

Optimized CRISPR/Cas9 Strategy for Homology-Directed Multiple Targeted Integration of Transgenes in CHO Cells

Sung Wook Shin¹ and Jae Seong Lee¹

¹Ajou University

May 5, 2020

Abstract

Site-specific integration has emerged as a promising strategy for precise Chinese hamster ovary (CHO) cell line engineering and predictable cell line development. CRISPR/Cas9 with homology-directed repair (HDR) pathway enables precise integration of transgenes into target genomic sites. However, inherent recalcitrance to HDR-mediated targeted integration (TI) of transgenes results in low targeting efficiency, thus requires selection process to acquire targeted integrant in CHO cells. Here we explored several parameters that influence the targeting efficiency using the promoter-trap based single or double knock-in (KI) monitoring system. A simple change in the donor template design by adding sgRNA recognition sequences strongly increased KI efficiency by 2.9–36 fold depending on integration sites and culture mode, compared with conventional circular donor plasmids. Furthermore, sequential and simultaneous KI strategies enabled the generation of double KI populations about 1–4% without the need of additional enrichment processes. This simple optimized strategy not only allowed efficient CRISPR/Cas9-mediated TI in CHO cells but also paved the way for the applicability of multiplexed KIs in one experimental step without the requirement of sequential and independent CHO cell line development procedures.

Chinese hamster ovary (CHO) cells are the most commonly used expression host of choice for manufacturing biopharmaceuticals. Development of recombinant CHO (rCHO) cell lines is critical for the biopharmaceutical industry for the production of recombinant therapeutic proteins (Walsh, 2018). Since the first use of CHO cells, stable cell line development (CLD) has been performed by untargeted random integration of transgenes followed by a selection and large-scale clone screening process. Although this traditional CLD method is easy to perform, the issue of clonal variation, particularly in transgene expression patterns, arises because of uncontrolled transgene integration sites, copy number variation, and transgene rearrangement upon random integration (Lee et al., 2018; Lee et al., 2019).

Site-specific integration of transgenes into designated genomic sites has been proposed as an alternative CLD method, as it can mitigate the high levels of heterogeneity seen in clonal rCHO cell lines, thus reducing CLD timelines (Lee et al., 2015b; Lee et al., 2019). The advent of programmable nuclease-mediated genome editing technology has accelerated targeted modification of the CHO genome by inducing a site-specific DNA double-strand break (DSB), which results in the activation of DNA damage repair pathways. Among various genome editing tools, CRISPR/Cas9 technology, which is based on the RNA-guided engineered nuclease, has been rapidly adopted in the CHO community due to its simple composition and high targeting efficiency (Lee et al., 2015a). Notably, CRISPR/Cas9 was successfully implemented to insert transgenes in the target CHO genomic sites in a precise manner with homology-directed repair (HDR) upon the creation of DSBs (Lee et al., 2015b; Lee et al., 2016).

The targeted integration (TI)-based CLD methods are proven to be effective in both academia and industry as they achieve reproducible transgene expression levels and stability (Grav et al., 2018; Zhang et al., 2015). Despite its many advantages, the low targeting efficiency of the current TI technology hinders multiplexed gene knock-ins (KIs), as well as high-throughput screening for identifying genomic hot spots yielding high

and stable transgene expression levels (Lee et al., 2019). Given the major clinical and commercial success of monoclonal antibody therapeutics and the emergence of next-generation difficult-to-express antibody products, including multi-specific antibody formats and fusion proteins (Spiess et al., 2015; Walsh, 2018), the demand for multiple KI based CLD will continue to grow because efficient expression of antibodies requires not only the balanced expression of heavy and light chains but also the co-expression of various engineering targets (Ho et al., 2013; Pybus et al., 2014). Selection-based strategy has previously been shown to target several single (Lee et al., 2015b) and multiple loci (Gaidukov et al., 2018) in CHO cells; however, sequential integration along with the use of different selection markers is not ideal for an industrial setting and is also time-consuming. We previously demonstrated a selection marker-free strategy to improve TI through fluorescent enrichment of cells with HDR-mediated genome editing; however, its targeting efficiency was low when challenged with integrating large transgenes (Lee et al., 2016).

