

# Efficient multiplexed targeted gene editing as a strategy to generate improved CHO host cell lines for biopharmaceutical manufacturing applications.

## Authors

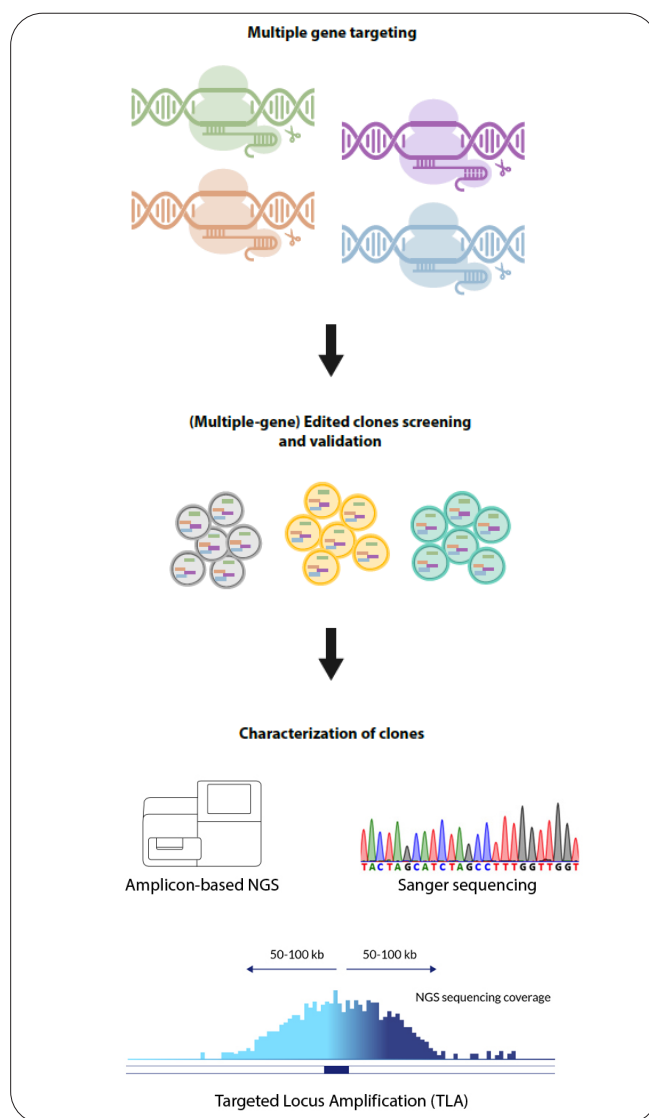
Neha Mishra<sup>1</sup>      Jesus Zurdo<sup>1</sup>  
 Natalia Loaiza<sup>2</sup>      Wietske Stevens<sup>2</sup>  
 Catherine Ingham<sup>1</sup>      Delphine Cougot<sup>1</sup>  
 Martijn Kelder<sup>2</sup>



## Abstract

The use of CRISPR-Cas9 tools for gene editing has opened new possibilities for the genetic optimization of biological systems with biotechnological or therapeutic applications. One particular area of interest has been the design of enhanced expression platforms used in the manufacturing of biopharmaceutical products. Chinese Hamster Ovary (CHO) cells have been traditionally a focus of interest for genetic manipulation to improve its robustness in culture, as well as product yield and quality, amongst other parameters. Although CRISPR platforms provide great flexibility and efficiency in gene editing, the generation of new host cell lines with the required phenotypic advantages and performance usually have complex editing requirements. The long-term success of host-optimization strategies depends in great measure on the ability to perform multiplexed gene editing with high efficiency and the parallel deployment of suitable quality control (QC) methodologies to accurately evaluate all the editing outcomes for every experiment. In this whitepaper we

## Graphical Abstract



demonstrate Revvity's workflow exemplifying the successful simultaneous knock-out of multiple endogenous loci in CHO cells. Phenotypic characterization of the resulting clones shows the absence of the targeted proteins in several of the selected clones, proving the efficiency of the multiplexed gene editing approach. One of the challenges in gene editing, particularly in cases where multiple and simultaneous edits are required, is the generation of unintended edits or genetic rearrangements, that could have negative consequences in the biology of the cells and their required performance.

