

RESEARCH ARTICLE OPEN ACCESS

A Flexible Hybrid Site-Specific Integration-Based Expression System in CHO Cells for Higher-Throughput Evaluation of Monoclonal Antibody Expression Cassettes

Alana C. Szkodny  | Kelvin H. Lee 

Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, Delaware, USA

Correspondence: Alana C. Szkodny (aszkodny@udel.edu) | Kelvin H. Lee (KHL@udel.edu)**Received:** 26 August 2024 | **Revised:** 20 December 2024 | **Accepted:** 31 December 2024**Funding:** This work was supported by funding from the National Science Foundation under grants 1624698 and 2100502, and in part by the financial assistance award 70NANB17H002 from the U.S. Department of Commerce National Institutes of Standards and Technology.**Keywords:** bioprocessing | cell line development | CHO cells | monoclonal antibodies | recombinase-mediated cassette exchange | site-specific integration

ABSTRACT

The implementation of site-specific integration (SSI) systems in Chinese hamster ovary (CHO) cells for the production of monoclonal antibodies (mAbs) can alleviate concerns associated with production instability and reduce cell line development timelines. SSI cell line performance is driven by the interaction between genomic integration location, clonal background, and the transgene expression cassette, requiring optimization of all three parameters to maximize productivity. Systematic comparison of these parameters has been hindered by SSI platforms involving low-throughput enrichment strategies, such as cell sorting. This study presents a recombinase-mediated cassette exchange (RMCE)-capable SSI system that uses only chemical selection to enrich for transgene-expressing RMCE pools in less than one month. The system was used to compare eight mAb expression cassettes containing two novel genetic regulatory elements, the *Azin1* CpG island and the Piggybac transposase 5' terminal repeat, in various orientations to improve the expression of two therapeutic mAbs from two genomic loci. Similar patterns of productivity and mRNA expression were observed across sites and mAbs, and the best performing cassette universally increased mAb productivity by 7- to 11-fold. This flexible system allows for higher-throughput comparison of expression cassettes from a consistent clonal and transcriptional background to optimize RMCE-derived cell lines for industrial production of mAbs.

1 | Introduction

Significant increases in the production capacity for therapeutic monoclonal antibodies (mAbs) in Chinese hamster ovary (CHO) cells have been made in part through improved workflows for the generation, selection, and isolation of high-producing cell lines containing mAb genes randomly integrated into the CHO genome, with the best performing systems now reaching 3–8 g/L consistently and reports of some processes reaching up to 10 g/L [1, 2]. However, the random nature of the gene insertions in these cell lines results in production instability driven primarily by copy

number loss and epigenetic silencing, necessitating lengthy clone screening procedures to identify optimal manufacturing cell lines that maintain their expression phenotype over extended culture [3–5].

The biopharmaceutical industry is now exploring site-specific integration (SSI)-based systems to mitigate these concerns by precisely inserting transgene cassettes at preselected genomic loci with known resistance to production instability. The current state-of-the-art SSI systems adopt a hybrid approach, using programmable nucleases like zinc finger nucleases or CRISPR/Cas9

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2025 The Author(s). *Biotechnology Journal* published by Wiley-VCH GmbH.

to insert a genetic landing pad (LP) at a chosen genomic locus to provide the capability for high-efficiency gene insertions using recombinases such as Flp/FRT, Cre/Lox, Bxb1, or Φ C31 [6–8]. Once a suitable LP-containing host cell line has been established, recombinase-mediated cassette exchange (RMCE) can quickly, efficiently, and reproducibly insert the transgene of choice into the chosen genomic locus, resulting in reduced population heterogeneity, improved production stability, and increased flexibility to study gene expression from a consistent clonal and transcriptional background [9–11].

Early work on SSI systems focused on identifying genomic “hotspots” through a variety of methods to find novel loci that can support high transgene expression and resist production instability [12–15]. However, the inclusion and orientation of genetic elements within transgene cassettes can influence expression even at favorable sites, requiring the careful design of expression vectors to maximize productivity. Evaluation of various Kozak sequences, promoters, insulators, and regulatory elements such as ubiquitous chromatin opening elements (UCOEs) and matrix attachment regions (MARs) in SSI systems have all demonstrated the benefits of cassette optimization, achieving up to 4-fold increases in transgene productivity [16–20]. Many systems have also included mechanisms to modulate gene copy number and ratio for multi-gene products such as mAbs, with linear increases in productivity observed with up to four inserted mAb gene copies [12, 14, 16, 21]. Despite these promising results, the unique LP designs, clone screening workflows, and pool enrichment approaches used in each system have made it difficult to translate observations across cell lines, integration locations, and products. In addition, pool enrichment has typically relied on low-throughput methods such as cell sorting, which has limited the throughput of these studies. While the best-performing industrial SSI systems for mAb expression can now reach 2–6 g/L across products [14, 22], these platforms have required lengthy screening procedures to identify prevalidated hotspots, engineered host cells to create stringent selection mechanisms, and complex LP systems to reach high titers, and the details of these processes are considered valuable intellectual property and are rarely disclosed [23–25].

This work presents the design, generation, and implementation of a novel LP/RMCE system in CHO-K1 host cells that can generate genotypically homogeneous RMCE pools in less than one month using only chemical selection for the rational comparison of gene expression cassettes from a consistent clonal and transcriptional background. By increasing throughput and reducing experimental timelines, this system allowed for the rapid comparison of up to eight different RMCE cargo designs for two different therapeutically relevant IgG₁ κ mAbs integrated at two genomic loci, *Rosa26* and *Dop1b*, for a total of 27 unique expression conditions. These cargo designs incorporated two recently identified novel genetic regulatory elements, the *Azin1* CpG island and the Piggybac transposase 5' terminal repeat, in varying copy numbers and orientations relative to the mAb genes [13]. One cargo design was identified that consistently resulted in productivity gains of 7- to 11-fold across integration sites and mAbs. Through characterization of these pools at the DNA and mRNA level, unique expression patterns were identified that suggest the underlying mechanisms driving increases in mAb production.

2 | Materials and Methods

2.1 | Cell Culture

All work described here was performed with a CHO-K1 cell line obtained from NIH NIAID. Suspension cell cultures were maintained in ActiPro medium (Cytiva) supplemented with 6 mM L-glutamine (Fisher Scientific) unless otherwise stated. All shaken cultures were grown at 37°C, 5% CO₂, and 80% humidity (Infors MultiTron, orbital diameter = 25.4 mm). Cultures in 96-deep well plates were maintained at 350 rpm with a 500 μ L working volume (Biotix, square-well, v-bottom), cultures in 24-deep well plates were maintained at 200 rpm with a 2 mL working volume (Axygen, square-well, v-bottom), and cultures in 50 mL spin tubes were maintained at 200 rpm with a 10 mL working volume (Chemglass NEST). Maintenance cultures were routinely passaged every 2–3 days in spin tubes. Static cultures were maintained at 37°C, 5% CO₂ (HeraCell). Cell counts were measured with a Vi-CELL XR cell viability analyzer (Beckman Coulter) via the trypan blue exclusion method.

2.2 | Molecular Cloning of LP and RMCE Cargo Cassettes

The LP gene cassette (Figure S1) was assembled using standard molecular cloning techniques and cloned into in-house plasmid backbones allowing for insertion of homology arms via Gibson assembly (*Rosa26*) or Golden Gate assembly (*Dop1b*) using protocols provided by New England Biolabs (NEB). LP elements were amplified from existing in-house plasmids and/or synthesized as gene fragments (gBlocks, Integrated DNA Technologies, IDT), and homology arms were amplified from CHO-K1 host genomic DNA (QIAamp DNA Mini kit, Qiagen). Oligos containing the *Dop1b* sgRNA sequence were annealed in TE buffer using an annealing protocol provided by IDT prior to assembly. RMCE cargo cassettes for the expression of monomeric teal fluorescent protein 1 (mTFP1, TFP), trastuzumab (TRMb), or adalimumab (ADMb) were built using a Golden Gate assembly approach as previously described [13, 26]. The coding sequences for the LC and HC of both mAbs were codon optimized for expression in CHO using the GenScript GenSmart Codon Optimization tool and purchased as synthesized gene fragments (IDT). All plasmids were maintained in TOP10 *Escherichia coli* (*E. coli*) (Thermo Fisher Scientific) and prepared using either QIAprep Spin Miniprep (Qiagen) or ZymoPURE Plasmid Miniprep (Zymo Research) kits according to the kit protocols. Plasmid sequences were verified using Sanger sequencing at the University of Delaware DNA Sequencing and Genotyping Center (UDSGC) prior to use.

2.3 | Cell Line Development

LP constructs were integrated into the genome of a CHO-K1 host cell line using Cas9-mediated integration. For integration at *Rosa26*, cells were transfected with a 1:4:2 molar ratio of Cas9:sgRNA:LP plasmid, using a hCas9 plasmid received from George Church (Addgene plasmid #41815), an in-house designed sgRNA [27], and LP plasmids described above. Cells were

transfected with 250 fmol of DNA per million cells using an SG Cell Line transfection kit (Lonza) on a 4D-Nucleofector unit (Lonza) using transfection code FF-137. Prior to transfection, DNA was aliquoted and dried in a SpeedVac (Savant) for 20 min at 45°C then resuspended in supplemented SG solution and allowed to rehydrate at 4°C overnight. Posttransfection, cells were selected with 8 µg/mL blasticidin S HCl (Gibco) starting on Day 5 and monitored every 3–4 days for growth, viability, and enhanced monomeric near-infrared fluorescent protein 670 (emiRFP670, RFP) expression [28]. Once cells recovered to > 95% viability and reached a constant RFP expression level, the selected pools were single-cell cloned into ClonaCell-CHO ACF semi-solid medium supplemented with ClonaCell-CHO ACF Supplement (Stem Cell Technologies) and 6 mM L-glutamine. The selected cell pool was diluted to 33 cells/mL in semi-solid medium, and 1.5 mL was added to each well of a 6-well plate to seed a total of 50 cells per well. This density was found to provide good separation between clones while maximizing the number of clones per well. Clones were allowed to expand under static conditions for 10–14 days. Approximately 100 colonies were manually picked and expanded into 96-deep well plates for RFP expression analysis and genomic DNA extraction using an in-house protocol (Supplemental Methods).

