(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 116406383 A
(43)申请公布日 2023.07.07

(21)申请号 202180074640.5
(22)申请日 2021.11.04
(30)优先权数据
2020-185619 2020.11.06 JP
(85)PCT国际申请进入国家阶段日
2023.05.04
(86)PCT国际申请的申请数据
PCT/JP2021/040503 2021.11.04
(87)PCT国际申请的公布数据
WO2022/097663 JA 2022.05.12
(71)申请人 日商基因编辑力股份有限公司
地址 日本福冈县
申请人 国立大学法人九州大学
(72)发明人 八木祐介 太田贤 中村崇裕
(74)专利代理机构 北京市中咨律师事务所 11247
专利代理师 曾祯 段承恩
(51)Int.Cl.
C07K 19/00(2006.01)

权利要求书2页 说明书16页
序列表10页 附图3页

(54)发明名称

FokI核酸酶结构域的突变体

(57)摘要

本发明发现4种核酸酶结构域突变体相较于野生型核酸酶结构域而言具有优异的核酸酶活性,在与各式各样的核酸结合结构域的组合中,能够提高基因组编辑效率。



CN 116406383 A

1.FokI蛋白质或其同源蛋白质的核酸酶结构域的突变体，该突变体具有下述(a)～(d)中任一项所记载的突变，通过该突变而核酸酶活性提升，

(a)选自由FokI蛋白质的386位或同源蛋白质的对应部位的氨基酸向Met的替换、FokI蛋白质的399位或同源蛋白质的对应部位的氨基酸向Arg的替换、FokI蛋白质的421位或同源蛋白质的对应部位的氨基酸向Tyr的替换、FokI蛋白质的424位或同源蛋白质的对应部位的氨基酸向Phe的替换、FokI蛋白质的432位或同源蛋白质的对应部位的氨基酸向Leu的替换、FokI蛋白质的433位或同源蛋白质的对应部位的氨基酸向Val的替换、FokI蛋白质的448位或同源蛋白质的对应部位的氨基酸向Asn的替换、FokI蛋白质的466位或同源蛋白质的对应部位的氨基酸向Ile的替换、FokI蛋白质的484位或同源蛋白质的对应部位的氨基酸向Ala的替换和FokI蛋白质的486位或同源蛋白质的对应部位的氨基酸向Leu的替换所组成的群组中的至少1种突变，

(b)选自由野生型蛋白质中的418位或同源蛋白质的对应部位的氨基酸向Pro的替换、FokI蛋白质的420位或同源蛋白质的对应部位的氨基酸向Leu的替换、FokI蛋白质的452位或同源蛋白质的对应部位的氨基酸向Val的替换、FokI蛋白质的473位或同源蛋白质的对应部位的氨基酸向Arg的替换、FokI蛋白质的486位或同源蛋白质的对应部位的氨基酸向His的替换、FokI蛋白质的490位或同源蛋白质的对应部位的氨基酸向Val的替换、FokI蛋白质的498位或同源蛋白质的对应部位的氨基酸向Leu的替换、FokI蛋白质的542位或同源蛋白质的对应部位的氨基酸向Lys的替换、FokI蛋白质的559位或同源蛋白质的对应部位的氨基酸向Arg的替换和FokI蛋白质的570位或同源蛋白质的对应部位的氨基酸向Ser的替换所组成的群组中的至少1种突变，

(c)选自由野生型蛋白质中的386位或同源蛋白质的对应部位的氨基酸向Met的替换、FokI蛋白质的395位或同源蛋白质的对应部位的氨基酸向Phe的替换、FokI蛋白质的399位或同源蛋白质的对应部位的氨基酸向Arg的替换、FokI蛋白质的421位或同源蛋白质的对应部位的氨基酸向Tyr的替换、FokI蛋白质的424位或同源蛋白质的对应部位的氨基酸向Phe的替换、FokI蛋白质的432位或同源蛋白质的对应部位的氨基酸向Leu的替换、FokI蛋白质的433位或同源蛋白质的对应部位的氨基酸向Val的替换、FokI蛋白质的448位或同源蛋白质的对应部位的氨基酸向Asn的替换、FokI蛋白质的484位或同源蛋白质的对应部位的氨基酸向Ala的替换和FokI蛋白质的486位或同源蛋白质的对应部位的氨基酸向Leu的替换所组成的群组中的至少1种突变，

(d)选自由野生型蛋白质中的385位或同源蛋白质的对应部位的氨基酸向Pro的替换、FokI蛋白质的395位或同源蛋白质的对应部位的氨基酸向Ala的替换、FokI蛋白质的427位或同源蛋白质的对应部位的氨基酸向Met的替换、FokI蛋白质的452位或同源蛋白质的对应部位的氨基酸向Val的替换、FokI蛋白质的460位或同源蛋白质的对应部位的氨基酸向Phe的替换、FokI蛋白质的473位或同源蛋白质的对应部位的氨基酸向Arg的替换、FokI蛋白质的481位或同源蛋白质的对应部位的氨基酸向His的替换、FokI蛋白质的490位或同源蛋白质的对应部位的氨基酸向Gln的替换、FokI蛋白质的542位或同源蛋白质的对应部位的氨基酸向Thr的替换、FokI蛋白质的559位或同源蛋白质的对应部位的氨基酸向Arg的替换和FokI蛋白质的570位或同源蛋白质的对应部位的氨基酸向Ser的替换所组成的群组中的至少1种突变。

2.如权利要求1所述的FokI蛋白质或其同源蛋白质的核酸酶结构域的突变体，该突变体具有FokI蛋白质的473位或同源蛋白质的对应部位的氨基酸向Arg的替换性突变，通过该突变而核酸酶活性提升了。

3.一种人工核酸切割酶，包含核酸结合结构域和权利要求1或2所述的核酸酶结构域的突变体。

4.如权利要求3所述的人工核酸切割酶，其中，所述核酸结合结构域是TALE、锌指、PPR或CRISPR-Cas。

5.一种编码权利要求1或2所述的核酸酶结构域的突变体、或者权利要求3或4所述的人工核酸切割酶的多核苷酸。

6.一种包含如权利要求5所述的多核苷酸的载体。

7.一种导入有权利要求5所述的多核苷酸、或权利要求6所述的载体的细胞。

8.一种基因组受到编辑的细胞或非人类生物的制造方法，包括将权利要求3所述的人工核酸切割酶、编码该人工核酸切割酶的多核苷酸或包含该多核苷酸的载体导入细胞或非人类生物中。

9.一种用于编辑细胞或生物的基因组的试剂盒，包含权利要求3所述的人工核酸切割酶、编码该人工核酸切割酶的多核苷酸或包含该多核苷酸的载体。

FokI核酸酶结构域的突变体

技术领域

[0001] 本发明涉及FokI核酸酶结构域的突变体、包含该突变体的人工核酸切割酶及它们的利用。

背景技术

[0002] 在基因组编辑中，要求以较高的机率对要改变的基因中的目标部位导入DNA的缺失、插入、替换等。用于基因组编辑的分子(基因组编辑工具)由结合至基因组上的目标序列的核酸结合结构域及切割目标序列附近的双链DNA的催化结构域所构成。在利用锌指蛋白质(ZFN)、转录激活因子样效应物(Transcription activator-like effector,TALE)蛋白质、三角状五肽重复(pentatricopeptide repeat,PPR)蛋白质等DNA结合蛋白质来构建基因组编辑工具的情况下，通常，使用属于Type IIS型限制性酶的FokI的具有切割双链DNA的活性的催化结构域(FokI-CD)。在FokI切割双链DNA时，必须在2个分子的FokI-CD间形成二聚体(非专利文献1)。因此，在融合FokI-CD所构建而得的基因组编辑工具中，也使用以要改变的目标部位周围的正义链及反义链作为目标序列的2个分子。

[0003] 对于在锌指核酸酶等融合FokI-CD所构建而得的基因组编辑工具，已进行了通过改变FokI-CD而提升利用基因组编辑工具的编辑效率。例如，已在ZFN中导入随机突变并使用大肠杆菌的人工进化系统而分离出基因组编辑效率提升的改良型FokI-CD(Sharky突变)(非专利文献2)。然而，融合带有Sharky突变的FokI-CD而得的TALEN，与野生型之间并未发现编辑效率的差异(非专利文献3)。

[0004] 现有技术文献

[0005] 非专利文献

[0006] 非专利文献1:Bitinaite J,et al.,(1998)Proc Natl Acad Sci U S A.95:10570-10575.

[0007] 非专利文献2:Guo J,et al.,(2010)J Mol Biol.400:96-107.

发明内容

[0008] 发明所要解决的课题

[0009] 本发明者们针对具有上述非专利文献2所记载的Sharky突变的FokI-CD进行进一步的性能评价，结果判明即便在融合PPR蛋白质作为核酸结合结构域的情况下，该FokI-CD相较于野生型而言也未发现基因组编辑效率的提升。因此，在具有Sharky突变的FokI-CD中，存在有在基因组编辑系统中无法与广泛的核酸结合结构域组合使用的问题。

[0010] 本发明是鉴于此种状况而完成，其目的在于提供在基因组编辑中可与各式各样的核酸结合结构域组合使用、且能够提升基因组编辑效率的核酸酶结构域突变体。

[0011] 用于解决课题的手段

[0012] 本发明者们为了解决上述课题，利用大肠杆菌的人工进化系统进行了高活性核酸酶结构域突变体的筛选，结果判明，与TALE结合了的4种核酸酶结构域突变体相较于野生型

核酸酶结构域而言具有优异的核酸酶活性，通过利用该核酸酶结构域突变体而基因组编辑效率提升。这些核酸酶结构域突变体之中，2种核酸酶结构域突变体共通地保持的在470位的突变对核酸酶活性的提升有较大的贡献。此外，该核酸酶结构域突变体在与PPR融合了的情况下，也发现基因组编辑效率的提升。由以上，本发明者们发现所制作而得的核酸酶结构域突变体具有优异的核酸酶活性，可与各式各样的核酸结合结构域进行组合而利用于基因组编辑，从而完成了本发明。

