bitinaite 等 (1998) *Proc. Natl. Acad. Sci. USA* 95:10,570-10,575。相应地, 为了本公开的目的, 将 Fok I 酶用于公开的融合蛋白中的部分视为裂解半结构域。因此, 为了使用锌指-或 TALE-Fok I 融合进行靶向双链裂解和/或靶向置换细胞序列, 各包含 Fok I 裂解半结构域的统一蛋白质可用于重新组成有催化活性的裂解结构域。可选地, 也可使用含有锌指结合结构域和两个 Fok I 裂解半结构域的统一多肽分子。在本公开的其它地方提供了使用锌指-或 TALE-Fok I 融合进行靶向裂解和靶向序列改变的参数。

[0148] 裂解结构域或裂解半结构域可为保持裂解活性, 或保持多聚化 (例如, 二聚化) 以形成功能裂解结构域的能力的蛋白质的任何部分。

[0149] 在整体并入本文的国际公布 WO 07/014275 中描述了示例性 IIS 型限制酶。另外的限制酶还含有可分离的结合和裂解结构域, 并且这些为本公开所涵盖。见, 例如, Roberts 等 (2003) *Nucleic Acids Res.* 31:418-420。

[0150] 在某些实施方案中, 例如美国专利公布第 20050064474 和 20060188987 及美国申请第 11/805,850 号 (2007 年 5 月 23 日提交) 中所述, 裂解结构域包含最小化或防止均二聚的一个或多个工程化裂解半结构域 (也称为二聚化结构域突变体), 其全部内容通过引用整体并

入本文。Fok I的位置446、447、479、483、484、486、487、490、491、496、498、499、500、531、534、537和538的氨基酸残基全部为影响Fok I裂解半结构域二聚化的靶标。形成专性杂二聚体的Fok I的示例性工程化裂解半结构域包括一对，其中第一裂解半结构域包括在Fok I的位置490和538的氨基酸残基处的突变而第二裂解半结构域包括在氨基酸残基486和499处的突变。

[0151] 因此，在一个实施方案中，490处的突变以Lys (K) 置换Glu (E)；538处的突变以Lys (K) 置换Iso (I)；486处的突变以Glu (E) 置换Gln (Q)；并且位置499处的突变以Lys (K) 置换Iso (I)。具体地，通过使一个裂解半结构域中的位置490 (E→K) 和538 (I→K) 突变以生成命名为“E490K:I538K”的工程化裂解半结构域并且通过使另一裂解半结构域中的位置486 (Q→E) 和499 (I→L) 突变以生成命名为“Q486E:I499L”的工程化裂解半结构域制备本文所述的工程化裂解半结构域。本文所述的工程化裂解半结构域为其中异常裂解减到最少或消除的专性杂二聚体突变体。见，例如，美国专利公布第2008/0131962和2011/0201055号，出于种种目的，其公开内容通过引用整体并入。在某些实施方案中，工程化裂解半结构域包含位置486、499和496 (相对于野生型FokI编号) 处的突变，例如以Glu (E) 残基置换位置486处的野生型Gln (Q) 残基，以Leu (L) 残基置换位置499处的野生型Iso (I) 残基和以Asp (D) 或Glu (E) 残基置换位置496处的野生型Asn (N) 残基的突变 (也分别称为“ELD”和“ELE”结构域)。在其它实施方案中，工程化裂解半结构域包含位置490、538和537 (相对于野生型FokI编号) 处的突变，例如以Lys (K) 残基置换位置490处的野生型Glu (E) 残基，以Lys (K) 残基置换位置538处的野生型Iso (I) 残基和以Lys (K) 残基或Arg (R) 残基置换位置537处的野生型His (H) 残基的突变 (也分别称为“KKK”和“KKR”结构域)。在其它实施方案中，工程化裂解半结构域包含位置490和537 (相对于野生型FokI编号) 处的突变，例如以Lys (K) 残基置换位置490处的野生型Glu (E) 残基和以Lys (K) 残基或Arg (R) 残基置换位置537处的野生型His (H) 残基的突变 (也分别称为“KIK”和“KIR”结构域)。(见美国专利公布第20110201055号)。

[0152] 可使用任何适合方法，例如如美国专利公布第20050064474和20080131962号中所述，通过定点诱变野生型裂解半结构域 (Fok I) 制备本文所述的工程化裂解半结构域。

[0153] 可选地，可使用所称的“裂解酶”技术，在体内在核酸靶位点处组装核酸酶 (见例如美国专利公布第20090068164号)。此类裂解酶的组成部分可在任一单独表达构建体上表达，或者可连接在一开放阅读框中，其中例如通过自我裂解2A肽或IRES序列将单独组成部分分隔开。组成部分可为单独锌指结合结构域或大范围核酸酶核酸结合结构域的结构域。

[0154] 在一些实施方案中，DNA结合结构域为来自与源自植物病原体黄单孢菌属的效应子相似TAL效应子的工程化结构域 (见Boch等，(2009) Science 326:1509-1512及Moscou和Bogdanove，(2009) Science326:1501) 和Ralstonia (见Heuer等(2007) Applied and Environmental Microbiology 73 (13):4379-4384)。同样，见PCT公布W02010/079430。

[0155] 使用之前，例如在如W0 2009/042163和20090068164中所述的基于酵母的染色体体系中，可筛选核酸酶 (例如，ZFN或TALEN) 的活性。使用本领域已知的方法可容易地设计核酸酶表达构建体。见，例如，美国专利公布20030232410、20050208489、20050026157、20050064474、20060188987、20060063231；和国际公布W0 07/014275。核酸酶的表达受组成型启动子或诱导型启动子，例如在棉子糖和/或半乳糖的存在下激活 (去阻遏) 而在葡萄糖的存在下阻遏的半乳糖激酶启动子的控制。

[0156] 递送

[0157] 可通过任何适合方式,包括(例如)通过注射ZFP-TF、TALE-TF蛋白或通过使用编码ZFN或TALEN的mRNA,向靶细胞递送蛋白质(例如,ZFP、TALE、CRISPR/Cas)、编码蛋白质的多核苷酸和包含本文所述的蛋白质和/或多核苷酸的组合物。适合的细胞包括但不限于真核和原核细胞和/或细胞系。此类细胞或由此类细胞产生的细胞系的非限制性实例包括COS、CHO(例如,CHO-S、CHO-K1、CHO-DG44、CHO-DUXB11、CHO-DUKX、CHOK1SV)、VERO、MDCK、WI38、V79、B14AF28-G3、BHK、HaK、NS0、SP2/0-Ag14、HeLa、HEK293(例如,HEK293-F、HEK293-H、HEK293-T)和perC6细胞以及昆虫细胞例如草地贪夜蛾(*Spodoptera fugiperda*)(Sf)或真菌细胞例如酵母属(*Saccharomyces*),毕赤酵母(*Pichia*)和裂殖酵母(*Schizosaccharomyces*)。在某些实施方案中,细胞系为CHO-K1、MDCK或HEK293细胞系。适合的细胞还包括干细胞,举例而言,例如胚胎干细胞、诱导多能干细胞、造血干细胞、神经元干细胞和间充质干细胞。

[0158] 例如,在美国专利第6,453,242、6,503,717、6,534,261、6,599,692、6,607,882、6,689,558、6,824,978、6,933,113、6,979,539、7,013,219和7,163,824号中描述了递送包含如本文所述的锌指蛋白的蛋白质的方法,所述专利全部的公开内容通过引用整体并入本文。

[0159] 也可使用含有编码锌指、TALE或CRISPR/Cas蛋白中的一种或多种的序列的载体递送如本文所述的锌指、TALE或CRISPR/Cas蛋白。可使用任何载体体系,包括但不限于质粒载体、逆转录病毒载体、慢病毒载体、腺病毒载体、痘病毒载体、疱疹病毒载体和腺相关病毒载体等。同样,见,美国专利第6,534,261、6,607,882、6,824,978、6,933,113、6,979,539、7,013,219和7,163,824号,其通过引用整体并入本文。此外,明显的是,任何这些载体可包含一个或多个编码锌指或TALE蛋白质的序列。因此,当向细胞引入一种或多种ZFP、TALE或CRISPR/Cas蛋白质时,可在相同载体或不同载体上携带编码ZFP、TALE或CRISPR/Cas蛋白质的序列。当使用多个载体时,每个载体均可包含编码一个或多个ZFP、TALE或CRISPR/Cas体系的序列。

[0160] 基于病毒和非病毒的常规基因转移方法可用于在细胞(例如,哺乳动物细胞)和靶组织中引入编码工程化ZFP、TALE或CRISPR/Cas体系的核酸。此类方法也可用于在体外向细胞施用编码ZFP、TALE或CRISPR/Cas体系的核酸。在某些实施方案中,施用编码ZFP、TALE或CRISPR/Cas体系的核酸用于体内或体外基因治疗用途。非病毒载体递送系统包括DNA质粒、裸露核酸和与递送媒介物例如脂质体或泊洛沙姆(poloxamer)复合的核酸。病毒载体递送系统包括在递送到细胞后具有附加体或整合基因组的DNA和RNA病毒。对于基因治疗过程的综述,见Anderson, *Science* 256:808-813 (1992); Nabel和Felgner, *TIBTECH* 11:211-217 (1993); Mitani和Caskey, *TIBTECH* 11:162-166 (1993); Dillon, *TIBTECH* 11:167-175 (1993); Miller, *Nature* 357:455-460 (1992); Van Brunt, *Biotechnology* 6(10):1149-1154 (1988); Vigne, *Restorative Neurology and Neuroscience* 8:35-36 (1995); Kremer和Perricaudet, *British Medical Bulletin* 51(1):31-44 (1995); Haddada等,在 *Current Topics in Microbiology and Immunology* 中, Doerfler和Böhm(编辑) (1995); 和Yu等, *Gene Therapy* 1:13-26 (1994)。

[0161] 核酸的非病毒递送方法包括电穿孔、脂质转染、显微注射、基因枪法、病毒颗粒、脂

质体、免疫脂质体、聚阳离子或脂质:核酸偶联物、裸DNA、裸RNA、人工病毒粒子和试剂增强的DNA摄取。使用(例如) Sonitron 2000系统(Rich-Mar)的声致穿孔法也可用于核酸的递送。在一个优选实施方案中,将一种或多种核酸作为mRNA递送。还优选使用加帽的mRNA以增加翻译效率和/或mRNA稳定性。尤其优选ARCA(抗-反向帽类似物)帽或其变体。见美国专利US7074596和US8153773,其通过引用并入本文。

[0162] 另外的示例性核酸递送系统包括Amaxa Biosystems (Cologne, Germany)、Maxcyte, Inc. (Rockville, Maryland)、BTX Molecular Delivery Systems (Holliston, MA)和Copernicus Therapeutics Inc.提供的核酸递送系统(见例如US6008336)。例如美国专利第5,049,386、4,946,787和4,897,355号中描述了脂质转染并且脂质转染试剂在市场上有售(例如,Transfectam<sup>TM</sup>和Lipofectin<sup>TM</sup>及Lipofectamine<sup>TM</sup> RNAiMAX)。适合多核苷酸的有效受体-识别脂质转染的阳离子和中性脂质包括Felgner、WO 91/17424、WO 91/16024的脂质。可向细胞(活体外施用)或靶组织(体内施用)递送。

[0163] 脂质:核酸复合物,包括靶向脂质体例如免疫脂质复合物的制备为本领域的技术人员众所周知(见,例如,Crystal, Science 270:404-410(1995); Blaese等, Cancer Gene Ther. 2:291-297(1995); Behr等, Bioconjugate Chem. 5:382-389(1994); Remy等, Bioconjugate Chem. 5:647-654(1994); Gao等, Gene Therapy 2:710-722(1995); Ahmad等, Cancer Res. 52:4817-4820(1992); 美国专利第4,186,183、4,217,344、4,235,871、4,261,975、4,485,054、4,501,728、4,774,085、4,837,028和4,946,787号)。

[0164] 另外的递送方法包括利用将待递送核酸包装到EnGeneIC递送媒介物(EDV)中。使用双特异性抗体将这些EDV特异性地递送到靶组织,其中所述抗体的一臂对靶组织具有特异性而另一臂对EDV具有特异性。抗体将EDV带到靶细胞表面,然后通过胞吞作用将EDV带入细胞。一旦进入细胞,就释放内含物(见MacDiarmid等(2009) Nature Biotechnology 27(7):643)。

[0165] 使用基于RNA或DNA病毒的系统递送编码工程化ZFP、TALE或CRISPR/Cas体系的核酸利用使病毒靶向体内的特定细胞并且将病毒有效载荷运输到核的高度进化过程。病毒载体可直接向患者(体内)施用或其可用于在体外处理细胞并且向患者(在活体外)施用经修饰的细胞。递送ZFP、TALE或CRISPR/Cas体系的基于病毒的常规系统包括但不限于用于基因转移的逆转录病毒、慢病毒、腺病毒、腺相关、牛痘和单纯性疱疹病毒载体。用逆转录病毒、慢病毒和腺相关病毒基因转移方法可能整合在宿主基因组中,常常导致插入的转基因长期表达。另外,已经在许多不同细胞类型和靶组织中观察到高转导效率。

[0166] 可通过并入外源包膜蛋白,扩增靶细胞的潜在靶群体改变逆转录病毒的向性。慢病毒载体是能够转导或感染非分裂细胞并且通常产生高病毒滴度的逆转录病毒载体。逆转录病毒基因转移系统的选择取决于靶组织。逆转录病毒载体由包装容量高达6-10kb外源序列的顺式作用长末端重复组成。最小的顺式作用LTR足以复制和包装载体,然后用于将治疗性基因整合到靶细胞中以提供永久性转基因表达。广泛使用的逆转录病毒载体包括基于小鼠白血病病毒(MuLV)、长臂猿白血病病毒(GaLV)、猿免疫缺陷病毒(SIV)、人免疫缺陷病毒(HIV)及其组合的那些(见,例如, Buchscher等, J. Virol. 66:2731-2739(1992); Johann等, J. Virol. 66:1635-1640(1992); Sommerfelt等, Virol. 176:58-59(1990); Wilson等, J. Virol. 63:2374-2378(1989); Miller等, J. Virol. 65:2220-2224(1991); PCT/US94/

05700)。

[0167] 在其中优选瞬时表达的应用中,可使用基于腺病毒的系统。基于腺病毒的载体在许多细胞类型中能够具有非常高的转导效率并不需要细胞分裂。用此类载体,已经获得高滴度和高水平的表达。在相对简单的系统中可大量生成这种载体。腺相关病毒(“AAV”)载体也用于例如在核酸和肽的体外生成中用核酸转导细胞,并且用于体内和活体外基因治疗过程(见,例如,West等,*Virology* 160:38-47 (1987);美国专利第4,797,368号;WO 93/24641;Kotin,*Human Gene Therapy* 5:793-801 (1994);Muzyczka,*J.Clin.Invest.*94:1351 (1994)。在许多出版物中描述了重组AAV载体的构建,包括美国专利第5,173,414号;Tratschin等,*Mol.Cell.Biol.*5:3251-3260 (1985);Tratschin等,*Mol.Cell.Biol.*4:2072-2081 (1984);Hermonat和Muzyczka,*PNAS* 81:6466-6470 (1984);和Samulski等,*J.Virol.*63:03822-3828 (1989)。

[0168] 至少6种病毒载体方法目前可用于临床试验中的基因转移,其利用牵涉由插入辅助细胞系中的基因互补缺陷型载体的方法以生成转导剂。

[0169] pLASN和MFG-S为已经用于临时试验的逆转录病毒载体的实例(Dunbar等,*Blood* 85:3048-305 (1995);Kohn等,*Nat.Med.*1:1017-102 (1995);Malech等,*PNAS* 94:22 12133-12138 (1997))。PA317/pLASN为用于基因治疗试验的首例治疗性载体。(Blaese等,*Science* 270:475-480 (1995))。对于MFG-S包装载体已经观察到50%或更高的转导效率。(Ellem等,*Immunol Immunother.*44 (1):10-20 (1997);Dranoff等,*Hum.Gene Ther.*1:111-2 (1997)。

[0170] 重组腺相关病毒载体(rAAV)是基于缺陷型和非病原性细小病毒腺相关2型病毒的有前景的可供选择的基因递送系统。所有载体源自仅保留了在转基因表达盒两侧的AAV 145bp反向末端重复的质粒。由于整合到转导细胞基因组中的有效基因转移和稳定的转基因递送是这种载体系统的关键特征。(Wagner等,*Lancet* 351:9117 1702-3 (1998),Kearns等,*Gene Ther.*9:748-55 (1996))。也可根据本发明使用其它AAV血清型,包括AAV1、AAV3、AAV4、AAV5、AAV6、AAV8AAV 8.2、AAV9和AAV rh10及假型AAV例如AAV2/8、AAV2/5和AAV2/6。

[0171] 复制缺陷型重组腺病毒载体(Ad)可在高滴度下生成并且容易感染许多不同细胞类型。大部分腺病毒载体经工程化,以致转基因置换Ad E1a、E1b和/或E3基因;随后使复制缺陷型载体在反式供给缺失基因功能的人293细胞中增殖。Ad载体可在体内转导多种类型的组织,包括非分裂、分化细胞例如存在于肝脏、肾脏和肌肉中的细胞。常规Ad载体承载容量大。临床试验中使用Ad载体的实例包括经肌肉内注射进行抗肿瘤免疫的多核苷酸疗法(Sterman等,*Hum.Gene Ther.*7:1083-9 (1998))。在临床试验中使用腺病毒载体进行基因转移的另外实例包括Rosenecker等,*Infection* 24:1 5-10 (1996);Sterman等,*Hum.Gene Ther.*9:7 1083-1089 (1998);Welsh等,*Hum.Gene Ther.*2:205-18 (1995);Alvarez等,*Hum.Gene Ther.*5:597-613 (1997);Topf等,*Gene Ther.*5:507-513 (1998);Sterman等,*Hum.Gene Ther.*7:1083-1089 (1998)。

[0172] 包装细胞用于形成能够感染宿主细胞的病毒粒子。此类细胞包括包装腺病毒的293细胞和包装逆转录病毒的 $\psi$ 2细胞或PA317细胞。通常由将核酸载体包装在病毒粒子中的生产细胞系生成用于基因治疗的病毒载体。所述载体通常含有包装和后续整合到宿主(若适用)所需的最小病毒序列,由编码待表达蛋白质的表达盒置换的其它病毒序列。由包装细胞系反式供给缺少的病毒功能。例如,用于基因治疗的AAV载体通常仅仅具有来自于AAV基

因组的包装和整合到宿主基因组中所需的反向末端重复 (ITR) 序列。病毒DNA包装在含有编码其它AAV基因,即rep和cap的辅助质粒,但是缺乏ITR序列的细胞系中。细胞系还受作为辅助因子的腺病毒感染。辅助病毒促进AAV载体的复制和AAV基因由辅助质粒表达。由于缺乏ITR序列,辅助质粒未大量包装。例如,可通过腺病毒比AAV更敏感的热处理减轻受腺病毒的污染。

[0173] 在许多基因治疗应用中,以对特定组织类型的高度特异性,递送基因治疗载体是可取的。相应地,可通过表达配体作为与病毒外表面上的病毒外壳蛋白的融合蛋白,将病毒载体修饰为对指定细胞类型具有特异性。选择所述配体对已知存在于目标细胞类型上的受体具有亲和力。例如,Han等,Proc.Natl.Acad.Sci.USA 92:9747-9751 (1995) 报道可将Moloney小鼠白血病病毒修饰为表达与gp70融合的人调蛋白 (heregulin),并且重组病毒感染表达人表皮生长因子受体的某些人乳腺癌细胞。这种原理可延伸到其它病毒-靶细胞对,其中靶细胞表达受体而病毒表达包含细胞表面受体的配体的融合蛋白。例如,丝状噬菌体可经工程化为展示对几乎所选任何细胞受体具有特异性结合亲和力的抗体片段 (例如,FAB或Fv)。虽然以上描述主要适用于病毒载体,但是可将相同原理应用于非病毒载体。此类载体可经工程化为含有利于被特定靶细胞摄取的特定摄取序列。

[0174] 可通过向患者个体施用,通常通过如下所述的全身施用 (例如,静脉内、腹膜内、肌肉内、皮下或颅内输注) 或局部应用,在体内递送基因治疗载体。可选地,可在活体外向细胞,例如从患者个体移植的细胞 (例如淋巴细胞、骨髓穿刺液、组织活检) 或万能供血者造血干细胞递送载体,接着通常在为已经并入载体的细胞选择后,将细胞再植入患者体内。

[0175] 用于诊断、研究或基因治疗的活体外细胞转染 (例如,通过将转染细胞再次输注到宿主生物中) 为本领域的技术人员众所周知。在一个优选实施方案中,从受试生物分离细胞,用ZFP、TALE或CRISPR/Cas体系核酸 (基因cDNA或mRNA) 转染,并且再次输注到受试生物 (例如,患者) 体内。在一个优选实施方案中,将一种或多种核酸作为mRNA递送。还优选使用加帽的mRNA以增加翻译效率和/或mRNA稳定性。尤其优选的是ARCA (抗-反向帽类似物) 帽或其变体。见美国专利7,074,596和8,153,773,其通过引用整体并入本文。适合活体外转染的各种细胞类型为本领域的技术人员众所周知 (见,例如,Freshney等,Culture of Animal Cells,A Manual of Basic Technique (1994年第3版)) 和本文引用以讨论如何从患者分离和培养细胞的参考文献)。

[0176] 在一个实施方案中,干细胞用于细胞转染和基因治疗的活体外过程。使用干细胞的优点是它们可在体外分化为其它细胞类型,或可引入哺乳动物 (例如细胞供体) 中,在这里它们将植入骨髓中。已知使用细胞因子例如GM-CSF、IFN- $\gamma$  和TNF- $\alpha$ 使CD34+细胞在体外分化成临床上重要的免疫细胞类型的方法 (见Inaba等,J.Exp.Med.176:1693-1702 (1992))。

[0177] 使用已知方法分离干细胞进行转导和分化。例如,通过用结合不必要的细胞,例如CD4+和CD8+ (T细胞)、CD45+ (panB细胞)、GR-1 (粒细胞) 和Iad (分化的抗原呈递细胞) 的抗体,淘选骨髓细胞,而从骨髓细胞分离干细胞 (见Inaba等,J.Exp.Med.176:1693-1702 (1992))。

[0178] 在一些实施方案中也可使用已经修饰的干细胞。例如,已经使其抗细胞凋亡的神经元干细胞可用作治疗组合物,其中干细胞还含有本发明的ZFPTF。例如,通过在干细胞中

使用BAX-或BAK-特异性TALEN或ZFN(见,美国专利公布第20100003756号)或在半胱天冬酶中破坏的干细胞中,例如再次使用半胱天冬酶-6特异性ZFN敲除BAX和/或BAK,可能出现对细胞凋亡的抗性。可用已知调控突变或野生型Htt的ZFP TF或TALE TF转染这些细胞。

[0179] 也可向生物直接施用含有治疗性ZFP核酸的载体(例如,逆转录病毒、腺病毒、脂质体等)以在体内进行细胞转导。可选地,可施用裸DNA。通过一般用于将分子引入与血液或组织细胞的最终接触的任何途径施用,包括但不限于注射、输注、局部应用和电穿孔。施用此类核酸的适合方法可用并且为领域的技术人员众所周知,并且,虽然可使用一种以上的途径施用特定组合物,但是常常一种特定途径可以提供比另一途径更直接且更有效的反应。

[0180] 例如,在美国专利第5,928,638号中公开了将DNA引入造血干细胞的方法。用于将转基因引入造血干细胞,例如CD34<sup>+</sup>细胞的载体包括35型腺病毒。

[0181] 适合将转基因引入免疫细胞(例如,T细胞)的载体包括非整合慢病毒载体。见,例如,Ory等(1996)Proc.Natl.Acad.Sci.USA 93:11382-11388;Dull等(1998)J.Virol.72:8463-8471;Zuffery等(1998)J.Virol.72:9873-9880;Follenzi等(2000)Nature Genetics 25:217-222。

[0182] 部分由施用的特定组合物以及由用于施用组合物的特定方法决定药学上可接受的载体。相应地,如下所述,存在可用药物组合物的多种适合配方(见,例如,Remington's Pharmaceutical Sciences,第17版,1989)。

[0183] 如上所述,公开的方法和组合物可用于任何类型的细胞,包括但不限于原核细胞、真菌细胞、古细菌细胞、植物细胞、昆虫细胞、动物细胞、脊椎动物细胞、哺乳动物细胞和人细胞。用于蛋白质表达的适合细胞系为本领域的技术人员已知并且包括但不限于COS、CHO(例如,CHO-S、CHO-K1、CHO-DG44、CHO-DUXB11)、VERO、MDCK、WI38、V79、B14AF28-G3、BHK、HaK、NS0、SP2/0-Ag14、HeLa、HEK293(例如,HEK293-F、HEK293-H、HEK293-T)、perC6、昆虫细胞例如草地贪夜蛾(Sf)和真菌细胞例如酵母属,毕赤酵母和裂殖酵母。也可使用这些细胞系的子代、变体和衍生物。

[0184] 应用

[0185] 公开的组合物和方法可用于需要调节Htt等位基因的任何应用,包括但不限于治疗和研究应用。

[0186] 阻遏Htt的ZFP TF或TALE TF可用作治疗剂的疾病和病状包括但不限于亨廷顿氏病。另外,包含对Htt突变等位基因有特异性的ZFN或TALEN的方法和组合物可用于治疗亨廷顿氏病的治疗剂。

[0187] 阻遏HD Htt等位基因的ZFP-TF或TALE TF也可连同激活神经营养因子,包括但不限于GDNF和BDNF的ZFP-TF或TALE-TF一起使用。这些ZFP或TALE(或编码这些ZFP或TALE的多核苷酸)可同时施用(例如,在相同药物组合物中)或可按任何顺序依次施用。

[0188] 治疗亨廷顿氏病的方法和组合物还包括干细胞组合物,其中已经使用Htt特异性ZFN或TALEN将干细胞中Htt等位基因的突变拷贝修饰为野生型Htt等位基因。

[0189] 本发明的方法和组合物还用于设计和实现体外和体内模型,例如,允许研究这些病症的三核苷酸重复症的动物模型。适合的体外模型的非限制性实例包括来自任何生物的细胞或细胞系,包括成纤维细胞。用作动物模型的适合动物的非限制性实例包括无脊椎动物(秀丽隐杆线虫(C.elegans)、果蝇(drosophila))、啮齿动物(例如,大鼠或小鼠)、灵长类

动物(例如,非人灵长类动物)。

### 实施例

[0190] 实施例1:Htt靶向的锌指蛋白转录因子(ZFP-TF)和ZFN的设计和构建

[0191] 基本上如美国专利第6,534,261号中所述,工程化靶向Htt的锌指蛋白。表1A和1B示出了示例性Htt靶向ZFP的DNA结合结构域的识别螺旋,而表2A和2B示出了这些ZFP的靶序列。

[0192] 设计具有一连续锌指阵列的ZFP以靶向完全在CAG重复区内的位点(图1B)。此类ZFP可以更高亲和力和/或更高的净占有率结合更长的突变簇,实现突变等位基因的选择性阻遏。还设计ZFN,靶向部分或完全位于CAG区外(图1A),并且因此将与野生型和突变等位基因同等结合,并以相似效率调控两个等位基因的表达的位点。将锌指蛋白设计为识别CAG区时,可呈“混合和匹配”组合采用一指和两指的一组模块。下面在表2C中示出了那些模块。

[0193] 表2C:用于靶向CAG重复的ZFP-TF中的锌指识别螺旋

靶位点	F2	F3	SEQ ID NO: (F2+F3)
CAGCAG	RSDNLSE	KRCNLRC	161
CAGCAG	RSDNLSE	KPYNLRT	162
CAGCAG	RSDNLSE	RLWNRKQ	163
CAGCAG	RSDNLSV	RRWNLRA	164
CAGCAG	RSDNLSV	RKWNRDS	165
CAGCAG	RSDNLSE	NTSPLML	166
CAGCAG	RSDNLSE	RRYNLVK	167
CTGCTG	RSDTLSE	RRWTLVG	168
GCAGCA	QSSDLR	QWSTRKR	169
GCAGCA	RSAHLR	QSGDLTR	170
GCAGCA	QSGDLTR	QSGDLTR	171
GCAGCA	QSGDLTR	QSSDLRR	172
GCTGCT	QSSDLR	QSSDLRR	173
GCTGCT	QSSDLR	HRSTRNR	174
AGC	MACCRYA	无	175
CAG	RSANLRE	无	176
CAG	RNADRKK	无	177
CTG	RSDVLSE	无	42
CTG	RS AVLSE	无	148
GCA	QSGDLTR	无	18
GCA	QSSDLRR	无	6
GCA	QNATRIK	无	178
GCT	QSSDLR	无	31
AAG	RSDNLRE	无	65

[0195] 除载体还含有编码一个或多个蛋白质相互作用结构域(也称为二聚化或蛋白质相互作用结构域)的序列,所述序列使得表达的蛋白质能够沿着与编码ZFP TF的序列可操作

地连接的三核苷酸重复簇多聚化外,也如上所述构建多聚化ZFP TF。见,图1D和图10。表3示出了与靶向CAG重复区的ZFP一起使用的二聚化结构域设计。图10C示出了设计为通过二聚化锌指(DZ),DZ1-DZ4之间的相互作用多聚化的4个ZFP单体支架的蛋白质序列。设计基于Mol.Syst.Biol.(2006)2:2006.2011中描述的工作。图10D示出了设计为通过卷曲螺旋(CC),CC1-CC7之间的相互作用多聚化的7个ZFP单体支架的蛋白质序列。CC#1的设计基于(J.Am.Chem.Soc.(2001),123:3151-3152)中描述的工作,而CC#2、CC#3和CC#4基于(J.Am.Chem.Soc.(2000),122:5658-5659)。CC和DZ结构域允许ZFP在CAG簇的大沟内聚合(图10B中有描绘)。通过选择指阵列和具有适当结合性质的二聚化结构域,将仅对疾病等位基因的扩增CAG簇发生有效结合。

[0196] 将ZFP-TF构建成包含核定位序列、靶向Htt等位基因的工程化锌指DNA结合结构域(表1A和1B)和来自于人KOX1蛋白质的KRAB阻遏结构域的融合蛋白。见图1A、1B和1D。设计的DNA结合结构域含有3-6个识别9-18bp序列的指模块(表2A和2B)。用大写字母指示靶位点中与ZFP识别螺旋接触的核苷酸;用小写字母指示非接触核苷酸。还构建了ZFP-ZFP-TF分子,其中两个ZFP DNA结合结构域用柔性接头融合并且与KRAB阻遏结构域融合(图1E)。DNA结合结构域选自表2A和2B。

[0197] 实施例2:人和小鼠细胞中Htt两个等位基因的阻遏

[0198] 为阻遏Htt的两个等位基因(非等位基因特异性),将ZFP设计为与Htt启动子和外显子1区结合,其中靶位点不完全在CAG重复内。见,图1A。为测试阻遏Htt的ZFP TF的活性,将ZFP TF转染至人细胞中并且使用实时RT-PCR监测Htt的表达。

[0199] 在补充了10%FBS的DMEM中培养人HEK293细胞(Graham等(1977)J Gen Virol 36:59-74)并且通过Amaxa **Nucleofector**<sup>®</sup>,按照生产商的说明,用1 $\mu$ g编码所示ZFP-KOX融合蛋白的质粒DNA转染 $1e^5$ 个细胞。

[0200] 孵育转染细胞2天,并且根据标准方案,分别使用Hs00918176\_m1和4352935E引物和探针(Applied Biosystems),通过实时PCR分析内源性人亨廷顿蛋白(Htt)和标准化对照 $\beta$ -肌动蛋白(ACTB)的水平。将Htt水平表示成标准化为模拟转染样品(设为1)的Htt/ACTB比例。

[0201] 如图2A所示,Htt靶向ZFP阻遏Htt表达。使用标准方案进行蛋白质印迹分析以确认Htt蛋白质水平降低(图2B);p65蛋白质用作负载对照。

[0202] 使用**Lipofectamine**<sup>®</sup> 2000试剂盒(Invitrogen)根据生产商的方案,将小鼠Htt特异性ZFP TF阻遏因子瞬时转染至Neuro2A细胞中(Klebe和Ruddle(1969)J.Cell Biol.43:69A)。转染48小时后,分别使用ABI **Taqman**<sup>®</sup> 引物/探针组Mm01213820m1和4352933E测量mHtt和ACTB mRNA水平。将ZFP转染样品的mHtt/ACTB比例标准化为GFP对照的mHtt/ACTB比例(设为1)。

[0203] 如图2C所示,ZFP阻遏小鼠Htt表达。另外,小鼠Htt特异性ZFP TF阻遏因子可阻遏源自Htt敲入小鼠的永生化纹状体细胞STHdh(Q111/Q7)中的小鼠Htt(Trettel等(2000)Hum.Mol.Genet 9:2799-2809)。见,图2D和2E。使用mMessage mMachine试剂盒(Ambion)生成所示ZFP的mRNA,并且如上所述使用Amaxa核转染仪转染0.1、0.5或2 $\mu$ g这些mRNA。如上所述,转染48小时后收获细胞进行mHtt和ACTB表达分析。与在Neuro2A细胞中相比,在纹状体

细胞中观察到更显著的阻遏并且可能是通过纹状体细胞中mRNA转染达到更高的转染效率的结果。

[0204] 实施例3:人和小鼠细胞中突变Htt的选择性阻遏

[0205] 为实现突变Htt等位基因的选择性阻遏,将ZFP设计为结合在CAG重复内。图1B示出了具有与阻遏结构域(例如,来自KOX1的KRAB结构域)连接的连续锌指阵列的此类ZFP的一种类型;这些ZFP可设计为具有适当亲和力,以致可仅在扩增CAG重复上建立转录阻遏所需的阈值占有率。图1C、1D和1E示出了可允许与扩增CAG重复特异性结合的ZFP设计的另外3个实例。

[0206] 如图1B所示设计的ZFP引入HEK293细胞中并且评估Htt表达。使用FugeneHD,使用标准方案将编码ZFP的构建体转染至HEK293细胞中。转染72小时后,分离总RNA并且分别使用Hs00918176\_m1和4352935E引物和探针(Applied Biosystems),通过实时PCR分析相对于内部对照 $\beta$ -肌动蛋白(ACTB)的内源性人亨廷顿蛋白(Htt)的水平。将ZFP转染样品的Htt/ACTB比例标准化为GFP对照的Htt/ACTB比例(设为1)。

[0207] 如图3A所示,设计为与HEK293细胞中的CAG重复(上链或下链)结合的ZFP阻遏因子(与KRAB阻遏结构域融合)(图1B)有效阻遏Htt表达。图3A描绘了转录的阻遏因子,其中在重复转染中测量了表达(图中的单独柱)并且完成多次实时PCR测定(误差线)。受单独ZFP不同水平的阻遏表明它们对CAG重复区具有不同亲和力。因为HEK293细胞中的Htt等位基因具有16和17个CAG,所以这个结果还表明“较弱”ZFP例如30640不会有效阻遏具有野生型(未扩增)CAG重复长度的Htt等位基因。

[0208] 为试验ZFP例如30640是否可以阻遏具有扩增CAG重复的Htt等位基因的转录,构建了受含有不同CAG重复长度的Htt启动子/外显子1片段控制的荧光素酶报告基因。首先,使用正向引物:5' GAAGATCTCACTTGGGGTCCTCAGGTCGTGCCGAC (SEQ ID NO:139)

[0209] 和反向引物:

[0210] 5' GTATCCAAGCTTCAGCTTTTCCAGGGTCGCTAGGCGGTCT (SEQ ID NO:140)

[0211] 从HEK293基因组DNA扩增人Htt启动子/外显子1片段。

[0212] 正向引物引入BglIII位点,反向引物将Htt的第一个ATG改变为TAG并且产生AvrII位点,并且还包含HindIII位点。PCR产物经BglIII和HindIII消化并连接到经相同酶消化的pRL-TK载体(Promega)以生成构建体pRL-Htt。然后使用正向引物:

[0213] 5' GCCTAGGCGACCCTGGAAAAGCTGATGAAGGCC (SEQ ID NO:141)

[0214] 和反向引物:5'

[0215] 5' GTATCCAAGCTTGAGCTGCAGCGGGCCAACTCAG (SEQ ID NO:142)

[0216] 用扩增CAG重复从HEK293基因组DNA或来自HD患者的基因组DNA扩增人Htt外显子1片段(编码序列减去第一个ATG)。

[0217] 正向引物引入AvrII位点,反向引物引入HindIII位点。PCR产物经AvrII和HindIII消化并且与经相同酶消化的pRL-Htt载体连接。通过测序鉴定具有10、17、23或47个CAG重复的克隆(pRL-Htt-CAG(x))。

[0218] 将pRL-Htt-CAG(x)报告基因(300ng)和pGL3-启动子报告基因(100ng,用作标准化对照,Progema)转染至有或无100ng的ZFP 30640表达载体的HEK293细胞中。转染24小时后测量萤火虫(pGL报告基因)和海肾(pRL报告基因)荧光素酶活性。将海肾荧光素酶水平标准

化为来自同一转染样品的萤火虫荧光素酶水平,并进一步标准化为“仅报告基因”样品的海肾/萤火虫比例。

[0219] 如图3B所示,荧光素酶报告基因受ZFP-TF 30640的阻遏随CAG重复的长度增加,表明具有与30640相似的DNA结合亲和力的ZFP可通过扩增的CAG重复阻遏Htt启动子活性,并且阻遏水平取决于CAG重复长度。

[0220] 除还试验了“强”ZFP-TF 30657,并且在所示剂量下试验了30640和30657外,图3C示出了与图3B所示的相似实验。在每个剂量水平下,30640对pRL-Htt-CAG47报告基因产生比pRL-Htt-CAG23报告基因多的阻遏(CAG重复长度依赖性阻遏),而30657以相似水平阻遏两个报告基因。关于pRL-Htt-CAG23报告基因,在每个剂量水平下30640产生少于30657的阻遏,概括了其对具有正常CAG重复长度的内源性Htt等位基因的活性差异(HEK293细胞,图3A);但是关于pRL-Htt-CAG47报告基因,在每个剂量水平下30640和30657产生相似阻遏,表明“较弱的”ZFP例如30640可通过扩增的CAG重复有效阻遏Htt启动子,最有可能是因为只有扩增的CAG靶标可允许由此类ZFP建立阻遏所需的阈值占有率。

[0221] 图3D显示ZFP-TF 30640和30657(与K0X1的KRAB阻遏结构域融合)可阻遏源自Hdh(Q111/Q7)敲入小鼠的永生化纹状体细胞中的敲入Htt等位基因(CAG111),证明驱动荧光素酶报告基因的CAG重复长度依赖性阻遏的ZFP例如30640也可阻遏由已经扩增了CAG重复的内源性Htt等位基因的表达。所示ZFP的mRNA使用mMessage mMachine试剂盒(Ambion)生成,并且使用Amaza核转染仪在指示剂量下转染至Hdh(Q111/Q7)中。为检测由野生型小鼠Htt等位基因的表达,在实时RT-PCR中使用正向引物CAGGTCGGCAGAGGAACC(SEQ ID NO:193)和反向引物TTCACACGGTCTTTCTTGGTGG(SEQ ID NO:194);为检测由敲入Htt等位基因的表达,使用正向引物GCCCGGCTGTGGCTGA(SEQ ID NO:195)和反向引物TTCACACGGTCTTTCTTGGTGG(SEQ ID NO:196)。

[0222] 图3E示出了在正常和突变Htt等位基因上分别具有15和70个CAG的HD患者来源的成纤维细胞系(GM21756, Coriell)中试验ZFP-TF 30640和30657的结果。首先建立基于SNP的等位基因特异性实时PCR测定以允许来自野生型或突变Htt等位基因的特定检测。Carroll等(Mol Ther. (2011) 19:2178-85)确定了SNP的定相(rs363099T/C);“T”在正常等位基因上而“C”在突变等位基因上。为检测由突变等位基因(099C)的表达,使用正向引物099C.F(5' AGTTTGGAGGGTTTCTC, SEQ ID NO:143)和反向引物099.R5(5' TCGACTAAAGCAGGATTCAGG, SEQ ID NO:144)通过实时PCR(SsoFast EvaGreen Supermix, Bio-Rad)扩增来自成纤维细胞的cDNA;退火/延伸温度为55.6°C。为检测由野生型等位基因(099T)的表达,使用正向引物099T.F(5' AGTTTGGAGGGTTTCTT, SEQ ID NO:145)、反向引物099.R5和3'磷酸化阻断剂寡聚099T.BL(5' AGGTTTCTCCGCTCAGC-3'磷酸, SEQ ID NO:146)进行成纤维细胞cDNA的实时PCR;退火/延伸温度为58.3°C。分别使用引物/探针Hs00918176\_m1和4352935E(Applied Biosystems),通过实时PCR分析人亨廷顿蛋白(hHtt,野生型和突变等位基因)总水平和标准化对照β-肌动蛋白(ACTB)水平。对于图3E中所示的实验,使用mMessage mMachine试剂盒(Ambion)生成所示ZFP的mRNA,如以上使用Amaza核转染仪转染1μg的mRNA。转染48小时后收获细胞;如上所述量化来自正常(CAG15, 099T)、突变(CAG70, 099C)Htt等位基因和总Htt(hHtt)的mRNA水平并且标准化为ACTB的水平;进一步将每种样品的Htt/ACTB比例标准化为模拟转染样品的Htt/ACTB比例。不出所料(基于其在

HEK293细胞中的活性,图3A),“强”CAG靶向的ZFP 30657阻遏两个等位基因。显示出对报告基因的CAG重复长度依赖性阻遏的ZFP 30640产生<10%的野生型等位基因阻遏,而阻遏突变等位基因>90%。每种样品中总Htt的水平与同样品中野生型和突变Htt的水平一致。

[0223] 还在正常成纤维细胞系以及在Htt基因中含有不同CAG重复长度的其它HD成纤维细胞系中试验了ZFP-30640(见图3F)。如上所述检测来自每个等位基因的Htt表达。在正常成纤维细胞系(CAG18/18)中未观察到Htt阻遏。相反,在CAG 15/67和CAG15/70细胞系中,在高和低转染30640mRNA剂量下均观察到优良等位基因分型;对于在突变等位基因(CAG 18/44和CAG18/45)上具有中间CAG重复长度的两个HD成纤维细胞系获得类似结果-其中在30640的高和低剂量下,扩增等位基因被阻遏约80%,而CAG18等位基因仍未受影响。总之,这些数据表明在更为普遍的疾病基因型例如CAG18/44和CAG18/45的情况下,等位基因特异性阻遏因子例如30640可保持很强的CAG等位基因长度选择性。

[0224] 使用蛋白质印迹分析证实,ZFP例如30640选择性下调了两个患者来源的成纤维细胞系中的突变Htt蛋白,确认了通过qPCR测定证实的等位基因特异性调节(见图3G)。在300ng剂量下通过mRNA转染(Amaxa核转染)将ZFP递送到4份 $1.5 \times 10^5$ 细胞复制品中并且在植入12孔板之前合并。48小时时,洗涤细胞并收获进行蛋白质提取物制备。将约2.5 $\mu$ g提取物负载到5%Tris-乙酸酯凝胶上并通过MAB2166(Millipore)检测。另外,将同样品负载到4-15%Tris-HCl凝胶上并使用标准方法转染,以通过抗B-肌动蛋白(1:20,000,Sigma)作为负载对照进行检测。基于测量Htt mRNA的qPCR研究,30640是靶向CAG重复的等位基因特异性阻遏因子;32528是靶向转录起始位点(TSS)的双等位基因阻遏因子,并且30657为在使用剂量下阻遏两个Htt等位基因的CAG-靶向阻遏因子。蛋白质印迹显示30640特异性降低了两个HD来源的细胞系中突变Htt(上方条带)的水平,而32528和30657相似地阻遏两个等位基因。

[0225] 实施例4:驱动Htt的等位基因特异性阻遏的附加CAG靶向ZFP设计

[0226] 图4A示出了在CAG18/45HD成纤维细胞系中试验ZFP-TF 30640、30643、30645和33074(全部靶向CAG重复并且使用KRAB阻遏结构域)的结果。使用所示Amaxa核转染仪转染不同量的ZFP mRNA,如上所述在转染24小时后测量突变Htt(右侧柱)、野生型Htt(中间柱)和总Htt(两个等位基因,左侧柱)的表达。在整个3 $\mu$ g-10ng ZFP mRNA剂量范围,ZFP 30640、30645和33074驱动等位基因特异性阻遏;而在30ng或更高剂量下30643似乎明显阻遏两个等位基因,并且在10ng剂量下开始表现出等位基因选择性。

[0227] 图4B示出了在CAG15/70HD成纤维细胞系中试验ZFP 30643、30648、30657和30658(全部靶向CAG重复并且使用KRAB阻遏结构域)的结果。使用所示Amaxa核转染仪转染不同量的ZFP mRNA,如上所述在转染24小时后测量突变Htt(右侧柱)、野生型Htt(中间柱)和总Htt(两个等位基因,左侧柱)的表达。与上一幅图中试验的ZFP(30640、30645和33074)相比,这些ZFP在较低剂量下驱动突变Htt特异性阻遏。这些结果表明,根据体内(例如,在HD患者脑部)可达到的ZFP表达水平,可使用适当的ZFP设计实现突变Htt的等位基因特异性阻遏。

[0228] 实施例5:含CAG的替代基因的阻遏

[0229] 使用实施例3(图3E)中分离的RNA,分析其它含CAG重复的基因的阻遏,并且图5中描绘了结果。使用实时PCR检查下列基因的表达水平并标准化为肌动蛋白:脊髓小脑共济失调蛋白2(“ATXN2”);动力蛋白(“DNM1”);F-box only蛋白11(“FBX011”);硝酸还原酶(“NAP”);起始识别复合物亚基4(“ORC4”);磷酸激酶(“PHK”);OCT3/4蛋白(“POU3”);含THAP

结构域的凋亡相关蛋白2 (“THAPII”); TATA结合蛋白 (“TBP”); 和斯钙素1 (“STC1”)的水平。另外,指出了CAG重复序列相对于转录起始位点 (TSS) 的位置,并且在图3F中表示为“TSS@”,其中“+”指示TSS下游的CAG重复的碱基位置,而“-”指示TSS上游的CAG重复的碱基位置。同样,对每个基因指示了CAG重复的数量 (“#CAG”)。

[0230] 数据证明30640对突变扩增Htt等位基因的阻遏有高度特异性,并且只有CAG重复相对靠近其各自的转录起始位点的含CAG重复的基因亚类可能是ZFP例如30640的阻遏靶标。

[0231] 实施例6:Htt的等位基因特异性ZFP阻遏因子的全基因组特异性

[0232] 转染HD成纤维细胞 (CAG18/45) 以通过微阵列分析 (图6) 研究CAG靶向ZFP的全基因组特异性。在指示剂量下通过mRNA转染 (Amaxa核转染) 递送ZFP-分别在75ng、15ng和15ng剂量下转染ZFP 30640、30645和33074,一式六份;转染GFP-Kox mRNA (150ng) 作为对照,并且用作载体将总量的经转染mRNA带到总计150ng样品中。转染24小时后如上所述通过等位基因特异性qPCR试剂测量由CAG18 (099T, 中间柱) 和CAG45 (099C, 右侧柱) 等位基因的表达,其中每个样品 (1-6) 均为生物复制品 (单独转染)。将Htt水平标准化为GAPDH水平。对全部3种ZFP观察到Htt的突变等位基因特异性阻遏。然后选择最相似的4个复制品进行微阵列分析 (Affymetrix HGU133plus2.0), 如下: GFP复制品1、3、4和6; 30640复制品2、3、5和6; 30645复制品2、3、5和6; 和33074复制品1、3、4和5用于微阵列分析。使用强大的多阵列平均 (RMA) 标准化来自每个探针组的原始信号; 使用T检验将ZFP转染样品与GFP转染样品进行比较; 对基因 (探针组) 产生“变化”命令, 相对于对照样品差异>2倍和T检验P值<0.05。基于该准则, 30640仅阻遏两个基因, 斯钙素1 (STC1) 和延伸突触结合蛋白样蛋白1 (ESYT1); 30645和33074各自分别仅阻遏一个基因, STC1和白细胞间素17受体A (IL17RA)。因为阵列上的Htt探针组检测野生型和突变Htt mRNA, 所以未检测到Htt呈阻遏 (>2倍阻遏) 基因。这个实验证明, 突变Htt特异性ZFP在驱动有效的Htt等位基因特异性阻遏的水平下表达时, 可以极高的全基因组特异性起作用。

[0233] 实施例7: HD神经干细胞 (NSC) 中的等位基因特异性阻遏

[0234] 用accutase使HD iPSC/ESC传代并且在涂有基质胶的板上的E8培养基 (Life Technologies) 中培养。使用StemPro神经诱导培养基 (Life Technologies) 驱动神经干细胞。简言之, 以200,000个细胞/孔将iPSC/ESC接种到涂有geltrex的6孔皿上并且当10-20%汇合时, 将培养基更换为StemPro神经诱导培养基。每2天更换培养基并且收获NSC并在第7天扩增。使用StemPro NSC SFM培养基 (Life Technologies) 培养NSC。使用核转染, 用1.5或0.5 $\mu$ g ZFP mRNA转染HD NSC (CAG17/69, 源自Coriell GM23225 iPSC)。转染48小时后, 收获细胞并且通过RT-PCR量化表达。使用基于SNP (rs1143646) 的基因分型测定#4351376 (Applied Biosystems) 进行Htt表达的等位基因特异性检测。在试验的ZFP剂量下, 30640引起突变Htt的等位基因特异性阻遏, 30643引起野生型Htt约50%阻遏和突变Htt约90%阻遏, 而30648阻遏两个等位基因 (图7); 这些ZFP的性状与在HD成纤维细胞中一致 (图4)。每个样品的总Htt水平 (中间柱) 与突变和野生型Htt水平一致。

[0235] 实施例8: 在分化HD神经元中的Htt阻遏

[0236] 在涂有geltrex的板上用accutase使HD NSC传代。通过将培养基更换为含有StemPRO NSC SFM培养基, 没有 (bFGF和EGF) 的神经分化培养基来诱导神经元分化。每3-4天

更换培养基,长达21天。通过在神经分化培养基中培养从NSC (CAG17/48,源自HD ESC得到神经元。在神经诱导后第15天,使用核转染用1.0或0.5 $\mu$ g ZFP mRNA转染细胞。转染48小时后,收获细胞并且通过RT-PCR量化基因表达。因为该患者系不含允许基于qPCR的野生型和突变Htt的等位基因特异性检测的SNP,所以仅可测量总Htt水平。因为我们证实30640和33074不阻遏HD成纤维细胞和NSC中的CAG18或CAG17等位基因,所以在30640和33074处理样品中观察到的总Htt水平与突变等位基因 (CAG48) 的等位基因特异性阻遏一致。在试验的ZFP剂量下,受30643和30648更有力的阻遏也与HD成纤维细胞中这些ZFP的性状一致 (图8)。

[0237] 实施例9: CAG靶向阻遏因子阻遏R6/2小鼠体内的突变Htt转基因

[0238] R6/2小鼠 (其携带具有约120个CAG重复的突变人Htt外显子1的转基因,见Mangiarini等, (1996) Cell 15:197) 接受立体定位、双侧纹状体注射 $3 \times 10^6$ 个受CMV启动子驱动编码ZFP 30640-KOX或GFP的重组AAV2/6载体基因组。在5周龄时为小鼠注射并且在8周龄时处死进行分子分析。从每个半球解剖左侧和右侧纹状体并且速冻。为评估突变Htt转基因的阻遏,用TRIzol Plus (Life Technologies) 从每个纹状体提取总RNA,接着使用大容量RT (Life Technologies) 进行cDNA合成。随后,如Benn等 (2008) Molecular Neurodegeneration:3,17) 先前所述,通过qPCR测量R6/2转基因表达并且标准化为3个参考基因 (Atp5b、Eif4a2、UbC) 的几何平均数。相对于4个经GFP处理的对照纹状体,我们在4个经ZFP处理的纹状体中观察到突变Htt转基因在统计上显著的阻遏 ( $P < 0.001$ ) (图9)。R6/2平均阻遏为经GFP处理的对照的64.9%。因为使用单一立体定向注射未实现纹状体的完全覆盖并且AAV2/6优先转导神经元细胞,所以观察到的阻遏倍数 (约35%) 很可能是在经AAV载体转导的细胞中实际阻遏的低估计值。

[0239] 实施例10: 使用具有二聚化/多聚化结构域的ZFP选择性阻遏突变Htt

[0240] 为了工程化锌指转录因子以更好地区别短CAG和长GAG重复,我们力图降低单独锌指转录因子的DNA结合亲和力和增加与CAG重复内相邻亚位点结合的融合蛋白的不同拷贝之间的相互作用强度。为了降低单独锌指转录因子的DNA结合亲和力,我们生成预计更少锌指和/或氨基酸序列以低于最佳的亲和力与DNA结合的锌指结构域。为了增加与CAG重复内相邻亚位点结合的融合蛋白的不同拷贝之间的相互作用强度,我们使各种二聚化结构域与我们的锌指转录因子融合。二聚化结构域可呈“平行”方式相互作用并且产生含有二聚化结构域的融合蛋白的“头与头”或“尾与尾”二聚体。一种潜在二聚化方案需要在“头与尾”方向结合的相同ZFP转录因子融合蛋白阵列,并且因此这种方案需要呈“反平行”方式相互作用的二聚化结构域。见,例如,McClain等 (2001) J. Am. Chem. Soc. 123:3151-3152) 和二聚化锌指肽 (Giesecke等 (2006), Molecular Systems Biology 2:2006.2011)。

[0241] 二聚化构建体CC1和CC2基于成对的反平行卷曲螺旋 (McClain等,同上, Ghosh等 2000) J Am Chem Soc 122:5658-5659)。二聚化构建体CC3和CC4是分别缺乏4个残基或7个残基的CC2截短形式。二聚化构建体DZ1、DZ2、DZ3和DZ4基于成对的二聚化锌指结构域 (Giesecke等,同上)。在每种情况下,所述对的一个成员与锌指DNA结合结构域的N端融合并且所述对的另一个成员与锌指DNA结合结构域的C端融合。富含甘氨酸和丝氨酸残基的短接头用于将二聚化结构域与锌指结合结构域融合。本发明另外的实施方案利用具有替代长度和/或氨基酸组成的接头。去除了一个或多个残基或有一个或多个甘氨酸或丝氨酸残基变为其它氨基酸残基的接头将降低这些接头的柔性并且可能导致长短CAG重复之间的区别改

进。

[0242] 为实现突变Htt等位基因的选择性阻遏,如图1D和图10A和4B中所示,设计ZFP。图10C和10D示出了多聚化结构域的序列。图11A和11B描绘了设计用于测量包含CC和DZ结构域的ZFP-TF分别阻遏其靶标的能力的实验结果。对于这些实验而言,将指示的ZFP构建体(50ng)与pRL-Htt-CAG17(200ng)、pGL3-Htt-CAG47(200ng)和pVax-SEAP(分泌型碱性磷酸酶,10ng,用作标准化对照)共转染至HEK293细胞中。转染24小时后测量荧光素酶活性和分泌型碱性磷酸酶活性。将每个样品的海肾荧光素酶(CAG17)/SEAP和萤火虫荧光素酶(CAG47)/SEAP比例标准化为仅报告基因样品的比例。除用pGL-启动子构建体(Promega)代替pRL-TK构建体外,以与pRL-Htt-CAG47报告基因相同的方式(见实施例3)构建pGL3-Htt-CAG47报告基因。

[0243] 如图11A所示,3种ZFP,当作为一个或多个含CC结构域的构建体试验时,与具有相同ZFP但是无CC结构域的构建体相比时,增强了对一个或两个报告基因的阻遏。图11B显示DZ1和DZ3结构域增强了32220对两个报告基因的阻遏。

[0244] 这些结果表明CC和DZ结构域一般可增加多聚化ZFP的亲合力并且DNA结合结构域和二聚化结构域的设计可产生最佳CAG重复长度区别。

[0245] 实施例11:ZFP-ZFP-Kox设计对突变Htt的选择性阻遏

[0246] 在具有ZFP-ZFP-KOX设计的HD成纤维细胞中试验ZFP TF。在这些实验中,两个ZFP DNA结合结构域经柔性接头(LRQKDAARGSAAMAERPFQ,SEQ ID NO:179)连接在一起并且与KOX阻遏结构域融合。所述接头位于保守性组氨酸和半胱氨酸之间。如上所述,使用ZFP mRNA在指示剂量下试验蛋白质。这些在CAG18/45(图12A)和CAG20/41(图12B)HD成纤维细胞系中的结果证明,以这种方式连接较低活性的ZFP DNA结合结构域可产生驱动等位基因特异性阻遏的复合ZFP。

[0247] 实施例12:小鼠细胞中Htt的激活

[0248] 还评估了本文所述ZFP激活Htt表达的能力。靶向小鼠Htt的+200至+467bp区域(相对于转录起始位点)的ZFP与NF B p65亚基的转录激活结构域融合。选择这个靶向区域,是因为在HD的各种敲入小鼠模型中,这个片段由来自于人Htt的相应序列(大部分外显子序列和一些内含子1序列)置换(Menalled等(2003) J.Comp.Neurol 4651:11-26;Wheeler等(2000) Hum Mol Genet 8:115-122),因此靶向该区域的ZFP可选择性激活那些动物体内的野生型等位基因,而非敲入等位基因。

[0249] 将ZFP转染至Neuro2A细胞中(在这些细胞中两个Htt等位基因均为野生型),如实施例2所述(重复转染和多次测定)测量小鼠Htt和ACTB mRNA水平。

[0250] 如图13所示,使用两种ZFP-TF检测Htt mRNA水平与模拟转染相比增加。见,图13A。通过蛋白质印迹确认Htt蛋白质水平增加。见,图13B。

[0251] 图13C中说明了敲入Htt等位基因的产生;序列比对(图13D)显示经置换的小鼠序列与相应人序列之间的偏差。图13E显示,当此类ZFP激活因子转染到源自HdhQ111/Q7敲入小鼠的永生生化纹状体细胞中时,仅选择性激活野生型Htt。

[0252] 实施例13:体内Htt表达的调控

[0253] 为了试验体内Htt特异性ZFP TF的功效,生成编码ZFP的AAV2载体。然后将这些基于AAV2的构建体递送到小鼠脑部。对于人Htt特异性ZFP TF而言,将AAV载体递送到R6.2小

鼠或BAC HD小鼠(C57B1/6或FVB/N品系)以评估对人转基因的阻遏以及HD样表现型的变化。对于小鼠Htt特异性ZFP(激活因子或阻遏因子)而言,将AAV载体递送到野生型小鼠(C57B1/6或FVB/N)或人Htt敲入小鼠HdhQ111/Q7、HdhQ140/Q7或HdhQ175/Q7)以评估内源性小鼠Htt表达的激活或阻遏。对于优先靶向CAG扩增等位基因的ZFP而言,将AAV载体递送到R6.2小鼠或人Htt敲入小鼠HdhQ111/Q7、HdhQ140/Q7或HdhQ175/Q7)以检查野生型与扩增Htt等位基因的选择性阻遏。处死后,通过Taqman实时RT-PCR分析脑组织的Htt表达,并且证明Htt基因受ZFP-TF调节。

[0254] 实施例14:神经营养因子和HD Htt等位基因特异性ZFP TF的共转染

[0255] 以上鉴定的Htt特异性ZFP TF与对脑神经营养因子有特异性的ZFP TF共转染。对使用的脑神经营养因子有特异性的ZFP TF对GDNF或BDNF有特异性。

[0256] 实施例15:靶向Htt的锌指核酸酶(ZFN)的设计和构建

[0257] 将靶向人Htt和小鼠Htt的ZFN设计为靶向在CAG重复两侧的序列,第一编码ATG、终止密码子附近以及早期外显子内的的序列。基本上如Urnov等(2005)Nature 435(7042):646-651、Perez等(2008)Nature Biotechnology 26(7):808-816和美国专利公布2008/0131962中所述,设计ZFN并且并入质粒或腺病毒载体中。

[0258] 实施例16:Htt特异性ZFN的裂解活性

[0259] 为测试裂解活性,将如上所述编码成对的人Htt特异性ZFN的质粒转染至K562细胞中。从美国模式培养物保藏所(American Type Culture Collection)获得K562细胞并且像推荐那样,使其在补充了10%合格胎牛血清(FCS,Cyclone)的F-12培养基(Invitrogen)中生长。使用TrypLE Select™蛋白酶从塑料器皿分离细胞。为了转染,将一百万个K562细胞与2g锌指核酸酶质粒和100 L Amaxa溶液T混合。在Amaxa Nucleofector II™中使用程序U-23转染细胞并且回收到1.4mL温热F-12培养基+10%FCS中。

[0260] 收获基因组DNA并且对涵盖预期裂解位点的Htt基因座部分进行PCR扩增。如下,使用来自Invitrogen的Accuprime HiFi聚合酶进行PCR:在94℃下最初3分钟变性后,经94℃下30秒变性步骤,接着是58℃下30秒退火步骤,接着是68℃下30秒延伸步骤进行30个PCR循环。完成30个循环后,在68℃下使反应孵育7分钟,然后无限期地在10℃下孵育。

[0261] 通过例如在美国专利公布第20080015164、20080131962和20080159996号中描述的Surveyor™核酸酶(Transgenomic)检查来自经K562Htt特异性ZFN处理的细胞的基因组DNA。

[0262] 以类似方式在Neuro-2a细胞中试验了编码成对小鼠Htt特异性ZFN的质粒。

[0263] 图14A和B显示,如先前所述按观察到的插入缺失量测定,ZFN能够靶向具有8-40%之间的基因修饰效率的Htt基因。

[0264] 实施例17:不同长度的三核苷酸重复的靶向整合

[0265] 在靶向整合方案中使用对如上所述CAG重复两侧的序列的裂解活性最强的Htt特异性ZFN将不同长度的CAG重复引入Htt的野生型拷贝中。构建含有50、80、109和180个重复CAG单元的供体。然后将这些供体转染至如上所述具有编码Htt特异性ZFN的质粒的K562细胞中。通过基因组DNA分离、PCR扩增(如上所述),接着是目标区域测序实现供体整合的验证。

[0266] 在K562细胞中鉴定的导致供体等位基因靶向整合到Htt等位基因中的ZFN用于将

可变长度供体核酸插入人胚胎干细胞 (hESC) 中。如上所述通过基因组DNA分离、PCR和测序验证供体整合成功。

[0267] 实施例18:野生型和/或突变Htt的破坏/敲除

[0268] 裂解早期外显子的ZFN可导致小的插入或缺失(插入-缺失),因此导致非同源末端连接(NHEJ),这样可产生一个或两个Htt等位基因破坏的细胞模型。

[0269] 如上所述制备指示的ZFN对并且使用如对实施例8描述的Cel I错配测试裂解活性。这些ZFN对靶向人Htt的早期外显子,并且因此用于敲除野生型或突变Htt等位基因。

[0270] 如图14A所示,ZFP对29627/29628、29631/29632(外显子12)和29637/29638(外显子18)裂解Htt基因并且因此可用于生成敲除细胞系。

[0271] 实施例19:野生型和HD Htt等位基因的表达标记

[0272] 对第一或最后一个编码外显子的裂解活性最强的ZFN用于标记具有不同报告蛋白的野生型和突变Htt等位基因。基于前导ZFN对的裂解位点设计每个报告基因(A和B)的供体DNA以允许报告基因的靶向整合以产生与Htt的框内融合。供体DNA与前导ZFN对共转染至K562细胞中,以选择产生最高频率整合的供体DNA构建体。

[0273] 如上所述制备ZFN对并且使用如对实施例8描述的Cel I错配测试裂解活性。使用的ZFN对靶向Htt编码序列的3'末端,并且因此可用于靶向野生型或突变Htt等位基因。如图14B所示,ZFP对25917/25916、25920/25921和25923/25922能够裂解Htt基因并且因此可用于引入报告基因标签。

[0274] 将为沿着相应ZFN的报告基因A选择的供体DNA构建体递送到源自携带突变Htt基因的受试者的细胞中(例如成纤维细胞、诱导多能细胞)。得到克隆并筛选用于报告基因A的靶向整合。期望杂合事件并且通过PCR鉴定靶向等位基因。选择含有单个报告基因标记的Htt等位基因和另一等位基因上未经修饰的ZFN靶序列的克隆;转染报告基因B和相应ZFN的供体构建体以用报告基因B标记第二等位基因。

[0275] 所得小鼠胚胎干细胞克隆含有经允许跟踪由每个等位基因的表达的两个不同标志标记的野生型Htt等位基因和突变等位基因;使用标准方案,用这些细胞产生三核苷酸重复症的小鼠模型。

[0276] 实施例20:抗Htt的活性TALE-TF蛋白质的构建

[0277] TALE DNA结合结构域与来自Kox1蛋白质(TALE TF)的KRAB阻遏结构域连接并且用于测试HD患者(CAG 20/41)来源的成纤维细胞中Htt基因的阻遏。如先前所述进行TALE蛋白质的构建(见共同拥有的美国专利公布20110301073和共同拥有的美国专利申请13/679,684,二者通过引用并入本文),并且以3种不同的C端构造构建:如US20110301073中所述的+63、+231和+278。为构建TALE TF表达质粒,除了用KRAB阻遏结构域置换在TALEN中使用的FokI结构域外,使用先前描述的TALEN表达质粒(见US20110301073)。下面示出了TALE蛋白质C端与KRAB结构域的连接,其中用下划线指示KRAB结构域序列。粗体和斜体文本指示三重flag标签,粗体文本指示核定位序列,“[重复]”指示TALE重复单元阵列的位置(全部重复+C端半重复),并且波浪下划线部分示出了KRAB结构域的序列:

[0278] TALE-C63-Kox1:  
MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPMDLRTLGYSSQQQEKIKPKVRSSTVAQHHEALVGHGF  
THAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDGTGQLLKI  
AKRGGVTAVEAVHAWRNALTGAPLN[repeats]GGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALD  
AVKKGLPHAPALIKRTNRRIPERTSHRVAGSGMDAKSLTAWSRTLVTFKDVFDFTREEWKLLDTAQOIVYRNVM  
LENYKNLVS LGYQLTKPDVILRLEKGEEPWLVEREIHQETHPDSETAFEIKSSV (SEQ ID NO:197)

[0279] TALE-C231-Kox1:  
MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPMDLRTLGYSSQQQEKIKPKVRSSTVAQHHEALVGHGF  
THAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDGTGQLLKI  
AKRGGVTAVEAVHAWRNALTGAPLN[repeats]GGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALD  
AVKKGLPHAPALIKRTNRRIPERTSHRVADHAQVVRVLGFFQCHSHPAQAFDDAMTQFGMSRHLLQLFRRVGV  
ELEARSGTLPASQRWDRI LQASGMKRAKPSPTSTQTPDQASLHAFADSLERDLDPSPTHEGDQRASSRKRSR  
SDRAVTGPSAQQSFEVRAPEQRDALHLPLSWRVKRPRTSIGGGLPDGSGMDAKSLTAWSRTLVTFKDVFDFTRE  
EWKLLDTAQOIVYRNVMLENYKNLVS LGYQLTKPDVILRLEKGEEPWLVEREIHQETHPDSETAFEIKSSV  
(SEQ ID NO:198)

TALE-C278-Kox1  
MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPMDLRTLGYSSQQQEKIKPKVRSSTVAQHHEALVGHGF  
THAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDGTGQLLKI  
AKRGGVTAVEAVHAWRNALTGAPLN[repeats]GGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALD  
AVKKGLPHAPALIKRTNRRIPERTSHRVADHAQVVRVLGFFQCHSHPAQAFDDAMTQFGMSRHLLQLFRRVGV  
ELEARSGTLPASQRWDRI LQASGMKRAKPSPTSTQTPDQASLHAFADSLERDLDPSPTHEGDQRASSRKRSR  
SDRAVTGPSAQQSFEVRAPEQRDALHLPLSWRVKRPRTSIGGGLPDPTAADLAASSTVMREQDEDPFAGAADD  
FPAFNEEELAWLMELLPGSGMDAKSLTAWSRTLVTFKDVFDFTREEWKLLDTAQOIVYRNVMLENYKNLVS LG  
YQLTKPDVILRLEKGEEPWLVEREIHQETHPDSETAFEIKSSV (SEQ ID NO:199)

[0280] 使用规范的RVD碱基对应关系(“TALE码”:对于A为NI,对于C为HD,对于G为NN(半重复中为NK),对于T为NG)实现碱基识别。在一些TALE TF中,蛋白质设计为结合DNA的有义(5’-3’)链,而在其它中,TALE TF设计为与反义(3’-5’)链结合。这组TALE TF设计为靶向Htt基因的CAG重复。TALE DNA结合蛋白常常优先与靶标5’末端的‘T’核苷酸碱基相互作用,所以因为靶标为CAG重复区,可以预计蛋白质与反义DNA链结合,并且因此在靶标5’3末端具有碱基‘T’的CTG重复序列可能具有更好的结合亲和力和特异性并且因此具有阻遏因子活性。

[0281] 下面表4中示出了试验的TALE TF的靶标和数字标识符。数字标识符标记“SBS#”,指示了对有义或反义链的特异性(“S/A”),以及靶标、重复单元或RVD的数量及C端类型。

[0282] 表4:Htt特异性TALE-TF

SBS#	S/A	靶标 (5'-3')	SEQ ID NO	RVD	C 端
102449	S	gcAGCAGCAGCAGCAGCAGca	200	17	+63
102450	S	gcAGCAGCAGCAGCAGca	201	14	+63
102451	S	gcAGCAGCAGCAGca	202	11	+63
102452	S	gcAGCAGCAGca	203	8	+63
102453	A	ctGCTGCTGCTGCTGCTGctg	204	17	+63
102454	A	ctGCTGCTGCTGCTGctg	205	14	+63
102455	A	ctGCTGCTGCTGctg	206	11	+63
102456	A	ctGCTGCTGctg	207	8	+63
102457	S	gcAGCAGCAGCAGCAGCAGca	200	17	+231
102458	S	gcAGCAGCAGCAGCAGca	201	14	+231
102459	S	gcAGCAGCAGCAGca	202	11	+231
102460	S	gcAGCAGCAGca	203	8	+231
102462	A	ctGCTGCTGCTGCTGCTGctg	205	14	+231
102463	A	ctGCTGCTGCTGctg	206	11	+231
102464	A	ctGCTGCTGctg	207	8	+231
102466	S	gcAGCAGCAGCAGCAGca	201	14	+278
102467	S	gcAGCAGCAGCAGca	202	11	+278
102468	S	gcAGCAGCAGca	203	8	+278
102469	A	ctGCTGCTGCTGCTGCTGctg	204	17	+278
102470	A	ctGCTGCTGCTGCTGctg	205	14	+278
102471	A	ctGCTGCTGCTGctg	206	11	+278
102472	A	ctGCTGCTGctg	207	8	+278

[0284] 然后测试表中TALE TF在HD患者(CAG 20/41)来源的成纤维细胞中的Htt阻遏,并且图15中示出了结果。在该实验中,用1000、100或10ng编码TALE-TF的mRNA转染细胞。3个一组示出了测定的每个TALE TF的结果,表示3个转染mRNA量。在每个分组中,还有3个样品:左侧柱指示总Htt表达,中间柱指示由CAG20Htt等位基因的表达,而右侧柱指示由CAG41Htt等位基因的表达。数据证明存在一些能够阻遏两个Htt等位基因的TALE TF(见例如102454),而其它TALE TF能够选择性抑制具有延伸CAG重复的突变Htt(见例如102451和102472)。