For integration at *Dop1b*, cells were transfected with a 1:2 molar ratio of Cas9:LP plasmids, using an in-house modified hCas9_D10A plasmid (a gift from George Church, Addgene plasmid #41816) that is compatible with the double-nick-donor CRISPR method [29]. This CRISPR method incorporates the LP and the sgRNA onto a single plasmid and has been shown to increase the frequency of on-target, SSI events. The site-specific sgRNA had been designed previously [13]. A basis of 5 µg of Cas9 plasmid DNA was used. Cells were transfected and selected as described above. Selected pools were single-cell cloned by single-cell deposition into 96-well plates using a Solentim VIPS (Advanced Instruments). Clones were expanded in XP Media CHO Growth A (Molecular Devices) supplemented with 6 mM L-glutamine under static conditions for 10–14 days. Colonies were monitored for growth with a Solentim Cell Metric X (Advanced Instruments), and approximately 100 colonies with significant expansion and verified as clonal were selected and expanded in 96-deep well plates for analysis as described above.

For both sites, RFP expression in each expanded clone was monitored with flow cytometry, and clones with appreciable outgrowth and an RFP-positive population greater than 80% were further expanded into 24-deep well plates. Clones that maintained their RFP fluorescence (10–20 clones per site) were tested with diagnostic junction PCRs from extracted genomic DNA to verify on-target LP integrations using four primer sets targeting the left homology arm, right homology arm, LP, and wild-type locus. Primer sets for the left and right homology arms were designed such that the amplified region included genomic DNA outside the homology arm and sequence within the LP to verify complete integration of the LP in the correct orientation. The LP primer set amplified a region including parts of all three genes to verify complete integration of the LP. The presence of an amplicon at the WT locus implied a single-copy integration. The PCRs were performed with either GoTaq Green (Promega) or NEBNext Ultra II Q5 Master Mix (NEB) using custom primers

(Table S1). Clones with correct diagnostic junction PCRs (5 clones at *Rosa26* and 10 clones at *Dop1b*) were tested for copy number using ddPCR as described below.

2.4 | RMCE Pool Generation and Batch Analysis

Stable RMCE pools were generated by co-transfecting the appropriate RMCE cargo plasmid with the pCAG-NLS-HA-Bxb1 plasmid (a gift from Pawel Pelczar, Addgene plasmid #51271) at a 1:6 molar ratio of Bxb1:cargo. Transfections were performed as described for cell line generation, using 2×10^6 cells per transfection and 250 fmol total DNA per million cells. Experiments using the mAb cargo constructs also included a “Null” transfection condition, which used the backbone RMCE cargo plasmid only for exchange, and a TFP cargo control to monitor expression of RMCE+ pools. All conditions were transfected in biological triplicate and maintained in 24 deep-well plates. On Day 4 posttransfection, cells were passaged to 1×10^6 cells/mL with the addition of 500 µg/mL geneticin (Gibco) and passaged every 3–4 days. On Day 11 posttransfection, cultures were split, with one culture continuing with geneticin selection and one culture receiving both geneticin and ganciclovir (GCV) (1 µM, Sigma) treatment. Cells on both selection agents were passaged every 3–4 days until Day 18, when all selection agents were removed.

On Day 14 posttransfection, cells receiving geneticin selection only were seeded at 0.5×10^6 cells/mL in 24 deep-well plates in the absence of selection for batch analysis. On Day 4 of the batch culture, cultures were harvested via centrifugation, and supernatant samples were retained for titer analysis. Genomic DNA and mRNA were extracted from the remaining cell pellets. mRNA was extracted using a Quick RNA-96 kit (Zymo Research) with on-column DNase digestion according to Zymo protocol. Genomic DNA was extracted with an in-house protocol (Supplemental Methods). Productivity (qP, pg/cell-day) for each culture was calculated by dividing the final Day 4 titer by the integral of viable cell density (IVCD), with IVCD determined using a trapezoidal approximation from Day 0 to 4 cell counts.

2.5 | Titer Determination

Harvested supernatant samples were diluted 2-fold in Dilution Buffer (Dulbecco’s phosphate buffered saline [PBS] + 0.1% bovine serum albumin) for titer analysis on an Octet 96e (Sartorius). Antibody standards were generated by diluting a 700 µg/mL antibody standard (Sartorius) to 50 µg/mL with Dilution Buffer + 50% ActiPro, then making 2× serial dilutions to 0.391 µg/mL. Protein A biosensors (Sartorius) were hydrated in dilution buffer + 50% ActiPro overnight before use. All samples and standards were run in technical triplicate in black 96-well polypropylene plates (Corning) using a high-sensitivity method (5-min read time, 1000 rpm, 30°C). In between samples, probes were regenerated in 10 mM glycine, pH 1.5, and neutralized in Dilution Buffer + 50% ActiPro. Binding rates for standards were fit to a “Dose Response-4PL Unweighted” equation, which was then used to determine sample concentrations.

2.6 | Droplet Digital PCR Analysis for Gene Copy Number and Gene Expression Analysis

Gene copy number and mAb gene expression were determined using custom multiplexed assays (Table S1) on a QX ONE ddPCR instrument (Bio-Rad). All primer-probe assays were purchased premixed at a primer:probe ratio of 3.6:1 (900 nM:250 nM, IDT). For gene copy number analysis, genomic DNA was digested with CviQ1 (LP clones) or HindIII (RMCE pools) overnight (NEB). Samples were analyzed using a ddPCR Supermix for Probes (No dUTP) kit (Bio-Rad) according to the kit protocol with *HPRT1* as a copy number reference ($2n = 1$). Gene expression of the mAb light chain and heavy chain genes was determined with a One-Step ddPCR Advanced Kit for Probes (Bio-Rad) according to the kit protocol. Approximately 0.25–0.5 ng of total mRNA was used per reaction. Antibody gene expression was normalized to the expression of *Rab10*. All mRNA reactions were run in technical duplicates.

2.7 | Flow Cytometry

Flow cytometry to measure cell fluorescence was performed on an Accuri C6 Plus with a CSampler Plus attachment (BD). Cells were diluted 2× in PBS prior to analysis. Healthy, unedited CHO-K1 cells were used to establish a cell gate using a forward scatter (FSC-A) versus side scatter (SSC-A) plot. Cells falling within that gate were further gated for singlets on an SSC-A versus SSC-H plot. Gates for fluorescent-positive cells were defined as the areas above the unedited cell populations using SSC-A versus FL-1 or FL-4 plots, where FL-1 corresponds to a 488 nm laser, 510/10 filter channel to detect TFP, and FL-4 corresponds to a 640 nm laser and a 675/25 filter channel to detect RFP. These fluorescent-positive gates contained < 1% of the healthy, unedited cell population at the time of definition. A minimum of 10,000 events were analyzed whenever possible.

2.8 | Whole Genome Resequencing

Whole genome resequencing, including genomic DNA extraction, library generation, and PacBio long-read sequencing, was performed by the UDSGC on a PacBio Sequel IIe instrument using two SMRTCells. Bioinformatics analysis was performed by the University of Delaware Center for Bioinformatics and Computational Biology (CBCB) Bioinformatics Core. Reads were mapped to the Chinese hamster PICRH genome assembly CriGri-PICRH-1.0, GCF_003668045.3 [30].

2.9 | Statistical Analyses

All statistical analyses were performed and plots were generated using R (v 4.0.4) and the tidyverse family of packages (v 1.3.0). Unless otherwise noted, error bars represent one standard deviation from the mean of n biological replicate cultures, with n indicated in figure captions. The statistical tests used are provided in the figure captions or main text, with $p < 0.05$ considered significant for all tests.

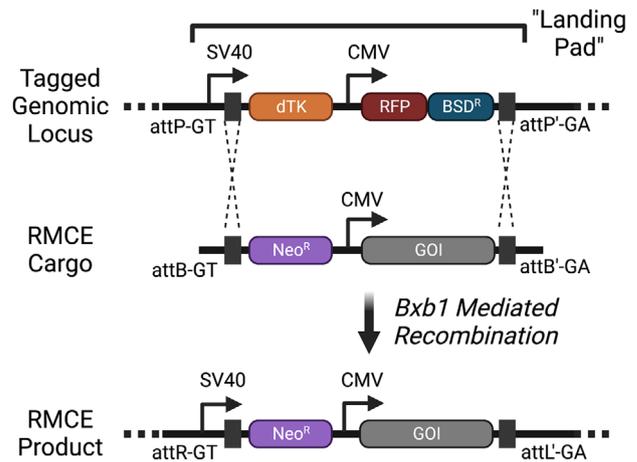


FIGURE 1 | Simplified schematic of the LP/RMCE system used. The landing pad contains a destabilized thymidine kinase (dTK) and an emiRFP670 (RFP)/blasticidin resistance (BSD^R) gene fusion, all flanked by orthogonal Bxb1 *attP*-GT/*attP*'-GA recognition sites. This construct was integrated at known genomic loci using CRISPR/Cas9. RMCE cargo cassettes, containing a neomycin resistance gene (Neo^R) and a gene of interest (GOI) flanked by orthogonal Bxb1 *attB*-GT/*attB*'-GA recognition sites, can be introduced into landing pad-containing cells, where the Bxb1 integrase mediates the removal of the landing pad and insertion of the cargo. Polyadenylation signals are not shown for simplicity. CMV, cytomegalovirus promoter; SV40, simian virus 40 promoter; RMCE, recombinase-mediated cassette exchange.

