Molecular Biology Reports (2023) 50:5889–5899
https://doi.org/10.1007/s11033-023-08529-8

ORIGINAL ARTICLE

Efficient site-specific integration in CHO-K1 cells using CRISPR/Cas9-modified donors

Mohammad Hassan Kheirandish[1,2] · Behnaz Rahmani[2,4] · Hossein Zarei Jaliani[3] · Farzaneh Barkhordari[2] · Mohammad Ali Mazlomi[1] · Fatemeh Davami[2]

Received: 2 February 2023 / Accepted: 16 May 2023 / Published online: 27 May 2023

Abstract

Background Conventional methods applied to develop recombinant CHO (rCHO) cell line as a predominant host for mammalian protein expression are limited to random integration approaches, which can prolong the process of getting the desired clones for months. CRISPR/Cas9 could be an alternative by mediating site-specific integration into transcriptionally active hot spots, promoting homogenous clones, and shortening the clonal selection process. However, applying this approach for the rCHO cell line development depends on an acceptable integration rate and robust sites for the sustained expression.

Methods and results In this study, we aimed at improving the rate of GFP reporter integration to the Chromosome 3 (Chr3) pseudo-attP site of the CHO-K1 genome via two strategies; these include the PCR-based donor linearization and increasing local concentration of donor in the vicinity of DSB site by applying the monomeric streptavidin (mSA)-biotin tethering approach. According to the results, compared to the conventional CRISPR-mediated targeting, donor linearization and tethering methods exhibited 1.6- and 2.4-fold improvement in knock-in efficiency; among on-target clones, 84% and 73% were determined to be single copy by the quantitative PCR, respectively. Finally, to evaluate the expression level of the targeted integration, the expression cassette of hrsACE2 as a secretory protein was targeted to the Chr3 pseudo-attP site by applying the established tethering method. The generated cell pool reached 2-fold productivity, as compared to the random integration cell line.

Conclusion Our study suggested reliable strategies for enhancing the CRISPR-mediated integration, introducing Chr3 pseudo-attP site as a potential candidate for the sustained transgene expression, which might be applied to promote the rCHO cell line development.

Keywords Site-specific integration · CHO cell line development · CRISPR/Cas9 · Donor design · Tethering approach

Abbreviations

CHO	Chinese hamster ovary
rCHO	Recombinant CHO
mSA	Monomeric streptavidin
DSB	Double-strand break
HDR	Homology-directed repair
NHEJ	Non-homologous end-joining
HITI	Homology-independent targeted integration
KI	Knock-in
ssODN	Single-stranded DNA oligonucleotide
RNP	Ribonucleoprotein
hrsACE2	Human recombinant soluble ACE2
FBS	Fetal bovine serum.

✉ Mohammad Ali Mazlomi
ma.mazlomi@gmail.com

✉ Fatemeh Davami
f.davami@gmail.com

[1] Medical Biotechnology Department, School of Advanced Technologies in Medicine, Tehran University of Medical Sciences, Tehran, Iran

[2] Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran

[3] Department of Medical Biotechnology, School of Medicine, Shahid Sadoughi University of Medical Sciences, Yazd, Iran

[4] Department of Biotechnology, Faculty of Medicine, Semnan University of Medical Sciences, Semnan, Iran

Introduction

Chinese hamster ovary (CHO) cells are a predominant host for the production of biotherapeutics developed in mammalian expression systems, owing to their human-like

glycosylation patterns and growth in defined serum-free media [1, 2].

Different approaches have been implemented to develop high-yielding rCHO cells. However, most of them are based on random integration methods that are challenging and require several rounds of clonal selection to elect high-producer clones from the heterogeneous cell pool. Moreover, due to the poor control on the integration sites, even suitable clones are potentially prone to expression instability as time passes [3]. To address these limitations, site-specific integration methods have come to the fore to target the transgene to the pre-determined transcriptionally active sites of the genome, known as the hot spots, for the purpose of achieving high producing cells and minimizing the unwanted effects on the genomic context [4].

CRISPR/Cas9 has recently been introduced as a breakthrough system to apply the desired editing into the genome; it could serve in rCHO cell line development. Following Cas9-induced double-strand break (DSB), if the repair template is provided, the homology-directed repair (HDR) pathway promotes the knock-in (KI) of the transgene precisely. This is conventionally accomplished by a plasmid-based donor comprising the transgene (e.g., expression cassette) flanked by arms homologous to the targeted site. However, whereas most mammalian cells favor error-prone non-homologous end-joining (NHEJ) over the HDR pathway, the rate of CRISPR-mediated integration is low and considerable efforts have been made to increase it [5, 6].

In the recent years, the donor design has been an alternative approach to promote CRISPR-based KI efficiency. Despite direct HDR-enhancing manipulations using NHEJ inhibiting or HDR stimulating molecules [7, 8], donor design approaches do not evoke any potential toxicity for cells. It has been reported that the in vivo cleavage of donor vector facilitates transgene integration. Based on this method, the donor template is cleaved in the cell parallel with the targeted locus using two sgRNA-targeted sites flanking the homology arms [9, 10]. However, there are some complexities when designing such donor vectors. To address this issue, subsequent studies benefited from in vitro donor linearization by PCR amplification or the use of restriction enzymes and achieved a higher rate of KI efficiency, as compared to the circular donor in mouse/human embryos and primary T cells [11, 12]. In another approach, increasing the local concentration of donor in the vicinity of DSB lesion by tethering it to Cas9 or sgRNA was applied to improve the CRISPR-mediated KI rate [13]. However, previous tethering approaches have often been employed for small insertions using ssODN donors [14].

Coupled with the knock-in rate, finding the safe harbor loci for the sustained and elevated levels of protein expression following cassette integration is necessary for accelerating the rCHO cell line development by site-specific integration approaches. Bi et al. introduced phiC31 recombinase target sites termed pseudo-attP as the potential safe harbor loci for robust recombinant expression [15]. We recently characterized novel pseudo-attP sites in CHO-K1 cells [16]; however, their functionality to express a recombinant secretory protein has not been examined.

Here, we aimed to improve the conventional CRISPR-mediated integration efficiency by applying two methods, including the PCR-based donor linearization and the monomeric streptavidin (mSA)-biotin tethering approach, to target the 1.8 kb GFP-harboring cassette within the Chr3 pseudo-attP site of CHO-K1 cell line. We assumed that by tethering the biotinylated donor to the mSA moiety of the Cas9-mSA fusion through the high-affinity covalent bond, the distance between donor and DSB site could be reduced, causing higher knock-in efficiency. At the next step, as a practical application, the mSA-biotin tethering approach was employed to target the Chr3 pseudo-attP site with the human recombinant soluble ACE2 (hrsACE2) expression cassette as a potential therapeutic for SARS-CoV-2, which undergoes the phase II of the clinical trial. Our data, thus, showed the potentiality of the mSA-biotin tethering approach for improving the knock-in efficiency, as compared to the conventional plasmid-based donor in CHO-K1 cells. Moreover, the Chr3 pseudo-attP locus in CHO cells could be considered as a safe harbor locus for the recombinant protein expression.

