

Multicopy Targeted Integration for Accelerated Development of High-Producing Chinese Hamster Ovary Cells

Daria Sergeeva, Gyun Min Lee, Lars Keld Nielsen, and Lise Marie Grav*



Cite This: *ACS Synth. Biol.* 2020, 9, 2546–2561



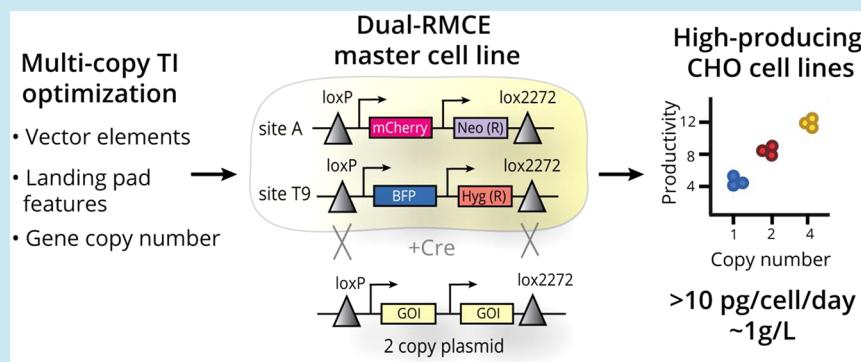
Read Online

ACCESS |

Metrics & More

Article Recommendations

Supporting Information



ABSTRACT: The ever-growing biopharmaceutical industry relies on the production of recombinant therapeutic proteins in Chinese hamster ovary (CHO) cells. The traditional timelines of CHO cell line development can be significantly shortened by the use of targeted gene integration (TI). However, broad use of TI has been limited due to the low specific productivity (q_p) of TI-generated clones. Here, we show a 10-fold increase in the q_p of therapeutic glycoproteins in CHO cells through the development and optimization of a multicity TI method. We used a recombinase-mediated cassette exchange (RMCE) platform to investigate the effect of gene copy number, 5' and 3' gene regulatory elements, and landing pad features on q_p . We evaluated the limitations of multicity expression from a single genomic site as well as multiple genomic sites and found that a transcriptional bottleneck can appear with an increase in gene dosage. We created a dual-RMCE system for simultaneous multicity TI in two genomic sites and generated isogenic high-producing clones with q_p of 12–14 pg/cell/day and product titer close to 1 g/L in fed-batch. Our study provides an extensive characterization of the multicity TI method and elucidates the relationship between gene copy number and protein expression in mammalian cells. Moreover, it demonstrates that TI-generated CHO cells are capable of producing therapeutic proteins at levels that can support their industrial manufacture.

KEYWORDS: Chinese hamster ovary cells, cell line development, recombinase-mediated cassette exchange, targeted integration, specific productivity

The foundation of the \$200 billion biopharmaceutical industry continues to be the production of complex therapeutic proteins such as monoclonal antibodies (mAbs), fusion proteins, hormones, enzymes, and coagulation factors in mammalian cells, with Chinese hamster ovary (CHO) cells being the leading production host.¹ For each therapeutic product, a stable cell line with a high product titer and a desirable protein quality must be developed.² The conventional cell line development process is long and unpredictable, spanning several months and involving the screening of several hundred cell clones for high productivity.³ To accelerate cell line development, more efficient and predictable methods for generating high-producing CHO cell lines are required.

For the past 30 years, random gene integration has been the method of choice for developing protein-producing CHO cell lines. Using this method, the gene of interest (GOI) is integrated into randomly occurring double-stranded breaks in

the genomic DNA, resulting in GOI integration into variable sites with an uncontrolled number of copies. This approach generates a set of clones with diverse and unpredictable phenotypes attributed to clonal variation,⁴ which necessitates tedious screening of clones and bioprocess reoptimization for the chosen clone.

Targeted gene integration (TI) has the potential to significantly reduce clonal variation, thus accelerating CHO cell line development.^{4,5} Integration of the transgene into a

Received: June 17, 2020

Published: August 24, 2020



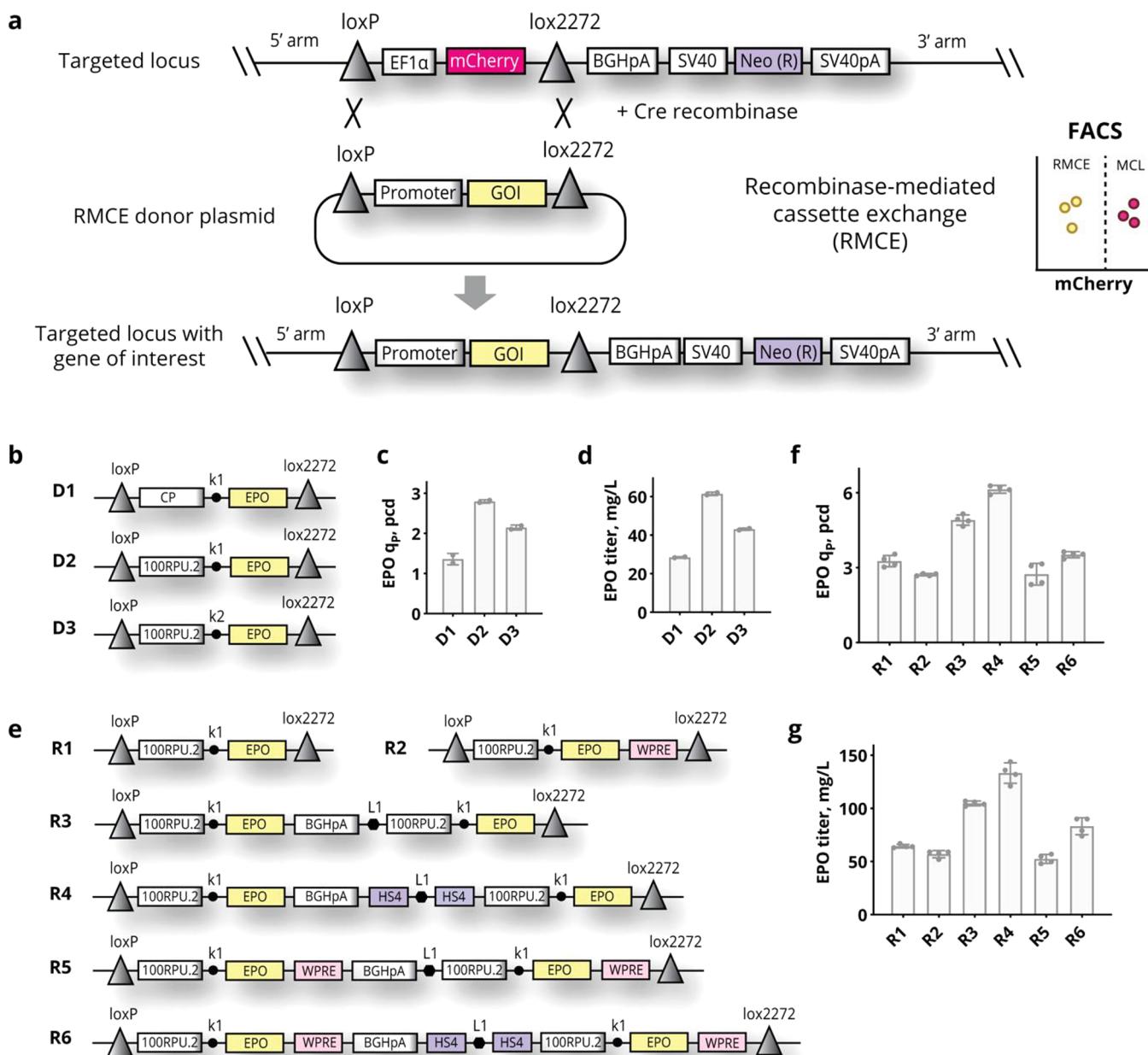


Figure 1. RMCE platform and screening of gene regulatory elements for increased protein production upon multicopy targeted integration. (a) Overview of the RMCE strategy. First, a landing pad with poly(A) trap is integrated into the specific genomic locus using CRISPR/Cas9-mediated targeted integration, resulting in the creation of a master cell line (MCL). Then, the MCL is transfected with a RMCE donor plasmid together with Cre recombinase, leading to the site-specific recombination between lox sites and exchange of EF1 α -mCherry cassette to promoter-GOI cassette in the targeted locus. mCherry-negative cells with recombinase-mediated integration are isolated using FACS. (b) Design of RMCE donor plasmids D1–D3 used for screening of 5' regulatory elements (promoters (CP, 100RPU.2) and Kozak sequences (k1, k2)). Cell pools D1–D3 were grown in batch culture (2 technical replicates); EPO specific productivity (c) and EPO titer on day 5 (d) are shown. (e) Design of RMCE donor plasmids R1–R6 used for screening of 3' regulatory elements (WPRE and HS4 insulator). Two clones for each cell design were grown in batch culture (2 technical replicates per clone); EPO specific productivity (f) and EPO titer on day 5 (g) are shown.

preselected genomic site enables predictable generation of isogenic cell lines with uniform phenotypes and consistent expression.^{6,7} Since the first reports of TI in mammalian cells in the late 1980s^{8,9} and early 1990s,^{10–12} several methods for targeted gene integration have been developed, which utilize site-specific recombinases such as Cre, Flp, and Bxb1 for single-site recombination or recombinase-mediated cassette exchange (RMCE) and programmable nucleases such as CRISPR/Cas9 (clustered regularly interspaced short palindromic repeat/CRISPR-associated protein 9).^{13–15} However,

while TI has been used in special applications in CHO,¹⁶ it has not been used widely in industry, largely due to low specific productivity (q_p) of TI-generated clones, caused by single-copy GOI integration, and the challenge to identify highly active transcriptional sites (hotspots) in the genome.

In recent years, the interest in TI has been rising in the biopharmaceutical industry, with several companies publishing reports on mAb production using RMCE systems. In 2013, Genentech established a Cre-based RMCE platform for mAb expression by the screening of transcriptionally active sites in

the CHO genome.¹⁷ Insertion of a single expression cassette containing heavy and light chain genes led to the generation of stable cell lines for five different mAbs with reproducible q_p of 3–4 pg/cell/day (pcd). Insertion of two mAb cassettes into the same genomic site doubled specific productivity with a maximum of 10 pcd achieved. Similarly, Pfizer reported the development of Flp-based RMCE¹⁸ and Bxb1-based RMCE¹⁹ platforms for mAb expression, reaching q_p of 3 pcd for single-copy mAb insertion. In 2020, Genentech showed that mAb expression can reach q_p of 20–50 pcd in the RMCE system by the integration of multiple copies of heavy and light chains into a single genomic site.²⁰ However, the commercial nature means that these studies fail to provide details on the design of the expression vectors and do not report any limitations of the methods, making them difficult to reproduce. Moreover, while these and other studies²¹ indicate a positive correlation between gene copy number and mAb production in TI-generated clones, this relationship can be obscured by the ratio of expression between light and heavy chains.

In this study, we focused on increasing q_p upon TI for the production of therapeutic proteins not limited to mAbs. We developed and extensively characterized a multicopy TI method for the production of biopharmaceutical proteins and used it to study the relationship between gene copy number and protein production to reveal if any expression bottlenecks appear with an increase in gene dosage. We selected two industrially relevant single-chain glycoproteins: the hormone erythropoietin (EPO) and the Fc-fusion protein etanercept (ETN), which have different biochemical properties and complex post-translational modifications, with ETN²² generally considered as a difficult-to-express protein. Compared to mAb-producing CHO cells generated by random integration that can reach q_p of 20–40 pcd,²³ EPO- and ETN-producing CHO cells generated by random integration usually have a lower q_p in the range 1–10 pcd.^{24,25}

In our previous works, we established a TI method for CHO cell line development based on the Cre-mediated RMCE, where EPO specific productivity of 1–3 pcd was achieved from a single-copy integration.^{7,26} We used this RMCE system as a starting point to develop and optimize the multicopy TI method described in the current paper. Here, we investigated the effect of different gene regulatory elements in the RMCE donor vector, landing pad features and gene copy number in order to increase q_p . We evaluated the limitations of multicopy expression from a single genomic site as well as multiple genomic sites and showed that a transcriptional bottleneck can appear when the copy number is increased. We developed a dual-RMCE master cell line with two landing pads for simultaneous multicopy TI in two genomic sites. With this system, we were able to site-specifically introduce up to four copies of GOI into the genome simultaneously, leading to the generation of high-producing EPO and ETN clones with q_p of 12–14 pcd and titer of nearly 1 g/L in shake flask cultivation, reaching and excelling q_p of high-producing clones generated by random integration. To support further use of multicopy TI by the scientific community, we made our vectors for the multicopy RMCE system available from Addgene.