[0285] 本文提到的所有专利、专利申请和出版物据此通过引用整体并入。

[0286] 虽然为了理解清楚的目的,已经通过举例说明和实施例的方式相当详细地提供了公开内容,但是对于本领域的技术人员显而易见的是,在不背离公开内容的精神或范围的前提下,可实践各种变化和修改。相应地,不得将前面的描述和实施例视为限制。

- [0001] 序列表
- [0002] <110> 桑格摩生物科学股份有限公司
- [0003] (SANGAMO BIOSCIENCES, INC.)
- [0004] <120> 治疗亨廷顿氏病的方法和组合物
- [0005] <130> 8325-0095.40
- [0006] <140> PCT/US2013/028348
- [0007] <141> 2013-02-28
- [0008] <150> 61/605,028
- [0009] <151> 2012-02-29
- [0010] <160> 261
- [0011] <170> PatentIn 3.5 版
- [0012] <210> 1
- [0013] <211> 7
- [0014] <212> PRT
- [0015] <213> 人工序列
- [0016] <220>
- [0017] <223> 人工序列的描述:合成肽
- [0018] <400> 1
- [0019] Arg Ser Asp Asp Leu Ser Arg
- [0020] 1 5
- [0021] <210> 2
- [0022] <211> 7
- [0023] <212> PRT
- [0024] <213> 人工序列
- [0025] <220>
- [0026] <223> 人工序列的描述:合成肽
- [0027] <400> 2
- [0028] Arg Asn Asp Asn Arg Thr Lys
- [0029] 1 5
- [0030] <210> 3
- [0031] <211> 7
- [0032] <212> PRT
- [0033] <213> 人工序列
- [0034] <220>
- [0035] <223> 人工序列的描述:合成肽
- [0036] <400> 3
- [0037] Arg Ser Asp Asp Leu Thr Arg
- [0038] 1 5

- [0039] <210> 4  
[0040] <211> 7  
[0041] <212> PRT  
[0042] <213> 人工序列  
[0043] <220>  
[0044] <223> 人工序列的描述:合成肽  
[0045] <400> 4  
[0046] Arg Ser Asp Asp Arg Lys Thr  
[0047] 1 5  
[0048] <210> 5  
[0049] <211> 7  
[0050] <212> PRT  
[0051] <213> 人工序列  
[0052] <220>  
[0053] <223> 人工序列的描述:合成肽  
[0054] <400> 5  
[0055] Arg Ser Ala Asp Leu Thr Arg  
[0056] 1 5  
[0057] <210> 6  
[0058] <211> 7  
[0059] <212> PRT  
[0060] <213> 人工序列  
[0061] <220>  
[0062] <223> 人工序列的描述:合成肽  
[0063] <400> 6  
[0064] Gln Ser Ser Asp Leu Arg Arg  
[0065] 1 5  
[0066] <210> 7  
[0067] <211> 7  
[0068] <212> PRT  
[0069] <213> 人工序列  
[0070] <220>  
[0071] <223> 人工序列的描述:合成肽  
[0072] <400> 7  
[0073] Arg Ser Asp Asn Leu Ala Arg  
[0074] 1 5  
[0075] <210> 8  
[0076] <211> 7  
[0077] <212> PRT

- [0078] <213> 人工序列  
[0079] <220>  
[0080] <223> 人工序列的描述:合成肽  
[0081] <400> 8  
[0082] Trp Arg Gly Asp Arg Val Lys  
[0083] 1 5  
[0084] <210> 9  
[0085] <211> 7  
[0086] <212> PRT  
[0087] <213> 人工序列  
[0088] <220>  
[0089] <223> 人工序列的描述:合成肽  
[0090] <400> 9  
[0091] Asp Arg Ser Asn Leu Ser Arg  
[0092] 1 5  
[0093] <210> 10  
[0094] <211> 7  
[0095] <212> PRT  
[0096] <213> 人工序列  
[0097] <220>  
[0098] <223> 人工序列的描述:合成肽  
[0099] <400> 10  
[0100] Thr Ser Gly Ser Leu Thr Arg  
[0101] 1 5  
[0102] <210> 11  
[0103] <211> 7  
[0104] <212> PRT  
[0105] <213> 人工序列  
[0106] <220>  
[0107] <223> 人工序列的描述:合成肽  
[0108] <400> 11  
[0109] Glu Arg Gly Thr Leu Ala Arg  
[0110] 1 5  
[0111] <210> 12  
[0112] <211> 7  
[0113] <212> PRT  
[0114] <213> 人工序列  
[0115] <220>  
[0116] <223> 人工序列的描述:合成肽

- [0117] <400> 12  
[0118] Arg Ser Asp Ala Leu Ser Arg  
[0119] 1 5  
[0120] <210> 13  
[0121] <211> 7  
[0122] <212> PRT  
[0123] <213> 人工序列  
[0124] <220>  
[0125] <223> 人工序列的描述:合成肽  
[0126] <400> 13  
[0127] Asp Arg Ser His Leu Ala Arg  
[0128] 1 5  
[0129] <210> 14  
[0130] <211> 7  
[0131] <212> PRT  
[0132] <213> 人工序列  
[0133] <220>  
[0134] <223> 人工序列的描述:合成肽  
[0135] <400> 14  
[0136] Arg Ser Asp His Leu Ser Arg  
[0137] 1 5  
[0138] <210> 15  
[0139] <211> 7  
[0140] <212> PRT  
[0141] <213> 人工序列  
[0142] <220>  
[0143] <223> 人工序列的描述:合成肽  
[0144] <400> 15  
[0145] Gln Ser Ser Asp Leu Thr Arg  
[0146] 1 5  
[0147] <210> 16  
[0148] <211> 7  
[0149] <212> PRT  
[0150] <213> 人工序列  
[0151] <220>  
[0152] <223> 人工序列的描述:合成肽  
[0153] <400> 16  
[0154] Thr Ser Gly Asn Leu Thr Arg  
[0155] 1 5

- [0156] <210> 17  
[0157] <211> 7  
[0158] <212> PRT  
[0159] <213> 人工序列  
[0160] <220>  
[0161] <223> 人工序列的描述:合成肽  
[0162] <400> 17  
[0163] Gln Ser Gly His Leu Gln Arg  
[0164] 1 5  
[0165] <210> 18  
[0166] <211> 7  
[0167] <212> PRT  
[0168] <213> 人工序列  
[0169] <220>  
[0170] <223> 人工序列的描述:合成肽  
[0171] <400> 18  
[0172] Gln Ser Gly Asp Leu Thr Arg  
[0173] 1 5  
[0174] <210> 19  
[0175] <211> 7  
[0176] <212> PRT  
[0177] <213> 人工序列  
[0178] <220>  
[0179] <223> 人工序列的描述:合成肽  
[0180] <400> 19  
[0181] Arg Ser Asp Val Leu Ser Thr  
[0182] 1 5  
[0183] <210> 20  
[0184] <211> 7  
[0185] <212> PRT  
[0186] <213> 人工序列  
[0187] <220>  
[0188] <223> 人工序列的描述:合成肽  
[0189] <400> 20  
[0190] Val Arg Ser Arg Leu Arg Arg  
[0191] 1 5  
[0192] <210> 21  
[0193] <211> 7  
[0194] <212> PRT

- [0195] <213> 人工序列  
[0196] <220>  
[0197] <223> 人工序列的描述:合成肽  
[0198] <400> 21  
[0199] Asp Arg Ser Thr Leu Arg Gln  
[0200] 1 5  
[0201] <210> 22  
[0202] <211> 7  
[0203] <212> PRT  
[0204] <213> 人工序列  
[0205] <220>  
[0206] <223> 人工序列的描述:合成肽  
[0207] <400> 22  
[0208] Asp Arg Ser Asp Leu Ser Arg  
[0209] 1 5  
[0210] <210> 23  
[0211] <211> 7  
[0212] <212> PRT  
[0213] <213> 人工序列  
[0214] <220>  
[0215] <223> 人工序列的描述:合成肽  
[0216] <400> 23  
[0217] Gln Ser Ser Thr Arg Ala Arg  
[0218] 1 5  
[0219] <210> 24  
[0220] <211> 7  
[0221] <212> PRT  
[0222] <213> 人工序列  
[0223] <220>  
[0224] <223> 人工序列的描述:合成肽  
[0225] <400> 24  
[0226] Arg Ser Asp Thr Leu Ser Glu  
[0227] 1 5  
[0228] <210> 25  
[0229] <211> 7  
[0230] <212> PRT  
[0231] <213> 人工序列  
[0232] <220>  
[0233] <223> 人工序列的描述:合成肽

- [0234] <400> 25  
[0235] His Arg Arg Ser Arg Trp Gly  
[0236] 1 5  
[0237] <210> 26  
[0238] <211> 7  
[0239] <212> PRT  
[0240] <213> 人工序列  
[0241] <220>  
[0242] <223> 人工序列的描述:合成肽  
[0243] <400> 26  
[0244] Arg Arg Asp Thr Leu Arg Ser  
[0245] 1 5  
[0246] <210> 27  
[0247] <211> 7  
[0248] <212> PRT  
[0249] <213> 人工序列  
[0250] <220>  
[0251] <223> 人工序列的描述:合成肽  
[0252] <400> 27  
[0253] Arg Ser Asp His Leu Ser Thr  
[0254] 1 5  
[0255] <210> 28  
[0256] <211> 7  
[0257] <212> PRT  
[0258] <213> 人工序列  
[0259] <220>  
[0260] <223> 人工序列的描述:合成肽  
[0261] <400> 28  
[0262] Gln Ser Ala His Arg Ile Thr  
[0263] 1 5  
[0264] <210> 29  
[0265] <211> 7  
[0266] <212> PRT  
[0267] <213> 人工序列  
[0268] <220>  
[0269] <223> 人工序列的描述:合成肽  
[0270] <400> 29  
[0271] Arg Ser Asp His Leu Ser Glu  
[0272] 1 5

- [0273] <210> 30  
[0274] <211> 7  
[0275] <212> PRT  
[0276] <213> 人工序列  
[0277] <220>  
[0278] <223> 人工序列的描述:合成肽  
[0279] <400> 30  
[0280] Gln Asn Ala His Arg Lys Thr  
[0281] 1 5  
[0282] <210> 31  
[0283] <211> 7  
[0284] <212> PRT  
[0285] <213> 人工序列  
[0286] <220>  
[0287] <223> 人工序列的描述:合成肽  
[0288] <400> 31  
[0289] Gln Ser Ser Asp Leu Ser Arg  
[0290] 1 5  
[0291] <210> 32  
[0292] <211> 7  
[0293] <212> PRT  
[0294] <213> 人工序列  
[0295] <220>  
[0296] <223> 人工序列的描述:合成肽  
[0297] <400> 32  
[0298] His Arg Ser Thr Arg Asn Arg  
[0299] 1 5  
[0300] <210> 33  
[0301] <211> 7  
[0302] <212> PRT  
[0303] <213> 人工序列  
[0304] <220>  
[0305] <223> 人工序列的描述:合成肽  
[0306] <400> 33  
[0307] Leu Arg Gln Asp Leu Lys Arg  
[0308] 1 5  
[0309] <210> 34  
[0310] <211> 7  
[0311] <212> PRT

- [0312] <213> 人工序列  
[0313] <220>  
[0314] <223> 人工序列的描述:合成肽  
[0315] <400> 34  
[0316] Asp Arg Ser His Leu Thr Arg  
[0317] 1 5  
[0318] <210> 35  
[0319] <211> 7  
[0320] <212> PRT  
[0321] <213> 人工序列  
[0322] <220>  
[0323] <223> 人工序列的描述:合成肽  
[0324] <400> 35  
[0325] Asp Arg Ser Asn Leu Thr Arg  
[0326] 1 5  
[0327] <210> 36  
[0328] <211> 7  
[0329] <212> PRT  
[0330] <213> 人工序列  
[0331] <220>  
[0332] <223> 人工序列的描述:合成肽  
[0333] <400> 36  
[0334] Leu Lys Gln Met Leu Ala Val  
[0335] 1 5  
[0336] <210> 37  
[0337] <211> 7  
[0338] <212> PRT  
[0339] <213> 人工序列  
[0340] <220>  
[0341] <223> 人工序列的描述:合成肽  
[0342] <400> 37  
[0343] Arg Ser Asp Ser Leu Ser Ala  
[0344] 1 5  
[0345] <210> 38  
[0346] <211> 7  
[0347] <212> PRT  
[0348] <213> 人工序列  
[0349] <220>  
[0350] <223> 人工序列的描述:合成肽

- [0351] <400> 38  
[0352] Arg Ser Asp Ala Leu Ser Thr  
[0353] 1 5  
[0354] <210> 39  
[0355] <211> 7  
[0356] <212> PRT  
[0357] <213> 人工序列  
[0358] <220>  
[0359] <223> 人工序列的描述:合成肽  
[0360] <400> 39  
[0361] Asp Arg Ser Thr Arg Thr Lys  
[0362] 1 5  
[0363] <210> 40  
[0364] <211> 7  
[0365] <212> PRT  
[0366] <213> 人工序列  
[0367] <220>  
[0368] <223> 人工序列的描述:合成肽  
[0369] <400> 40  
[0370] Asp Arg Ser Ala Leu Ala Arg  
[0371] 1 5  
[0372] <210> 41  
[0373] <211> 7  
[0374] <212> PRT  
[0375] <213> 人工序列  
[0376] <220>  
[0377] <223> 人工序列的描述:合成肽  
[0378] <400> 41  
[0379] Gln Ser Gly His Leu Ser Arg  
[0380] 1 5  
[0381] <210> 42  
[0382] <211> 7  
[0383] <212> PRT  
[0384] <213> 人工序列  
[0385] <220>  
[0386] <223> 人工序列的描述:合成肽  
[0387] <400> 42  
[0388] Arg Ser Asp Val Leu Ser Glu  
[0389] 1 5

- [0390] <210> 43  
[0391] <211> 7  
[0392] <212> PRT  
[0393] <213> 人工序列  
[0394] <220>  
[0395] <223> 人工序列的描述:合成肽  
[0396] <400> 43  
[0397] Thr Ser Gly His Leu Ser Arg  
[0398] 1 5  
[0399] <210> 44  
[0400] <211> 7  
[0401] <212> PRT  
[0402] <213> 人工序列  
[0403] <220>  
[0404] <223> 人工序列的描述:合成肽  
[0405] <400> 44  
[0406] Lys Leu Cys Asn Arg Lys Cys  
[0407] 1 5  
[0408] <210> 45  
[0409] <211> 7  
[0410] <212> PRT  
[0411] <213> 人工序列  
[0412] <220>  
[0413] <223> 人工序列的描述:合成肽  
[0414] <400> 45  
[0415] His Arg Thr Ser Leu Thr Asp  
[0416] 1 5  
[0417] <210> 46  
[0418] <211> 7  
[0419] <212> PRT  
[0420] <213> 人工序列  
[0421] <220>  
[0422] <223> 人工序列的描述:合成肽  
[0423] <400> 46  
[0424] Arg Ser Ala His Leu Ser Arg  
[0425] 1 5  
[0426] <210> 47  
[0427] <211> 7  
[0428] <212> PRT

- [0429] <213> 人工序列  
[0430] <220>  
[0431] <223> 人工序列的描述:合成肽  
[0432] <400> 47  
[0433] Gln Ser Gly Asn Leu Ala Arg  
[0434] 1 5  
[0435] <210> 48  
[0436] <211> 7  
[0437] <212> PRT  
[0438] <213> 人工序列  
[0439] <220>  
[0440] <223> 人工序列的描述:合成肽  
[0441] <400> 48  
[0442] Leu Lys His His Leu Thr Asp  
[0443] 1 5  
[0444] <210> 49  
[0445] <211> 7  
[0446] <212> PRT  
[0447] <213> 人工序列  
[0448] <220>  
[0449] <223> 人工序列的描述:合成肽  
[0450] <400> 49  
[0451] Tyr Arg Trp Leu Arg Asn Asn  
[0452] 1 5  
[0453] <210> 50  
[0454] <211> 7  
[0455] <212> PRT  
[0456] <213> 人工序列  
[0457] <220>  
[0458] <223> 人工序列的描述:合成肽  
[0459] <400> 50  
[0460] Arg Ser Asp His Leu Ser Gln  
[0461] 1 5  
[0462] <210> 51  
[0463] <211> 7  
[0464] <212> PRT  
[0465] <213> 人工序列  
[0466] <220>  
[0467] <223> 人工序列的描述:合成肽

- [0468] <400> 51  
[0469] Arg Ser Ala Val Arg Lys Asn  
[0470] 1 5  
[0471] <210> 52  
[0472] <211> 7  
[0473] <212> PRT  
[0474] <213> 人工序列  
[0475] <220>  
[0476] <223> 人工序列的描述:合成肽  
[0477] <400> 52  
[0478] Trp Ser Thr Ser Leu Arg Ala  
[0479] 1 5  
[0480] <210> 53  
[0481] <211> 7  
[0482] <212> PRT  
[0483] <213> 人工序列  
[0484] <220>  
[0485] <223> 人工序列的描述:合成肽  
[0486] <400> 53  
[0487] Gln Arg Thr His Leu Thr Gln  
[0488] 1 5  
[0489] <210> 54  
[0490] <211> 7  
[0491] <212> PRT  
[0492] <213> 人工序列  
[0493] <220>  
[0494] <223> 人工序列的描述:合成肽  
[0495] <400> 54  
[0496] Arg Ser Ser His Leu Ser Arg  
[0497] 1 5  
[0498] <210> 55  
[0499] <211> 7  
[0500] <212> PRT  
[0501] <213> 人工序列  
[0502] <220>  
[0503] <223> 人工序列的描述:合成肽  
[0504] <400> 55  
[0505] Thr Ser Gly Ser Leu Ser Arg  
[0506] 1 5

- [0507] <210> 56  
[0508] <211> 7  
[0509] <212> PRT  
[0510] <213> 人工序列  
[0511] <220>  
[0512] <223> 人工序列的描述:合成肽  
[0513] <400> 56  
[0514] Thr Arg Gln Asn Arg Asp Thr  
[0515] 1 5  
[0516] <210> 57  
[0517] <211> 7  
[0518] <212> PRT  
[0519] <213> 人工序列  
[0520] <220>  
[0521] <223> 人工序列的描述:合成肽  
[0522] <400> 57  
[0523] Asp Gln Ser Thr Leu Arg Asn  
[0524] 1 5  
[0525] <210> 58  
[0526] <211> 7  
[0527] <212> PRT  
[0528] <213> 人工序列  
[0529] <220>  
[0530] <223> 人工序列的描述:合成肽  
[0531] <400> 58  
[0532] Arg Ser Ala Ala Leu Ser Arg  
[0533] 1 5  
[0534] <210> 59  
[0535] <211> 7  
[0536] <212> PRT  
[0537] <213> 人工序列  
[0538] <220>  
[0539] <223> 人工序列的描述:合成肽  
[0540] <400> 59  
[0541] Arg Ser Asp Ala Leu Ala Arg  
[0542] 1 5  
[0543] <210> 60  
[0544] <211> 7  
[0545] <212> PRT

- [0546] <213> 人工序列  
[0547] <220>  
[0548] <223> 人工序列的描述:合成肽  
[0549] <400> 60  
[0550] Arg Ser Asp Asn Leu Ser Glu  
[0551] 1 5  
[0552] <210> 61  
[0553] <211> 7  
[0554] <212> PRT  
[0555] <213> 人工序列  
[0556] <220>  
[0557] <223> 人工序列的描述:合成肽  
[0558] <400> 61  
[0559] Lys Arg Cys Asn Leu Arg Cys  
[0560] 1 5  
[0561] <210> 62  
[0562] <211> 7  
[0563] <212> PRT  
[0564] <213> 人工序列  
[0565] <220>  
[0566] <223> 人工序列的描述:合成肽  
[0567] <400> 62  
[0568] Trp Arg Ser Cys Arg Ser Ala  
[0569] 1 5  
[0570] <210> 63  
[0571] <211> 7  
[0572] <212> PRT  
[0573] <213> 人工序列  
[0574] <220>  
[0575] <223> 人工序列的描述:合成肽  
[0576] <400> 63  
[0577] Gln Trp Ser Thr Arg Lys Arg  
[0578] 1 5  
[0579] <210> 64  
[0580] <211> 7  
[0581] <212> PRT  
[0582] <213> 人工序列  
[0583] <220>  
[0584] <223> 人工序列的描述:合成肽

- [0585] <400> 64  
[0586] Arg Arg Trp Thr Leu Val Gly  
[0587] 1 5  
[0588] <210> 65  
[0589] <211> 7  
[0590] <212> PRT  
[0591] <213> 人工序列  
[0592] <220>  
[0593] <223> 人工序列的描述:合成肽  
[0594] <400> 65  
[0595] Arg Ser Asp Asn Leu Arg Glu  
[0596] 1 5  
[0597] <210> 66  
[0598] <211> 7  
[0599] <212> PRT  
[0600] <213> 人工序列  
[0601] <220>  
[0602] <223> 人工序列的描述:合成肽  
[0603] <400> 66  
[0604] Tyr Arg Trp Leu Leu Arg Ser  
[0605] 1 5  
[0606] <210> 67  
[0607] <211> 7  
[0608] <212> PRT  
[0609] <213> 人工序列  
[0610] <220>  
[0611] <223> 人工序列的描述:合成肽  
[0612] <400> 67  
[0613] Arg Lys Asp Ala Leu Val Ala  
[0614] 1 5  
[0615] <210> 68  
[0616] <211> 7  
[0617] <212> PRT  
[0618] <213> 人工序列  
[0619] <220>  
[0620] <223> 人工序列的描述:合成肽  
[0621] <400> 68  
[0622] Arg Arg Ala Asp Leu Ser Arg  
[0623] 1 5

- [0624] <210> 69  
[0625] <211> 7  
[0626] <212> PRT  
[0627] <213> 人工序列  
[0628] <220>  
[0629] <223> 人工序列的描述:合成肽  
[0630] <400> 69  
[0631] Leu Ala Gln Pro Arg Asn Lys  
[0632] 1 5  
[0633] <210> 70  
[0634] <211> 7  
[0635] <212> PRT  
[0636] <213> 人工序列  
[0637] <220>  
[0638] <223> 人工序列的描述:合成肽  
[0639] <400> 70  
[0640] Arg Ser Asp His Leu Ser Ala  
[0641] 1 5  
[0642] <210> 71  
[0643] <211> 7  
[0644] <212> PRT  
[0645] <213> 人工序列  
[0646] <220>  
[0647] <223> 人工序列的描述:合成肽  
[0648] <400> 71  
[0649] Arg Ser Ala Asp Arg Thr Arg  
[0650] 1 5  
[0651] <210> 72  
[0652] <211> 7  
[0653] <212> PRT  
[0654] <213> 人工序列  
[0655] <220>  
[0656] <223> 人工序列的描述:合成肽  
[0657] <400> 72  
[0658] Thr Ser Ser Asp Arg Lys Lys  
[0659] 1 5  
[0660] <210> 73  
[0661] <211> 7  
[0662] <212> PRT

- [0663] <213> 人工序列  
[0664] <220>  
[0665] <223> 人工序列的描述:合成肽  
[0666] <400> 73  
[0667] Arg Ser Ala Asn Leu Ala Arg  
[0668] 1 5  
[0669] <210> 74  
[0670] <211> 7  
[0671] <212> PRT  
[0672] <213> 人工序列  
[0673] <220>  
[0674] <223> 人工序列的描述:合成肽  
[0675] <400> 74  
[0676] His His Ser Ala Arg Arg Cys  
[0677] 1 5  
[0678] <210> 75  
[0679] <211> 7  
[0680] <212> PRT  
[0681] <213> 人工序列  
[0682] <220>  
[0683] <223> 人工序列的描述:合成肽  
[0684] <400> 75  
[0685] Asp Asn Ser Ser Arg Thr Arg  
[0686] 1 5  
[0687] <210> 76  
[0688] <211> 7  
[0689] <212> PRT  
[0690] <213> 人工序列  
[0691] <220>  
[0692] <223> 人工序列的描述:合成肽  
[0693] <400> 76  
[0694] His Lys Gln His Arg Asp Ala  
[0695] 1 5  
[0696] <210> 77  
[0697] <211> 7  
[0698] <212> PRT  
[0699] <213> 人工序列  
[0700] <220>  
[0701] <223> 人工序列的描述:合成肽

- [0702] <400> 77  
[0703] Arg Arg Thr Asp Leu Arg Arg  
[0704] 1 5  
[0705] <210> 78  
[0706] <211> 7  
[0707] <212> PRT  
[0708] <213> 人工序列  
[0709] <220>  
[0710] <223> 人工序列的描述:合成肽  
[0711] <400> 78  
[0712] Asp Pro Ser Asn Arg Val Gly  
[0713] 1 5  
[0714] <210> 79  
[0715] <211> 7  
[0716] <212> PRT  
[0717] <213> 人工序列  
[0718] <220>  
[0719] <223> 人工序列的描述:合成肽  
[0720] <400> 79  
[0721] Arg Ser Ala Asp Leu Ser Arg  
[0722] 1 5  
[0723] <210> 80  
[0724] <211> 7  
[0725] <212> PRT  
[0726] <213> 人工序列  
[0727] <220>  
[0728] <223> 人工序列的描述:合成肽  
[0729] <400> 80  
[0730] Arg Asn Asp Asp Arg Lys Lys  
[0731] 1 5  
[0732] <210> 81  
[0733] <211> 7  
[0734] <212> PRT  
[0735] <213> 人工序列  
[0736] <220>  
[0737] <223> 人工序列的描述:合成肽  
[0738] <400> 81  
[0739] Gln Lys Val Thr Leu Ala Ala  
[0740] 1 5

- [0741] <210> 82  
[0742] <211> 7  
[0743] <212> PRT  
[0744] <213> 人工序列  
[0745] <220>  
[0746] <223> 人工序列的描述:合成肽  
[0747] <400> 82  
[0748] Arg Arg Ser Thr Leu Arg Ser  
[0749] 1 5  
[0750] <210> 83  
[0751] <211> 7  
[0752] <212> PRT  
[0753] <213> 人工序列  
[0754] <220>  
[0755] <223> 人工序列的描述:合成肽  
[0756] <400> 83  
[0757] Asp Arg Ser Ala Leu Ser Arg  
[0758] 1 5  
[0759] <210> 84  
[0760] <211> 7  
[0761] <212> PRT  
[0762] <213> 人工序列  
[0763] <220>  
[0764] <223> 人工序列的描述:合成肽  
[0765] <400> 84  
[0766] Gln Ser Gly Ser Leu Thr Arg  
[0767] 1 5  
[0768] <210> 85  
[0769] <211> 7  
[0770] <212> PRT  
[0771] <213> 人工序列  
[0772] <220>  
[0773] <223> 人工序列的描述:合成肽  
[0774] <400> 85  
[0775] Leu Lys Trp Asn Leu Arg Thr  
[0776] 1 5  
[0777] <210> 86  
[0778] <211> 7  
[0779] <212> PRT

- [0780] <213> 人工序列  
[0781] <220>  
[0782] <223> 人工序列的描述:合成肽  
[0783] <400> 86  
[0784] Leu Lys Trp Asp Arg Gln Thr  
[0785] 1 5  
[0786] <210> 87  
[0787] <211> 7  
[0788] <212> PRT  
[0789] <213> 人工序列  
[0790] <220>  
[0791] <223> 人工序列的描述:合成肽  
[0792] <400> 87  
[0793] Asp Arg Ser His Leu Ser Arg  
[0794] 1 5  
[0795] <210> 88  
[0796] <211> 28  
[0797] <212> DNA  
[0798] <213> 未知  
[0799] <220>  
[0800] <223> 未知序列的描述:人或小鼠靶序列  
[0801] <400> 88  
[0802] acgctgcgcc ggcggaggcg gggccgcg 28  
[0803] <210> 89  
[0804] <211> 28  
[0805] <212> DNA  
[0806] <213> 未知  
[0807] <220>  
[0808] <223> 未知序列的描述:人或小鼠靶序列  
[0809] <400> 89  
[0810] agccggccgt ggactctgag ccgaggtg 28  
[0811] <210> 90  
[0812] <211> 28  
[0813] <212> DNA  
[0814] <213> 未知  
[0815] <220>  
[0816] <223> 未知序列的描述:人或小鼠靶序列  
[0817] <400> 90  
[0818] gtggc gatgc ggggggcgtg gtgaggta 28

[0819] <210> 91  
[0820] <211> 28  
[0821] <212> DNA  
[0822] <213> 未知  
[0823] <220>  
[0824] <223> 未知序列的描述:人或小鼠靶序列  
[0825] <400> 91  
[0826] ccgggacggg tccaagatgg acgcccgc 28  
[0827] <210> 92  
[0828] <211> 28  
[0829] <212> DNA  
[0830] <213> 未知  
[0831] <220>  
[0832] <223> 未知序列的描述:人或小鼠靶序列  
[0833] <400> 92  
[0834] ccgtcccggc agccccacg gcgccttg 28  
[0835] <210> 93  
[0836] <211> 28  
[0837] <212> DNA  
[0838] <213> 未知  
[0839] <220>  
[0840] <223> 未知序列的描述:人或小鼠靶序列  
[0841] <400> 93  
[0842] cgggtccaag atggacggcc gctcaggt 28  
[0843] <210> 94  
[0844] <211> 28  
[0845] <212> DNA  
[0846] <213> 未知  
[0847] <220>  
[0848] <223> 未知序列的描述:人或小鼠靶序列  
[0849] <400> 94  
[0850] ctgctgctgc tgctggaagg acttgagg 28  
[0851] <210> 95  
[0852] <211> 28  
[0853] <212> DNA  
[0854] <213> 未知  
[0855] <220>  
[0856] <223> 未知序列的描述:人或小鼠靶序列  
[0857] <400> 95

[0858] tcagatggga cggcgctgac ctggctgg 28  
[0859] <210> 96  
[0860] <211> 28  
[0861] <212> DNA  
[0862] <213> 未知  
[0863] <220>  
[0864] <223> 未知序列的描述:人或小鼠靶序列  
[0865] <400> 96  
[0866] ctgccatgga cctgaatgat gggaccca 28  
[0867] <210> 97  
[0868] <211> 28  
[0869] <212> DNA  
[0870] <213> 未知  
[0871] <220>  
[0872] <223> 未知序列的描述:人或小鼠靶序列  
[0873] <400> 97  
[0874] gtggtctggg agctgtcgct gatgggcg 28  
[0875] <210> 98  
[0876] <211> 28  
[0877] <212> DNA  
[0878] <213> 未知  
[0879] <220>  
[0880] <223> 未知序列的描述:人或小鼠靶序列  
[0881] <400> 98  
[0882] ccgaagggcc tgattcagct gttacccc 28  
[0883] <210> 99  
[0884] <211> 28  
[0885] <212> DNA  
[0886] <213> 未知  
[0887] <220>  
[0888] <223> 未知序列的描述:人或小鼠靶序列  
[0889] <400> 99  
[0890] aacttgcaag taacagaaga ctcatcct 28  
[0891] <210> 100  
[0892] <211> 28  
[0893] <212> DNA  
[0894] <213> 未知  
[0895] <220>  
[0896] <223> 未知序列的描述:人或小鼠靶序列

[0897] <400> 100  
[0898] cttgtacagc tgtgagggtg agcataat 28  
[0899] <210> 101  
[0900] <211> 28  
[0901] <212> DNA  
[0902] <213> 未知  
[0903] <220>  
[0904] <223> 未知序列的描述:人或小鼠靶序列  
[0905] <400> 101  
[0906] gccatggtgg gagagactgt gaggcggc 28  
[0907] <210> 102  
[0908] <211> 28  
[0909] <212> DNA  
[0910] <213> 未知  
[0911] <220>  
[0912] <223> 未知序列的描述:人或小鼠靶序列  
[0913] <400> 102  
[0914] ctcagcaggt ggtgacctg tggacatt 28  
[0915] <210> 103  
[0916] <211> 28  
[0917] <212> DNA  
[0918] <213> 未知  
[0919] <220>  
[0920] <223> 未知序列的描述:人或小鼠靶序列  
[0921] <400> 103  
[0922] gcgctcagca ggtggtgacc ttgtggac 28  
[0923] <210> 104  
[0924] <211> 28  
[0925] <212> DNA  
[0926] <213> 未知  
[0927] <220>  
[0928] <223> 未知序列的描述:人或小鼠靶序列  
[0929] <400> 104  
[0930] atggtgggag agactgtgag gcggcagc 28  
[0931] <210> 105  
[0932] <211> 28  
[0933] <212> DNA  
[0934] <213> 未知  
[0935] <220>

- [0936] <223> 未知序列的描述:人或小鼠靶序列  
[0937] <400> 105  
[0938] tgggagagac tgtgaggcgg cagctggg 28  
[0939] <210> 106  
[0940] <211> 28  
[0941] <212> DNA  
[0942] <213> 未知  
[0943] <220>  
[0944] <223> 未知序列的描述:人或小鼠靶序列  
[0945] <400> 106  
[0946] atggcgctca gcaggtggtg accttgtg 28  
[0947] <210> 107  
[0948] <211> 28  
[0949] <212> DNA  
[0950] <213> 未知  
[0951] <220>  
[0952] <223> 未知序列的描述:人或小鼠靶序列  
[0953] <400> 107  
[0954] cagcagcagc agcagcagca gcagcagc 28  
[0955] <210> 108  
[0956] <211> 28  
[0957] <212> DNA  
[0958] <213> 未知  
[0959] <220>  
[0960] <223> 未知序列的描述:人或小鼠靶序列  
[0961] <400> 108  
[0962] ctgctgctgc tgctgctgct ggaaggac 28  
[0963] <210> 109  
[0964] <211> 28  
[0965] <212> DNA  
[0966] <213> 未知  
[0967] <220>  
[0968] <223> 未知序列的描述:人或小鼠靶序列  
[0969] <400> 109  
[0970] tgctgctgct gctgctgctg ctggaagg 28  
[0971] <210> 110  
[0972] <211> 28  
[0973] <212> DNA  
[0974] <213> 未知

- [0975] <220>  
[0976] <223> 未知序列的描述:人或小鼠靶序列  
[0977] <400> 110  
[0978] agcagcagca gcagcagcag cagcagca 28  
[0979] <210> 111  
[0980] <211> 28  
[0981] <212> DNA  
[0982] <213> 未知  
[0983] <220>  
[0984] <223> 未知序列的描述:人或小鼠靶序列  
[0985] <400> 111  
[0986] tgctgctgct gctgctgctg ctggaagg 28  
[0987] <210> 112  
[0988] <211> 28  
[0989] <212> DNA  
[0990] <213> 未知  
[0991] <220>  
[0992] <223> 未知序列的描述:人或小鼠靶序列  
[0993] <400> 112  
[0994] ggctggcttt tgcggaagg ggcggggc 28  
[0995] <210> 113  
[0996] <211> 28  
[0997] <212> DNA  
[0998] <213> 未知  
[0999] <220>  
[1000] <223> 未知序列的描述:人或小鼠靶序列  
[1001] <400> 113  
[1002] gaattgacag gcggatgcgt cgtcctct 28  
[1003] <210> 114  
[1004] <211> 28  
[1005] <212> DNA  
[1006] <213> 未知  
[1007] <220>  
[1008] <223> 未知序列的描述:人或小鼠靶序列  
[1009] <400> 114  
[1010] attctgcggg tctggcgtgg cctcgtct 28  
[1011] <210> 115  
[1012] <211> 28  
[1013] <212> DNA