Materials and methods

Vector construction

The Cas9/sgRNA vector (all-in-one) (GenBank accession no. OQ579018) contained sgRNA and Cas9-2A-mCherry expression cassettes. The sgRNA (5′-ATGCAGAACTAG AGTACAGC-3′) targets the phiC31 pseudo-attP intergenic site located on chromosome 3 of CHO-K1 genome (GenBank accession no. APMK01032147.1) (Fig. 1a). To construct the Cas9-mSA/sgRNA vector (GenBank accession no. OQ579019), the monomeric streptavidin (mSA) gene [17] was genetically fused to the C-terminus of Cas9 by quick-change PCR. As depicted in Supplementary Fig. 1, The mSA sequence was amplified by PCR using a set of primers to synthesize the chimeric mSA fragment containing the all-in-one backbone. The PCR product was gel purified, serving as a mega primer in the quick-change PCR step to complete the rest of the all-in-one plasmid by using the high-fidelity CloneAMP premix (Takara Bio Inc. Japan). The PCR product of the quick-change PCR was then digested with the DpnI restriction enzyme to eliminate the template vector and transformed into the

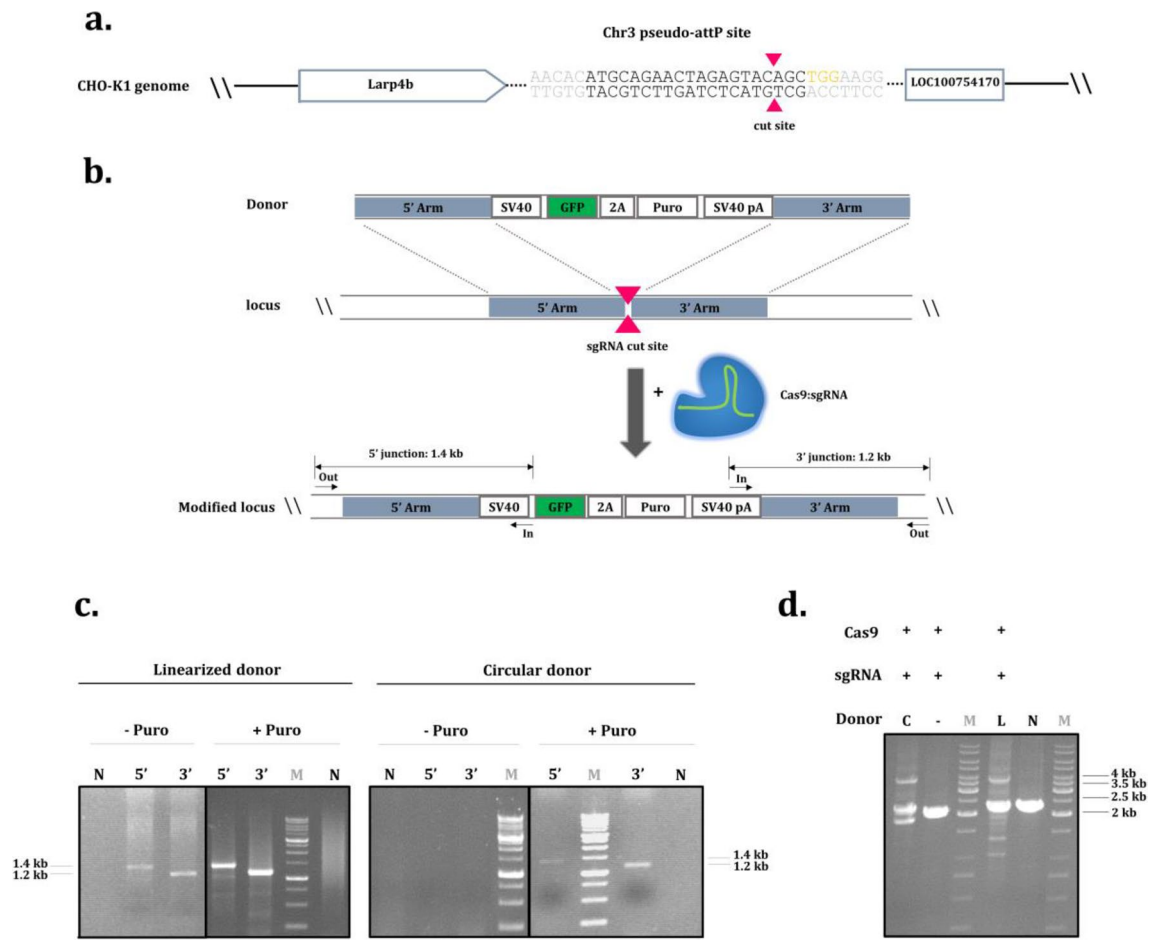

Fig. 1 CRISPR-mediated targeted knock-in into the Chr3 pseudo-attP site of CHO-K1 genome. **a** The location of the Chr3 pseudo-attP site on the CHO-K1 genome with the nearest upstream and downstream genes. The sgRNA target site was also illustrated with the PAM sequence (depicted in yellow). **b** Schematic overview of targeted integration into the CHO-K1 genome. A donor template containing the GFP-2A-Puro cassette flanked by 5′ and 3′ homology arms (1 kb) was used to target the pseudo-attP site by CRISPR/Cas9. The annealing sites of 5′/3′ junction PCR and out-out PCR primers were indicated. **c** 1% agarose gel of 5′ and 3′ junction PCR products of stable cell pools before and after puromycin selection following co-transfection with linearized or circular donor and Cas9/sgRNA vector. The expected band size for 5′ and 3′ junction PCR products was 1.4 kb and 1.2 kb, respectively. M, 1 kb DNA marker, and N, negative control (parental CHO-K1 genome). **d** 1% agarose gel of out-out PCR products of the stable cell pools. The expected band size for the out-out PCR product was 3.9 kb. The wild-type CHO band is 2.2 kb. *C* circular donor, and *L* linearized donor

DH5α strain of *E. coli* cells. The resulting colonies were confirmed by colony PCR and Sanger sequencing. The construct was purified by an endotoxin-free plasmid DNA miniprep kit (BioBasic, Canada).

A plasmid donor (GenBank accession no. OQ579020) containing GFP and puromycin expression cassettes (pSV40-GFP-2A-Puro-SV40pA) flanked by 1 kb left and right arms homologous to the Chr3 pseudo-attP site (Fig. 1a) was employed as the template in the PCR reactions to generate linearized and biotinylated donors using non-biotinylated and biotinylated primers, respectively (Supplementary Table 1). To create the hrsACE2 donor (GenBank accession no. OQ579021), the hrsACE2 cassette with N-terminal signal peptide and His-tag was substituted with GFP in the plasmid donor by quick-change PCR and then linearized by biotinylated primers. Each PCR product was cleaned up using the PCR purification kit (Gene Transfer Pioneers. Iran).