3 | Results

3.1 | Establishment of LP Cell Lines Enables Rapid Gene Insertion Via RMCE

The LP/RMCE system used in this study consists of various reporter and selection genes for establishing LP host cell lines and subsequent RMCE pools (Figure 1). The LP gene cassette (Figure S1) includes an SV40 promoter-driven thymidine kinase (TK) gene with a destabilization tag attached to the C-terminus to reduce protein half-life, ensuring that cells undergoing an RMCE event are not selected against due to TK protein persistence following LP removal [31]. The TK gene is followed by an emiRFP670 (RFP)—blasticidin S deamidase (BSD) fusion gene driven by a CMV promoter [28]. Recombinase recognition sites (wild-type *attP*-GT and mutant *attP*'-GA) for RMCE using the serine recombinase Bxb1 are included 3' of the SV40 promoter and 3' of the RFP-BSD^R polyA tail [32]. RMCE cargo cassettes include a neomycin resistance gene (Neo^R) lacking a promoter and a gene (or genes) of interest (GOI) driven by a CMV promoter, all flanked by *attB*-GT/*attB*'-GA recombinase recognition sites. The positioning of the 5' *attP*-GT site downstream of the SV40 promoter creates a “promoter trap” that, combined with the use of two orthogonal recombinase recognition sites, promotes insertion of the RMCE cargo in the correct orientation [32–34]. The RFP-BSD fusion protein allows for monitoring of LP-positive cells with flow cytometry and enrichment of LP-positive cells by treatment with blasticidin. Expression of the TK gene in LP-positive cells confers sensitivity to the pro-drug GCV, which is converted to a cytotoxic product by the TK enzyme. This gene acts as a negative selection marker to eliminate LP-positive cells post-

RMCE. The Neo^R gene provides resistance to geneticin, which is used as a positive selection marker for cells that undergo a correctly oriented, on-target RMCE event.

The LP cassette was integrated into the genome of a CHO-K1 host cell line using CRISPR/Cas9-mediated insertion at two preselected genomic loci, *Rosa26* and *Dop1b*. The *Rosa26* locus is a well-known “safe harbor” in human cells [35, 36], and has been shown to support stable expression of transgenes in CHO [12]. High expression from the *Dop1b* locus was recently identified through a high-throughput pooled screen of hotspots in CHO using Piggybac transposase-mediated insertion [13]. Postintegration, cell pools were selected with blasticidin to enrich for LP-positive cells, then single-cell cloned to isolate clones with single-copy, on-target insertions of the LP. RFP expression was monitored using flow cytometry for approximately 100 expanded clones from each pool. Clones that maintained high levels of RFP expression were expanded further and tested for on-target integration with diagnostic PCRs across the genome/LP junctions, within the LP, and across the wild-type genomic locus, as well as copy number verification of the TK and RFP-BSD genes using ddPCR. Five clones for *Rosa26* and seven clones for *Dop1b* were identified to have single-copy, on-target integrations in the correct orientation. From these identified clones, Clone C9 (R26-C9) was chosen for LP expression at *Rosa26*, and clone P1C11 (D1b-P1C11) was chosen for LP expression at *Dop1b* based on their high RFP expression and appreciable growth rate (Figure S2).

To verify the RMCE capabilities of each of these cell lines, each clone was co-transfected with an mTFP1-expressing cargo plasmid and a plasmid expressing Bxb1 in biological triplicate. Cells were selected with geneticin to enrich for RMCE+ cells, and further selected with GCV to eliminate any remaining cells containing the LP. RFP and TFP expression were monitored throughout this process, with a progressive enrichment of TFP-expressing cells and depletion of RFP-expressing cells observed throughout the course of selection (Figure 2). The population of RFP+/TFP+ cells observed early in the process is due to transient expression of TFP from the cargo plasmid and the persistence of RFP in cells. After 21 days posttransfection, pools contained > 90% TFP-expressing cells and < 1% RFP-expressing cells for both cell lines. We hypothesize that the slight increase in the TFP-/RFP- population after Day 21 arises from a small population of low-expressing cells whose TFP levels decrease back to baseline levels after selection is removed. The biological replicates showed nearly identical selection profiles, suggesting that RMCE is a highly reproducible process. Copy number analysis for an R26-C9 pool showed single-copy insertions of the TFP gene, with almost complete removal of the LP genes (Figure S3).

3.2 | Genetic Regulatory Elements and mAb Gene Positioning in RMCE Cargo Cassettes Increase mAb Productivity and mRNA Transcription

To optimize the RMCE cargos to support high mAb productivity, two previously identified genetic regulatory elements, the upstream CpG island from the *Azin1* locus (CpG) and the Piggybac transposase 5' inverted terminal repeat (PB5TR), were included in expression cassettes to study their effect on mAb gene expression [13, 37]. Elements were placed upstream and/or

internal to the LC and HC genes, with one cassette including a second copy of the PB5TR. The order of mAb genes was also varied to determine any position effects. Trastuzumab and adalimumab were selected as model IgG1 κ mAbs based on their clinical relevance, amino acid sequence similarity, and the availability of structural data. Trastuzumab is also considered an “easy-to-express” mAb [38]. These eight cargo designs (Figure 3) were created for trastuzumab and adalimumab expression, and then co-transfected into R26-C9 and D1b-P1C11 cell lines with the Bxb1 expression plasmid in biological triplicate. The RMCE pools were selected for 10 days on geneticin and regularly analyzed with flow cytometry to monitor gene exchange through loss of RFP expression (Figure S4). Pools were then tested for expression in a rapid, small-scale 4-day batch process in 24 deep-well plates without selection pressure during production. All pools showed similar growth rates and high viabilities (> 95%) throughout culture (Figure S5A and S5C), suggesting that the observed differences in volumetric titer were due to differences in cell-specific productivity (Figure S5B and S5D).

For trastuzumab expressed from *Rosa26* (Figure 4A, Figure S5B), low but measurable expression was observed with the baseline constructs lacking any regulatory elements (A and B). The introduction of elements upstream of the mAb genes (E and F) led to an increase in productivity, with further improvement achieved by adding a second copy of the PB5TR internal to the mAb genes (G and H). The highest productivity was observed for the constructs containing both elements in the internal position (C and D), representing more than 8-fold increases in qP from the equivalent baseline constructs. These trends were observed for both “LC-first” (cargos A, C, E, and G) and “HC-first” (cargos B, D, F, and H) gene orientations, although all HC-first constructs showed lower expression as compared to their LC-first counterparts. Construct C resulted in the highest productivity of 0.69 pg/cell-day from the small-scale batch culture. Previous data for trastuzumab expression with these constructs at *Dop1b* suggested that LC-first cargos were preferred at this site [13], so only cargos A, C, and G were tested in the D1b-P1C11 cell line (Figure 4B, Figure S5B). Construct A resulted in the lowest productivity, with progressive increases observed with constructs G and C, respectively, with construct C increasing productivity by more than 7-fold compared to construct A. However, overall productivities in D1b-P1C11 were 50%–60% lower than the equivalent condition in R26-C9, as evidenced by construct C, which also resulted in the highest productivity in D1b-P1C11 but only reached an absolute productivity of 0.35 pg/cell-day in small-scale batch cultures.

mRNA expression analysis for the trastuzumab HC and LC genes showed a strong positive correlation between transcription and productivity, with low productivity constructs showing 1–2 \times higher gene expression relative to *Rab10*, and the highest-expressing constructs at *Rosa26* showing 5–10 \times higher gene expression relative to *Rab10* (Figure 4C). The lower productivity observed at *Dop1b* can be attributed to the overall lower transcription observed at this site as compared to *Rosa26*, with transcripts expressed approximately 2-fold less than the corresponding conditions at *Rosa26* and only 0.5–3.5 \times greater expression than *Rab10*. In all cases across both sites, the HC transcript was expressed higher than the LC transcript, with the positioning of the mAb genes playing the largest role in