RESULTS

Optimal 5' and 3' Gene Regulatory Elements for Increased Protein Production upon Multicopy Targeted Integration. To increase q_p upon targeted integration in CHO cells, we first focused on finding strong gene regulatory

elements that can support high protein expression in the RMCE system. As a platform for screening, we used the previously described CHO-S derived MCL AD1 with a landing pad for poly(A) trap-based RMCE integrated into the genomic site A, which is a proven safe harbor site for TI supporting high and stable protein expression.²⁶ The landing pad in MCL AD1 has a mCherry fluorescent marker expressed from an EF1 α promoter surrounded by loxP and lox2272 recombination sites and a BGH poly(A) tail outside the lox sites, thus allowing a RMCE of EF1 α -mCherry to GOI with different 5' regulatory sequences (Figure 1a). To find optimal 5' genetic elements, we designed RMCE donor plasmids D1–D3 with two different promoters (CP and 100RPU.2) and two Kozak sequences (k1 and k2) driving expression of the EPO gene (Figure 1b). CP is a composite promoter consisting of mouse CMV enhancer, minimal human EF1 α promoter, and human T cell leukemia virus 1 untranslated region. CP was previously found to be stronger than CMV, SV40, and EF1 α promoters when integrated into site A in the RMCE system.²⁶ The 100RPU.2 promoter is a *de novo*-designed synthetic promoter composed of CHO-specific transcription factor regulatory elements, and it showed a twofold higher expression level compared to the CMV promoter in transient transfection and stable CHO-K1 pools generated by random integration.²⁷ Kozak sequence k1 is a consensus mammalian translation initiation site (TIS) CGCCACC.²⁸ Kozak sequence k2 is a TCGGTC TIS found to be the strongest in yeast and showed almost 50% higher expression of eGFP in CHO cells when compared to the consensus Kozak.²⁹ The combination of CP and k1 showed the best performance in our poly(A)-trap RMCE system previously.²⁶ We cotransfected Cre recombinase together with each D1–D3 donor plasmid into MCL AD1 and bulk sorted mCherry-negative cells 6 days after transfection. The cell pools were verified to have the desired one-copy insert by genomic PCR of the target region and copy number analysis by digital PCR (dPCR) (Supporting Figure 1a,b). Generated cell pools were grown in batch culture, and EPO q_p , titer, and mRNA were evaluated (Figure 1c,d, Supporting Figure 1e). EPO q_p and titer of D2 cells with 100RPU.2 promoter were doubled (2.8 pcd, 61 mg/L) compared to D1 cells with CP promoter (1.4 pcd, 28 mg/L) (multiple testing, adj. p = 0.001). Kozak sequence k2 had about 30% lower EPO expression in D3 cells (2.1 pcd, 43 mg/L) compared to Kozak k1 in D2 cells (multiple testing, adj. p = 0.01), which was attributed to lower mRNA levels. Hence, we chose a combination of 100RPU.2 and k1 as optimal 5' gene regulatory elements for GOI expression in the RMCE system.

To evaluate 3' gene regulatory elements for multicopy RMCE integration, we designed RMCE donor plasmids R1–R6 (Figure 1e), which had one copy (R1, R2) or two copies (R3–R6) of the expression cassette 100RPU.2_k1_EPO. We assessed the effects of Woodchuck hepatitis virus post-transcriptional regulation element (WPRE) and chicken β -globin insulator HS4. WPRE has been known to affect the transport of mRNA and increase transgene expression in viral vectors,³⁰ and enhance mAb production in HEK 293E cells.³¹ The chicken HS4 insulator can act as a barrier to block promoter interference,³² can prevent DNA methylation and maintain open chromatin state,³³ and has been used for the construction of multigenic plasmids for expression in mammalian cells.³⁴ WPRE was inserted in the 3' untranslated region (UTR) of the EPO gene between the stop codon and the BGH poly(A) tail in plasmids R2, R5, and R6. Two 250-bp

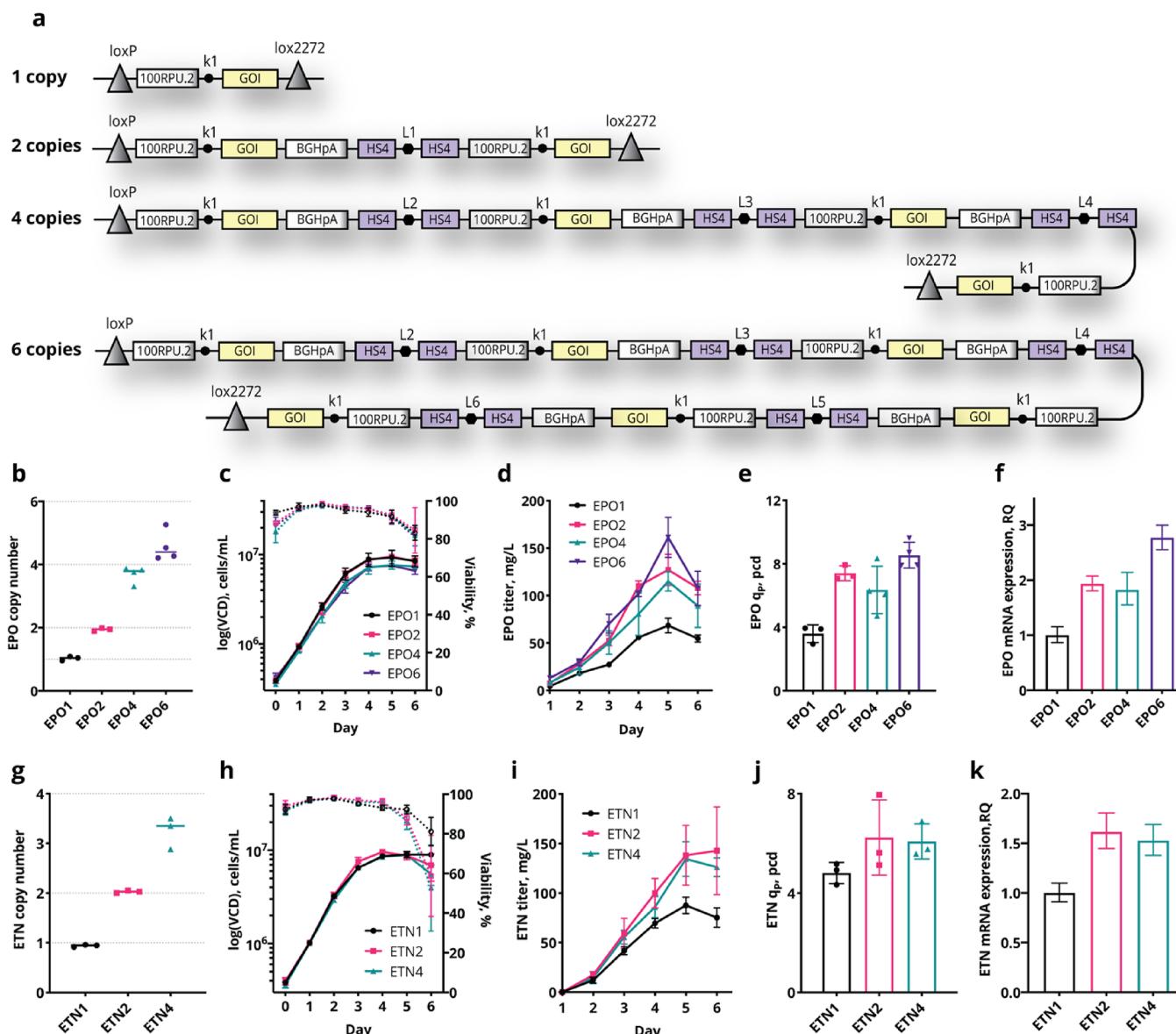


Figure 2. Targeted integration of one, two, four, and six copies of EPO and ETN into a single genomic site. (a) Schematic representation of RMCE donor plasmids with up to six GOI expression units. Each expression unit was flanked by a specific USER linker (L1–L6). (b–f) Characterization of EPO1, EPO2, EPO4, and EPO6 clones created by targeted integration of RMCE donor plasmids into AD1 master cell line (MCL). Three (EPO1, EPO2) or four (EPO4, EPO6) clones for each cell design were grown in batch culture. (b) Copy number analysis of EPO1–EPO6 clones, showing median value for biological replicates. (c) Viable cell densities (VCD) and viability of EPO1–EPO6 clones, showing mean with SD for biological replicates. (d) EPO titer, showing mean with SD for biological replicates. (e) EPO specific productivity, showing mean with SD for biological replicates. (f) EPO mRNA expression, showing mean and SD for biological replicates. (g–k) Characterization of ETN1, ETN2, and ETN4 clones created by targeted integration of RMCE donor plasmids into MCL AD1. Three clones for each cell design were grown in batch culture. (g) Copy number analysis of ETN1–ETN4 clones, showing median value for biological replicates. (h) VCD and viability of ETN1–ETN4 clones, showing mean and SD for biological replicates. (i) ETN titer, showing mean and SD for biological replicates. (j) ETN specific productivity, showing mean and SD for biological replicates. (k) ETN mRNA expression, showing mean and SD for biological replicates.

core HS4 insulators were placed between poly(A) tail and 100RPU.2 promoter in plasmids R4 and R6 to separate two EPO expression cassettes. Each RMCE donor plasmid R1–R6 was cotransfected with Cre recombinase into MCL AD1; mCherry-negative cells were single-cell sorted by FACS, and clones were verified by genomic PCR of the target region and copy number analysis (Supporting Figure 1c,d). Two clones for each cell design R1–R6 were cultivated in batch mode, and EPO q_p , titer, and mRNA were evaluated (Figure 1f,g, Supporting Figure 1f). We observed that the introduction of WPRE in each case (i.e., R2 vs R1, R5 vs R3, and R6 vs R4)

had a significant negative effect on specific productivity (multiple testing, adj. $p < 10^{-10}$). The HS4 insulator significantly enhanced specific productivity in two copy expression vectors (i.e., R4 vs R3 and R6 vs R5) (multiple testing, adj. $p < 3 \times 10^{-9}$). The effect of HS4 and WPRE was also reflected in changes of EPO mRNA levels (Supporting Figure 1f). The best two copy plasmid R4 (with insulator and no WPRE) enabled cells to almost double (1.9-fold increase) EPO expression compared to the best single copy plasmid R1. Thus, the use of an insulator was crucial to block the

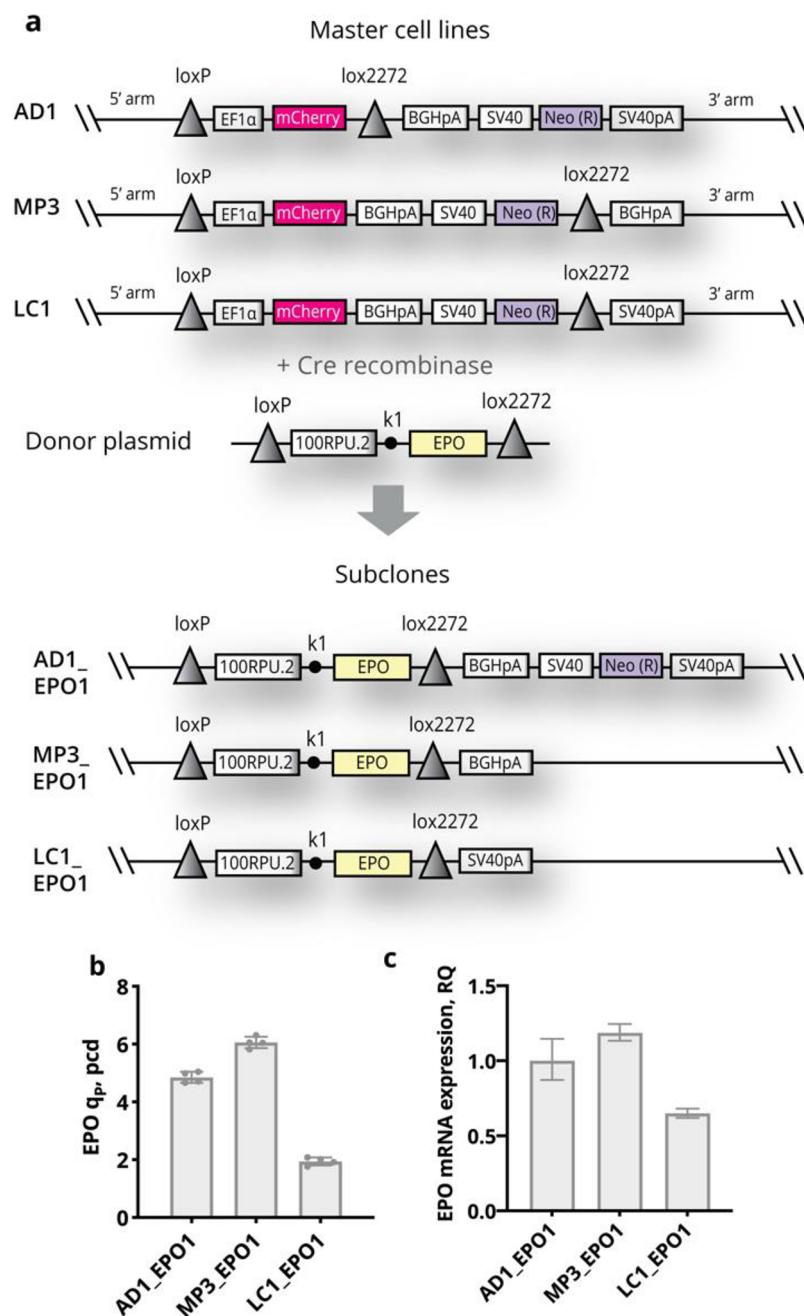


Figure 3. Optimization of landing pad features: poly(A) tail and antibiotic selection marker NeoR. (a) Three master cell lines (AD1, MP3, and LC1) were created by CRISPR/Cas9-mediated targeted integration of landing pads into genomic site A. MCL AD1 had NeoR cassette outside the loxP and lox2272 sites, with BGH poly(A) outside lox sites. MCL MP3 had NeoR expression cassette moved inside loxP and lox2272 sites, with BGH poly(A) outside lox sites. MCL LC1 had NeoR expression cassette moved inside loxP and lox2272 sites, with SV40 poly(A) outside lox sites. The MCLs were used to generate subclones using a single copy (EPO1) RMCE donor plasmid. Two EPO-producing subclones per MCL were grown in batch culture (two technical replicates). EPO specific productivity (b) and EPO mRNA expression (c) are shown with mean and SD for biological replicates.

interference between expression cassettes, and plasmid design R4 was optimal for building multicopy plasmids.