Cell culture and transfection

The adherent CHO-K1 cells (INCB, Iran) were cultured with DMEM/F12 medium (Gibco, USA) supplemented with 10% (v/v) FBS (Gibco) and 1% penicillin/streptomycin (v/v) (Gibco), incubated in a 37 °C humidified incubator with 5% CO_2.

Lipofectamine 3000 reagent (Thermo fisher scientific, USA) was used for transfection based on the manufacturer's instructions. Briefly, CHO-K1 cells were seeded into the 24-well plate at the confluency of 90% and after overnight

incubation, co-transfected with 1 µg of Cas9/sgRNA or Cas9-mSA/sgRNA vector and corresponding donor at a mass ratio of 1:1.

Dot blot assay

The expression of active Cas9-mSA fusion was analyzed by dot blot immunoassay. The CHO-K1 cells were transfected by Cas9-mSA/sgRNA or Cas9/sgRNA (as a negative control) constructs. 24 h post-transfection, the cells were collected by centrifuge and resuspended in dH_2O. The cells were lyzed by sonication (5 parts with 50% amplitude), and the total protein concentration of each sample was estimated by Bradford assay. An equal quantity of samples was spotted on the nitrocellulose membrane in duplicate. The membrane was blocked by skim milk (3% w/v in PBS) for 2 h. The membrane was washed 4 times with the wash buffer (0.1% tween 20 v/v in PBS) and then incubated with double biotinylated DNA for 2 h. The washing was repeated, and the membrane was incubated with HRP-conjugated streptavidin (1:1000) for 1 h. After 5 times of washing, the membrane was developed by adding the diaminobenzidine (DAB) solution. The PCR product of the biotinylated donor was also blotted against HRP-conjugated streptavidin on the membrane to confirm biotinylation using DAB substrate.

Stable cell line generation and clonal selection

Forty-eight-hour post-transfection, cells were seeded into the 6-well plates at the confluency of approximately 30%; following overnight incubation, they were treated with 3.5 µg/ml of puromycin antibiotic (BioBasic, Canada) for nearly two weeks to generate the stable cell line. The genomic DNA of the resulting cell pools was extracted using the genomic DNA extraction kit (Gene Transfer Pioneers) and analyzed by 5′/3′ junction PCRs to amplify 5′ and 3′ junctions of the targeted locus and transgene using Taq DNA polymerase master mix (Ampliqon, Denmark) according to the following PCR program: 95 °C for 5 min, 30 cycles of 95 °C for 20 s, 57 °C for 30 s, 72 °C for 1 min 25 s (5′ junction) and 1 min 15 s (3′ junction), and final 72 °C for 10 min. The out-out PCR assay was also performed with primers designed for the outer regions of each of the homology arms to verify whole transgene integration by using the high-fidelity Super PCR master mix (YTA, Iran) based on the following PCR program: 95 °C for 5 min, 30 cycles of 95 °C for 20 s, 58 °C for 30 s, 72 °C for 4 min, and final 72 °C for 10 min. The PCR primers are listed in Supplementary Tables 1 and their binding sites are indicated in Fig. 1b and Fig. 5a.

Limiting dilution was performed to achieve individual clones from the cell pool. To do so, cells were detached by trypsin and diluted in antibiotic-free media with 15% FBS. Diluted cells were seeded at the density of 1 cell per each well of 96-well plates. Following approximately 14 days, each well, including single-cell clones, was transferred to the 24-well plates and then expanded to 6-well plates for further analysis.

Knock-in efficiency analysis

The genomic DNA was extracted from the recovered clones by the previously described protocol [18]. To describe briefly, cell pellets were provided by centrifuging and diluted in 20 µl of NaOH solution (0.2 M). After 10 min of incubation at 75 °C, cell lysates were obtained by adding 180 µl of tris-HCL buffer (0.04 M, pH 7.8). For the knock-in analysis, 2–3 µl of the resultant mixture was used as the template in 5′/3′ junction PCRs, and PCR products were visualized on 1% agarose gel.

The knock-in efficiency for each group was estimated by analyzing single-cell clones recovered from three individual stable cell pools. The efficiency was expressed based on the percentage of 5′/3′ junction PCR-positive clones, and the statistical significance was determined using a two-sided Fisher's exact test. Data with $p < 0.05$ were considered significant.

Copy number analysis and sequencing

The copy number of the inserted transgene (GFP reporter and hrsACE2) was determined by relative quantitative real-time PCR (qRT-PCR) using RealQ plus 2× master mix green (Ampliqon, Denmark) on 7500 ABI real-time PCR equipment (Thermo fisher scientific, USA). All qRT-PCRs were run in duplicate and no template controls (NTC) were included. Primers were designed for the target gene (GFP reporter or hrsACE2) and reference gene (Beta-actin) (Supplementary Table 1) and validated by melt curve analysis. Amplification was performed on the genomic DNA of each 5′/3′ junction PCR-positive clone with the following program: 1 cycle of 95 °C for 15 min, and 40 cycles of 95 °C for 20 s and 60 °C for 35 s. Copy number was estimated by the ΔΔCt method using a representative clone with a single copy of transgene as the calibrator.

To consider any possible mutations within the boundary sites of the locus and homology arms, 5′ and 3′ junction PCR amplicons of a selected clonal cell from each group were sent for Sanger sequencing using 5′ junction forward and 3′ junction reverse primers, respectively (Supplementary Table 1).

Stability analysis and growth profile determination

To evaluate GFP expression stability over a determined period of time, two single-copy clones from each group were randomly selected and passaged two times a week within

60 days without antibiotic pressure. The mean fluorescent intensity (MFI) values and the percentage of GFP-positive cells were calculated at passages 0, 10, and 20 by flow cytometry using BD FACSCalibur system (BD biosciences, USA). Data were analyzed by FlowJo software v10. Non-transfected CHO-K1 cells were used as a negative control to set the gating.

For growth profiles analysis, a selected clonal cell from each group and wild-type CHO-K1 cells were seeded at the density of 2.5×10^4 cells per each well of 12-well plates, and the viable cell density was measured daily for 8 days. Two wells were assessed for each day.

hrsACE2 targeted integration and protein expression analysis

Following co-transfection of mSA/sgRNA vector and biotinylated donor containing hrsACE2 cassette into the CHO-K1 cells, the stable cell pool was generated by puromycin, and its genomic DNA was analyzed by 5′/3′ junction PCRs under the following condition: 95 °C for 5 min, 30 cycles of 95 °C for 20 s, 57 °C for 30 s, 72 °C for 1 min 50 s (5′ junction) and 1 min 15 s (3′ junction), and final 72 °C for 10 min. The clonal selection was performed on the stable cell pool and knock-in efficiency was calculated as described above.

