

一种 CHO 细胞基因组内 NW_006882456-1 稳定表达蛋白质的应用

技术领域

本发明涉及基因技术领域，尤其是涉及一种用于稳定表达蛋白质的 CHO 细胞重组基因。

背景技术

中国仓鼠卵巢细胞(Chinese Hamster Ovary cell, CHO)作为生物制药领域的主力细胞系，已经被开发出许多不同种类的 CHO 细胞系，包括可以用来扩大基因拷贝数的细胞系。然而，转基因拷贝数的增多与目的蛋白质产率提高并不是明确的正相关关系。且即使蛋白质表达增多，多数 CHO 细胞的表达水平不稳定。当前普遍使用的构建稳定转染细胞的方法耗时耗力，主要是因为需要重复大量的单克隆筛选过程，所以当前在细胞系构建领域普遍期待可以开发出一种能在短时间内，获得高表达及稳定表达的细胞的方法，并且能够确保构建出来的重组细胞系与传统方法相比，有相同的质量水平，以确保监管机构的批准。

传统构建外源蛋白质表达细胞系的方法是通过外源基因随机整合到细胞基因组上，再经过系列的高表达单克隆细胞筛选，获得外源蛋白质高表达细胞系。由于位点效应差异的多样性，随机整合产生的重组细胞表达水平各异，因此需要后期花费很长的时间和很多的步骤，用于挑选高表达单克隆细胞。通过随机整合获得的单克隆细胞无法保证多肽/蛋白质在细胞传代中稳定表达，每进行一次重组细胞构建均需要反复进行单克隆筛选，增加了生物制药的研发成本。

位点效应妨碍着传统随机整合构建重组细胞系的效率，反复的高表达单克隆筛选费时费力，且成本昂贵。如何克服位点效应，利用定点整合技术，快速高效地获得稳定表达单克隆细胞已经在学术界被讨论了多年，一直未有突破性进展。

发明内容

针对现有技术存在的上述问题，本发明申请人提供了一种 CHO 细胞基因组内 NW_006882456-1 稳定表达蛋白质的应用。本发明申请在 CHO 细胞基因组内的固定位置，导入不同蛋白质基因，并进行稳定表达；此外在实现该定点整

合的过程中，无需多次挑选单克隆以获得更高的表达细胞株，节省了大量的时间。

本发明的技术方案如下：

一种 CHO 细胞基因组内某位点稳定表达蛋白质的应用，所述 CHO 细胞基因组内用于稳定表达蛋白质的位点位于 CHO 细胞基因 NW_006882456.1 的第 1020651 碱基处；

所述位点附近 1020585-1020690 范围内可被 CRISPR/Cas9 技术识别的 5' NNNNNNNNNNNNNNNNNNNNNNGG3' 为靶序列。

所述蛋白质为分子量小于 160KDa 的蛋白质。

所述蛋白质为多肽、功能性蛋白质、抗体、融合蛋白质中的一种。

所述靶序列为 CHO 细胞基因 NW_006882456.1 的第 1020651 碱基处附近第 1020640-1020662 碱基处。

进一步所述靶序列为 5'-ACCCTTGTGCCCCAAAGACAGGG-3'。

所述靶序列为 5'-ACCCTGTCTTTGGGGCACAAGGG-3'。

所述靶序列为 5'-GTCTTTGGGGCACAAGGGTGAGG-3'。

所述靶序列为 5'-TTTGGGGCACAAGGGTGAGGTGG-3'。

所述靶序列为 5'-TTGGGGCACAAGGGTGAGGTGGG-3'。

所述靶序列为 5'-CAAGGGTGAGGTGGGGTAAATGG-3'。

所述靶序列为 5'-GGTGAGGTGGGGTAAATGGATGG-3'。

所述靶序列为 5'-AGGTGGGGTAAATGGATGGTTGG-3'。

所述靶序列为 5'-GGGGTAAATGGATGGTTGGTTGG-3'。

所述靶序列为 5'-TGGGGCACAAGGGTGAGGTGGGG-3'。

所述靶序列为 5'-ACAGAGAAACCCTGTCTTTGGGG-3'。

所述靶序列为 5'-CACAGAGAAACCCTGTCTTTGGG-3'。

所述靶序列为 5'-AGTGAGTTCTAAGACAGCCAGGG-3'。

所述靶序列为 5'-ACACAGAGAAACCCTGTCTTTGG-3'。

所述靶序列为 5'-GTTTCTCTGTGTAACAGCCCTGG-3'。

所述靶序列为 5'-ACTCACTCTGTAGACCAGGCTGG-3'。

一种靶向载体，该靶向载体含有所述可用于位点编辑的能够被 CRISPR/Cas9 技术识别的 5' NNNNNNNNNNNNNNNNNNNNNNGG3' 的靶序列。