Targeted Integration of up to Six Gene Copies in a Single Genomic Site. To further increase recombinant protein expression upon TI, we constructed multicopy plasmids expressing four and six copies of GOI for RMCE-based integration into a single genomic site. To assemble such long and repetitive constructs, we used USER cloning, a flexible method for the plasmid assembly of up to seven DNA fragments in one reaction enabled by single-stranded DNA

overhangs (linkers).³⁵ USER linkers flanked by unique restriction sites (L1–L6) were placed between each GOI expression cassette to ease the plasmid and clone verification (Figure 2a). The four- and six-copy plasmids were confirmed by sequencing and restriction analysis (Supporting Figure 2), no undesirable recombination events between repetitive elements were observed. We integrated the multicopy plasmids expressing one (EPO1), two (EPO2), four (EPO4), and six (EPO6) copies of erythropoietin into MCL AD1 using the described RMCE method. The RMCE efficiency decreased

from 1.8% (EPO1) to 0.06% (EPO6) as the size of the RMCE insert increased (Supporting Table 1), but it was high enough to select desired clones. The single-cell sorted cells were expanded and verified by genomic PCR (Supporting Figure 3a-f,j) and copy number analysis (Figure 2b). While most of EPO1 and EPO2 mCherry-negative clones had correctly inserted expression cassettes in site A, we faced some difficulties verifying the integration of EPO4 and EPO6 plasmids in MCL AD1. Although genomic PCR of the 5' and 3' junctions of EPO4 and EPO6 cassettes confirmed that plasmids were integrated into site A (Supporting Figure 3e,f), PCR amplification of the whole four- and six-copy expression cassettes in the landing pad (insert PCR) consistently resulted in amplicons with a smaller than expected size (2.1 kb instead of 7.4 kb (EPO4) and 10.9 kb (EPO6)) (Supporting Figure 3j). The observed amplicons were sequence-specific to the landing pad with EPO cassette in site A (Supporting Sequence Data 1). These PCR artifacts can be explained by difficulties of PCR amplification of repetitive DNA sequences resulting from polymerase disengagements, which have been reported previously.³⁶ To analyze GOI copy number of multicopy clones we used dPCR as a reportedly accurate method for copy number analysis.³⁷ While EPO1, EPO2, and EPO4 clones had the desired EPO copy number (Figure 2b), we noticed that dPCR measurements of six-copy EPO6 clones consistently showed lower copy number than expected. This inaccuracy can be attributed to a high partitioning error³⁸ or to incomplete separation of tandem gene copies by restriction digestion prior to dPCR analysis. We also cannot exclude possible recombination of repetitive genetic elements happening inside the CHO cell that could lead to a decrease in copy number in EPO6 clones.

We selected three EPO1 and EPO2 clones and four EPO4 and EPO6 clones, which had integration of RMCE cassettes confirmed by genomic insert PCR (EPO1) or junction PCR (EPO2, EPO4, EPO6) and copy number analysis (EPO1, EPO2, EPO4), and cultivated these clones in batch mode to evaluate EPO expression. The growth of the clones was uniform irrespective of inserted EPO copies (Figure 2c). A twofold increase in titer and q_p was observed for EPO2 clones compared to EPO1, but no further increase was observed for EPO4 and EPO6 clones (Figure 2d,e). EPO mRNA expression levels followed the same trend as q_p (Figure 2f).

To confirm that the observed effects were not specific to EPO, we next constructed multicopy plasmids ETN1, ETN2, ETN4, and ETN6 encoding etanercept and cotransfected them into MCL AD1 together with Cre recombinase to generate RMCE clones. The efficiency of RMCE dropped with the size of the plasmid similarly to EPO plasmids (Supporting Table 1), but no ETN6 clones could be FACS sorted, possibly due to the big insert size (15 kb). We verified ETN1, ETN2, and ETN4 clones by genomic junction PCR (Supporting Figure 3g,h,i) and copy number analysis (Figure 2g), although ETN4 copy number was lower than expected. Moreover, we observed similar PCR artifacts for insert PCR of ETN4 as discussed above (Supporting Figure 3j). We cultivated three clones for each cell design ETN1, ETN2, and ETN4 in batch mode to assess ETN expression. The growth, ETN titer, q_p and mRNA levels of ETN clones followed the same trend as was noticed for EPO clones (Figure 2h–k), with no increase in ETN productivity observed for ETN4 clones.

The single-site multicopy targeted integration of plasmids with more than two GOI copies proved challenging due to

difficulties in clone verification, possible gene recombination, and transcriptional limitation. Thus, we looked for other ways to increase recombinant protein expression in the RMCE system.

Optimal Landing Pad Features. To increase protein expression in our RMCE system, we optimized features of the landing pad, evaluating the effect of a poly(A) tail outside recombination sites and the position of the antibiotic selection marker. Previous investigations of the global translation in the recombinant CHO cells have shown that nonessential and highly abundant mRNA encoding an antibiotic selection marker could occupy about 5% of the translatome and its elimination could improve the production of recombinant proteins.³⁹ Because the cultivation of RMCE-derived subclones does not require the presence of an antibiotic in the medium, we designed new landing pads to be able to remove the NeoR antibiotic selection marker after RMCE. In MCL AD1 used in the initial experiments, NeoR expression cassette was placed outside of lox sites in the landing pad; hence, NeoR was stably expressed in all RMCE-derived subclones. The new landing pads MP3 and LC1 (Figure 3a) had the NeoR expression cassette moved inside the lox sites, allowing it to be exchanged together with EF1 α -mCherry to GOI expression cassettes after RMCE, thus eliminating NeoR expression. We placed BGH poly(A) tail outside of lox sites in the landing pad MP3, and SV40 poly(A) tail in the landing pad LC1 to assess the effect of these poly(A) tails on recombinant protein expression. Using CRISPR/Cas9-mediated targeted integration⁴⁰ we inserted a single copy of the landing pads MP3 or LC1 into the genomic site A of CHO-S cells, creating corresponding MCLs. The integration of the landing pads was confirmed by junction PCR and copy number analysis (Supporting Figure 4a,b). We used MCLs MP3 and LC1 to generate RMCE subclones expressing one copy of EPO by cotransfected EPO1 plasmid and Cre recombinase (Figure 3a). Resulting clones were verified by genomic PCR (data not shown) and copy number analysis (Supporting Figure 4c). We cultivated EPO-expressing subclones of MCLs AD1, MP3, and LC1 in batch mode and measured EPO expression (Figure 3b,c). To investigate the effect of NeoR cassette removal after RMCE, we compared subclones AD1_EPO1 and MP3_EPO1, which both had BGH poly(A) tail outside of lox sites, but kept or removed NeoR after RMCE, accordingly. EPO mRNA expression and q_p both increased by ~20% in MP3-derived clones compared to AD1-derived clones (multiple comparison, p . adj. = 10^{-5}), showing that depletion of the NeoR marker can increase recombinant protein expression in the RMCE system. To study the effect of poly(A) tail on EPO expression, we compared MP3_EPO1 and LC1_EPO1 subclones, which had BGH poly(A) or SV40 poly(A) outside of lox sites, accordingly. The use of SV40 poly(A) tail resulted in a twofold decrease in EPO mRNA expression and a threefold decrease in q_p when comparing MP3- and LC1-derived subclones (multiple comparison, p . adj. = 2×10^{-10}). Thus, we conclude that the design of the landing pad in MCL MP3 supports the highest recombinant protein expression upon RMCE due to removal of NeoR marker and usage of BGH poly(A) tail.

Simultaneous Integration of Multicopy Plasmids in Two Genomic Sites Using the Dual-RMCE System. As demonstrated earlier, single-site targeted integration of >2-copy RMCE plasmids is challenging due to difficulties in verification and transcriptional limitation. As a solution to this, we designed a dual-RMCE system for simultaneous plasmid

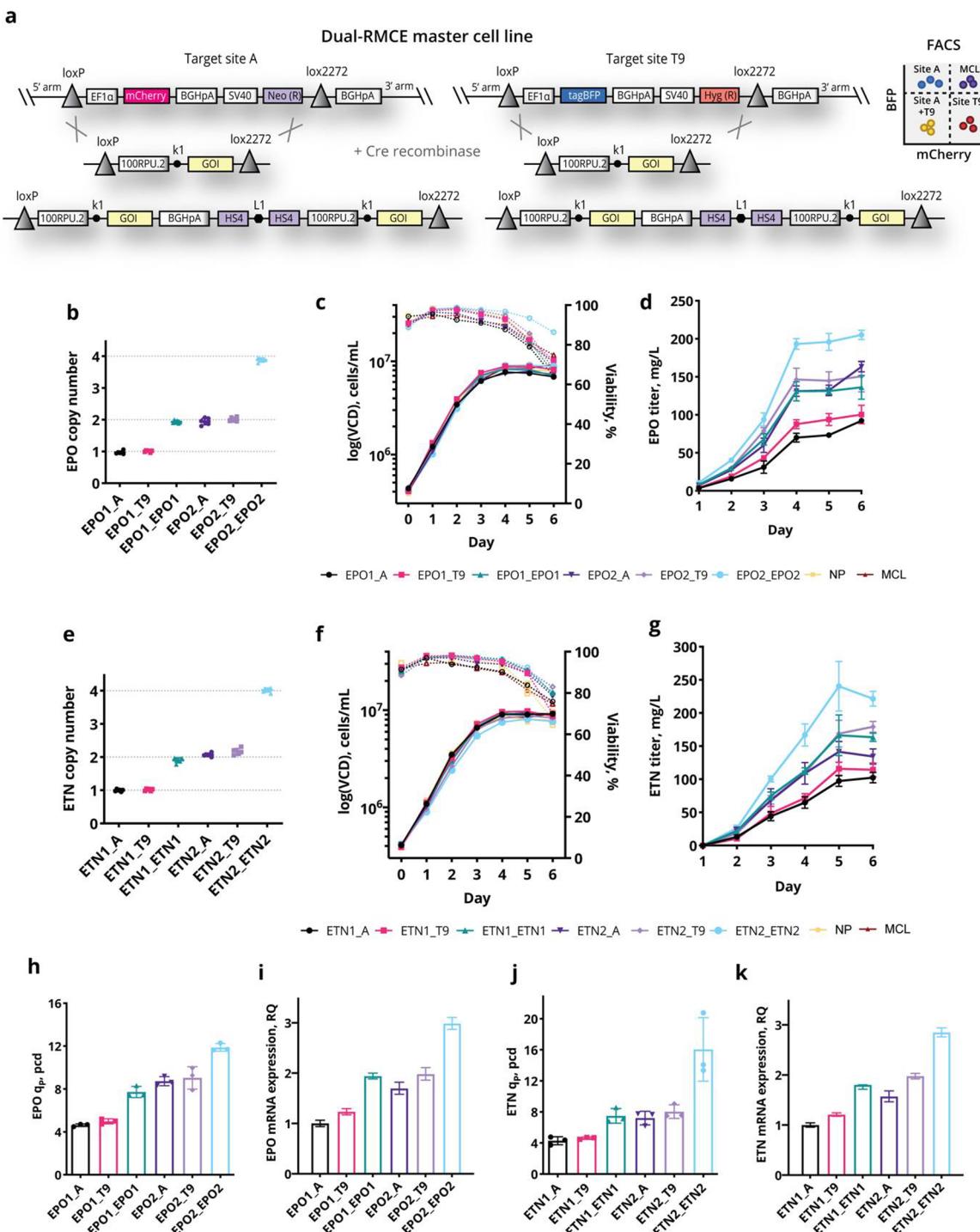


Figure 4. Simultaneous multicopy integration of GOI in a dual-RMCE system. (a) Overview of the dual-RMCE system. The dual-RMCE master cell line has two landing pads integrated in two different genomic sites: loxP-EF1 α -mCherry-NeoR-lox2272 landing pad in site A and loxP-EF1 α -tagBFP-HygR-lox2272 landing pad in site T9. When Cre recombinase is transfected together with a single RMCE donor plasmid (with one GOI copy or two GOI copies), RMCE can occur in site A only, in site T9 only, or both sites simultaneously. To select the cells with integration in one specific site or both sites, FACS is used. (b–k) Characterization of EPO and ETN clones created by targeted integration of RMCE donor plasmids EPO1, EPO2, ETN1, and ETN2 into the dual-RMCE master cell line (MCL). Three clones for each cell design were grown in batch culture with the MCL and three nonproducer (NP) clones. (b, e) Copy number analysis, showing median value for three biological replicates. (c, f) VCD and viability, showing mean and SD for three biological replicates. (d, g) Titer, showing mean and SD for three biological replicates. (h, j) Specific productivity, showing mean and SD for three biological replicates. (i, k) mRNA expression, showing mean and SD for three biological replicates.

