Journal of Biotechnology 221 (2016) 49–54

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Short communication

Comparison of surrogate reporter systems for enrichment of cells with mutations induced by genome editors

Zuyong He[a,1], Xuan Shi[a,1], Meirui Liu[a], Guangjie Sun[a], Chris Proudfoot[b], C. Bruce A. Whitelaw[b], Simon G. Lillico[b,**], Yaosheng Chen[a,*]

[a] State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-sen University, Guangzhou 510006, PR China
[b] The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush Campus, Edinburgh EH25 9RG, UK

ARTICLE INFO

Article history:
Received 8 August 2015
Received in revised form 11 January 2016
Accepted 12 January 2016
Available online 14 January 2016

Keywords:
ZFN
CRISPR/Cas9
Surrogate reporter
SIGLEC-1

ABSTRACT

Genome editors are powerful tools that allow modification of the nuclear DNA in eukaryotic cells both *in vitro* and *in vivo*. *In vitro* modified cells are often phenotypically indistinguishable from unmodified cells, hampering their isolation for analysis. Episomal reporters encoding fluorescent proteins can be used for enrichment of modified cells by flow cytometry. Here we compare two surrogate reporters, RGS and SSA, for the enrichment of porcine embryonic fibroblasts containing mutations induced by ZFNs or CRISPR/Cas9. Both systems were effective for enrichment of edited porcine cells with the RGS reporter proving more effective than the SSA reporter. We noted a higher-fold enrichment when editing events were induced by Cas9 compared to those induced by ZFNs, allowing selection at frequencies as high as 70%.

Programmable nuclease-based technologies such as zinc-finger nucleases (ZFNs) (Urnov et al., 2010) and the bacterial clustered regularly interspaced short palindromic repeats (CRISPR)-associated (Cas) system (Cong et al., 2013) are powerful and versatile tools for genome editing. They can both induce site-specific DNA double-strand breaks (DSBs) in the genome of eukaryotic cells by utilising either the non-specific nuclease *Fok*I or Cas9 respectively. While ZFNs direct localisation of the *Fok*I dimer to the target sequence by specific DNA binding by zinc-finger domains, the CRISPR/Cas9 system relies on Watson/Crick base pairing between a small guide RNA and a 20-nucleotide target site flanked by a protospacer-adjacent motif (PAM) to direct the Cas9 nuclease (Sternberg et al., 2014). The error-prone NHEJ pathway repairs many nuclease-mediated DSBs, resulting in the inclusion of small insertions or deletions (indels) at the repair site. These may result in frame shift if the target site is situated in protein coding sequence, and thereby functional gene disruption. Although application of designer nucleases have enabled targeted genome editing in numerous eukaryotic species and cell types (Hai et al., 2014; He et al., 2014; Ma et al., 2014), there is often difficulty in the isolation of mutant cells post-modification due to grossly similar phenotypes between modified and wild-type counterparts, hampering their application to a certain extent. Recently, two different episomal surrogate reporter systems have been successfully used to enrich cells containing mutations induced by programmable nucleases. One surrogate reporter, called RGS (Kim et al., 2011), encodes a monomeric mRFP-eGFP fusion protein with a cloning site for introduction of the proposed target site located between the two reading frames. In its native conformation the mRFP is expressed from the constitutive promoter while the eGFP sits out of frame downstream (Fig. 1B). Consequently, upon transfection cells express mRFP but not eGFP. DSB creation by engineered nucleases at the introduced target site between the two FPs results in NHEJ and frame shift mutations within the construct, resulting in a proportion of cells expressing both mRFP and eGFP (Fig. 1C) which can be detected and/or selected by FACS. The second reporter system also utilises a fluorescent switch. Named SSA (for single-strand annealing) this construct contains a non-functional GFP with a duplicated central domain containing the nuclease target site (Doyon et al., 2013). In this instance generation of a DSB can restore GFP function when repair is via the SSA pathway (Fig. 1E).

To our knowledge, there is no report of a direct comparison between these two reporters for the enrichment of genome-modified primary cells. Competition between different DNA repair pathways has been documented, and which pathway predominates may be influenced by cell type and the nature of the broken ends

* Corresponding author at: The No.3 Road of Higher Education Mega Center North, School of Life Sciences, PR China.
** Corresponding author.
E-mail addresses: simon.lillico@roslin.ed.ac.uk (S.G. Lillico), chyaosh@mail.sysu.edu.cn (Y. Chen).
[1] These authors contribute equally to this work.