Here we also show how Targeted Locus Amplification (TLA) can effectively be used to verify the accuracy of gene inactivation in all four loci in one single multiplex experiment, with most of the editing events resulting in small insertions and deletions, as intended. Interestingly, 2 loci also showed large deletions (>300bp) which were not detected by amplicon-based sequencing. Furthermore, TLA also detected two different edited alleles per loci and no wildtype (WT) alleles, confirming 100% targeting efficiency in all loci in the selected clone. Remarkably, no rearrangements directly involving the targeted loci were detected, confirming the efficiency of approaches with simultaneous multiple gRNAs. These results illustrate that simultaneous knock-out of multiple loci using CRISPR-Cas9 and multiple gRNAs is an effective strategy to enhance CHO cell protein production levels and shows that TLA constitutes a comprehensive and unbiased approach to assess the complete landscape of targeted, CRISPR-Cas9-mediated gene editing outcomes in CHO cell lines.

## Introduction

To satisfy the unmet medical needs of millions of patients waiting for available or improved biologic drugs, the pharmaceutical industry must enhance productivity and efficiency in their manufacturing process. The fine-tuned genetic engineering of Chinese Hamster Ovary (CHO) cells, the most widely used expression system for complex recombinant proteins, is one of the key strategies to enhance production efficiency and quality, which has also been accelerated in parallel by advances and wider availability of novel tools for targeted genome editing<sup>1</sup>.

Several published studies have reported successful gene editing approaches to enhanced performance and productivity of CHO cell lines. For example, improved IgG yields and culture performance under stress has been achieved by inactivating pro-apoptotic genes Bak

and Bax, in CHO cells<sup>2</sup>. Furthermore, targeted knock-out (KO) approaches have also proved valuable to optimize glycosylation profiles in CHO cells<sup>1</sup>. Several studies have targeted FUT8 to generate knock-out cell lines that can produce monoclonal antibodies (mAbs) with improved potency and stronger antibody-dependent cellular cytotoxicity<sup>3-6</sup>. More recently, it has also been reported that targeted gene editing can facilitate downstream purification by drastically reducing the amount of co-purifying host cell proteins, which is considered difficult to remove<sup>7</sup>.

The presence of cell-host related impurities can negatively affect the yield and quality of biopharmaceuticals. Process impurities are typically eliminated using several downstream process steps involving chromatography and diafiltration primarily, but occasionally some impurities are difficult to remove, either because they co-purify with the biotherapeutic product or somehow interact with it. Even small amounts of such impurities can potentially trigger or exacerbate immunogenic reactions in patients, and regulatory agencies tend to pay special attention to controlling and minimising levels of impurities in therapeutics.

As a proof-of-concept study we decided to focus on the elimination of several protein impurities that had been identified during bioreactor runs performed with CHO cells in culture. We selected specifically four of these proteins and proceeded to knock out simultaneously all these gene targets. We performed both phenotypic and genotypic validation of the edits and show that the knockout of targeted genes results in homozygous knockout clones with reduced impurities compared to the parental CHOSOURCE CHO-K1 GS KO cell line.

The design and optimization of manufacturing bioproduction processes is a demanding and expensive task. Therefore, switching to a new expression cell host requires evidence of substantial performance advantages. These, in turn, are likely to be linked to complex synergistic genetic manipulations involving multiple genes, as suggested in some of these studies<sup>7</sup>. This has been made possible by the use of new tools like CRISPR-Cas9 that have enabled faster, cheaper, easier, and more efficient alterations compared with earlier gene editing platforms<sup>8</sup>. Despite these advantages, CRISPR-Cas9 KO gene editing approaches are known to rely on non-homologous end joining (NHEJ) repair, which can result in undesired sequence variants ranging from single base pair changes to large structural

rearrangements<sup>9,10</sup>. As a result, the application of more extensive gene editing approaches must always be accompanied with a comprehensive quality control (QC) strategy to ensure the success and accuracy of the gene editing approach and the quality of the resulting products.

Traditionally, gene editing outcomes are characterized using PCR and amplicon sequencing of the targeted region, which provide only limited resolution into the locus of interest. Furthermore, unexpected events, such as large deletions or translocations occurring in the vicinity of the edited locus, cannot be comprehensively evaluated with this approach and are often missed or difficult to interpret. This is why exclusive reliance on these methods presents challenges when attempting to describe, in a robust and hypothesis-free way, all genetic variants (including structural rearrangements) in and around the different targeted loci.