所述靶向载体的制作方法为：将 PSK-u6-gRNA 质粒进行 BbsI 酶切，与含有 BbsI 粘性末端的靶序列 DNA 片段连接，得到靶向载体。

一种用于 CHO 细胞表达外源基因的重组供体载体，该重组供体载体含有所述位点 NW_006882456.1 的第 1020651 碱基处上下游 500bp-800bp 同源臂。

所述的重组供体载体的制作方法为：将蛋白质基因插入到质粒 5'arm 和 3'arm 中间的区域，使该核苷酸序列位于启动子的下游并受其调控，得到重组 CHO 细胞表达质粒。

所述启动子为：CMV（人类巨细胞病毒来源的强哺乳动物表达启动子）、EF-1a（人延长因子 1 α 来源的强哺乳动物表达启动子）、SV40（猿猴空泡病毒 40 来源的哺乳动物表达启动子）、PGK1（磷酸甘油酸酯激酶基因来源的哺乳动物启动子）、UBC（人类泛素 C 基因来源的哺乳动物启动子）、human beta actin（ β -肌动蛋白基因来源的哺乳动物启动子）、CAG（强杂交哺乳动物启动子）中的一种。

一种用于稳定表达外源蛋白质的 CHO 重组细胞系，该重组细胞系构建过程中使用了所述的载体或质粒。

一种 CHO 细胞基因表达蛋白质的方法，其特征在于，包括以下步骤：

（1）用 Cas9 及所述靶向载体和重组供体载体，三种载体共同转化 CHO 细胞，得重组 CHO 细胞；

（2）重组 CHO 细胞在平板上培养，收集上清检测表达水平，以及对贴壁重组 CHO 细胞进行悬浮驯化；

（3）悬浮驯化的重组 CHO 细胞在小摇瓶内进行培养，并检测蛋白质的表达水平。

本发明申请还提供一种 CHO 细胞基因组中稳定表达位点的选择：

1)构建带荧光标签的慢病毒，并推算其滴度。将 igk-luc 基因整合到 pLVX-CMV-MCS-T2A-Zsgreen 载体的多克隆位点后，再同时利用 pSPAX2、pMD2G 两个质粒进行三质粒转染至 HEK-293T 细胞，于 48 小时及 72 小时两次取细胞上清液，收集后超速离心上清液获取慢病毒。

2)将 CHO 细胞铺在 6 孔板上培养过夜，第二天将慢病毒稀释并以较低的 MOI（MOI<1）（每个细胞对应病毒颗粒）来感染 CHO 细胞。感染 96 小时后，通过流式细胞分选仪分选，将荧光强度最亮的部分细胞直接接种到 96 孔板内。一周后等细胞长成单克隆集落时，荧光显微镜下观察细胞，选择形态和生长正

常，且最亮的集落细胞，转移到 24 孔板内扩大培养。培养细胞达汇合度近 90% 时再转移到 6 孔板内培养，最后再扩大至 10cm 的培养皿中培养，冻存一部分细胞，剩余细胞继续扩大培养。

3) 利用染色体步移技术 Lenti-X Integration Site Analysis Kit (Clontech: 631263) 寻找到慢病毒所有的 CHO 细胞基因整合位点。

以荧光强度最亮、细胞形态和生长速度正常的若干个细胞系为材料，采用 ADraI, SspI, HpaI 三个限制性内切酶对基因组 DNA 进行过夜酶切。其中基因组 DNA 有 2.5 μ g，限制性内切酶有 80U，配成 100 μ L 反应体系。置于 37°C 下过夜酶切（16-18 小时）。

利用 DNA 回收试剂盒将酶切产物纯化回收。将酶切后的基因组 DNA 4.8 μ L，加上 1.9 μ L 染色体步移接头 GenomeWalker Adaptor (25 μ M) 和 0.5 μ L T4 连接酶，配成 8 μ L 体系进行连接实验。置于 16°C 环境下，进行过夜连接。将连接体系 70°C 下加热 5 分钟，使连接酶失活。各个体系内加入 32 μ L TE buffer 配置成相应的文库。