- [1014] <213> 未知  
[1015] <220>  
[1016] <223> 未知序列的描述:人或小鼠靶序列  
[1017] <400> 115  
[1018] gtgacgtcat gccggcggag acgaggcc 28  
[1019] <210> 116  
[1020] <211> 28  
[1021] <212> DNA  
[1022] <213> 未知  
[1023] <220>  
[1024] <223> 未知序列的描述:人或小鼠靶序列  
[1025] <400> 116  
[1026] gtgcgtcccc tgacgtcatg ccggcgga 28  
[1027] <210> 117  
[1028] <211> 28  
[1029] <212> DNA  
[1030] <213> 未知  
[1031] <220>  
[1032] <223> 未知序列的描述:人或小鼠靶序列  
[1033] <400> 117  
[1034] gccgcgaggg ttgccgggac gggcccaa 28  
[1035] <210> 118  
[1036] <211> 28  
[1037] <212> DNA  
[1038] <213> 未知  
[1039] <220>  
[1040] <223> 未知序列的描述:人或小鼠靶序列  
[1041] <400> 118  
[1042] ccgcgagggt tgccgggacg ggcccaag 28  
[1043] <210> 119  
[1044] <211> 28  
[1045] <212> DNA  
[1046] <213> 未知  
[1047] <220>  
[1048] <223> 未知序列的描述:人或小鼠靶序列  
[1049] <400> 119  
[1050] catcgggcag gaagccgtca tggcaacc 28  
[1051] <210> 120  
[1052] <211> 28

- [1053] <212> DNA
- [1054] <213> 未知
- [1055] <220>
- [1056] <223> 未知序列的描述:人或小鼠靶序列
- [1057] <400> 120
- [1058] tcctgcccga tgggacagac cctgaaga 28
- [1059] <210> 121
- [1060] <211> 28
- [1061] <212> DNA
- [1062] <213> 未知
- [1063] <220>
- [1064] <223> 未知序列的描述:人或小鼠靶序列
- [1065] <400> 121
- [1066] gtactgagca atgctgtagt cagcaatc 28
- [1067] <210> 122
- [1068] <211> 53
- [1069] <212> PRT
- [1070] <213> 人工序列
- [1071] <220>
- [1072] <223> 人工序列的描述:合成多肽
- [1073] <400> 122
- [1074] Thr Lys Cys Val His Cys Gly Ile Val Phe Leu Asp Glu Val Met Tyr
- [1075] 1                    5                    10                    15
- [1076] Ala Leu His Met Ser Cys His Gly Phe Arg Asp Pro Phe Glu Cys Asn
- [1077]                    20                    25                    30
- [1078] Ile Cys Gly Tyr His Ser Gln Asp Arg Tyr Glu Phe Ser Ser His Ile
- [1079]                    35                    40                    45
- [1080] Val Arg Gly Glu His
- [1081]                    50
- [1082] <210> 123
- [1083] <211> 53
- [1084] <212> PRT
- [1085] <213> 人工序列
- [1086] <220>
- [1087] <223> 人工序列的描述:合成多肽
- [1088] <400> 123
- [1089] Phe Lys Cys Glu His Cys Arg Ile Leu Phe Leu Asp His Val Met Phe
- [1090] 1                    5                    10                    15
- [1091] Thr Ile His Met Gly Cys His Gly Phe Arg Asp Pro Phe Lys Cys Asn



[1131]	Asp Arg Tyr Glu Phe Ser Ser His Ile Val Arg Gly Glu His
[1132]	100                          105                          110
[1133]	<210> 126
[1134]	<211> 110
[1135]	<212> PRT
[1136]	<213> 人工序列
[1137]	<220>
[1138]	<223> 人工序列的描述:合成多肽
[1139]	<400> 126
[1140]	Phe Lys Cys Glu His Cys Arg Ile Leu Phe Leu Asp His Val Met Phe
[1141]	1                          5                          10                          15
[1142]	Thr Ile His Met Gly Cys His Gly Phe Arg Asp Pro Phe Lys Cys Asn
[1143]	20                          25                          30
[1144]	Met Cys Gly Glu Lys Cys Asp Gly Pro Val Gly Leu Phe Val His Met
[1145]	35                          40                          45
[1146]	Ala Arg Asn Ala His Gly Glu Lys Pro Phe Tyr Cys Glu His Cys Glu
[1147]	50                          55                          60
[1148]	Ile Thr Phe Arg Asp Val Val Met Tyr Ser Leu His Lys Gly Tyr His
[1149]	65                          70                          75                          80
[1150]	Gly Phe Arg Asp Pro Phe Glu Cys Asn Ile Cys Gly Tyr His Ser Gln
[1151]	85                          90                          95
[1152]	Asp Arg Tyr Glu Phe Ser Ser His Ile Val Arg Gly Glu His
[1153]	100                          105                          110
[1154]	<210> 127
[1155]	<211> 30
[1156]	<212> PRT
[1157]	<213> 人工序列
[1158]	<220>
[1159]	<223> 人工序列的描述:合成多肽
[1160]	<400> 127
[1161]	Ala Gln Leu Glu Lys Glu Leu Gln Ala Leu Glu Lys Lys Leu Ala Gln
[1162]	1                          5                          10                          15
[1163]	Leu Glu Trp Glu Asn Gln Ala Leu Glu Lys Glu Leu Ala Gln
[1164]	20                          25                          30
[1165]	<210> 128
[1166]	<211> 30
[1167]	<212> PRT
[1168]	<213> 人工序列
[1169]	<220>

[1170] <223> 人工序列的描述:合成多肽  
 [1171] <400> 128  
 [1172] Ala Gln Leu Lys Lys Lys Leu Gln Ala Asn Lys Lys Glu Leu Ala Gln  
 [1173] 1                    5                    10                    15  
 [1174] Leu Lys Trp Lys Leu Gln Ala Leu Lys Lys Lys Leu Ala Gln  
 [1175]                    20                    25                    30  
 [1176] <210> 129  
 [1177] <211> 30  
 [1178] <212> PRT  
 [1179] <213> 人工序列  
 [1180] <220>  
 [1181] <223> 人工序列的描述:合成多肽  
 [1182] <400> 129  
 [1183] Glu Gln Leu Glu Lys Lys Leu Gln Ala Leu Glu Lys Lys Leu Ala Gln  
 [1184] 1                    5                    10                    15  
 [1185] Leu Glu Trp Lys Asn Gln Ala Leu Glu Lys Lys Leu Ala Gln  
 [1186]                    20                    25                    30  
 [1187] <210> 130  
 [1188] <211> 29  
 [1189] <212> PRT  
 [1190] <213> 人工序列  
 [1191] <220>  
 [1192] <223> 人工序列的描述:合成肽  
 [1193] <400> 130  
 [1194] Ala Leu Lys Lys Glu Leu Gln Ala Asn Lys Lys Glu Leu Ala Gln Leu  
 [1195] 1                    5                    10                    15  
 [1196] Lys Trp Glu Leu Gln Ala Leu Lys Lys Glu Leu Ala Gln  
 [1197]                    20                    25  
 [1198] <210> 131  
 [1199] <211> 26  
 [1200] <212> PRT  
 [1201] <213> 人工序列  
 [1202] <220>  
 [1203] <223> 人工序列的描述:合成肽  
 [1204] <400> 131  
 [1205] Glu Gln Leu Glu Lys Lys Leu Gln Ala Leu Glu Lys Lys Leu Ala Gln  
 [1206] 1                    5                    10                    15  
 [1207] Leu Glu Trp Lys Asn Gln Ala Leu Glu Lys  
 [1208]                    20                    25

[1209]	<210>	132		
[1210]	<211>	25		
[1211]	<212>	PRT		
[1212]	<213>	人工序列		
[1213]	<220>			
[1214]	<223>	人工序列的描述:合成肽		
[1215]	<400>	132		
[1216]		Glu Leu Gln Ala Asn Lys Lys Glu Leu Ala Gln Leu Lys Trp Glu Leu		
[1217]		1                  5                          10                          15		
[1218]		Gln Ala Leu Lys Lys Glu Leu Ala Gln		
[1219]		20                          25		
[1220]	<210>	133		
[1221]	<211>	23		
[1222]	<212>	PRT		
[1223]	<213>	人工序列		
[1224]	<220>			
[1225]	<223>	人工序列的描述:合成肽		
[1226]	<400>	133		
[1227]		Glu Gln Leu Glu Lys Lys Leu Gln Ala Leu Glu Lys Lys Leu Ala Gln		
[1228]		1                  5                          10                          15		
[1229]		Leu Glu Trp Lys Asn Gln Ala		
[1230]		20		
[1231]	<210>	134		
[1232]	<211>	23		
[1233]	<212>	PRT		
[1234]	<213>	人工序列		
[1235]	<220>			
[1236]	<223>	人工序列的描述:合成肽		
[1237]	<400>	134		
[1238]		Gln Ala Asn Lys Lys Glu Leu Ala Gln Leu Lys Trp Glu Leu Gln Ala		
[1239]		1                  5                          10                          15		
[1240]		Leu Lys Lys Glu Leu Ala Gln		
[1241]		20		
[1242]	<210>	135		
[1243]	<211>	5		
[1244]	<212>	PRT		
[1245]	<213>	人工序列		
[1246]	<220>			
[1247]	<223>	人工序列的描述:合成肽		

- [1248] <400> 135  
[1249] Thr Gly Glu Lys Pro  
[1250] 1 5  
[1251] <210> 136  
[1252] <211> 6  
[1253] <212> PRT  
[1254] <213> 人工序列  
[1255] <220>  
[1256] <223> 人工序列的描述:合成肽  
[1257] <400> 136  
[1258] Thr Gly Gly Gln Arg Pro  
[1259] 1 5  
[1260] <210> 137  
[1261] <211> 5  
[1262] <212> PRT  
[1263] <213> 人工序列  
[1264] <220>  
[1265] <223> 人工序列的描述:合成肽  
[1266] <400> 137  
[1267] Thr Gly Gln Lys Pro  
[1268] 1 5  
[1269] <210> 138  
[1270] <211> 6  
[1271] <212> PRT  
[1272] <213> 人工序列  
[1273] <220>  
[1274] <223> 人工序列的描述:合成肽  
[1275] <400> 138  
[1276] Thr Gly Ser Gln Lys Pro  
[1277] 1 5  
[1278] <210> 139  
[1279] <211> 35  
[1280] <212> DNA  
[1281] <213> 人工序列  
[1282] <220>  
[1283] <223> 人工序列的描述:合成引物  
[1284] <400> 139  
[1285] gaagatctca cttggggtcc tcaggtcgtg ccgac 35  
[1286] <210> 140

- [1287] <211> 41  
[1288] <212> DNA  
[1289] <213> 人工序列  
[1290] <220>  
[1291] <223> 人工序列的描述:合成引物  
[1292] <400> 140  
[1293] gtatccaagc ttcagctttt ccagggtcgc ctaggcggtc t 41  
[1294] <210> 141  
[1295] <211> 33  
[1296] <212> DNA  
[1297] <213> 人工序列  
[1298] <220>  
[1299] <223> 人工序列的描述:合成引物  
[1300] <400> 141  
[1301] gcctaggcga ccctggaaaa gctgatgaag gcc 33  
[1302] <210> 142  
[1303] <211> 37  
[1304] <212> DNA  
[1305] <213> 人工序列  
[1306] <220>  
[1307] <223> 人工序列的描述:合成引物  
[1308] <400> 142  
[1309] gtatccaagc ttgagctgca gcgggcccaa actcag 37  
[1310] <210> 143  
[1311] <211> 17  
[1312] <212> DNA  
[1313] <213> 人工序列  
[1314] <220>  
[1315] <223> 人工序列的描述:合成引物  
[1316] <400> 143  
[1317] agtttggagg gtttctc 17  
[1318] <210> 144  
[1319] <211> 22  
[1320] <212> DNA  
[1321] <213> 人工序列  
[1322] <220>  
[1323] <223> 人工序列的描述:合成引物  
[1324] <400> 144  
[1325] tcgactaaag caggatttca gg 22

- [1326] <210> 145  
[1327] <211> 17  
[1328] <212> DNA  
[1329] <213> 人工序列  
[1330] <220>  
[1331] <223> 人工序列的描述:合成引物  
[1332] <400> 145  
[1333] agtttggagg gtttctt 17  
[1334] <210> 146  
[1335] <211> 18  
[1336] <212> DNA  
[1337] <213> 人工序列  
[1338] <220>  
[1339] <223> 人工序列的描述:合成引物  
[1340] <400> 146  
[1341] agggtttctc cgctcagc 18  
[1342] <210> 147  
[1343] <211> 7  
[1344] <212> PRT  
[1345] <213> 人工序列  
[1346] <220>  
[1347] <223> 人工序列的描述:合成肽  
[1348] <400> 147  
[1349] Gln Ser Gly Ala Leu Ala Arg  
[1350] 1 5  
[1351] <210> 148  
[1352] <211> 7  
[1353] <212> PRT  
[1354] <213> 人工序列  
[1355] <220>  
[1356] <223> 人工序列的描述:合成肽  
[1357] <400> 148  
[1358] Arg Ser Ala Val Leu Ser Glu  
[1359] 1 5  
[1360] <210> 149  
[1361] <211> 7  
[1362] <212> PRT  
[1363] <213> 人工序列  
[1364] <220>

- [1365] <223> 人工序列的描述:合成肽  
[1366] <400> 149  
[1367] Arg Ser Asp Asp Leu Ser Lys  
[1368] 1 5  
[1369] <210> 150  
[1370] <211> 7  
[1371] <212> PRT  
[1372] <213> 人工序列  
[1373] <220>  
[1374] <223> 人工序列的描述:合成肽  
[1375] <400> 150  
[1376] Asp Asn Ser Asn Arg Ile Lys  
[1377] 1 5  
[1378] <210> 151  
[1379] <211> 7  
[1380] <212> PRT  
[1381] <213> 人工序列  
[1382] <220>  
[1383] <223> 人工序列的描述:合成肽  
[1384] <400> 151  
[1385] Asp Asp Ser His Arg Lys Asp  
[1386] 1 5  
[1387] <210> 152  
[1388] <211> 7  
[1389] <212> PRT  
[1390] <213> 人工序列  
[1391] <220>  
[1392] <223> 人工序列的描述:合成肽  
[1393] <400> 152  
[1394] Arg Ser Asp His Leu Thr Gln  
[1395] 1 5  
[1396] <210> 153  
[1397] <211> 7  
[1398] <212> PRT  
[1399] <213> 人工序列  
[1400] <220>  
[1401] <223> 人工序列的描述:合成肽  
[1402] <400> 153  
[1403] Arg Ser Asp Ser Leu Leu Arg

- [1404] 1 5
- [1405] <210> 154
- [1406] <211> 7
- [1407] <212> PRT
- [1408] <213> 人工序列
- [1409] <220>
- [1410] <223> 人工序列的描述:合成肽
- [1411] <400> 154
- [1412] Arg Arg Asp Trp Leu Pro Gln
- [1413] 1 5
- [1414] <210> 155
- [1415] <211> 28
- [1416] <212> DNA
- [1417] <213> 未知
- [1418] <220>
- [1419] <223> 未知序列的描述:人或小鼠靶序列
- [1420] <400> 155
- [1421] cgcactcgcc gcgagggttg ccgggacg 28
- [1422] <210> 156
- [1423] <400> 156
- [1424] 000
- [1425] <210> 157
- [1426] <211> 28
- [1427] <212> DNA
- [1428] <213> 未知
- [1429] <220>
- [1430] <223> 未知序列的描述:人或小鼠靶序列
- [1431] <400> 157
- [1432] agcagcagca gcagcagcag cagcagca 28
- [1433] <210> 158
- [1434] <211> 28
- [1435] <212> DNA
- [1436] <213> 未知
- [1437] <220>
- [1438] <223> 未知序列的描述:人或小鼠靶序列
- [1439] <400> 158
- [1440] ctgctgctgc tgctgctgct gctggaag 28
- [1441] <210> 159
- [1442] <211> 28

- [1443] <212> DNA
- [1444] <213> 未知
- [1445] <220>
- [1446] <223> 未知序列的描述:人或小鼠靶序列
- [1447] <400> 159
- [1448] cctgtccaga gggtcgcggt acctccct 28
- [1449] <210> 160
- [1450] <211> 28
- [1451] <212> DNA
- [1452] <213> 未知
- [1453] <220>
- [1454] <223> 未知序列的描述:人或小鼠靶序列
- [1455] <400> 160
- [1456] tgccggacct ggcagcggcg gtggtggc 28
- [1457] <210> 161
- [1458] <211> 14
- [1459] <212> PRT
- [1460] <213> 人工序列
- [1461] <220>
- [1462] <223> 人工序列的描述:合成肽
- [1463] <400> 161
- [1464] Arg Ser Asp Asn Leu Ser Glu Lys Arg Cys Asn Leu Arg Cys
- [1465] 1                    5                    10
- [1466] <210> 162
- [1467] <211> 14
- [1468] <212> PRT
- [1469] <213> 人工序列
- [1470] <220>
- [1471] <223> 人工序列的描述:合成肽
- [1472] <400> 162
- [1473] Arg Ser Asp Asn Leu Ser Glu Lys Pro Tyr Asn Leu Arg Thr
- [1474] 1                    5                    10
- [1475] <210> 163
- [1476] <211> 14
- [1477] <212> PRT
- [1478] <213> 人工序列
- [1479] <220>
- [1480] <223> 人工序列的描述:合成肽
- [1481] <400> 163





- [1560] <220>  
[1561] <223> 人工序列的描述:合成肽  
[1562] <400> 172  
[1563] Gln Ser Gly Asp Leu Thr Arg Gln Ser Ser Asp Leu Arg Arg  
[1564] 1 5 10  
[1565] <210> 173  
[1566] <211> 14  
[1567] <212> PRT  
[1568] <213> 人工序列  
[1569] <220>  
[1570] <223> 人工序列的描述:合成肽  
[1571] <400> 173  
[1572] Gln Ser Ser Asp Leu Ser Arg Gln Ser Ser Asp Leu Arg Arg  
[1573] 1 5 10  
[1574] <210> 174  
[1575] <211> 14  
[1576] <212> PRT  
[1577] <213> 人工序列  
[1578] <220>  
[1579] <223> 人工序列的描述:合成肽  
[1580] <400> 174  
[1581] Gln Ser Ser Asp Leu Ser Arg His Arg Ser Thr Arg Asn Arg  
[1582] 1 5 10  
[1583] <210> 175  
[1584] <211> 7  
[1585] <212> PRT  
[1586] <213> 人工序列  
[1587] <220>  
[1588] <223> 人工序列的描述:合成肽  
[1589] <400> 175  
[1590] Met Ala Cys Cys Arg Tyr Ala  
[1591] 1 5  
[1592] <210> 176  
[1593] <211> 7  
[1594] <212> PRT  
[1595] <213> 人工序列  
[1596] <220>  
[1597] <223> 人工序列的描述:合成肽  
[1598] <400> 176





[1677]	50	55	60
[1678]	Val His Met Ala Arg Asn Ala His Ser Gly Val Pro		
[1679]	65	70	75
[1680]	<210> 183		
[1681]	<211> 133		
[1682]	<212> PRT		
[1683]	<213> 人工序列		
[1684]	<220>		
[1685]	<223> 人工序列的描述:合成多肽		
[1686]	<400> 183		
[1687]	Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Leu Arg		
[1688]	1	5	10
[1689]	Gly Ala Ala His His Cys Gln His Cys Asp Met Tyr Phe Ala Asp Asn		
[1690]	20	25	30
[1691]	Ile Leu Tyr Thr Ile His Met Gly Cys His Ser Cys Asp Asp Val Phe		
[1692]	35	40	45
[1693]	Lys Cys Asn Met Cys Gly Glu Lys Cys Asp Gly Pro Val Gly Leu Phe		
[1694]	50	55	60
[1695]	Val His Met Ala Arg Asn Ala His Gly Glu Lys Pro Thr Lys Cys Val		
[1696]	65	70	75
[1697]	His Cys Gly Ile Val Phe Leu Asp Glu Val Met Tyr Ala Leu His Met		
[1698]	85	90	95
[1699]	Ser Cys His Gly Phe Arg Asp Pro Phe Glu Cys Asn Ile Cys Gly Tyr		
[1700]	100	105	110
[1701]	His Ser Gln Asp Arg Tyr Glu Phe Ser Ser His Ile Val Arg Gly Glu		
[1702]	115	120	125
[1703]	His Ser Gly Val Pro		
[1704]	130		
[1705]	<210> 184		
[1706]	<211> 57		
[1707]	<212> PRT		
[1708]	<213> 人工序列		
[1709]	<220>		
[1710]	<223> 人工序列的描述:合成多肽		
[1711]	<400> 184		
[1712]	Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Leu Arg		
[1713]	1	5	10
[1714]	Gly Ala Ala Gly Gly Ala Gln Leu Glu Lys Glu Leu Gln Ala Leu Glu		
[1715]	20	25	30

[1716]	Lys Lys Leu Ala Gln Leu Glu Trp Glu Asn Gln Ala Leu Glu Lys Glu
[1717]	35 40 45
[1718]	Leu Ala Gln Gly Gly Ser Gly Val Pro
[1719]	50 55
[1720]	<210> 185
[1721]	<211> 57
[1722]	<212> PRT
[1723]	<213> 人工序列
[1724]	<220>
[1725]	<223> 人工序列的描述:合成多肽
[1726]	<400> 185
[1727]	Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Leu Arg
[1728]	1 5 10 15
[1729]	Gly Ala Ala Gly Gly Glu Gln Leu Glu Lys Lys Leu Gln Ala Leu Glu
[1730]	20 25 30
[1731]	Lys Lys Leu Ala Gln Leu Glu Trp Lys Asn Gln Ala Leu Glu Lys Lys
[1732]	35 40 45
[1733]	Leu Ala Gln Gly Gly Ser Gly Val Pro
[1734]	50 55
[1735]	<210> 186
[1736]	<211> 53
[1737]	<212> PRT
[1738]	<213> 人工序列
[1739]	<220>
[1740]	<223> 人工序列的描述:合成多肽
[1741]	<400> 186
[1742]	Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Leu Arg
[1743]	1 5 10 15
[1744]	Gly Ala Ala Gly Gly Glu Gln Leu Glu Lys Lys Leu Gln Ala Leu Glu
[1745]	20 25 30
[1746]	Lys Lys Leu Ala Gln Leu Glu Trp Lys Asn Gln Ala Leu Glu Lys Gly
[1747]	35 40 45
[1748]	Gly Ser Gly Val Pro
[1749]	50
[1750]	<210> 187
[1751]	<211> 50
[1752]	<212> PRT
[1753]	<213> 人工序列
[1754]	<220>

[1755] <223> 人工序列的描述:合成多肽  
 [1756] <400> 187  
 [1757] Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Leu Arg  
 [1758] 1 5 10 15  
 [1759] Gly Ala Ala Gly Gly Glu Gln Leu Glu Lys Lys Leu Gln Ala Leu Glu  
 [1760] 20 25 30  
 [1761] Lys Lys Leu Ala Gln Leu Glu Trp Lys Asn Gln Ala Gly Gly Ser Gly  
 [1762] 35 40 45  
 [1763] Val Pro  
 [1764] 50  
 [1765] <210> 188  
 [1766] <211> 52  
 [1767] <212> PRT  
 [1768] <213> 人工序列  
 [1769] <220>  
 [1770] <223> 人工序列的描述:合成多肽  
 [1771] <400> 188  
 [1772] Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Leu Ala  
 [1773] 1 5 10 15  
 [1774] Ala Gly Gly Glu Gln Leu Glu Lys Lys Leu Gln Ala Leu Glu Lys Lys  
 [1775] 20 25 30  
 [1776] Leu Ala Gln Leu Glu Trp Lys Asn Gln Ala Leu Glu Lys Lys Leu Ala  
 [1777] 35 40 45  
 [1778] Gln Gly Val Pro  
 [1779] 50  
 [1780] <210> 189  
 [1781] <211> 55  
 [1782] <212> PRT  
 [1783] <213> 人工序列  
 [1784] <220>  
 [1785] <223> 人工序列的描述:合成多肽  
 [1786] <400> 189  
 [1787] Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Leu Ala  
 [1788] 1 5 10 15  
 [1789] Ala Gly Gly Glu Gln Leu Glu Lys Lys Leu Gln Ala Leu Glu Lys Lys  
 [1790] 20 25 30  
 [1791] Leu Ala Gln Leu Glu Trp Lys Asn Gln Ala Leu Glu Lys Lys Leu Ala  
 [1792] 35 40 45  
 [1793] Gln Gly Gly Ser Gly Val Pro



- [1833] <220>
- [1834] <223> 人工序列的描述:合成引物
- [1835] <400> 193
- [1836] caggtccggc agaggaacc 19
- [1837] <210> 194
- [1838] <211> 22
- [1839] <212> DNA
- [1840] <213> 人工序列
- [1841] <220>
- [1842] <223> 人工序列的描述:合成引物
- [1843] <400> 194
- [1844] ttcacacggt ctttcttggt gg 22
- [1845] <210> 195
- [1846] <211> 16
- [1847] <212> DNA
- [1848] <213> 人工序列
- [1849] <220>
- [1850] <223> 人工序列的描述:合成引物
- [1851] <400> 195
- [1852] gcccggtgt ggctga 16
- [1853] <210> 196
- [1854] <211> 22
- [1855] <212> DNA
- [1856] <213> 人工序列
- [1857] <220>
- [1858] <223> 人工序列的描述:合成引物
- [1859] <400> 196
- [1860] ttcacacggt ctttcttggt gg 22
- [1861] <210> 197
- [1862] <211> 175
- [1863] <212> PRT
- [1864] <213> 人工序列
- [1865] <220>
- [1866] <223> 人工序列的描述:合成多肽
- [1867] <400> 197
- [1868] Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp
- [1869] 1 5 10 15
- [1870] Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val
- [1871] 20 25 30

[1872]	Gly Ile His Gly Val Pro Met Val Asp Leu Arg Thr Leu Gly Tyr Ser
[1873]	35 40 45
[1874]	Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala
[1875]	50 55 60
[1876]	Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile
[1877]	65 70 75 80
[1878]	Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys
[1879]	85 90 95
[1880]	Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile
[1881]	100 105 110
[1882]	Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu
[1883]	115 120 125
[1884]	Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr
[1885]	130 135 140
[1886]	Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val Glu
[1887]	145 150 155 160
[1888]	Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn
[1889]	165 170 175
[1890]	<210> 198
[1891]	<211> 175
[1892]	<212> PRT
[1893]	<213> 人工序列
[1894]	<220>
[1895]	<223> 人工序列的描述:合成多肽
[1896]	<400> 198
[1897]	Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp
[1898]	1 5 10 15
[1899]	Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val
[1900]	20 25 30
[1901]	Gly Ile His Gly Val Pro Met Val Asp Leu Arg Thr Leu Gly Tyr Ser
[1902]	35 40 45
[1903]	Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala
[1904]	50 55 60
[1905]	Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile
[1906]	65 70 75 80
[1907]	Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys
[1908]	85 90 95
[1909]	Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile
[1910]	100 105 110

[1911]	Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu
[1912]	115 120 125
[1913]	Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr
[1914]	130 135 140
[1915]	Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val Glu
[1916]	145 150 155 160
[1917]	Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn
[1918]	165 170 175
[1919]	<210> 199
[1920]	<211> 175
[1921]	<212> PRT
[1922]	<213> 人工序列
[1923]	<220>
[1924]	<223> 人工序列的描述:合成多肽
[1925]	<400> 199
[1926]	Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp
[1927]	1 5 10 15
[1928]	Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val
[1929]	20 25 30
[1930]	Gly Ile His Gly Val Pro Met Val Asp Leu Arg Thr Leu Gly Tyr Ser
[1931]	35 40 45
[1932]	Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala
[1933]	50 55 60
[1934]	Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile
[1935]	65 70 75 80
[1936]	Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys
[1937]	85 90 95
[1938]	Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile
[1939]	100 105 110
[1940]	Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu
[1941]	115 120 125
[1942]	Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr
[1943]	130 135 140
[1944]	Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val Glu
[1945]	145 150 155 160
[1946]	Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn
[1947]	165 170 175
[1948]	<210> 200
[1949]	<211> 21

- [1950] <212> DNA  
[1951] <213> 未知  
[1952] <220>  
[1953] <223> 未知序列的描述:靶序列  
[1954] <400> 200  
[1955] gcagcagcag cagcagcagc a 21  
[1956] <210> 201  
[1957] <211> 18  
[1958] <212> DNA  
[1959] <213> 未知  
[1960] <220>  
[1961] <223> 未知序列的描述:靶序列  
[1962] <400> 201  
[1963] gcagcagcag cagcagca 18  
[1964] <210> 202  
[1965] <211> 15  
[1966] <212> DNA  
[1967] <213> 未知  
[1968] <220>  
[1969] <223> 未知序列的描述:靶序列  
[1970] <400> 202  
[1971] gcagcagcag cagca 15  
[1972] <210> 203  
[1973] <211> 12  
[1974] <212> DNA  
[1975] <213> 未知  
[1976] <220>  
[1977] <223> 未知序列的描述:靶序列  
[1978] <400> 203  
[1979] gcagcagcag ca 12  
[1980] <210> 204  
[1981] <211> 21  
[1982] <212> DNA  
[1983] <213> 未知  
[1984] <220>  
[1985] <223> 未知序列的描述:靶序列  
[1986] <400> 204  
[1987] ctgctgctgc tgctgctgct g 21  
[1988] <210> 205

- [1989] <211> 18  
[1990] <212> DNA  
[1991] <213> 未知  
[1992] <220>  
[1993] <223> 未知序列的描述:靶序列  
[1994] <400> 205  
[1995] ctgctgctgc tgctgctg 18  
[1996] <210> 206  
[1997] <211> 15  
[1998] <212> DNA  
[1999] <213> 未知  
[2000] <220>  
[2001] <223> 未知序列的描述:靶序列  
[2002] <400> 206  
[2003] ctgctgctgc tgctg 15  
[2004] <210> 207  
[2005] <211> 12  
[2006] <212> DNA  
[2007] <213> 未知  
[2008] <220>  
[2009] <223> 未知序列的描述:靶序列  
[2010] <400> 207  
[2011] ctgctgctgc tg 12  
[2012] <210> 208  
[2013] <211> 29  
[2014] <212> DNA  
[2015] <213> 未知  
[2016] <220>  
[2017] <223> 未知序列的描述:人或小鼠靶序列  
[2018] <400> 208  
[2019] ctgctgctgc tgctgctgct gctggaagg 29  
[2020] <210> 209  
[2021] <211> 28  
[2022] <212> DNA  
[2023] <213> 未知  
[2024] <220>  
[2025] <223> 未知序列的描述:人或小鼠靶序列  
[2026] <400> 209  
[2027] gcggcgagtgc cgtcccgtga cgtcatgc 28

[2028]	<210>	210
[2029]	<211>	170
[2030]	<212>	PRT
[2031]	<213>	人工序列
[2032]	<220>	
[2033]	<223>	人工序列的描述:合成多肽
[2034]	<400>	210
[2035]	Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro	
[2036]	1	5 10 15
[2037]	Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala	
[2038]		20 25 30
[2039]	Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro	
[2040]		35 40 45
[2041]	His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg	
[2042]		50 55 60
[2043]	Thr Ser His Arg Val Ala Gly Ser Gly Met Asp Ala Lys Ser Leu Thr	
[2044]		65 70 75 80
[2045]	Ala Trp Ser Arg Thr Leu Val Thr Phe Lys Asp Val Phe Val Asp Phe	
[2046]		85 90 95
[2047]	Thr Arg Glu Glu Trp Lys Leu Leu Asp Thr Ala Gln Gln Ile Val Tyr	
[2048]		100 105 110
[2049]	Arg Asn Val Met Leu Glu Asn Tyr Lys Asn Leu Val Ser Leu Gly Tyr	
[2050]		115 120 125
[2051]	Gln Leu Thr Lys Pro Asp Val Ile Leu Arg Leu Glu Lys Gly Glu Glu	
[2052]		130 135 140
[2053]	Pro Trp Leu Val Glu Arg Glu Ile His Gln Glu Thr His Pro Asp Ser	
[2054]		145 150 155 160
[2055]	Glu Thr Ala Phe Glu Ile Lys Ser Ser Val	
[2056]		165 170
[2057]	<210>	211
[2058]	<211>	338
[2059]	<212>	PRT
[2060]	<213>	人工序列
[2061]	<220>	
[2062]	<223>	人工序列的描述:合成多肽
[2063]	<400>	211
[2064]	Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro	
[2065]	1	5 10 15
[2066]	Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala	



[2106] Ser Val  
 [2107] <210> 212  
 [2108] <211> 384  
 [2109] <212> PRT  
 [2110] <213> 人工序列  
 [2111] <220>  
 [2112] <223> 人工序列的描述:合成多肽  
 [2113] <400> 212  
 [2114] Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro  
 [2115] 1 5 10 15  
 [2116] Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala  
 [2117] 20 25 30  
 [2118] Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro  
 [2119] 35 40 45  
 [2120] His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg  
 [2121] 50 55 60  
 [2122] Thr Ser His Arg Val Ala Asp His Ala Gln Val Val Arg Val Leu Gly  
 [2123] 65 70 75 80  
 [2124] Phe Phe Gln Cys His Ser His Pro Ala Gln Ala Phe Asp Asp Ala Met  
 [2125] 85 90 95  
 [2126] Thr Gln Phe Gly Met Ser Arg His Gly Leu Leu Gln Leu Phe Arg Arg  
 [2127] 100 105 110  
 [2128] Val Gly Val Thr Glu Leu Glu Ala Arg Ser Gly Thr Leu Pro Pro Ala  
 [2129] 115 120 125  
 [2130] Ser Gln Arg Trp Asp Arg Ile Leu Gln Ala Ser Gly Met Lys Arg Ala  
 [2131] 130 135 140  
 [2132] Lys Pro Ser Pro Thr Ser Thr Gln Thr Pro Asp Gln Ala Ser Leu His  
 [2133] 145 150 155 160  
 [2134] Ala Phe Ala Asp Ser Leu Glu Arg Asp Leu Asp Ala Pro Ser Pro Thr  
 [2135] 165 170 175  
 [2136] His Glu Gly Asp Gln Arg Arg Ala Ser Ser Arg Lys Arg Ser Arg Ser  
 [2137] 180 185 190  
 [2138] Asp Arg Ala Val Thr Gly Pro Ser Ala Gln Gln Ser Phe Glu Val Arg  
 [2139] 195 200 205  
 [2140] Ala Pro Glu Gln Arg Asp Ala Leu His Leu Pro Leu Ser Trp Arg Val  
 [2141] 210 215 220  
 [2142] Lys Arg Pro Arg Thr Ser Ile Gly Gly Gly Leu Pro Asp Pro Thr Pro  
 [2143] 225 230 235 240  
 [2144] Thr Ala Ala Asp Leu Ala Ala Ser Ser Thr Val Met Arg Glu Gln Asp