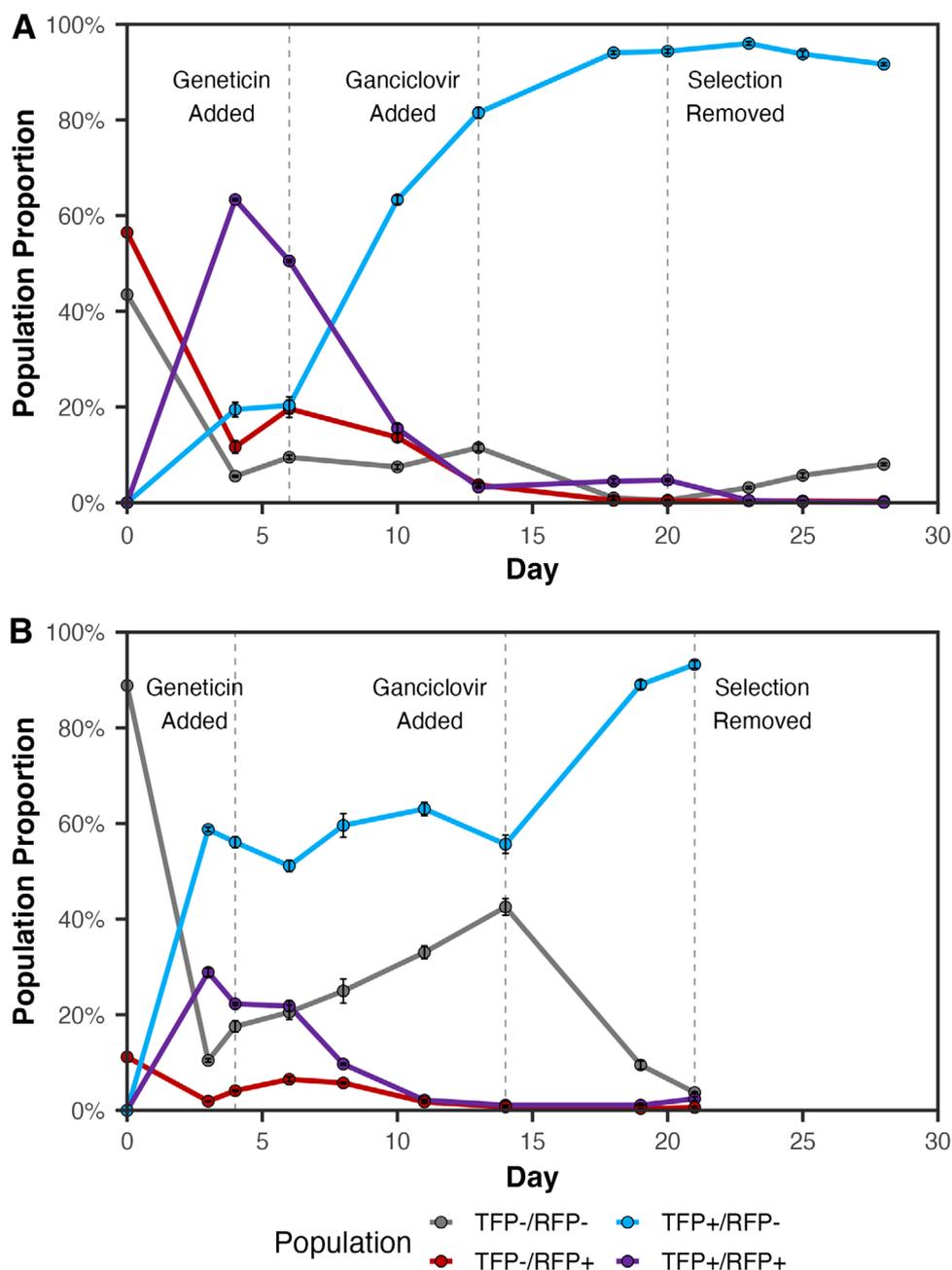


FIGURE 2 | Example selection curves for the LP/RMCE system. Plots show the exchange of the emiRFP670 (RFP)-containing landing pad with an RMCE cargo construct expressing mTFP1 (TFP) at (A) *Rosa26* (R26-C9) and (B) *Dop1b* (D1b-P1C11). Cells were selected with 500 $\mu\text{g/mL}$ of geneticin starting 4–6 days posttransfection, then 1 μM of ganciclovir was added 7–10 days later. After one week of selection with both chemicals, all selection was removed. Population phenotypes are defined as the presence (+) or absence (–) of fluorescence as measured by flow cytometry. Error bars represent one standard deviation from the mean of biological triplicates. RMCE, recombinase-mediated cassette exchange.

determining the relative transcript ratio between the two genes. Expressing the HC upstream of the LC led to increased HC transcription at the expense of LC transcription, increasing the HC:LC transcript ratio, which in turn reduced productivity. At both sites, construct C resulted in the highest gene transcription, which likely drove the high productivities observed with this construct.

Overall productivities for adalimumab were lower than for trastuzumab at both genomic locations by more than 3-fold, but the relative impact of the regulatory elements on mAb

expression was similar across mAbs (Figure S5D). For the LC-first constructs, the same rank order ($A < E < G < C$) was observed; however, very large differences in productivity were seen between the corresponding HC-first constructs (Figures 5A and B). Three of the four HC-first constructs had titers outside of the assay standard curve range. As was seen with trastuzumab, cargos containing internal regulatory elements (C and D) resulted in the best overall productivity, with construct C reaching small-scale batch productivities of 0.23 pg/cell-day in R26-C9 and 0.19 pg/cell-day in D1b-P1C11. mRNA expression analysis of LC and HC genes showed that, contrary to trastuzumab,

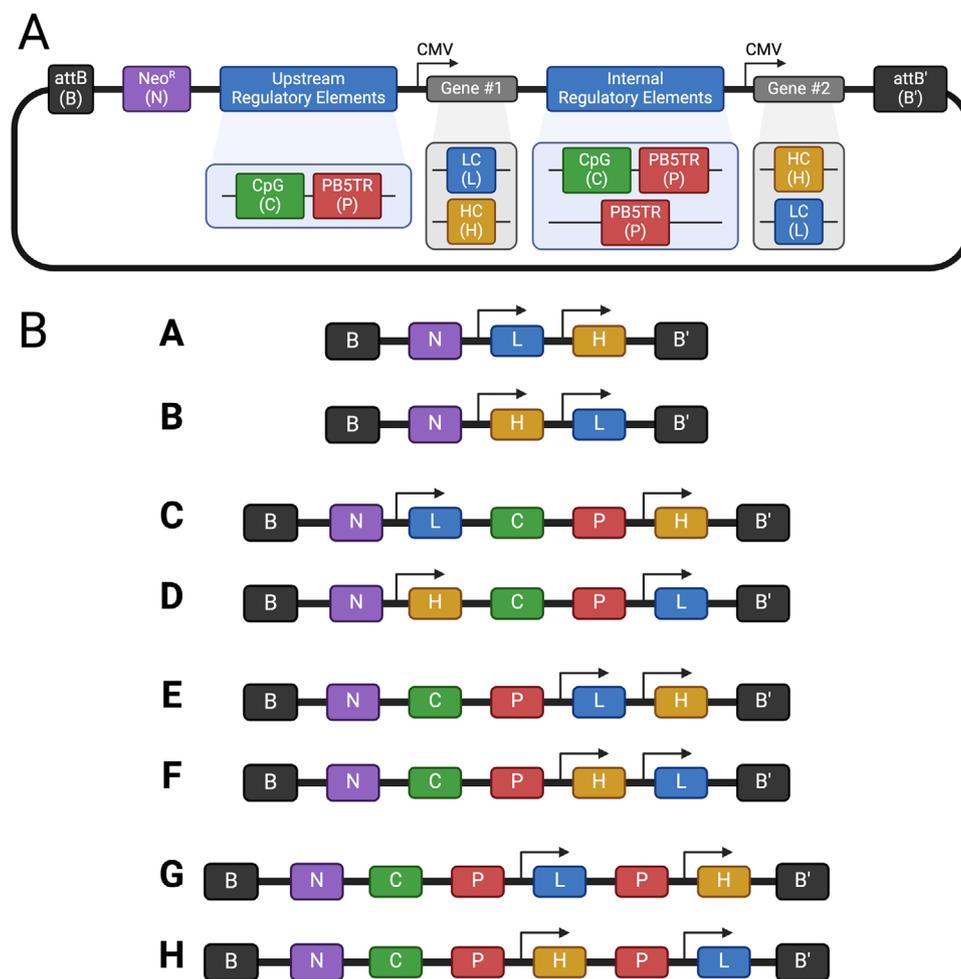


FIGURE 3 | RMCE cargo plasmid designs used for optimization of mAb expression. (A) A general schematic of the dual-gene RMCE cargo plasmids, which include genetic regulatory elements placed either upstream or internal to the mAb genes. Both mAb genes used a SV40 late poly(A) tail, and the neomycin resistance gene used an synthetic poly(A) tail (not shown). (B) Simplified schematics of each cargo used for expression optimization, using a subset of designs presented in Hilliard and Lee, 2023. For each combination of genetic regulatory elements, cargos were designed with the LC or HC in either the Gene #1 or Gene #2 position. Plasmid elements include: Bxb1 *attB*-GT recognition site (B), neomycin resistance gene (Neo^R , N) the upstream genomic CpG island from the *Azin1* locus (CpG, C), the Piggybac transposase 5' terminal repeat (PB5TR, P), cytomegalovirus promoter (CMV, arrows), mAb light chain (LC, L), mAb heavy chain (HC, H), and Bxb1 *attB'*-GA recognition site (B'). RMCE, recombinase-mediated cassette exchange.

LC-first adalimumab constructs expressed higher amounts of LC transcripts compared to HC transcripts, with the trend reversed for HC-first constructs, suggesting that excess HC mRNA is detrimental to expression (Figures 5C and D). Again, constructs that led to higher productivities also had higher overall levels of transcription, but overall transcription levels remained low, with the best performing conditions only resulting in 1.5–2.5× higher gene expression than *Rab10*. Interestingly, construct C was the best performing condition at both sites for trastuzumab and adalimumab expression, but the transcription trends leading to high expression were mAb-dependent.

3.3 | Copy Number Analysis and Whole Genome Resequencing Reveals Distribution of Genotype Patterns in RMCE Pools

The copy number of the LC and HC genes in the RMCE pools was determined using ddPCR, with *HPRT1* as a single-copy housekeeping gene reference. Copy numbers were greater

than one for all pools generated, with many pools showing approximately two copies of each gene but others reaching copy numbers as high as five (Figures 6 and 7). The differences in copy number between construct designs could not account for the observed increases in mRNA expression and productivity, as evidenced from the HC-first adalimumab pools that show some of the highest copy numbers (between two and five) but failed to produce quantifiable amounts of antibody. In general, copy numbers were more variable in the D1b-PIC11 cell line and in pools expressing adalimumab. A geneticin- and GCV-selected pool expressing trastuzumab from construct C at *Rosa26* was sequenced using PacBio long-read sequencing to characterize the genotypes present. While most reads contained evidence of the desired, single-copy, on-target insertion, some low-frequency instances of partial recombination, plasmid duplications, and random insertions were evident but were not the dominant genotype (Figure S6 and Supplemental Data). While these situations of complex, multi-copy integrations are rare, their presence may result in an increased apparent copy number in the pool, depending on the frequency of occurrence.