For protein expression analysis, the stable cell pool and a single-cell clone with a single copy of hrsACE2 were transferred to T75 flasks. When each cell line reached a confluency of approximately 100%, the cell culture supernatant (20 ml) containing the secreted protein was collected to be purified by the Ni-NTA column. After the column was equilibrated by phosphate buffer (50 mM NaH2PO4, 300 mM NaCl and 10 mM imidazole; pH 8), the supernatant was loaded onto it. The column was then washed with five-column volume of wash buffer (same as equilibration buffer) and proteins were eluted by elution buffer (50 mM NaH2PO4, 300 mM NaCl and 300 mM imidazole; pH 8). The purity of protein fractions was analyzed by 8% SDS-PAGE.

Estimation of the protein concentration

The concentration of the hrsACE2 protein purified from the cell culture supernatant of the stable cell pool and a clonal cell with a single copy of hrsACE2 was estimated by densitometric analysis (Quantity One, Bio-Rad, USA). The standard curve was plotted using the serially diluted bovine serum albumin (BSA) with known concentration.

Results

Evaluation of plasmid donor linearization on the knock-in rate in CHO-K1 cells

To examine whether switching from an HDR-based circular donor to the linearized donor—the transgene with flanked homology arms without any bacterial elements—could improve the knock-in efficiency in the CHO-K1 cell line, a transgene including expression cassettes of GFP reporter and puromycin selection marker was targeted to the Chr3 pseudo-attP site of the CHO-K1 genome (Fig. 1b). The CHO-K1 cells were co-transfected with Cas9/sgRNA vector and either linearized or circular donor. The targeted integration of the cassette was assessed on the genomic DNA of each cell pool 3 days post-transfection and following puromycin selection by 5′/3′ junction PCRs. Before antibiotic selection, only the cell pool developed by the linearized donor was 5′/3′ junction PCR-positive; however, stable cell pools of both donors revealed on-target cassette integration (Fig. 1c). The expected band size for 5′ and 3′ junction PCRs was 1.4 kb and 1.2 kb, respectively. Intact transgene integration was also confirmed for stable cell pools by out-out PCR using locus-specific forward and reverse primers, resulting in a predicted band size of 3.9 kb (Fig. 1d). Since a proportion of cells in stable cell pools showed either mono-allelic or random integration, the wild-type CHO band (2.2 kb) was also observed. The negative control (cell pool which was developed by transfecting an all-in-one vector without any donor) only exhibited a 2.2 kb band.

After verifying stable cell pools, single-cell cloning was performed using limiting dilution. The recovered single-cell clones from this step which expressed the GFP reporter were analyzed by 5′ and 3′ junction PCRs. Individual clones with a low growth rate or negative GFP expression were excluded for further analysis. The knock-in efficiency analysis was performed in triplicate; in each replicate, clonal cells obtained from verified stable pools were evaluated for targeted integration. Agarose gels of 5′/3′ junction PCR products of clonal cells are shown in Supplementary Figs. 2 and 3. According to the results, by using the linearized donor, an average of 35.3% of single-cell clones in each replicate were 5′/3′ junction PCR-positive. Meanwhile, in the control group (plasmid-based donor), an average of 21.9% of recovered clones were 5′/3′ junction PCR-positive (Table 1).

Investigating the effect of mSA-biotin tethering strategy on the knock-in rate

To reach a higher knock-in rate, we took the advantage of the tethering approach. According to the recent evidence,

Table 1 Summary of targeted knock-in

KI strategy	Total clones	5′ junction PCR-positive	3′ junction PCR-positive	5′/3′ junction PCR-positive	KI efficiency (%)
Donor linearization	49	19	28	17	34.69
	41	14	18	14	34.14
	43	18	21	16	37.2
mSA-biotin tethering	39	20	26	20	51.28
	45	27	29	23	51.11
	38	25	27	21	55.26
Control	36	13	17	8	22.22
	49	14	27	10	20.4
	39	9	16	9	23.07

reducing the distance between the donor template and the DSB site, once the break is induced by Cas9 endonuclease, may improve the efficiency of CRISPR-mediated knock-in. We benefited from the robust interaction between recently introduced monomeric streptavidin (mSA) and biotin. The interaction is created through the binding of the biotinylated donor to mSA-fused Cas9 (Fig. 2a). The mSA presents high stability and solubility compared to other previously engineered monomers. The sequence of mSA was genetically fused to the C-terminus of Cas9 with a flexible linker using the restriction-free cloning method. The Sanger sequencing verified the construct (Supplementary Fig. 4). Dot blot immunoassay was then performed following the transfection of Cas9-mSA to the cells. Our