In this study, we describe a simple optimized strategy that allows efficient CRISPR/Cas9- mediated TI of transgenes in a transient setup in CHO-K1 cells. To measure targeting efficiency, we used a promoter-trap based KI monitoring CHO-K1 cell line in which an enhanced green fluorescent protein (EGFP) expression cassette without a promoter region was integrated at the specific genomic site (Figure 1a). Co-transfection of a Cas9/single guide RNA (sgRNA) expression vector and donor plasmid, allowed TI of the promoter into the upstream of EGFP coding sequences, and this then restored EGFP expression (Figure 1a). Because the EGFP+ cells represent TI of the promoter, the percentage of HDR events can be quantitatively measured by flow cytometry. To select an efficient sgRNA target site, we designed two sgRNAs, and used sgRNA1 which showed higher editing efficiency that was detected by the T7E1 assay (Figure S1).

To test and optimize TI with the monitoring system, we determined KI efficiency of three exogenous constitutive promoters, including short variant of CMV (148 bp), human EF1 α (1167 bp) and Chinese hamster EF1 α promoter (1665 bp), 3 and 6 days after electroporation. (Figure 1b). The EGFP+ populations were visible 48 hours after transfection regardless of the promoter lengths. However, the percentage of EGFP+ populations was marginal when transfected with the Cas9 vector and individual donor plasmid. After subculturing on day 6, the EGFP+ populations were not detectable without sgRNA expression; however, they were noticeable with sgRNA expression (Figure 1b). Both electroporation and lipofection achieved detectable TI, although higher overall KI efficiency was seen using electroporation (Figure S2). Therefore, unless otherwise stated, we proceeded to use electroporation in the subsequent experiments.

The mean fluorescence intensity (MFI) of the EGFP+ populations showed different strengths of three exogenous promoters among which the EF1 α promoters exhibited higher EGFP expression levels (Figure 1c). This result is consistent with the previous study where EGFP expression levels were compared in targeted integrants (Lee et al., 2018). Genomic DNA based 5' Junction PCR analysis of the EGFP+ sorting pool further verified TI (Figure 1d). In addition, we successfully acquired EGFP expressing clonal cell lines upon limiting dilution of transfected pools of cells (Supplementary Table S1). The above results demonstrated that the monitoring system enables the detection of true positive KI events in a transient system.

With the promoter-trap based monitoring cell line, TI of three promoters was achieved in a transient setup. The KI efficiencies were comparable among promoters from 5 to 8% (Figure 1b). However, the efficiency was lower than that seen from the selection-based strategy (Lee et al., 2015b). Given the expectation that the efficiency should be lower when applying multiplex gene KIs or inserting longer transgenes, we modified the HDR donor vector design to enhance transient KI efficiency. Compared with a conventional circular donor (CCD), the modified donor, termed double cut donor (DCD), contained two sgRNA recognition sequences, which were identical to the genomic site targeting sgRNA/PAM sequence and flanked homology arms in the donor vector (Figure 2a). By simply adding sgRNA recognition sequences, genomic DSBs and *in vivo* linearization of target promoter sequences flanked by homologous sequences can be synchronized. *In vivo* cleavable donors were used to insert large gene of interests (GOIs) in a homology-independent manner (Cristea et al., 2013; He et al., 2016). However, non-homologous end-joining based KI approaches often cause incomplete integration with mutagenic junctions and the insertion of backbone fragments; thus, HDR-based KI approaches incorporating cleavable donors have been developed (Yao et al., 2017; Zhang et al.,

2017). Although previous studies have shown that the DCD functions efficiently in human cells (Zhang et al., 2017), its effect on transient KI efficiency remains unknown in CHO cells. To address this, we transfected Cas9/sgRNA1 expression vectors with either CCD or DCD of the three different promoters in the CHO-K1 monitoring cell line. The use of DCDs resulted in increased KI efficiencies of up to about 13-18%, which were seen 6 days after transfection, in the case of adherent cell line (Figure 2b). It corresponded to a 2.9–4.5 fold increase in the KI populations compared with CCDs. Interestingly, the effect of DCD on KI efficiency was more pronounced in the case of serum-free adapted suspension monitoring CHO cell line than that seen in the serum adherent monitoring cell line, corresponding to a 3.9-10.1 fold increase in the KI populations compared with CCDs (Figure 2c).