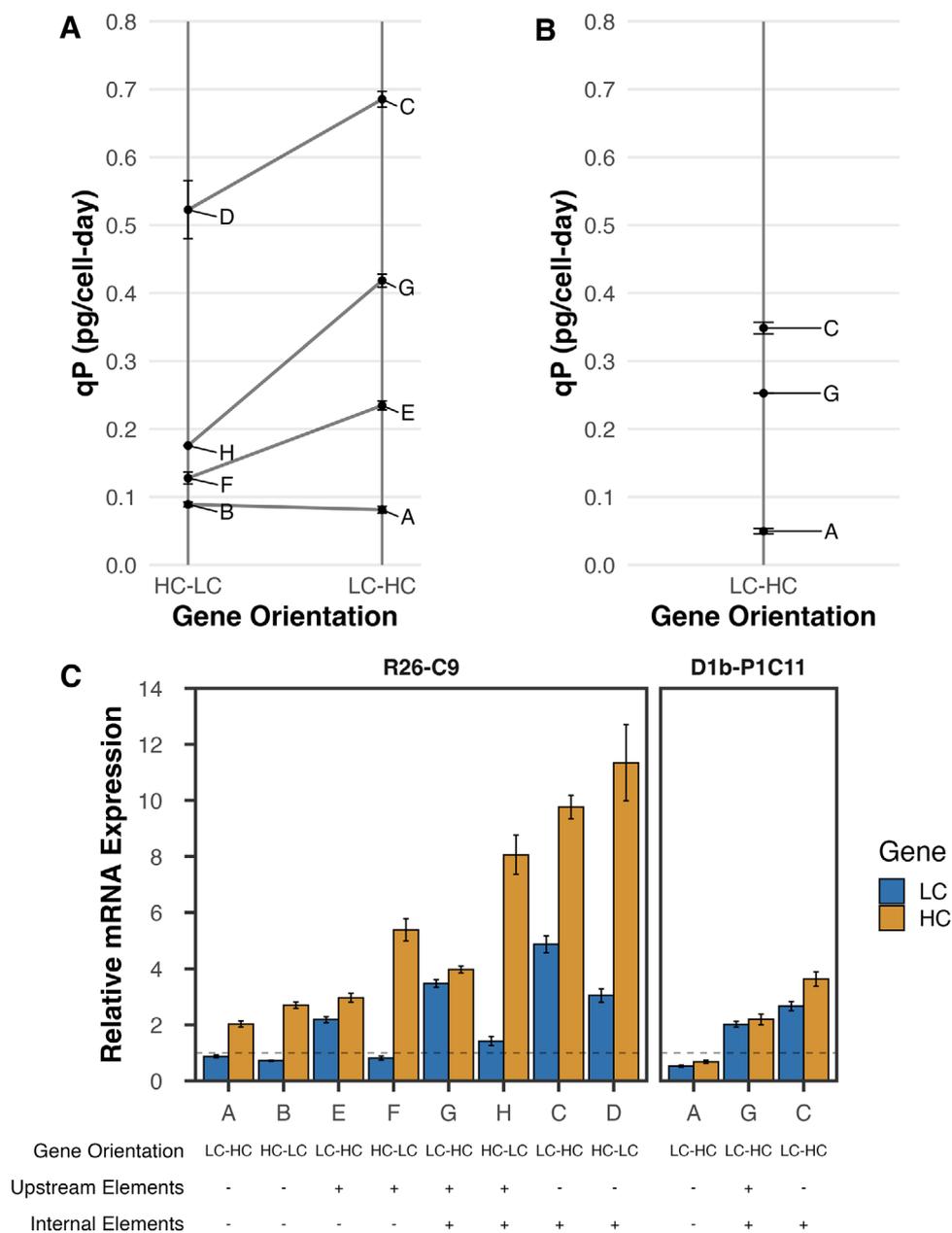


FIGURE 4 | Trastuzumab productivity and mRNA expression for RMCE cargo designs across genomic loci. Trastuzumab productivity is expressed from cargos integrated at (A) *Rosa26* and (B) *Dop1b*; (C) mRNA expression of the LC and HC at both genomic loci, with cargo pairs plotted in the order of increasing productivity. mRNA expression was measured using multiplexed ddPCR assays in technical duplicate, and expression was calculated relative to *Rab10*, as indicated by the dotted line. Error bars indicate one standard deviation from the mean of biological triplicate RMCE pools. RMCE, recombinase-mediated cassette exchange.

4 | Discussion

The hybrid LP/RMCE system reported here builds upon existing targeted integration systems and includes features that improve the identification of LP-positive host cells and selectively enrich for cells with on-target RMCE events. Following the establishment of an LP host cell line, the efficient, rapid, and reproducible nature of RMCE allows for pools to be generated in less than one month. In addition, the use of chemical selection agents only makes the system amenable to higher throughput experimental designs using small-scale (24 deep-well plate) culture systems, which enabled the systematic comparison of up to eight mAb

expression cassettes at two genomic loci for two therapeutically relevant mAbs. In this work, 81 cell pools were generated to test 27 distinct conditions in biological triplicate. Evaluating these cassettes from a consistent clonal background in pools brings significant benefits as compared to previous work with random integration, where lengthy single-cell cloning and clone evaluation are necessary to account for clonal heterogeneity and can dramatically reduce throughput.

The comparison across mAbs and cell lines in this study highlighted striking similarities between the patterns of productivity gains achieved for each RMCE cassette design. For all

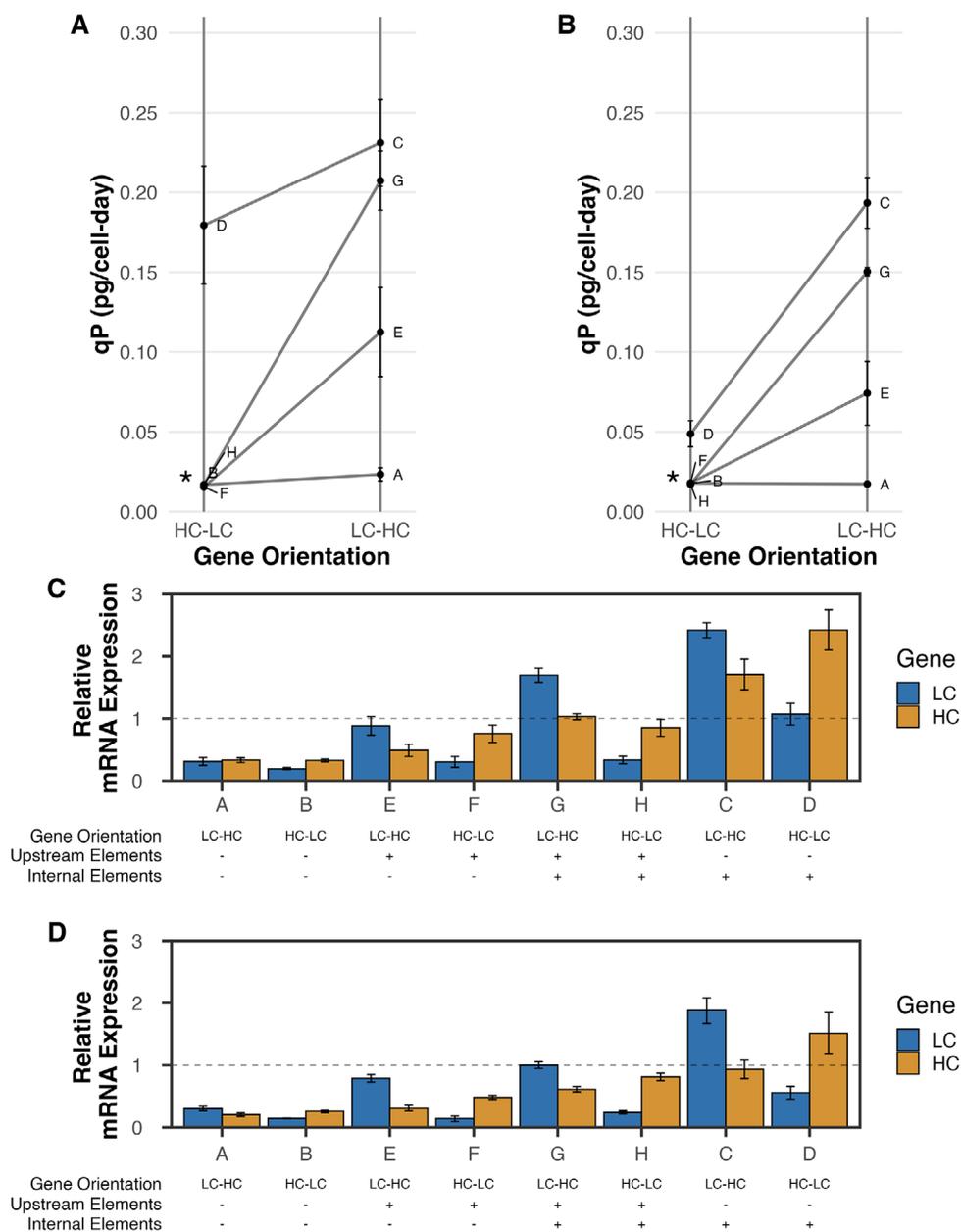


FIGURE 5 | Adalimumab productivity and mRNA expression for RMCE cargo designs across genomic loci. Adalimumab productivity is expressed from cargos integrated at (A) *Rosa26* and (B) *Dop1b*; mRNA expression of the LC and HC at (C) *Rosa26* and (D) *Dop1b*, with cargo pairs plotted in the order of increasing productivity. mRNA expression was measured using multiplexed ddPCR assays in technical duplicate, and expression was calculated relative to *Rab10*, as indicated by the dotted line. Error bars indicate one standard deviation from the mean of biological triplicate RMCE pools. Conditions B, F, and H (marked with asterisks) had titer measurements outside the range of the Octet standard curve. RMCE, recombinase-mediated cassette exchange.

