Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA

Xiquan Liang, Jason Potter*, Shantanu Kumar, Namritha Ravinder, Jonathan D. Chesnut

Thermo Fisher Scientific, 5781 Van Allen Way, Carlsbad, CA 92008, USA

A R T I C L E   I N F O

Article history:
Received 15 June 2016
Received in revised form 9 November 2016
Accepted 10 November 2016
Available online 11 November 2016

Keywords:
CRISPR
Cas9
gRNA
Genome editing
Knock-in
Homologous recombination

A B S T R A C T

While CRISPR-based gene knock out in mammalian cells has proven to be very efficient, precise insertion of genetic elements via the cellular homology directed repair (HDR) pathway remains a rate-limiting step to seamless genome editing. Under the conditions described here, we achieved up to 56% targeted integration efficiency with up to a six-nucleotide insertion in HEK293 cells. In induced pluripotent stem cells (iPSCs), we achieved precise genome editing rates of up to 45% by co-delivering the Cas9 RNP and donor DNA. In addition, the use of a short double stranded DNA oligonucleotide with 3' overhangs allowed integration of a longer FLAG epitope tag along with a restriction site at rates of up to 50%.

We propose a model that favors the design of donor DNAs with the change as close to the cleavage site as possible. For small changes such as SNPs or short insertions, asymmetric single stranded donor molecules with 30 base homology arms 3' to the insertion/repair cassette and greater than 40 bases of homology on the 5' end seems to be favored. For larger insertions such as an epitope tag, a dsDNA donor with protruding 3' homology arms of 30 bases is favored. In both cases, protecting the ends of the donor DNA with phosphorothioate modifications improves the editing efficiency.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

The recent advances in CRISPR-mediated genome engineering have enabled researchers to efficiently introduce double-strand breaks (DSBs) in genomic DNA using Cas9 and an appropriate gRNA (Cho et al., 2013; Jiang et al., 2013; Liang et al., 2015; Mali et al., 2013; Wang et al., 2013). Most DSBs are then repaired by either the non-homologous end joining (NHEJ) pathway or the homology-directed repair (HDR) pathway.

In mammalian cells, the NHEJ pathway is predominant and error-prone, which results in disruptive insertions or deletions (indels) at targeted loci. This allows for the efficient creation of gene knockouts. Alternative repair pathways include the use of sister chromatids or an exogenous donor DNA template via components of the HDR pathway. Given that the efficiency of repairing Cas9 induced breaks by HDR is relatively low (Ran et al., 2013; Rong et al., 2014), several methods have been explored to improve the utility of this approach for genome editing. The co-delivery of gRNA and an ssDNA donor into Cas9-expressing human pluripotent stem cells (hPSCs) generated homozygous knock-in clones at a rate of up to 10% (González et al., 2014). The synchronization of cells at M phase with nocodazole prior to nucleofection of Cas9 ribonucleoproteins (RNPs) and a ssDNA donor resulted in up to 38% HDR in HEK293T cells and 1.6% in hESCs, compared to the controls of 26% and ~0% in un-synchronized HEK293T cells and hESCs respectively (Lin et al., 2014). The delivery of Cas9 RNPs and a ssDNA donor into primary T cells via electroporation generated genomic knock-in modifications with up to 20% efficiency (Schumann et al., 2015).

Recently, several attempts have been made to improve HDR efficiency by biochemically altering the HDR or NHEJ pathways. For example, the treatment of various mammalian cells with Scr7, a DNA ligase IV inhibitor, resulted in an up to 19-fold increase in HDR efficiency (Maruyama et al., 2015). The simultaneous suppression of both KU70 and DNA ligase IV with siRNAs improved the efficiency of HDR 4–5 fold in human and mouse cell lines (Chu et al., 2015). Likewise, the HDR enhancer RS-1 increased the knock-in efficiency in rabbit embryos both in vitro and in vivo by 2–5 fold (Song et al., 2016). Most recently, the use of asymmetric ssDNA donors of optimal length increased the rate of HDR in human cells up to 60% for a single nucleotide substitution (Richardson et al., 2016).

Abbreviations: CRISPR, clustered regularly interspaced short palindromic repeats; CAS9, CRISPR associated protein; gRNA, guide RNA; FAM, protospacers adja-cent motif; RNP, ribonucleoprotein; HDR, homology directed repair; hESCs, human embryonic stem cells; NHEJ, non-homologous end joining; SNP, single-nucleotide polymorphism.

* Corresponding author.

E-mail address: jason.potter@thermofisher.com (J. Potter).

http://dx.doi.org/10.1016/j.jbiotec.2016.11.011
0168-1656/© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
In this study, we examined alternative approaches to improve HDR without impairment of other cellular DNA repair machinery. By optimizing the design and delivery of gRNA, Cas9 nuclelease and donor DNA, we achieved up to 56% precise genome editing efficiency in HEK293 and 45% in iPSC cell lines after normalizing for cutting efficiency of the gRNAs. When optimizing for overall integration efficiency it was found that the proximity of the DSB to the target modification, phosphorothioate protection of the donor, asymmetric design of the ssDNA donor, and electroporation conditions are all critical factors to be considered. Furthermore, the alternate design of a short dsDNA oligonucleotide with 3’ overhangs improved the insertion efficiency of epitope tags into the genome.