对文库进行 2 轮巢式 PCR，将 LTR 区域同临近的基因组区域扩增出来。相关的 PCR 反应操作步骤可以参照 Lenti-X Integration Site Analysis Kit (Clontech: 631263) 试剂盒说明书进行。

最后将 PCR 产物进行电泳，将主要条带切胶回收后送测序。在获得每个细胞系所有慢病毒整合信息后，选择只有单拷贝慢病毒整合的 CHO 细胞系的相关信息，并将其序列信息在 BLAST 上同 CHO-K1 基因组信息加以比对，找出高表达的整合位点。

由于数据库更新，经序列比对，原所述位点信息由 CHO 细胞基因 NW_006882456.1 的第 1020651 碱基处，变更为 CHO 细胞基因 NW_003614241.1 的第 148091 碱基处。所述范围由 1020585-1020690 变更为 148052 -148157。

本发明有益的技术效果在于：

本发明采用定点整合方法，通过把目的基因定点整合到稳定表达区域，能很好的克服随机整合带来的整合位点不确定性的问题，有效的规避了反复多轮的高表达单克隆筛选，能有效的减少生物制药构建稳定表达细胞系的研发时间，降低成本。

本发明在 CHO 细胞基因的定点位置，导入蛋白质基因，并进行稳定表达。

附图说明

图 1 为本发明示意图；

图 2 为导入 NGGH 75KDa 基因的 CHO 细胞基因鉴定结果；

图 3 为 OoPCR_fwd 及 OoPCR_rev 对导入 NGGH 75KDa 基因的 CHO 细胞基因进行的测序；

图 4 为不同传代下的细胞表达 HSA 的情况；

图 5 为不同传代下的细胞表达 NGGH 的情况；

图 6 为每日每个重组 CHO 细胞分泌抗体蛋白的质量情况。

具体实施方式

下面结合附图和实施例，对本发明进行具体描述。

图 1 为用于整合到该位点的供体质粒示意图和如何通过同源重组定点整合到该位点的模拟示意图。其中 GOI 为我们的目的基因，其通过 5'arm 和 3'arm 两个同源重组臂，在 4 μ g/mL 的 puromycin 筛选压力下，定点整合到靶点位置。此外 5'arm 上游的序列为负筛选标签，可以用来去除掉随机整合的单克隆细胞，确保最终获得的为定点整合的重组 CHO 细胞系。

实施例 1

高表达位点的选择；

NW_006882456.1 (NW_003614241.1) (6 号染色体)，1020651 (148091) 碱基处整合，Zsgreen1 基因被整合在了该处。对于这个荧光细胞，进行了不少于 50 代次的传代培养，通过流式细胞仪检测其荧光的表达水平。第 50 代次荧光细胞仍有较好的绿色荧光蛋白表达水平，荧光信号在细胞的传代过程中能稳定的保留。

此外对于这个荧光细胞还进行了悬浮驯化，并对其悬浮驯化后的荧光蛋白表达水平利用流式细胞仪再次进行了检测。检测结果显示，悬浮达到 50 代次的重组 CHO 细胞中，超过 95% 以上的细胞在悬浮后依然保持了绿色荧光蛋白的表达，可以认为该位点是极其稳定，不会因细胞传代丢失定点整合的荧光蛋白基因。

实施例 2

具体靶点的选择;

根据就近原则, 将序列:

5'CCAGCCTGGTCTACAGAGTGAGTTCTAAGACAGCCAGGGCTGTTAC
ACAGAGAAACCCTGTCTTTGGGGCACAAGGGTGAGGTGGGGTAAATGGA
TGGTTGGTTGG 3'

输入 CRISPRater 系统中, 预测并挑选出脱靶效率较低的靶序列。其中参数设置如下所示: 1) NGG 后开头 15bp 的最大错配碱基数为 0; 2) NGG 后全部 21bp 的错配碱基数为 2。

根据如上操作后, 根据其评分选择分数为 0.62 的如下序列作为靶序列:

5'-ACCCTTGTGCCCCAAAGACAGGG-3';

根据 CRISPRater 系统, LOW efficacy (score<0.56); MEDIUM efficacy (0.56<=score<=0.74); HIGH efficacy (score>0.74)。