[2145]		245		250		255
[2146]	Glu Asp Pro Phe Ala Gly Ala Ala Asp Asp Phe Pro Ala Phe Asn Glu					
[2147]		260		265		270
[2148]	Glu Glu Leu Ala Trp Leu Met Glu Leu Leu Pro Gln Gly Ser Gly Met					
[2149]		275		280		285
[2150]	Asp Ala Lys Ser Leu Thr Ala Trp Ser Arg Thr Leu Val Thr Phe Lys					
[2151]		290		295		300
[2152]	Asp Val Phe Val Asp Phe Thr Arg Glu Glu Trp Lys Leu Leu Asp Thr					
[2153]		305		310		315
[2154]	Ala Gln Gln Ile Val Tyr Arg Asn Val Met Leu Glu Asn Tyr Lys Asn					
[2155]		325		330		335
[2156]	Leu Val Ser Leu Gly Tyr Gln Leu Thr Lys Pro Asp Val Ile Leu Arg					
[2157]		340		345		350
[2158]	Leu Glu Lys Gly Glu Glu Pro Trp Leu Val Glu Arg Glu Ile His Gln					
[2159]		355		360		365
[2160]	Glu Thr His Pro Asp Ser Glu Thr Ala Phe Glu Ile Lys Ser Ser Val					
[2161]		370		375		380
[2162]	<210>	213				
[2163]	<211>	174				
[2164]	<212>	PRT				
[2165]	<213>	人工序列				
[2166]	<220>					
[2167]	<223>	人工序列的描述:合成多肽				
[2168]	<400>	213				
[2169]	Gly Ser Gly Gly Thr Lys Cys Val His Cys Gly Ile Val Phe Leu Asp					
[2170]	1	5		10		15
[2171]	Glu Val Met Tyr Ala Leu His Met Ser Cys His Gly Phe Arg Asp Pro					
[2172]		20		25		30
[2173]	Phe Glu Cys Asn Ile Cys Gly Tyr His Ser Gln Asp Arg Tyr Glu Phe					
[2174]		35		40		45
[2175]	Ser Ser His Ile Val Arg Gly Glu His Leu Arg Gln Lys Asp Ala Ala					
[2176]		50		55		60
[2177]	Arg Ser Arg Ser Gly Met Asp Ala Lys Ser Leu Thr Ala Trp Ser Arg					
[2178]		65		70		75
[2179]	Thr Leu Val Thr Phe Lys Asp Val Phe Val Asp Phe Thr Arg Glu Glu					
[2180]		85		90		95
[2181]	Trp Lys Leu Leu Asp Thr Ala Gln Gln Ile Val Tyr Arg Asn Val Met					
[2182]		100		105		110
[2183]	Leu Glu Asn Tyr Lys Asn Leu Val Ser Leu Gly Tyr Gln Leu Thr Lys					

[2184]	115	120	125
[2185]	Pro Asp Val Ile Leu Arg Leu Glu Lys Gly Glu Glu Pro Trp Leu Val		
[2186]	130	135	140
[2187]	Glu Arg Glu Ile His Gln Glu Thr His Pro Asp Ser Glu Thr Ala Phe		
[2188]	145	150	155
[2189]	Glu Ile Lys Ser Ser Val Asp Tyr Lys Asp Asp Asp Asp Lys		
[2190]	165	170	
[2191]	<210> 214		
[2192]	<211> 174		
[2193]	<212> PRT		
[2194]	<213> 人工序列		
[2195]	<220>		
[2196]	<223> 人工序列的描述:合成多肽		
[2197]	<400> 214		
[2198]	Gly Ser Gly Gly Thr Lys Cys Val His Cys Gly Ile Val Phe Leu Asp		
[2199]	1	5	10
[2200]	Glu Val Met Tyr Ala Leu His Met Ser Cys His Gly Phe Arg Asp Pro		
[2201]	20	25	30
[2202]	Phe Glu Cys Asn Ile Cys Gly Tyr His Ser Gln Asp Arg Tyr Glu Phe		
[2203]	35	40	45
[2204]	Ser Ser His Ile Val Arg Gly Glu His Leu Arg Gln Lys Asp Ala Ala		
[2205]	50	55	60
[2206]	Arg Ser Arg Ser Gly Met Asp Ala Lys Ser Leu Thr Ala Trp Ser Arg		
[2207]	65	70	75
[2208]	Thr Leu Val Thr Phe Lys Asp Val Phe Val Asp Phe Thr Arg Glu Glu		
[2209]	85	90	95
[2210]	Trp Lys Leu Leu Asp Thr Ala Gln Gln Ile Val Tyr Arg Asn Val Met		
[2211]	100	105	110
[2212]	Leu Glu Asn Tyr Lys Asn Leu Val Ser Leu Gly Tyr Gln Leu Thr Lys		
[2213]	115	120	125
[2214]	Pro Asp Val Ile Leu Arg Leu Glu Lys Gly Glu Glu Pro Trp Leu Val		
[2215]	130	135	140
[2216]	Glu Arg Glu Ile His Gln Glu Thr His Pro Asp Ser Glu Thr Ala Phe		
[2217]	145	150	155
[2218]	Glu Ile Lys Ser Ser Val Asp Tyr Lys Asp Asp Asp Asp Lys		
[2219]	165	170	
[2220]	<210> 215		
[2221]	<211> 174		
[2222]	<212> PRT		

[2223]	<213>	人工序列
[2224]	<220>	
[2225]	<223>	人工序列的描述:合成多肽
[2226]	<400>	215
[2227]	Gly Ser Gly Gly His His Cys Gln His Cys Asp Met Tyr Phe Ala Asp	
[2228]	1	5 10 15
[2229]	Asn Ile Leu Tyr Thr Ile His Met Gly Cys His Gly Tyr Glu Asn Pro	
[2230]		20 25 30
[2231]	Phe Glu Cys Asn Ile Cys Gly Tyr His Ser Gln Asp Arg Tyr Glu Phe	
[2232]		35 40 45
[2233]	Ser Ser His Ile Val Arg Gly Glu His Leu Arg Gln Lys Asp Ala Ala	
[2234]		50 55 60
[2235]	Arg Ser Arg Ser Gly Met Asp Ala Lys Ser Leu Thr Ala Trp Ser Arg	
[2236]		65 70 75 80
[2237]	Thr Leu Val Thr Phe Lys Asp Val Phe Val Asp Phe Thr Arg Glu Glu	
[2238]		85 90 95
[2239]	Trp Lys Leu Leu Asp Thr Ala Gln Gln Ile Val Tyr Arg Asn Val Met	
[2240]		100 105 110
[2241]	Leu Glu Asn Tyr Lys Asn Leu Val Ser Leu Gly Tyr Gln Leu Thr Lys	
[2242]		115 120 125
[2243]	Pro Asp Val Ile Leu Arg Leu Glu Lys Gly Glu Glu Pro Trp Leu Val	
[2244]		130 135 140
[2245]	Glu Arg Glu Ile His Gln Glu Thr His Pro Asp Ser Glu Thr Ala Phe	
[2246]		145 150 155 160
[2247]	Glu Ile Lys Ser Ser Val Asp Tyr Lys Asp Asp Asp Asp Lys	
[2248]		165 170
[2249]	<210>	216
[2250]	<211>	231
[2251]	<212>	PRT
[2252]	<213>	人工序列
[2253]	<220>	
[2254]	<223>	人工序列的描述:合成多肽
[2255]	<400>	216
[2256]	Gly Ser Gly Gly Phe Lys Cys Glu His Cys Arg Ile Leu Phe Leu Asp	
[2257]	1	5 10 15
[2258]	His Val Met Phe Thr Ile His Met Gly Cys His Gly Phe Arg Asp Pro	
[2259]		20 25 30
[2260]	Phe Lys Cys Asn Met Cys Gly Glu Lys Cys Asp Gly Pro Val Gly Leu	
[2261]		35 40 45

[2262]	Phe Val His Met Ala Arg Asn Ala His Gly Glu Lys Pro Phe Tyr Cys
[2263]	50 55 60
[2264]	Glu His Cys Glu Ile Thr Phe Arg Asp Val Val Met Tyr Ser Leu His
[2265]	65 70 75 80
[2266]	Lys Gly Tyr His Gly Phe Arg Asp Pro Phe Glu Cys Asn Ile Cys Gly
[2267]	85 90 95
[2268]	Tyr His Ser Gln Asp Arg Tyr Glu Phe Ser Ser His Ile Val Arg Gly
[2269]	100 105 110
[2270]	Glu His Leu Arg Gln Lys Asp Ala Ala Arg Ser Arg Ser Gly Met Asp
[2271]	115 120 125
[2272]	Ala Lys Ser Leu Thr Ala Trp Ser Arg Thr Leu Val Thr Phe Lys Asp
[2273]	130 135 140
[2274]	Val Phe Val Asp Phe Thr Arg Glu Glu Trp Lys Leu Leu Asp Thr Ala
[2275]	145 150 155 160
[2276]	Gln Gln Ile Val Tyr Arg Asn Val Met Leu Glu Asn Tyr Lys Asn Leu
[2277]	165 170 175
[2278]	Val Ser Leu Gly Tyr Gln Leu Thr Lys Pro Asp Val Ile Leu Arg Leu
[2279]	180 185 190
[2280]	Glu Lys Gly Glu Glu Pro Trp Leu Val Glu Arg Glu Ile His Gln Glu
[2281]	195 200 205
[2282]	Thr His Pro Asp Ser Glu Thr Ala Phe Glu Ile Lys Ser Ser Val Asp
[2283]	210 215 220
[2284]	Tyr Lys Asp Asp Asp Asp Lys
[2285]	225 230
[2286]	<210> 217
[2287]	<211> 153
[2288]	<212> PRT
[2289]	<213> 人工序列
[2290]	<220>
[2291]	<223> 人工序列的描述:合成多肽
[2292]	<400> 217
[2293]	Gly Ser Gly Gly Ala Gln Leu Lys Lys Lys Leu Gln Ala Asn Lys Lys
[2294]	1 5 10 15
[2295]	Glu Leu Ala Gln Leu Lys Trp Lys Leu Gln Ala Leu Lys Lys Lys Leu
[2296]	20 25 30
[2297]	Ala Gln Gly Gly Leu Arg Gln Lys Asp Ala Ala Arg Ser Arg Ser Gly
[2298]	35 40 45
[2299]	Met Asp Ala Lys Ser Leu Thr Ala Trp Ser Arg Thr Leu Val Thr Phe
[2300]	50 55 60

[2301]	Lys Asp Val Phe Val Asp Phe Thr Arg Glu Glu Trp Lys Leu Leu Asp
[2302]	65 70 75 80
[2303]	Thr Ala Gln Gln Ile Val Tyr Arg Asn Val Met Leu Glu Asn Tyr Lys
[2304]	85 90 95
[2305]	Asn Leu Val Ser Leu Gly Tyr Gln Leu Thr Lys Pro Asp Val Ile Leu
[2306]	100 105 110
[2307]	Arg Leu Glu Lys Gly Glu Glu Pro Trp Leu Val Glu Arg Glu Ile His
[2308]	115 120 125
[2309]	Gln Glu Thr His Pro Asp Ser Glu Thr Ala Phe Glu Ile Lys Ser Ser
[2310]	130 135 140
[2311]	Val Asp Tyr Lys Asp Asp Asp Asp Lys
[2312]	145 150
[2313]	<210> 218
[2314]	<211> 152
[2315]	<212> PRT
[2316]	<213> 人工序列
[2317]	<220>
[2318]	<223> 人工序列的描述:合成多肽
[2319]	<400> 218
[2320]	Gly Ser Gly Gly Ala Leu Lys Lys Glu Leu Gln Ala Asn Lys Lys Glu
[2321]	1 5 10 15
[2322]	Leu Ala Gln Leu Lys Trp Glu Leu Gln Ala Leu Lys Lys Glu Leu Ala
[2323]	20 25 30
[2324]	Gln Gly Gly Leu Arg Gln Lys Asp Ala Ala Arg Ser Arg Ser Gly Met
[2325]	35 40 45
[2326]	Asp Ala Lys Ser Leu Thr Ala Trp Ser Arg Thr Leu Val Thr Phe Lys
[2327]	50 55 60
[2328]	Asp Val Phe Val Asp Phe Thr Arg Glu Glu Trp Lys Leu Leu Asp Thr
[2329]	65 70 75 80
[2330]	Ala Gln Gln Ile Val Tyr Arg Asn Val Met Leu Glu Asn Tyr Lys Asn
[2331]	85 90 95
[2332]	Leu Val Ser Leu Gly Tyr Gln Leu Thr Lys Pro Asp Val Ile Leu Arg
[2333]	100 105 110
[2334]	Leu Glu Lys Gly Glu Glu Pro Trp Leu Val Glu Arg Glu Ile His Gln
[2335]	115 120 125
[2336]	Glu Thr His Pro Asp Ser Glu Thr Ala Phe Glu Ile Lys Ser Ser Val
[2337]	130 135 140
[2338]	Asp Tyr Lys Asp Asp Asp Asp Lys
[2339]	145 150

[2340] <210> 219  
 [2341] <211> 148  
 [2342] <212> PRT  
 [2343] <213> 人工序列  
 [2344] <220>  
 [2345] <223> 人工序列的描述:合成多肽  
 [2346] <400> 219  
 [2347] Gly Ser Gly Gly Glu Leu Gln Ala Asn Lys Lys Glu Leu Ala Gln Leu  
 [2348] 1 5 10 15  
 [2349] Lys Trp Glu Leu Gln Ala Leu Lys Lys Glu Leu Ala Gln Gly Gly Leu  
 [2350] 20 25 30  
 [2351] Arg Gln Lys Asp Ala Ala Arg Ser Arg Ser Gly Met Asp Ala Lys Ser  
 [2352] 35 40 45  
 [2353] Leu Thr Ala Trp Ser Arg Thr Leu Val Thr Phe Lys Asp Val Phe Val  
 [2354] 50 55 60  
 [2355] Asp Phe Thr Arg Glu Glu Trp Lys Leu Leu Asp Thr Ala Gln Gln Ile  
 [2356] 65 70 75 80  
 [2357] Val Tyr Arg Asn Val Met Leu Glu Asn Tyr Lys Asn Leu Val Ser Leu  
 [2358] 85 90 95  
 [2359] Gly Tyr Gln Leu Thr Lys Pro Asp Val Ile Leu Arg Leu Glu Lys Gly  
 [2360] 100 105 110  
 [2361] Glu Glu Pro Trp Leu Val Glu Arg Glu Ile His Gln Glu Thr His Pro  
 [2362] 115 120 125  
 [2363] Asp Ser Glu Thr Ala Phe Glu Ile Lys Ser Ser Val Asp Tyr Lys Asp  
 [2364] 130 135 140  
 [2365] Asp Asp Asp Lys  
 [2366] 145  
 [2367] <210> 220  
 [2368] <211> 146  
 [2369] <212> PRT  
 [2370] <213> 人工序列  
 [2371] <220>  
 [2372] <223> 人工序列的描述:合成多肽  
 [2373] <400> 220  
 [2374] Gly Ser Gly Gly Gln Ala Asn Lys Lys Glu Leu Ala Gln Leu Lys Trp  
 [2375] 1 5 10 15  
 [2376] Glu Leu Gln Ala Leu Lys Lys Glu Leu Ala Gln Gly Gly Leu Arg Gln  
 [2377] 20 25 30  
 [2378] Lys Asp Ala Ala Arg Ser Arg Ser Gly Met Asp Ala Lys Ser Leu Thr

[2379]	35	40	45
[2380]	Ala Trp Ser Arg Thr Leu Val Thr Phe Lys Asp Val Phe Val Asp Phe		
[2381]	50	55	60
[2382]	Thr Arg Glu Glu Trp Lys Leu Leu Asp Thr Ala Gln Gln Ile Val Tyr		
[2383]	65	70	75
[2384]	Arg Asn Val Met Leu Glu Asn Tyr Lys Asn Leu Val Ser Leu Gly Tyr		
[2385]	85	90	95
[2386]	Gln Leu Thr Lys Pro Asp Val Ile Leu Arg Leu Glu Lys Gly Glu Glu		
[2387]	100	105	110
[2388]	Pro Trp Leu Val Glu Arg Glu Ile His Gln Glu Thr His Pro Asp Ser		
[2389]	115	120	125
[2390]	Glu Thr Ala Phe Glu Ile Lys Ser Ser Val Asp Tyr Lys Asp Asp Asp		
[2391]	130	135	140
[2392]	Asp Lys		
[2393]	145		
[2394]	<210> 221		
[2395]	<211> 55		
[2396]	<212> PRT		
[2397]	<213> 人工序列		
[2398]	<220>		
[2399]	<223> 人工序列的描述:合成多肽		
[2400]	<400> 221		
[2401]	Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Leu Ala		
[2402]	1	5	10
[2403]	Ala Gly Gly Glu Gln Leu Glu Lys Lys Leu Gln Ala Leu Glu Lys Lys		
[2404]	20	25	30
[2405]	Leu Ala Gln Leu Glu Trp Lys Asn Gln Ala Leu Glu Lys Lys Leu Ala		
[2406]	35	40	45
[2407]	Gln Gly Gly Ser Gly Val Pro		
[2408]	50	55	
[2409]	<210> 222		
[2410]	<211> 141		
[2411]	<212> PRT		
[2412]	<213> 人工序列		
[2413]	<220>		
[2414]	<223> 人工序列的描述:合成多肽		
[2415]	<400> 222		
[2416]	Gly Ser Gly Gly Ala Leu Lys Lys Glu Leu Gln Ala Asn Lys Lys Glu		
[2417]	1	5	10

[2418]	Leu Ala Gln Leu Lys Trp Glu Leu Gln Ala Leu Lys Lys Glu Leu Ala
[2419]	20 25 30
[2420]	Gln Gly Ser Gly Met Asp Ala Lys Ser Leu Thr Ala Trp Ser Arg Thr
[2421]	35 40 45
[2422]	Leu Val Thr Phe Lys Asp Val Phe Val Asp Phe Thr Arg Glu Glu Trp
[2423]	50 55 60
[2424]	Lys Leu Leu Asp Thr Ala Gln Gln Ile Val Tyr Arg Asn Val Met Leu
[2425]	65 70 75 80
[2426]	Glu Asn Tyr Lys Asn Leu Val Ser Leu Gly Tyr Gln Leu Thr Lys Pro
[2427]	85 90 95
[2428]	Asp Val Ile Leu Arg Leu Glu Lys Gly Glu Glu Pro Trp Leu Val Glu
[2429]	100 105 110
[2430]	Arg Glu Ile His Gln Glu Thr His Pro Asp Ser Glu Thr Ala Phe Glu
[2431]	115 120 125
[2432]	Ile Lys Ser Ser Val Asp Tyr Lys Asp Asp Asp Asp Lys
[2433]	130 135 140
[2434]	<210> 223
[2435]	<211> 141
[2436]	<212> PRT
[2437]	<213> 人工序列
[2438]	<220>
[2439]	<223> 人工序列的描述:合成多肽
[2440]	<400> 223
[2441]	Gly Ser Gly Gly Ala Leu Lys Lys Glu Leu Gln Ala Asn Lys Lys Glu
[2442]	1 5 10 15
[2443]	Leu Ala Gln Leu Lys Trp Glu Leu Gln Ala Leu Lys Lys Glu Leu Ala
[2444]	20 25 30
[2445]	Gln Gly Ser Gly Met Asp Ala Lys Ser Leu Thr Ala Trp Ser Arg Thr
[2446]	35 40 45
[2447]	Leu Val Thr Phe Lys Asp Val Phe Val Asp Phe Thr Arg Glu Glu Trp
[2448]	50 55 60
[2449]	Lys Leu Leu Asp Thr Ala Gln Gln Ile Val Tyr Arg Asn Val Met Leu
[2450]	65 70 75 80
[2451]	Glu Asn Tyr Lys Asn Leu Val Ser Leu Gly Tyr Gln Leu Thr Lys Pro
[2452]	85 90 95
[2453]	Asp Val Ile Leu Arg Leu Glu Lys Gly Glu Glu Pro Trp Leu Val Glu
[2454]	100 105 110
[2455]	Arg Glu Ile His Gln Glu Thr His Pro Asp Ser Glu Thr Ala Phe Glu
[2456]	115 120 125

[2457]	Ile Lys Ser Ser Val Asp Tyr Lys Asp Asp Asp Asp Lys
[2458]	130 135 140
[2459]	<210> 224
[2460]	<211> 139
[2461]	<212> PRT
[2462]	<213> 人工序列
[2463]	<220>
[2464]	<223> 人工序列的描述:合成多肽
[2465]	<400> 224
[2466]	Gly Ser Ala Leu Lys Lys Glu Leu Gln Ala Asn Lys Lys Glu Leu Ala
[2467]	1 5 10 15
[2468]	Gln Leu Lys Trp Glu Leu Gln Ala Leu Lys Lys Glu Leu Ala Gln Gly
[2469]	20 25 30
[2470]	Ser Gly Met Asp Ala Lys Ser Leu Thr Ala Trp Ser Arg Thr Leu Val
[2471]	35 40 45
[2472]	Thr Phe Lys Asp Val Phe Val Asp Phe Thr Arg Glu Glu Trp Lys Leu
[2473]	50 55 60
[2474]	Leu Asp Thr Ala Gln Gln Ile Val Tyr Arg Asn Val Met Leu Glu Asn
[2475]	65 70 75 80
[2476]	Tyr Lys Asn Leu Val Ser Leu Gly Tyr Gln Leu Thr Lys Pro Asp Val
[2477]	85 90 95
[2478]	Ile Leu Arg Leu Glu Lys Gly Glu Glu Pro Trp Leu Val Glu Arg Glu
[2479]	100 105 110
[2480]	Ile His Gln Glu Thr His Pro Asp Ser Glu Thr Ala Phe Glu Ile Lys
[2481]	115 120 125
[2482]	Ser Ser Val Asp Tyr Lys Asp Asp Asp Asp Lys
[2483]	130 135
[2484]	<210> 225
[2485]	<211> 139
[2486]	<212> PRT
[2487]	<213> 人工序列
[2488]	<220>
[2489]	<223> 人工序列的描述:合成多肽
[2490]	<400> 225
[2491]	Gly Ser Ala Leu Lys Lys Glu Leu Gln Ala Asn Lys Lys Glu Leu Ala
[2492]	1 5 10 15
[2493]	Gln Leu Lys Trp Glu Leu Gln Ala Leu Lys Lys Glu Leu Ala Gln Gly
[2494]	20 25 30
[2495]	Ser Gly Met Asp Ala Lys Ser Leu Thr Ala Trp Ser Arg Thr Leu Val

[2496]	35	40	45
[2497]	Thr Phe Lys Asp Val Phe Val Asp Phe Thr Arg Glu Glu Trp Lys Leu		
[2498]	50	55	60
[2499]	Leu Asp Thr Ala Gln Gln Ile Val Tyr Arg Asn Val Met Leu Glu Asn		
[2500]	65	70	75
[2501]	Tyr Lys Asn Leu Val Ser Leu Gly Tyr Gln Leu Thr Lys Pro Asp Val		
[2502]	85	90	95
[2503]	Ile Leu Arg Leu Glu Lys Gly Glu Glu Pro Trp Leu Val Glu Arg Glu		
[2504]	100	105	110
[2505]	Ile His Gln Glu Thr His Pro Asp Ser Glu Thr Ala Phe Glu Ile Lys		
[2506]	115	120	125
[2507]	Ser Ser Val Asp Tyr Lys Asp Asp Asp Asp Lys		
[2508]	130	135	
[2509]	<210> 226		
[2510]	<211> 60		
[2511]	<212> DNA		
[2512]	<213> 人工序列		
[2513]	<220>		
[2514]	<223> 人工序列的描述:合成寡核苷酸		
[2515]	<220>		
[2516]	<221> misc_feature		
[2517]	<222> (1) .. (60)		
[2518]	<223> 该序列可能涵盖15-20个重复“CAG”单元		
[2519]	<400> 226		
[2520]	cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60		
[2521]	<210> 227		
[2522]	<211> 105		
[2523]	<212> DNA		
[2524]	<213> 人工序列		
[2525]	<220>		
[2526]	<223> 人工序列的描述:合成多核苷酸		
[2527]	<400> 227		
[2528]	cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60		
[2529]	cagcagcagc agcagcagca gcagcagcag cagcagcagc agcag 105		
[2530]	<210> 228		
[2531]	<211> 117		
[2532]	<212> DNA		
[2533]	<213> 人工序列		
[2534]	<220>		

- [2535] <223> 人工序列的描述:合成多核苷酸  
[2536] <220>  
[2537] <221> misc\_feature  
[2538] <222> (1) .. (117)  
[2539] <223> 该序列可能涵盖36-39个重复“CAG”单元  
[2540] <400> 228  
[2541] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60  
[2542] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 117  
[2543] <210> 229  
[2544] <211> 180  
[2545] <212> DNA  
[2546] <213> 人工序列  
[2547] <220>  
[2548] <223> 人工序列的描述:合成多核苷酸  
[2549] <400> 229  
[2550] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60  
[2551] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120  
[2552] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 180  
[2553] <210> 230  
[2554] <211> 150  
[2555] <212> DNA  
[2556] <213> 人工序列  
[2557] <220>  
[2558] <223> 人工序列的描述:合成多核苷酸  
[2559] <400> 230  
[2560] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60  
[2561] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120  
[2562] cagcagcagc agcagcagca gcagcagcag 150  
[2563] <210> 231  
[2564] <211> 240  
[2565] <212> DNA  
[2566] <213> 人工序列  
[2567] <220>  
[2568] <223> 人工序列的描述:合成多核苷酸  
[2569] <400> 231  
[2570] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60  
[2571] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120  
[2572] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 180  
[2573] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 240

- [2574] <210> 232
- [2575] <211> 327
- [2576] <212> DNA
- [2577] <213> 人工序列
- [2578] <220>
- [2579] <223> 人工序列的描述:合成多核苷酸
- [2580] <400> 232
- [2581] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60
- [2582] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120
- [2583] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 180
- [2584] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 240
- [2585] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 300
- [2586] cagcagcagc agcagcagca gcagcag 327
- [2587] <210> 233
- [2588] <211> 540
- [2589] <212> DNA
- [2590] <213> 人工序列
- [2591] <220>
- [2592] <223> 人工序列的描述:合成多核苷酸
- [2593] <400> 233
- [2594] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60
- [2595] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120
- [2596] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 180
- [2597] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 240
- [2598] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 300
- [2599] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 360
- [2600] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 420
- [2601] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 480
- [2602] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 540
- [2603] <210> 234
- [2604] <211> 30
- [2605] <212> DNA
- [2606] <213> 人工序列
- [2607] <220>
- [2608] <223> 人工序列的描述:合成寡核苷酸
- [2609] <400> 234
- [2610] cagcagcagc agcagcagca gcagcagcag 30
- [2611] <210> 235
- [2612] <211> 141

- [2613] <212> DNA  
[2614] <213> 人工序列  
[2615] <220>  
[2616] <223> 人工序列的描述:合成多核苷酸  
[2617] <400> 235  
[2618] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60  
[2619] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120  
[2620] cagcagcagc agcagcagca g 141  
[2621] <210> 236  
[2622] <211> 51  
[2623] <212> DNA  
[2624] <213> 人工序列  
[2625] <220>  
[2626] <223> 人工序列的描述:合成寡核苷酸  
[2627] <400> 236  
[2628] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca g 51  
[2629] <210> 237  
[2630] <211> 69  
[2631] <212> DNA  
[2632] <213> 人工序列  
[2633] <220>  
[2634] <223> 人工序列的描述:合成寡核苷酸  
[2635] <400> 237  
[2636] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60  
[2637] cagcagcag 69  
[2638] <210> 238  
[2639] <211> 333  
[2640] <212> DNA  
[2641] <213> 人工序列  
[2642] <220>  
[2643] <223> 人工序列的描述:合成多核苷酸  
[2644] <400> 238  
[2645] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60  
[2646] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120  
[2647] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 180  
[2648] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 240  
[2649] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 300  
[2650] cagcagcagc agcagcagca gcagcagcag cag 333  
[2651] <210> 239

- [2652] <211> 45  
[2653] <212> DNA  
[2654] <213> 人工序列  
[2655] <220>  
[2656] <223> 人工序列的描述:合成寡核苷酸  
[2657] <400> 239  
[2658] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcag 45  
[2659] <210> 240  
[2660] <211> 210  
[2661] <212> DNA  
[2662] <213> 人工序列  
[2663] <220>  
[2664] <223> 人工序列的描述:合成多核苷酸  
[2665] <400> 240  
[2666] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60  
[2667] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120  
[2668] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 180  
[2669] cagcagcagc agcagcagca gcagcagcag 210  
[2670] <210> 241  
[2671] <211> 300  
[2672] <212> DNA  
[2673] <213> 人工序列  
[2674] <220>  
[2675] <223> 人工序列的描述:合成多核苷酸  
[2676] <400> 241  
[2677] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60  
[2678] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120  
[2679] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 180  
[2680] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 240  
[2681] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 300  
[2682] <210> 242  
[2683] <211> 201  
[2684] <212> DNA  
[2685] <213> 人工序列  
[2686] <220>  
[2687] <223> 人工序列的描述:合成多核苷酸  
[2688] <400> 242  
[2689] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60  
[2690] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120

- [2691] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 180
- [2692] cagcagcagc agcagcagca g 201
- [2693] <210> 243
- [2694] <211> 135
- [2695] <212> DNA
- [2696] <213> 人工序列
- [2697] <220>
- [2698] <223> 人工序列的描述:合成多核苷酸
- [2699] <400> 243
- [2700] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60
- [2701] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120
- [2702] cagcagcagc agcag 135
- [2703] <210> 244
- [2704] <211> 132
- [2705] <212> DNA
- [2706] <213> 人工序列
- [2707] <220>
- [2708] <223> 人工序列的描述:合成多核苷酸
- [2709] <400> 244
- [2710] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60
- [2711] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120
- [2712] cagcagcagc ag 132
- [2713] <210> 245
- [2714] <211> 54
- [2715] <212> DNA
- [2716] <213> 人工序列
- [2717] <220>
- [2718] <223> 人工序列的描述:合成寡核苷酸
- [2719] <400> 245
- [2720] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcag 54
- [2721] <210> 246
- [2722] <211> 207
- [2723] <212> DNA
- [2724] <213> 人工序列
- [2725] <220>
- [2726] <223> 人工序列的描述:合成多核苷酸
- [2727] <400> 246
- [2728] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60
- [2729] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120

- [2730] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 180
- [2731] cagcagcagc agcagcagca gcagcagc 207
- [2732] <210> 247
- [2733] <211> 144
- [2734] <212> DNA
- [2735] <213> 人工序列
- [2736] <220>
- [2737] <223> 人工序列的描述:合成多核苷酸
- [2738] <400> 247
- [2739] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60
- [2740] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120
- [2741] cagcagcagc agcagcagca gcag 144
- [2742] <210> 248
- [2743] <211> 60
- [2744] <212> DNA
- [2745] <213> 人工序列
- [2746] <220>
- [2747] <223> 人工序列的描述:合成寡核苷酸
- [2748] <400> 248
- [2749] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60
- [2750] <210> 249
- [2751] <211> 123
- [2752] <212> DNA
- [2753] <213> 人工序列
- [2754] <220>
- [2755] <223> 人工序列的描述:合成多核苷酸
- [2756] <400> 249
- [2757] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60
- [2758] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120
- [2759] cag 123
- [2760] <210> 250
- [2761] <211> 9
- [2762] <212> PRT
- [2763] <213> 未知
- [2764] <220>
- [2765] <223> 未知序列的描述:未知家族肽
- [2766] <400> 250
- [2767] Leu Ala Gly Leu Ile Asp Ala Asp Gly
- [2768] 1 5

- [2769] <210> 251  
[2770] <211> 48  
[2771] <212> DNA  
[2772] <213> 人工序列  
[2773] <220>  
[2774] <223> 人工序列的描述:合成寡核苷酸  
[2775] <400> 251  
[2776] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagc 48  
[2777] <210> 252  
[2778] <211> 360  
[2779] <212> DNA  
[2780] <213> 人工序列  
[2781] <220>  
[2782] <223> 人工序列的描述:合成多核苷酸  
[2783] <400> 252  
[2784] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60  
[2785] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120  
[2786] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 180  
[2787] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 240  
[2788] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 300  
[2789] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 360  
[2790] <210> 253  
[2791] <211> 75  
[2792] <212> DNA  
[2793] <213> 人工序列  
[2794] <220>  
[2795] <223> 人工序列的描述:合成寡核苷酸  
[2796] <400> 253  
[2797] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60  
[2798] cagcagcagc agcagcag 75  
[2799] <210> 254  
[2800] <211> 108  
[2801] <212> DNA  
[2802] <213> 人工序列  
[2803] <220>  
[2804] <223> 人工序列的描述:合成多核苷酸  
[2805] <400> 254  
[2806] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60  
[2807] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagc 108

- [2808] <210> 255
- [2809] <211> 117
- [2810] <212> DNA
- [2811] <213> 人工序列
- [2812] <220>
- [2813] <223> 人工序列的描述:合成多核苷酸
- [2814] <400> 255
- [2815] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60
- [2816] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcag 117
- [2817] <210> 256
- [2818] <211> 72
- [2819] <212> DNA
- [2820] <213> 人工序列
- [2821] <220>
- [2822] <223> 人工序列的描述:合成寡核苷酸
- [2823] <400> 256
- [2824] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60
- [2825] cagcagcagc ag 72
- [2826] <210> 257
- [2827] <211> 126
- [2828] <212> DNA
- [2829] <213> 人工序列
- [2830] <220>
- [2831] <223> 人工序列的描述:合成多核苷酸
- [2832] <400> 257
- [2833] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60
- [2834] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120
- [2835] cagcag 126
- [2836] <210> 258
- [2837] <211> 87
- [2838] <212> DNA
- [2839] <213> 人工序列
- [2840] <220>
- [2841] <223> 人工序列的描述:合成寡核苷酸
- [2842] <400> 258
- [2843] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60
- [2844] cagcagcagc agcagcagca gcagcag 87
- [2845] <210> 259
- [2846] <211> 120

- [2847] <212> DNA
- [2848] <213> 人工序列
- [2849] <220>
- [2850] <223> 人工序列的描述:合成多核苷酸
- [2851] <400> 259
- [2852] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 60
- [2853] cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag 120
- [2854] <210> 260
- [2855] <211> 20
- [2856] <212> DNA
- [2857] <213> 人工序列
- [2858] <220>
- [2859] <223> 人工序列的描述:合成寡核苷酸
- [2860] <400> 260
- [2861] gcagcagcag cagcagcagc 20
- [2862] <210> 261
- [2863] <211> 21
- [2864] <212> DNA
- [2865] <213> 人工序列
- [2866] <220>
- [2867] <223> 人工序列的描述:合成寡核苷酸
- [2868] <400> 261
- [2869] cagcagcagc agcagcagca g 21