2. Materials and methods

2.1. Materials

GeneArt™ Platinum™ Cas9 Nuclease, GeneArt™ CRISPR gRNA Design Tool, GeneArt™ Precision gRNA Synthesis Kit, GripTite™ HEK293 cells, DMEM medium, Fetal Bovine Serum (FBS), TrypLE™ Express Enzyme, Jump-In™ GripTite™ HEK293 Kit, Gibco® Human Episomal iPSC Line, MEF feeder cells, RPMI 1640 medium, DMEM/F-12, Knockout™ Serum Replacement, Non-Essential Amino Acid solution, basic fibroblast growth factor, Collagenase IV, Genelute™ Lentivirus expressing Cas9 nuclease and basicidin marker, 2% E-Gel® EX Agarose Gels, ViraPower® kit, TranscriptAid™ T7 High Yield Transcription Kit, MEGAclear® Transcription Clean-Up Kit, Zero Blunt® TOPO® PCR Cloning Kit, PureLink® Pro Quick96 Plasmid Purification Kit, Qubit® RNA BR Assay Kit, Neon® Transfection System10 μL Kit, and Phusion High-Fidelity PCR Master Mix were from Thermo Fisher Scientific. Monoclonal Cas9 antibody was purchased from Diagenode. ROCK inhibitor Y-27632 was purchased from EMD Millipore. The DNA oligonucleotides used for gRNA synthesis or donors were from Thermo Fisher Scientific (Supplementary Table S1).

2.2. Synthesis of gRNA

DNA oligonucleotides used for gRNA synthesis were designed with the GeneArt® CRISPR gRNA Design Tool. The gRNAs were then synthesized using the GeneArt™ Precision gRNA Synthesis Kit. The concentration of gRNA was determined by the Qubit® RNA BR Assay Kit.

2.3. Genomic cleavage and detection (GCD) assay

The genomic cleavage efficiency was measured by the GeneArt® Genomic Cleavage Detection kit according to manufacturer’s instructions. The primer sequences for PCR amplification of each genomic locus are described in Supplementary Table S1. Cells were analyzed at 48 to 72 hours post transfection. The cleavage efficiencies were calculated based on the relative agarose gel band intensity, which was quantified using an Alphalmager® gel documentation system running AlphaView®, Version 3.4.0.0. ProteinSimple (San Jose, CA, USA).

2.4. Cell culture

GripTite™ HEK293 cells were maintained in DMEM medium supplemented with 10% FBS. Feeder-dependent human epithelial iPSC were cultured on mitotically inactivated MEF feeder cells in human ESC (hESC) media containing 20% Knockout™ Serum Replacement, 10 μM Non-Essential Amino Acid solution, 55 μM 2-Mercaptoethanol, and 4 ng/mL basic fibroblast growth factor (bFGF) in DMEM/F-12. All cultures were maintained in a 5% CO₂, 37°C humidified incubator. iPSC cultures were maintained with daily media changes and were passaged regularly using Collagenase IV.

2.5. Generation of stable cell lines

The Jump-In™ system was used to prepare GripTite™ HEK293 stable cell expressing EmGFP (Thermo Fisher Scientific). The gene sequence of EmGFP is described in Supplementary Table S2. To create a disrupted EmGFP mutant stable cell line, a gRNA targeting the 5’-ctgctgaccacctcactccag-3’ sequence in EmGFP gene was synthesized. The resulting gRNA (300 ng) was incubated with 1.5 μg of GeneArt® Platinum™ Cas9 nuclease and was used to transfect wild type EmGFP cells via electroporation. The single cell clonal isolation was carried out by limiting dilution. The EmGFP loci from non-glowing cells were amplified by PCR using a forward primer 5’-atggtgagcggaggagctg-3’ and a reverse primer 5’-gtctctcttgtagctgtccc-3’ and the resulting PCR products were subjected to TOPO cloning and sequencing. From these clones, a disrupted EmGFP stable cell line containing a deletion of 5’-CACCTT-3’ was identified. In the homologous recombination assay, the gain of EmGFP function was determined by flow cytometric analysis.

To generate a HEK293FT stable cell line expressing eBFP, an EBFP ORF was synthesized by GeneArt® custom DNA synthesis (Thermo Fisher Scientific) and then cloned into pDONAR221 vector. Using Gateway® recombination technology (Thermo Fisher Scientific), the EBFP ORF was transferred to pLenti6.2-DEST Gateway Vector and then verified by sequencing. Lentiviruses were generated using ViraPower® kit as described in the manual. To generate a stable cell line, HEK293FT cells were transduced with 0.1 MOI of Lentivirus expressing EBFP. Three days post transduction cells were selected on 5 μg/mL blasticidine antibiotics for 2 weeks. Cells expressing EBFP were then collected and diluted to 0.8 cells/mL in complete medium and plated into 96 well plates for single cell cloning. After 2 weeks clones were isolated and verified for EBFP expression by flow cytometry. In a reporter assay using a stable eBFP-expressing HEK293 cell line, the substitution of C to T in eBFP gene converts His66 to Tyr66, generating a GFP variant (Supplementary Table S2).

2.6. Homologous recombination assays

To create homologous recombination (HR) assays, a series of gRNAs flanking the insertion site within the EmGFP gene were designed and synthesized (Supplementary Table S1). Each individual gRNA was combined with GeneArt™ Platinum™ Cas9 Nuclease to form the Cas9 protein/gRNA ribonucleoprotein complexes (Cas9 RNPs). The Cas9 RNPs were then used to transfect cells via Neon® electroporation. The genomic cleavage efficiency was then evaluated using the GeneArt® Genomic Cleavage Detection kit at 48 hours post transfection. The gRNAs with highest editing efficiencies and also in close proximity to the insertion site were selected for the subsequent HR assays. For donor design of a single-stranded oligonucleotide, typically the mutation site was positioned at the center flanked by 30 to 50 nucleotides on each side. For asymmetric donor design, 30 nucleotides of homology were placed on either the left or right arm relative to the desired change and 50 or 67 nucleotides on the opposite arm. Donor oligonucleotides with homology to the PAM or non-PAM strands were tested. For a dsDNA donor oligonucleotide with single-stranded overhangs, oligonucleotides were designed with the insertion element, such as a FLAG tag, at the 3’ or 5’ ends and the respective homology arms at the opposite ends. Once annealed, the tag is within the dsDNA region and the homology arms are single stranded. By annealing two single-stranded oligonucleotides at 95°C for 3 minutes, dsDNA donor molecules with either a 5’ protrusion or a 3’ protrusion were generated. To measure homologous recombination efficiency, the donor DNA was either co-transfected with Cas9 RNPs or delivered
sequentially into cells via electroporation. At 48 h post transfection, the gain of EmGFP function in reporter cell lines was determined by flow cytometric analysis with an Attune<sup>®</sup> NxT Acoustic Focusing Cytometer (Thermo Fisher Scientific). Alternatively, the genomic loci were PCR-amplified using the corresponding primers and then subjected to GeneArt<sup>®</sup> Genomic Cleavage Detection assay or restriction enzyme digestion. The resulting PCR products were also subjected to TOPCloning. Typically, 96 colonies were randomly picked for sequencing. The sequencing data were analyzed using Vector NTI Advance<sup>®</sup> 11.5 software (Thermo Fisher Scientific).