To maximize the targeting efficiency, we determined an optimal ratio of the DCD to Cas9/sgRNA expression vector. To test this, KI efficiencies were assessed in cells transfected with five different ratios of constructs (Figure S3). We found that the use of approximately equal amounts of two constructs, applied in Figure 2b and 2c, led to the highest KI efficiency, while increasing the level of either Cas9/sgRNA expression vector or DCD had negative effects on KI efficiency.

Based on the significant increase in KI efficiencies with DCDs, we postulated whether or not multiplexed gene KIs could be achieved in CHO cells without the need of an additional enriching process. To monitor and quantify double KI events, we established a double-KI monitoring CHO-K1 cell line by inserting a promoter-less TagRFP657 expression cassette into another locus, *C1GALT1C1* (*COSMC*), in the parental single KI monitoring cell line (Figure S4a). A sequence from the mouse *Rosa26* locus (Chu et al., 2015) was included upstream of the TagRFP657 coding sequence as an sgRNA target site. 5'/3' Junction PCR and out-out PCR was conducted to evaluate the correct integration at the *COSMC* locus (Figure S4b). We then evaluated the impact of DCDs to enable double TI by introducing two Cas9/sgRNA expression vectors and two human EF1 α promoter donor plasmids harboring 650 – 800 bp homology arms (Figure 3a). This homology arm length corresponds to the truncated TagRFP657 and EGFP coding sequences, as well as the range of the highest HDR with double cut donor, which leads to the transgene integration mediated by almost all complete HDRs (Zhang et al., 2017). In parallel, we tested CCDs for double TI. For the generation of double KI integrants, we rationally designed two strategies, simultaneous and sequential KI (Figure 3b). In the simultaneous KI strategy, all vector constructs needed for the TI were simultaneously introduced by a one-step transfection. Although simultaneous KI strategy is simple and fast, efficiency can be reduced as a result of the maximum quantity of vector required for the transfection being divided at individual sites. In contrast, the sequential KI strategy targets two loci one by one through two rounds of sequential transfection. In the sequential KI strategy, second transfection was conducted at passage 2 (time interval day 6). Regardless of the target site order, either site 1 followed by site 2 or site 2 followed by site 1, both EGFP+/TagRFP657+ populations were detected in a range of approximately 1–2% when using DCDs on day 12, in the adherent monitoring cell line (Figure 3c). Conversely, when using CCDs, EGFP+/TagRFP657+ populations were rarely detected, demonstrating the limitation of the conventional KI method. The KI efficiency at the *COSMC* locus was lower than that at the non-coding region, which highlights how genomic location affects KI efficiency. In the simultaneous KI strategy, double KI events were also distinctly detected only when using DCDs (Figure 3d). Although KI efficiency at individual sites was lower than that of sequential KI in the adherent monitoring cell line, the level of double-KI integrants was comparable at 1–2%. In the serum-free adapted suspension double-KI monitoring CHO-K1 cell line, the use of DCDs led to significant increase in KI efficiency, 7.3 and 36 fold increase in EGFP+ and TagRFP657+ populations respectively, compared with CCDs, which results in detection of double-targeted integrants about 2-4% on day 6 (Figure 3d).

In summary, we provide an efficient CRISPR/Cas9 mediated KI strategy that can be successfully applied to integrate single or double GOIs into target genomic sites in CHO cells. The promoter-trap based single or double KI monitoring CHO-K1 cell lines were established to detect and quantify KI events in a simple manner. Depending on the culture mode, making a change in donor DNA template design by adding sgRNA recognition sequences increased targeting efficiency by 2.9–10.1 fold in the non-coding region and 13.5-36 fold at the *COSMC* locus (Figure 2 and 3). This then led to simultaneous double KIs in a transient set up.

The simple modification of donor plasmid can be more general than changing the widely used Cas9/sgRNA vector; thus it will be easily applicable to existing CRISPR/Cas9 mediated TI for generating rCHO cell lines. In this way, we generated rCHO cell lines by TI without the need of selection and enriching processes and were also able to avoid the sequential CLD process for multiple KIs. Ultimately, this platform can significantly reduce CLD timelines and save efforts to engineer CHO cell lines.

