F-2084

Use of Surrogate Reporter Vectors to Enrich for CRISPR & TALEN Modified Cells

Natasha Roark, Jian-Ping Yang, Veronica Blackston, Jonathon Chesnut, Life Technologies, 5781 Van Allen Way, Carlsbad, CA, USA, 92009

ABSTRACT

The ability to effectively model human diseases is often hindered by the difficulty in altering nucleotide sequences and influencing gene expression in a targeted fashion. Engineered transcriptional activator-like effector nucleases (TALENs) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs) have emerged as attractive tools for inducing genetic modifications as both can be designed de novo to target specific DNA sequences and modify DNA at that locus. Given their modular structure and specificity, TALENs and CRISPRs allow for the control of cellular gene expression in a precise, predictable and robust manner. An impediment to the many applications involving genome editing tools such as these is the ability to select or enrich for genetically modified cells, as they are generally phenotypically indistinguishable from their wild types. This is further complicated by the fact that often only a small population of cells will contain the desired mutation, necessitating the screening of a multitude of clones. This difficulty is compounded in hard to culture cell types such as induced pluripotent stem cells (iPSCs) and embryonic stem cells (ESCs).

For this purpose we constructed surrogate reporter vectors containing the target sequence and encoding an out of frame fluorescent protein and CD4 membrane protein. A TALEN or CRISPR induced double-stranded break will be repaired by non-homologous end joining (NHEJ), resulting in a frameshift mutation that will render both the fluorescent protein and CD4 in-frame in a fraction of the modified cells. This enables us to enrich for genetically modified cells using flow cytometry and antibody-coated magnetic beads, as only cells containing the TALEN or CRISPR modified reporter plasmids will express CD4 or fluorescent proteins. Cells were co-transfected with the aforementioned reporter vector and corresponding TALEN pair or CRISPR guide RNA with Cas9 and were subsequently enriched for CD4 and fluorescent protein expression using magnetic beads and flow cytometry, respectively. An endonuclease-based mismatch detection assay was then used to quantify the percentage of gene modification in pre and post-enrichment samples. Preliminary findings show that antibody coated beads and fluorescent based cell sorting yield up to 25-fold and 4-fold enrichment, respectively. The fold enrichment is inversely proportional to the number of clones one must then screen to obtain the desired mutation, i.e. with four-fold enrichment the number of clones screened can be reduced by one-fourth. This project demonstrates that utilizing a surrogate reporter vector is an effective and reliable tool enabling one to selectively enrich genetically modified cells, obviating the need to isolate a large number of clones. The results from this project have a broad range of practical applications for modeling human disease including the development of knock-in, knockout and transgenic organisms, particularly for difficult to culture cell-types such as iPSCs and ESCs.

INTRODUCTION

Efficient and accurate cell and genomic engineering is imperative to the advancement of modeling and eventually curing disease. TALENs and CRISPRs are extremely useful tools to this pursuit as they can be designed de novo to target and alter specific nucleotide sequences in a precise and predicable manner. Advancements in genome editing are often hindered by the difficulty in obtaining a homogenous population of cells with the desired mutation. Contributing factors to this are difficult to transfect cell types as well as loci that are less amenable to gene modification. Surrogate reporter vectors provide a means to mitigate this difficulty by allowing investigators to enrich for cells containing the desired mutation by antibody coated magnetic beads or fluorescence activated cell sorting (FACS).

MATERIALS AND METHODS

The surrogate reporter vectors were constructed using a variety of molecular cloning techniques, then were purified and confirmed by sequence analysis. TALEN and CRISPR vectors were similarly constructed and subsequently confirmed by restriction digests and sequence analysis, respectively. TALEN transfected cells were transfected with forward and reverse TALENs and surrogate reporter vectors containing their target binding sequence. CRISPR transfected cells were transfected with CRISPR guide RNA, the CAS-9 protein and surrogate reporter vectors containing the CRISPR target sequence of interest. Cells were also transfected with the surrogate reporter vector and irrelevant TALENs/CRISPR guide RNA as well as positive fluorescent controls. Fluorescent protein expression of the reporter vector was observed and recorded by fluorescent microscopy and cells were harvested 72 hours post transfection.

Cells were harvested and a sample was obtained for later analysis to determine the proportion of gene modification prior to enrichment. Cells were then enriched by either fluorescence activated cell sorting or via antibody coated magnetic beads (Dynabeads®). Fluorescent protein expression of cells that were enriched with Dynabeads® magnetic beads was additionally analyzed by flow cytometry. Additionally, CD4 expression was confirmed using immunohistochemistry. After enrichment using FACS or Dynabeads® magnetic beads, the genomic DNA of the pre-enrichment, post-enrichment and irrelevant TALN transfected cells was extracted and the DNA was amplified using primers that flank the locus of the CRISPR or TALN-pair binding sites. Once amplified the PCR amplicons were denatured and re-annealed through thermocycling. This step yields a predictable ratio of homoduplex mutated, homoduplex wild type and heteroduplex DNA. This DNA was then digested with an enzyme that recognizes the NHEJ-induced indels resulting from the double stranded breaks caused by CRISPR and TALEN cleavage. The cleavage products were then run on a gel and analyzed using gel analysis software. Band intensity was quantified the cleavage efficiency quantified. The accuracy of the locus specific cleavage (LSC) assay results was additionally validated by cloning the PCR amplified fragments into pCR™2.1-TOPO® vector. The plasmid was then transformed, and individual colonies purified and sequenced to verify the presence of the mutation in the correct site. The proportion of sequence-confirmed modified cells was compared with that of the LSC assay.