trastuzumab conditions and all adalimumab LC-first conditions, inclusion of the CpG/PB5TR elements upstream of the mAb genes led to a 3–5× improvement in productivity, with another 2× improvement obtained by adding a second copy of the PB5TR in between the mAb genes. Moving both regulatory elements in between the two genes resulted in the greatest gains in productivity, with a 7–11× increase compared to the baseline constructs. For adalimumab HC-first constructs at both sites, only construct D resulted in titers greater than the Octet's limit of detection. In all cases, LC-first constructs gave higher titers than the equivalent HC-first cassettes. These results show that the HC:LC transcript ratio plays an equally important role to overall

transcription in determining productivity and supports previous observations that productivity is primarily driven by HC mRNA and translation rate [39, 40]. However, many high-expressing mAb systems are also characterized by excess LC polypeptide, which has been predicted to drive efficient folding and assembly of HCs [41]. Work in transient [41, 42] and stable [21, 43] expression systems has found that modulating LC expression to reach LC:HC polypeptide ratios of greater than one can increase titers. While polypeptide content was not directly measured in this work, the consistent observation of HC:LC transcript ratios greater than one across all trastuzumab conditions hints that LC may be limiting in these cultures, but sufficient overall

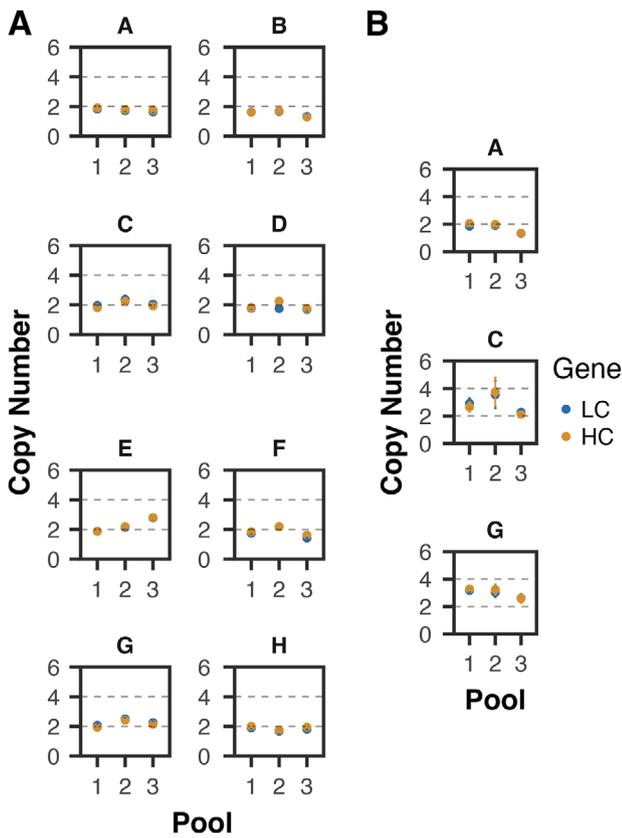


FIGURE 6 | Copy number of trastuzumab-expressing pools at (A) *Rosa26* and (B) *Dop1b* as measured by multiplexed ddPCR assays, using *HPRT1* as a reference gene. The construct identifier is shown above each plot, with the copy number for each biological replicate pool shown. Error bars represent the Poisson error for each pool, as calculated by the QX ONE software.

expression can still support high titers. The underlying cause of these elevated transcript ratios may lie in features inherent to the DNA sequences introduced during codon optimization, which can influence mRNA stability and translation rate. Further engineering of the trastuzumab coding sequences with the goal of achieving a HC:LC transcript ratio below one may provide useful information about the interaction between transgene DNA sequence and expression. Although the transcription pattern for adalimumab more closely mimicked previous results, with HC:LC transcript ratios of less than one being necessary for measurable expression, the low mRNA expression of both chains also suggests the presence of DNA sequence features that limit transcription or intracellular responses to protein production stress that led to downregulation of LC and HC transcription.

While relative productivity differences were similar across conditions, significant differences in absolute productivity were observed across cell lines and products, with D1b-PIC11 showing consistently lower productivity for both mAbs, and adalimumab showing lower productivity than trastuzumab. The differences between the cell lines emphasize the importance of clone selection when choosing a LP master host, as the absolute performance of RMCE pools is inherently limited by the clonal background. It has been shown previously that even genotypically

homogeneous cell pools can still show high levels of phenotypic heterogeneity, and the stochastic nature of gene expression may necessitate multiple rounds of sorting and cloning to isolate cells with a permanently shifted phenotype toward high protein expression [44, 45]. The selection and reporter systems used for generating stable clones can also impact phenotypic distributions, as evidenced by differences observed in gene expression upon pool enrichment with different antibiotics [46, 47]. As a result, it is crucial to choose a stringent selection system that effectively selects for not just on-target LP integrations, but also cells with high protein expression capacity, which has been achieved by using a surface-displayed reporter protein, as these proteins follow the same folding and secretion pathways as mAbs. Well-characterized membrane-anchored markers such as CD4 have found recent use [13, 48], with other work using generic mAb expression cassettes for initial identification of high-producing genomic loci, then employing a “tag-and-exchange” approach to generate the LP cell line directly from these promising clones [14, 15].

It is especially interesting to compare the results obtained here with the D1b-PIC11 cell line and those reported by Hilliard and Lee, who generated their LP cell line at *Dop1b* in the same CHO-K1 host cell line used for this work, but with a different selection mechanism, reporter gene, and single-cell cloning procedure [13]. The cell line used by Hilliard achieved an approximately 3-fold higher productivity for the expression of trastuzumab using construct C in an identical culture format. While the *Dop1b* locus does have the potential to drive high transgene expression, the difference in performance between the D1b-PIC11 cell line studied here and the cell line generated by Hilliard demonstrates the clonal variation possible at this site. Characterization of RMCE cargo construct performance in multiple LP clones could help further characterize this variation. This observation also emphasizes that high-producing cell lines are not created by optimal integration locations alone and may require multi-faceted approaches designed to select individual clones with high protein production potential as defined by many complex and interacting factors.

5 | Conclusion

This work presents an LP/RMCE system capable of rapid and reproducible insertion of transgenes at defined genomic loci for the systematic comparison of expression cassettes from a consistent clonal background. The use of chemical selection methods only for RMCE pool enrichment allows for higher throughput experimentation and eliminates the need for lengthy and low-throughput clone screening workflows, as would be necessary with random integration or de novo Cas9-mediated insertion methods. The flexibility and utility of this system were demonstrated through the evaluation of 27 distinct conditions for expression optimization of two therapeutic mAbs from two genomic loci using various orientations of two novel genetic regulatory elements, the *Azin1* CpG island and the Piggybac 5' ITR, within the mAb expression cassettes. A universal improvement in mRNA expression and productivity was observed from the inclusion of these regulatory elements, with 7- to 11-fold increases in productivity achieved when both elements are placed in between the LC and HC genes. These results may

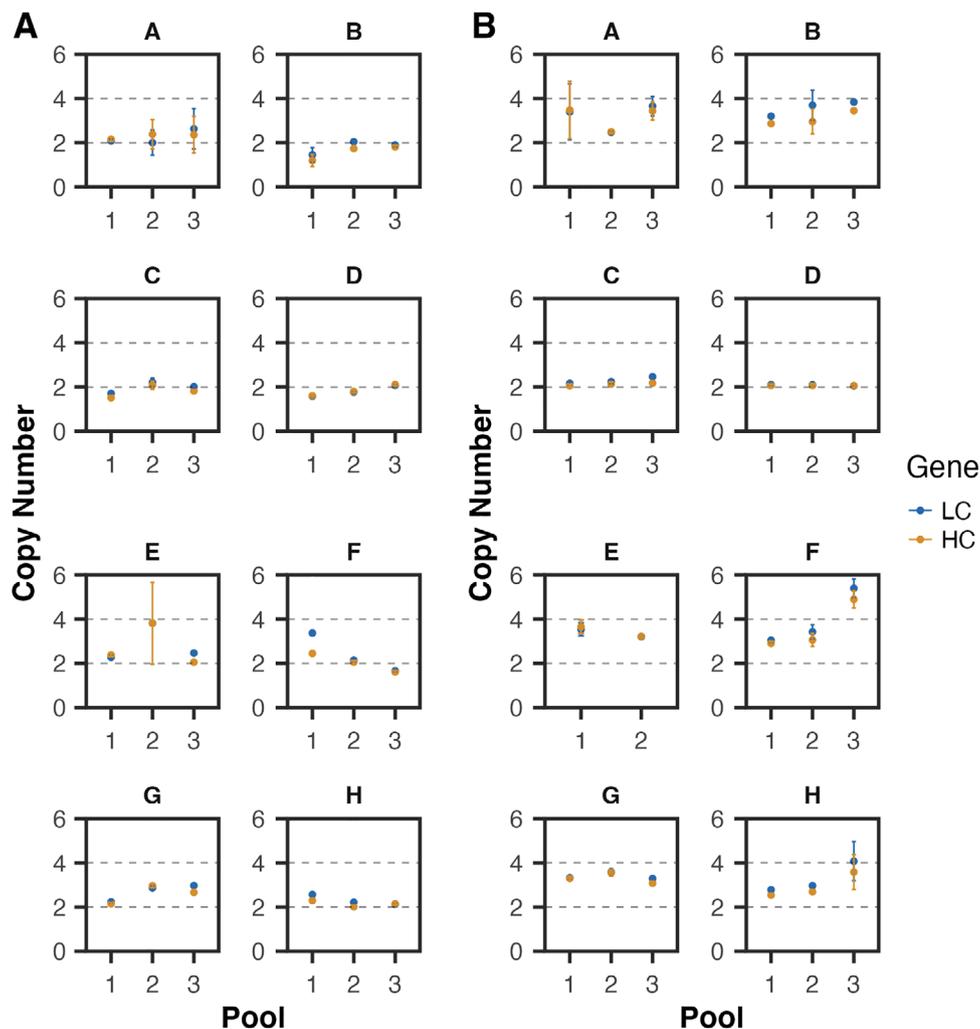


FIGURE 7 | Copy number of adalimumab-expressing pools at (A) *Rosa26* and (B) *Dop1b* as measured by multiplexed ddPCR assays, using *HPRT1* as a reference gene. The construct identifier is shown above each plot, with the copy number for each biological replicate pool shown. Error bars represent the Poisson error for each pool, as calculated by the QX ONE software.