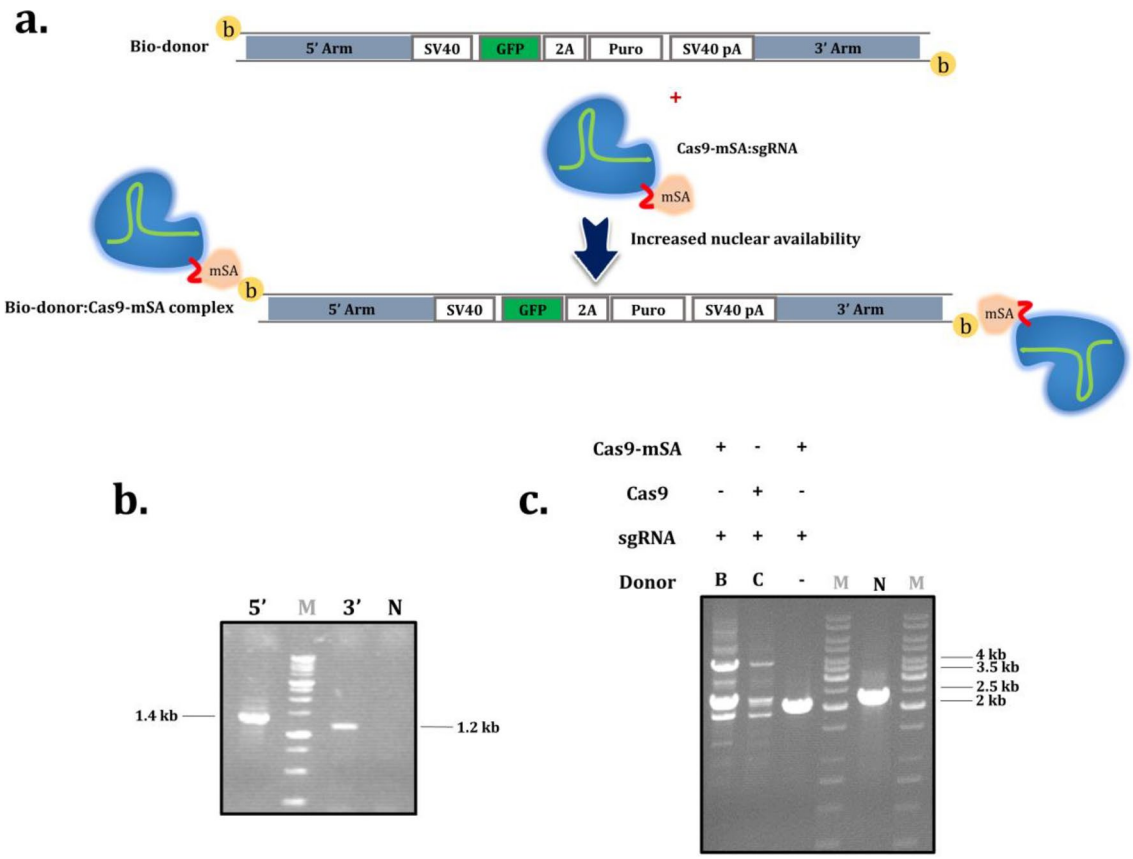

Fig. 2 The mSA-biotin tethering strategy. **a** Schematic overview of mSA-biotin tethering strategy. **b** 1% agarose gel of 5′ and 3′ junction PCR products of stable cell pools following co-transfection with the biotinylated donor and Cas9-mSA/sgRNA vector. M, 1 kb DNA marker, and N, negative control (parental CHO-K1 genome). **c** 1% agarose gel of out-out PCR products of stable cell pools. *C* circular donor, and *B* biotinylated donor

findings revealed the intact ability of the mSA moiety to bind to the biotin molecules (Supplementary Fig. 5a). The double biotinylated donor was amplified by the PCR and biotinylation was also confirmed by dot blot (Supplementary Fig. 5b).

Stable puromycin-resistant cell pools were generated following the transfection of cells with Cas9-mSA/sgRNA construct and biotinylated donor. The genomic DNA of stable cell pools was extracted and validated by 5′/3′ junction PCRs (Fig. 2d) and the out-out PCR assay (Fig. 2e). Then clonal selection was performed for three individual stable cell pools. The recovered single-cell clones which expressed the GFP protein were analyzed by 5′/3′ junction PCRs. The results as depicted in Supplementary Fig. 6 indicated that by employing the tethering approach, averagely 52.5% of single-cell clones were 5′ and 3′ junction PCR-positive (Table 1).

Copy number analysis

The relative copy number of GFP transgene for randomly selected clonal cells of each group was determined by quantitative PCR to evaluate the possible effects of donor linearization and tethering approach on copy number. Clonal cells with negative out-out PCR results were excluded for further assay (data not shown). Our results showed that all single-cell clones resulting from conventional plasmid-based donor harbored one copy of the GFP transgene (Fig. 3a). In the linearized donor group, with the exception of two clones (clones 41 and 56), other clones were single copy (Fig. 3b). By exploiting the tethering approach, 11 out of 15 clones were single copy but four clones (clones 67, 70, 85 and, 102) were random integrants (Fig. 3c).

Genotyping, expression stability, and growth profile analysis

The obtained clones were further characterized in terms of important parameters for industrial approaches. Clone 2 from the circular donor group, clone 37 from the linearized donor group, and clone 92 from the tethering group were nominated for sequencing. The results did not reveal any insertions or deletions (indels) at 5′/3′ junction sites for all clones (Fig. 4a). In addition, the cell growth profile of these selected clones and wild-type CHO-K1 cells was monitored for 8 days. As shown in Fig. 4b, no significant difference in growth profiles was concluded. Finally, we analyzed the long-term stability of the GFP expression for 20 passages. Clones 2 and 3 from the circular donor group, clones 37 and 57 from the linearized donor group, and clones 92 and 93 from the tethering group were selected; the MFI values and the percentage of GFP-expressing cells were determined by flow cytometry. During 20 passages, the percentage of GFP-positive cells was between 96.5 and 99.8% for all selected clones; however, there was a minor drop in GFP MFI for clones 3 and 37 from the circular and linearized donor groups, respectively (Fig. 4c and Supplementary Fig. 7).

Targeting the hrsACE2 cassette into the Chr3 pseudo-attp site and protein expression evaluation

To evaluate the potentiality of Chr3 pseudo-attP locus for the recombinant expression of a secretory protein, the hrsACE2 cassette was targeted to the locus using the established mSA-biotin tethering strategy (Fig. 5a). The stable cell pool was created and verified by 5′/3′ junction PCRs; as depicted in Fig. 5b, the desired bands of nearly 1.8 kb and 1.2 kb were visualized on 1% agarose gel for 5′ and 3′ junction PCRs, respectively. Following clonal selection, we first evaluated the efficiency of hrsACE2 integration by the 5′/3′ junction PCR analysis of single-cell clones. Although, compared to the GFP-harboring transgene, the cassette size increased

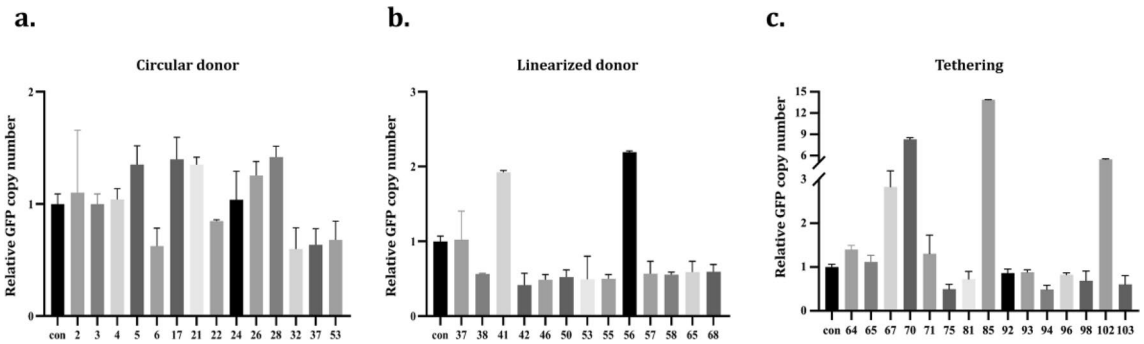

Fig. 3 Copy number analysis by real-time PCR. **a, b**, and **c** The relative copy number of GFP transgene was determined for the clonal cells of each group. A representative clone with a single copy of GFP was used as the calibrator. The error bars represent the standard deviations ($n=2$)