2.7. Electroporation

Typically, 1 × 10<sup>5</sup> GripTite<sup>®</sup> HEK293 cells or iPSCs were used per electroporation using Neon<sup>®</sup> Transfection System 10 μL Kit (Thermo Fisher Scientific). To optimize the electroporation conditions, the preprogrammed Neon<sup>®</sup> 24-well optimization protocol was tested according to the manufacturer’s instructions. To make up a master mix of 24 reactions, 8 μL of 3 mg/ml GeneArt<sup>®</sup> Platinum<sup>®</sup> Cas9 Nuclease was added to 240 μL of Resuspension Buffer R provided in the kit, followed by addition of 4.8 μg of gRNA (up to 24 μl). Upon mixing, the sample was incubated at room temperature for 10 min to form Cas9 RNP complexes. Meanwhile, 2.4 × 10<sup>5</sup> cells were transferred to a sterile 1.5 mL tube and centrifuged at 1000 × g for 5 min. The supernatant was carefully aspirated and the cell pellet was washed once with 1 mL of DPBS without Ca<sup>2+</sup> and Mg<sup>2+</sup>. The sample was again centrifuged and the supernatant was carefully aspirated. Resuspension Buffer R containing the Cas9 RNPs was then used to resuspend the cell pellets. A 10 μL cell suspension was used for each of the preprogrammed Neon<sup>®</sup> 24-well optimization protocols. The electroporated cells were transferred to 24 or 48-well plates containing 0.5 mL of the corresponding growth medium and then incubated for 48 h in a 5% CO<sub>2</sub> incubator. The cells were washed with DPBS and then lysed in lysis buffer, followed by the genomic cleavage and detection assay as described above. Upon optimization of electroporation conditions, higher doses of Cas9 protein (1.5–2 μg) and gRNA (300–500 ng) were used to improve the genome editing efficiency.

For each homologous recombination assay: 1.5 μg of Cas9 protein and 360 ng of gRNA in a volume less than 1 μL was added to Resuspension Buffer R to a final volume of 7 μL. The gRNA could be diluted in Buffer R if the concentration was too high. Upon mixing, the sample was incubated at room temperature for 5–10 min to form Cas9 RNPs. Meanwhile, GripTite<sup>®</sup> HEK293 cells expressing either eBFP or disrupted EmGFP were detached from culture flask with TrypLE<sup>®</sup> Express Enzyme and then counted. Aliquots of 1 × 10<sup>6</sup> cells were washed once with DPBS without Ca<sup>2+</sup> and Mg<sup>2+</sup> and the cell pellets were resuspended in 50 μL of Resuspension Buffer R. A 5 μL aliquot of cell suspension was mixed with the 7 μL of Cas9 RNPs. For sequential delivery of Cas9 RNPs and DNA donor, 10 μL of cell suspension containing Cas9 RNP was electroporated at 1150 V, with a 20 ms pulse width, for 2 pulses. The electroporated cells were transferred to 300 μL of Resuspension Buffer R or DPBS. Upon centrifugation at 2000 × g for 5 min, the supernatant was carefully aspirated and the cell pellet was resuspended in Buffer R to a final volume of 11 μL, followed by the addition of 1 μL of 10 pmol/μL (0.3 μg/μL) ssDNA donor. Alternatively, 1 μL of 10 pmol/μL short dsDNA donor with and without single-stranded overhangs was added. An aliquot of 10 μL cell suspension containing donor DNA was used for electroporation using the same instrument settings. Upon electroporation, the cells were transferred to a 48-well plate containing 0.5 mL culture media. For sequential delivery, the viability of HEK293 cells was around 50%. For co-transfection of Cas9 RNPs with donor DNA, 0.5 μL of 20 pmol/μL (0.6 μg/μL) ssDNA donor was directly added to the 12 μL of cell suspension containing Cas9 RNPs. Alternatively, 0.5 μL of 20 pmol/μL short ssDNA donor with and without single-stranded overhangs was added. An aliquot of 10 μL cell suspension containing Cas9 RNPs and donor DNA was used for electroporation. Samples lacking one component (either gRNA or donor DNA) served as controls. In addition to ssDNA donor, a 400 bp double-stranded DNA fragment was also tested, which was amplified from the wild type EmGFP gene using a pair of forward 5′-atggtgcggaagcgcgagtagctp-3′ and reverse 5′-gttcctcttgaaagcgtatgccc -3′ primers. For each assay, 300 ng to 500 ng dsDNA was used. At 48 h post transfection, the cells were analyzed by flow cytometry. Alternatively, the genomic loci were PCR-amplified with the corresponding primers. The resulting PCR fragments were analyzed using the GeneArt<sup>®</sup> Genomic Cleavage Detection assay or restriction digestion, as well as cloning and then sequencing.