Targeted Locus Amplification (TLA), combined with Next Generation Sequencing (NGS), has become the method of choice for the comprehensive genetic QC of genetically engineered cell lines for biopharmaceutical production, including both transgene integration as well as targeted editing outcomes using CRISPR-Cas9 approaches<sup>11-13</sup>. In one single experiment, TLA can identify any potential sequence variant (i.e. single nucleotide or structural variations) in integrated vector sequences or targeted gene editing in an unbiased manner<sup>14</sup>, unlike PCR-based approaches which are largely hypothesis-driven (i.e. rely on “predicted” edited sites).

We use TLA to verify the accuracy of CRISPR-Cas9-induced edits in four different loci targeted simultaneously in CHO cells. In addition, we demonstrate the potential of TLA to highlight unpredicted outcomes in the targeted region while also providing further genetic information in a larger genomic context, unlike amplicon-sequencing methods. The comprehensive and solid genetic data generated in one single TLA experiment provides robust assurance of the quality of the gene editing in CHO cell lines and constitutes a strong pillar to support the optimization strategies to enable rational design of mammalian expression systems based on targeted gene editing.

## Methods

### Cell culture

CHOSOURCE CHO-K1 GS KO cells (Revvity Discovery) were used to generate the knock-out (KO) cell lines. Cells were grown as suspension culture in CD OptiCHO (Gibco,

ThermoFisher Scientific) supplemented with 4mM L-Glutamine (Sigma-Aldrich). Cultures were maintained at 37°C, 5% CO<sub>2</sub>, 85% humidity in a shaking incubator at 125 rpm.

### Gene copy number analysis by digital droplet PCR (ddPCR)

Copy number of the target genes was determined using a ddPCRbased assay. Assays (primer/probe sets) were designed against the target genes using the nucleotide sequence available on Ensembl (Revvity CHOSOURCE CHO-K1 GS-KO genome: CHO-K1 GS assembly CHOK1GS\_HDv1).  $\beta$ 2-Microglobulin was used as reference gene. Copy number for each target gene was quantified using the QX200 droplet reader (Bio-Rad). Data analysis was performed on the QuantaSoft software (Bio-Rad).

### Gene editing

CRISPR RNAs (crRNAs) targeting the genes of interest were designed using Revvity’s guide design algorithm. Five crRNAs per target were tested and the best design per target gene was selected. For generating the KO pool, four targets were simultaneously knocked out using Revvity’s proprietary transfection method with Cas9 protein, tracrRNA and the 4 crRNAs selected.

CRISPR-Cas9 editing efficiency was assessed using a DNA mismatch detection assay using T7E1 endonuclease as well as a TIDE-like in-house bioinformatic tool to analyze Sanger sequencing reads. First, cells were lysed in Direct-PCR lysis buffer (Viagen Biotech) supplemented with Proteinase K (Sigma Aldrich) and the cell lysate was used as template for PCR. Primers were designed to amplify around each of the targeted sequences (600-700bp total amplicon length). For the T7E1 assay (New England Biolabs), the PCR products were denatured, rehybridized and incubated with or without T7E1 for 30 min at 37°C. The reactions were then loaded on a 2% agarose gel and visualized on a gel imager.

To assess the editing efficiency, Sanger sequencing traces from the PCR products were analyzed using a TIDE-like bioinformatics tool developed in-house. The tool allows for the detection of indels and determines the percentage of the wild-type allele detected in the samples.

### Clonal isolation

Transfected pools were subjected to limiting dilution cloning in 384-well plates using Revvity’s standard protocol. To ensure monoclonality, plates were scanned on the day of seeding (day 0), day 1 and day 2. Two weeks after seeding, plates were scanned to identify clone outgrowth.

Clones were expanded to 96-wells and screened for the presence of edited alleles on the four targeted genes. Identified KO clones were expanded, banked, and subjected to additional validation.

### Amplicon-based NGS

Screening of edited clones was performed using a targeted-NGS approach. Primers were designed to amplify up to 300bp around the targeted sequences. PCR was performed, and the products purified as per Revvity standard process. Sequencing was performed on the Illumina MiSeq system and data analyzed using an in-house bioinformatics tool.

### Targeted Locus Amplification

Selected clones were processed and analyzed by Cergentis (Utrecht, Netherlands) as described in 14. Each locus of interest was targeted with two different primer sets each. TLA products were sequenced on an Illumina sequencer generating paired-end, 2x150 bp reads. Mapping was performed using BWA-MEM<sup>15</sup> version 0.7.15-r1140, settings mem -B 7, with the Chinese hamster genome sequence (GCF\_003668045.1 assembly) as reference genome.