MATERIALS AND METHODS

sgRNA design and plasmid construction

Cas9/sgRNA expression vectors and donor plasmids used in this study are listed in the Supplementary Table S2. Target site-specific sgRNA sequences were designed by the CRISPOR bioinformatics tool (version 4.95) (Concordet and Haeussler, 2018). Duplex of single stranded oligos were cloned into Cas9-T2A-mCherry and sgRNA expression vectors (Addgene #64324) (Chu et al., 2015). Donor plasmids were constructed via the uracil-specific excision reagent (USER) cloning method (Lee et al., 2015b) with uracil-containing primers listed in Supplementary Table S3. The primers used for DCDs contain sgRNA/PAM sequences for the backbone fragment. The PCR products were purified from 1% agarose TAE gel using NucleoSpin® Gel and PCR Cleanup kit (Macherey-Nagel, Duren, Germany). Four PCR amplicons including homology arms (5'HA and 3'HA), backbone (CCD or DCD), and promoter (short variant of CMV or human EF-1 α promoter or Chinese hamster EF-1 α promoter) were assembled with USER enzyme (New England Biolabs, Ipswich, MA); 3' homology arm lengths were optimized for the truncation of EGFP (~0.8kb) and TagRFP657 (~0.65kb) expression in donor plasmid. All templates for the PCR reactions are listed in Supplementary Table S4. The reaction mixtures were then transformed into *E.coli* One Shot® Mach1™ competent cells (Life technologies, Thermo Scientific, Rockford, IL) according to standard procedures. Plasmids were harvested from a single colony on LB agar plates containing ampicillin. All constructs were purified using NucleoBond Xtra Midi EF (Macherey-Nagel) according to the manufacturer's instructions, and validated by sanger sequencing by primers listed in Supplementary Table S3.

Generation of KI monitoring cell lines

The promoter-trap based KI monitoring CHO-K1 cell line harboring promoter-less EGFP expression cassette was generated in the previous study (Lee et al., 2018). Briefly, targeted integrants expressing GFP in the non-coding region of the CHO-K1 genome were transfected with two sgRNA expression vectors targeting entire CMV promoter regions, upstream of EGFP coding sequences, and then targeted clonal cells were isolated (Lee et al., 2018). The double KI monitoring CHO-K1 cell line was generated from the aforementioned single KI monitoring cell line. The single KI monitoring cell line was co-transfected with Cas9/sgRNA expression vector targeting *COSMC* locus and donor plasmid containing puromycin and promoter-less TagRFP657 expression cassette as described in Figure S4A. After two weeks of puromycin selection, a limiting dilution was conducted for single cell cloning. The selected clones were screened by out-out and 5'/3' Junction PCR (Figure S4B), and maintained in culture media with puromycin (3 μ g/ml) and G418 (500 μ g/ml). The adherent cell lines were adapted to serum free suspension culture. Cells were seeded at 5.0×10^5 cells/ml in 6-well clear flat bottom not treated cell culture plate (Falcon) with suspension culture media, Power CHO medium (Lonza) supplemented with 8 mM L-Glutamine (Hyclone). The fetal bovine serum (FBS, Hyclone) gradually decreased from 3% to 0% until viable cell density reached up to $1.0 - 1.5 \times 10^6$ cells/ml.

Cell culture and transfection

The adherent CHO-K1 cell lines were maintained in Dulbecco's modified Eagle's medium (DMEM, Gibco) supplemented with 10% (v/v) FBS (Hyclone). Cells were maintained as monolayer cultures in 25 cm² T-flasks (Nunc) with a working volume of 5 ml. The serum free suspension adapted CHO-K1 cells were maintained in Power CHO medium (Lonza) supplemented with 8 mM L-Glutamine (Hyclone) and cultivated as suspension culture in 125 ml Erlenmeyer flasks (Corning) with a working volume of 20 ml. All cells