RESULTS

Figure 1. TALEN Pair

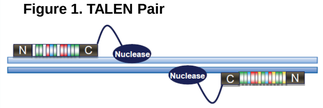

Figure 1 is a schematic representation of a TALEN pair (forward and reverse) bound to a target DNA sequence and fused to a Fok1 nuclease domain. TALENs are comprised of tandem, polymorphic amino acid repeats with each repeat recognizing a specific contiguous nucleotide on the target sequence. Binding specificity is determined by the repeat variable di-residues (RVDs) at positions 12 and 13.

Figure 2. CRISPR/Cas9

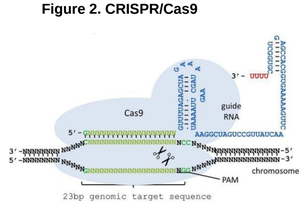

Figure 2 illustrates the CRISPR/CAS system. A tracrRNA (trans-activating crRNA) and CRISPR fusion comprises CRISPR guide RNA (gRNA). The guide RNA confers sequence specificity, while Cas9 has the nuclease activity.

Figure 3. Surrogate Reporter Vectors

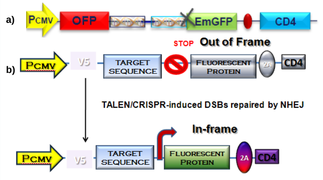

Figure 3 diagrams the basic components of a surrogate reporter vectors. Each reporter contains the TALEN / CRISPR target sequence and an out of frame fluorescent protein, 2A peptide and CD4 membrane protein. In CRISPR/TALEN modified cells containing the reporter, both the fluorescent and CD4 protein coding sequences will be rendered in-frame and thus expressed. The reporter depicted in figure a codes for an additional in-frame fluorescent protein.

Figure 4. Enrichment of Modified Cells

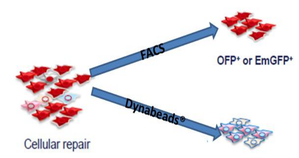

Figure 4 diagrams the method by which fluorescent and CD4-expressing cells are enriched via fluorescent activated cell sorting and antibody coated magnetic beads, respectively.

Figure 5. Locus Specific Cleavage Assay

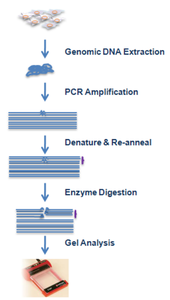

Figure 5 outlines the steps of the LSC Assay, the method by which endogenous gene modification by TALEN & CRISPR modified cells is quantified.

Figure 6. Fluorescent Image of CRISPR/CAS9/Surrogate reporter Transfected cells

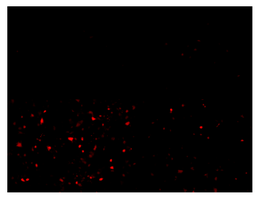

Figure 6 shows OFP expression of HEK293-ft cells co-transfected with Cas9 ,CRISPR gRNA targeting the AAVS1 locus and the surrogate reporter vector. Clockwise from top-left: Reporter/CAS9/Irrelevant gRNA, Reporter/CAS9/AAVS1 gRNA 24, 48, and 72 hours post transfection.

Figure 7. LSC Assay Results

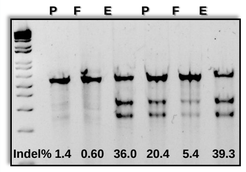

Figure 7 depicts the LSC Assay results for the CRISPR/CAS9 transfected cells pre (P) and post (E) magnetic bead enrichment as well as the flow though (F).

Figure 8. LSC Assay Results

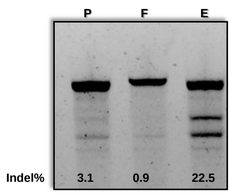

Figure 8 depicts the LSC Assay results for the CRISPR/CAS9 transfected cells pre (P) and post enrichment (E) as well as the flow though (F).

Figure 9. Flow Cytometry Analysis of Surrogate Reporter OFP Expression

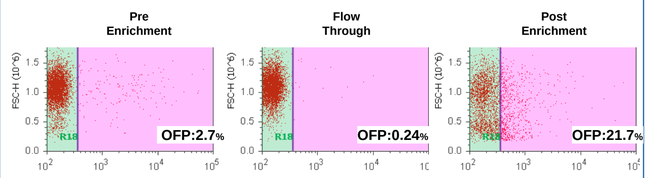

Figure 9 shows analysis of OFP expression of Dynabeads® magnetic beads enriched cells using flow cytometry.