[0013] 本发明涉及高活性核酸酶结构域突变体、包含该核酸酶结构域突变体的人工核酸切割酶及它们的利用，更详细而言，提供以下发明。

[0014] [1]一种FokI蛋白质或其同源蛋白质的核酸酶结构域的突变体，该突变体具有下述(a)～(d)中任一项所记载的突变，通过该突变而核酸酶活性提升，

[0015] (a)选自由FokI蛋白质的386位或同源蛋白质的对应部位的氨基酸向Met的替换、FokI蛋白质的399位或同源蛋白质的对应部位的氨基酸向Arg的替换、FokI蛋白质的421位或同源蛋白质的对应部位的氨基酸向Tyr的替换、FokI蛋白质的424位或同源蛋白质的对应部位的氨基酸向Phe的替换、FokI蛋白质的432位或同源蛋白质的对应部位的氨基酸向Leu的替换、FokI蛋白质的433位或同源蛋白质的对应部位的氨基酸向Val的替换、FokI蛋白质的448位或同源蛋白质的对应部位的氨基酸向Asn的替换、FokI蛋白质的466位或同源蛋白质的对应部位的氨基酸向Ile的替换、FokI蛋白质的484位或同源蛋白质的对应部位的氨基酸向Ala的替换和FokI蛋白质的486位或同源蛋白质的对应部位的氨基酸向Leu的替换所组成的群组中的至少1种突变，

[0016] (b)选自由野生型蛋白质中的418位或同源蛋白质的对应部位的氨基酸向Pro的替换、FokI蛋白质的420位或同源蛋白质的对应部位的氨基酸向Leu的替换、FokI蛋白质的452位或同源蛋白质的对应部位的氨基酸向Val的替换、FokI蛋白质的473位或同源蛋白质的对应部位的氨基酸向Arg的替换、FokI蛋白质的486位或同源蛋白质的对应部位的氨基酸向His的替换、FokI蛋白质的490位或同源蛋白质的对应部位的氨基酸向Val的替换、FokI蛋白质的498位或同源蛋白质的对应部位的氨基酸向Leu的替换、FokI蛋白质的542位或同源蛋白质的对应部位的氨基酸向Lys的替换、FokI蛋白质的559位或同源蛋白质的对应部位的氨基酸向Arg的替换和FokI蛋白质的570位或同源蛋白质的对应部位的氨基酸向Ser的替换所组成的群组中的至少1种突变，

[0017] (c)选自由野生型蛋白质中的386位或同源蛋白质的对应部位的氨基酸向Met的替换、FokI蛋白质的395位或同源蛋白质的对应部位的氨基酸向Phe的替换、FokI蛋白质的399位或同源蛋白质的对应部位的氨基酸向Arg的替换、FokI蛋白质的421位或同源蛋白质的对应部位的氨基酸向Tyr的替换、FokI蛋白质的424位或同源蛋白质的对应部位的氨基酸向Phe的替换、FokI蛋白质的432位或同源蛋白质的对应部位的氨基酸向Leu的替换、FokI蛋白质的433位或同源蛋白质的对应部位的氨基酸向Val的替换、FokI蛋白质的448位或同源蛋白质的对应部位的氨基酸向Asn的替换、FokI蛋白质的484位或同源蛋白质的对应部位的氨基酸向Ala的替换和FokI蛋白质的486位或同源蛋白质的对应部位的氨基酸向Leu的替换所组成的群组中的至少1种突变，

[0018] (d)选自由野生型蛋白质中的385位或同源蛋白质的对应部位的氨基酸向Pro的替换、FokI蛋白质的395位或同源蛋白质的对应部位的氨基酸向Ala的替换、FokI蛋白质的427

位或同源蛋白质的对应部位的氨基酸向Met的替换、FokI蛋白质的452位或同源蛋白质的对应部位的氨基酸向Val的替换、FokI蛋白质的460位或同源蛋白质的对应部位的氨基酸向Phe的替换、FokI蛋白质的473位或同源蛋白质的对应部位的氨基酸向Arg的替换、FokI蛋白质的481位或同源蛋白质的对应部位的氨基酸向His的替换、FokI蛋白质的490位或同源蛋白质的对应部位的氨基酸向Gln的替换、FokI蛋白质的542位或同源蛋白质的对应部位的氨基酸向Thr的替换、FokI蛋白质的559位或同源蛋白质的对应部位的氨基酸向Arg的替换和FokI蛋白质的570位或同源蛋白质的对应部位的氨基酸向Ser的替换所组成的群组中的至少1种突变。

[0019] [2]如[1]所述的FokI蛋白质或其同源蛋白质的核酸酶结构域的突变体，该突变体具有FokI蛋白质的473位或同源蛋白质的对应部位的氨基酸向Arg的替换性突变，通过该突变而核酸酶活性提升了。

[0020] [3]一种人工核酸切割酶，包含核酸结合结构域和[1]或[2]所述的核酸酶结构域的突变体。

[0021] [4]如[3]所述的人工核酸切割酶，其中，前述核酸结合结构域是TALE、锌指(zinc finger)、PPR或CRISPR-Cas。

[0022] [5]一种编码[1]或[2]所述的核酸酶结构域的突变体、或者[3]或[4]所述的人工核酸切割酶的多核苷酸。

[0023] [6]一种包含[5]所述的多核苷酸的载体。

[0024] [7]一种导入有[5]所述的多核苷酸或[6]所记载的载体的细胞。

[0025] [8]一种基因组受到编辑的细胞或非人类生物的制造方法，包括将[3]所述的人工核酸切割酶、编码该人工核酸切割酶的多核苷酸或包含该多核苷酸的载体导入细胞或非人类生物中。

[0026] [9]一种用于编辑细胞或生物的基因组的试剂盒，包含[3]所述的人工核酸切割酶、编码该人工核酸切割酶的多核苷酸或包含该多核苷酸的载体。

[0027] 发明的效果

[0028] 依据本发明的核酸酶结构域突变体相较于野生型核酸酶结构域而言，可显示出优异的核酸酶活性。依据将本发明的核酸酶结构域与TALE或PPR等核酸结合结构域进行融合所制作的人工核酸切割酶，能够进行有效率的基因组编辑。

附图说明

[0029] 图1是示出4种FokI核酸酶结构域突变体的氨基酸序列的图。示出了相当于FokI核酸酶结构域的第384个至第579个氨基酸序列。下划线部分表示与野生型FokI不同的突变处。

[0030] 图2是示出以大肠杆菌的人工进化系统(B2P)中的生存率为指标，评价FokI核酸酶结构域突变型的核酸酶活性的结果的图。生存率表示将添加了阿拉伯糖的平板的菌落数除以未添加阿拉伯糖的平板中所生长的菌落数而得的值。WT为将野生型FokI核酸酶结构域融合至TALEN-A效应物所得的，图中的数值表示利用野生型FokI核酸酶的情况的生存率。

[0031] 图3是示出以大肠杆菌的人工进化系统(B2P)中的生存率为指标，评价FokI核酸酶结构域中的G473R突变对核酸酶活性所带来的影响的结果的图。生存率表示将添加了阿拉

伯糖的平板的菌落数除以未添加阿拉伯糖的平板中所生长的菌落数而得的值。WT为将野生型FokI核酸酶结构域融合至TALEN-A效应物所得的,图中的数值表示利用野生型FokI核酸酶的情况及利用已导入G473回复突变的M50(M50(G473))的情况的生存率。

[0032] 图4是示出使用以B2M基因或GFP基因作为目标序列的TALEN的靶标/报告物测定的结果的图。图表示出FokI核酸酶结构域突变体相对于野生型FokI核酸酶结构域的活性的增加率(倍)。值为3次独立实验的结果。

[0033] 图5是示出使用以B2M基因或GFP基因作为目标序列的PPR的靶标/报告物测定的结果的图。图表示出相对于野生型FokI核酸酶结构域的M50的活性(nLUC/fLUC)的增加率(倍)。值表示3次独立实验的平均值及标准误差。*是以P<0.1、**是以P<0.05相对于野生型FokI效应物的报告物活性显示出显著差异。

具体实施方式

[0034] <核酸酶结构域突变体>

[0035] 本发明提供FokI蛋白质或其同源蛋白质的核酸酶结构域的突变体。

[0036] 本发明中的“FokI蛋白质”为在自然界中在海床黄杆菌(Flavobacterium okeanokoites)中所发现的IIS型限制性酶。FokI蛋白质具有在双链DNA中,在与一条链上的识别部位相距9个核苷酸的位置、在与另一条链上的识别部位相距13个核苷酸的位置具有催化切割的作用。FokI蛋白质在N末端侧存在有核酸结合结构域,在C末端侧存在有核酸酶结构域(DNA切割结构域)。

[0037] 将典型的野生型FokI蛋白质的氨基酸序列示于序列号1,将其核酸酶结构域的氨基酸序列示于序列号2。在序列号1中,384～579位相当于核酸酶结构域。本发明的核酸酶结构域突变体是在FokI蛋白质的核酸酶结构域导入特定的突变而得的,相较于野生型核酸酶结构域而言,核酸酶活性提升。

[0038] 本发明的FokI核酸酶结构域突变体的一个方案为具有后述实施例的M47中所导入的突变的突变体,具体而言,是具有选自由FokI蛋白质的386位的氨基酸向Met的替换、FokI蛋白质的399位的氨基酸向Arg的替换、FokI蛋白质的421位的氨基酸向Tyr的替换、FokI蛋白质的424位的氨基酸向Phe的替换、FokI蛋白质的432位的氨基酸向Leu的替换、FokI蛋白质的433位的氨基酸向Val的替换、FokI蛋白质的448位的氨基酸向Asn的替换、FokI蛋白质的466位的氨基酸向Ile的替换、FokI蛋白质的484位的氨基酸向Ala的替换及FokI蛋白质的486位的氨基酸向Leu的替换所组成的群组中的至少1种突变的核酸酶结构域突变体。

[0039] 本发明的FokI核酸酶结构域突变体的另一个方案为具有后述实施例的M48中所导入的突变的突变体,具体而言,是具有选自由野生型蛋白质中的418位的氨基酸向Pro的替换、FokI蛋白质的420位的氨基酸向Leu的替换、FokI蛋白质的452位的氨基酸向Val的替换、FokI蛋白质的473位的氨基酸向Arg的替换、FokI蛋白质的486位的氨基酸向His的替换、FokI蛋白质的490位的氨基酸向Val的替换、FokI蛋白质的498位的氨基酸向Leu的替换、FokI蛋白质的542位的氨基酸向Lys的替换、FokI蛋白质的559位的氨基酸向Arg的替换、FokI蛋白质的570位的氨基酸向Ser的替换所组成的群组中的至少1种突变的核酸酶结构域突变体。