## Results

### Gene copy number analysis of target genes

Gene copy number analysis was performed to determine the number of alleles of the selected genes present in the CHOSOURCE CHO-K1S GS KO parental cell line. ddPCR assays were designed for each of the four target genes as well as for the reference  $\beta$ 2-Microglobulin gene. The analysis revealed that each gene is present in two copies in the parental cell line (Table 1) indicating that two alleles for each gene need to be edited.

Table 1: Copy number of target genes assessed by digital droplet PCR

Gene	Copy number of targeted genes
A	2
B	2
C	2
D	2

### Screening crRNA for gene editing efficiency

A crRNA screen was performed to identify the crRNA for each target gene with the best editing efficiency. For this, 5 crRNAs per target gene were transfected in the parental CHO cell line using Revvity's proprietary transfection protocol. A T7E1 mismatch detection analysis was performed on PCR products to confirm the presence of indels in the target region. An example of the T7E1 mismatch assay results obtained for gene A is shown in Figure 1. The presence of cleaved bands in lanes where the PCR product was treated with the T7E1 enzyme (+) indicates the presence of indels and therefore confirmed editing in these pools.

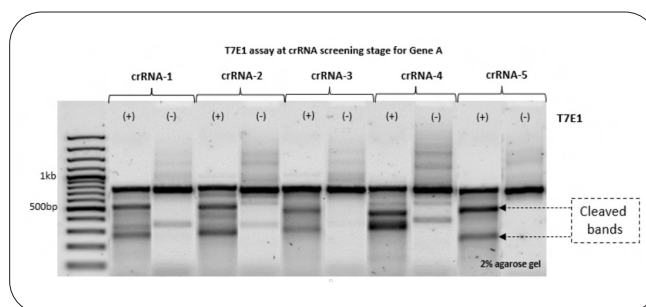


Figure 1: example of T7E1 endonuclease assay obtained during the screening of crRNA targeting gene A

In addition to the T7E1 mismatch assay, the PCR product was subjected to Sanger sequencing and the editing efficiency was assessed using a TIDE style analysis, using an in-house bioinformatic tool. A summary of all the editing efficiencies from the crRNA screening performed for all the genes is shown in Table 2.

Table 2: Editing efficiencies (expressed in %) obtained from crRNA screening for the four target genes. The crRNAs selected for the rest of the study are highlighted in green.

gRNA	% Efficiency			
	Gene A	Gene B	Gene C	Gene D
1	33.2	79.3	75.4	26.9
2	68.7	50.5	78.6	25.7
3	31.7	64	53.8	29.3
4	22.2	64.9	42.2	12.3
5	42.2	41.9	68.2	42.8

## Generation and characterization of an edited pool

Following identification of the best crRNA for each gene, the parental CHOSOURCE CHO-K1 GS KO cells were co-transfected using Revvity's optimized transfection conditions. All four genes of interest were targeted simultaneously. Efficiency of the editing in the four target genes was assessed in the transfected pool using Sanger sequencing and an in-house bioinformatic tool. As shown in Figure 2, high editing efficiency was achieved for each of the target loci. The transfected pool was subject to limiting dilution cloning to isolate clones.

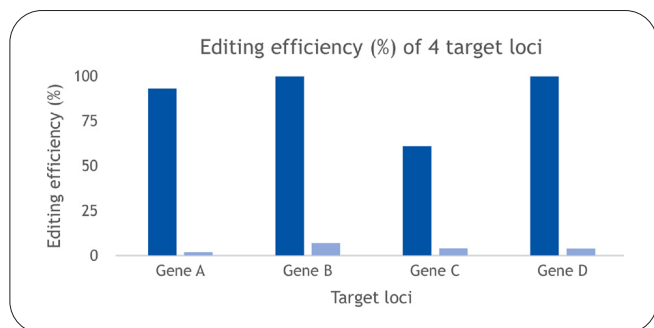


Figure 2: Editing efficiency (%) obtained in the 4 target loci in the transfected pool. For each gene, a non-transfected control sample was used to assess the background of the analysis.