2.8. Delivery optimization of Cas9 RNP and donor DNA

To measure HDR efficiency, we engineered GripTite<sup>®</sup> HEK293 and HEK293PT stable cell lines expressing EmGFP and eBFP respectively (Supplementary Fig. S1). Cas9 RNPs were subsequently used to target the fluorogenic region of EmGFP to generate a disrupted EmGFP stable cell line containing a deletion of six nucleotides (Supplementary Fig. S1A). The deletion of Thr63 and Phe64 residues resulted in ablation of EmGFP activity, which could be restored by introducing an exogenous wild type donor DNA molecule. The donors introduced the desired change only without disrupting the corresponding PAM. As depicted in Supplementary Fig. S1B, when the disrupted EmGFP stable cells were transfected with Cas9 RNP and ssDNA donor, a significant number of EmGFP-positive cells were observed. Conversely, in the absence of ssDNA donor (Cas9 RNP alone) or gRNA (Cas9/donor), almost no EmGFP-expressing cells were detected. For measurement of homologous recombination activity using the eBFP expressing HEK293 stable cells, a single nucleotide transition of “C” to “T” converts a His to a Tyr at residue 66, resulting in the conversion of eBFP into the closely related GFP (Supplementary Fig. S1C). As shown in Supplementary Fig. S1D, when eBFP-expressing cells were transfected with Cas9 RNP plus ssDNA donor, a significant number of GFP-positive cells were detected. As expected, predominantly eBFP-positive cells, but very few GFP-positive cells were detected in the absence of gRNA. The time-lapse video for HDR was recorded every 2 h for a total of 72 h (Supplementary Video S1).

After validating our HDR assay systems, we optimized the delivery of Cas9 RNP and donor DNA. Most Neon<sup>®</sup> optimization programs worked well for delivery of Cas9 RNP into HEK293 cells (Supplementary Fig. S2). The Pstd program with the voltage set at 1150 V, pulse width set at 20 ms, and 2 pulses was used for the subsequent study. Initially, we co-delivered Cas9 RNP with a 97 base single-stranded PAM or non-PAM oligonucleotide into HEK293 cells. The PAM ssDNA oligonucleotide donor was defined as the strand containing the PAM (NGC) sequence (Fig. 1A). Since the program used for Cas9 RNP delivery might not apply to the delivery of donor DNA, we tested the sequential delivery of Cas9 RNP and donor DNA. The Cas9 RNP was first delivered into HEK293 cells via electroporation, the electroporated cells were then washed once with Resuspension Buffer R and the cell pellets were resuspended in Buffer R containing ssDNA or dsDNA donor. The cell suspension was then delivered using the Neon<sup>®</sup> 24-well optimization protocol (Supplementary Figs. S3 and S4).
3. Results

3.1. Effect of sequential delivery of nuclease and donor DNA

We observed approximately 5% and 6% EmGFP-positive cells using the PAM or non-PAM ssDNA oligonucleotides, respectively, when co-delivered with cas9 RNPs (Fig. 1B). The sequential delivery of Cas9 RNP followed by donor DNA resulted in more than a two-fold increase in EmGFP-positive cells regardless of the use of ssDNA or dsDNA donor (Fig. 1, panels B, D). The reverse sequential delivery of ssDNA donor first and then Cas9 RNP exhibited a similar effect (Fig. 1B). However, two consecutive electroporations without the intermediate wash step only showed mild improvement over the co-delivery of Cas9 RNP and ssDNA donor. The use of non-PAM strand donor exhibited slightly higher HDR efficiency than the PAM strand donor (Fig. 1B). The effect of sequential delivery was also observed in another reporter cell line system in which eBFP was converted to GFP by a single nucleotide substitution (Fig. 1E). The dosage titrations of ssDNA donor indicated that the optimal amount of ssDNA oligonucleotide was approximately 10 pmol or 0.2–0.5 μg per 10 μL reaction (Fig. 1C). Upon optimization, we observed approximately 15% EmGFP-positive cells using the non-PAM ssDNA oligonucleotide (Fig. 1C and Supplementary Fig. S3). The HEK293 cell viability for co-delivery of Cas9 RNPs and
donor DNA was approximately 85%, whereas the cell viability for sequential delivery of Cas9 RNPs and donor decreased to approximately 50%. In line with what we previously reported, the optimal electroporation conditions are highly dependent on cell type and should be determined experimentally (Liang et al., 2015).

3.2. Effects of oligonucleotide length and modification on HDR

It has been reported that relatively short single-stranded oligonucleotides containing 25–61 bases that are homologous to the target sequence were capable of correcting a single point mutation (Igoucheva et al., 2001). The use of phosphorothioate modification of nucleotides has also been shown to prevent degradation of oligonucleotide therapeutic agents in serum and cells (Brown et al., 1994; Eckstein, 2000; Orlando et al., 2010). Here we examined the effect of oligonucleotide length and modification on HDR efficiency in our system. The oligonucleotides were chemically synthesized and PAGE-purified with and without phosphorothioate modification at both the 5’ and 3’ ends and a phosphate at the 5’ end with a total length that varied from 40 to 100 bases. The desired mutation was positioned at the center of the oligonucleotide. As shown in Fig. 2A and B, the optimal ssDNA oligonucleotide length was approximately 80 bases with 36–40 bases of homology on each side of the edit. Oligonucleotides with shorter homology arms reduced the HDR efficiency significantly, whereas the longer homology arms also showed a slightly decreased efficiency. The phosphorothioate modifications improved the efficiency for both the 6-base insertion and one-base substitution. Using an 80-base modified oligonucleotide, we observed approximately 45% GFP-positive cells while introducing a single nucleotide substitution. To confirm the HDR result obtained from flow cytometry, the PCR fragments were cloned and 96 clones were randomly picked for sequencing regardless of GFP status. As shown in Fig. 2C and D, approximately 6% of the BFP- > GFP colonies contained the wild type sequence (BFP+), indicating that the overall genome modification efficiency was nearly 94%. Among the remaining 94% modified cells, approximately 54% of them harbored indels (BFP-). Approximately 40% of the clones contained the correct point mutation to make them GFP+, which was in agreement with the flow cytometry assay.