[0040] 本发明的FokI核酸酶结构域突变体的另一个方案为具有后述实施例的M49中所导

入的突变的突变体，具体而言，是具有选自由野生型蛋白质中的386位的氨基酸向Met的替换、FokI蛋白质的395位的氨基酸向Phe的替换、FokI蛋白质的399位的氨基酸向Arg的替换、FokI蛋白质的421位的氨基酸向Tyr的替换、FokI蛋白质的424位的氨基酸向Phe的替换、FokI蛋白质的432位的氨基酸向Leu的替换、FokI蛋白质的433位的氨基酸向Val的替换、FokI蛋白质的448位的氨基酸向Asn的替换、FokI蛋白质的484位的氨基酸向Ala的替换、FokI蛋白质的486位的氨基酸向Leu的替换所组成的群组中的至少1种突变的核酸酶结构域突变体。

[0041] 本发明的FokI核酸酶结构域突变体的另一个方案为具有后述实施例的M50中所导入的突变的突变体，具体而言，是具有选自由野生型蛋白质中的385位的氨基酸向Pro的替换、FokI蛋白质的395位的氨基酸向Ala的替换、FokI蛋白质的427位的氨基酸向Met的替换、FokI蛋白质的452位的氨基酸向Val的替换、FokI蛋白质的460位的氨基酸向Phe的替换、FokI蛋白质的473位的氨基酸向Arg的替换、FokI蛋白质的481位的氨基酸向His的替换、FokI蛋白质的490位的氨基酸向Gln的替换、FokI蛋白质的542位的氨基酸向Thr的替换、FokI蛋白质的559位的氨基酸向Arg的替换、FokI蛋白质的570位的氨基酸向Ser的替换所组成的群组中的至少1种突变的核酸酶结构域突变体。

[0042] 作为本发明的FokI核酸酶结构域突变体中的优选的突变的例子，可列举FokI蛋白质的473位的氨基酸向Arg的替换。

[0043] 在各突变体中，可具有2种以上突变的组合，也可具有3种以上突变的组合，也可具有4种以上突变的组合，也可具有5种以上(例如6种以上、7种以上、8种以上、9种以上、10种以上)突变的组合。

[0044] 此外，本发明的核酸酶结构域突变体也可为天然的FokI蛋白质的同源蛋白质的核酸酶结构域的突变体(以下，称为“同源蛋白质核酸酶结构域突变体”)。在此处，所谓“同源蛋白质”，是指在将其核酸酶结构域与天然的FokI蛋白质的核酸酶结构域进行比较的情况下，具有至少70％以上，优选为80％以上，更优选为90％以上(例如95％以上、96％以上、97％以上、98％以上、99％以上)的氨基酸序列的同源性的蛋白质。在同源蛋白质核酸酶结构域突变体中，与上述FokI核酸酶结构域突变体中的突变部位相对应的部位的氨基酸被替换成与FokI核酸酶结构域突变体中的突变部位相同的氨基酸。同源蛋白质也可源自于海床黄杆菌(Flavobacterium okeanokoites)以外的生物。

[0045] “氨基酸序列的同源性”可利用序列分析软件(例如BLAST；http://blast.ncbi.nlm.nih.gov/Blast.cgi)加以特定。此外，与FokI蛋白质的特定的氨基酸“相对应的部位”的氨基酸可特定为利用上述氨基酸序列分析软件(例如BLAST；http://blast.ncbi.nlm.nih.gov/Blast.cgi)与FokI蛋白质的氨基酸序列进行对齐时，与其突变部位的氨基酸处于同等地位的氨基酸。在经由软件的氨基酸序列的分析中，可利用例如预设的参数设定。

[0046] 本发明的核酸酶结构域突变体的一个优选方案为具有FokI蛋白质的473位或同源蛋白质的对应部位的氨基酸向Arg的替换的突变体。

[0047] 此外，本发明的核酸酶结构域突变体(FokI核酸酶结构域突变体、同源蛋白质核酸酶结构域突变体)也可以是除了上述突变部位以外，进一步导入了1个或多个氨基酸突变(替换、缺失、添加及/或插入)的突变体。在此处，所谓“多个”，并无特别限制，通常为2～50

个,优选为2～30个,更优选为2～20个,进一步优选为2～10个(例如2～8个、2～4个、2个)。作为有用的突变的例子,可列举例如异源二聚体形成型突变(Doyon Y,et al.(2011)Nat Methods 8:74-79)。

[0048] 作为部位特异性地导入突变的方法,可利用例如Kunkel法(Kunkel,T.A.Proc Natl Acad Sci USA(1985),82(2):488-492)、ODA法(Hashimoto-Gotoh et al.,(1995)Gene 152:271-276)、SOE(splicing-by-overlap-extention)-PCR法(Ho,S.N.et al.,(1989)Gene77:51-59)等公知的方法。也可利用市售的定点诱变试剂盒。

[0049] 本发明的核酸酶结构域突变体(FokI核酸酶结构域突变体、同源蛋白质核酸酶结构域突变体)通过突变的导入,相较于野生型蛋白质而言,核酸酶活性提升。核酸酶活性的提升优选为10％以上,更优选为30％以上,更优选为50％以上(例如70％以上、100％以上、150％以上、200％以上)。核酸酶活性的提升可通过例如实施例5(图4、5)所记载的报告物测定予以评价。具体而言,可通过下列方式进行评价:将表达TALEN或PPR等核酸结合结构域与本发明的核酸酶结构域突变体的融合蛋白质的效应物质粒、以及包含该核酸结合结构域的识别序列且若融合蛋白质识别出该识别序列并发生DNA切割时便会表达报告物的报告物质粒导入培养细胞中,测定该培养细胞中的报告物活性,与表达包含野生型核酸酶结构域的融合蛋白质的情况进行比较。在此报告物测定系统中,可使用例如HEK293-T细胞作为培养细胞,在各质粒中,作为用于表达融合蛋白质或报告物的启动子,可使用例如CMV启动子,作为核酸结合结构域的识别序列,可使用例如人类B2M基因或GFP基因上的识别序列。

[0050] 在本发明的核酸酶结构域突变体中,也可包含经修饰的氨基酸及/或非天然氨基酸。作为经修饰的氨基酸,并无限定,可列举例如甲基化、酯化、酰胺化、乙酰化、烷基化、卤化等。经修饰的氨基酸及非天然氨基酸可通过公知的方法予以导入。

[0051] <人工核酸切割酶>

[0052] 本发明提供包含核酸结合结构域和上述核酸酶结构域突变体的人工核酸切割酶。本发明的人工核酸切割酶系经由核酸结合结构域结合至核酸上的目标序列(核酸结合结构域的识别序列),通过核酸酶结构域在该目标序列的附近(目标切割部位)切割核酸。因此,本发明的人工核酸切割酶可作为序列特异性核酸切割酶而发挥功能。

[0053] 在本说明书中,在“核酸”中包括DNA和RNA的全部。本发明的人工核酸切割酶所切割的核酸主要为DNA。在DNA包括双链DNA和单链DNA两者。作为DNA,并无特别限制,可列举例如真核生物核基因组DNA、粒线体DNA、质体DNA、原核生物基因组DNA、噬菌体DNA或质粒DNA等。优选本发明的人工核酸切割酶切割基因组上的双链DNA。

[0054] 本发明的人工核酸切割酶的目标序列为核酸上的任意的序列。在以基因组DNA作为目标的情况下,目标序列也可设定成任意的基因区域或基因外区域。目标序列的长度并无特别限制,例如为10～30个碱基。在使用包含CRISPR/Cas的核酸结合结构域的情况下,Cas需要识别间隔物前体邻近基序(protospacer adjacent motif,PAM)序列,因而将位于PAM序列的附近的序列设定为目标序列。

[0055] 在通过本发明的人工核酸酶进行DNA的双链切割的情况下,优选的系设定夹着间隔物序列的2个目标序列。间隔物序列的长度并无特别限定,例如为1～20个碱基。本领域技术人员可适宜设定由所期望的长度及碱基序列所构成的目标序列。2个目标序列可以互为回文序列,或者也可为非回文序列。在2个目标序列为非回文序列的情况下,使用以各序列

作为目标的2种人工核酸切割酶。

[0056] 本发明的核酸切割酶中的核酸结合结构域只要是特异性地结合至任意的核酸序列(目标序列)的蛋白质结构域即可,可列举例如锌指、TALE、CRISPR/Cas(Cas蛋白质与指导RNA的复合体)、三角状五肽重复(pentatricopeptide repeat,PPR)等。在本发明的人工核酸切割酶中,核酸酶结构域突变体与核酸结合结构域可直接连接,或者也可经由接头进行连接。接头的链长并无特别限制,例如为1～20个氨基酸。

[0057] 包含锌指的核酸结合结构域优选可包含2个以上锌指,并无限定,也可包含例如由3～9个锌指所构成的锌指阵列。已知1个锌指系识别连续的3个碱基,例如,由3～9个锌指所构成的锌指阵列可识别9～27个碱基。关于ZFN的例子可参照文献(国际公开2003/087341号)。

[0058] 包含TALE的核酸结合结构域优选可包含6个以上重复模块,并无限定,也可包含例如由9～27个重复模块所构成的TALE。已知重复模块具有以34个氨基酸为1单位的重复结构,1个模块识别1个碱基。关于TALEN的例子参照文献(国际公开2012/104729号、国际公开2011/072246号、国际公开2015/019672号)。

[0059] 包含PPR的核酸结合结构域优选可包含6个以上重复模块,并无限定,也可包含例如由9～27个重复模块所构成的PPR。已知重复模块具有以35个氨基酸为1单位的重复结构,1个模块系识别1个碱基。关于PPR的例子参照文献(Nakamura et al.,(2012)Plant Cell Physiol 53:1171-1179、国际公开2014/175284号)。

[0060] 包含CRISPR-Cas的核酸结合结构域通过构成CRISPR-Cas的指导RNA中的目标化RNA(crRNA)识别目标序列。即,目标化RNA与目标序列具有彼此互补的碱基序列,可进行杂交。Cas由于与指导RNA形成复合体,故通过指导RNA的作用而被诱导至目标序列。在本发明中,可利用指导RNA与Cas的复合体作为核酸结合结构域,在此情况下,本发明的核酸酶突变体通常融合至Cas。在本发明中,通过核酸酶突变体进行核酸的切割,因而CRISPR/Cas中的Cas不一定需要保持核酸酶活性。Cas优选为使核酸酶活性失活了的Cas(例如dCas)。也可以通过改变Cas蛋白质(例如突变的导入)而改变PAM识别(Benjamin,P.et al.,(2015)Nature523:481-485、Hirano,S.et al.,(2016)Molecular Cell 61:886-894、国际公开2018/221685号),由此,可扩大目标序列的候选范围。