integration in two genomic sites to create CHO cell lines expressing multiple copies of GOI from independent locations. Our dual-RMCE system has two landing pads integrated into two predefined chromosomal loci enabling simultaneous

RMCE in both genomic sites within a single transfection (Figure 4a). To build the dual-RMCE system, we used MCL MP3 having loxP-EF1 α -mCherry-NeoR-lox2272 landing pad integrated into site A as a parental cell line. Using CRISPR/

Cas9, we integrated a second landing pad loxP-EF1 α -tagBFP-HygR-lox2272 into site T9 of MCL MP3, creating the dual-RMCE MCL BPS (Supporting Figure 5a,b). Genomic site T9 was previously identified as a safe harbor site for TI and showed high and stable expression of recombinant proteins in CHO cells.²⁶ The use of landing pads encoding different fluorescent markers (mCherry and tagBFP) enables selection of clones with RMCE integration in either or both target sites using FACS. The transfection with Cre recombinase and a single RMCE donor plasmid into the dual-RMCE MCL will result in a mixture of cells, where mCherry-negative/tagBFP-positive cells will have the donor plasmid integrated into site A, mCherry-positive/tagBFP-negative cells will have integration in site T9, and double-negative cells will have integration in both site A and site T9. Thus, by performing a single transfection and selecting the desired population using FACS, it is possible to generate cell lines with GOI integrated into single or dual sites, controlling the GOI copy number.

To test our dual-RMCE system for expression of therapeutic proteins, we generated dual-RMCE subclones expressing one, two, and four copies of EPO and ETN as well as nonproducing (NP) clones serving as a control. NP clones were created by cotransfection of Cre recombinase together with RMCE donor plasmid having three stop codons and short synthetic poly(A) tail; mCherry-negative/tagBFP-negative cells were single-cell sorted by FACS and clones were verified to have integration in both site A and site T9 using genomic PCR (Supporting Figure 5c). To create EPO-producing cell lines, we transfected one-copy EPO1 plasmid together with Cre recombinase into dual-RMCE MCL to generate one-copy clones in site A (EPO1_A), one-copy clones in site T9 (EPO1_T9), and two-copy clones with the one-copy plasmid integrated into both site A and T9 (EPO1_EPO1). Similarly, we transfected two-copy EPO2 plasmid into dual-RMCE MCL to generate two-copy clones in site A (EPO2_A), two-copy clones in site T9 (EPO2_T9) and four-copy clones with two-copy plasmid integrated into both site A and site T9 (EPO2_EPO2). RMCE efficiency measured by FACS was 2–3% for one-copy clones EPO1_A and EPO1_T9; 0.6% for two copy clones EPO1_EPO1, EPO2_A, and EPO2_T9; and 0.01% for four-copy clones EPO2_EPO2 (Supporting Table 2). We single-cell sorted and verified the clones using genomic junction PCR (Supporting Figure 6) and copy number analysis (Figure 4b) without any issues with the verification of four-copy clones. Using the same setup and transfecting ETN1 or ETN2 donor plasmids, we generated ETN-producing cell lines: one-copy clones ETN1_A and ETN1_T9; two-copy clones ETN1_ETN1, ETN2_A, and ETN2_T9; and four-copy clones ETN2_ETN2. RMCE efficiency was similar to the efficiency of EPO integration (Supporting Table 2); clones were confirmed to have desired integration using genomic junction PCR (Supporting Figure 7) and copy number analysis (Figure 4e). Thus, we showed that the dual-RMCE system enables simultaneous integration of up to four copies of therapeutic proteins site-specifically into the CHO genome.

Dual-RMCE System Supports High Expression of EPO and ETN. To assess the expression capability of the dual-RMCE system, we selected three dual-RMCE subclones for each EPO-expressing cell design (EPO1_A, EPO1_T9, EPO1_EPO1, EPO2_A, EPO2_T9, EPO2_EPO2) and ETN-expressing cell design (ETN1_A, ETN1_T9, ETN1_ETN1, ETN2_A, ETN2_T9, ETN2_ETN2) and cultivated them in batch mode together with three NP clones and

the dual-RMCE MCL. The growth of EPO and ETN clones was uniform and similar to the growth of NP clones and MCL, only the ETN2_ETN2 clones had slightly lower VCD (Figure 4c,f). Titer (Figure 4d,g), q_p (Figure 4h,j), and mRNA levels (Figure 4i,k) of both EPO and ETN were continuously increasing with the number of GOI copies integrated. Expression of EPO and ETN in site A and site T9 was similar, with about 7% higher q_p in site T9 for both proteins (insignificant, F-test, $p = 0.578$). Two-copy EPO and ETN clones showed an expected ~2-fold increase in titer, q_p , and mRNA compared to one-copy clones: q_p of one-copy clones EPO1_A, EPO1_T9 and ETN1_A, ETN1_T9 was around 4.6 pcd, q_p of two-copy clones EPO1_EPO1, EPO2_A, EPO2_T9 and ETN1_ETN1, ETN2_A, ETN2_T9 was around 8 pcd (Figure 4h,j). Instead of an ideal 4-fold change in expression in four-copy clones, the increase in q_p of EPO2_EPO2 clones was only 2.6-fold (12 pcd), while the q_p of ETN2_ETN2 clones was increased only to 3.2-fold (14 pcd, excluding one outlier clone with 21 pcd, which showed q_p of 14 pcd in the follow-up experiments). This can indicate the possible appearance of transcription, translation, or secretion bottlenecks with the increased copy number load. mRNA expression of EPO and ETN suggests that the protein production in EPO2_EPO2 and ETN2_ETN2 clones was limited on the level of transcription (Figure 4i,k). Overall, four-copy clones of both EPO and ETN showed the same high specific productivity of 12–14 pcd, and both EPO2_EPO2 and ETN2_ETN2 reached over 200 mg/L in batch culture with no process optimization.

The increase in protein production puts an additional load on the secretory pathway that may influence folding and post-translational modifications of the secreted proteins. To evaluate if the increase in q_p affected the protein quality, we analyzed ETN in the supernatants of four-copy clones (ETN2_ETN2) versus one-copy clones (ETN1_A). ETN is a homodimer with 13 intrachain and 3 interchain disulfide bonds, two N-glycosylation sites in the TNFR domain, one N-glycosylation site in the Fc-domain, and 13 potential O-glycosylation sites on each monomer.⁴¹ We analyzed ETN N-glycans by LC-MS and found the same N-glycan structures in ETN1_A and ETN2_ETN2 samples, with ETN2_ETN2 having only a minor decrease in N-glycan maturation (Supporting Figure 8a). It is known that when ETN is expressed in CHO cells, it can be secreted as a dimer held together by disulfide bonds, but there is a high probability of disulfide bond scrambling, leading to the formation of high-molecular-weight multimers.⁴¹ Western blot analysis of ETN1_A and ETN2_ETN2 supernatants during batch culture showed no difference between one-copy and four-copy clones in the distribution of dimers and multimers (Supplementary Figure 8b). We conclude that increased protein expression in high-producing four-copy clones generated by the dual-RMCE system did not affect the protein quality.

Dual-RMCE Subclones Show Long-Term Stability and High Protein Production in Fed-Batch Culture. Industrial production of biopharmaceuticals requires an assessment of growth and titer stability of the cell lines and optimization of cultivation conditions to ensure consistent and high product yield upon bioprocess. To evaluate the long-term stability of recombinant protein expression and growth of dual-RMCE subclones, we cultivated two cell lines for each EPO- and ETN-expressing cell designs (EPO1_A, EPO1_T9, EPO1_EPO1, EPO2_A, EPO2_T9, EPO2_EPO2, ETN1_A, ETN1_T9, ETN1_ETN1, ETN2_A, ETN2_T9, ETN2_ETN2) and

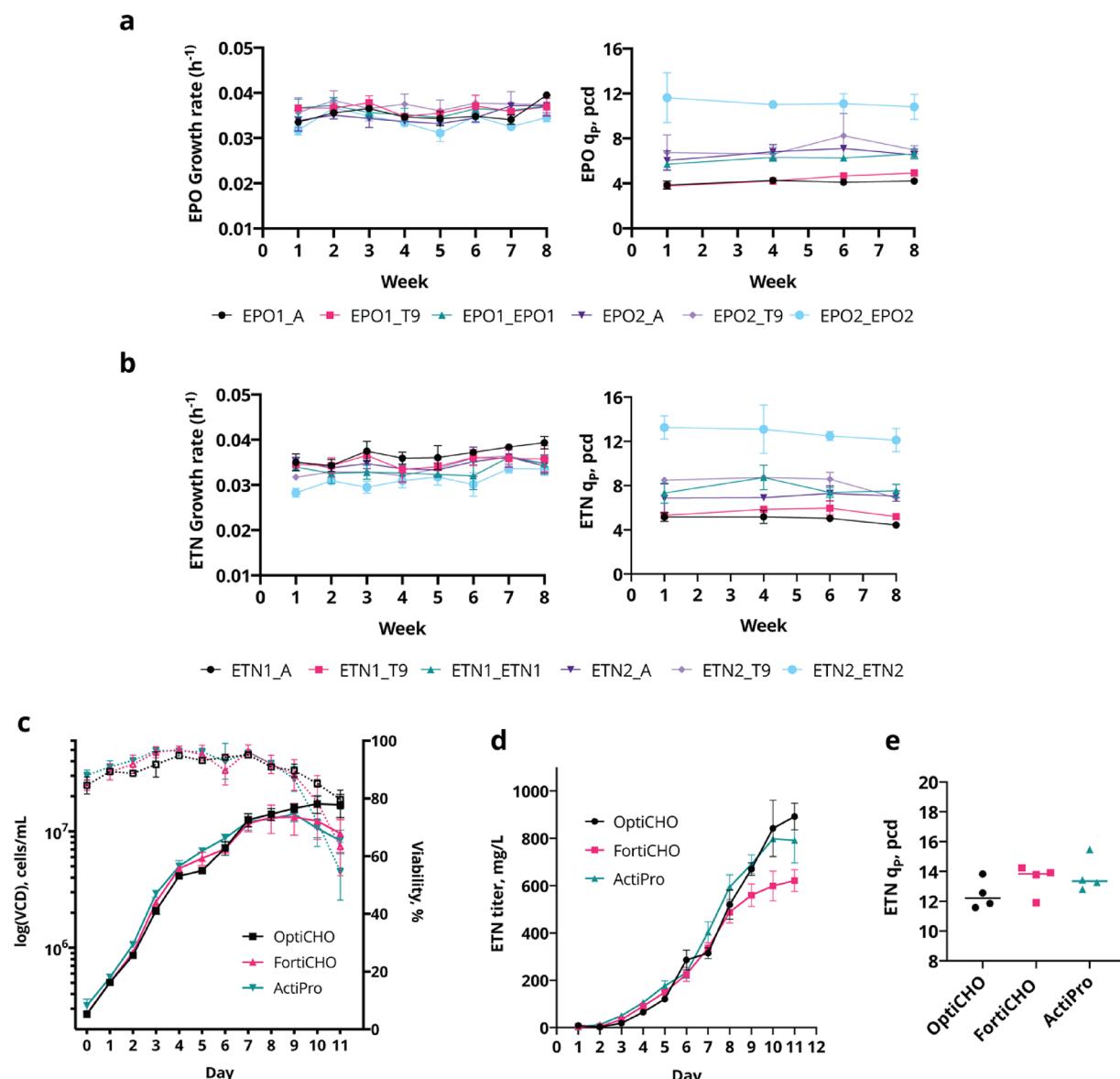


Figure 5. Long-term stability and fed-batch culture performance of multicopy clones generated by the dual-RMCE system. (a) Stability of the dual-RMCE EPO clones, represented by growth rate and specific productivity over 8 weeks, showing mean and SD for two biological replicates for each cell design. (b) Stability of the dual-RMCE ETN clones, represented by growth rate and specific productivity over 8 weeks, showing mean and SD for two biological replicates for each cell design. (c–e) Fed-batch cultivation of two ETN2_ETN2 clones in different production media (OptiCHO, FortiCHO, ActiPro), 2 technical replicates per clone. (c) VCD and viability, showing mean and SD for biological replicates. (d) ETN titer, showing mean and SD for biological replicates. (e) ETN specific productivity, showing median value for two clones with two technical replicates.