were incubated in a humidified 5% CO₂ at 37°C without shaking (adherent), or with 120 rpm shaking (suspension). Viable cell concentration was measured by a Countess II FL automated cell counter (Invitrogen) with the trypan blue dye exclusion method. Two transfection methods were used in this study; electroporation and lipofection. Electroporation was conducted using NEPA21 electroporator (Nepagene, Japan). According to the manufacturer's instructions, adherent CHO-K1 cells were washed twice with Opti-MEM (Gibco) but not serum-free suspension CHO-K1. 1.0×10⁶ cells were pelleted and re-suspended in 100 μl mixture of Opti-MEM (Gibco) and 10 μg DNA of Cas9/sgRNA expression vector and donor plasmid at a ratio 1:1 (w:w). Poring pulse and transfer pulse condition were as follows: 135 V, 7.5 milliseconds (ms) pulse length, two pulses, 50 ms pulse interval, 10% decay rate with + polarity and 20 V, 50 ms pulse length, five pulses, 50 ms pulse interval, 40% decay rate with +/- polarity respectively. After the electroporation, cells were immediately seeded in 6 well plates with pre-warmed culture media. Lipofection was conducted using Lipofectamine 3000 (Invitrogen) as per the manufacturer's instructions.

Evaluation of KI efficiency by flow cytometry

72 hours post-transfection, cells were sub-cultured to 6 well plates at a seeding concentration of 2×10⁵ cells/ml. The remaining cells were re-suspended in PBS supplemented with 10% FBS and analyzed by FACSCalibur (Becton Dickinson, Franklin Lakes, NJ) to quantify the EGFP+ and/or TagRFP657+ cells. Similarly, after another 3 days, cells were re-analyzed by flow cytometry and the remaining cells were transfected for the sequential strategy. Cells were gated by forward scatter versus side scatter plots, and fluorescence threshold was set based on mock (no DNA control) to be about 0.1 %.

Genomic DNA extraction and Junction PCR

Genomic DNAs were purified from pellets of KI cell pools using GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific) and served as templates for PCR using the Phusion High-Fidelity PCR Master Mix (Thermo Fisher Scientific). To verify TI, 5' Junction PCR was conducted by touchdown PCR (98°C for 30 s; 10×: 98°C for 10 s, 72–62°C [-1degC/cycle] for 30 s, 72degC for 45 s [short variant of CMV, EF-1α], 1 μM [HEΦ-1α]. 30×: 98°C for 10 s, 62°C for 30 s, 72°C for 45 s [short variant of CMV, EF-1α], 1 μM [HEΦ-1α]. 72°C for 10 min). For the generation of double KI monitoring CHO-K1 cell line, out-out and 5'/3' Junction PCR was performed by touchdown PCR (98°C for 30 s; 10×: 98°C for 10 s, 70–60°C [-1degC/cycle] for 30 s, 72degC for 2 min 10 s [out-out], 1 min 20 s [5' Junction], 1 min [3' Junction]; 30x: 98degC for 10 s, 60degC for 30 s, 72degC for 2 min 10 s [out-out], 1 min 20 s [5' Junction], 1 min [3' Junction]; 72degC for 10 min). PCR primers are listed in Supplementary Table S3.

Acknowledgments

This work was supported by the NRF funded by the Korean government (2018R1C1B6001423 and 2019R1A6A1A11051471). The authors declare no competing financial interests.

References

- Chu, V. T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., & Kuhn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nature Biotechnology*, *33* (5), 543-548. <https://doi.org/10.1038/nbt.3198>
- Concordet, J. P., & Haeussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Research*, *46* (W1), W242-W245. <https://doi.org/10.1093/nar/gky354>
- Cristea, S., Freyvert, Y., Santiago, Y., Holmes, M. C., Urnov, F. D., Gregory, P. D., & Cost, G. J. (2013). In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration. *Biotechnology and*

Bioengineering, 110 (3), 871-880. <https://doi.org/10.1002/bit.24733>

Gaidukov, L., Wroblewska, L., Teague, B., Nelson, T., Zhang, X., Liu, Y., ... Weiss, R. (2018). A multi-landing pad DNA integration platform for mammalian cell engineering. *Nucleic Acids Research*, 46 (8), 4072-4086. <https://doi.org/10.1093/nar/gky216>