[0061] 作为用于本发明的CRISPR/Cas,可列举例如CRISPR-Cas9等2类/II型CRISPR/Cas,CRISPR-Cpf1(Cas12a)、CRISPR-Cas12b、CRISPR-CasX(Cas12e)及CRISPR-Cas14等2类/V型CRISPR/Cas,CRISPR-Cas13a等2类/VI型CRISPR/Cas,CRISPR-Cas3等1类/I型CRISPR/Cas,但并不限于这些。关于CRISPR/Cas的例子参照文献(国际公开2014/093712号、国际公开2013/176772号、国际公开2013/142578号、国际公开2016/205711号、Strecker J,et al.,Nature Communications10:212(2019)、Liu JJ,et al.,Nature 566:218-223(2019)、国际公开2018/225858号)。

[0062] <基因组受到编辑的细胞或非人类生物的制造方法>

[0063] 本发明提供基因组受到编辑的细胞或非人类生物的制造方法,包括将上述人工核酸切割酶、编码该人工核酸切割酶的多核苷酸、包含该多核苷酸的载体导入细胞或非人类生物中。

[0064] 若将本发明的人工核酸切割酶导入细胞(也包含非人类生物的体内的细胞)中,则

核酸结合结构域结合至核酸上的目标序列，核酸酶结构域突变体在目标切割部位切割该核酸。紧接于此切割之后，进行经由非同源末端连接修复(NHEJ)或同源重组修复(HR)等的修复，基因组受到编辑。在成为主要修复途径的经由NHEJ的修复中，在该切割部位插入1个以上突变，该核酸受到改变。另一方面，通过使用后述供体DNA，便可通过在目标切割部位的周围区域所发生的同源重组修复(HDR)，而在目标DNA区域插入供体DNA内的所期望的DNA。

[0065] 导入细胞中的本发明的人工核酸切割酶可为蛋白质的形态，也可为多核苷酸的形态，也可为包含该多核苷酸的载体的形态。在使用CRISPR-Cas作为人工核酸切割酶的核酸结合结构域的情况下，可进一步采用RNA与蛋白质的复合体的形态。多核苷酸可为DNA，也可为RNA，也可为了在细胞中高表达而进行密码子最佳化。

[0066] 在采用表达载体的形态的情况下，包含可作用地结合至应表达的DNA的1个以上调控元件。在此处，所谓“可作用地结合”，是指上述DNA能够进行表达地结合至调控元件。作为“调控元件”，可列举启动子、增强子、内部核糖体进入位点(IRES)及其他表达调控元件(例如转录终止信号(例如多聚腺苷化信号、多聚U序列))。作为调控元件，根据目的，例如可以是指向在多样的宿主细胞中的DNA的结构性表达的，也可以是仅在特定的细胞、组织或器官中指向DNA的表达的。此外，也可以是仅在特定的时期指向DNA的表达的，也可以是指向能够人为地诱导的DNA的表达的。作为启动子，可列举例如polIII启动子(例如U6及H1启动子)、polII启动子(例如反转录病毒的劳斯肉瘤病毒(Rous sarcoma virus，RSV)LTR启动子、巨细胞病毒(CMV)启动子、SV40启动子、二氢叶酸还原酶启动子、β-肌动蛋白启动子、磷酸甘油激酶(PGK)启动子及EF1α启动子)、polI启动子或它们的组合。只要是本领域技术人员，便可根据所导入的细胞的种类等，选择适当的表达载体。

[0067] 在本发明的基因组受到编辑的细胞或非人类生物的制造方法中，除了上述人工核酸切割酶以外，也可将供体DNA导入细胞中。由此，可利用在目标切割部位的周围区域所发生的同源重组修复(HDR)，在目标DNA区域中插入所期望的DNA。在同源重组修复的过程中，需要目标DNA区域的碱基序列与供体DNA的碱基序列的同源性，将供体DNA用于包含目标切割部位的目标DNA区域的模板修复，导致遗传信息从供体DNA移动至目标DNA区域。由此，可使目标DNA区域的碱基序列发生变化(例如插入、缺失、替换等)。因此，供体DNA包含与目标DNA区域内的碱基序列具有较高的同一性的2个碱基序列(同源性臂)及配置于它们之间的所期望的DNA(用于插入目标DNA区域中的DNA)。

[0068] 同源性臂只要是足以进行同源重组的程度的大小即可，此外，可依供体DNA的形态、链长而有所变动，在双链供体DNA的情况下，例如为500～1000个碱基对，在单链供体DNA的情况下，例如为30～1000个碱基。此外，同源性臂只要以足以进行同源重组的程度与目标DNA区域内的碱基序列具有同一性，即便没有100％的同一性也可。例如，各自具有95％以上、优选为97％以上、更优选为99％以上、进一步优选为99.9％以上的同一性。

[0069] 此外，存在于同源性臂之间的所期望的DNA的长度并无特别限定，可利用各式各样尺寸的。在所期望的DNA中存在要事后去除的碱基序列的情况下，也可在该碱基序列的两端添加例如重组酶的识别序列(例如loxP序列、FRT序列)。被该识别序列所夹住的碱基序列可通过使重组酶(例如Cre重组酶、FLP重组酶)进行作用而予以去除。此外，在用于确认DNA敲入的成功等的目的下，也可例如将选择标志物序列(例如荧光蛋白质或药剂耐性基因等)装入所期望的DNA中。此外，也可使用可作用地结合1个以上调控元件的基因作为所期望的

DNA。

[0070] 供体DNA可为直链状DNA，也可为环状DNA。此外，可为单链DNA，也可为双链DNA。

[0071] 本发明的人工核酸切割酶(上述各种形态)向细胞的导入可通过例如电穿孔、显微注射、DEAE-葡聚糖法、脂质体转染法、纳米粒子介导的转染法、病毒介导的核酸传递法等公知的方法来进行。在进行向生物体内的直接施与的情况下，可通过注射等非经口施与及经口施与来进行。在此情况下，优选视需要使其与对目标部位具有指向性的传递系统进行组合。在使用供体DNA的情况下，该供体DNA也可以以同样的方法导入细胞或生物体内。

[0072] 导入本发明的人工切割酶的细胞可为原核生物的细胞(原核细胞)或真核生物的细胞(真核细胞)中的任何种细胞，并无特别限定。作为原核细胞，可列举例如大肠杆菌、放线菌、古细菌等。作为真核细胞，可列举例如动物细胞、植物细胞、藻细胞、真菌细胞。作为动物细胞，可列举例如哺乳动物细胞，以及鱼类、鸟类、爬虫类、两栖类、昆虫类的细胞。

[0073] 在动物细胞中，包含例如构成动物的个体的细胞、构成从动物摘出的器官/组织的细胞、源自于动物的组织的培养细胞等。具体而言，可列举例如卵母细胞或精子等生殖细胞；各阶段的胚胎的胚细胞(例如1细胞期胚、2细胞期胚、4细胞期胚、8细胞期胚、16细胞期胚、桑椹期胚等)；诱导多能性干(iPS)细胞或胚性干(ES)细胞等干细胞；成纤维细胞、造血细胞、神经元、肌细胞、骨细胞、肝细胞、胰脏细胞、脑细胞、肾细胞等体细胞等。作为用于基因组编辑动物的制作的卵母细胞，可利用受精前及受精后的卵母细胞，优选为受精后的卵母细胞，即受精卵。特别优选受精卵为原核期胚。卵母细胞可将经冻结保存后的加以解冻而使用。但是，在从伦理的观点而言无法容许的情况下，排除对人类的生殖细胞或胚胎的细胞的应用。

[0074] 哺乳动物是包括人类及非人类哺乳动物的概念。作为非人类哺乳动物的例子，可列举牛、野猪、家猪、绵羊、山羊等偶蹄类，马等奇蹄类，小鼠、大鼠、豚鼠、仓鼠、松鼠等啮齿类，兔等兔目，犬、猫、雪貂等食肉类等。非人类哺乳动物可为家畜或伴侣动物(宠物)，也可为野生动物。

[0075] 作为植物细胞，可列举例如谷物类、油料作物、饲料作物、水果、蔬菜类的细胞。在植物细胞中，包含例如构成植物的个体的细胞、构成从植物分离出的器官/组织的细胞、源自于植物的组织的培养细胞等。作为植物的器官/组织，可列举例如叶、茎、茎尖(生长点)、根、块茎、愈伤组织等。作为植物的例子，可列举水稻、玉米、香蕉、花生、向日葵、西红柿、油菜、烟草、小麦、大麦、马铃薯、大豆、棉花、康乃馨等，也包含其繁殖材料(例如种子、块根、块茎等)。

[0076] 在本发明的基因组受到编辑的细胞或非人类生物的制造方法中，也可将核酸结合结构域不同的2种以上上述人工核酸切割酶导入细胞中。例如，在基因组上的要编辑的区域中不存在回文序列的情况下，也可使用各自以不同的序列作为目标的2种人工核酸切割酶。在此情况下，也可进一步在2种人工核酸切割酶的核酸酶结构域突变体中导入促进该核酸酶结构域突变体的异源二聚体形成的进一步的突变。由此，可增加人工核酸切割酶的识别序列，同时使脱靶的机率降低。

[0077] 〈试剂盒〉

[0078] 本发明提供用于编辑细胞或生物的基因组的试剂盒，包含上述人工核酸切割酶、编码该人工核酸切割酶的多核苷酸、包含该多核苷酸的载体。该试剂盒有时进一步包含一

种或多种补充试剂，作为补充试剂，可列举例如稀释缓冲液、再构成溶液、洗净缓冲液、核酸导入试剂、蛋白质导入试剂、对照试剂，但并不限于这些。通常，试剂盒附有操作说明书。

[0079] 实施例

[0080] (实施例1.B2P筛选的原理)

[0081] 在B2P筛选系统中使用下列2种质粒：具有致死基因ccdB基因的报告物质粒，及表达TALEN-A的效应物质粒(Chen Z,Zhao H.(2005)Nucleic Acids Res.33:e154)。在报告物质粒的ccdB基因的下游，将TALEN-A所识别的序列(表1)各自插入正义链侧及反义侧(若从正义链侧来看，则成为“目标序列-间隔物序列-目标序列的互补序列”)(表2)。另外，表2中，大写字母表示TALEN-A所识别的序列，小写字母表示2个序列之间的间隔物序列。

[0082] 表1

[0083]

名称	目标序列
TALEN-A	AGCCGAAATCATCGCAG(序列号:3)
GFP_TALEN(L)	CAGCGTGTCCGGCGA(序列号:4)
GFP_TALEN(R)	TTGCCGTAGGTGGCA(序列号:5)
B2M_TALEN(L)	CCAAAGATTCAGGT(序列号:6)
B2M_TALEN(R)	GACTTTCCATTCTC(序列号:7)