N2_ETN2) for two months by passaging the cells three times a week, with titer measurements in the beginning (week 1), middle (week 4 and 6), and end (week 8) of the cultivation. None of the clones decreased in growth rate or lost specific productivity (Figure 5a,b). The growth rate of ETN clones slightly increased by week 8, leading to a small decrease in q_p , which may be an adaptation to culture conditions in repeated passages. Overall, all dual-RMCE subclones retained their growth and productivity over 80 generations, proving their stability.

We aimed to increase the volumetric productivity of four-copy clones ETN2_ETN2 and EPO2_EPO2 by bioprocess optimization to show their suitability in an industrial setup. We cultivated two ETN2_ETN2 clones in the fed-batch mode in shake flasks for 11 days with three different basal media:

OptiCHO, FortiCHO, and ActiPro. In all media, cell lines reached high viable cell densities above 10 million cells/mL (Figure 5c), with OptiPro supporting higher VCD at the end of the culture compared to ActiPro and FortiCHO. It resulted in the final ETN titer of ~900 mg/L in OptiCHO, ~800 mg/L in ActiPro, and ~620 mg/L in FortiCHO media (Figure 5d). ETN q_p in all three media was in the range of 12–14 pcd (Figure 5e), the same as that previously noted in batch culture with CD CHO media. Similarly, EPO2_EPO2 clones cultivated in the fed-batch mode with ActiPro media showed high cell densities, high final titer ~980 mg/L, and median q_p of 14.6 pcd (Supporting Figure 9). Thus, high ETN and EPO titers in fed-batch cultivation were achieved by the increase in the duration of the culture and final cell concentrations and were not attributed to the changes in the specific productivity

of the cell lines. With further bioprocess development aimed at the maximization of cell mass of these clones, it would be possible to reach the scale of >1g/L protein production in the fed-batch culture.

■ DISCUSSION

The overarching aim of this study was to increase q_p of therapeutic proteins using the TI system for accelerated and predictable generation of high-producing CHO cell lines. There are four main engineering strategies that can be used to increase q_p upon TI: (i) selecting hotspots in the genome, (ii) using strong gene regulatory elements in the expression vector, (iii) increasing gene copy number, and (iv) improving the secretory capacity of the cell. In this study, we combined these strategies to improve the expression of complex therapeutic proteins EPO and ETN in a RMCE system, demonstrating a 10-fold increase in q_p compared to our initial RMCE setup. In the process, we have explored the importance of different elements in the expression cassette as well as the capacity for introducing multiple gene copies in a single RMCE or multiple RMCE sites. By developing and optimizing the dual-RMCE system for multicopy integration, we created cell lines that reached q_p of 12–14 pcd for both EPO and ETN, exceeding q_p of the best high-producing clones generated by random integration in our laboratory. Interestingly, both EPO and ETN were expressed at similar q_p , although their protein size and post-translational modifications are different.

To increase q_p upon RMCE, we first explored the effect of gene regulatory elements in the RMCE expression vectors on protein production. We found that the strong synthetic promoter 100RPU.2 supports high expression of recombinant proteins upon TI, resulting in a 2-fold increase in q_p compared to the best promoter (CP) from our previous study (Figure 1c). The 100RPU.2 promoter was also protein- and site-independent: the expression of 100RPU.2-EPO (EPO1) and 100RPU.2-ETN (ETN1) constructs in target sites A and T9 was equally high (Figure 4h,j). This illustrates the promise of synthetic biology approaches for designing gene regulatory elements to improve biopharmaceuticals production. In this study, we tested just a few regulatory elements to maximize expression from multicopy vectors due to a small list of mammalian gene regulatory elements characterized to date, selecting only the elements with reported ability to increase protein expression in mammalian cells. Some genetic elements that may positively affect gene expression and stability such as epigenetic regulatory elements S/MAR and UCOE⁴² were not included in our study due to long-term stability of genomic sites A and T9 shown previously.²⁶ Additionally, the large size of these epigenetic elements (over 1 kb) would negatively affect the integration efficiency of RMCE plasmids. Two regulatory elements tested in our study with previously reported ability to increase protein expression (WPRE that was able to improve protein expression in viral vectors³⁰ and Kozak sequence k2 that showed increased expression of eGFP in CHO cells²⁹) did not perform well, suggesting a significant degree of a construct-, cell-, or gene-specificity. Construct-, cell-, and gene-specificity of WPRE have been previously noted in other studies.^{43–45} Interestingly, both post-transcriptional elements WPRE and Kozak sequence k2 affected transgene mRNA levels (Supporting Figure 1e,f), correlating with specific productivity. A similar effect of 5'UTR sequence on mRNA levels has been observed in yeasts, showing a high correlation between protein and mRNA levels for different 5'UTR

sequences in position −10 to −1.⁴⁶ Because both WPRE and Kozak sequences are located in UTR, we hypothesize that their presence can affect the overall structure of mRNA and so may change its stability. The molecular mechanisms of such effects are yet to be uncovered. Another tested element, the chicken HS4 insulator, had a positive effect on the expression from two-copy vectors (Figure 1f,g) presumably by blocking transcriptional interference, thus proving the importance of insulator usage in multicopy expression cassettes. Vector engineering remains one of the main strategies to boost therapeutic protein expression upon TI, with advances in mammalian synthetic biology and machine learning being crucial for the discovery of novel elements.⁴⁷ However, the reported crosstalk between the integration sites and vector regulatory elements might contribute to additional levels of transgene regulation upon TI,^{26,48} so evaluation of newly discovered vector elements in each target site is required.

The hypothesis that q_p strongly correlates with GOI copy number, although intuitive, was not supported by studies in CHO cell lines generated by random integration. Transgene vector rearrangements and genomic positions appear to have a greater influence on the level of expression than the number of integrated gene copies.²³ With the development of TI methods for mammalian cell line generation, it has become possible to test this hypothesis under controlled conditions with minimal clonal variation. Recent reports of multicopy mAb expression in TI systems showed that mAb production can be linearly increased with the copy number²¹ or increased in a nonlinear fashion²⁰ and affected by the ratio and position of heavy and light chains in RMCE expression cassette.²⁰ To assess the effect of gene copy number on protein expression independently from the effect of chain ratio, we expressed single-chain proteins EPO and ETN in our RMCE system. It allowed us to study the correlation between copy number and productivity in a more defined manner and reveal bottlenecks in protein production that occur with an increase in gene copy number.

There are two ways of controlling GOI copy number upon TI: one is multicopy GOI integration into a single genomic site, another is GOI integration into multiple genomic sites. Here, we evaluated both strategies in the RMCE system in order to improve q_p and identified their limitations. First, we demonstrated that the introduction of optimized two-copy expression vectors in a single site leads to a twofold increase in q_p and is feasible for both EPO and ETN (Figure 2). However, the insertion of larger four- and six-copy plasmids in a single site was more difficult, limited by a capacity to validate the clones and the low integration efficiency. Overall, the challenge of insert verification makes TI of large multicopy plasmids less preferable for routine CHO cell line development until more robust methods of clone verification are developed. The observed low integration efficiency can be attributed to two factors: the transfection efficiency of large multicopy plasmids and RMCE efficiency. The transfection efficiency of large plasmids in mammalian cells is known to be low,⁴⁹ so improved methods for plasmid delivery such as minicircle DNA may increase TI efficiency.⁵⁰ The principle of RMCE places a limitation on TI efficiency due to the requirement of two recombination events happening simultaneously in two recombination sites. The use of a single recombination site for TI can increase integration efficiency of multicopy plasmids, which was demonstrated for mAbs production in CHO cells.²¹ However, the reported mAb expression levels were low (maximum 30 mg/L), and no q_p estimates were shown in

the mentioned study, making it difficult to assess the applicability of such TI system for industrial production of biopharmaceuticals. Moreover, TI using a single recombination site has a disadvantage in the integration of bacterial plasmid DNA together with GOI that can affect cell growth and cause transcriptional silencing of transgenes.⁵⁰

Importantly, when integrating four and six copies of GOI into a single RMCE site, we saw no increase in protein expression in EPO4, EPO6, or ETN4 clones, which was accompanied by low transgene mRNA level (Figure 2). There are several potential causes of the observed transcriptional limitation, including recombination or silencing of repetitive elements in the genome⁵¹ or transcriptional interference between tandemly assembled gene copies,^{52,53} which was not prevented by the use of insulators. Similar to our study, accumulative TI of scFv-Fc antibody into a single genomic site, which led to the generation of four- and six-copy clones, showed the appearance of promoter interference with an increase in the number of expression cassettes.⁵⁴ Follow-up experiments are required to reveal the cause of low transgene expression in the EPO4, EPO6, and ETN4 clones.

To remove possible promoter interference between multiple expression cassettes in our RMCE system, we created a dual-RMCE MCL for multicopy integration into separate genomic sites. First, we showed that the integration of two copies in either of our two target sites or one copy in both sites can generate cell lines with almost double the productivity (Figure 4). Then, we created four-copy cell lines EPO2_EPO2 and ETN2_ETN2 with two-copy plasmids being integrated into two genomic sites simultaneously. We observed that the protein expression of both EPO and ETN did not increase proportionally to the gene copy number, as the q_p increase from one to four copies was only 2.6- and 3.2-fold, respectively (Figure 4h,j). The decrease in GOI expression efficiency in four-copy clones was reflected in transgene mRNA levels (Figure 4i,k), pointing to a transcriptional bottleneck. The transcriptional bottleneck was not caused by a transcriptional interference of integrated cassettes in their target sites, because two-copy clones showed a 2-fold increase in q_p individually. Moreover, the verification of clones confirmed that no recombination occurred between the gene copies. Thus, opposed to the common notion that high protein expression in CHO cells is hindered by secretion bottlenecks,⁵⁵ we observed that a transcriptional bottleneck can appear with an increase in gene dosage.

What can limit transgene mRNA expression when its copy number is increased? A possible reason for ineffective transcription could be a depletion of required transcription factors.⁵⁶ Such depletion of the transcription machinery would have caused changes in global transcription and, consequently, cell growth. However, the growth of the clones was not hampered with an increase in protein production (Figure 4c,f; Figure 5a,b), which suggests that the transcription machinery was not in limited supply. It could be that dual-RMCE clones showing uniform growth adapted to avoid utilizing all available resources toward recombinant protein production, and the selective repression of transgene expression occurred as a feedback regulation to keep the cell growth unchanged. Various post-transcriptional mechanisms can lead to a negative regulation of transgene mRNA expression and turnover, altering its transport and stability. The selectivity of such regulation would be controlled by specific RNA-binding proteins and/or noncoding RNAs. Curiously, experiments in

yeasts have shown that mRNA of genes with an increased copy number expressed from endogenous loci can be selectively targeted and degraded by RNA interference.⁵⁷ A similar mechanism might be responsible for a decrease in mRNA levels of multicopy transgenes in CHO cells. To understand the cause of transcriptional bottlenecks observed upon multicopy expression in our system, further systematic studies of the global cellular response are required.

In summary, our study is the first comprehensive characterization of multicopy targeted integration for the production of complex therapeutic proteins in CHO cells. The use of the dual-RMCE system enabled accelerated and predictable generation of stable CHO cell lines with high productivity and high titer, thus demonstrating the suitability of TI for industrial production of therapeutic proteins. Due to minimal clonal variation, it was possible to elucidate the relationship between gene copy number and protein expression and reveal a transcriptional bottleneck appearing with an increase in gene dosage. We have made our multicopy RMCE system available for the scientific community and believe that it can aid a better understanding of the response to recombinant protein production in mammalian cells and help develop new strategies for mammalian cell engineering toward the efficient production of biopharmaceuticals.