## Isolation and characterization of KO clones

Following the limited dilution cloning of the transfected pool, a total of 298 clones were isolated and screened using an amplicon-based NGS approach to identify clones carrying out-of-frame indels on both alleles for the four target genes. Gene B could not be assessed at this stage due to a technical issue with the sequencing of the NGS amplicon. Interestingly, the analysis of the NGS data revealed that in most of the clones, only one indel was detected for one or

more of the target genes instead of the two indels expected. There are a few different hypotheses to consider that may explain these results: 1) the same indel has been created on both alleles; 2) the editing event has created a large indel which has removed one or both primer-binding sites used for the NGS analysis; 3) the double-strand breaks created by targeting four genes simultaneously have led to chromosomal rearrangements.

To further evaluate the clones, we performed Sanger sequencing of all the edited clones identified from the NGS screen including those that may contain out-of-frame indels in gene B as this could not be assessed with NGS for technical reasons. The sequence traces were analyzed by a TIDE-like internal bioinformatics tool. The analysis confirmed a total of 17 clones with all four target genes knocked out. The results also verified the data obtained from the NGS screen for genes A, C and D. A summary of the results from the NGS and Sanger sequencing analysis from one of the clones is shown in Table 3 below.

Although the data from the NGS screen was confirmed by Sanger sequencing, it was still unclear whether the simultaneous gene editing of the four different targets had led to complex genetic events i.e., chromosomal rearrangements, that are undetectable by sequencing methods such as Sanger and amplicon-based NGS. Therefore, to further characterize the KO clones and to determine all the editing outcomes, two different clones were analysed by TLA (described in the section below).

Table 3: Table summarizing data obtained from amplicon-based NGS and Sanger sequencing in one of the KO clones (Sample 1).

Gene	Copy Number	Amplicon-based		Sanger sequencing	
		detected indels	Percentage	detected indels	Percentage
A	2	+1		+1	91.9
B	2	*NA	*NA	-2	47.5
				-5	44.2
C	2			-2	48.8
				+1	45.9
D	2	+2	98.6	+2	94

\*NA - not analysed

### Phenotypic validation of the edited clones

Following the genotypic validation described in the previous section, a functional validation was also performed to confirm the inactivation of the four targeted genes in those clones isolated from this study. A mass spectrometry-based analysis was performed on the edited clones and parental cell line to determine the amount of each of the proteins in the KO clones compared to the levels observed in the parental cell line. Mass spectrometry analysis confirmed the absence of the proteins encoded by the corresponding edited genes in clones 1 and 2 (Table 4). This further validated that both alleles of each of the targeted genes had indeed been knocked out successfully (they are homozygous knockouts). Two additional clones (clones 3 and 4), showed reduced amounts of two of the proteins compared to those observed in the parental GS KO cell line. This observation

was presumed to be likely due to the clones being heterozygous KO, which was confirmed by further analysis using Sanger sequencing (data not shown). These clones were not further analysed by TLA.

Interestingly, we observed how the successful knock out of the four target genes caused the elimination of another protein (Encoded by Gene X) that was not originally targeted in this study, compared to the parental cell line. This observation seem to support a synergistic interplay amongst different genes and further supports the use of multiplexing approaches in gene editing to achieve more profound phenotypic desirable effects compared to single (one-at-a-time) gene knockout approaches. In this particular case, additional removal of impurities was achieved in the edited cell lines, with considerable advantages in their use as hosts for biotherapeutic production.

Table 4: Table summarizing percentage of protein in the edited clones. Clones 1 and 2 have been analysed by both Mass Spectrometry and TLA, whereas Clones 3 and 4 have been analysed by mass spectrometry only.

Gene	% protein in parental cell line	% protein in KO cell line			
		Clone 1	Clone 2	Clone 3	Clone 4
A	0.21	0	0	0	0
B	0.46	0	0	0.04	0
C	0.30	0	0	0	0
D	0.20	0	0	0.14	0.11
X	0.05	0	0	0.02	0

### Targeted Locus Amplification (TLA)

The four edited loci were analyzed by TLA-NGS in two selected clones (sample 1 and sample 2), with the main findings summarized in Table 5. Although two samples were

analyzed, the plots shown correspond to only sample 1 for illustration purposes.

Table 5: Summary of events detected in each of the targeted loci in sample 1.