[0084] 表2

[0085]

名称	目标序列
TALEN-A	AGCCGAAATCATCGCAGccgctgccgcgagctcCTGCGATGATTTCGGCT(序列号:8)
GFP_TALEN	CAGCGTGTCCGGCGAGggcgagggcgaTGCCACCTACGGCAA(序列号:9)
B2M_TALEN	CCAAAGATTCAGGTttactcacgtcatccagcaGAGAATGGAAAGTC(序列号:10)

[0086] ccdB基因可通过阿拉伯糖诱导性启动子加以表达控制，若在培养基中添加阿拉伯糖，则经报告物转化的大肠杆菌无法生存(将此质粒称为“TALEN-A报告物质粒”)。在效应物质粒中，TALEN-A在N-末端侧融合有麦芽糖结合蛋白质，可通过IPTG诱导MBP-TALEN-A融合蛋白质(将此质粒称为“TALEN-A效应物质粒”)。若将此2种质粒转化至大肠杆菌(XL1-Blue)中，则转化体在阿拉伯糖存在下无法生存，但在阿拉伯糖及IPTG存在下可生存。这是由于MBP-TALEN-A在位于TALEN-A报告物质粒的目标序列中导入双链DNA切割，因而经切割的质粒DNA无法进行DNA复制，报告物质粒从转化体中消失，失去ccdB基因的致死效果。若在具有TALEN-A报告物的大肠杆菌中导入TALEN-A效应物的突变体文库，并在阿拉伯糖及IPTG存在下进行筛选，则具有活性更高的TALEN-A效应物的大肠杆菌可生存，而具有活性较弱的TALEN-A效应物的大肠杆菌无法生存。若从已生存的大肠杆菌中提取出TALEN-A效应物的质粒DNA并再次转化至具有TALEN-A报告物的大肠杆菌中，则具有活性最高的TALEN-A效应物

的大肠杆菌进一步被浓缩。若重复此种循环,则可分离出包含活性最高的TALEN-A效应物的质粒。

[0087] (实施例2.突变体文库的构建)

[0088] TALEN-A效应物的突变体文库如下构建。以TALEN-A效应物质粒DNA当作模板,使用PrimeSTAR(注册商标)HS DNA Polymerase(TaKaRa)进行PCR扩增,通过去除FokI-CD部分并导入Esp3I部位而构建目的载体。将MBP-TALEN-A基因的FokI-CD部分的碱基序列使用GeneMorph II Random Mutagenesis kit(Agilent Technologies公司)进行扩增,调制在FokI-CD的碱基序列中导入有随机突变的插入片段。使经Esp3I切割的目的载体与插入片段进行连接而构建TALEN-A效应物的突变体文库。评价随机突变的导入率,结果在FokI-CD部分的碱基序列(595bp)中平均导入了4处替换。

[0089] (实施例3.突变体文库的筛选)

[0090] 在阶段1的第1次筛选中,将所构建的突变体文库的质粒DNA 1μg以电穿孔法(Bio-Rad)转化至具有TALEN-A报告物的大肠杆菌中。电穿孔后,将大肠杆菌移至30mL的SOB中并于37℃恢复培养1小时。以最终浓度成为1mM的方式加入IPTG并于30℃保温培养4小时之后,接种至包含100mg/L的氯霉素(Cm)及0.2%阿拉伯糖的LB琼脂平板中。此时,通过将菌液的一部分各自接种至添加及未添加阿拉伯糖的LB+Cm琼脂平板中并对菌落数进行比较而求出生存率。第1次筛选的总菌落数为2×10^7个,在有阿拉伯糖的培养基中所获得的菌落为6×10^6个。因此,生存率为30%。使有阿拉伯糖的培养基上的菌落悬浮于SOB培养基中,使用NucleoSpin(注册商标)Plasmid(TaKaRa)从菌液中提取出质粒DNA。

[0091] 第2次筛选也使用所提取出的质粒DNA,与第1次同样地进行。此时的总菌落数为6.8×10^8个,生存率为58%。在第3次筛选中,生存率成为100%。在此条件下,无法选拔出更好的TALEN-A效应物,因而在第4次筛选中,将IPTG的诱导时间从4小时缩短成2小时,使选拔条件更加严格。第4次的生存率为20%,但在相同条件下进行第5次筛选,结果生存率达到90%。第6～8次将IPTG诱导时间设成1小时,结果生存率分别为1.3%、10%、6.5%。即便重复次数,生存率也不提升,从第1次的独立克隆数6×10^6个及8次筛选的生存率所算出的独立克隆数为55个以下,因而阶段1在第8次筛选结束。

[0092] 将以阶段1的第8次筛选中所获得的质粒DNA当作模板并进一步导入随机突变而得的FokI-CD插入片段插入目的载体中,而构建突变再导入文库。将此文库转化至具有TALEN-A报告物的大肠杆菌中,并开始进行阶段2的筛选。第1次是以1mM IPTG 1小时,第2～3次是30分钟,第4次是15分钟,第5～6次是以0.1mM IPTG 15分钟,分别进行诱导并进行筛选。筛选的初次诱导于30℃进行,第2次以后的诱导皆于37℃进行。从第1次的独立克隆数2.8×10^5个及6次筛选的生存率所算出的独立克隆数成为1个。

[0093] 将以阶段2的第6次筛选中所获得的质粒DNA当作模板并进一步导入随机突变而得的FokI-CD插入片段插入目的载体中,而构建突变再导入文库。将此文库转化至具有TALEN-A报告物的大肠杆菌中,并开始进行阶段3的筛选。第1次是20μM IPTG,第2～3次是10μM IPTG,第4次是1μM IPTG,第5～6次是无IPTG,分别诱导15分钟并进行筛选。筛选的初次诱导于30℃进行,第2次以后的诱导皆于37℃进行。从第1次的独立克隆数1.5×10^5个及6次筛选的生存率所算出的独立克隆数成为1个,故结束阶段3的筛选。

[0094] 将以阶段3的第6次筛选中所获得的质粒DNA当作模板并进一步导入随机突变而得

的FokI-CD插入片段插入目的载体中，而构建突变再导入文库。将此文库转化至具有TALEN-A报告物的大肠杆菌中，并开始进行阶段4的筛选。全部在相同条件(在0.5μM IPTG存在下，于37℃诱导5分钟)下进行5次筛选。从第1次的独立克隆数8.0×10^4个及6次筛选的生存率所算出的独立克隆数成为6个，故结束阶段4的筛选。

[0095] 将在阶段4中，第5次中所获得的质粒DNA转化至大肠杆菌XL1 Blue株中，从所获得的菌落中提取出质粒DNA，并确认FokI-CD部分的碱基序列。其结果获得4种突变型FokI(M47、M48、M49、M50)。在M47、M48、M49中各自包含10处氨基酸替换，在M50中包含11处氨基酸替换(图1)。

[0096] 将所分离出的4种突变型FokI中所发现的氨基酸替换与现有的突变型FokI中所包含的氨基酸替换进行比较。就锌指核酸酶而言显示出高活性的Sharkey突变由S418P替换及K441E替换所构成(上述非专利文献2)。S418P替换也存在于M48中，但K441E替换不存在。在非专利文献2中也已验证Q481H及N527D突变的效果，但相比于Sharkey突变而言，其效果较小。Q481H替换也存在于M50中。此外，已报导FokI仅在不同分子间形成二聚体的异源二聚体形成型突变，Q486E/I499L/N496D及E490K/I538K/H537R(Doyon Y,et al.,(2011)Nat Methods.8:74-79)。在M47、M48、M49中，在Q486导入替换，在M47及M49中为Q486L替换，在M48中为Q486H替换。在M50中存在E490Q替换，与Doyon等人的文献的E490K替换不同。同样的异源二聚体形成型突变也已在文献(Ramalingam S,et al.,(2011)J Mol Biol.405:630-641、Szczepek M,et al.,(2007)Nat Biotechnol.25:786-793)中报导，但除了Doyon等人的文献与本实施例中所分离出的4种突变体所共通的残基以外，并不存在共通的氨基酸替换。Q481A替换已被报导为使脱靶突变减少的突变(Miller JC,et al.,(2019)Nat Biotechnol.37:945-952)，但在M47及M49中为Q486L替换，在M48中为Q486H替换，因而并不一致。

[0097] (实施例4.利用大肠杆菌B2P的突变型FokI的评价)

[0098] (1)首先，通过大肠杆菌的B2P确认4种突变型FokI(M47、M48、M49、M50)相比于野生型FokI而言，生存率提升何种程度。将具有野生型及4种突变型FokI(M47、M48、M49、M50)的TALEN-A效应物质粒DNA 10ng通过电穿孔法(Bio-Rad)转化至具有TALEN-A报告物的大肠杆菌中。加入1mL SOB并于37℃恢复培养1小时后，加入包含0.1mM IPTG的6mL SOB并于37℃保温培养15分钟而诱导TALEN-A。将转化体接种至添加及未添加阿拉伯糖的平板中并于37℃培养1夜之后，对菌落数进行计测。生存率设定为将添加了阿拉伯糖的平板的菌落数除以未添加阿拉伯糖的平板的菌落数而得的值。

[0099] 其结果是，表达具有野生型FokI的TALEN-A效应物的大肠杆菌的生存率为0.0001％，相对于此，在使用M47、M48、M49或M50作为FokI的大肠杆菌中，生存率分别为29％、37％、45％、28％(图2)。突变型FokI-TALEN-A效应物相比于野生型而言，活性上升21万倍(M47、M50)～33万倍(M49)。

[0100] (2)4种突变型FokI(M47、M48、M49、M50)之中，在M48及M50中共通地包含G473R突变。G473R突变为在筛选的最初阶段(阶段1)已鉴定出的突变，邻近属于FokI核酸酶的活性中心的K469残基。于是，通过大肠杆菌的B2P研究G473R突变在M50的核酸酶的活性中的重要性。首先，构建以PCR法导入将M50的R473残基恢复成野生型G473残基的回复突变(M50(G473))而得的TALEN-A效应物质粒。将分别具有野生型FokI、M50及M50(G473)的TALEN-A效

应物质粒DNA 10ng通过电穿孔法(Bio-Rad)转化至具有TALEN-A报告物的大肠杆菌。加入1mL SOB并于37℃恢复培养1小时后,加入包含0.1mM IPTG的6mL SOB并于37℃保温培养15分钟而诱导TALEN-A。将转化体接种至添加及未添加阿拉伯糖的平板中并于37℃培养1夜之后,对菌落进行计数。生存率设定为将添加了阿拉伯糖的平板的菌落数除以未添加阿拉伯糖的平板的菌落数而得的值。