Grav, L. M., Sergeeva, D., Lee, J. S., Marin de Mas, I., Lewis, N. E., Andersen, M. R., ... Kildegaard, H. F. (2018). Minimizing Clonal Variation during Mammalian Cell Line Engineering for Improved Systems Biology Data Generation. *ACS Synthetic Biology*, 7 (9), 2148-2159. <https://doi.org/10.1021/acssynbio.8b00140>

He, X., Tan, C., Wang, F., Wang, Y., Zhou, R., Cui, D., ... Feng, B. (2016). Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology dependent and independent DNA repair. *Nucleic Acids Research*, 44 (9), e85. <https://doi.org/10.1093/nar/gkw064>

Ho, S. C., Koh, E. Y., van Beers, M., Mueller, M., Wan, C., Teo, G., ... Yang, Y. (2013). Control of IgG LC:HC ratio in stably transfected CHO cells and study of the impact on expression, aggregation, glycosylation and conformational stability. *Journal of Biotechnology*, 165 (3-4), 157-166. <https://doi.org/10.1016/j.jbiotec.2013.03.019>

Lee, J. S., Grav, L. M., Lewis, N. E., & Fastrup Kildegaard, H. (2015a). CRISPR/Cas9-mediated genome engineering of CHO cell factories: Application and perspectives. *Biotechnology Journal*, 10 (7), 979-994. <https://doi.org/10.1002/biot.201500082>

Lee J. S., Grav, L. M., Pedersen, L. E., Lee, G. M., & Kildegaard, H. F. (2016). Accelerated homology-directed targeted integration of transgenes in Chinese hamster ovary cells via CRISPR/Cas9 and fluorescent enrichment. *Biotechnology and Bioengineering*, 113 (11), 2518-2523. <https://doi.org/10.1002/bit.26002>

Lee, J. S., Kallehauge, T. B., Pedersen, L. E., & Kildegaard, H. F. (2015b). Site-specific integration in CHO cells mediated by CRISPR/Cas9 and homology-directed DNA repair pathway. *Scientific Reports*, 5 , 8572. <https://doi.org/10.1038/srep08572>

Lee, J. S., Kildegaard, H. F., Lewis, N. E., & Lee, G. M. (2019). Mitigating Clonal Variation in Recombinant Mammalian Cell Lines. *Trends in Biotechnology*, 37 (9), 931-942. <https://doi.org/10.1016/j.tibttech.2019.02.007>

Lee, J. S., Park, J. H., Ha, T. K., Samoudi, M., Lewis, N. E., Palsson, B. O., ... Lee, G. M. (2018). Revealing Key Determinants of Clonal Variation in Transgene Expression in Recombinant CHO Cells Using Targeted Genome Editing. *ACS Synthetic Biology*, 7 (12), 2867-2878. <https://doi.org/10.1021/acssynbio.8b00290>

Pybus, L. P., Dean, G., West, N. R., Smith, A., Daramola, O., Field, R., ... James, D. C. (2014). Model-directed engineering of "difficult-to-express" monoclonal antibody production by Chinese hamster ovary cells. *Biotechnology and Bioengineering*, 111 (2), 372-385. <https://doi.org/10.1002/bit.25116>

Spiess, C., Zhai, Q., & Carter, P. J. (2015). Alternative molecular formats and therapeutic applications for bispecific antibodies. *Molecular Immunology*, 67 (2 Pt A), 95-106. <https://doi.org/10.1016/j.molimm.2015.01.003>

Walsh, G. (2018). Biopharmaceutical benchmarks 2018. *Nature Biotechnology*, 36 (12), 1136-1145. <https://doi.org/10.1038/nbt.4305>

Yao, X., Wang, X., Hu, X., Liu, Z., Liu, J., ... Yang, H. (2017). Homology-mediated end joining-based targeted integration using CRISPR/Cas9. *Cell Research*, 27 (6), 801-814. <https://doi.org/10.1038/cr.2017.76>

Zhang, J. P, Li, X. L., Li, G. H., Chen, W., Arakaki, C., Botimer, G. D., ... Zhang, X. B. (2017). Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biology*, 18 (1), 35. <https://doi.org/10.1186/s13059-017-1164-8>

Zhang, L., Inness, M. C., Han, S., Moffat, M., Jones, H., Zhang, B., ... Young, R. J. (2015). Recombinase-mediated cassette exchange (RMCE) for monoclonal antibody expression in the commercially relevant