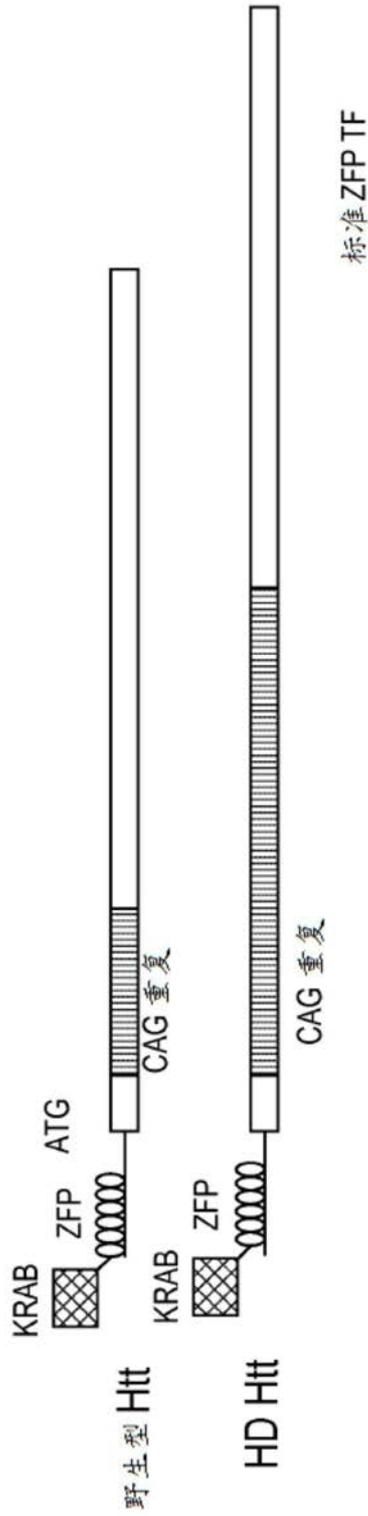


图1A

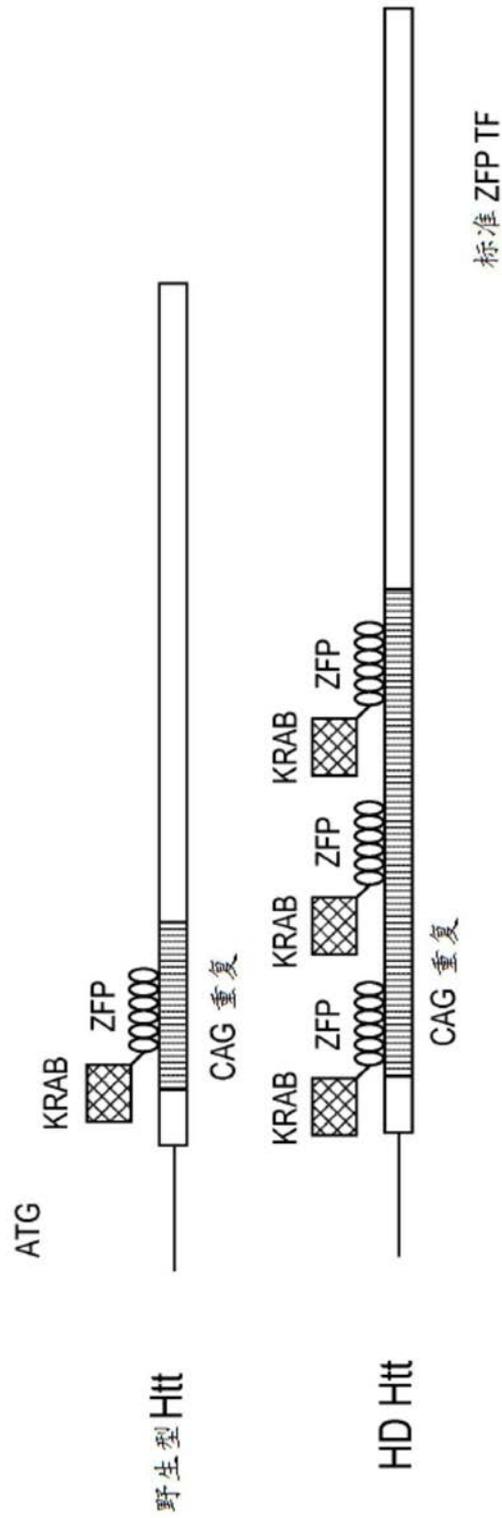


图1B

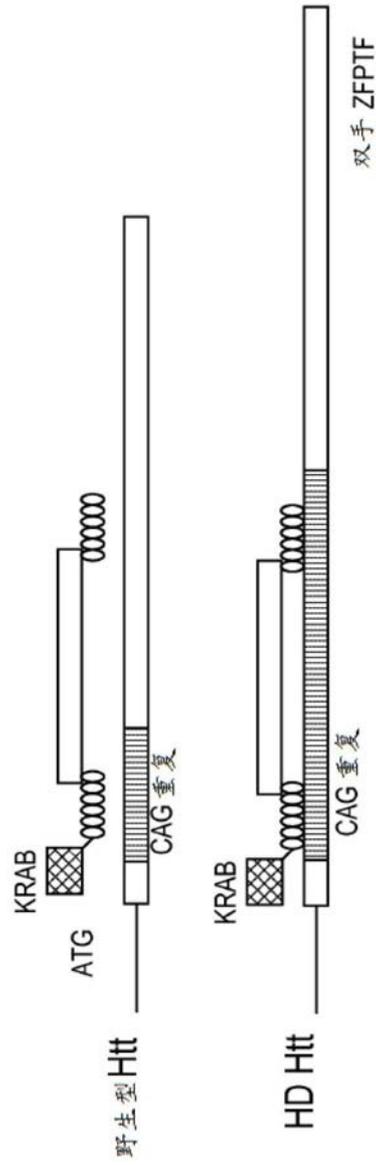


图1C

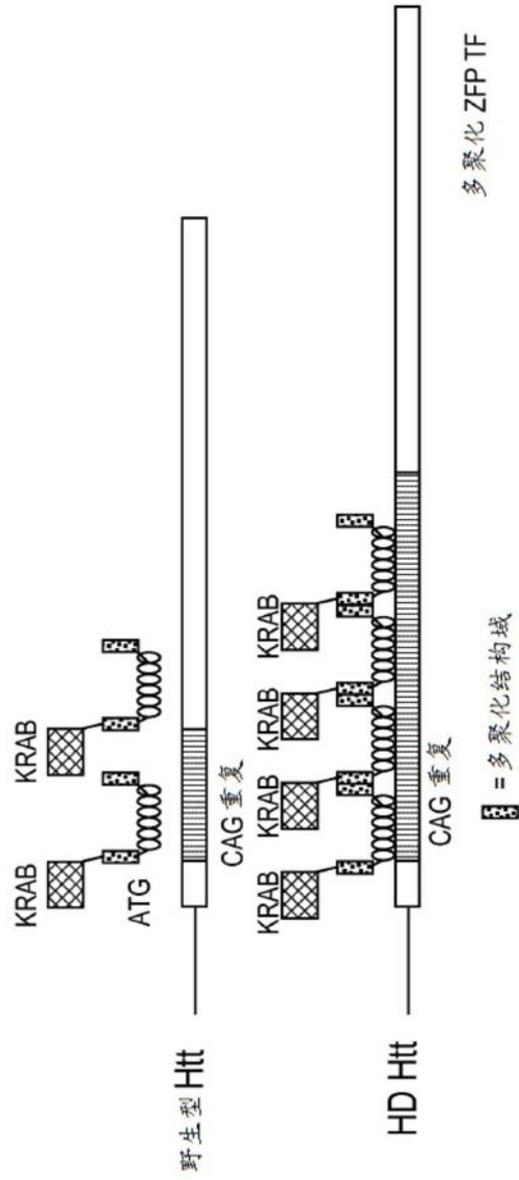


图1D

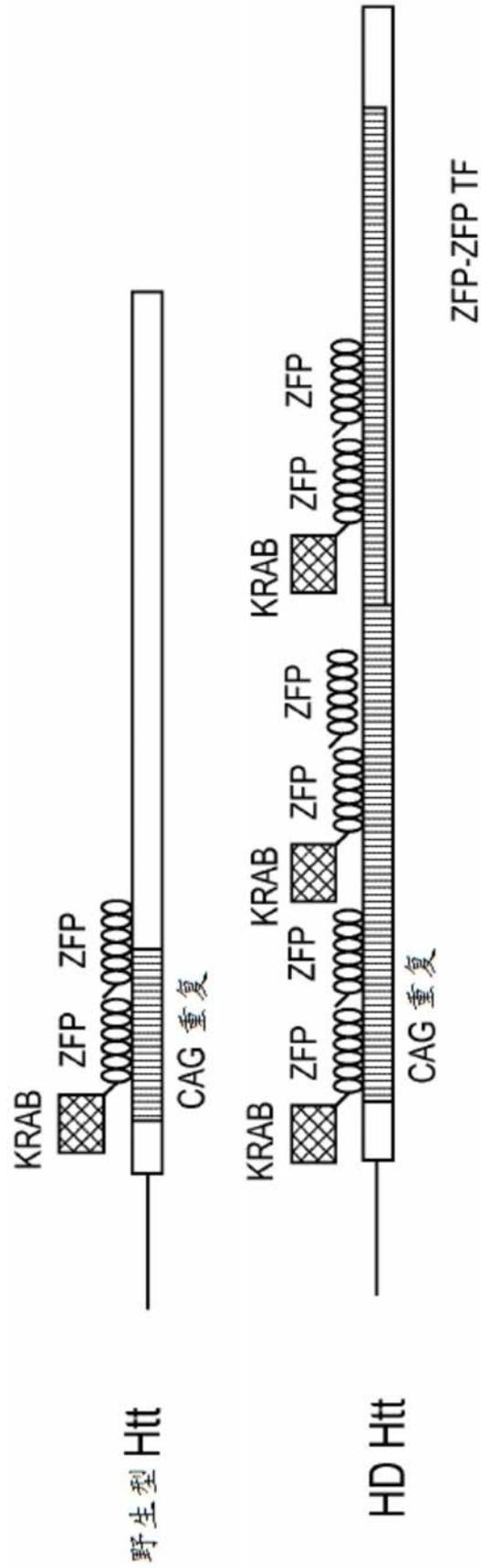


图1E

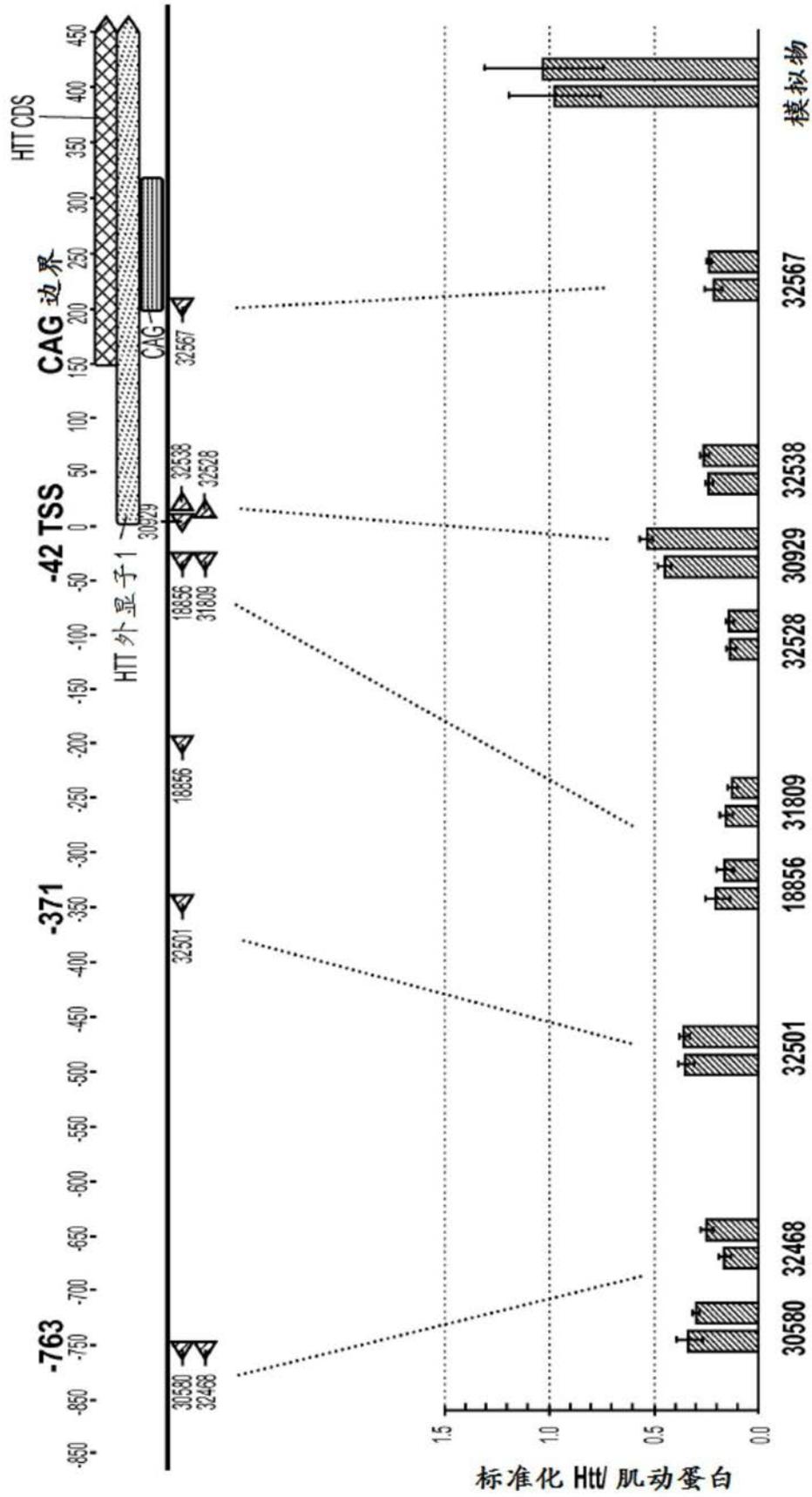


图2A

溶解产物的1:3稀释液

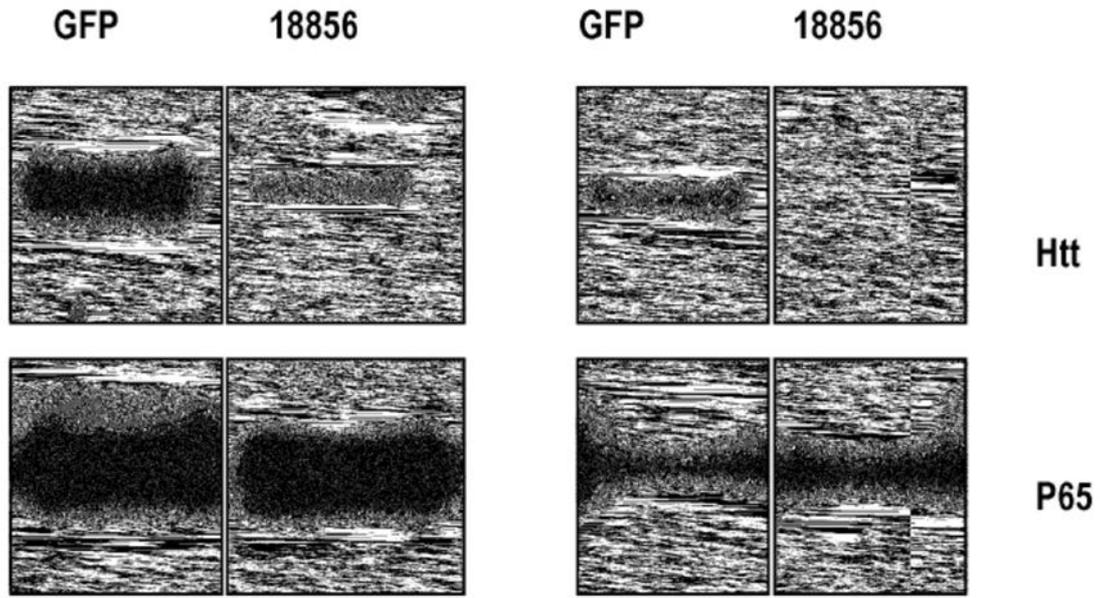


图2B

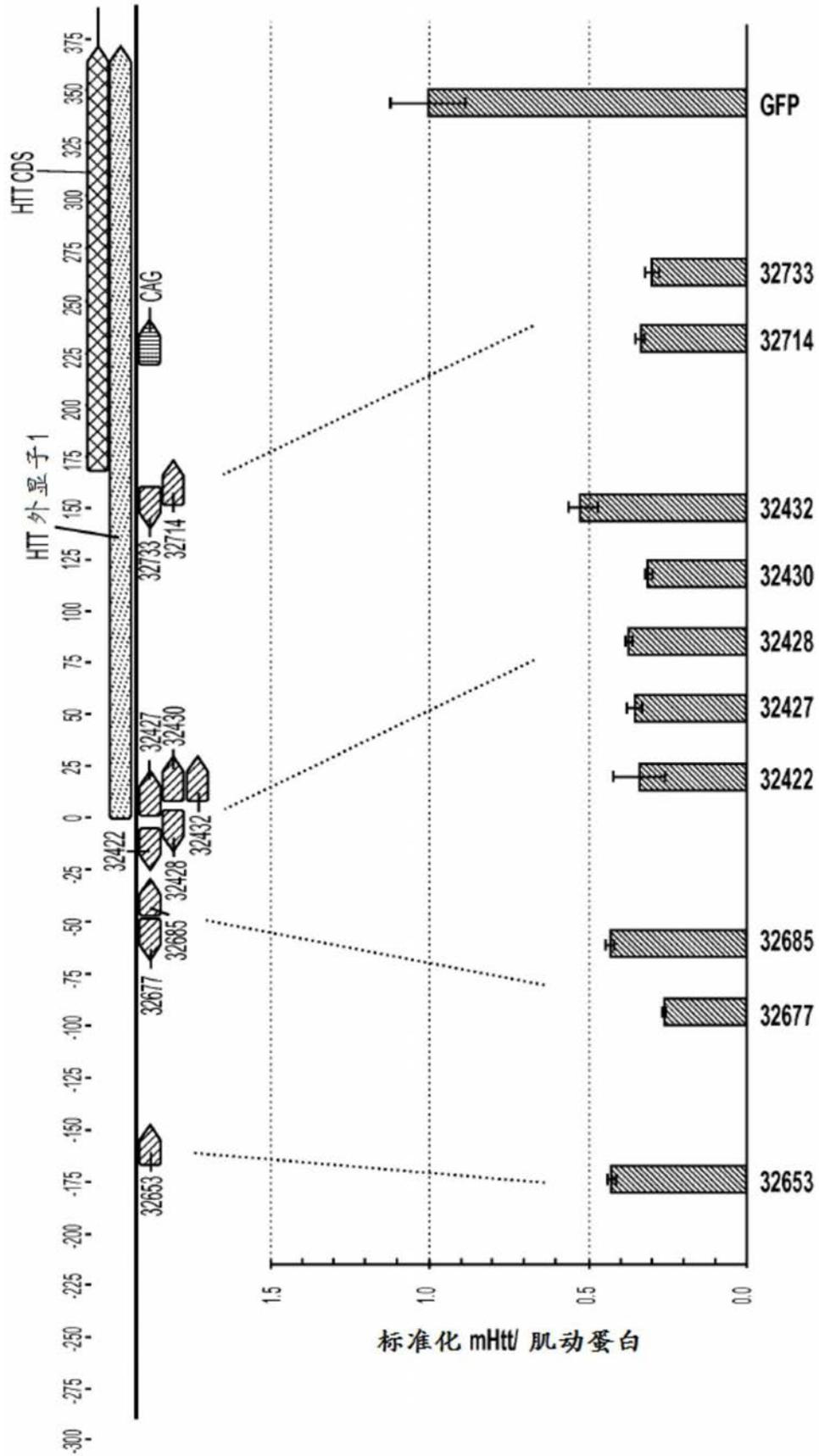


图2C

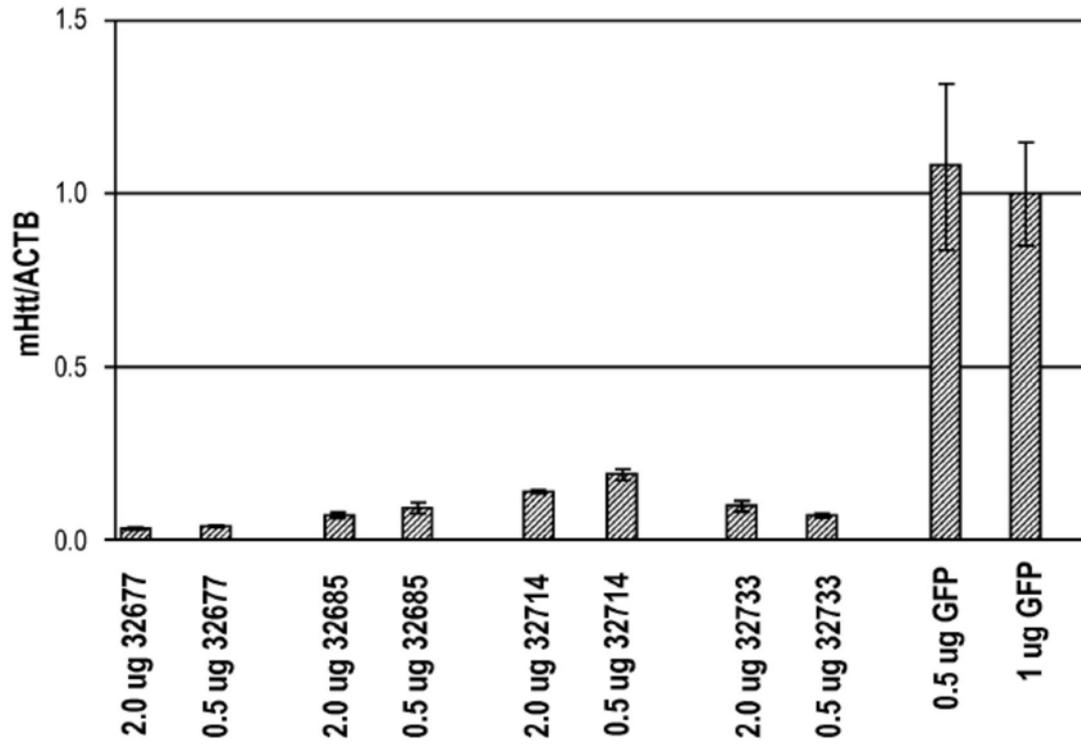


图2D

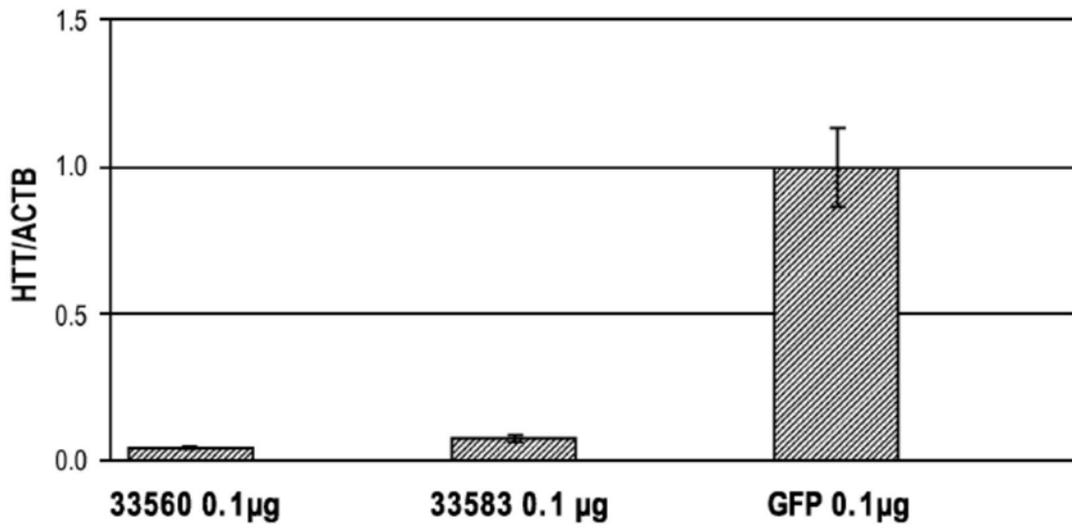


图2E

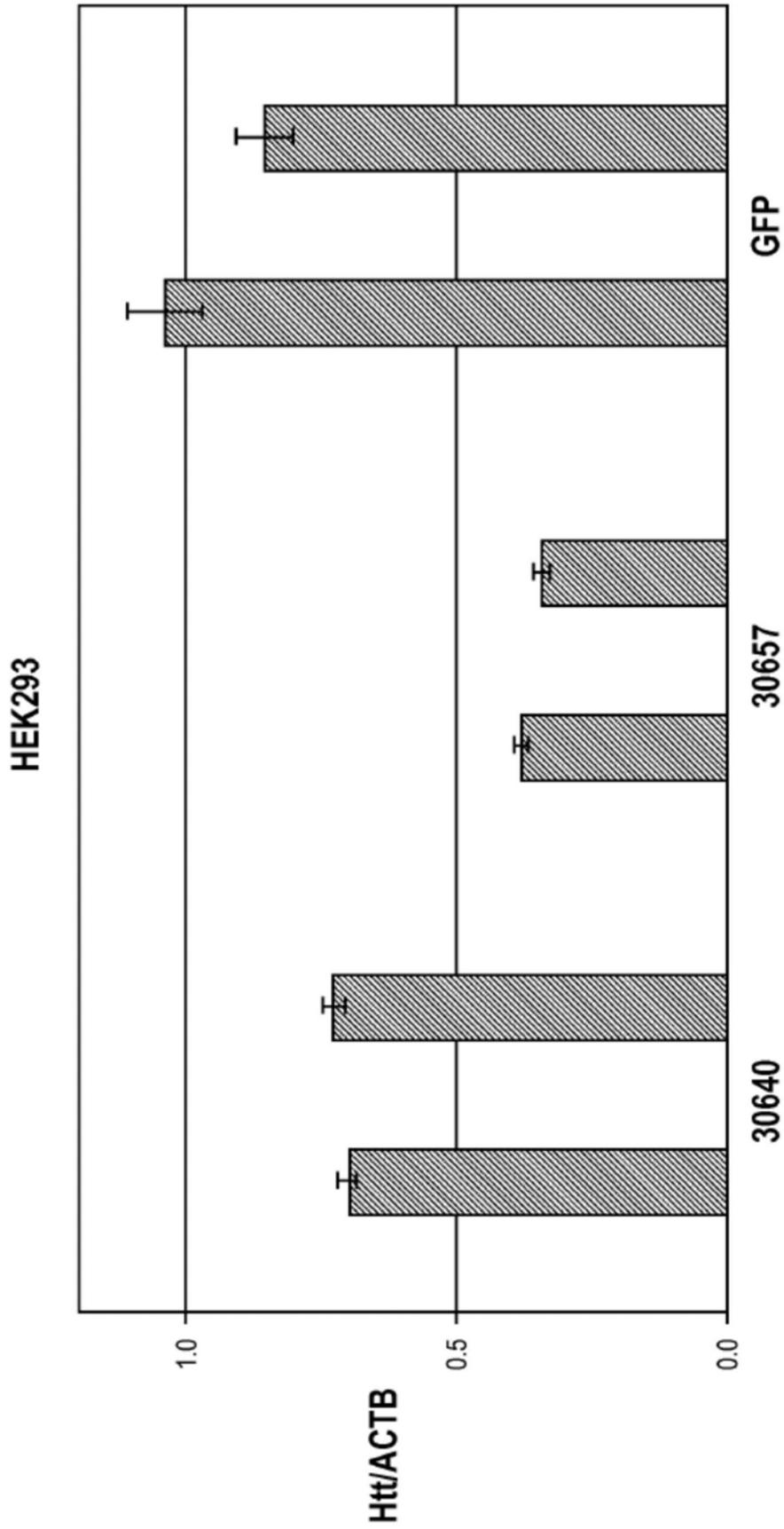


图3A

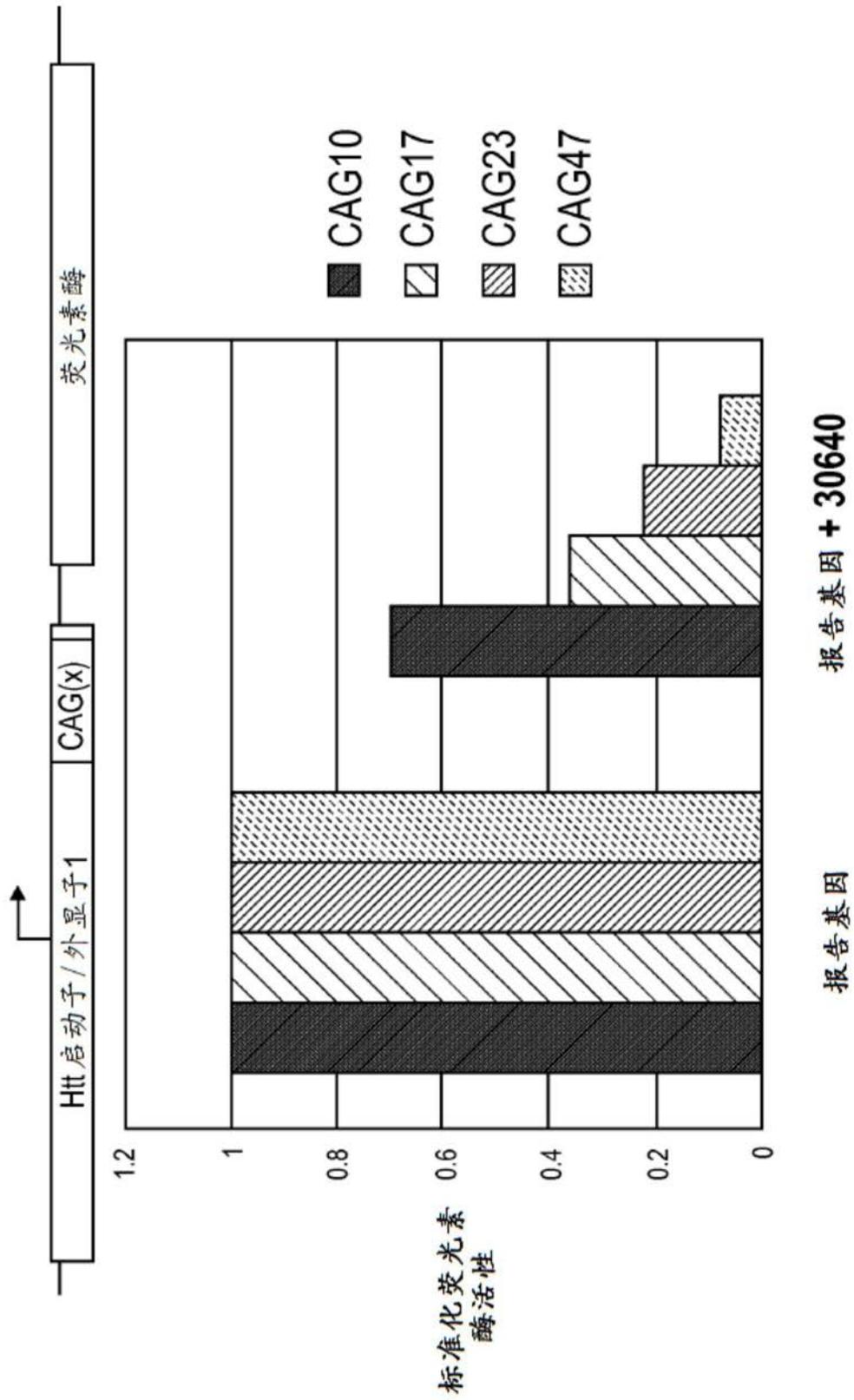


图3B

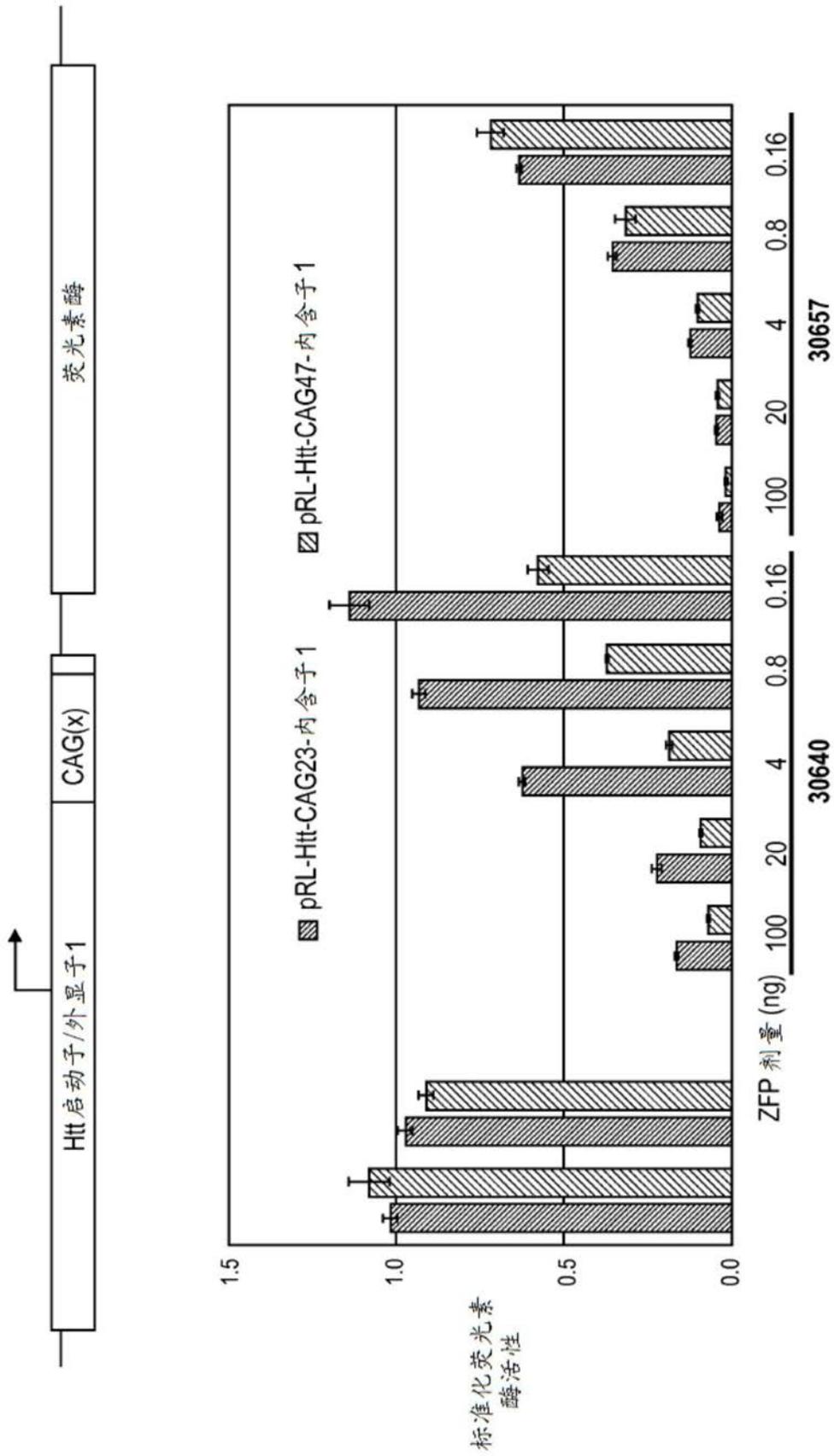


图3C

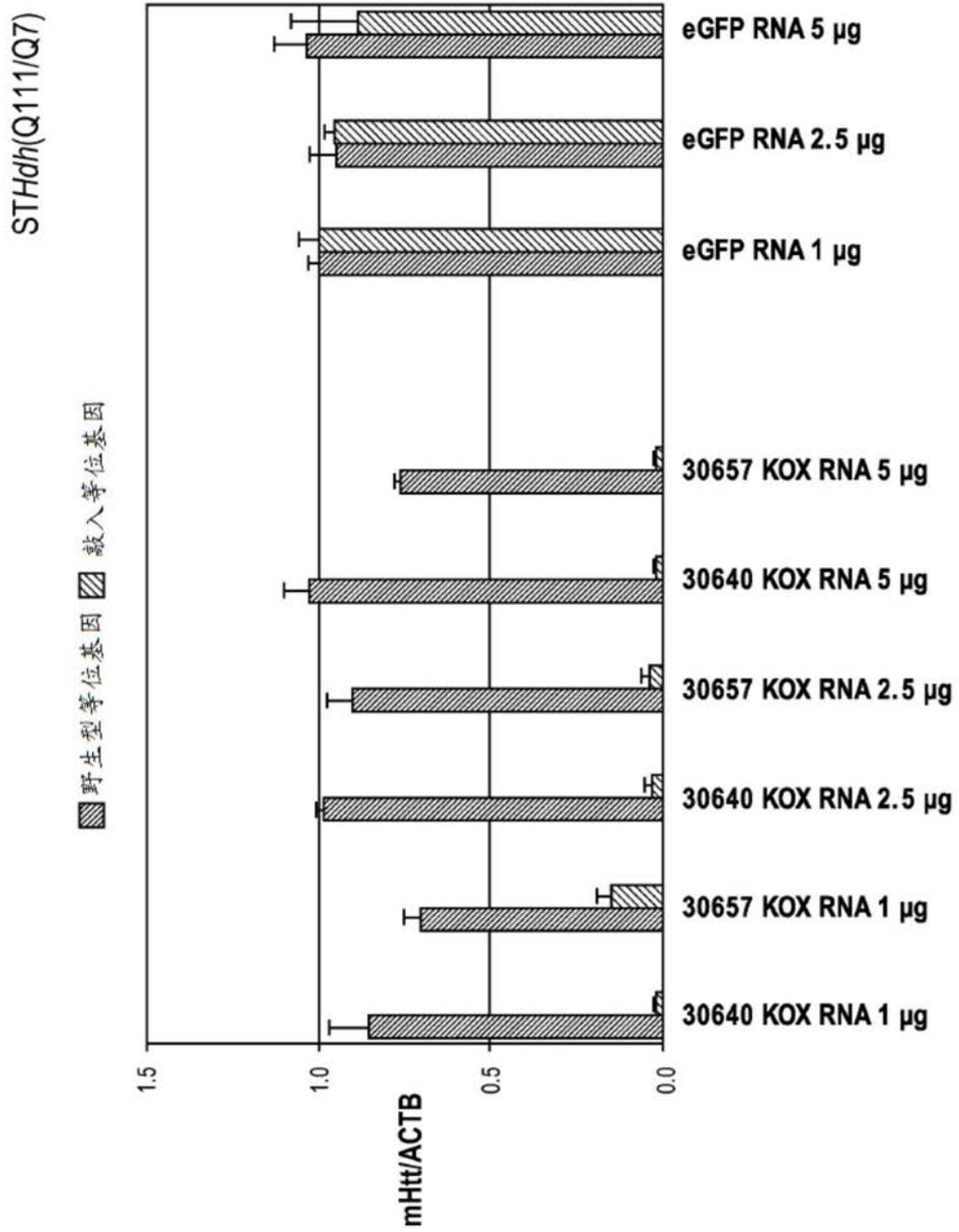


图3D

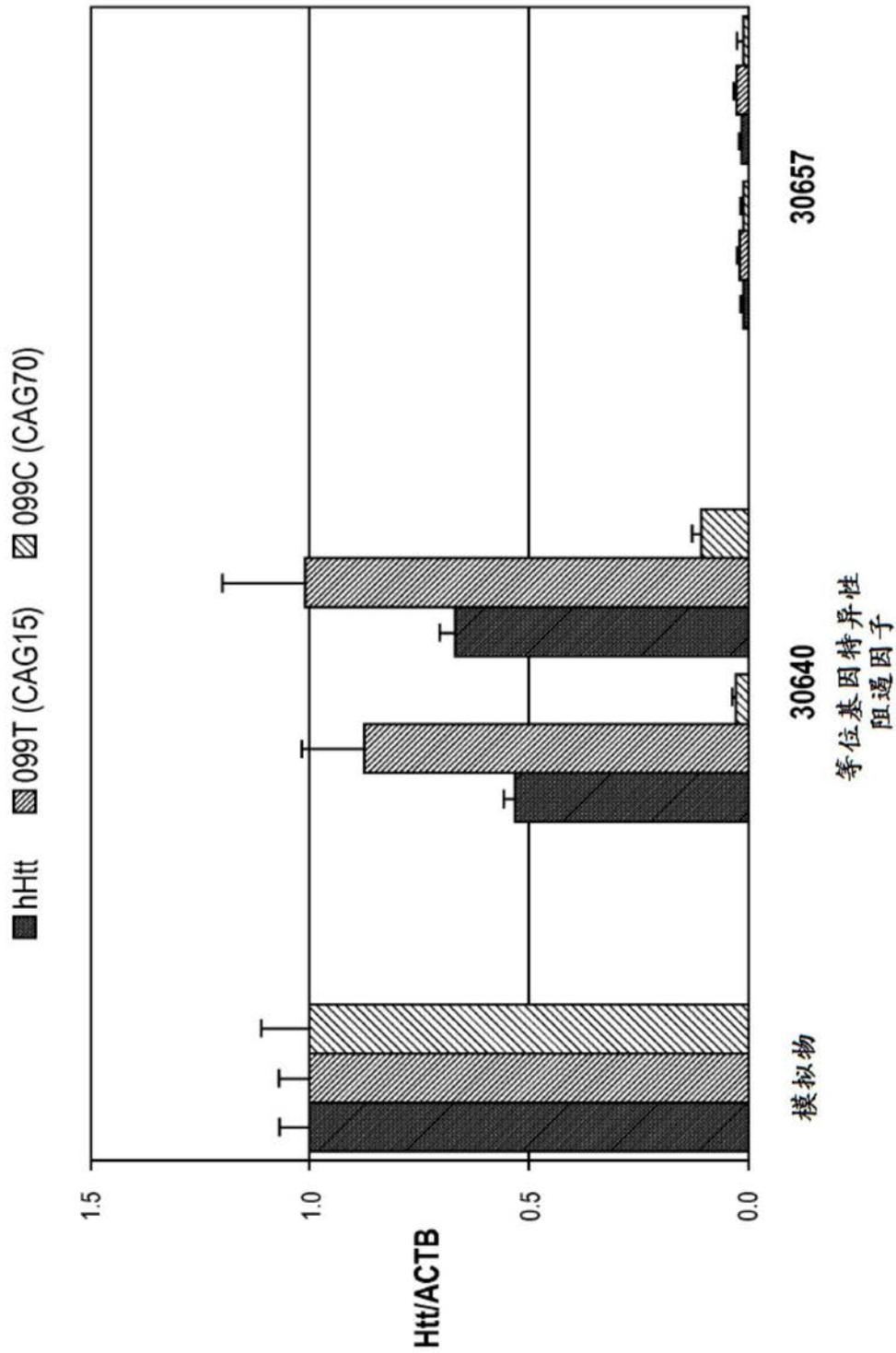


图3E

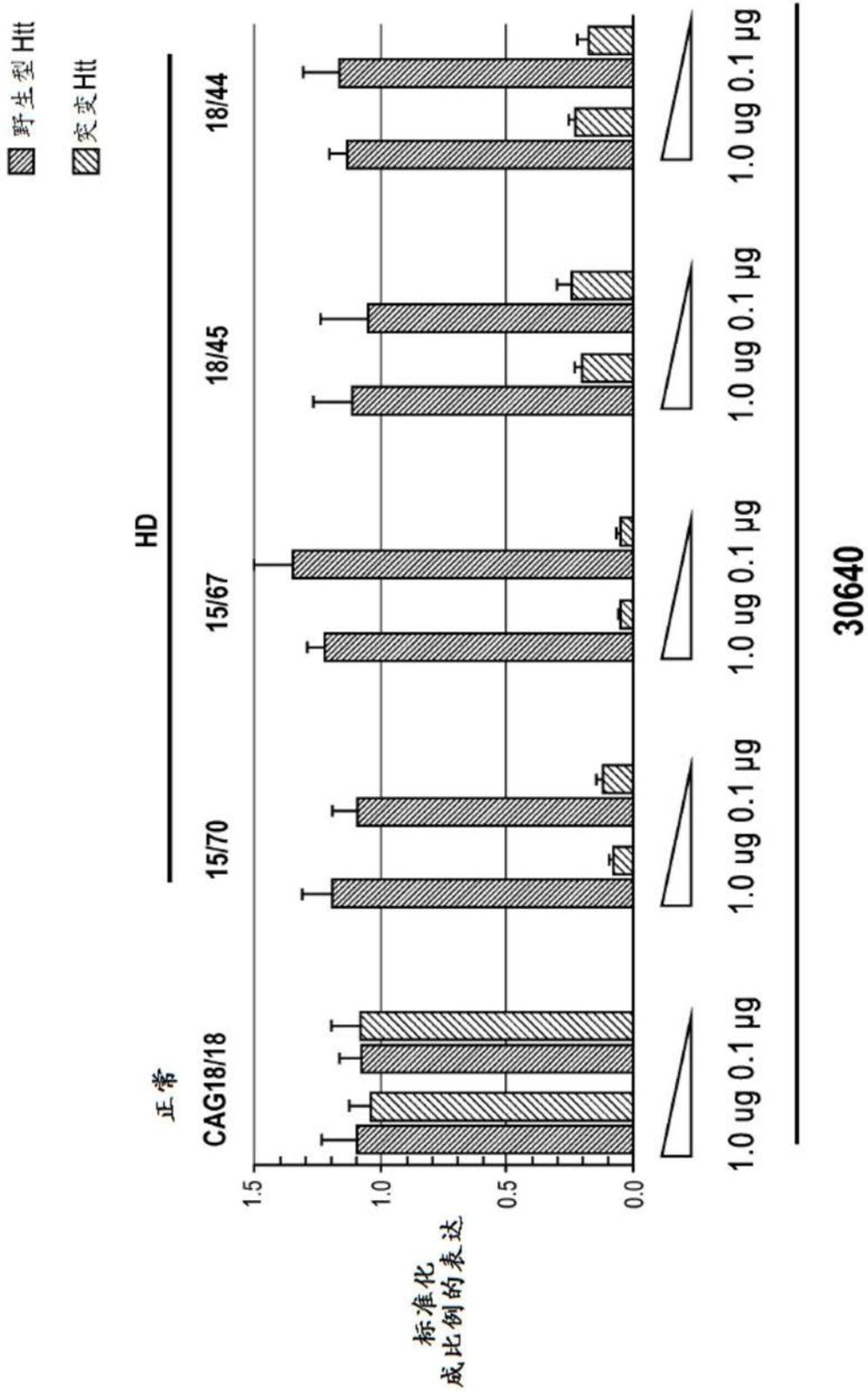


图3F

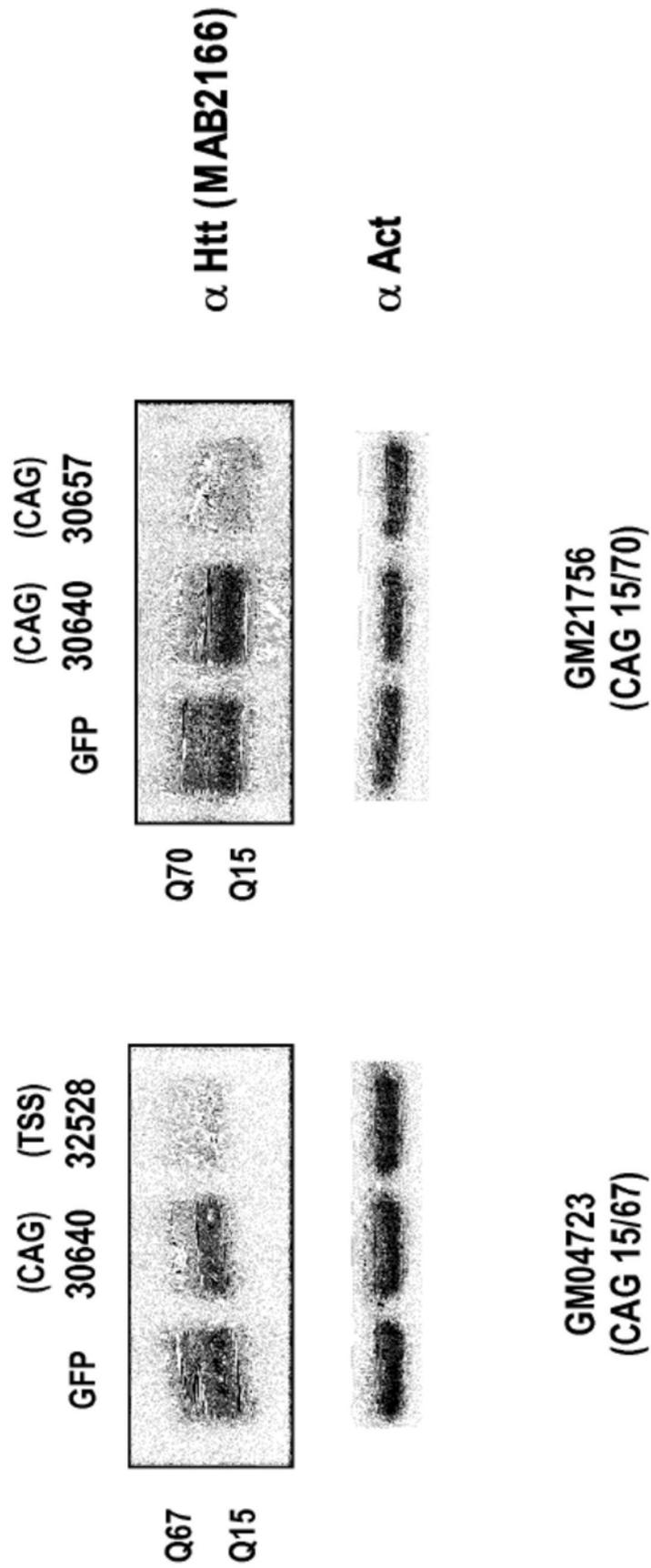


图3G

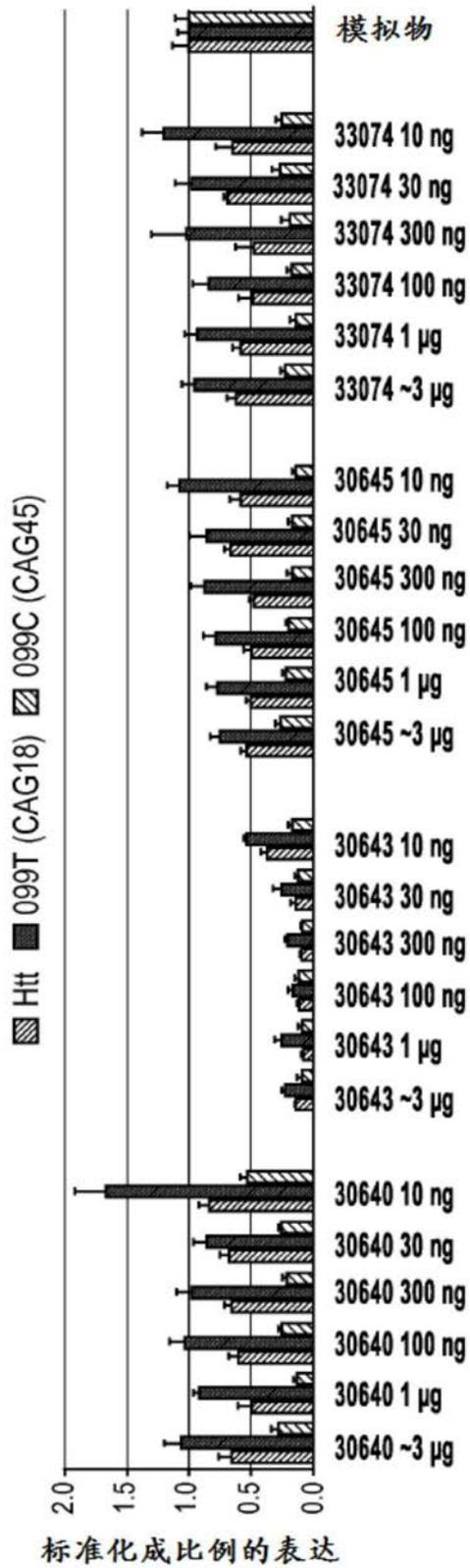


图4A

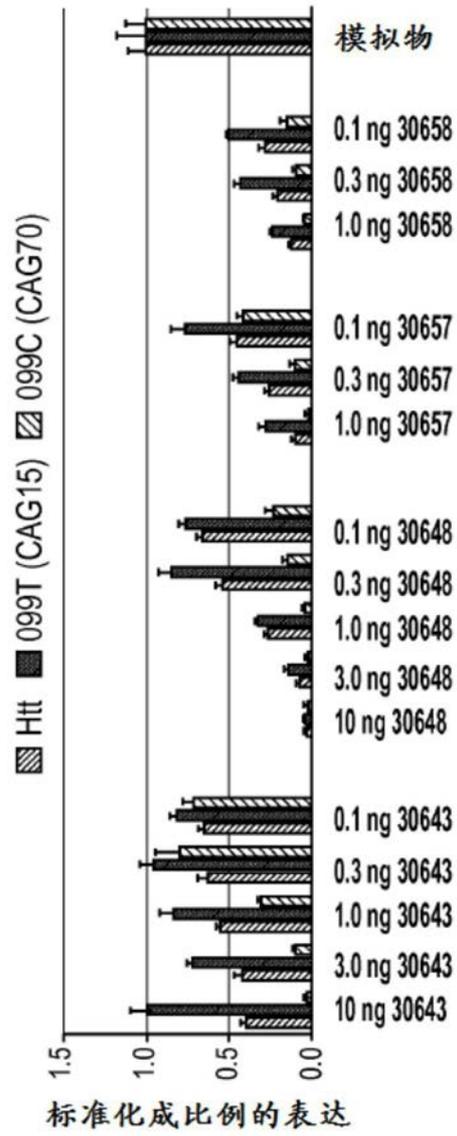


图4B

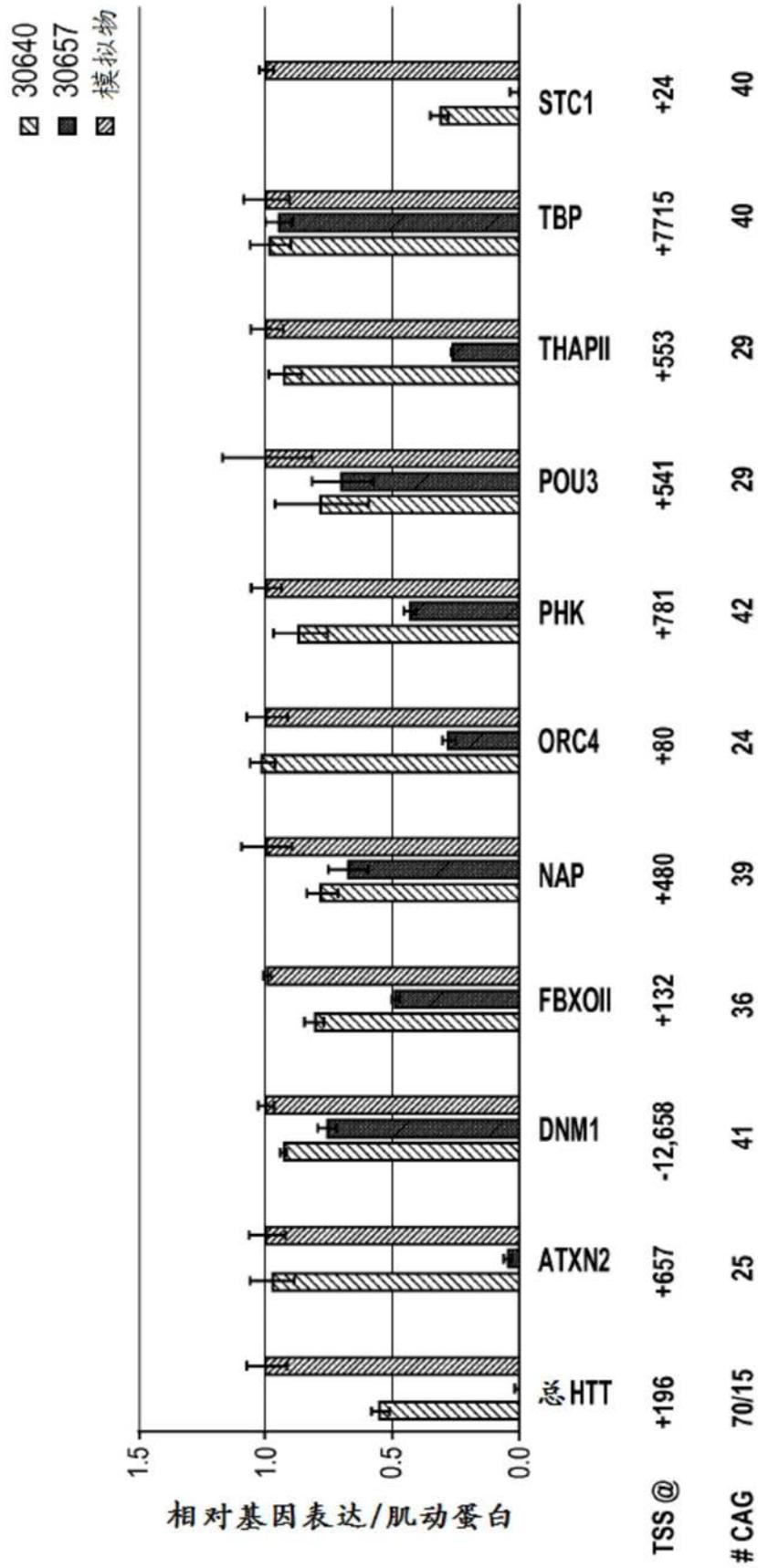


图5

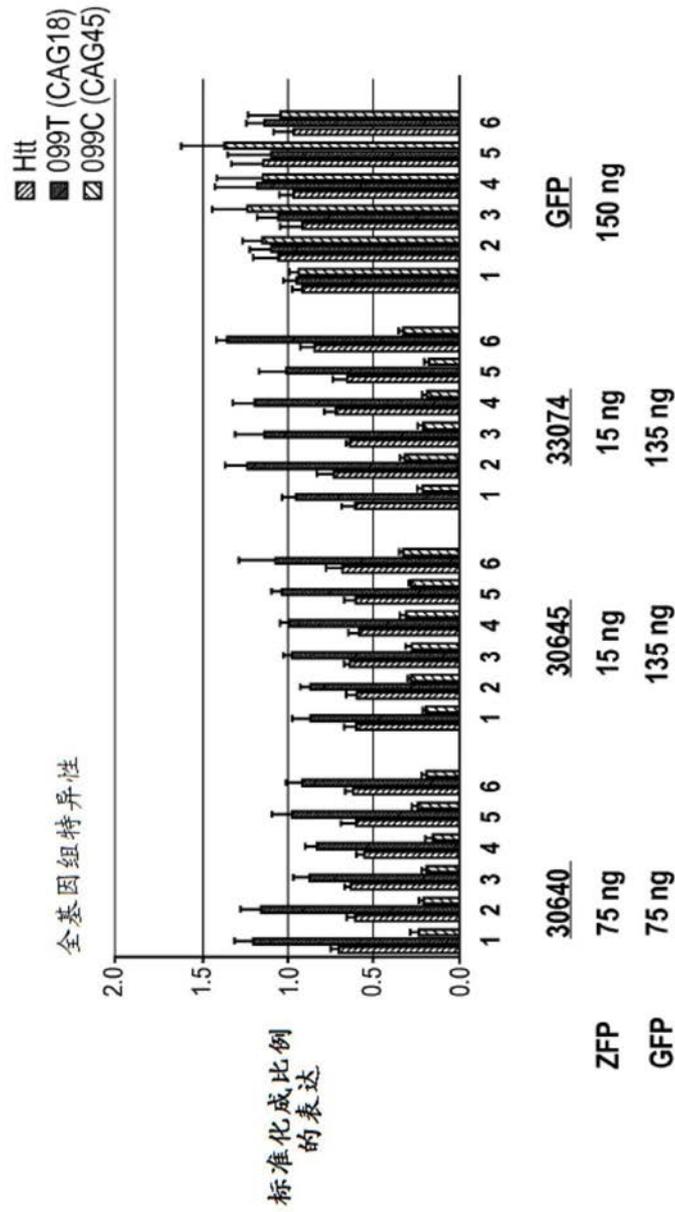


图6A

ZFP	# 下调基因的数量	# 上调基因的数量	试验的细胞系
30640	2	0	HD 成纤维细胞
30645	1	0	HD 成纤维细胞
33074	1	0	HD 成纤维细胞

图6B

在NSC中的阻遏

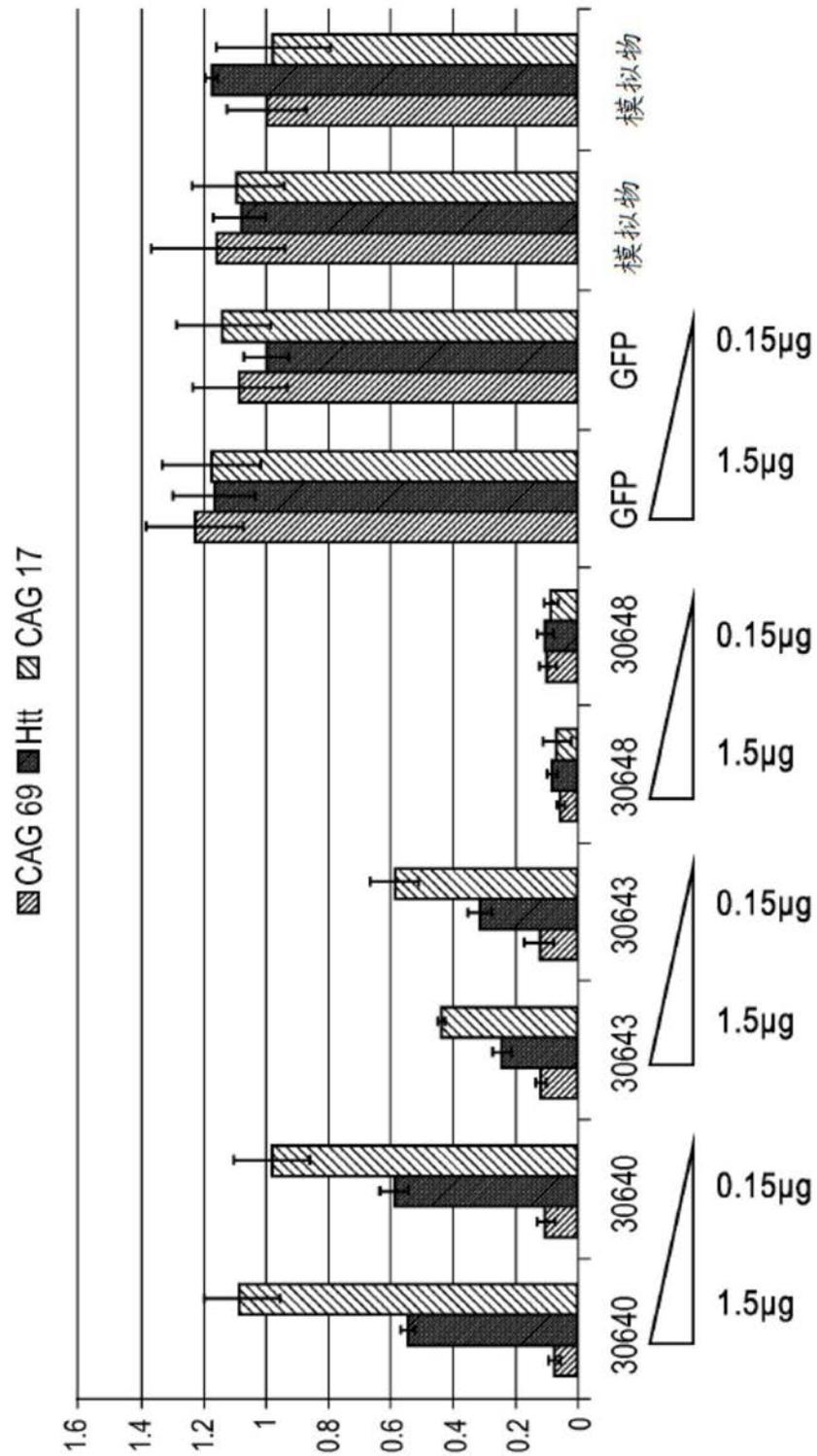


图7

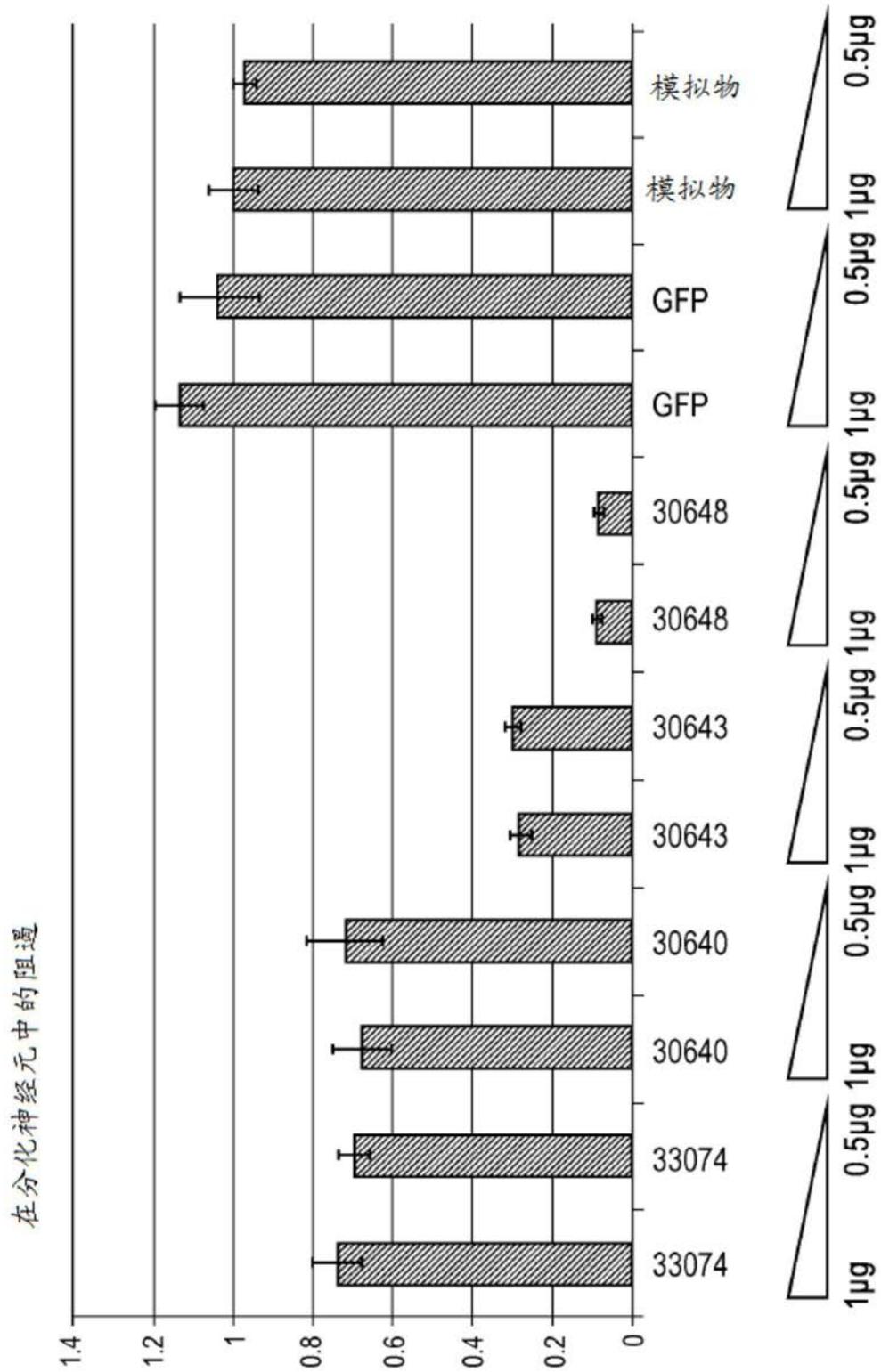


图8

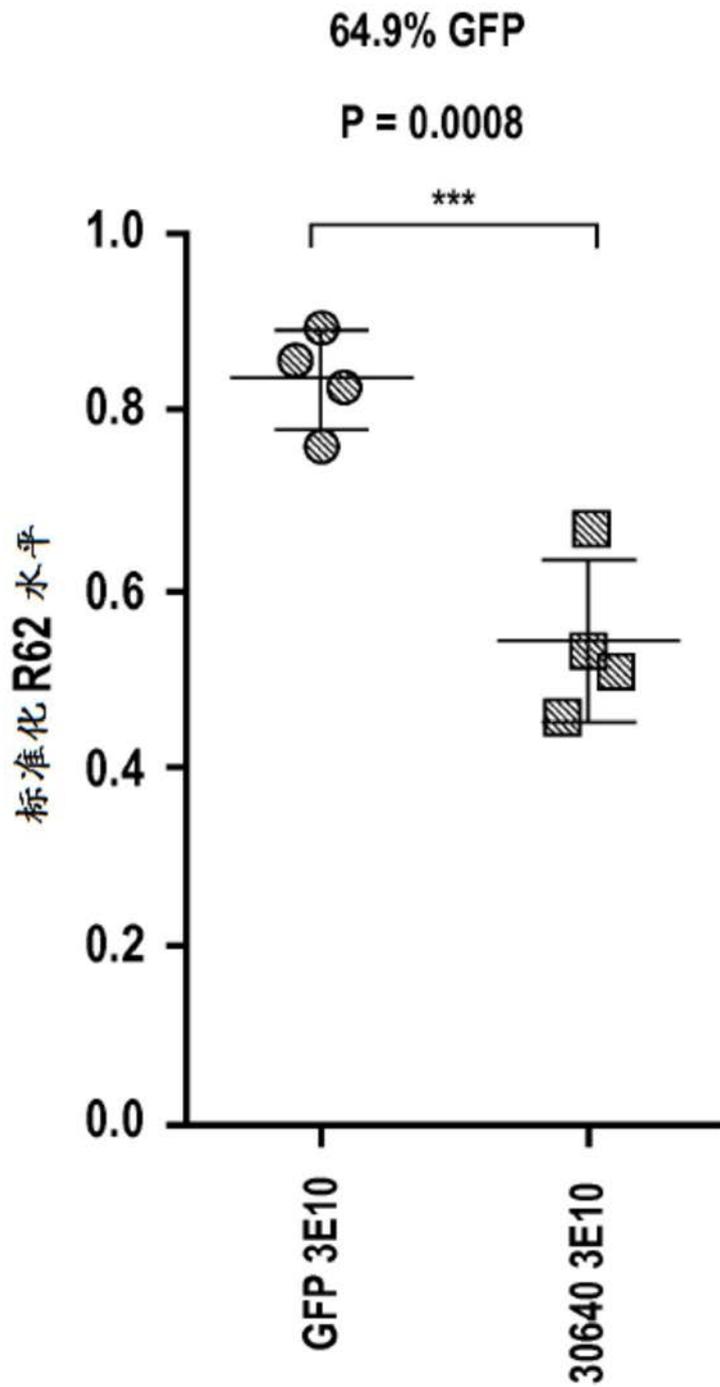


图9

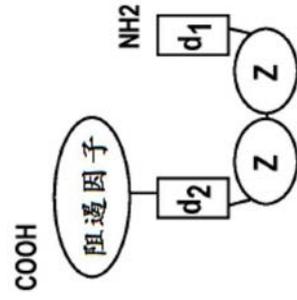


图10A

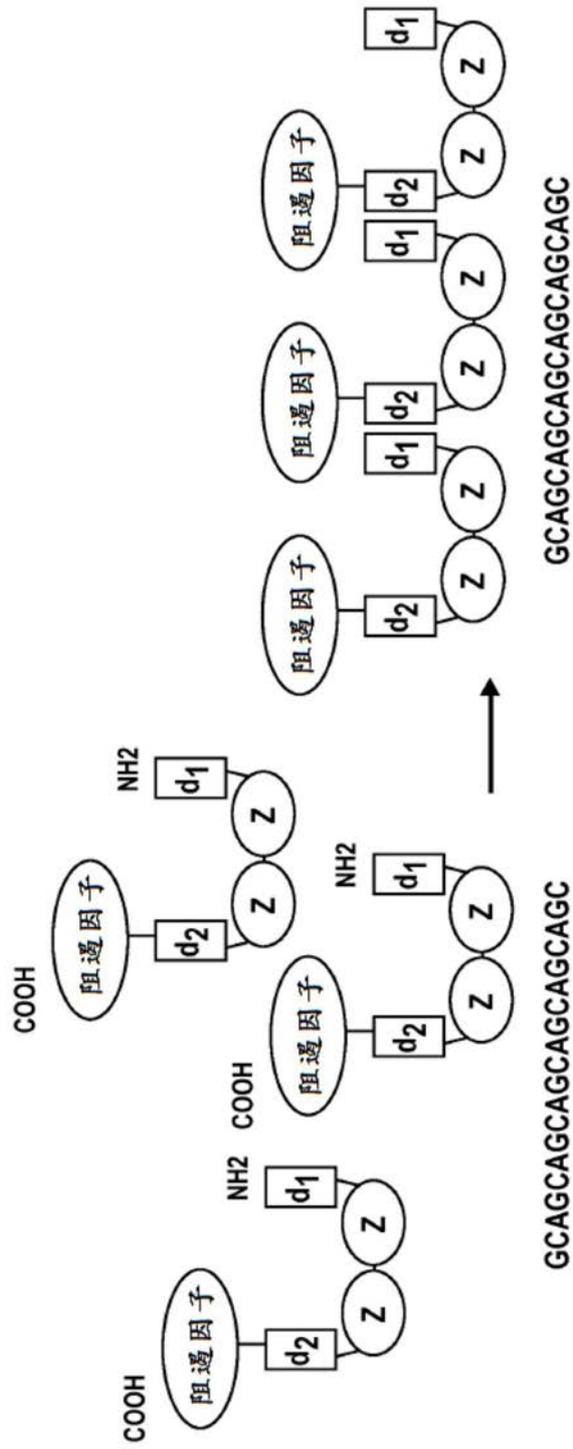


图10B

DZ #1:  
MAPKKRKGIGVLRGAATKCVHCGIVFLDEVMYALHMSCHGFRDPFECNICGYHSQDRYEFSSHIVRGEHSGVP [ZFP]  
GSGGTKCVHCGIVFLDEVMYALHMSCHGFRDPFECNICGYHSQDRYEFSSHIVRGEHLRQKDAARSRSGMDAKSLTAWSR  
LVTFKDVVDFFTREEWKLLDTAQOIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEPEPWLVEREIQETHPDSETAFEIKSSVDYKDDDDK

DZ #2  
MAPKKRKGIGVLRGAAFKCEHCRILFLDHVMFTIHMCGHGFDPFKCNMCKGKCDGVPVGLFVHMARNAHSGVP [ZFP]  