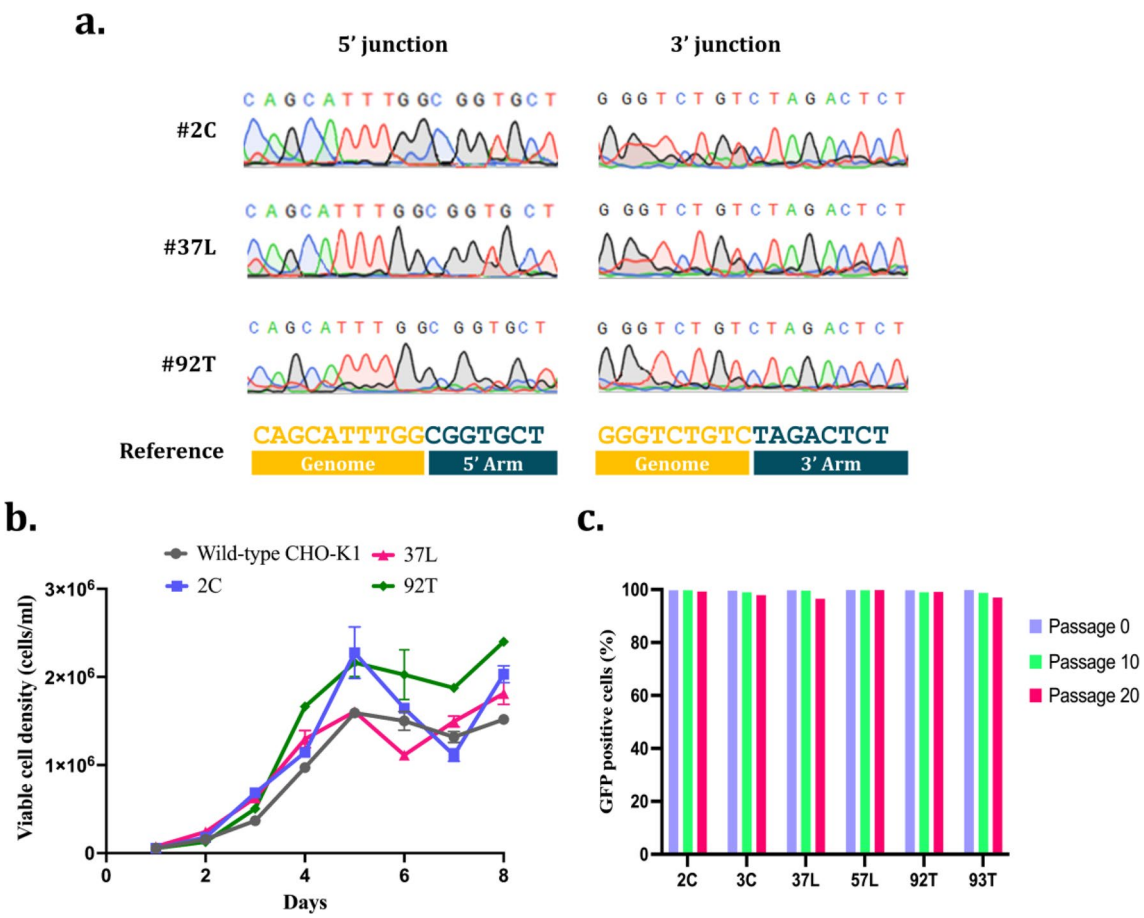

Fig. 4 Sequencing, growth profile, and expression stability of selected clones. **a** One single-copy clone from the circular donor group (2C), linearized donor group (37L), and tethering group (92T) was selected, and their 5′/3′ junction PCR amplicons were sent for sequencing. **b** Cell growth profiles of selected clones. The viable cell density of GFP-expressing clones and wild-type CHO-K1 was monitored for 8 days. **c** Two clones from the circular donor group (2C and 3C), linearized donor group (37L and 57L), and tethering group (92T and 93T) were selected and analyzed by flow cytometry. The percentage of GFP-positive cells during 20 passages was reported

from 1.8 kb to 3.4 kb, the knock-in rate was still high and approximately 42% of the recovered clones showed on-target integration (8 clones out of 19) (Supplementary Fig. 8). Afterward, for protein expression assessments, cell culture supernatant (20 ml) from the stable cell pool and a clonal cell with a single copy of hrsACE2 (clone 2) were collected separately and passed through Ni-NTA Sepharose resin. The obtained samples were analyzed on 8% SDS-PAGE gel (Fig. 5c). The glycosylated hrsACE2 protein showed a band size of approximately 120 kDa on the gel. The concentration of hrsACE2 protein purified from the cell culture supernatant of the targeted integration cell pool and clone 2 was calculated by densitometric assay and compared with the hrsACE2 purified from the supernatant of the random integration cell pool (Supplementary Fig. 9). By considering the fraction and initial sample volumes, productivity was calculated to be ~0.8 µg/ml for the targeted integration clone 2, ~0.6 µg/ml for the targeted integration cell pool and ~0.3 µg/ml for the random integration cell pool.

Discussion

Although the conventional method used to develop the recombinant CHO cell is mainly based on random integration, reaching the desired clone with a comparable protein expression level could be shortened by applying the CRISPR/Cas9-mediated integration into the genome's hot spot(s). Bypassing the time-limiting steps of clonal selection primarily depends on (i) the efficiency of targeted integration and (ii) the potentiality of the target site for the stable and high-level expression. However, the low rate of HDR in CHO cells [19] and the necessity for finding the hot spot sites are still a challenge in rCHO cell line development with this technology. Here, we tried to observe the challenges by employing two donor design approaches to target the Chr3 pseudo-attP site to increase the knock-in efficiency and establish a safe harbor site in CHO cells.

Plasmid-based donors flank the gene of interest with usually long homology arms and execute the gene integration