3.3. Double Strand Breaks in the immediate vicinity of the altered locus facilitates HDR

In the design of gRNAs for homologous recombination, introduction of the cleavage site in close proximity to the altered locus has been recommended (Inui et al., 2014). However, the ability to accomplish this would depend on the availability of a PAM site near the altered locus. To test this, we designed a set of 12 gRNAs flanking the 6-base insertion site in EmGFP (Fig. 3A). The gRNAs were enzymatically synthesized by the GeneArt™ Precision gRNA Synthesis Kit and then complexed with Platinum™ Cas9 nuclease protein. The resulting Cas9 RNPs were delivered to cells by electroporation and the genome cleavage efficiencies were determined at 48 h post transfection. As can be seen in Fig. 3B, some gRNAs were more active than others. To evaluate the effect of distance between the DSB and the altered locus on HDR, we delivered Cas9 RNP and ssDNA or dsDNA donor into cells sequentially and then determined the percentage of EmGFP-positive cells using flow cytometry. As depicted in Fig. 3C, the gRNAs (−7, −3, +3, and +5) in close proximity to the insertion site produced the highest percentages of EmGFP-positive cells. Although the +3 gRNA was closer to the insertion site than the +5 gRNA, the +3 gRNA exhibited lower HDR efficiency than the +5 gRNA, likely because the genome cleavage efficiency of the +3 gRNA was two-fold lower than that of the +5 gRNA (Fig. 3B and C). Under optimal conditions, we observed more than 30% EmGFP-positive cells using the −3 or +5 gRNAs, which represented more than 200-fold increase in knock-in efficiency over the donor-only negative control.

3.4. Asymmetric PAM and non-PAM ssDNA donors facilitate HDR

A recent report showed that an asymmetric ssDNA donor, complementary to the target strand with 36-bases on the PAM-distal side and a 91-base extension on the PAM-proximal side of the break, enhanced HDR efficiency (Richardson et al., 2016). It was proposed that when Cas9 cleaved the target loci, the 3’ end of the PAM-distal strand could dissociate from the RNP/dNA complex and initiate HDR by annealing to a donor complementary to this exposed sequence, suggesting that a donor designed in this manner would be preferred. However, we observed only a slight difference in HDR efficiency between the symmetric PAM (corresponding to the non-target strand in Richardson’s paper) and non-PAM (corresponding to the target strand in Richardson’s paper) strands (Fig. 1B). To further understand the mechanism of ssDNA donor-mediated HDR, we designed gRNAs to introduce DSBs both upstream (−3 gRNA) and downstream (+5 gRNA) of the insertion site in our emGFP cell line (Fig. 4A). Furthermore, we designed a set of asymmetric PAM strand and non-PAM strand ssDNA donors with 30-bases on one homology arm and 51-bases or 67-bases on the other homology arm (Fig. 4B and Supplementary Fig. 5B: D). The PAM strand was defined as the strand containing the NGG PAM sequence. The symmetric ssDNA donors served as controls. The GFP+ cell percentage determined by flow cytometry was plotted separately for each individual gRNA (Supplementary Fig. 5). For clarification, only a subset of asymmetric ssDNA donors were shown in Fig. 4B and C with the percentage of GFP+ cells normalized to the cleavage efficiencies (Fig. 3B).

When either the −3 or +5 gRNAs were used to generate DSB with its PAM site located upstream or downstream (respectively) of the insertion site (Fig. 4A), the asymmetric PAM strand ssDNA donor 67–30 and non-PAM strand donor 67–30 yielded the highest HDR efficiencies compared to the 30–67 designs (shown in Fig. 4B,C). This suggests that a 30 base 3' homology arm is favored over longer arms of 67 bases on the 3’ end. This fits the model described in Fig. 4A where the resected DSB allows access to either 3’ overhang for annealing. Finally, the symmetrical donors shown in Fig. 4C (and in Supplementary Fig. S5) show an intermediate efficiency, suggesting the optimal 3’ homology arm length in this model system could be near 30 bases but not as much as 50 bases.

The results were further validated using the same reporter system as described in Richardson’s paper, using a gRNA targeting the eBFP gene. Using symmetric donors with 35 bases of homology on the 3’ end (PAM 65–35 or non-PAM 65–35) resulted in approximately 52% and 48% HDR efficiency respectively (Supplementary Fig. S6A, panel a, and in Fig. S6C). Using asymmetric donors with 65 bases of homology on the 3’ end (PAM 35–65 or non-PAM 35–65) were less effective. Similar results were seen when using Cas9 mRNA and asymmetric donors (Supplemental Fig. S7).

Since these reporter systems are targeting a single copy of an artificial gene, we tested the asymmetric design to insert a 6 base HindIII restriction site into the beta-actin gene in iPSC cells. (Supplementary Fig. S8). Consistent with the GFP results, we observed that the best insertion efficiency of up to 25% (45% when normalized to the cutting efficiency) was with asymmetric donors that had a short 30 base 3’ homology arm and long 60 base 5’ arm (PAM 60-30 or non-PAM 60-30) compared to about 13% for symmetrical donors with 45 base homology arms. Overall, the use of either the asymmetric PAM strand or non-PAM strand ssDNA donor, which harbors approximately 60–67 bases of homology on the 5’ end and 30–35 bases of homology on the 3’ end, resulted in the highest efficiency of HDR regardless of
which genomic strand contained the PAM or whether the DSB was upstream or downstream of the edit site, inferring a common intermediate for HDR. Contrary to the proposed model of Richardson, we observed no bias in a donor design favoring the genomic strand that is released by the Cas9 complex. Instead it appears that once Cas9 is released, both template strands are resected and can initiate repair by annealing to the appropriate ssDNA donor. Although the mechanism is not understood and requires further study, it may rely on an interaction of the 5' resection rate of the cut template strand (Zierhut and Diffley, 2008) and the stability of donor annealing. It may be that the natural 5' resection results in 3' overhangs that are ~30–40 bases which allows for initiation of stable annealing and repair with the donor directly. With shorter arms, the donor annealing may not be stable enough to promote efficient initiation of repair. With longer arms, the long 3' donor arm may need to be trimmed, await further resectioning of the template, or invade the partially resected template strand.