GSGGTKCVHCGIVFLDEVMYALHMSCHGFRDPFECNICGYHSQDRYEFSSHIVRGEHLRQKDAARSRSGMDAKSLTAWSR  
LVTFKDVVDFFTREEWKLLDTAQOIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEPEPWLVEREIQETHPDSETAFEIKSSVDYKDDDDK

DZ #3  
MAPKKRKGIGVLRGAAFKCEHCRILFLDHVMFTIHMCGHGFDPFKCNMCKGKCDGVPVGLFVHMARNAHSGVP [ZFP]  
GSGGHCHCHCDMYFADNILYTIHMCGHYENPFECNICGYHSQDRYEFSSHIVRGEHLRQKDAARSRSGMDAKSLTAWSR  
TREEWKLLDTAQOIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEPEPWLVEREIQETHPDSETAFEIKSSVDYKDDDDK

DZ #4  
MAPKKRKGIGVLRGAAHCHCHCDMYFADNILYTIHMCHSCDDVFKCNMCKGKCDGVPVGLFVHMARNAHGKPTKCVHCGIVFLD  
EVMYALHMSCHGFRDPFECNICGYHSQDRYEFSSHIVRGEHSGVP [ZFP] GSGGFKCEHCRILFLDHVMFTIHMCGHGFDPFKCNM  
GEKCDGVPVGLFVHMARNAHGKPFYCEHCEITFRDVVMYSLHKGYHGFRDPFECNICGYHSQDRYEFSSHIVRGEHLRQKDAARSRSGM  
DAKSLTAWSR  
TLVTFKDVVDFFTREEWKLLDTAQOIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEPEPWLVEREIQETHPDSETAF  
EIKSSVDYKDDDDK

图10C

CC#1:  
MAPKKRKVGIHGVLRGAAGAQLEKELQALEKKLAOLEWENOALEKELAOGGSGVP[ZFP]GSGAOLKKKLOANKKELAOLKWKLQALK  
KLAOGGLRQKDAARSRSGMDAKSLTAWSRTLVTFKDVFVDFTREEWKLLDTAOOIVYRNVMLENYKNLVSLGYOLTKPDVILRLEKGEEP  
WLVEREIHOETHPDSETAFEIKSSVDYKDDDDK

CC#2:  
MAPKKRKVGIHGVLRGAAGAQLEKELQALEKKLAOLEWKNOALEKKLAOGGSGVP[ZFP]GSGALKKELAOLKWELOALK  
ELAOGGLRQKDAARSRSGMDAKSLTAWSRTLVTFKDVFVDFTREEWKLLDTAOOIVYRNVMLENYKNLVSLGYOLTKPDVILRLEKGEEP  
LVEREIHOETHPDSETAFEIKSSVDYKDDDDK

CC#3:  
MAPKKRKVGIHGVLRGAAGAQLEKELQALEKKLAOLEWKNOALEKKLAOGGSGVP[ZFP]GSGELOANKKELAOLKWELOALKKELAOGGL  
RQKDAARSRSGMDAKSLTAWSRTLVTFKDVFVDFTREEWKLLDTAOOIVYRNVMLENYKNLVSLGYOLTKPDVILRLEKGEEPWLVEREI  
HOETHPDSETAFEIKSSVDYKDDDDK

CC#4:  
MAPKKRKVGIHGVLRGAAGAQLEKELQALEKKLAOLEWKNOALEKKLAOGGSGVP[ZFP]GSGOANKKELAOLKWELOALKKELAOGGLRQKD  
AARSRSGMDAKSLTAWSRTLVTFKDVFVDFTREEWKLLDTAOOIVYRNVMLENYKNLVSLGYOLTKPDVILRLEKGEEPWLVEREIHOE  
THPDSETAFEIKSSVDYKDDDDK.