http://dx.doi.org/10.1016/j.jbiotec.2016.01.009
0168-1656/

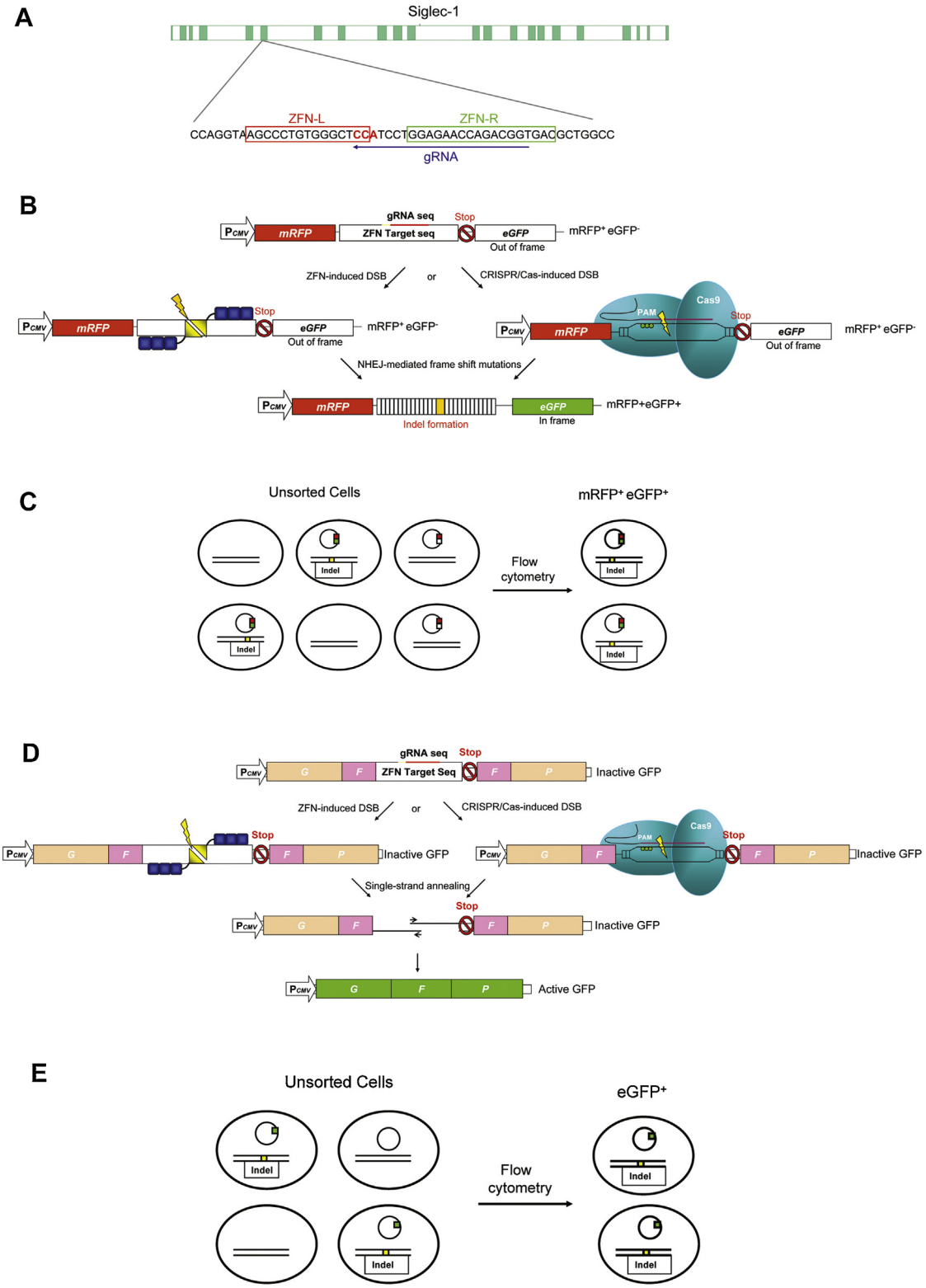

Fig. 1. ZFN and gRNA design, and the principle of the use of surrogate reporters for enrichment of nuclease-modified cells. (A) The schematic diagram of target sites of ZFN and gRNA on the exon 6 of porcine *SIGLEC-1* gene. Colored arrow lines indicate the sequence used for the guide segment of gRNA. The NGG nucleotide protospacer adjacent motif (PAM) sequences in red. Colour boxed sequences denote the DNA binding regions of the ZFN proteins. (B) The working mechanism of the RGS reporter. mRFP is constitutively expressed by the CMV promoter, whereas eGFP is not expressed without ZFN or CRISPR/Cas9 activity because the eGFP sequence is out of frame and there is a stop codon before eGFP. If a double-strand break (DSB) is introduced into the target sequence by the programmable nucleases, the break is repaired by error-prone nonhomologous end joining (NHEJ), which often results in indels. This indel formation can cause frame shifts, making either of the eGFP genes in frame and expressed. (C) A schematic illustrating the enrichment of nuclease-induced mutations in mRFP$^+$eGFP$^+$ cell population sorted by flow cytometry. Reporter plasmids and chromosomal target loci are illustrated. Mutations are shown as yellow spots. (D) The working mechanism of the SSA reporter. The SSA reporter consists of a sequence encoding GFP, and ZFN or CRISPR/Cas9 target site which disrupts the expression of GFP. The unique 5′ GFP sequence, middle repeated GFP sequence, and unique 3′ GFP sequence are designated as G, F, and P respectively. Following cleavage with the programmable nucleases and SSA-mediated repair, the functional GFP open reading frame is reconstituted by loss of sequences between the two identical 5′ and 3′ F sequences. (E) Schematic illustrates enrichment of nuclease-induced mutations in eGFP$^+$ cells sorted by flow cytometry. Reporter plasmids and chromosomal target loci are illustrated. Mutations are shown as yellow spots. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)