3.5. Short double-stranded DNA donor with single-stranded overhangs facilitates highly efficient HDR

The work of asymmetric ssDNA donors described above suggested that only about 30 bases at the 3' end were needed for sufficient single stranded DNA annealing. To extend this concept, we hypothesized that a dsDNA donor harboring single-stranded overhangs would facilitate HDR to higher levels than with blunt ends. To test this hypothesis, we designed and generated a series of donor molecules with either blunt end, 5' end protrusion or 3' end protrusion by annealing two small single-stranded oligonucleotides. A single-stranded DNA donor was used as a control. The 5' and 3' ends of the oligonucleotides were protected with two consecutive phosphorothioate-modified bases (Supplementary Table S1). For proof of concept, we inserted a 30 nucleotide FLAG epitope tag along with an EcoRI site into the BFP gene stably expressed in HEK293 cells. The gRNA was designed to target the sense DNA strand. The length of single-stranded overhangs varied from 6 nucleotides to 30 nucleotides. The oligonucleotides were denatured and annealed prior to transfection to form the structures displayed in Fig. 5A. The Cas9 RNP and donor DNA were delivered sequentially to HEK293 cells by electroporation and editing was measured after 48 h by the GCD assay. Co-delivery of Cas9 RNPs and donor DNA gave lower HDR efficiencies. As shown in Fig. 5A, approximately 75% cleavage efficiencies were observed with various donor configurations. When the PCR fragments were subjected to restriction digestion with EcoRI to identify properly inserted constructs, only the donor DNA molecules containing 30-base single-stranded overhangs at the 3' end produced the expected digested fragments (30-3' in Fig. 5A). The double stranded donor with 30-base 3' overhangs was inserted with efficiencies above 35% (~44% when normalized to the cutting efficiency) while the ssDNA donor was inserted with approximately 20% efficiency. Upon close examination of the length of the single-stranded overhang, we found that the optimal length of the 3' single-stranded overhang was 30 to 36 nucleotides with a digestion efficiency of approximately 40% (Fig. 5B). Phosphorothioate modification at both 5' and 3' ends of the donor (30-3') exhibited approximately 42% digestion efficiency.
Fig. 3. DSB in close proximity to the insertion site enhanced HDR. (A) A series of gRNAs were designed and synthesized flanking the insertion site (↓) targeting either the top strand (▼) or bottom strand (▲). The number and ± signs indicate the position of DSB upstream (−) or downstream (+) of the insertion site (0). (B) A series of gRNAs were associated with Cas9 nuclease separately and the resulting Cas9 RNPs were transfected into disrupted EmGFP stable cell lines. The percentages of Indel were evaluated at 48 h post transfection. (C) A series of Cas9 RNPs along with a 400 bp dsDNA donor or 97-base PAM ssDNA donor were sequentially delivered to disrupted EmGFP stable cell lines. Samples in the absence of donor (+gRNA) or gRNA (-gRNA) served as controls. The percentages of GFP-positive cells were determined by flow cytometry at 48 h post transfection.

Fig. 4. Asymmetric ssDNA donors enhanced HDR. (A) Two separate gRNAs flanking the insertion site (↓, 0) were designed and synthesized with double-stranded breaks (DSB) occurring at position −3 and +5 separately (▲). Upon end resection of DSB, the 3′ recessive ends were generated in two opposite orientations, which could anneal to either PAM (a) or non-PAM (b) ssDNA donors (red). The PAM ssDNA oligonucleotide is defined as the strand containing the NGG PAM sequence. (B) A series of ssDNA donors were designed with a variable number of nucleotides on the left arm (−) and right arm (+) of the insertion site. Both the PAM and non-PAM strands were tested. The Cas9 RNP (1.5 ug Cas9 nuclease, 360 ng gRNA) and ssDNA donors (10 pmol) were sequentially delivered to disrupted EmGFP stable HEK293 cell lines. At 48 h post transfection, the% Indel was determined by the GCD assay (Fig. 3B), whereas the percentages of EmGFP-positive cells were determined by flow cytometry. The bar graphs represented the normalized HDR efficiency (% EmGFP+ cells/% Indel) with averages of three individual experiments.
Fig. 5. Insertion of a FLAG tag along with an EcoRI site using dsDNA donor with single-stranded overhangs. (A) Various donor DNA molecules containing a 30-base FLAG tag along with an EcoRI site were designed and synthesized, including single-stranded DNA donor (ssDNA), blunt-end dsDNA donor (blunt), dsDNA donor with 5′ overhang (5′Δ), dsDNA donor with 3′ overhang (3′Δ). The length of overhangs were 6 nucleotides (6), 15 nucleotides (15) or 30 nucleotides (30). The 3′ and 5′ ends of the oligonucleotides harbored two consecutive phosphorothioate-modified bases (Supplementary Table S1). The short dsDNA donors with and without overhangs were prepared by annealing two short DNA oligonucleotides. The Cas9 RNP targeting the eBFP gene and various forms of DNA donors were sequentially delivered to HEK293 cells expressing eBFP. At 48 h post transfection, the eBFP locus was PCR-amplified. The resulting PCR fragments were analyzed by the genomic cleavage and detection assays to determine the percentage of Indel or subjected to restriction digestion with EcoRI to determine the percentage of digestion. (B) The dsDNA donors with 15, 24, 30, 36, or 45-base 3′ overhang were sequentially delivered with Cas9 RNP to HEK293 cells expressing eBFP. Alternatively, a dsDNA donor with 30-base 3′ overhang but without phosphorothioate modification (30−3′n) was used. The percentage of digestion with EcoRI was determined at 48 h post transfection. (C) Cas9 RNP and various amount of ssDNA donor or dsDNA donor with 30-base 3′ overhangs were sequentially delivered to HEK293 cells expressing eBFP. The eBFP loci were PCR-amplified. The resulting PCR fragments were analyzed by EcoRI digestion. (D) The PCR fragments were cloned into E.coli and 192 clones were randomly picked for sequencing. The relative percentage of wild type (wt), NHEJ, and HDR clones derived from either ssDNA donor (ssDNA) or dsDNA with 3′ overhangs (3′ overhang) was plotted. The white rectangles represented the population of clones that contained the insert but with a point mutation. Examples of edited clones were shown in (E) (not representing the actual percentages of NHEJ and HDR. The underlined sequences represented the FLAG tag along with an EcoRI site.