图10D

CC2-ZFP-KOX:  
MAPKKRKKVGHGVLAAAGGEOLEKKLQALEWKNQALEKKLAQGGGVP [ZFP] GSGGALKKELOANKKELAQLKWELQALKKELAQGS  
GMDAKSLTAWSRRLVTFKDVFDFTREWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEPPWLVEREIHQETHPDSETAFEI  
KSSVDYKDDDDK

CC5-ZFP-KOX:  
MAPKKRKKVGHGVLAAAGGEOLEKKLQALEWKNQALEKKLAQGV [ZFP] GSGGALKKELOANKKELAQLKWELQALKKELAQGS  
GMDAKSLTAWSRRLVTFKDVFDFTREWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEPPWLVEREIHQETHPDSETAF  
EIKSSVDYKDDDDK

CC6-ZFP-KOX:  
MAPKKRKKVGHGVLAAAGGEOLEKKLQALEWKNQALEKKLAQGGGVP [ZFP] GSALKKELOANKKELAQLKWELQALKKELAQGS  
GMDAKSLTAWSRRLVTFKDVFDFTREWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEPPWLVEREIHQETHPDSETAF  
EIKSSVDYKDDDDK

CC7-ZFP-KOX:  
MAPKKRKKVGHGVLAAAGGEOLEKKLQALEWKNQALEKKLAQGV [ZFP] GSALKKELOANKKELAQLKWELQALKKELAQGS  
GMDAKSLTAWSRRLVTFKDVFDFTREWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEPPWLVEREIHQETHPDSETA  
FEIKSSVDYKDDDDK

图10D (续)

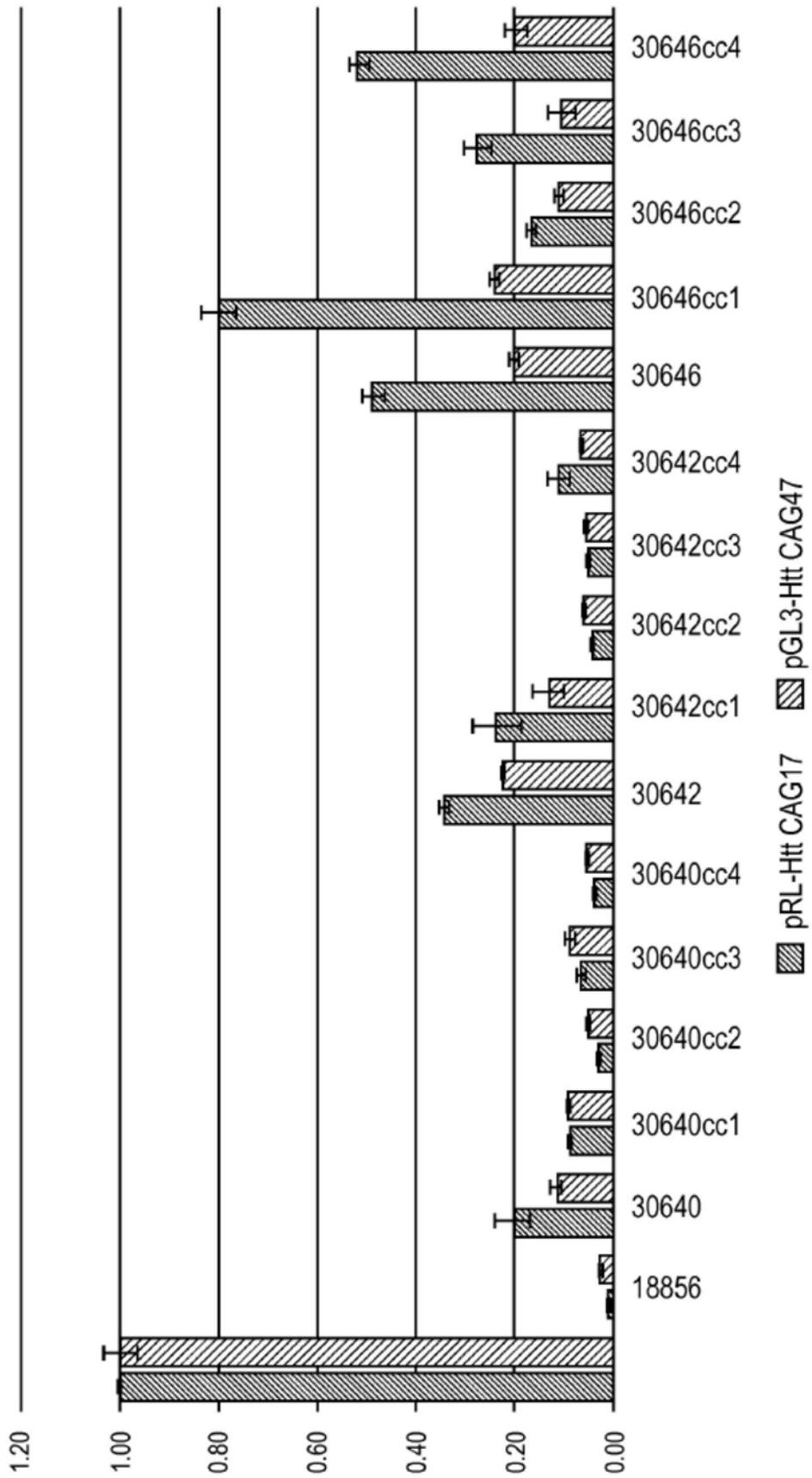


图11A

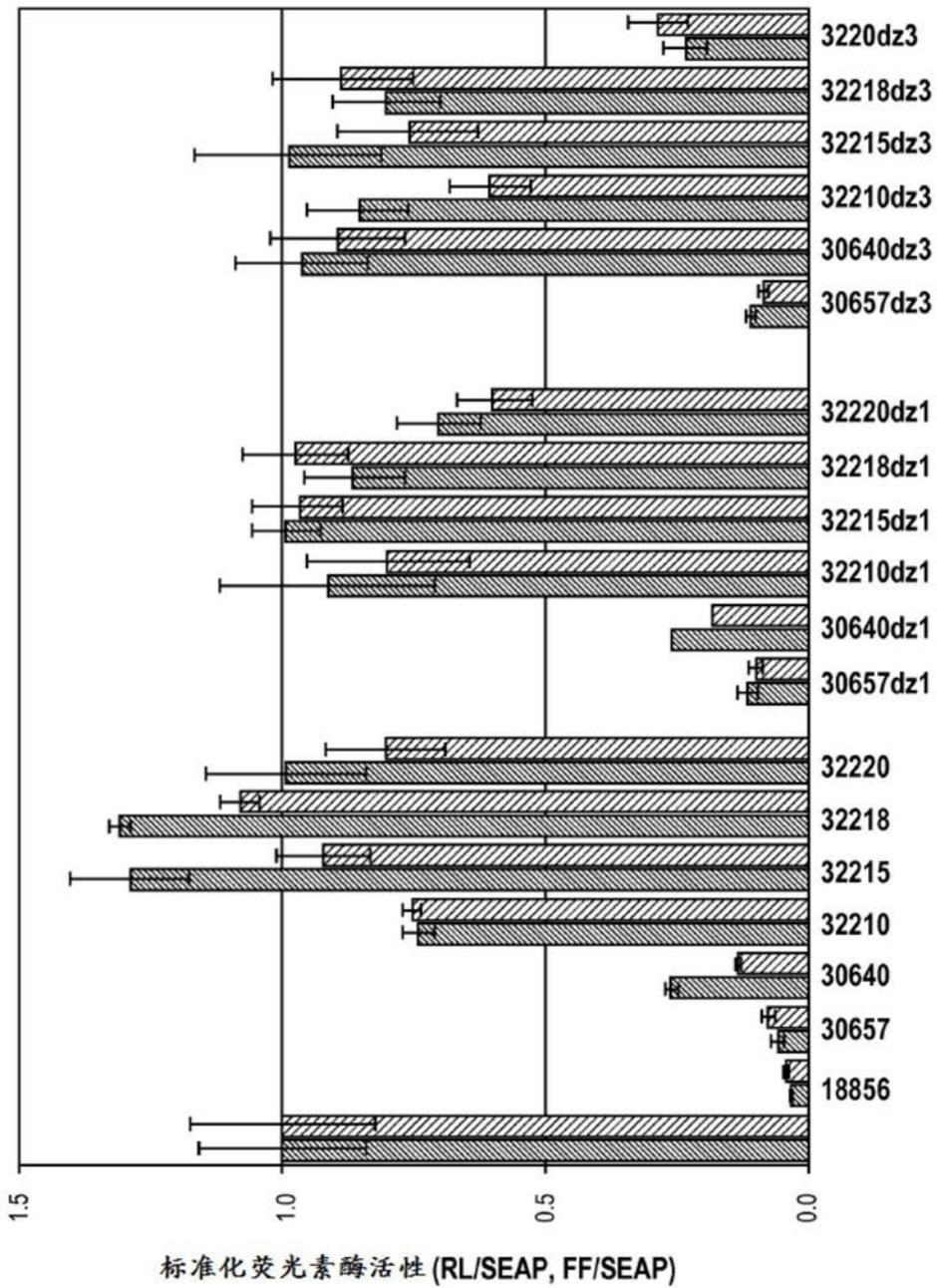


图11B

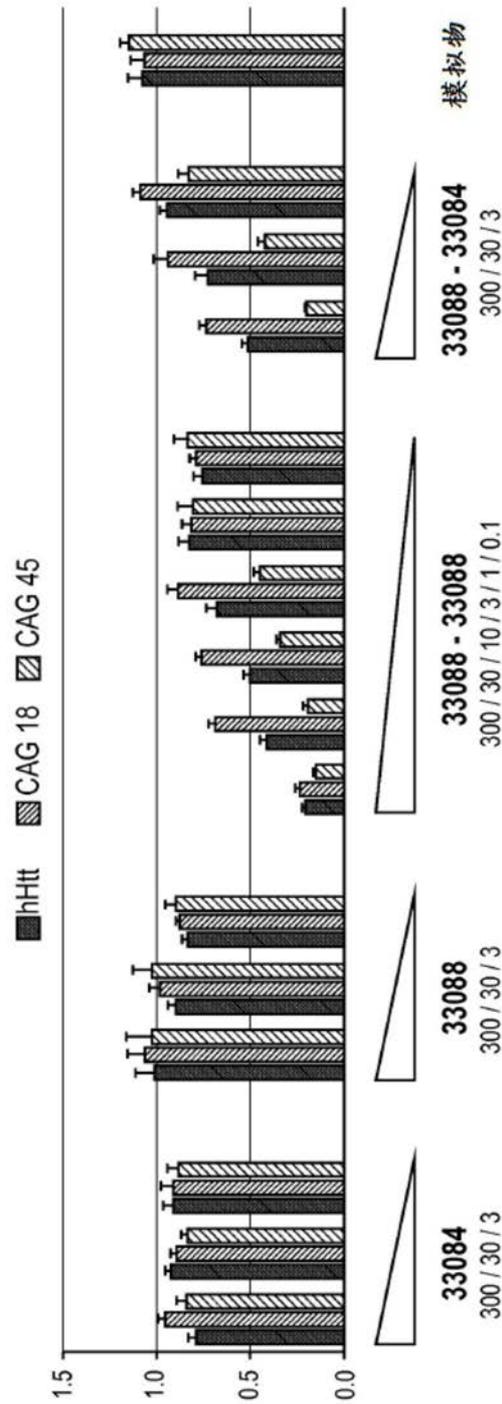


图12A

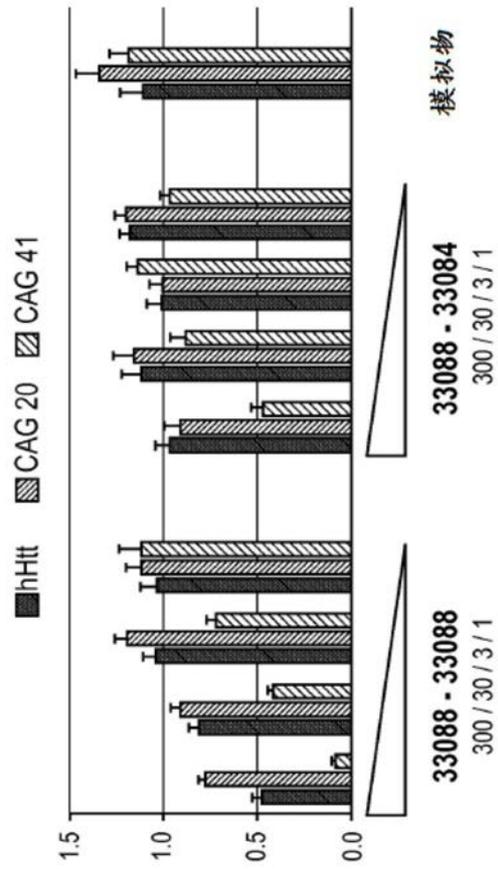


图12B

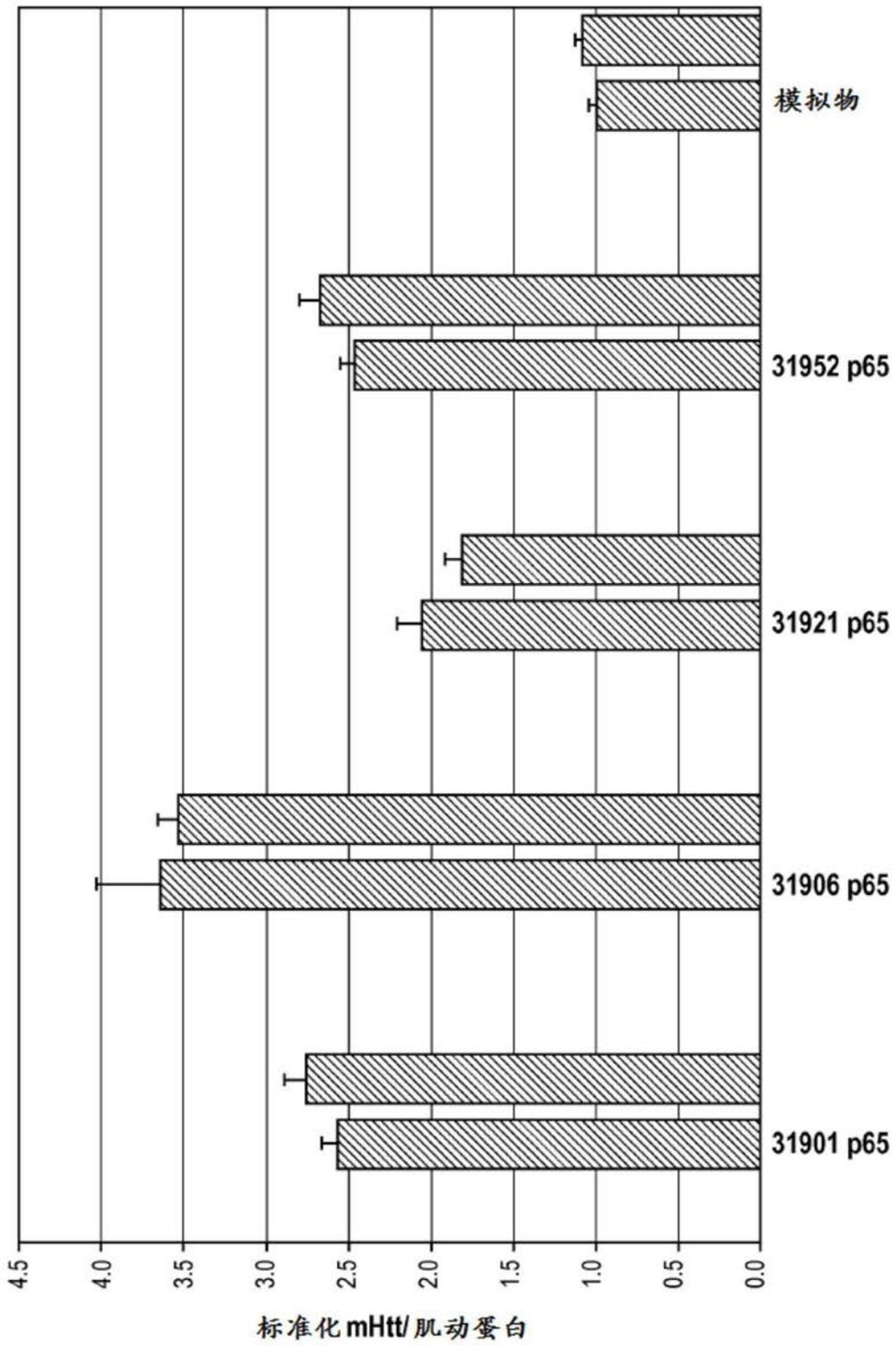


图13A

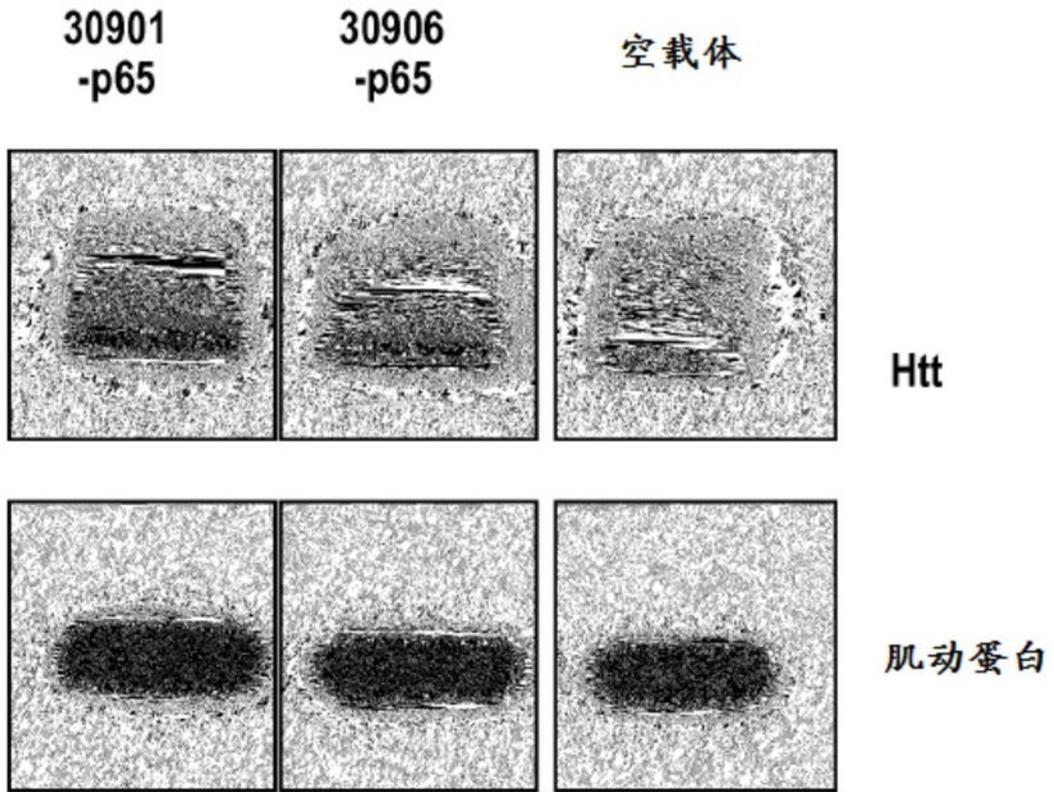


图13B

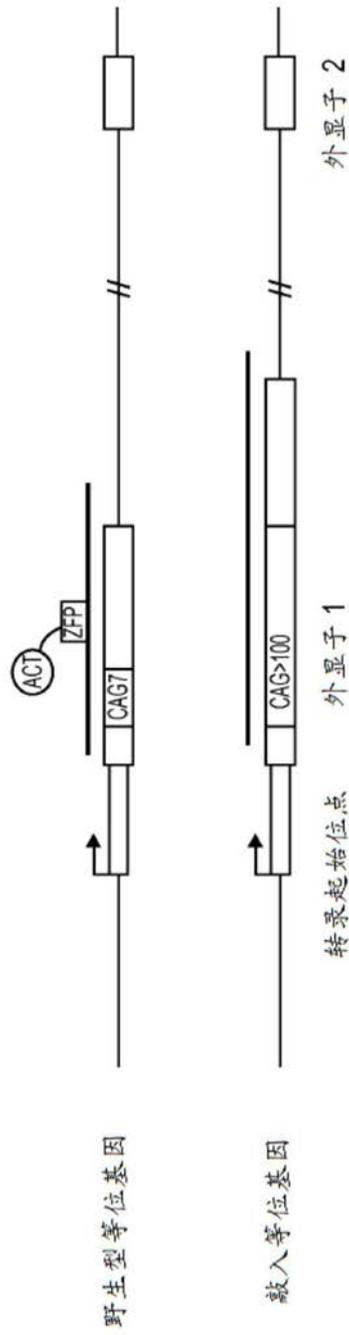


图13C



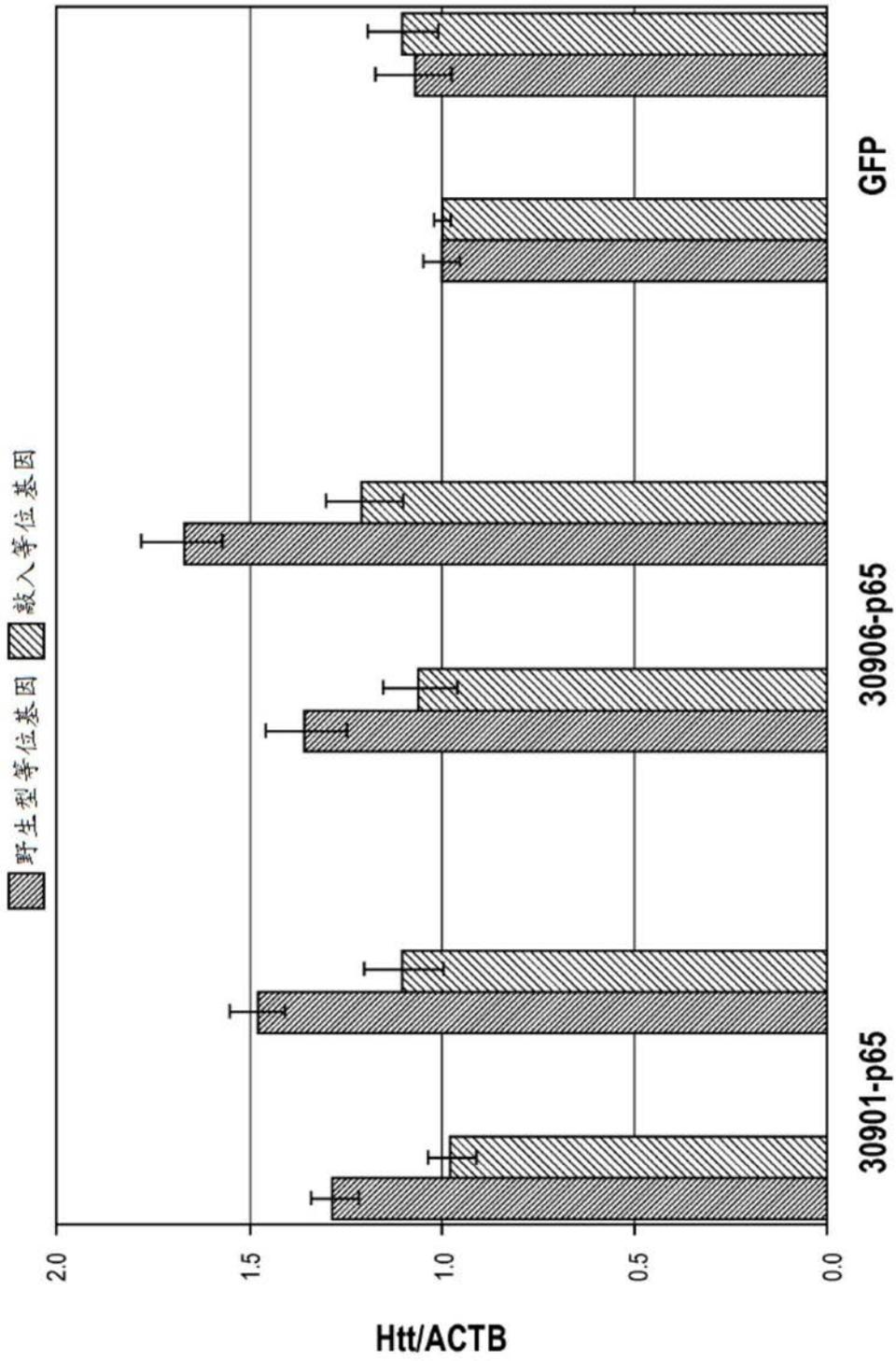
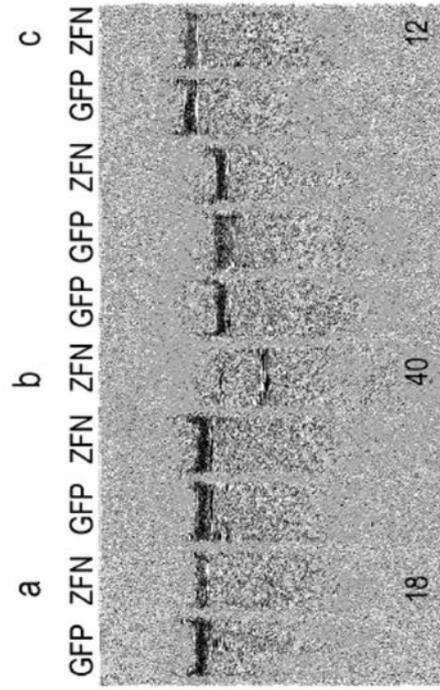


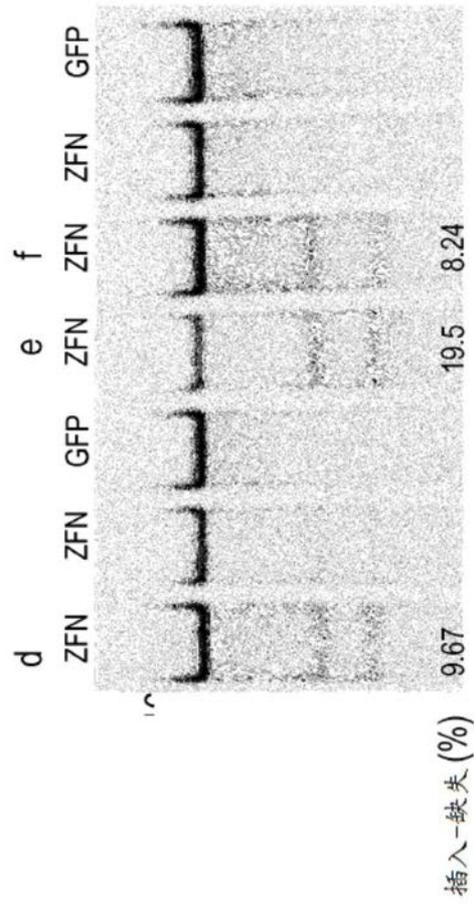
图13E



a: 29627/29628  
b: 29631/29632  
c: 29637/29638

插入-缺失(%)

图14A



d: 25917/25910  
e: 25921/25920  
f: 25923/25922

图14B

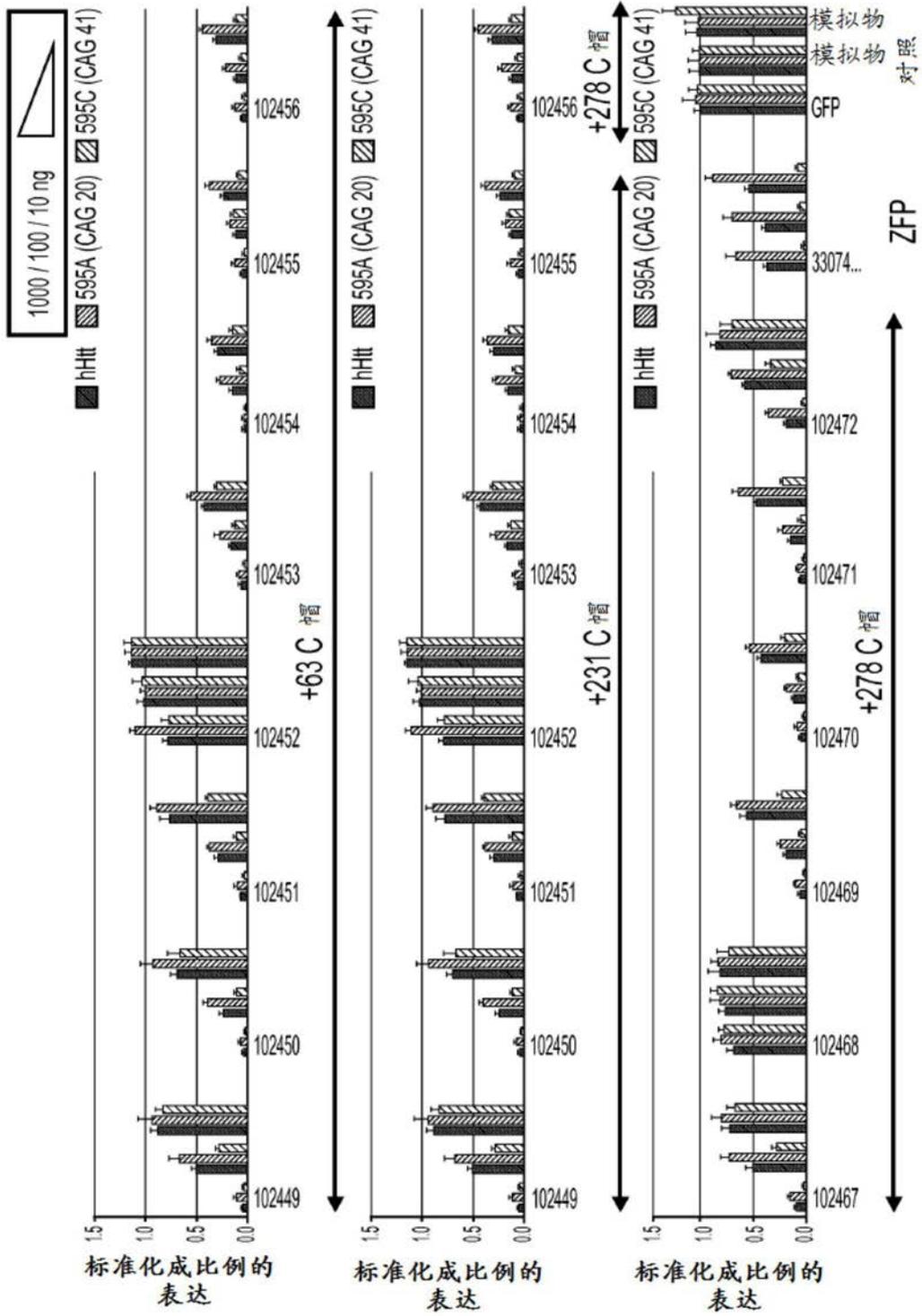


图15