

## I 基本情報

研究開発課題名： (日本語) 新規高核局在性 Cas9 による高効率 *in vivo* ゲノム編集法の開発  
(英語) Application of novel highly nuclear-localized Cas9 to *in vivo* genome editing

研究開発実施期間：2018年10月15日～2021年3月31日

研究開発代表者 氏名：(日本語) 水野 聖哉  
(英語) Mizuno Seiya

研究開発代表者 所属機関・部署・役職：  
(日本語) 筑波大学・医学医療系・准教授  
(英語) University of Tsukuba・Faculty of Medicine・Associate Professor

## II 研究開発の概要

研究開発の成果およびその意義等

CRISPR-Cas9 を利用した *in vivo* ゲノム編集での遺伝子治療の臨床治験が開始されており、その効果と安全性に大きな注目が集まっている。ゲノム DNA は核にあるが、Cas9 タンパクは原核生物に由来し、真核細胞においては核局在性をもたない。少ない Cas9 タンパクもしくはそれをコードする RNA または DNA にて、効果的に *in vivo* ゲノム編集を実施するためには、高核移行性の Cas9 の開発が重要である。ゲノム編集の安全性を脅かす要因に、標的外のゲノム領域で変異を生じさせるオフターゲットと並んで、オンターゲットサイトにおける大規模なゲノム領域欠損がある。

我々は、Cas9 に大変強力な核局在性を付与する新たな因子を発見した。この因子を付加した Cas9 (Cas9-FmC) は、培養細胞や受精卵で強く核に局在し、そのゲノム編集効率も上昇させた。Cas9-FmC が *in vivo* ゲノム編集の効率を上昇させることができれば、少ない投与量で最大限の治療(編集)効果を得ることができる。そこで、①Cas9-FmC の最適化、②Cas9-FmC および最適化 Cas9-FmC の各成体内組織における細胞内局在の評価、③Cas9-FmC および最適化 Cas9-FmC での *in vivo* ゲノム編集、④オンターゲットにおける編集効率と安全性も確認できる評価系の開発も実施した。

まず、①Cas9-FmC の最適化について記載する。我々は Cas9 タンパク質にマウス Cdt1 タンパク質の一部 (29-132AA) を結合させた Cas9 タンパク質 (以後、Cas9-FmC と記載) が真核細胞の核に強く局在する可能性を見出した。そこで、この Cas9-FmC に FLAG 配列を付加したタンパク質と任意の guide RNA の両者を発現するカセットを有するベクターである pX330-FmC を構築した。この pX330-FmC を HEK293T 細胞に導入し、FLAG に対する抗体を用いた免疫染色により Cas9 の局在を確認した結果、Cas9-FmC は高度に核に局在した。前述の通り、FmC のサイズは 104 アミノ酸であるため、そのコーディング配列は 312 塩基対と核移行性を付加する機能をもつ遺伝子配列としては比較的大きい。遺伝子治療のためにウィルスベクター等に Cas9 を搭載する

ためには、サイズダウンする必要があると考えた。そこで、FmC 内の一部だけを Cas9 に付加した Cas9-XmC と Cas9-YmC と Cas9-ZmC を構築した。そして、これらの Cas9 に FLAG を付加したタンパク質と任意の guide RNA の両者をそれぞれ発現する pX330-XmC、pX330-YmC、pX330-ZmC ベクターを構築した。これらの pX330-XmC、pX330-YmC、pX330-ZmC ベクターを HEK293T 細胞に導入し、FLAG に対する抗体を用いた免疫染色を行った結果、pX330-ZmC を導入した細胞でシグナルが核に検出された。次に、Cas9-FmC や Cas9-ZmC において、Cas9 の切断活性が阻害されていないかを検討するために、Cyp19a1 遺伝子の一部を切断する guide RNA 配列を pX330-FmC と pX330-ZmC にそれぞれ組込んで、その切断活性を *in vitro* で評価したところ、両方で切断活性が確認された。なお、更なるサイズダウンを目指して、ZmC を更に分割した Z1mC や Z2mC を Cas9 に付加したタンパク質の細胞内局在を上述の方法で確認したが、核への強い局在は確認されなかった。この結果から、Cas9 を核に局在させることが可能な最小の Cdt1 領域が決定され、かつ、この付加は Cas9 の切断活性を障害しないことを明らかにした。