whereas the donor DNA without phosphorothioate modification (30−3′n) lowered the efficiency to around 27% (Fig. 5B). Furthermore, upon titration of the ssDNA donor and dsDNA donor with 30-base 3′ end protrusions, we found that the optimal concentration of DNA donors in the transfection reaction was approximately 1 μM (Fig. 5C). Under these conditions, we measured greater than
40% digestion efficiency with the 3′-protruded dsDNA donor compared to 25% using a ssDNA donor. When we performed sequencing analysis of 192 clones, 13% of the clones were wild type and 48.8% of the clones contained indels suggesting that in these clones the cleavage was repaired by NHEJ. Although 39% of the clones contained the insert, about 4% of the clones also harbored a point mutation (Fig. 5D, white rectangle). Excluding all the errors and wild type clones, 34% of the clones harbored the correct insertion.

We also analyzed the edited locus where a ssDNA oligonucleotide served as donor. In this case, approximately 9% of the clones were wild type, 61% of the clones were NHEJ, 9% of the clones harbored the insertion but with a single base mutation, and 21% of the clones contained the correct insert (Fig. 5D). The point mutations in the clones containing the insert are most likely due to an error in the synthesis of the DNA oligonucleotide which has been reported to be 1–3 errors/kb. The synthesis of longer oligonucleotides has also been shown to be more error prone than shorter ones which may explain why the long 90 base ssDNA donor resulted in more mutations than the dsDNA donor made from two 60 base oligonucleotides (Xiong et al., 2004).

In order to understand the effect of the polarity of dsDNA donor with single stranded overhangs, we inserted a FLAG epitope tag along with an EcoRI site into a separate locus where the +5 gRNA was targeting the bottom strand of EmGFp gene (Supplementary Fig. 5S). The Cas9 RNPs were first delivered into GrippTite™ HEK293 cells, followed by the delivery of ssDNA donor, or short dsDNA donors with 32-base single stranded overhangs at either 5′ or 3′ end (Supplementary Fig. S5A). Samples in the absence of gRNA served as controls. At 48 h post transfection, the genomic loci were amplified by PCR. The resulting PCR fragments were subjected to restriction digestion with EcoRI (data not shown) or subjected to sequencing analysis. The integration efficiency of the FLAG epitope tag along with the EcoRI site was approximately 34% using dsDNA donor with 3′ overhangs and 20% using ssDNA donor (Supplementary Fig. S5B). The dsDNA donor with 5′ overhangs resulted in a barely detectable integration product. The results indicated that the polarity of single stranded overhangs remained the same regardless of how the DSBs were introduced by Cas9 RNPs.

4. Discussion

We have demonstrated that the design and delivery of gRNAs, Cas9 nuclease, and donor molecules are critical to achieve high HDR efficiencies. Ideally, in order to achieve high editing efficiency, the double-stranded break induced by Cas9 nuclease should be in close proximity to the edit site, as just a few additional bases further up or downstream can make a significant difference in editing efficiency. One limitation of the CRISPR system for precise editing is exposed here since the location of a potential DSB site, and consequently the efficiency of donor insertion into the genome, is dictated by the availability of PAM sites in proximity to the intended edit. Furthermore, even though a gRNA target site happens to be in the immediate vicinity of the edit locus, it is not guaranteed to have high modification efficiency. The gRNA activity may depend on the nature of the gRNA sequence, chemical modification, and propensity of the DSB to be repaired perfectly, as well as accessibility of the targeted genomic locus. Finally, the off-target cutting potential for each gRNA must also be considered. In this regard, alternate tools such as TALENs mutated to lack the 5′ T Targeting requirement have an inherent advantage over CRISPR since they can be programmed to target virtually anywhere in the genome without PAM restrictions (Lamb et al., 2013). Some protocols recommend disrupting the PAM with a silent mutation in the donor to prevent recutting of the edited target by the active Cas9 complex (Kim et al., 2014). Using Cas9 RNPs, we were able to achieve high editing efficiencies with donors designed to only introduce the desired mutation (either 6 base insertions or 1 base substitution) and we did not disrupt the PAM site so changing the PAM may not be required. The Cas9 RNP has a short half-life in cells so it has diminished opportunity to recruit as compared to plasmid derived Cas9 which can persist at high levels for several days (Liang et al., 2015). However, our observation may be due to efficient disruption of the seed recognition sequence. A 6 base insertion is expected to disrupt the seed sequence and so a PAM change should be unnecessary. A single base substitution, especially if it is distant from the PAM may be more tolerated and a PAM change could improve the editing efficacy.