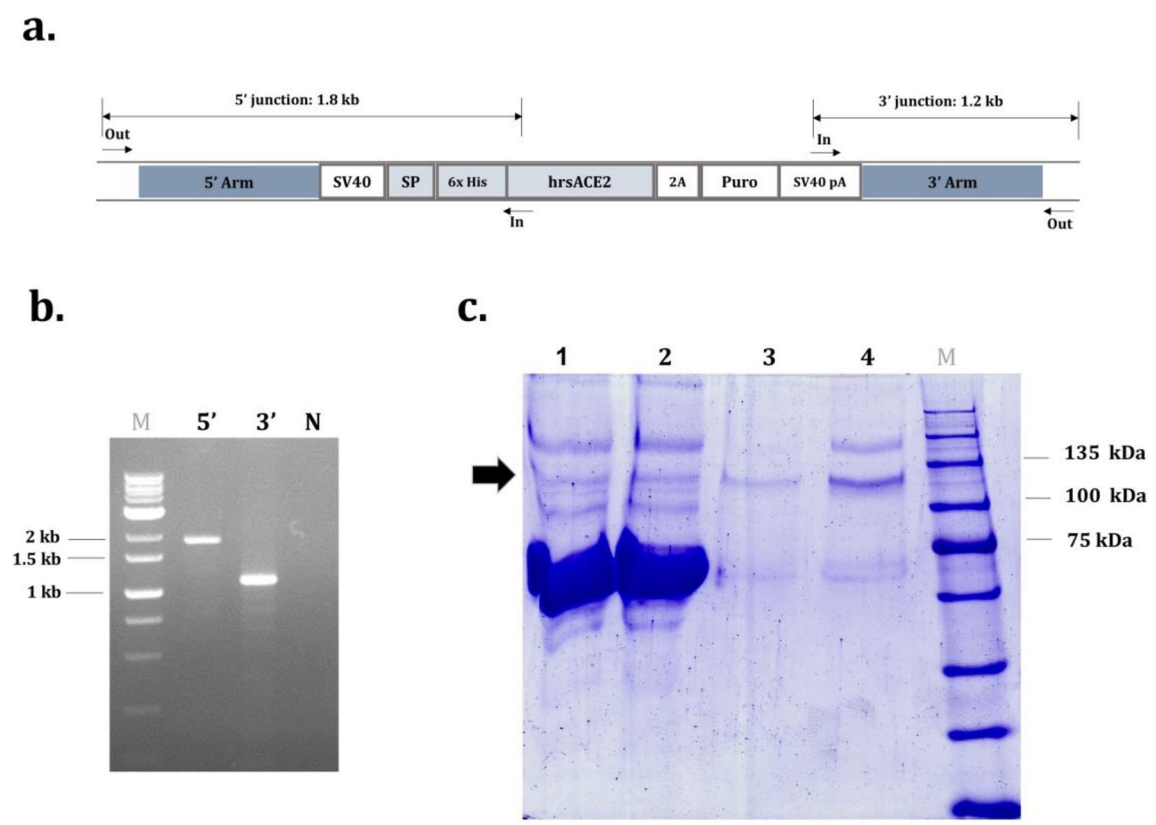

Fig. 5 Targeting the hrsACE2 cassette and protein expression analysis. **a** Schematic overview of the hrsACE2 targeted Chr3 pseudo-attP locus. SP, signal peptide. The annealing sites of 5′/3′ junction PCR and out-out PCR primers were indicated. **b** 1% agarose gel of 5′/3′ junction PCR products of stable cell pool following hrsACE2 integration. The expected band size for 5′ and 3′ junction PCR products was 1.8 kb and 1.2 kb, respectively. *M* 1 kb DNA marker, and *N* negative control (parental CHO-K1 genome). **c** 8% SDS-PAGE gel result of hrsACE2 purification from the cell culture supernatant of targeted integration clone 2. Lane 1; initial sample, lane 2; flow-through, lane 3; elution 2, lane 4; elution 1. M, protein size marker

by the HDR mechanism precisely, albeit with low efficiency [6]. As an alternative perspective to increase the KI efficiency in CHO cells, conventional donors were converted to PCR-amplified linearized ones. This accessible method has recently achieved promising results in different cell types [11, 12, 20]. We, therefore, assessed this methodology in the CHO-K1 cells to target the 1.8 kb transgene into the Chr3 pseudo-attP site. By cleaving the plasmid donor out of the homology arms, the KI efficiency was improved to about 35.3%, as compared to the control group with the KI efficiency of 21.9%. However, this improvement was not statistically significant ($p > 0.05$). These findings could be correlated with the previous results in CHO cells with a similar donor, but using zinc finger nuclease genome editing tool [21]. As described there, this can be due to the susceptibility of the linearized donor to the various nucleases in the cell owing to its free termini.

To achieve better KI efficiency, we aimed to provide sufficient repair templates in the vicinity of DSB site and decrease the distance between the DSB and donor. This was accomplished by tethering the donor to the Cas9 endonuclease. Most previous tethering studies have applied ssODN donors for the insertion of a few nucleotides [14, 22]. Here, we used the above-mentioned linearized double-stranded donor, which was biotinylated by PCR and tethered to the monomeric streptavidin-fused Cas9. In the current research, the monomeric streptavidin with the highest affinity to the biotin among non-tetrameric streptavidins was used [17]. Its monomeric structure enabled the formation of fusion in our multielement construct (Fig. 1b). We added mSA to the C-terminus of the Cas9 via a linker by the restriction-free cloning. This method does not add any extra nucleotides to the N- or C-terminal regions of Cas9. Dot blot analysis verified that mSA in the fusion form (Cas9-mSA) kept its affinity to the biotin. The KI efficiency of the tethering group was determined to be 52.5% on average by precisely targeting transgene into the CHO cell hot spot, representing a 2.4-fold increase over the conventional circular donor ($p < 0.05$).

We also validated our mSA-biotin tethering approach by assessing its efficiency to insert a larger transgene: a 3.4 kb hrsACE2-harboring cassette. The knock-in rate reached to

42%, which seemed to be still high and comparable to the GFP reporter integration; however, this minor variation in efficiency could be as a result of the size dependency of targeting efficiency [20]. In addition to CHO cells, this approach might be promising for cells with a low rate of knock-in, including non-dividing or quiescent cells [23].

It was assumed that when the linearized template is presented, homology-independent targeted integration (or HITI) at induced dsDNA break could play a role in the overall knock-in rate, though trivial. As a result, it may cause random indels at the locus and transgene boundary sites. However, our sequence analysis did not reveal any indel mutation in these sites. Furthermore, to analyze the effects of donor linearization and the mSA-biotin tethering approach on the rate of the non-HDR insertions of transgene at the endogenous breaks (i.e., random integration), the copy number of selected clones showing on-target integration was determined by the real-time PCR. Based on the findings, despite plasmid-based donor, linearization of the donor by either PCR amplification or use of the tethering approach induced few random integrants among single-cell clones; however, most of them were still single copy (84 and 73% of clonal cells, respectively). Therefore, owing to the high targeted efficiency, especially in the tethering approach, it could be hardly considered as a limitation since most of the achieved clones showed no random integration event.

Since CHO cells are the preferred mammalian host for producing the recombinant protein with complex post-translational modifications, at the next step, we evaluated the hrsACE2 expression as a glycosylated secretory protein following targeting to the Chr3 pseudo-attP site. Though the productivity of the targeted integration clone was far from the industrial-developed cell line, the obtained findings showed that the productivity of the targeted integration cell pool was increased by 2-fold, as compared to the random integration cell pool. Expectedly, by moving toward industrial conditions with regard to the bioprocess mode and producing CHO, the productivity could be increased remarkably. Moreover, while most nominated clones maintained the GFP expression during multiple passages, two clones (clone 3 from circular donor group and clone 37 from linearized donor group) exhibited a slight drop in MFI in passage 20. It seems that, this phenomenon rather than the gene loss could be as result of epigenetic gene silencing such as promoter methylation [24, 25].