Figure legends

Figure 1. The occurrence of CRISPR/Cas9-mediated targeted integration was analyzed using promoter-trap based CHO-K1 monitoring cell line. (a) Schematic outline of the promoter-trap based CHO-K1 monitoring cell line. The monitoring cell line contains the promoter-less EGFP expression cassette in the non-coding region. After co-transfection of Cas9/sgRNA expression vector and promoter donor targeting the EGFP upstream region, EGFP expressing cells can be generated by the HDR pathway. (b) Analysis of knock-in efficiency of three different promoters. The promoter-trap based monitoring CHO-K1 cell lines were co-transfected with Cas9 or Cas9/sgRNA1 expression vector and three exogenous constitutive promoter donors including a short variant of CMV, EF-1 α and CHEF-1 α promoter. The EGFP+ percentage (%) was determined by flow cytometry three and six days after electroporation. (c) Mean fluorescence intensity (MFI) of the EGFP+ populations on day 6. MFI reflects the different strength of knock-in promoters. In (b) and (c), results are shown as the average values \pm standard error of the mean (SEM) in three independent experiments. (d) 5' Junction PCR validation of EGFP+ sorting pool cells. The transfected pool of cells were FACS sorted for EGFP+ populations, followed by genomic DNA purification. Promoter specific primer sets were used to amplify each 5' Junction region. The expected amplicon size of short variant of CMV (lane 2), EF1 α promoter (lane 4) and CHEF1 α promoter (lane 5) was 1313 bp, 1389 bp and 1841 bp respectively. The mock was no DNA control (lane 1,3).

Figure 2. Comparison of conventional circular donor (CCD) and double cut donor (DCD) on knock-in efficiency. (a) Schematic description of conventional CCD and DCD. The CCD is normally used plasmid donor, whereas DCD is a modified form where both 5' and 3' homology arms are flanked by two sgRNA recognition sequences which are identical to the target genomic sgRNA/PAM sequence. Comparison of knock-in efficiency of CCD and DCD in (b) adherent monitoring cell lines and (c) in serum-free suspension adapted monitoring cell lines. Cas9 or Cas9/sgRNA1 expression vector and three different promoter donors (CCD or DCD) were co-transfected to the promoter-trap based monitoring CHO-K1 cell line. The EGFP+ percentage (%) was measured by flow cytometry three and six days after electroporation. In (b) and (c), results are shown as the average values \pm SEM in three independent experiments, and the representative histograms are shown.

Figure 3. Generation of double-targeted integrants using double cut donors (DCDs) in the simultaneous and sequential knock-in (KI) strategies. (a) Schematic outline of double-KI monitoring CHO-K1 cell line. The promoter-less EGFP and TagRFP657 cassettes were integrated into the non-coding region (site 1) and *COSMC* locus (site 2), respectively. The second monitoring cassette includes mouse Rosa26 sgRNA targeting sequence upstream of TagRFP657 coding sequence. Co-transfection with the Cas9/sgRNA1 expression vector and donor plasmid targeting site 1 and/or Cas9/sgRNA_mouse Rosa26 expression vector and donor plasmid targeting site 2, results in single or double targeted integration. (b) Two strategies for double targeted integration. The simultaneous double-KI strategy delivers all vector constructs into cells through a single transfection process. (Cas9/sgRNA1, Cas9/sgRNA_mouse Rosa26, EF1 α donor targeting non-coding region, and EF1 α donor targeting *COSMC* locus). The sequential double-KI strategy creates double-KI pool cells by one by one targeting at 6-day time intervals through two rounds of targeted integration (Cas9/sgRNA1 and EF1 α donor targeting non-coding region for site 1 and Cas9/sgRNA_mouse Rosa26 and EF1 α donor targeting *COSMC* locus for site 2 independently). Flow cytometry analysis of double-KI efficiency in (c) the sequential KI strategy and (d) the simultaneous KI strategy. The EGFP+ percentage (%) was determined by flow cytometry on day 6 and/or day 12. In (d), both adherent and suspension monitoring cell lines were used. In (c) and (d), results are shown as the average values \pm SEM in three independent experiments, and the representative dot plots are shown.