次に、②Cas9-FmC および最適化 Cas9-FmC(Cas9-ZmC)の各成体内組織における細胞内局在の評価について記載する。Cas9-FmC と Cas9-ZmC と S-Cas9(SV40-NLS を付加した Cas9)をそれぞれ全身で発現するマウスの開発を実施した。当初の計画では、ストレートに各 Cas9 を全身で強く発現させる予定であったが、核に強く局在する Cas9 が全身で強発現することで、マウスの発育や稔性に影響を及ぼす可能性が考えられたために、Cre 依存的に各 Cas9 を発現するマウスを作出した。具体的には、CAG プロモーター配列の直下に loxP で挟まれた EGFP-rabbit Globin polyA(以後、rGpA と記載)を配置し、その下流にそれぞれの各 Cas9 と rGpA を繋いだカセットを ROSA26 遺伝子座にノックインするための相同アーム配列の間に設置した。これらのベクターをマウス受精卵ゲノム編集で ROSA 遺伝子座にノックインしたマウスの作出を試みた結果、目的の 3 系統が全て作出された。その後、Floxed-EGFP-Cas9-FmC マウスと Floxed-EGFP-Cas9-ZmC マウスは無事に系統化された。Floxed-EGFP-Cas9-FmC ヘテロノックインマウス及び Floxed-EGFP-Cas9-ZmC ヘテロノックインマウスから採取した精子と野生型マウスとの卵子をそれぞれ体外受精することで得た受精卵に *in vitro* transcription で作成した *Cre* mRNA をエレクトロポレーションで導入した。その結果、EGFP が染色体から脱落(loop-out)し、Cas9 を全身で強発現するマウスが誕生した。その後、これらのマウスをホモ化した。そのなかで異常な表現型を示す個体は観察されなかった。以上の結果から、当初懸念していた、Cas9-FmC や Cas9-ZmC の強発現はマウスの発育や稔性に影響を与えないことが明らかとなった。上記の方法で作出した Cas9-FmC 全身発現マウスと Cas9-ZmC 全身発現マウスから 18 組織(脳、肝臓、胸腺、空腸、骨格筋、子宮、心臓、腎臓、精巣、大腸、直腸、胃、皮膚、肺、卵巣、脾臓、膀胱、膵臓)を採材・切片化し、抗 Cas9 抗体で免疫染色することで、各組織における Cas9-FmC と Cas9-ZmC の細胞内局在を確認した。その結果、いずれの組織においても Cas9-FmC と Cas9-ZmC は核に存在していた。以上の結果から、Cas9-FmC と Cas9-ZmC はいずれの臓器/組織を構成する細胞においても核に強局在することが明らかとなった。

次に、③Cas9-FmC および最適化 Cas9-FmC(Cas9-ZmC)での *in vivo* ゲノム編集について記載する。上述の Cas9 発現マウスの筋肉に sgRNA を発現させることで *in vivo* ゲノム編集を実施し、その効率を確認した。筋肉への高い効率で導入可能なことが既に確認されているアデノ随伴ウイルス(adeno-associated virus、以後 AAV と記載)6 をウイルスベクターとして採用した。導入効率をモニターするために、AAV6 ウイルスベクターには sgRNA 発現ユニットに加えて EGFP 発現ユニットも付加した。標的の遺伝子は *Dmd* 遺伝子とし、第 23 エクソン上に点変異を導入するデザインと、第 23 エクソン自身を欠損させるデザインにした。点変異導入デザインのための AAV6 ベクターには、sgRNA 発現ユニットおよび EGFP 発現ユニットに加えて、HDR のためのドナー部位も付加した。エクソン欠損デザインでは、第 22 イントロンと第 23 イントロンに対する二つの sgRNA 発現ユニットを搭載した AAV6 ベクターとした。まず、野生型マウス(生後 2 日目および成獣)の腓腹筋にこれらの AAV6 ベクターを導入し、その活性を評価した結果、幼若マウスでも成獣マウスでも EGFP の蛍光が観察された。そこで、第 23 エクソンを欠損させるデザインとした AAV6 ベクターを成獣野生型マウ

ス(n=1)と上述の成獣 Cas9-FmC 全身発現マウス(n=4)と成獣 Cas9-ZmC 全身発現マウス(n=4)の腓腹筋へと投与した。その結果、全ての Cas9-FmC と Cas9-ZmC マウスで目的通りの欠失変異が誘導されていることが PCR 解析により判明した。また、点変異導入のための AAV6 ベクターを上述の成獣 Cas9-FmC 全身発現マウス(n=4)と成獣 Cas9-ZmC 全身発現マウス(n=5)の腓腹筋へと投与した。その結果、全ての Cas9-FmC と Cas9-ZmC マウスでなんらかのゲノム編集変異が誘導された。以上の結果から、Cas9-FmC および Cas9-ZmC は in vivo ゲノム編集を引き起こす能力があることが明らかとなった。