Overall, in the present study, we have devised and compared different strategies to increase the knock-in rate into the Chr3 pseudo-attP site of the CHO genome as a potential candidate for sustained transgene expression, aiming to promote rCHO cell line development. As a growing body of the recently published literature from the biopharmaceutical companies shows their interest in site-specific integrations [26, 27], we hope that the present findings would be applied in the CHO cell line development as the prevailing mammalian host for the production of biopharmaceuticals.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1007/s11033-023-08529-8.

Acknowledgements Not applicable.

Author contributions MHK contributed to conceptualization-supporting, data curation-lead, formal analysis-lead, investigation-lead, methodology-lead, validation-lead, visualization-lead, and writing—original draft-lead. BR contributed to formal analysis-supporting, investigation-supporting, and methodology-supporting**.** HZ contributed to formal analysis-supporting and investigation-supporting. FB contributed to investigation-supporting. MAM contributed to conceptualization-supporting, funding acquisition-lead, resources-lead, supervision-lead. FD contributed to conceptualization-lead, project administration-lead, resources-supporting, supervision-lead, and writing—review & editing-lead.

Funding This research was part of a Ph.D. dissertation supported by the Tehran University of Medical Sciences, Tehran, Iran (Grant No. 99-1-148-47958).

Data availability The flow cytometry raw data (FCS files) are openly available on FlowRepository.org (FR-FCM-Z645). All other data are included in the manuscript and its supplementary information file.

Declarations

Competing interest The authors declare that they have no competing interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical approval This article does not contain any studies with human participants performed by any of the authors

Consent for publication All authors approved the final version of the manuscript and consent for publication.

References

1. Walsh G, Walsh E (2022) Biopharmaceutical benchmarks. Nat Biotechnol 40:1722–1760
2. Tihanyi B, Nyitray L (2021) Recent advances in CHO cell line development for recombinant protein production. Drug Discov Today Technol. https://doi.org/10.1016/j.ddtec.2021.02.003
3. Shin SW, Lee JS (2020) CHO cell line development and engineering via site-specific integration: challenges and opportunities. Biotechnol Bioprocess Eng. https://doi.org/10.1007/s12257-020-0093-7
4. Lee JS, Kildegaard HF, Lewis NE, Lee GM (2019) Mitigating clonal variation in recombinant mammalian cell lines. Trends Biotechnol 37(9):931–942
5. Banan M (2020) Recent advances in CRISPR/Cas9-mediated knock-ins in mammalian cells. J Biotechnol 308:1–9
6. Yeh CD, Richardson CD, Corn JE (2019) Advances in genome editing through control of DNA repair pathways. Nat Cell Biol 21(12):1468–1478

7. Riesenberg S, Maricic T (2018) Targeting repair pathways with small molecules increases precise genome editing in pluripotent stem cells. Nat Commun 9(1):2164
8. Lin S, Staahl BT, Alla RK, Doudna JA (2014) Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. elife 3:e04766
9. Yao X, Wang X, Hu X, Liu Z, Liu J, Zhou H et al (2017) Homology-mediated end joining-based targeted integration using CRISPR/Cas9. Cell Res 27(6):801–814
10. Zhang J-P, Li X-L, Li G-H, Chen W, Arakaki C, Botimer GD et al (2017) Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. Genome Biol 18(1):1–18
11. Yao X, Zhang M, Wang X, Ying W, Hu X, Dai P et al (2018) Tild-CRISPR allows for efficient and precise gene knockin in mouse and human cells. Dev Cell 45(4):526-36e5
12. Roth TL, Puig-Saus C, Yu R, Shifrut E, Carnevale J, Li PJ et al (2018) Reprogramming human T cell function and specificity with non-viral genome targeting. Nature 559(7714):405–409
13. Savic N, Ringnalda FC, Lindsay H, Berk C, Bargsten K, Li Y et al (2018) Covalent linkage of the DNA repair template to the CRISPR-Cas9 nuclease enhances homology-directed repair. elife 7:e33761
14. Aird EJ, Lovendahl KN, St. Martin A, Harris RS, Gordon WR (2018) Increasing Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template. Commun Biology 1(1):54
15. Bi Y, Hua Z, Ren H, Zhang L, Xiao H, Liu X et al (2018) ΦC31 integrase-mediated isolation and characterization of novel safe harbors for transgene expression in the pig genome. Int J Mol Sci. https://doi.org/10.3390/ijms19010149
16. Damavandi N, Raigani M, Joudaki A, Davami F, Zeinali S (2017) Rapid characterization of the CHO platform cell line and identification of pseudo attP sites for PhiC31 integrase. Protein Expr Purif 140:60–64
17. Lim KH, Huang H, Pralle A, Park S (2013) Stable, high-affinity streptavidin monomer for protein labeling and monovalent biotin detection. Biotechnol Bioeng 110(1):57–67
18. Freire CAM (2017) Genome editing via CRISPR/Cas9 targeted integration in CHO cells. MA Thesis Instituto Superior Técnico, Lisbon, Portugal
19. Cristea S, Freyvert Y, Santiago Y, Holmes MC, Urnov FD, Gregory PD et al (2013) In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration. Biotechnol Bioeng 110(3):871–880
20. Paix A, Folkmann A, Goldman DH, Kulaga H, Grzelak MJ, Rasoloson D et al (2017) Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks. Proc Nat Acad Sci 114(50):E10745–E10754
21. Orlando SJ, Santiago Y, DeKelver RC, Freyvert Y, Boydston EA, Moehle EA et al (2010) Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. Nucleic Acids Res 38(15):e152
22. Lee K, Mackley VA, Rao A, Chong AT, Dewitt MA, Corn JE et al (2017) Synthetically modified guide RNA and donor DNA are a versatile platform for CRISPR-Cas9 engineering. elife 6:e25312
23. Genovese P, Schiroli G, Escobar G, Di Tomaso T, Firrito C, Calabria A et al (2014) Targeted genome editing in human repopulating haematopoietic stem cells. Nature 510(7504):235–240
24. O'Callaghan PM, Racher AJ (2015) Building a cell culture process with stable foundations: searching for certainty in an uncertain world. In: Al-Rubeai M (ed) Animal cell culture. Springer, Cham, pp 373–406
25. Kim M, O'Callaghan PM, Droms KA, James DC (2011) A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies. Biotechnol Bioeng 108(10):2434–2446
26. Inniss MC, Bandara K, Jusiak B, Lu TK, Weiss R, Wroblewska L et al (2017) A novel Bxb1 integrase RMCE system for high fidelity site-specific integration of mAb expression cassette in CHO cells. Biotechnol Bioeng 114(8):1837–1846
27. Carver J, Ng D, Zhou M, Ko P, Zhan D, Yim M et al (2020) Maximizing antibody production in a targeted integration host by optimization of subunit gene dosage and position. Biotechnol Prog 36(4):e2967