METHODS

Plasmid Construction. Plasmids created and used in this study are listed in Supporting Table 3 (with links to Addgene for selected plasmids). All plasmids were constructed by uracil-specific excision reagent (USER) cloning using flexible assembly sequence tags.³⁵ Fragments for USER assembly (DNA bricks) were generated by PCR amplification with Phusion U Hot Start DNA polymerase (Thermo Fisher Scientific) and uracil-containing primers (IDT). Landing pad plasmids used for the generation of MCLs MP3, LC1 and dual-RMCE BP5 included 5' homology arm for targeted site (site A or site T9), loxP (ATAACTTCGTATAGCATACT-TATACGAAGTTAT), EF1 α promoter, mCherry or tagBFP gene, BGH poly(A) tail, SV40 promoter, NeoR or HygR gene, lox2272 (ATAACTTCGTATAGGATACTTTATAC-GAAGTTAT), BGH poly(A) or SV40 poly(A) tail, 3' homology arm and ZsGreen1-DR expression cassette (CMV_ZsGreen1-DR_BGHpA). RMCE donor plasmids were assembled by USER cloning by combining up to 7 DNA bricks in one reaction. Features of donor plasmids are described in Supporting Table 3 and depicted in Figure 1 and Figure 2. They contained one or more copies of GOI (EPO or ETN) and other genetic elements (promoter, poly(A) tail, WPRE, insulator) flanked by loxP site at the 5' end and lox2272 at the 3' end. For cloning of multicopy plasmids EPO4, EPO6, ETN4, and ETN6, smaller plasmids HS4_100RPU.2_GOI_BGHpA_HS4 were created first, and these were used as a template for PCR amplification of expression units for the USER assembly of full multicopy plasmids. Each expression unit was flanked by specific restriction sites that were introduced on 5' and 3' ends of USER linkers (listed in Supporting Table 4) to ease the following plasmid and clone verification. RMCE donor plasmid for generation of NP clones had 3 stop codons and short synthetic poly(A) tail,⁵⁸ flanked by loxP and lox2272 sites. Plasmids and sequences used for the amplification of genetic elements have been described previously: site A and site T9 homology arms, mCherry, NeoR, EF1 α , SV40 promoters and

poly(A) tails, ZsGreen1-DR expression cassette, *Cricetulus griseus* codon-optimized EPO were amplified from plasmids described in Pristovšek et al.;²⁶ the ETN sequence was described in Pristovšek et al.²⁴ The sequence of 100RPU.2 promoter (aka 100 RPU construct 2) was described in Brown et al.²⁷ and was synthesized as gBlock (IDT). tagBFP was amplified from plasmid Addgene#60903; WPRE was amplified from plasmid Addgene#17492, and chicken HS4 insulator was amplified from plasmid Addgene#36880. Plasmid backbone was cloned from pJ204 plasmid (DNA 2.0). Assembled DNA bricks were transformed into *E. coli* Mach1 competent cells (Thermo Fisher Scientific) for cloning of landing pad plasmids and one- and two-copy RMCE donor plasmids. For the cloning of multicopy RMCE donor plasmids (EPO4, EPO6, ETN4 and ETN6) Stbl3 competent cells were used (Thermo Fisher Scientific). All constructs were verified by restriction digestion and Sanger sequencing. Multicopy plasmids EPO4, EPO6, ETN4, and ETN6 were digested prior to sequencing with linker-specific restriction enzymes, generating fragments with separate expression units. These units were then isolated from agarose gel and sequenced. For the transfection in CHO cells, plasmids were purified using NucleoBond Xtra Midi EF kit (Macherey-Nagel) according to manufacturer's instructions.

Cell Cultivation. CHO-S cells (Thermo Fisher Scientific) were maintained in CD CHO medium supplemented with 8 mM L-glutamine (Thermo Fisher Scientific) and cultivated in 125 mL Erlenmeyer shake flasks (Corning), incubated at 37 °C, 5% CO₂ at 120 rpm in humidified incubator, and passaged every 2–3 days. Viable cell density and viability were monitored using the NucleoCounter NC-200 Cell Counter (ChemoMetec), using Vial-Cassettes and “Viability and Cell Count Method 2” assay.

Master Cell Lines Generation Using CRISPR/Cas9. CHO-S cells (1 × 10⁶ cells/mL) were transfected with the sgRNA plasmid and landing pad plasmid for the corresponding integration sites together with a GFP_2A_Cas9 plasmid at a ratio of 1:1:1 (w:w:w) in 6-well plates using FreeStyle MAX transfection reagent (Thermo Fisher Scientific) according to the manufacturer's recommendations. Three days after transfection cells were seeded at 0.3 × 10⁶ cells/mL in 6-well plates in medium containing G418 (600 µg/mL; Sigma-Aldrich) for MCLs MP3 and LC1 generation or Hygromycin (1000 µg/mL; Sigma-Aldrich) for the dual-RMCE MCL BPS generation. During antibiotic selection, the antibiotic-containing medium was exchanged every 2–3 days. After 9–12 days of selection in suspension, cell pools were subjected to single cell sorting by FACS, where the mCherry-positive/ZsGreen1-DR-negative cells were isolated for the generation of MCLs MP3 and LC1. ZsGreen1-DR was used as a marker to exclude cells that had random integration of the landing pad plasmid. For the generation of dual-RMCE MCL BPS, the loxP-EF1α-mCherry-NeoR-lox2272-BGHpA landing pad (PL0627) was first integrated in the site A resulting in creation of MCL MP3, and then the loxP-EF1α-tagBFP-HygR-lox2272-BGHpA landing pad (PL0626) was integrated into the site T9 of MCL MP3. Generation of MCL AD1 was described previously.²⁶ Detailed protocol for CRISPR/Cas9-mediated targeted integration in CHO cells is described in.⁴⁰

Fluorescence Activated Cell Sorting (FACS). For clonal isolation, cell pools were single cell sorted using a FACSJazz cell sorter (BD Biosciences) equipped with 405, 488, and 561 nm lasers. Single cells were seeded in flat-bottom Corning 384-

well plates (Sigma-Aldrich) in 30 µL of CD CHO medium, supplemented with 8 mM L-glutamine, 1.5% HEPES (Gibco), and 1× antibiotic–antimycotic (Gibco). Ten to 14 days after single cell sorting, plates were visualized using Celigo image cytometer (Nexcelom Bioscience), and the entire volume of subconfluent clones was transferred to 180 µL of CD CHO medium supplemented with 1× antibiotic–antimycotic (Gibco) in flat-bottom 96-well plates using an epMotion 5070 liquid handling workstation (Eppendorf). Subsequently, verified cells were expanded in suspension to 12-well plates, then to 6-well plates and shake flasks.

RMCE Subclones Generation. MCLs at concentration 1 × 10⁶ cells/mL were transfected with RMCE donor plasmid and Cre-recombinase plasmid in 3:1 ratio (w:w) in 6-well plates using FreeStyle MAX transfection reagent (Thermo Fisher Scientific). Cells were passed two times after transfection and then were FACS sorted, using the respective MCLs as a gating control. To generate D1–D3 cell pools, mCherry-negative cells were bulk sorted as described before.²⁶ For the generation of R1–R6, EPO1–EPO6, and ETN1–ETN4 cell lines and MP3 and LC1 RMCE subclones, mCherry-negative cells were single cell sorted as described above. For the generation of dual-RMCE subclones, the following sorting scheme was applied: mCherry-negative/tagBFP-positive cells were sorted for selecting clones with recombinase-mediated integration in site A (EPO1_A, EPO2_A, ETN1_A, ETN2_A); mCherry-positive/tagBFP-negative cells were sorted for selecting clones with recombinase-mediated integration in site T9 (EPO1_T9, EPO2_T9, ETN1_T9, ETN2_T9), and mCherry-negative/tagBFP-negative cells were sorted for selecting clones with simultaneous recombinase-mediated integration in sites A and T9 (EPO1_EPO1, ETN1_ETN1, EPO2_EPO2, ETN1_ETN1, NP). Subsequently, clones were expanded and verified by PCR of the targeted regions and copy number analysis.

Genomic PCR Amplification of Targeted Regions. Targeted integration was verified by PCR amplification of genomic DNA of targeted locus using specific primers binding outside of the integration site in the genomic locus (5' and 3' OUT primers) and primers specific to the 5' and 3' ends of the integrated cassette (5' and 3' IN primers). Insert PCR was performed by combining forward 5' OUT primer with reverse 3' OUT primer. 5' and 3' junction PCR was performed by pairing forward 5' OUT primer with reverse 5' IN primer and forward 3' IN primer with reverse 3' OUT primer, respectively. Schematic representation of PCR set up is shown on Supporting Figure 3a. PCR primers for junction and insert PCR are listed in Supporting Table 5. For junction and insert PCR genomic DNA was extracted from the cell pellets harvested from 96-well plates using QuickExtract DNA extraction solution (Lucigen) according to the manufacturer's instructions. PCR was carried out using 2× Phusion Master Mix (Thermo Fisher Scientific) in touchdown PCR: 98 °C for 30 s; 10X : 98 °C for 10 s, (Tm + 10) – Tm °C [-1°/cycle] for 30 s, 72 °C for 30 s or more (6 min for the insert PCR for EPO4, EPO6, and ETN4 clones); 25X : 98 °C for 10 s, Tm °C for 30 s, 72 °C for 30 s or more (6 min for the insert PCR for EPO4, EPO6, and ETN4 clones); 72 °C for 5 min. One microliter of genomic DNA was used as a PCR template in 10–20 µL of total PCR reaction volume. PCR products were visualized on a 1% agarose gel, and only clones showing correct amplicons were selected.

Gene Copy Number Analysis by qPCR. Copy number analysis of MCLs and RMCE subclones designed to have one copy of GOI was performed by qPCR on QuantStudio 5 Real-Time PCR System (Applied Biosystems). Genomic DNA was purified using a GeneJet Genomic DNA extraction kit (Thermo Fisher Scientific) according to the manufacturer's instructions and diluted to 10 ng/μL. TaqMan Gene Expression Master Mix (Thermo Fisher Scientific), custom-made TaqMan assays for GOI (mCherry, tagBFP, EPO, ETN), and one-copy endogenous reference gene C1GALT1C1 (*Cosmc*)⁵⁹ were used in duplex assay (one GOI and one reference gene). Sequences of TaqMan probes and primers are listed in Supporting Table 6. qPCR amplification was performed under the following conditions: 50 °C for 2 min, 95 °C for 10 min; 40× : 95 °C for 15 s, 60 °C for 1 min. Each experiment included no template control and three technical replicates for each sample. Results were analyzed in QuantStudio Design and Analysis software. A comparative C_t method was applied to calculate GOI copy number by comparing $\Delta\Delta C_t$ values of analyzed clones to $\Delta\Delta C_t$ values of one-copy GOI clones generated previously.⁷ Only clones with a single copy of GOI (RQ 0.7–1.3) were expanded and banked.

Gene Copy Number Analysis by Digital PCR. RMCE subclones designed to have two, four, and six copies of GOI as well as cell pools D1–D3 were analyzed by digital PCR (dPCR) on QuantStudio 3D Digital PCR System (Applied Biosystems). Genomic DNA was purified using GeneJet Genomic DNA extraction kit (Thermo Fisher Scientific) according to manufacturer's instructions. Prior dPCR DNA of multicopy subclones was digested using restriction enzyme Eco32I (EcoRV) by incubating 1000 ng of genomic DNA with 2 μL FastDigest Eco32I (Thermo Fisher Scientific) at 37 °C for 20 h, then diluted to 6 ng/uL with TE buffer. DNA was mixed with QuantStudio 3D Digital PCR Master Mix v2 and two TaqMan assays (endogenous one-copy gene *Cosmc* and GOI, Supporting Table 6) and loaded on QuantStudio 3D Digital PCR 20K Chips v2 (Thermo Fisher Scientific). dPCR amplification was performed under the following conditions: 96 °C for 10 min; 39×: 60 °C for 2 min, 98 °C for 30 s; 60 °C for 2 min. Results were evaluated using Applied Biosystems AnalysisSuite. GOI copy number was calculated as a ratio of GOI copies/μL to *Cosmc* copies/μL.

Batch Cultivation. For batch cultivation, cells were seeded at 4×10^5 cells/mL in 40 mL CD CHO medium, supplemented with 8 mM L-glutamine and 2 μL/mL antyclumping agent in 125 mL Erlenmeyer shake flasks. Cells were incubated in a humidified incubator at 37 °C, 5% CO₂ at 120 rpm. VCD and viability were monitored daily using the NucleoCounter NC-200 Cell Counter; samples for titer measurements were harvested daily, and samples for RNA expression analysis were harvested on day 3. Cultures were discontinued on day 6.

Titer Measurements and q_p Calculation. Supernatants during a batch, fed-batch, and stability test were collected by harvesting cell culture and centrifugation at 200g 5 min and stored at –80 °C. EPO titers were measured with an Octet RED96 system (ForteBio) using Streptavidin biosensors (ForteBio), functionalized with the anti-EPO VHH biotin conjugate (Life Technologies, cat. no. 7103372100), as previously described.⁶⁰ Absolute EPO titers were calculated using a 7-point calibration curve generated from a dilution series of commercially available EPO (Genscript, cat. no.