If the Cas9 RNPs are efficiently delivered to the nucleus for generation of double stranded breaks and the donor molecules are readily available at the time of DNA cutting and repair, the HDR efficiency can be significantly improved. The HDR frequencies depend on the dose of donor DNA molecules with the optimal delivery concentration being approximately 1 μM. The optimal length of the ssDNA donor is approximately 70 to 100 nucleotides with a 35–50 base homology arm on either side of the edit site. While longer ssDNA donors can be used we did see an increased incidence of errors with longer oligos, most likely due to the chemical synthesis process. Moreover, the protection of donor DNA with phosphorothioate modification can improve HDR efficiency. The delivery conditions for Cas9 RNPs and donors are also crucial as we observe that sequential delivery of Cas9 RNPs and donor DNA facilitates HDR. This may be due to the Cas9 protein having non-specific DNA binding activity, leading to decreased transfection efficiency when paired with the donor. However, sequential delivery may not be applicable to cells that are sensitive to multiple rounds of electroporation. For example, we choose co-delivery in iPSCs since we observed that two rounds of electroporation resulted in >50% cell death. In iPSC, the co-delivery of Cas9 RNPs and ssDNA donor produced up to 24% HDR efficiency. The use of stable Cas9-expressing cells can be beneficial for genome editing because the delivery of Cas9 nuclease is not necessary, resulting in increased transfection efficiency of gRNA and/or donor DNA. For example, we observed precise genome editing rates of up to 40% in Cas9-expressing iPSCs for a single nucleotide substitution at multiple genomic loci (Supplemental Fig. S10). However, extra effort is required to generate the stable cell lines expressing Cas9 nuclease with the added risk for a higher off-target effect. Co-delivery of Cas9 mRNA with a gRNA and donor can also be an effective alternative.

The donor design and configuration also contribute to the editing efficiency. A recent report showed that asymmetric design of ssDNA donors promoted HDR by overlapping the Cas9 cut site with 36-bases on the PAM-distal side and with a 91-base extension on the PAM-proximal side of the break. A donor DNA complementary to the non-target strand stimulated HDR frequencies up to 2.6-fold greater than those obtained with a donor DNA complementary to the target strand (Richardson et al., 2016). However, we observe that both the asymmetric PAM strand (corresponding to the non-target strand in Richardson’ paper) and non-PAM strand (corresponding to the target strand in Richardson’ paper) enhance HDR regardless of the orientation of the Cas9 nuclease compared to a standard symmetrical donor designs. This discrepancy could be due to their use of donors with 3′ homology arms of different lengths. They had highest efficiencies with a 36 base 3′ homology arm to the target strand which agrees with our observation. However, when testing the corresponding non-target donor, they appeared to only test donors with a short 21 base 3′ homology arm.

We have shown that unprotected short (<30 bases) homology arms have significantly reduced editing efficiency. Thus, we propose that after Cas9 nuclease cleaves, both sides of the double-stranded break are recognized by the DNA repair machinery equally.

In this model, a repertoire of cellular proteins involved in DNA repair is recruited to the cleaved ends to repair the damaged DNA.
via either NHEJ or the HDR pathway (Fig. 6). In order for HDR-mediated donor insertion to occur, cellular exonucleases excise back the 5′ ends thereby generating 3′ overhangs on either side of the break (Nimonkar et al., 2008), which can anneal to the 3′ end of either a PAM or non-PAM ssDNA donor. In either case, the 3′ protruding end of the break will anneal with the complementary ssDNA donor and then be extended by DNA polymerase with the 5′ region of the ssDNA donor serving as template (Fig. 6A). It is unclear how the other 3′ protruding end of the break joins with the newly-extended dsDNA donor region to repair the lesion but it may be similar to synthesis dependent strand annealing (SDSA) (Kan et al., 2014). It appears that 30–36 nucleotides are sufficient for the initial ssDNA annealing. The polarity and spacing of the template 3′ protruding ends are further confirmed by the use of short dsDNA donors with single stranded overhangs. Interestingly, short dsDNA molecules, ideally with 30 to 36 base single-stranded overhangs at the 3′ ends, appear to be used efficiently in the HDR pathway regardless of whether the gRNA targets the top or bottom strand. In the model, the single strand homologous region of the dsDNA donor can readily anneal to both sides of the resected template DNA (Fig. 6B). This design allows for larger insertions such as epitope tags and may allow for even larger tags such as GFP to be added. These results support a model whereby after the DSB is made a common “3′ recessed ends” intermediate is formed by the HDR machinery and can be used with many complementary donor DNA molecules containing matching 3′ homology arms. This model and associated data favor the design of donor DNAs with the insertion or SNP repair element as close to the cleavage site as possible and 3′ protruding single strand homology arms of approximately 30–36 bases for larger donor molecules. For smaller single stranded donor molecules, 30–35 base arms 3′ to the insertion/repair cassette and greater than 40 bases on the 5′ end seems to be favored.

Acknowledgments

We acknowledged Mahalakshmi Sridharan, Sridhar Ranganathan, Sanjay Kumar, Hillary Barbour, Jarrod Clark, Wen Chen, Jason Carter, Jian-Ping Yang, Michael Poderyczyk and Natasha Roark for providing technical support and inputs on the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jbiotec.2016.11.011.

References


学霸图书馆

www.xuebalib.com

本文献由“学霸图书馆-文献云下载”收集自网络，仅供学习交流使用。

学霸图书馆（www.xuebalib.com）是一个“整合众多图书馆数据库资源，提供一站式文献检索和下载服务”的24小时在线不限IP图书馆。

图书馆致力于便利、促进学习与科研，提供最强文献下载服务。

图书馆导航：

图书馆首页 文献云下载 图书馆入口 外文数据库大全 疑难文献辅助工具