support the theory that baseline gene expression is determined by integration location, clonal background, and transgene DNA sequence, yet the inclusion of regulatory elements can create isolated environments around the transgenes that can locally impact transcription. The observation of similar trends across two integration locations despite differences in local genomic architecture and transcriptional influences suggests that the increases in productivity are not driven solely by interactions with neighboring genomic sequences. High productivities with this system required both high overall transcription of mAb genes and an optimal HC:LC ratio, which appeared to be mAb-dependent and possibly driven by features inherent to the DNA sequence or mAb structure.

Author Contributions

Alana C. Szkodny: Conceptualization; Data curation; Formal analysis; Methodology; Writing—original draft; Writing—review and editing. **Kelvin H. Lee:** Conceptualization; Funding acquisition; Methodology; Resources; Supervision; Writing—original draft; Writing—review and editing.

Acknowledgments

The authors would like to thank Bruce Kingham and Erin Bernberg at the University of Delaware DNA Sequencing and Genotyping Center for their support in planning and performing the PacBio sequencing. We would also like to thank Madolyn MacDonald and Shawn Polson at the University of Delaware Center for Bioinformatics and Computational Biology Bioinformatics Core for their support and expertise in analyzing the PacBio data. The CHO-K1 host cell line used for this work was obtained from the NIH (NIAID). This work was supported by funding from the National Science Foundation under grants 1624698 and 2100502, and in part by the financial assistance award 70NANB17H002 from the U.S. Department of Commerce and National Institutes of Standards and Technology. Figures 1, 3, and S2 were created with Biorender. DNA sequencing maps (Figures S1 and S6) were generated on SnapGene.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Plasmid sequence information is freely available upon request. DNA sequences for all major elements used in this work (TK, RFP-BSD, Neo^R, TFP, trastuzumab light chain and heavy chain, adalimumab light chain and heavy chain, CpG, and PB5TR) can be found in the

supplemental DNA Sequences file in FASTA format. The data that support the findings of this study and the plasmid DNA sequences used in this work are available from the corresponding author upon reasonable request.

References

1. Y.-M. Huang, W. Hu, E. Rustandi, K. Chang, H. Yusuf-Makagiansar, and T. Ryll, "Maximizing Productivity of CHO Cell-Based Fed-Batch Culture Using Chemically Defined Media Conditions and Typical Manufacturing Equipment," *Biotechnology Progress* 26, no. 5 (2010): 1400–1410, <https://doi.org/10.1002/btpr.436>.
2. B. Kiss, U. Gottschalk, and M. Pohlscheidt, eds., *New Bioprocessing Strategies: Development and Manufacturing of Recombinant Antibodies and Proteins* (Cham, Switzerland: Springer International Publishing, 2018), <https://doi.org/10.1007/978-3-319-97110-0>.
3. L. M. Barnes, C. M. Bentley, and A. J. Dickson, "Stability of Protein Production From Recombinant Mammalian Cells," *Biotechnology and Bioengineering* 81, no. 6 (2003): 631–639, <https://doi.org/10.1002/bit.10517>.
4. M. Kim, P. M. O'Callaghan, K. A. Droms, and D. C. James, "A Mechanistic Understanding of Production Instability in CHO Cell Lines Expressing Recombinant Monoclonal Antibodies," *Biotechnology and Bioengineering* 108, no. 10 (2011): 2434–2446, <https://doi.org/10.1002/bit.23189>.
5. F. M. Wurm, "Production of Recombinant Protein Therapeutics in Cultivated Mammalian Cells," *Nature Biotechnology* 22, no. 11 (2004): 1393–1398, <https://doi.org/10.1038/nbt1026>.
6. N. K. Hamaker and K. H. Lee, "Site-Specific Integration Ushers in a New Era of Precise CHO Cell Line Engineering," *Current Opinion in Chemical Engineering* 22 (2018): 152–160, <https://doi.org/10.1016/j.coche.2018.09.011>.
7. J. S. Lee, L. M. Grav, N. E. Lewis, and H. F. Kildegaard, "CRISPR/Cas9-Mediated Genome Engineering of CHO Cell Factories: Application and Perspectives," *Biotechnology Journal* 10, no. 7 (2015): 979–994, <https://doi.org/10.1002/biot.201500082>.
8. S. Turan, C. Zehe, J. Kuehle, J. Qiao, and J. Bode, "Recombinase-Mediated Cassette Exchange (RMCE) — A Rapidly-Expanding Toolbox for Targeted Genomic Modifications," *Gene* 515, no. 1 (2013): 1–27, <https://doi.org/10.1016/j.gene.2012.11.016>.
9. L. M. Grav, D. Sergeeva, J. S. Lee, et al., "Minimizing Clonal Variation During Mammalian Cell Line Engineering for Improved Systems Biology Data Generation," *ACS Synthetic Biology* 7, no. 9 (2018): 2148–2159, <https://doi.org/10.1021/acssynbio.8b00140>.
10. J. S. Lee, H. F. Kildegaard, N. E. Lewis, and G. M. Lee, "Mitigating Clonal Variation in Recombinant Mammalian Cell Lines," *Trends in Biotechnology* 37, no. 9 (2019): 931–942, <https://doi.org/10.1016/j.tibtech.2019.02.007>.
11. K. Srirangan, M. Loignon, and Y. Durocher, "The Use of Site-Specific Recombination and Cassette Exchange Technologies for Monoclonal Antibody Production in Chinese Hamster Ovary Cells: Retrospective Analysis and Future Directions," *Critical Reviews in Biotechnology* 40, no. 6 (2020): 833–851, <https://doi.org/10.1080/07388551.2020.1768043>.
12. L. Gaidukov, L. Wroblewska, and B. Teague, et al., "A Multi-Landing Pad DNA Integration Platform for Mammalian Cell Engineering," *Nucleic Acids Research* 46, no. 8 (2018): 4072–4086, <https://doi.org/10.1093/nar/gky216>.
13. W. Hilliard and K. H. Lee, "A Compendium of Stable Hotspots in the CHO Genome," *Biotechnology and Bioengineering* 120, no. 8 (2023): 2133–2143, <https://doi.org/10.1002/bit.28390>.
14. D. Ng, M. Zhou, D. Zhan, et al., "Development of a Targeted Integration Chinese Hamster Ovary Host Directly Targeting Either One or Two Vectors Simultaneously to a Single Locus Using the Cre/Lox Recombinase-Mediated Cassette Exchange System," *Biotechnology Progress* 37, no. 4 (2021): e3140, <https://doi.org/10.1002/btpr.3140>.
15. L. Zhang, M. C. Inniss, S. Han, et al., "Recombinase-Mediated Cassette Exchange (RMCE) for Monoclonal Antibody Expression in the Commercially Relevant CHOK1SV Cell Line," *Biotechnology Progress* 31, no. 6 (2015): 1645–1656, <https://doi.org/10.1002/btpr.2175>.
16. N. Gödecke, S. Herrmann, H. Hauser, A. Mayer-Bartschmid, M. Trautwein, and D. Wirth, "Rational Design of Single Copy Expression Cassettes in Defined Chromosomal Sites Overcomes Intracloonal Cell-to-Cell Expression Heterogeneity and Ensures Robust Antibody Production," *ACS Synthetic Biology* 10, no. 1 (2021): 145–157, <https://doi.org/10.1021/acssynbio.0c00519>.
17. C. Oliviero, S. C. Hinz, J. P. Bogen, et al., "Generation of a Host Cell Line Containing a MAR-Rich Landing Pad for Site-Specific Integration and Expression of Transgenes," *Biotechnology Progress* 38, no. 4 (2022): e3254, <https://doi.org/10.1002/btpr.3254>.
18. Y. D. Patel, A. J. Brown, J. Zhu, et al., "Control of Multigene Expression Stoichiometry in Mammalian Cells Using Synthetic Promoters," *ACS Synthetic Biology* 10, no. 5 (2021): 1155–1165, <https://doi.org/10.1021/acssynbio.0c00643>.
19. N. Pristovšek, S. Nallapareddy, L. M. Grav, et al., "Systematic Evaluation of Site-Specific Recombinant Gene Expression for Programmable Mammalian Cell Engineering," *ACS Synthetic Biology* 8, no. 4 (2019): 758–774, <https://doi.org/10.1021/acssynbio.8b00453>.
20. D. Sergeeva, G. M. Lee, L. K. Nielsen, and L. M. Grav, "Multicopy Targeted Integration for Accelerated Development of High-Producing Chinese Hamster Ovary Cells," *ACS Synthetic Biology* 9, no. 9 (2020): 2546–2561, <https://doi.org/10.1021/acssynbio.0c00322>.
21. J. Carver, D. Ng, M. Zhou, et al., "Maximizing Antibody Production in a Targeted Integration Host by Optimization of Subunit Gene Dosage and Position," *Biotechnology Progress* 36, no. 4 (2020): e2967, <https://doi.org/10.1002/btpr.2967>.
22. M. Feary, M. A. Moffat, G. F. Casperson, M. J. Allen, and R. J. Young, "CHOK1SV GS-KO SSI Expression System: A Combination of the Fer1L4 Locus and Glutamine Synthetase Selection," *Biotechnology Progress* 37, no. 4 (2021): e3137, <https://doi.org/10.1002/btpr.3137>.
23. G. Chen, R. Babb, and J. P. Fandl, "Enhanced Expression and Stability Regions (United States Patent US7771997B2)," Google Patents, 2010, <https://patents.google.com/patent/US7771997B2/en>.
24. C. K. D. Ng, Y. G. Crawford, A. Shen, et al., "Targeted Integration of Nucleic Acids (World Intellectual Property Organization Patent WO2019126634A2)," Google Patents, 2019, <https://patents.google.com/patent/WO2019126634A2/en>.
25. Y. Shen, D. Burakov, G. Chen, and J. P. Fandl, "CHO Integration Sites and Uses Thereof (United States Patent US9816110B2)," Google Patents, 2017, <https://patents.google.com/patent/US9816110B2/en>.
26. H. Ai, J. N. Henderson, S. J. Remington, and R. E. Campbell, "Directed Evolution of a Monomeric, Bright and Photostable Version of Clavularia Cyan Fluorescent Protein: Structural Characterization and Applications in Fluorescence Imaging," *Biochemical Journal* 400, no. 3 (2006): 531–540, <https://doi.org/10.1042/BJ20060874>.
27. N. K. Hamaker and K. H. Lee, "A Site-Specific Integration Reporter System That Enables Rapid Evaluation of CRISPR/Cas9-Mediated Genome Editing Strategies in CHO Cells," *Biotechnology Journal* 15, no. 8 (2020): 2000057, <https://doi.org/10.1002/biot.202000057>.
28. M. E. Matlashov, D. M. Shcherbakova, J. Alvelid, et al., "A Set of Monomeric Near-infrared Fluorescent Proteins for Multicolor Imaging across Scales," *Nature Communications* 11, no. 1 (2020): 1–12, <https://doi.org/10.1038/s41467-019-13897-6>.
29. N. K. Hamaker and K. H. Lee, "High-Efficiency and Multi-locus Targeted Integration in CHO Cells Using CRISPR-Mediated Donor Nicking and DNA Repair Inhibitors," *Biotechnology and Bioengineering* 120, no. 9 (2023): 2419–2440, <https://doi.org/10.1002/bit.28393>.
30. W. Hilliard, M. L. MacDonald, and K. H. Lee, "Chromosome-Scale Scaffolds for the Chinese Hamster Reference Genome Assembly to Facili-