Z02975-50) in range 1.95–125 mg/L. ETN titers were measured by fluorescent polarization assay with ValitaTiter plates (Valitacell) following manufacturer protocol using Spark plate reader (Tecan). Absolute ETN titers were calculated using a 10-point calibration curve generated from a dilution series of Enbrel (Pfizer) in range 3.125–150 mg/L. EPO and ETN standards and samples, when necessary, were diluted with spent CD CHO media. Specific productivity (q_p , pg/cell/day) was calculated using titer and integral of viable cell (IVC) values for exponential phase of cultivation (from day 1 to day 4 for batch cultures, from day 1 to day 5 for fed-batch). q_p was calculated as a slope of the linear regression line through data points of daily IVC (x-axis) and titer (y-axis) measurements. IVC value for each day was calculated as shown in the equation:

$$\text{IVC} = \frac{\text{VCD} + \text{VCD}_{-1}}{2} + \text{IVC}_{-1}$$

where VCD is the viable cell density measured on a specific day, VCD₋₁ is the viable cell density recorded on the previous day, and IVC₋₁ is the IVC calculated for the previous day.

Relative mRNA Expression Analysis by RT-qPCR. Total RNA was extracted using RNeasy Mini Plus Kit (QIAGEN) following the manufacturer's protocol. Samples were treated with DNase and cDNA was synthesized using Maxima First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). RT-qPCR was performed on the QuantStudio 5 Real-Time PCR System using TaqMan Multiplex Master Mix (Thermo Fisher Scientific) in a triplex assay (GOI and two reference genes) using the following amplification conditions: 50 °C for 2 min, 95 °C for 10 min; 40× : 95 °C for 15 s, 60 °C for 1 min. Sequences of custom-made Taqman probes for EPO and ETN, reference genes *Gnb1* and *Fkbp1a*⁶¹ as well as primers are listed in Supporting Table 6. Relative mRNA expression levels of GOIs were calculated using Applied Biosystems Analysis Software by comparative C_t method. Each experiment included a “no template” control and three technical replicates for each sample.

Analysis of ETN Protein Quality by Western Blot. SDS-PAGE of culture supernatants was performed in nonreduced conditions using Novex 4–12% Bis-Tris protein gels. Around 100 ng of ETN were loaded per lane according to ETN titer measurements. Proteins were transferred to iBlot2 nitrocellulose membranes using iBlot2 dry blotting system (Thermo Fisher Scientific). After blocking with 5% milk in PBS with 0.1% Tween buffer, the membranes were incubated with IgG Fc Rabbit antihuman HRP-conjugated antibodies (Thermo Fisher Scientific, cat. no. 10475535) and were visualized using Amersham ECL Prime Western Blotting Detection Reagent on Amersham Imager 600 (GE Healthcare).

N-Glycan Analysis. For N-glycan analysis, ETN was purified from supernatant samples harvested on day 6 in batch culture. Supernatants were centrifuged at 6000g for 5 min and purified by protein A chromatography using HiTrap MabSelect Xtra columns (GE Healthcare) on ÄKTA pure protein purification system (GE Healthcare). N-glycans were released from purified ETN and derivatized using GlycoWorks RapiFluor-MS N-Glycan kit (Waters) following the manufacturer's instructions. Labeled N-glycans were analyzed by LC-MS using Thermo Ultimate 3000 HPLC with fluorescence detector coupled to Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific). Separation was performed on ACQUITY UPLC Glycan BEH Amide 130 Å column

(Waters). The amount of N-glycans was measured by integrating the areas under the normalized fluorescence spectrum peaks with Thermo Xcalibur software, giving a normalized relative amount of glycans.

Stability Test. Prior to the stability test, clones were recovered from cryopreservation for 1 week by passaging every 2–3 days in 125 mL shake flasks in CD CHO medium supplemented with 8 mM L-glutamine, 1% antibiotic–antimycotic, and 2 μ L/mL antyclumping agent. For stability test cells were seeded at 3×10^5 cells/mL in 6-well plates (Corning) and passed every 2–3 days at the same concentration for 8 weeks (80 generations) without antibiotic selection pressure. VCD and viability were monitored using the NucleoCounter NC-200 Cell Counter. Growth rate (μ , h^{-1}) was determined using the following equation:

$$\mu = \frac{\ln(VCD) - \ln(VCD_s)}{\Delta t}$$

where VCD is the viable cell density measured on a specific day, VCD_s is the seeding density, and Δt is time in hours.

Fed-Batch Cultivation. Prior fed-batch cultivation cells were adapted to corresponding media 3 passages before inoculation. Fed-batch cultures were inoculated at 3×10^5 cells/mL in 30 mL of OptiCHO (Thermo Fisher Scientific), FortiCHO (Thermo Fisher Scientific) or ActiPro (GE Healthcare) medium, supplemented with 8 mM L-glutamine, 1% Antibiotic-antimycotic and 1 μ L/mL antyclumping agent in 125 mL Erlenmeyer shake flasks. Cells were incubated in a humidified incubator at 37 °C, 5% CO₂ at 120 rpm, on day 3 temperature was shifted to 33 °C. Accelerated feeding was started on day 3 and continued for 4 days with Cell Boost 7a (GE Healthcare) at 1.5, 3, 4.5, and 6% of working volume and Cell Boost 7b (GE Healthcare) at 0.15, 0.3, 0.45, and 0.6% of working volume. Cultures were maintained at a minimum of 24 mM glucose by supplementing with a 400 mM glucose concentrate and at a minimum of 4 mM glutamine by supplementing with a 200 mM glutamine concentrate. Metabolite concentrations were determined daily using Bioprofile 400 (Nova Biomedical). VCD and viability were monitored daily using the NucleoCounter NC-250 Cell Counter (ChemoMetec). Cell cultures were terminated when viability was below 80%.

Statistics. All statistics was performed in R. Data were fitted using linear models, and the multcomp package was used to perform the contrasts described. Adjusted *p*-values corrected for multiple testing are reported.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acssynbio.0c00322>.

Verification of D1–D3 cell pools and R1–R6 clones; multicopy plasmid verification (EPO4, EPO6); verification of EPO1–EPO6 and ETN1–ETN4 clones; verification of MP3 and LC1 master cell lines and clones AD1_EPO1(R1), MP3_EPO1 and LC1_EPO1; verification of dual-RMCE BPS master cell line and NP clones; verification of dual-RMCE EPO clones; verification of dual-RMCE ETN clones; analysis of ETN protein quality by N-glycan analysis and Western blot; fed-batch cultivation of two EPO2_EPO2 clones; RMCE integration efficiency for cell designs EPO1–

EPO6 and ETN1-ETN6; RMCE integration efficiency for dual-RMCE EPO and ETN clones; plasmids used and created in this study; list of extended USER linkers with restriction sites; primers used for verification of targeted integration; TaqMan assays; supporting sequence data (PDF)

AUTHOR INFORMATION

Corresponding Author

Lise Marie Grav — *The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby 2800, Denmark*;  orcid.org/0000-0001-8042-7108; Email: lgrav@biosustain.dtu.dk

Authors

Daria Sergeeva — *The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby 2800, Denmark*;  orcid.org/0000-0003-2779-945X

Gyun Min Lee — *The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby 2800, Denmark*; *Department of Biological Sciences, KAIST, Daejeon 34141, Republic of Korea*;  orcid.org/0000-0002-9876-7378

Lars Keld Nielsen — *The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby 2800, Denmark*; *Australian Institute for Bioengineering and Nanotechnology, University of Queensland, Brisbane 4072, Australia*;  orcid.org/0000-0001-8191-3511

Complete contact information is available at: <https://pubs.acs.org/10.1021/acssynbio.0c00322>

Author Contributions

D.S., L.M.G., G.M.L., and L.K.N. designed the study. D.S. performed experiments and analyzed the data. L.K.N. performed statistical analysis. L.M.G., G.M.L., and L.K.N. supervised the project. D.S. drafted the manuscript. L.M.G., G.M.L., and L.K.N. reviewed and corrected the manuscript.

Notes

The authors declare the following competing financial interest(s): A patent application has been filed by third parties on the promoter 100RPU.2 used in this paper, permission for deposition of plasmids containing the promoter sequence has been provided by AstraZeneca and the University of Sheffield.

ACKNOWLEDGMENTS

The authors thank Helene Fastrup Kildegaard for the initiation of the project, Saranya Nallapareddy for assistance with landing pad plasmid cloning, Nathon Charanyanonda Petersen and Karen Kathrine Brøndum for assistance with the FACS, Elham Maria Javidi and Marianne Decker for performing fed-batch cultures, Helle Munck Petersen for assistance with the protein purification, Anders Holmgård Hansen for assistance with N-glycan analysis, and Sara Petersen Bjørn for valuable discussions. We greatly acknowledge funding support from the Novo Nordisk Foundation (NNF10CC1016517). D.S. is receiving funding from NNF Copenhagen Bioscience PhD programme (NNF16CC0020908).

REFERENCES

- Walsh, G. (2018) Biopharmaceutical benchmarks 2018. *Nat. Biotechnol.* 36, 1136–1145.