根据 CRISPRater 评定系统, 位点 NW_006880285.1 (NW_003614241.1) 附近 1020585-1020690 (148052 -148157) 范围内所有靶序列均获得了 0.56 以上的分数, 均处于中等有效或高效范围, 可作为 CRISPR/Cas9 技术识别的 5' NNNNNNNNNNNNNNNNNNNNNNNNNNNNGG3'靶序列。

实施例 3 启动子的选择

将不同的启动子替换上述 CMV (人类巨细胞病毒来源的强哺乳动物表达启动子) 启动子位置处, 包含了 EF-1a (人延长因子 1 α 来源的强哺乳动物表达启动子)、SV40 (猿猴空泡病毒 40 来源的哺乳动物表达启动子)、PGK1 (磷酸甘油酸酯激酶基因来源的哺乳动物启动子)、UBC (人类泛素 C 基因来源的哺乳动物启动子)、human beta actin (β -肌动蛋白基因来源的哺乳动物启动子) 或 CAG (强杂交哺乳动物启动子) 等常见启动子。经检测, 以上启动子均能调控下游的红色荧光蛋白基因序列, 并表达出相应红色荧光蛋白。

实施例 4

将人血清白蛋白基因(HSA, 68KDa)定点整合在特定位点: 为了后期进行构建 CRISPR/Cas9 介导的同源重组, 需要构建 sgRNA 和 Donor Plasmid, 方法如下:

1、SgRNA 的构建, 首先合成如下序列

sgRNA-1fwd 5' TTTGACCCTTGTGCCCCAAAGACAGT 3'

sgRNA-1rev 5' TAAAACTGTCTTTGGGGCACAAGGGT 3'

1) 将 PSK-u6-gRNA 质粒进行 BbsI 酶切, 将切后的载体回收;

2) 合成的片段退火成带有粘性末端的双链

sgRNA-1fwd (100 μ M) 4 μ L

sgRNA-1rev (100 μ M) 4 μ L

10 $\times$ NEB buffer2 2 μ L

H₂O 10 μ L

20 μ L

水浴 95 $^{\circ}$ C 5min, 水浴锅中自然降至室温;

3) 片段连接, 重组质粒的构建

回收的线性化载体 50ng

退火片段 1 μ L

10 $\times$ T4 ligase buffer 1 μ L

T4 ligase 1 μ L

H₂O

室温连接 1h 以上,
或者 4 $^{\circ}$ C 连接过夜

至 10 μ L;

4) 连接, 转化;

5) 挑取克隆, PCR 鉴定, 使用的鉴定引物为:

M13- 合成的 R 的引物, 可以扩增出条带的为阳性克隆。

2、Donor plasmid 构建

donor plasmid 具体信息如图 1 所示: 除 GOI 之外, 其余部分均为合成; 靶点上下游 600bp 序列为 donor plasmid 的左右同源臂序列信息, GOI 为 HSA 基因通过已有的诺维赞公司 C115 试剂盒完成 HSA 整合到 donor plasmid 上这个过程。

3、将 Cas9 (丹麦科技大学 Dr. Helene F Kildegaard 捐赠)、SgRNA 及 Donor plasmid 3 个质粒共同转染 37 $^{\circ}$ C、5% CO₂ 条件下培养的 CHO 细胞, 三种质粒摩尔比为 1:1:1, 转染试剂为 Lipofectamine 3000 (Thermo Fisher Scientific), 具体转染方法参考说明书, 之后细胞添加 4 μ g/mL puromycin 进行

筛选，过程共计 10 天；再用 MoFloXDP FACS machine (Beckman Coulter) 进行单克隆细胞分选，选择不含任何荧光的细胞，接种到 96 孔板内。

4、细胞生长 2 周后，取一部分进行鉴定，通过做 PCR 鉴定，做 5'junction PCR，3'Junction PCR 及 out-out PCR 进行鉴定。将阳性细胞保留。

实施例 5

将胰高血糖素样肽-1-人血清白蛋白融合蛋白基因(NGGH, 75KDa)定点整合在特定位点：为了后期进行构建 CRISPR/Cas9 介导的同源重组，需要构建 sgRNA 和 Donor Plasmid，方法如下：

1、构建 sgRNA 同实施例 4。

2、Donor plasmid 构建。

donor plasmid 具体信息如图 1 所示：除 GOI 之外，其余部分均为合成；靶点上下游 600bp 序列为 donor plasmid 的左右同源臂序列信息，GOI 为 NGGH 基因通过已有的诺维赞公司 C115 试剂盒完成 NGGH 整合到 donor plasmid 上这个过程。