[0101] 其结果是,表达具有M50的TALEN-A效应物的大肠杆菌的生存率为24%,而表达具有M50(G473)的TALEN-A效应物的大肠杆菌的生存率为0.05%以下(图3)。由此,得知G473R突变在M50的核酸酶活性中的重要性。

[0102] (实施例5.突变型FokI在动物培养细胞中的利用报告物测定的评价)

[0103] (1)接着,为了调查突变型FokI是否使基因组编辑效率提升,使用被用作评价基因组编辑效率的常规方法的SSA(单链退火,Single Strand Annealing)报告物来评价使突变型FokI连接至TALE而得的效应物的性能。

[0104] 为了使4种突变型FokI(M47、M48、M49、M50)可在动物细胞中效率良好地表达,合成经最佳化成人类的密码子的基因。使4种经最佳化成人类的密码子的突变型FokI-CD融合至以GFP基因作为目标的TALEN(GFP_TALEN)或以人类B2M基因作为目标的TALEN(B2M_TALEN)的C-末端侧,并使用Golden-Gate法插入具有CMV启动子的表达载体中,而构建动物细胞用效应物质粒。另外,所使用的TALEN所识别的序列示于上述表1。

[0105] 具有上述TALEN所识别的目标序列的SSA报告物为以下构成。在CMV启动子的下游依序连接萤火虫荧光素酶(fLUC)的5'侧1131bp的序列、目标序列、fLUC的3'侧1317bp的序列。在fLUC的5'侧及3'侧有795bp重复的序列,在于这些重复序列之间发生同源重组的情况下,会生成正常的fLUC并进行表达。目标序列示于上述表2。

[0106] SSA报告物测定如下进行。将1500个人类培养细胞HEK293-T细胞预培养1日,使用FuGENE(注册商标)Transfection Reagent(Promega)转染2.0ng的报告物质粒DNA、10ng的效应物质粒DNA、0.5ng的参考质粒DNA。在5%CO_2存在下于37℃培养2日后,使用Dual-Glo(注册商标)Luciferase Assay System(Promega)测定fLUC及rLUC的活性。对各样品按4个系列进行测定3次。其结果是,4种突变型FokI全部相较于野生型而言皆显示出较高的活性(图4)。

[0107] (2)使用可以高灵敏度检测出基因组编辑的Fn(Firefly luciferase&NanoLUC)报告物研究对PPR使用突变型FokI是否也提升基因组编辑效率。

[0108] 使已最佳化成人类的密码子的野生型FokI或M50融合至以人类B2M基因或GFP基因作为目标的1对PPR(B2M_PPR、GFP_PPR)的C-末端侧,插入具有CMV启动子的表达载体中而构建动物细胞用效应物质粒。另外,所使用的PPR所识别的序列示于表3。

[0109] 表3

[0110]

名称	目标序列
B2M_PPR(L)	AATGGAAAGTCAAA(序列号:15)
B2M_PPR(R)	AGCAATTCAGGAAA(序列号:16)
GFP_PPR(L)	ACGACTTCTTCAAGT(序列号:17)
GFP_PPR(R)	CCTGGACGTAGCCTT(序列号:18)

[0111] 具有上述PPR所识别的目标序列的Fn报告物为以下构成。在CMV启动子的下游依序连接萤火虫荧光素酶(fLUC)、目标序列、NanoLuc(nLUC)而构建Fn报告物质粒。在Fn报告物系统中,fLUC与目标序列框内连接,通常,直至fLUC为止会被翻译,而nLUC由于在上游有终止密码子而不会被翻译。因此,通常仅检测出fLUC的活性。另一方面,若目标序列受到编辑而产生Indel,则会在目标序列产生移码,fLUC与nLUC进行融合而得的报告物基因会进行表达,可检测出fLUC及nLUC两者的活性。另外,将插入Fn报告物系统中的人类B2M基因及GFP基因的目标序列示于表4。

[0112] 表4

[0113]

名称	目标序列
B2M目标	TACTCCAAAGATTCAGGTTTACTCACGTCATCCAG CAGAGAATGGAAAGTCAAATTTCCTGAATTGCTATG TGTCTGGGTTTCATCCATCCGACAT(序列号:19)
GFP目标	CAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGAT GCCACCTACGGCAAGCTGACCCTGAAGTTCATCTG CACCACCGGCAAGCTGCCCGTGCCCTGGCCCACC CTCGTGACCACCCTGACCTACGGCGTGCAGTGCTT CAGCCGCTACCCCGACCACATGAAGCAGCACGACT TCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAG GAGC(序列号:20)

[0114] Fn报告物测定如下进行。将1500个人类培养细胞HEK293-T细胞预培养1日。将2.5ng的Fn报告物质粒DNA以及分别针对PPR(L)及PPR(R)各者的50ng的效应物质粒DNA使用FuGENE(注册商标)Transfection Reagent(Promega)转染至经预培养的HEK293-T细胞。在5%CO_2存在下于37℃培养1夜后,使用Nano-Glo(注册商标)Luciferase Assay System(Promega)测定fLUC及nLUC的活性。对各样品按4个系列进行测定,以将nLUC除以fLUC而得的值评价活性。

[0115] 其结果是,在以B2M基因作为目标的情况下,M50显示出野生型FokI的2.2倍的活性,在以GFP基因作为目标的情况下,M50显示出野生型FokI的1.5倍的活性(图5)。由以上结果,已判明M50即便连接至PPR,相较于野生型FokI而言也显示出较高的性能。

[0116] 产业可利用性

[0117] 如以上所说明,本发明的新的核酸酶结构域突变体在与各式各样的核酸结合结构

域的组合中，相较于野生型FokI核酸酶结构域而言具有优异的活性。本发明的人工核酸切割酶作为基因组编辑工具不仅在基础研究中有用，在包含医疗、农业、工业在内的各式各样的产业应用中也有用。

[0118] 序列表自由文字

[0119] 序列号3

[0120] <223>TALEN-A的目标序列

[0121] 序列号4

[0122] <223>GFP_TALEN(L)的目标序列

[0123] 序列号5

[0124] <223>GFP_TALEN(R)的目标序列

[0125] 序列号6

[0126] <223>B2M_TALEN(L)的目标序列

[0127] 序列号7

[0128] <223>B2M_TALEN(R)的目标序列

[0129] 序列号8

[0130] <223>TALEN-A的目标部位

[0131] 序列号9

[0132] <223>GFP_TALEN的目标部位

[0133] 序列号10

[0134] <223>B2M_TALEN的目标部位

[0135] 序列号11

[0136] <223>M47突变体

[0137] 序列号12

[0138] <223>M48突变体

[0139] 序列号13

[0140] <223>M49突变体

[0141] 序列号14

[0142] <223>M50突变体

[0143] 序列号15

[0144] <223>B2M_PPR(L)的目标序列

[0145] 序列号16

[0146] <223>B2M_PPR(R)的目标序列

[0147] 序列号17

[0148] <223>GFP_PPR(L)的目标序列

[0149] 序列号18

[0150] <223>GFP_PPR(R)的目标序列

[0151] 序列号19

[0152] <223>B2M_PPR的目标部位

[0153] 序列号20

[0154] <223>GFP_PPR的目标部位

序列表

<110> 日商基因编辑力股份有限公司
国立大学法人九州大学

<120> FokI核酸酶结构域的突变体

<130> IBPF21-544WO

<150> JP 2020-185619

<151> 2020-11-06

<160> 20

<170> PatentIn 版本 3.5

<210> 1

<211> 579

<212> PRT

<213> 海床黄杆菌(Flavobacterium okeanokoites)

<400> 1

```
Met Val Ser Lys Ile Arg Thr Phe Gly Trp Val Gln Asn Pro Gly Lys
1               5                  10                  15
Phe Glu Asn Leu Lys Arg Val Val Gln Val Phe Asp Arg Asn Ser Lys
            20                  25                  30
Val His Asn Glu Val Lys Asn Ile Lys Ile Pro Thr Leu Val Lys Glu
        35                  40                  45
Ser Lys Ile Gln Lys Glu Leu Val Ala Ile Met Asn Gln His Asp Leu
    50                  55                  60
Ile Tyr Thr Tyr Lys Glu Leu Val Gly Thr Gly Thr Ser Ile Arg Ser
65                  70                  75                  80
Glu Ala Pro Cys Asp Ala Ile Ile Gln Ala Thr Ile Ala Asp Gln Gly
                85                  90                  95
Asn Lys Lys Gly Tyr Ile Asp Asn Trp Ser Ser Asp Gly Phe Leu Arg
            100                 105                 110
Trp Ala His Ala Leu Gly Phe Ile Glu Tyr Ile Asn Lys Ser Asp Ser
        115                 120                 125
Phe Val Ile Thr Asp Val Gly Leu Ala Tyr Ser Lys Ser Ala Asp Gly
    130                 135                 140
Ser Ala Ile Glu Lys Glu Ile Leu Ile Glu Ala Ile Ser Ser Tyr Pro
145                 150                 155                 160
Pro Ala Ile Arg Ile Leu Thr Leu Leu Glu Asp Gly Gln His Leu Thr
                165                 170                 175
Lys Phe Asp Leu Gly Lys Asn Leu Gly Phe Ser Gly Glu Ser Gly Phe
            180                 185                 190
```

Thr Ser Leu Pro Glu Gly Ile Leu Leu Asp Thr Leu Ala Asn Ala Met
195 200 205
Pro Lys Asp Lys Gly Glu Ile Arg Asn Asn Trp Glu Gly Ser Ser Asp
210 215 220
Lys Tyr Ala Arg Met Ile Gly Gly Trp Leu Asp Lys Leu Gly Leu Val
225 230 235 240
Lys Gln Gly Lys Lys Glu Phe Ile Ile Pro Thr Leu Gly Lys Pro Asp
245 250 255
Asn Lys Glu Phe Ile Ser His Ala Phe Lys Ile Thr Gly Glu Gly Leu
260 265 270
Lys Val Leu Arg Arg Ala Lys Gly Ser Thr Lys Phe Thr Arg Val Pro
275 280 285
Lys Arg Val Tyr Trp Glu Met Leu Ala Thr Asn Leu Thr Asp Lys Glu
290 295 300
Tyr Val Arg Thr Arg Arg Ala Leu Ile Leu Glu Ile Leu Ile Lys Ala
305 310 315 320
Gly Ser Leu Lys Ile Glu Gln Ile Gln Asp Asn Leu Lys Lys Leu Gly
325 330 335
Phe Asp Glu Val Ile Glu Thr Ile Glu Asn Asp Ile Lys Gly Leu Ile
340 345 350
Asn Thr Gly Ile Phe Ile Glu Ile Lys Gly Arg Phe Tyr Gln Leu Lys
355 360 365
Asp His Ile Leu Gln Phe Val Ile Pro Asn Arg Gly Val Thr Lys Gln
370 375 380
Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His Lys
385 390 395 400
Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg
405 410 415
Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe Phe
420 425 430
Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser Arg Lys
435 440 445
Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly Val
450 455 460
Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly
465 470 475 480
Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn Gln Thr Arg Asn
485 490 495
Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser Val