最後に、④オンターゲットにおける編集効率と安全性も確認できる評価系の開発について記載する。ゲノム編集では各体細胞において異なる変異がそれぞれ誘導されるために、PCR やサンガーシーケンスなどのバルク解析で生じた変異を網羅的に解析することができない。また2箇所切断で標的領域の切除を狙うゲノム編集では、それぞれの標的部位に indel 変異が導入される場合もあり、この変異が同一アレル上、すなわち cis に導入されたのかどうかは、一般的な short read の次世代シーケンスでは解析できない。そこで一分子ロングシーケンスが可能なナノポアシーケンサーを利用した系の確立を目指した。上述の通り、in vivo ゲノム編集では極めて多種多様な変異アレルが産出されると予想され、その陽性コントロールを用意することが困難であるために、およそ 1~8 程度のアレルが産出される受精卵ゲノム編集で誕生したファウンダーマウスを系確立のために用いた。ナノポアシーケンサーは従来の次世代シーケンスデバイスとは異なり、小型かつ安価で導入コストがかからないという利点があるが、リード時のエラー率が高いことが弱点である。そこで、機械学習によりエラー率を補正するソフトウェアを開発した。このソフトウェアを用いて、点変異誘導マウスにおけるゲノム編集標的部位を解析した結果、デザイン通りの点変異はもちろんのこと、意図しない数塩基対単位の挿入/欠損変異や意図しない 100 から数千塩基対以上の大幅な欠損を 1 塩基単位の正確性で検出できることが明らかとなった。また、2つのゲノム部位を同時に切断させることで目的の領域を欠失させるゲノム編集を受精卵で施したマウスを対象にした解析においても、意図した変異と意図しない変異の両方を 1 塩基単位の正確性で検出できることが明らかとなった。更に、複数のマウスのサンプルを混ぜたマルチプレックス解析を実施したところ、1回のランで 281 アレルを検出することに成功した。以上の結果から、我々が新たに開発したソフトウェアを用いることで、ゲノム編集標的部位に生じた変異を広範囲かつ網羅的に正確に評価できることが明らかとなった。

Clinical trials for gene therapy using in vivo genome editing with CRISPR-Cas9 have begun, and its effectiveness and safety have attracted considerable attention. Although genomic DNA is located in the nucleus, Cas9 protein is derived from prokaryotes and does not have nuclear localization signal. In order to effectively perform in vivo genome editing with a small amount of Cas9 protein or RNA or DNA encoding Cas9, it is important to develop the highly nuclear-localized Cas9. A threat to the safety of genome editing is large-scale loss of genomic regions at on-target sites, as well as off-target mutations.

We have discovered a new factor, FmC, that confers a very strong nuclear localization to Cas9. This FmC-added Cas9, Cas9-FmC, strongly localized to the nucleus in cultured cells and fertilized eggs, and also increased its genome editing efficiency. If Cas9-FmC can increase the efficiency of in vivo genome editing, the therapeutic (editing) effect can be maximized with a small dose. Therefore, we carried out four tasks: (1) optimization of Cas9-FmC, (2) evaluation of subcellular localization of Cas9-FmC and optimized Cas9-FmC in each adult tissue, (3) in vivo genome editing with Cas9-FmC and optimized Cas9-FmC, and (4) development of an evaluation system that can also confirm editing efficiency and safety on-target.

(1) Since the size of FmC is 104 amino acids, its coding sequence is 312 base pairs, which is relatively large for a gene sequence of nuclear localization signal. We constructed Cas9-XmC, Cas9-YmC, and Cas9-ZmC, in which only a part in FmC was added to Cas9. These subcellular localizations were evaluated in HEK293T cells, and the results showed that Cas9-ZmC strongly localized to the nucleus. From these results, we determined the smallest Cdt1 region where Cas9 can be localized to the nucleus.

(2) We then generated mice that strongly expressed Cas9-FmC and Cas9-ZmC throughout the body, respectively. These homozygous knocked-in mutants showed no abnormal phenotype, indicating that strong expression of Cas9-FmC or Cas9-ZmC does not affect the growth or fertility of the mice. We collected and sectioned 18 tissues from Cas9-FmC and Cas9-ZmC ubiquitous expressing mice, and confirmed the subcellular localization of Cas9-FmC and Cas9-ZmC in each tissue by immunostaining with anti-Cas9 antibody. The results showed that Cas9-FmC and Cas9-ZmC were present in the nucleus in all tissues. These results indicate that Cas9-FmC and Cas9-ZmC are strongly localized to the nucleus in the cells of both organs/tissues.

(3) The results of in vivo (skeletal muscle) genome editing using AAV6 vector expressing sgRNA and the Cas9-expressing mice described above showed that genetic mutations were induced in the target sites in all mice treated with AAV6 vector.

(4) Since genome editing induces different mutations in each somatic cell, bulk analysis such as PCR and Sanger sequencing cannot be used to comprehensively analyze the mutations occurring. Therefore, we established a system using a nanopore sequencer capable of single molecule long sequencing. Our analysis system was able to comprehensively detect not only point mutations as designed, but also unintentional small indel mutation and unintentional deletions of 100 to several thousand base pairs with single-base accuracy.