Figure 10. LSC Assay of FACS Enriched CRISPR Modified Cells

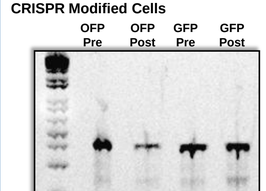

Figure 10 shows LSC assay results for CRISPR modified cells enriched using GFP or OFP expressing reporter vectors. 4-fold and 3-fold enrichment is observed.

Figure 11. LSC Assay of FACS Enriched TALEN Modified Cells

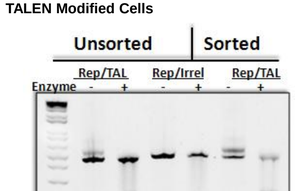

Figure 11shows LSC assay results for TALEN modified cells enriched using reporter vector GFP expression. 4-fold enrichment is observed.

Figure 12. Fluorescent Image of TALEN/ Surrogate Reporter Transfected Cells

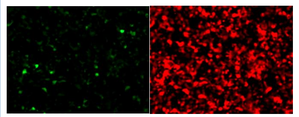

Figure 12 shows GFP and OFP Expression of the surrogate reporter of TALEN transfected cells. OFP represents the transfection efficiency (in-frame) while GFP represents modification via TALENs at the IL2 locus.

Figure 13. LSC Assay Enrichment Results

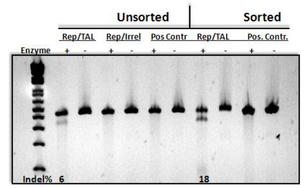

Figure 13 depicts the enrichment results quantified by the LSC assay. Enrichment using Dynabeads gives a 3-fold enrichment of cells modified at the IL2 locus.

Figure 14. Sequencing of Pre-sorted PCR Amplicons

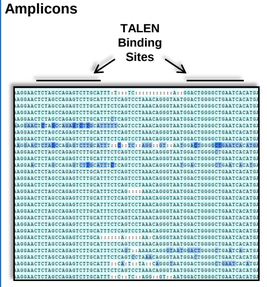

Figure 14 shows sequencing data from TOPO®-cloned PCR amplicons of TALEN modified genomic DNA pre-enrichment. 7/29 or 24% have deletions.

Figure 15. Sequencing of Sorted PCR Amplicons

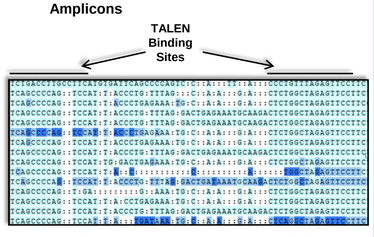

Figure 15 shows sequencing data from TOPO®-cloned PCR amplicon s of TALEN modified genomic DNA post –enrichment. 10/15 or 67% have deletions. A 3-fold enrichment was attained.

Figure 16. Immunohistochemistry of Reporter Vector Expression

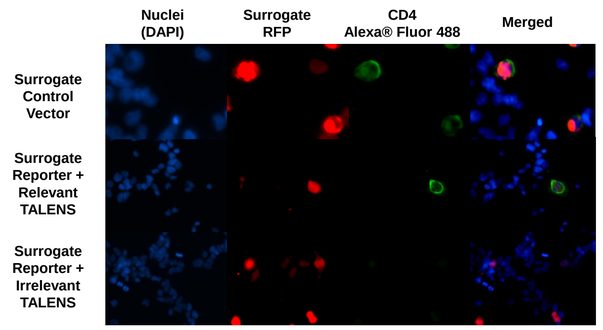

Figure 16 shows immunohistochemistry staining of cells transfected with the surrogate reporter vector, containing an in-frame RFP and out of frame GFP, and with relevant and irrelevant TALENs. A positive control vector , containing both RFP and GFP in-frame, was also stained.

CONCLUSIONS

These data suggest that surrogate reporter vectors are an efficient way to select for cells that have been genetically modified at a specific locus using either CRISPRs or TALENs. Twenty-five-fold and four-fold enrichment has been achieved using Dynabeads® magnetic beads and fluorescence activated cell sorting, respectively. This corresponds to a 4-fold and 25-fold decrease in the number of clones that must be carried out and sequenced to obtain the desired mutation.

The proportion of gene modification has been determined using the locus specific cleavage assay, flow cytometry and sequencing. These data suggest that each method of analysis are consistent and provide an accurate estimation of the proportion of genetically modified cells in a given population. Additionally, the expression of proteins associated with TALEN and CRISPR induced mutations have been confirmed using immunohistochemistry and fluorescent microscopy.

Surrogate reporter vectors are valuable tools for the enrichment and isolation of genetically modified cells and will be extremely useful in the development of knock-outs, knock-ins and transgenic organisms, particularly in difficult to transfect cell lines and loci complex to target.

REFERENCES

1. A. Bogdanove, D. Voytas, TAL Effectors: Customizable Proteins for DNA Targeting, *Science* **333**, 1843-1846 (2011).
2. Prashant Mali *et al,* RNA-Guided Human Genome Engineering via Cas9, *Science* ***339***, *823-826 (2013)*.

ACKNOWLEDGEMENTS

California Institute for Regenerative Medicine (CIRM), Grant# TB1-01186
California State University, San Marcos

TRADEMARKS/LICENSING