500 505 510

Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn Tyr
515 520 525

Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn Gly Ala
530 535 540

Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys Ala
545 550 555 560

Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly Glu
565 570 575

Ile Asn Phe

<210> 2

<211> 196

<212> PRT

<213> 海床黄杆菌(Flavobacterium okeanokoites)

<400> 2

Gln Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His
1 5 10 15

Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala
20 25 30

Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe
35 40 45

Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser Arg
50 55 60

Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly
65 70 75 80

Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile
85 90 95

Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn Gln Thr Arg
100 105 110

Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser
115 120 125

Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn
130 135 140

Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn Gly
145 150 155 160

Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys
165 170 175

Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly
180 185 190

Glu Ile Asn Phe
195

<210> 3
<211> 17
<212> DNA
<213> 人工序列
<220>
<223> TALEN-A的目标序列
<400> 3
agccgaaatc atcgcag 17

<210> 4
<211> 15
<212> DNA
<213> 人工序列
<220>
<223> GFP_TALEN(L)的目标序列
<400> 4
cagcgtgtcc ggcga 15

<210> 5
<211> 15
<212> DNA
<213> 人工序列
<220>
<223> GFP_TALEN(R)的目标序列
<400> 5
ttgccgtagg tggca 15

<210> 6
<211> 14
<212> DNA
<213> 人工序列
<220>
<223> B2M_TALEN(L)的目标序列
<400> 6
ccaaagattc aggt 14

<210> 7
<211> 14
<212> DNA
<213> 人工序列
<220>

<223> B2M_TALEN(R)的目标序列
<400> 7
gactttccat tctc 14
<210> 8
<211> 50
<212> DNA
<213> 人工序列
<220>
<223> TALEN-A的目标部位
<400> 8
agccgaaatc atcgcagccg ctgccgcgag ctcctgcgat gatttcggct 50
<210> 9
<211> 42
<212> DNA
<213> 人工序列
<220>
<223> GFP_TALEN的目标部位
<400> 9
cagcgtgtcc ggcgagggcg agggcgatgc cacctacggc aa 42
<210> 10
<211> 47
<212> DNA
<213> 人工序列
<220>
<223> B2M_TALEN的目标部位
<400> 10
ccaaagattc aggtttactc acgtcatcca gcagagaatg gaaagtc 47
<210> 11
<211> 196
<212> PRT
<213> 人工序列
<220>
<223> M47突变体
<400> 11

Gln	Leu	Met	Lys	Ser	Glu	Leu	Glu	Glu	Lys	Lys	Ser	Glu	Leu	Arg	Arg
1				5				10					15		
Lys	Leu	Lys	Tyr	Val	Pro	His	Glu	Tyr	Ile	Glu	Leu	Ile	Glu	Ile	Ala
			20					25				30			
Arg	Asn	Ser	Thr	Gln	Tyr	Arg	Ile	Phe	Glu	Met	Lys	Val	Met	Glu	Phe

35 40 45

Leu Val Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser Arg
50 55 60
Asn Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly
65 70 75 80
Val Ile Ile Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile
85 90 95
Gly Gln Ala Asp Ala Met Leu Arg Tyr Val Glu Glu Asn Gln Thr Arg
100 105 110
Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser
115 120 125
Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn
130 135 140
Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn Gly
145 150 155 160
Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys
165 170 175
Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly
180 185 190
Glu Ile Asn Phe
195

<210> 12
<211> 196
<212> PRT
<213> 人工序列
<220>
<223> M48突变体
<400> 12

Gln Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His
1 5 10 15
Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala
20 25 30
Arg Asn Pro Thr Leu Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe
35 40 45
Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser Arg
50 55 60
Lys Pro Asp Gly Val Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly
65 70 75 80
Val Ile Val Asp Thr Lys Ala Tyr Ser Arg Gly Tyr Asn Leu Pro Ile

85 90 95
Gly Gln Ala Asp Glu Met His Arg Tyr Val Val Glu Asn Gln Thr Arg
100 105 110
Asn Lys Leu Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser
115 120 125
Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn
130 135 140
Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Lys Gly
145 150 155 160
Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Arg
165 170 175
Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Ser Lys Phe Asn Asn Gly
180 185 190
Glu Ile Asn Phe
195

<210> 13
<211> 196
<212> PRT
<213> 人工序列
<220>
<223> M49突变体
<400> 13

Gln Leu Met Lys Ser Glu Leu Glu Glu Lys Lys Phe Glu Leu Arg Arg
1 5 10 15
Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala
20 25 30
Arg Asn Ser Thr Gln Tyr Arg Ile Phe Glu Met Lys Val Met Glu Phe
35 40 45
Leu Val Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser Arg
50 55 60
Asn Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly
65 70 75 80
Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile
85 90 95
Gly Gln Ala Asp Ala Met Leu Arg Tyr Val Glu Glu Asn Gln Thr Arg
100 105 110
Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser
115 120 125
Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn

130 135 140

Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn Gly
145 150 155 160
Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys
165 170 175
Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly
180 185 190
Glu Ile Asn Phe
195

<210> 14
<211> 196
<212> PRT
<213> 人工序列
<220>
<223> M50突变体
<400> 14

Gln Pro Val Lys Ser Glu Leu Glu Glu Lys Lys Ala Glu Leu Arg His
1 5 10 15
Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala
20 25 30
Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Met Val Met Glu Phe
35 40 45
Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser Arg
50 55 60
Lys Pro Asp Gly Val Ile Tyr Thr Val Gly Ser Pro Phe Asp Tyr Gly
65 70 75 80
Val Ile Val Asp Thr Lys Ala Tyr Ser Arg Gly Tyr Asn Leu Pro Ile
85 90 95
Gly His Ala Asp Glu Met Gln Arg Tyr Val Gln Glu Asn Gln Thr Arg
100 105 110
Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser
115 120 125
Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn
130 135 140
Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Thr Gly
145 150 155 160
Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Arg
165 170 175
Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Ser Lys Phe Asn Asn Gly

180 185 190

Glu Ile Asn Phe
195

<210> 15
<211> 14
<212> DNA
<213> 人工序列
<220>
<223> B2M_PPR(L)的目标序列
<400> 15
aatggaaagt caaa 14
<210> 16
<211> 14
<212> DNA
<213> 人工序列
<220>
<223> B2M_PPR(R)的目标序列
<400> 16
agcaattcag gaaa 14
<210> 17
<211> 15
<212> DNA
<213> 人工序列
<220>
<223> GFP_PPR(L)的目标序列
<400> 17
acgacttctt caagt 15
<210> 18
<211> 15
<212> DNA
<213> 人工序列
<220>
<223> GFP_PPR(R)的目标序列
<400> 18
cctggacgta gcctt 15
<210> 19
<211> 96
<212> DNA
<213> 人工序列