- tate the Study of the CHO Epigenome,” *Biotechnology and Bioengineering* 117, no. 8 (2020): 2331–2339, <https://doi.org/10.1002/bit.27432>.
31. X. Li, X. Zhao, Y. Fang, et al., “Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter,” *Journal of Biological Chemistry* 273, no. 52 (1998): 34970–34975, <https://doi.org/10.1074/jbc.273.52.34970>.
 32. P. Ghosh, A. I. Kim, and G. F. Hatfull, “The Orientation of Mycobacteriophage Bxb1 Integration Is Solely Dependent on the Central Dinucleotide of attP and attB,” *Molecular Cell* 12, no. 5 (2003): 1101–1111, [https://doi.org/10.1016/S1097-2765\(03\)00444-1](https://doi.org/10.1016/S1097-2765(03)00444-1).
 33. M. C. Inniss, K. Bandara, B. Jusiak, et al., “A Novel Bxb1 Integrase RMCE System for High Fidelity Site-Specific Integration of mAb Expression Cassette in CHO Cells,” *Biotechnology and Bioengineering* 114, no. 8 (2017): 1837–1846, <https://doi.org/10.1002/bit.26268>.
 34. J. Qiao, A. Oumard, W. Wegloehner, and J. Bode, “Novel Tag-and-Exchange (RMCE) Strategies Generate Master Cell Clones With Predictable and Stable Transgene Expression Properties,” *Journal of Molecular Biology* 390, no. 4 (2009): 579–594, <https://doi.org/10.1016/j.jmb.2009.05.012>.
 35. S. Irion, H. Luche, P. Gadue, H. J. Fehling, M. Kennedy, and G. Keller, “Identification and Targeting of the ROSA26 Locus in Human Embryonic Stem Cells,” *Nature Biotechnology* 25, no. 12 (2007): 1477–1482, <https://doi.org/10.1038/nbt1362>.
 36. B. P. Zambrowicz, A. Imamoto, S. Fiering, L. A. Herzenberg, W. G. Kerr, and P. Soriano, “Disruption of Overlapping Transcripts in the ROSA Bgeo 26 Gene Trap Strain Leads to Widespread Expression of β -Galactosidase in Mouse Embryos and Hematopoietic Cells,” *Proceedings of the National Academy of Sciences* 94, no. 8 (1997): 3789–3794, <https://doi.org/10.1073/pnas.94.8.3789>.
 37. W. C. Hilliard and K. H. Lee, “Chinese Hamster Ovary Cell Genomic Hotspots for Recombinant Protein Production (World Intellectual Property Organization Patent WO2024073464),” WIPO, 2024, <https://patents.google.com/patent/WO2024073464A2/en>.
 38. V. Le Fourn, P.-A. Girod, M. Buceta, A. Regamey, and N. Mermod, “CHO Cell Engineering to Prevent Polypeptide Aggregation and Improve Therapeutic Protein Secretion,” *Metabolic Engineering* 21 (2014): 91–102, <https://doi.org/10.1016/j.ymben.2012.12.003>.
 39. C. J. Lee, G. Seth, J. Tsukuda, and R. W. Hamilton, “A Clone Screening Method Using mRNA Levels to Determine Specific Productivity and Product Quality for Monoclonal Antibodies,” *Biotechnology and Bioengineering* 102, no. 4 (2009): 1107–1118, <https://doi.org/10.1002/bit.22126>.
 40. P. M. O’Callaghan, J. McLeod, L. P. Pybus, et al., “Cell Line-Specific Control of Recombinant Monoclonal Antibody Production by CHO Cells,” *Biotechnology and Bioengineering* 106, no. 6 (2010): 938–951, <https://doi.org/10.1002/bit.22769>.
 41. S. Schlatter, S. H. Stansfield, D. M. Dinnis, A. J. Racher, J. R. Birch, and D. C. James, “On the Optimal Ratio of Heavy to Light Chain Genes for Efficient Recombinant Antibody Production by CHO Cells,” *Biotechnology Progress* 21, no. 1 (2005): 122–133, <https://doi.org/10.1021/bp049780w>.
 42. J. Li, C. Zhang, T. Jostock, and S. Dübel, “Analysis of IgG Heavy Chain to Light Chain Ratio With Mutant Encephalomyocarditis Virus Internal Ribosome Entry Site,” *Protein Engineering, Design and Selection* 20, no. 10 (2007): 491–496, <https://doi.org/10.1093/protein/gzm038>.
 43. S. C. L. Ho, E. Y. C. Koh, M. van Beers, et al., “Control of IgG LC:HC Ratio in Stably Transfected CHO Cells and Study of the Impact on Expression, Aggregation, Glycosylation and Conformational Stability,” *Journal of Biotechnology* 165, no. 3 (2013): 157–166, <https://doi.org/10.1016/j.jbiotec.2013.03.019>.
 44. J. Pichler, S. Galosy, J. Mott, and N. Borth, “Selection of CHO Host Cell Subclones With Increased Specific Antibody Production Rates by Repeated Cycles of Transient Transfection and Cell Sorting,” *Biotechnology and Bioengineering* 108, no. 2 (2011): 386–394, <https://doi.org/10.1002/bit.22946>.
 45. W. Pilbrough, T. P. Munro, and P. Gray, “Intraclonal Protein Expression Heterogeneity in Recombinant CHO Cells,” *PLoS ONE* 4, no. 12 (2009): e8432, <https://doi.org/10.1371/journal.pone.0008432>.
 46. C. Guo, F. K. Fordjour, S. J. Tsai, J. C. Morrell, and S. J. Gould, “Choice of Selectable Marker Affects Recombinant Protein Expression in Cells and Exosomes,” *Journal of Biological Chemistry* 297, no. 1 (2021): 100838, <https://doi.org/10.1016/j.jbc.2021.100838>.
 47. J. H. M. Yeo, S. C. L. Ho, M. Mariati, et al., “Optimized Selection Marker and CHO Host Cell Combinations for Generating High Monoclonal Antibody Producing Cell Lines,” *Biotechnology Journal* 12, no. 12 (2017): 1700175, <https://doi.org/10.1002/biot.201700175>.
 48. M. Baumann, E. Gludovacz, N. Sealover, et al., “Preselection of Recombinant Gene Integration Sites Enabling High Transcription Rates in CHO Cells Using Alternate Start Codons and Recombinase Mediated Cassette Exchange,” *Biotechnology and Bioengineering* 114, no. 11 (2017): 2616–2627, <https://doi.org/10.1002/bit.26388>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.