3、将 Cas9、SgRNA 及 Donor plasmid 3 个质粒共同转染 37℃、5%、CO₂ 条件下培养的 CHO 细胞，三种质粒摩尔比为 1:1:1，转染试剂为 Lipofectamine 3000 (Thermo Fisher Scientific)，具体转染方法参考说明书。之后细胞添加 4 μg/mL puromycin 进行筛选，过程共计 10 天。再用 MoFloXDP FACS machine (Beckman Coulter) 进行单克隆细胞分选，选择不含任何荧光的细胞，接种到 96 孔板内。

4、细胞生长 2 周后，取一部分进行鉴定。通过做 PCR 鉴定，做 5'junction PCR，3'Junction PCR 及 out-out PCR 进行鉴定。将阳性细胞保留。

实施例 6

将抗体基因(Avastin, 160KDa)定点整合在特定位点：为了后期进行构建 CRISPR/Cas9 介导的同源重组，需要构建 sgRNA 和 Donor Plasmid，方法如下：

1、sgRNA 构建同实施例 4。

2、Donor plasmid 构建。

donor plasmid 具体信息如图 1 所示：除 GOI 之外，其余部分均为合成；靶点上下游 600bp 序列为 donor plasmid 的左右同源臂序列信息，GOI 为 Avastin 基因

通过已有的诺维赞公司 C115 试剂盒完成抗体基因整合到 donor plasmid 上这个过程。

3、将 Cas9（丹麦科技大学 Dr. Helene F Kildegaard 捐赠）、SgRNA 及 Donor plasmid 3 个质粒共同转染 37°C、5%、CO₂ 条件下培养的 CHO 细胞，三种质粒摩尔比为 1:1:1，转染试剂为 Lipofectamine 3000 (Thermo Fisher Scientific)，具体转染方法参考说明书。之后细胞添加 4 µg/mL puromycin 进行筛选，过程共计 10 天。再用 MoFloXDP FACS machine (Beckman Coulter) 进行单克隆细胞分选，选择不含任何荧光的细胞，接种到 96 孔板内。

4、细胞生长 2 周后，取一部分进行鉴定。通过做 PCR 鉴定，做 5'junction PCR，3'Junction PCR 及 out-out PCR 进行鉴定。将阳性细胞保留。

图 2 为导入 NGGH 75KDa 基因的 CHO 基因鉴定结果，其中泳道 1-3 为三个单克隆细胞的 5'junction PCR 结果，泳道 4-5 为 3'junction PCR 结果，均有显著条带，证实有基因敲入。

图 3 为利用 OoPCR_fwd 及 OoPCR_rev 进行测序，确定交界处（5'junction 及 3'junction 的 5'上游处与基因组的交界处）的序列准确，测序结果验证了 GOI 被精准的插入到了靶点位置区域。

测试例：

在对实施例 4-6 制得的这三个细胞系进行 ELISA 检测，观察是否有目的蛋白质表达，且是否为稳定的长期表达。

检测的方法：四者的检测均通过 ELISA 方法进行，在 6 孔板内培养挑选到的所有阳性细胞，检验其是否有目的蛋白质长期稳定表达，使用的试剂盒为 Abclonal 公司的 Human Albumin ELISA Kit (RK00157) 和 Human IgG(Total) ELISA Kit(RK00393)进行实验。

图 4 与图 5 为 HSA 及 NGGH 在不同传代下的细胞表达情况，纵坐标为每个细胞每天分泌蛋白质质量。

可以看到图中 NGGH 和 HSA 在平板内可以在 50 代次以内稳定地表达相应基因，且挑选到的 3 个 NGGH 定点整合细胞系及 2 个 HSA 定点整合细胞系目的蛋白质表达水平趋于接近。

图 6 为每日每个重组 CHO 细胞分泌抗体蛋白的质量。显然，细胞在不同传

代条件下均能稳定地表达分泌出相应的蛋白，体现了很好的稳定性，与之前的荧光细胞结果相互吻合。结果显示该位点可以进行 CRISPR/Cas9 介导的定点整合，并可以稳定表达出相应的蛋白质。

在选取了 5'-ACCCTTGTGCCCCAAAGACAGGG-3' 序列进行如上测试后，取得了良好结果，因此提出权利要求 5-19 所述靶序列皆能成功构建出定点整合稳定表达细胞系，且均能稳定表达出目的蛋白。