<220>

<223> B2M_PPR的目标部位

<400> 19

tactccaaag attcaggttt actcacgtca tccagcagag aatggaaagt caaatttcct 60

gaattgctat gtgtctgggt ttcatccatc cgacat 96

<210> 20

<211> 212

<212> DNA

<213> 人工序列

<220>

<223> GFP_PPR的目标部位

<400> 20

caagttcagc gtgtccggcg agggcgaggg cgatgccacc tacggcaagc tgaccctgaa 60

gttcatctgc accaccggca agctgcccgt gccctggccc accctcgtga ccaccctgac 120

ctacggcgtg cagtgcttca gccgctaccc cgaccacatg aagcagcacg acttcttcaa 180

gtccgccatg cccgaaggct acgtccagga gc 212

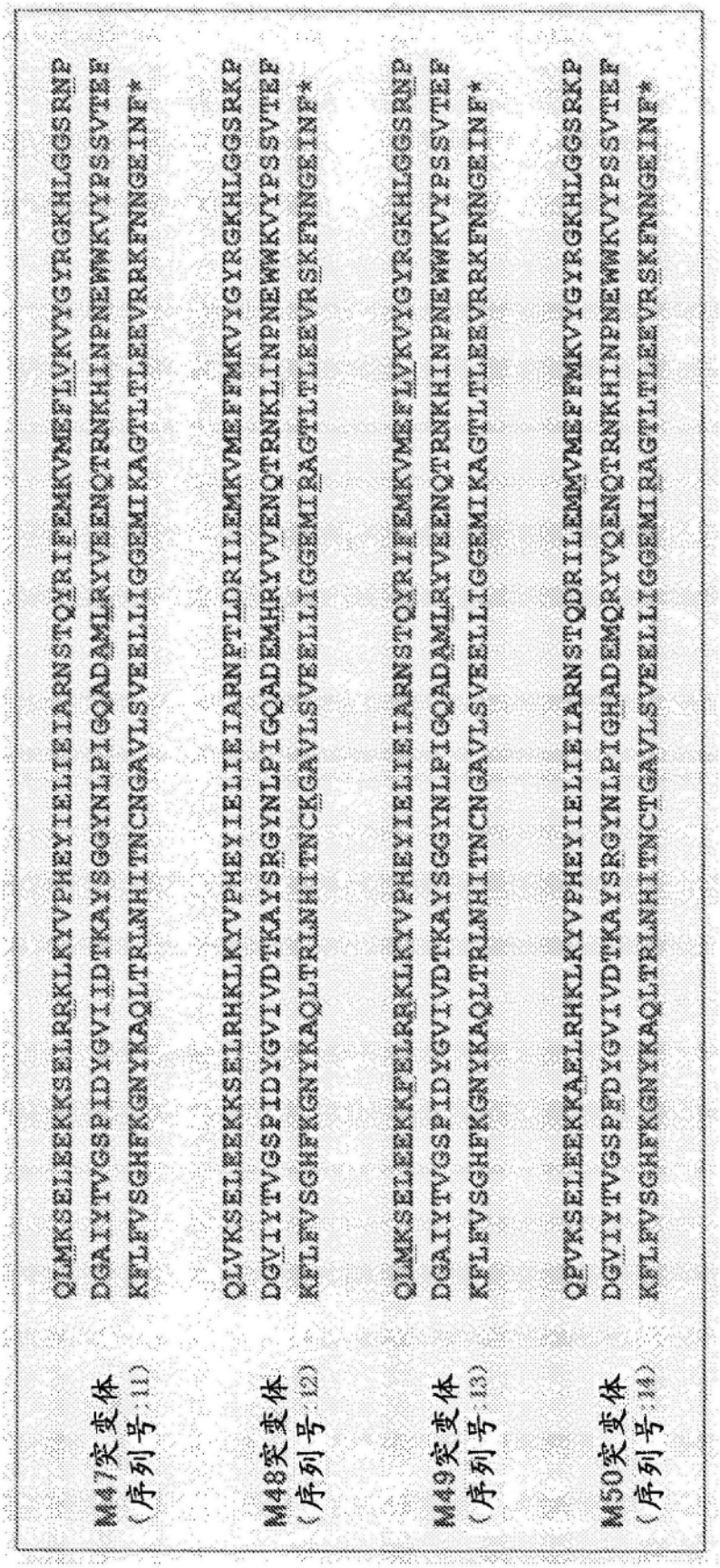

M47突变体
（序列号:11）
QIMKSELEEKKSELRRKLKYVPHEYIELIEIARNSTQYRIIEMKVMEFLVKVYGYRGKHLGGSRKP
DGAIYTVGSPIDYGVIIDTKAYSGGYNLPIGQADAMLRYVEENQTRNKHINPNEWWKVYPSSVTEF
KFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELIIGGEMIKAGTLTLEEVRRKFNNGEINF*
M48突变体
（序列号:12）
QIVKSELEEKKSELRHKLKYVPHEYIELIEIARNPTLDRILEMKVMEFMKVYGYRGKHLGGSRKP
DGVIYTVGSPIDYGVIVDTKAYSRGYNLPIGQADEMHRYVVENQTRNKLINPNEWWKVYPSSVTEF
KFLFVSGHFKGNYKAQLTRLNHITNCKGAVLSVEELIIGGEMIRAGTLTLEEVRSKFNNGEINF*
M49突变体
（序列号:13）
QIMKSELEEKKFELRRKLKYVPHEYIELIEIARNSTQYRIIEMKVMEFLVKVYGYRGKHLGGSRNP
DGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADAMLRYVEENQTRNKHINPNEWWKVYPSSVTEF
KFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELIIGGEMIKAGTLTLEEVRRKFNNGEINF*
M50突变体
（序列号:14）
QPVRSELEEKKAELRHKLKYVPHEYIELIEIARNSTQDRILEMMVMEPFMKVYGYRGKHLGGSRKP
DGVIYTVGSPFDYGVIVDTKAYSRGYNLPIGHADEMQRYVQENQTRNKHINPNEWWKVYPSSVTEF
KFLFVSGHFKGNYKAQLTRLNHITNCTGAVLSVEELIIGGEMIRAGTLTLEEVRSKFNNGEINF*

图1

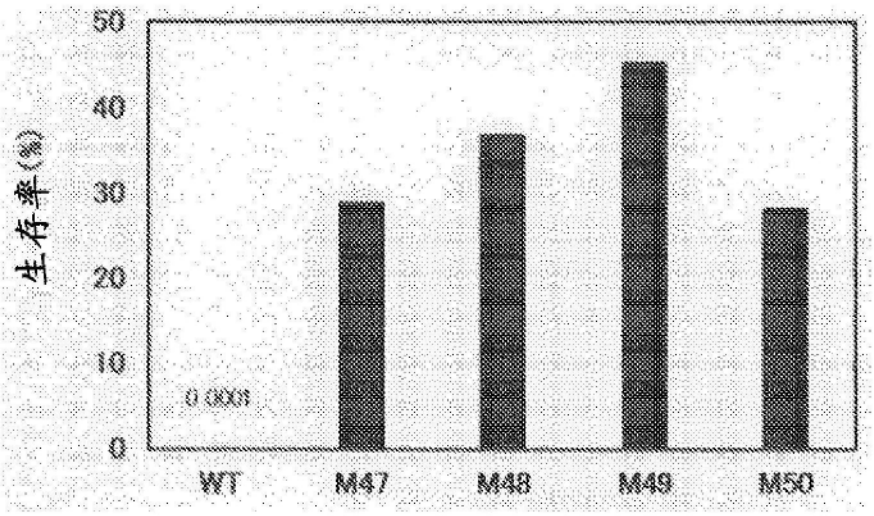

生存率(%)
50
40
30
20
10
0
0.000
WT
M47
M48
M49
M50

图2

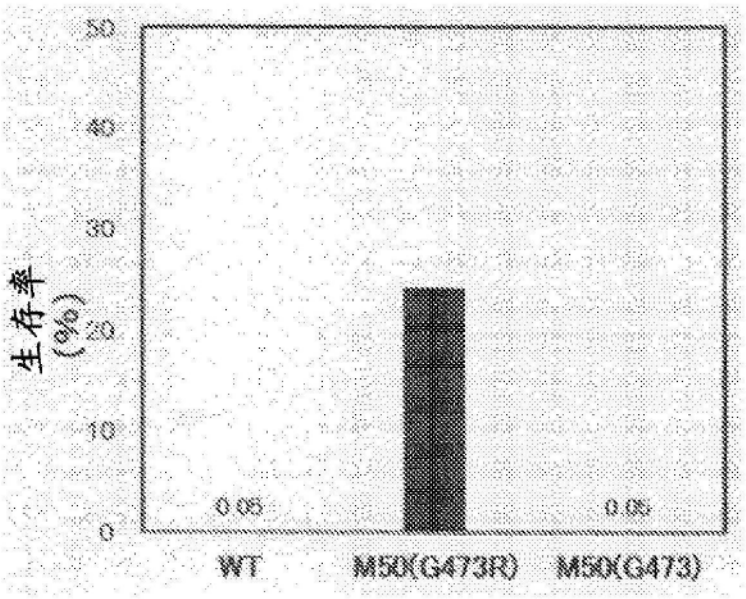

生存率(%)
50
40
30
20
10
0
0.05
0.05
WT
M50(G473R)
M50(G473)

图3

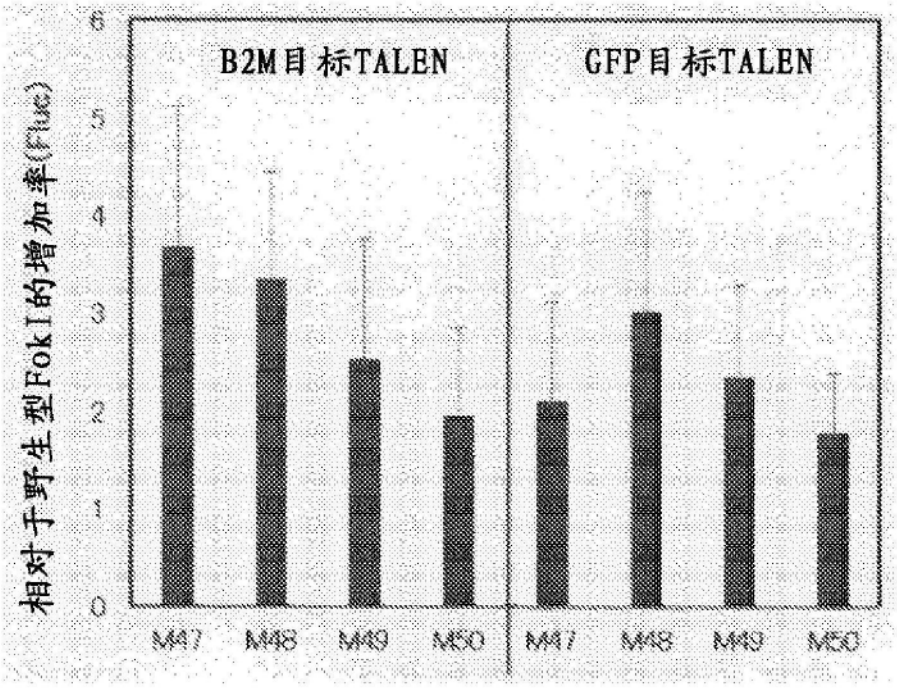

B2M目标TALEN
GFP目标TALEN
相对于野生型FokI的增加率(Fluc)
0
1
2
3
4
5
6
M47
M48
M49
M50
M47
M48
M49
M50

图4

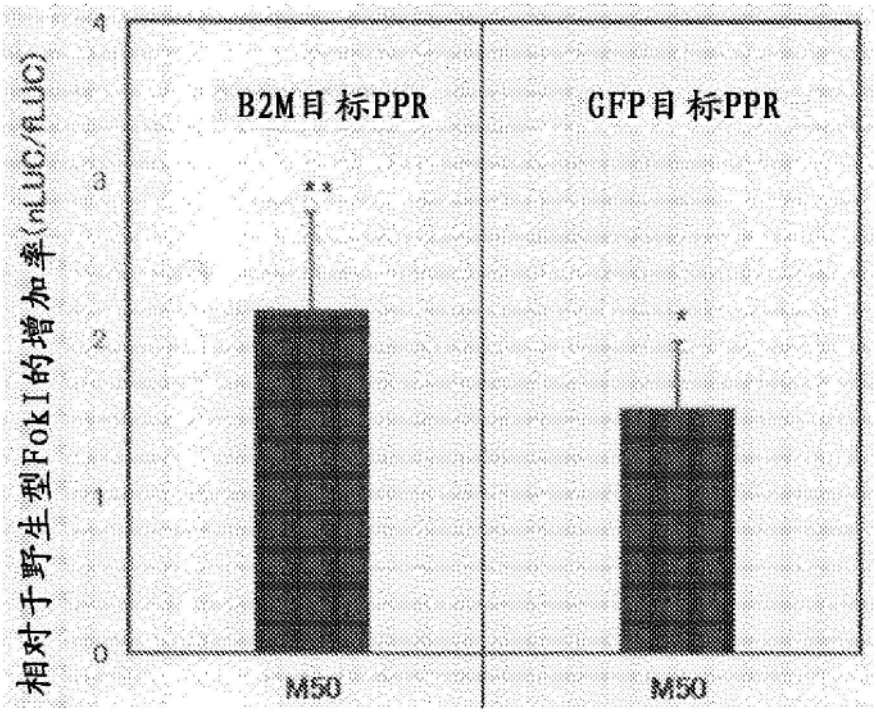

B2M目标PPR
GFP目标PPR
相对于野生型FokI的增加率(nLUC/fLUC)
0
1
2
3
4
**
*
M50
M50

图5