- (2) Noh, S. M., Shin, S., and Lee, G. M. (2019) Cell Line Development for Therapeutic Protein Production. In *Cell Culture Engineering* (Lee, G. M., and Kildegaard, H. F., Ed.), pp 23–47, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.
- (3) Seth, G., Charaniya, S., Wlaschin, K. F., and Hu, W.-S. (2007) In pursuit of a super producer—alternative paths to high producing recombinant mammalian cells. *Curr. Opin. Biotechnol.* 18, 557–64.
- (4) Lee, J. S., Kildegaard, H. F., Lewis, N. E., and Lee, G. M. (2019) Mitigating Clonal Variation in Recombinant Mammalian Cell Lines. *Trends Biotechnol.* 37, 931–942.
- (5) Kelley, B. (2020) Developing therapeutic monoclonal antibodies at pandemic pace. *Nat. Biotechnol.* 38, 540–545.
- (6) Fukushige, S., and Sauer, B. (1992) Genomic targeting with a positive-selection lox integration vector allows highly reproducible gene expression in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 89, 7905–7909.
- (7) Grav, L. M., Sergeeva, D., Lee, J. S., Marin de Mas, I., Lewis, N. E., Andersen, M. R., Nielsen, L. K., Lee, G. M., and Kildegaard, H. F. (2018) Minimizing Clonal Variation during Mammalian Cell Line Engineering for Improved Systems Biology Data Generation. *ACS Synth. Biol.* 7, 2148–2159.
- (8) Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A., and Kucherlapati, R. S. (1985) Insertion of DNA sequences into the human chromosomal β -globin locus by homologous recombination. *Nature* 317, 230–234.
- (9) Thomas, K. R., and Capecchi, M. R. (1987) Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* 51, 503–512.
- (10) Sauer, B., and Henderson, N. (1990) Targeted insertion of exogenous DNA into the eukaryotic genome by the Cre recombinase. *New Biol.* 2, 441–449.
- (11) O’Gorman, S., Fox, D. T., and Wahl, G. M. (1991) Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science* 251, 1351–1355.
- (12) Schlake, T., and Bode, J. (1994) Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. *Biochemistry* 33, 12746–12751.
- (13) Hamaker, N. K., and Lee, K. H. (2018) Site-specific Integration Ushers in a New Era of Precise CHO Cell Line Engineering. *Curr. Opin. Chem. Eng.* 22, 152–160.
- (14) Wirth, D., Gama-Norton, L., Riemer, P., Sandhu, U., Schucht, R., and Hauser, H. (2007) Road to precision: recombinase-based targeting technologies for genome engineering. *Curr. Opin. Biotechnol.* 18, 411–419.
- (15) Turan, S., Zehe, C., Kuehle, J., Qiao, J., and Bode, J. (2013) Recombinase-mediated cassette exchange (RMCE) — A rapidly-expanding toolbox for targeted genomic modifications. *Gene* 515, 1–27.
- (16) Wiberg, F. C., Rasmussen, S. K., Frandsen, T. P., Rasmussen, L. K., Tengbjerg, K., Coljee, V. W., Sharon, J., Yang, C.-Y., Bregenholt, S., Nielsen, L. S., Haurum, J. S., and Tolstrup, A. B. (2006) Production of target-specific recombinant human polyclonal antibodies in mammalian cells. *Biotechnol. Bioeng.* 94, 396–405.
- (17) Crawford, Y., Zhou, M., Hu, Z., Joly, J., Snedecor, B., Shen, A., and Gao, A. (2013) Fast identification of reliable hosts for targeted cell line development from a limited-genome screening using combined φ C31 integrase and CRE-Lox technologies. *Biotechnol. Prog.* 29, 1307–1315.
- (18) Zhang, L., Inniss, M. C., Han, S., Moffat, M., Jones, H., Zhang, B., Cox, W. L., Rance, J. R., and Young, R. J. (2015) Recombinase-mediated cassette exchange (RMCE) for monoclonal antibody expression in the commercially relevant CHOK1SV cell line. *Biotechnol. Prog.* 31, 1645–1656.
- (19) Inniss, M. C., Bandara, K., Jusiak, B., Lu, T. K., Weiss, R., Wroblewska, L., and Zhang, L. (2017) A novel Bxb1 integrase RMCE system for high fidelity site-specific integration of mAb expression cassette in CHO Cells. *Biotechnol. Bioeng.* 114, 1837–1846.
- (20) Carver, J., Ng, D., Zhou, M., Ko, P., Zhan, D., Yim, M., Shaw, D., Snedecor, B., Laird, M. W., Lang, S., Shen, A., and Hu, Z. (2020) Maximizing antibody production in a targeted integration host by optimization of subunit gene dosage and position. *Biotechnol. Prog.* 36, No. e2967.
- (21) Gaidukov, L., Wroblewska, L., Teague, B., Nelson, T., Zhang, X., Liu, Y., Jagtap, K., Mamo, S., Tseng, W. A., Lowe, A., Das, J., Bandara, K., Baijuraj, S., Summers, N. M., Lu, T. K., Zhang, L., and Weiss, R. (2018) A multi-landing pad DNA integration platform for mammalian cell engineering. *Nucleic Acids Res.* 46, 4072–4086.
- (22) O’Callaghan, P. M., Berthelot, M. E., Young, R. J., Graham, J. W. A., Racher, A. J., and Aldana, D. (2015) Diversity in host clone performance within a Chinese hamster ovary cell line. *Biotechnol. Prog.* 31, 1187–1200.
- (23) Noh, S. M., Shin, S., and Lee, G. M. (2018) Comprehensive characterization of glutamine synthetase-mediated selection for the establishment of recombinant CHO cells producing monoclonal antibodies. *Sci. Rep.* 8, 5361.
- (24) Pristovsek, N., Hansen, H. G., Sergeeva, D., Borth, N., Lee, G. M., Andersen, M. R., and Kildegaard, H. F. (2018) Using Titer and Titer Normalized to Confluence Are Complementary Strategies for Obtaining Chinese Hamster Ovary Cell Lines with High Volumetric Productivity of Etanercept. *Biotechnol. J.* 13, No. e1700216.
- (25) Ley, D., Seresht, A. K., Engmark, M., Magdenoska, O., Nielsen, K. F., Kildegaard, H. F., and Andersen, M. R. (2015) Multi-omic profiling of EPO-producing Chinese hamster ovary cell panel reveals metabolic adaptation to heterologous protein production. *Biotechnol. Bioeng.* 112, 2373–2387.
- (26) Pristovsek, N., Nallapareddy, S., Grav, L. M., Hefzi, H., Lewis, N. E., Rubgjerg, P., Hansen, H. G., Lee, G. M., Andersen, M. R., and Kildegaard, H. F. (2019) Systematic Evaluation of Site-Specific Recombinant Gene Expression for Programmable Mammalian Cell Engineering. *ACS Synth. Biol.* 8, 758–774.
- (27) Brown, A. J., Gibson, S. J., Hatton, D., and James, D. C. (2017) In silico design of context-responsive mammalian promoters with user-defined functionality. *Nucleic Acids Res.* 45, 10906–10919.
- (28) Kozak, M. (1987) At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J. Mol. Biol.* 196, 947–950.
- (29) Petersen, S. D., Zhang, J., Lee, J. S., Jakociunas, T., Grav, L. M., Kildegaard, H. F., Keasling, J. D., and Jensen, M. K. (2018) Modular 5'-UTR hexamers for context-independent tuning of protein expression in eukaryotes. *Nucleic Acids Res.* 46, No. e127.
- (30) Xu, Z.-L., Mizuguchi, H., Mayumi, T., and Hayakawa, T. (2003) Woodchuck hepatitis virus post-transcriptional regulation element enhances transgene expression from adenovirus vectors. *Biochim. Biophys. Acta, Gen. Subj.* 1621, 266–271.
- (31) Kim, K.-S., Kim, M. S., Moon, J. H., Jeong, M. S., Kim, J., Lee, G. M., Myung, P.-K., and Hong, H. J. (2009) Enhancement of recombinant antibody production in HEK 293E cells by WPRE. *Biotechnol. Bioprocess Eng.* 14, 633.
- (32) Yahata, K., Maeshima, K., Sone, T., Ando, T., Okabe, M., Imamoto, N., and Imamoto, F. (2007) cHS4 insulator-mediated alleviation of promoter interference during cell-based expression of tandemly associated transgenes. *J. Mol. Biol.* 374, 580–590.
- (33) Mutskov, V. J., Farrell, C. M., Wade, P. A., Wolffe, A. P., and Felsenfeld, G. (2002) The barrier function of an insulator couples high histone acetylation levels with specific protection of promoter DNA from methylation. *Genes Dev.* 16, 1540–1554.
- (34) Duportet, X., Wroblewska, L., Guye, P., Li, Y., Eyquem, J., Rieders, J., Rimchala, T., Batt, G., and Weiss, R. (2014) A platform for rapid prototyping of synthetic gene networks in mammalian cells. *Nucleic Acids Res.* 42, 13440–13451.
- (35) Lund, A. M., Kildegaard, H. F., Petersen, M. B. K., Rank, J., Hansen, B. G., Andersen, M. R., and Mortensen, U. H. (2014) A versatile system for USER cloning-based assembly of expression vectors for mammalian cell engineering. *PLoS One* 9, No. e96693.
- (36) Hommelsheim, C. M., Frantzeskakis, L., Huang, M., and Ülker, B. (2015) PCR amplification of repetitive DNA: a limitation to genome editing technologies and many other applications. *Sci. Rep.* 4, 5052.

- (37) Hindson, B. J., Ness, K. D., Masquelier, D. A., Belgrader, P., Heredia, N. J., Makarewicz, A. J., Bright, I. J., Lucero, M. Y., Hiddessen, A. L., Legler, T. C., Kitano, T. K., Hodel, M. R., Petersen, J. F., Wyatt, P. W., Steenblock, E. R., Shah, P. H., Bousse, L. J., Troup, C. B., Mellen, J. C., Wittmann, D. K., Erndt, N. G., Cauley, T. H., Koehler, R. T., So, A. P., Dube, S., Rose, K. A., Montesclaros, L., Wang, S., Stumbo, D. P., Hedges, S. P., Romine, S., Milanovich, F. P., White, H. E., Regan, J. F., Karlin-Neumann, G. A., Hindson, C. M., Saxonov, S., and Colston, B. W. (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal. Chem.* 83, 8604–8610.
- (38) Basu, A. S. (2017) Digital Assays Part I: Partitioning Statistics and Digital PCR. *SLAS Technol.* 22, 369–386.
- (39) Kallehauge, T. B., Li, S., Pedersen, L. E., Ha, T. K., Ley, D., Andersen, M. R., Kildegaard, H. F., Lee, G. M., and Lewis, N. E. (2017) Ribosome profiling-guided depletion of an mRNA increases cell growth rate and protein secretion. *Sci. Rep.* 7, 40388.
- (40) Sergeeva, D., Camacho-Zaragoza, J. M., Lee, J. S., and Kildegaard, H. F. (2019) CRISPR/Cas9 as a Genome Editing Tool for Targeted Gene Integration in CHO Cells. In *CRISPR Gene Editing. Methods in Molecular Biology* vol 1961 (Luo, Y., Ed.), pp 213–232, Humana Press, New York.
- (41) Cho, I. H., Lee, N., Song, D., Jung, S. Y., Bou-Assaf, G., Sosic, Z., Zhang, W., and Lyubarskaya, Y. (2016) Evaluation of the structural, physicochemical, and biological characteristics of SB4, a biosimilar of etanercept. *MAbs* 8, 1136–1155.
- (42) Harraghy, N., Calabrese, D., Fisch, I., Girod, P.-A., LeFourn, V., Regamey, A., and Mermod, N. (2015) Epigenetic regulatory elements: Recent advances in understanding their mode of action and use for recombinant protein production in mammalian cells. *Biotechnol. J.* 10, 967–978.
- (43) Klein, R., Ruttkowski, B., Knapp, E., Salmons, B., Günzburg, W. H., and Hohenadl, C. (2006) WPRE-mediated enhancement of gene expression is promoter and cell line specific. *Gene* 372, 153–161.
- (44) Schambach, A., Wodrich, H., Hildinger, M., Bohne, J., Kräusslich, H. G., and Baum, C. (2000) Context dependence of different modules for posttranscriptional enhancement of gene expression from retroviral vectors. *Mol. Ther.* 2, 435–445.
- (45) Mariati, Mariati, Ho, S. C. L., Yap, M. G. S., and Yang, Y. (2010) Evaluating post-transcriptional regulatory elements for enhancing transient gene expression levels in CHO K1 and HEK293 cells. *Protein Expression Purif.* 69, 9–15.
- (46) Dvir, S., Velten, L., Sharon, E., Zeevi, D., Carey, L. B., Weinberger, A., and Segal, E. (2013) Deciphering the rules by which 5'-UTR sequences affect protein expression in yeast. *Proc. Natl. Acad. Sci. U. S. A.* 110, E2792–801.
- (47) de Jongh, R. P. H., van Dijk, A. D. J., Julsing, M. K., Schaap, P. J., and de Ridder, D. (2020) Designing Eukaryotic Gene Expression Regulation Using Machine Learning. *Trends Biotechnol.* 38, 191–201.
- (48) Maricque, B. B., Chaudhari, H. G., and Cohen, B. A. (2019) A massively parallel reporter assay dissects the influence of chromatin structure on cis-regulatory activity. *Nat. Biotechnol.* 37, 90.
- (49) Kreiss, P. (1999) Plasmid DNA size does not affect the physicochemical properties of lipoplexes but modulates gene transfer efficiency. *Nucleic Acids Res.* 27, 3792–3798.
- (50) Wang, X., Kawabe, Y., Hada, T., Ito, A., and Kamihira, M. (2018) Cre-Mediated Transgene Integration in Chinese Hamster Ovary Cells Using Minicircle DNA Vectors. *Biotechnol. J.* 13, No. e1800063.
- (51) Garrick, D., Fiering, S., Martin, D. I. K., and Whitelaw, E. (1998) Repeat-induced gene silencing in mammals. *Nat. Genet.* 18, 56–59.
- (52) Shearwin, K., Callen, B., and Egan, J. (2005) Transcriptional interference - a crash course. *Trends Genet.* 21, 339–345.
- (53) Eszterhas, S. K., Bouhassira, E. E., Martin, D. I. K., and Fiering, S. (2002) Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. *Mol. Cell. Biol.* 22, 469–479.
- (54) Wang, X., Kawabe, Y., Kato, R., Hada, T., Ito, A., Yamana, Y., Kondo, M., and Kamihira, M. (2017) Accumulative scFv-Fc antibody gene integration into the hprt chromosomal locus of Chinese hamster ovary cells. *J. Biosci. Bioeng.* 124, 583–590.
- (55) Hansen, H. G., Přistovský, N., Kildegaard, H. F., and Lee, G. M. (2017) Improving the secretory capacity of Chinese hamster ovary cells by ectopic expression of effector genes: Lessons learned and future directions. *Biotechnol. Adv.* 35, 64–76.
- (56) Brewster, R. C., Weinert, F. M., Garcia, H. G., Song, D., Rydenfelt, M., and Phillips, R. (2014) The transcription factor titration effect dictates level of gene expression. *Cell* 156, 1312–1323.
- (57) Cruz, C., and Houseley, J. (2014) Endogenous RNA interference is driven by copy number. *eLife* 3, No. e01581.
- (58) Levitt, N., Briggs, D., Gil, A., and Proudfoot, N. J. (1989) Definition of an efficient synthetic poly(A) site. *Genes Dev.* 3, 1019–1025.
- (59) Yang, Z., Halim, A., Narimatsu, Y., Jitendra Joshi, H., Steentoft, C., Schjoldager, K. T.-B. G., Alder Schulz, M., Sealover, N. R., Kayser, K. J., Paul Bennett, E., Levery, S. B., Vakhrushev, S. Y., and Clausen, H. (2014) The GalNAc-type O-Glycoproteome of CHO cells characterized by the SimpleCell strategy. *Mol. Cell. Proteomics* 13, 3224–3235.
- (60) Kol, S., Kallehauge, T. B., Adema, S., and Hermans, P. (2015) Development of a VHH-Based Erythropoietin Quantification Assay. *Mol. Biotechnol.* 57, 692–700.
- (61) Brown, A. J., Gibson, S., Hatton, D., and James, D. C. (2018) Transcriptome-Based Identification of the Optimal Reference CHO Genes for Normalisation of qPCR Data. *Biotechnol. J.* 13, 1700259.