Sample	Targeted loci	Breakpoint reads	Indels	Percentage of breakpoints in locus (%)	% WT reads	Detected by ampliconNGS	Detected by Sanger sequencing
1	GeneA	2	-457	25	x	x	x
			+1	75		✓	✓
	GeneB	2	-5	48	x	No data	x
			-2	52		No data	✓
	GeneC	2	-2	48	x	x	x
			+1	52		✓	✓
	GeneD	2	+2	60	x	✓	✓
			-300	40		x	x

X: not detected; ✓: detected

For sample 1, coverage was found in the corresponding genomic loci (Figure 3), and closer inspection of the TLA coverage peaks (see locus coverage plots of Figure 4 as example) shows that two modifications (one per allele) were found in each gRNA targeted gene locus, consisting mostly of small deletions or insertions but also larger previously undetected deletions in genes A and D of sample 1. Of note, this allele-specific level of resolution was missed by the amplicon sequencing approaches reported in Table 3, which only detected the most predominant outcome.

No wild-type reads were found at any of the gRNA targeted sites. Specific breakpoint sequences were identified for each modified allele (data not shown). Based on TLA analysis, we reconstructed the different loci as shown in Figure 5 and could confirm the absence of complex chromosomal rearrangements between the targeted loci possibly induced by the simultaneous gene editing approach.

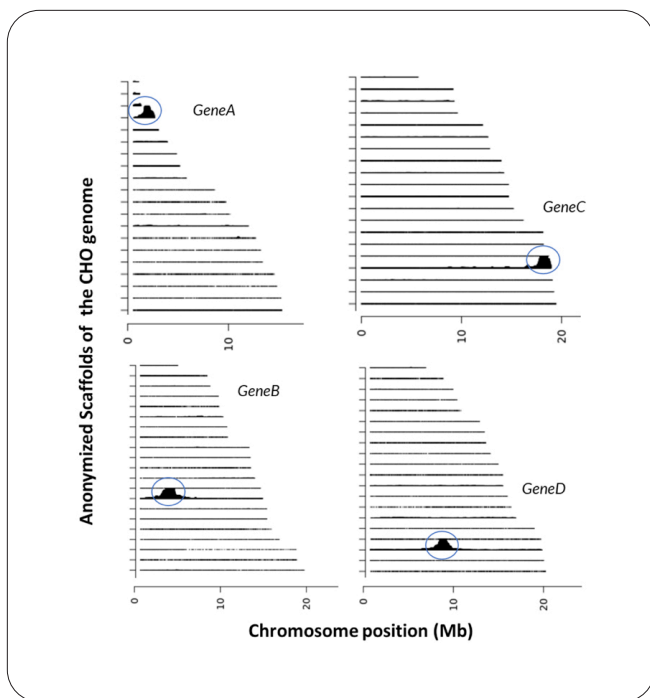


Figure 3. Whole genome coverage plots for Genes A, B, C and D targeted simultaneously with four different gRNAs in CHO cells. TLA sequence coverage across the top 50 scaffolds of the Chinese Hamster genome using primer specific primer pairs for each locus. The anonymized scaffolds are indicated on the y-axis, the scaffold position on the x-axis. Similar coverage plots were obtained for both samples analyzed.

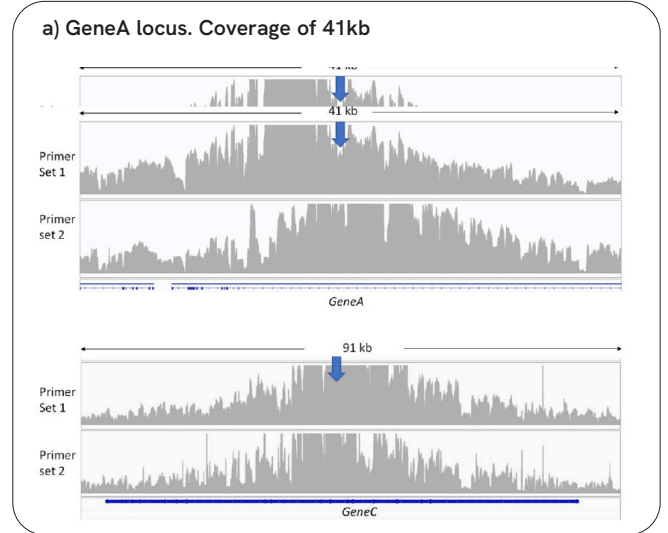


Figure 4. Locus-wide coverage in the corresponding scaffold for two locus analyzed for illustrative purposes. Examples based on sample 1. Similar data was generated for the 4 loci analyzed and both samples.

TLA sequence coverage (in grey) across the different targeted loci. The blue arrow indicates the location of the gRNA. No WT reads are identified. Y-axes are limited to 1,000x (top) and 2,000x (bottom).

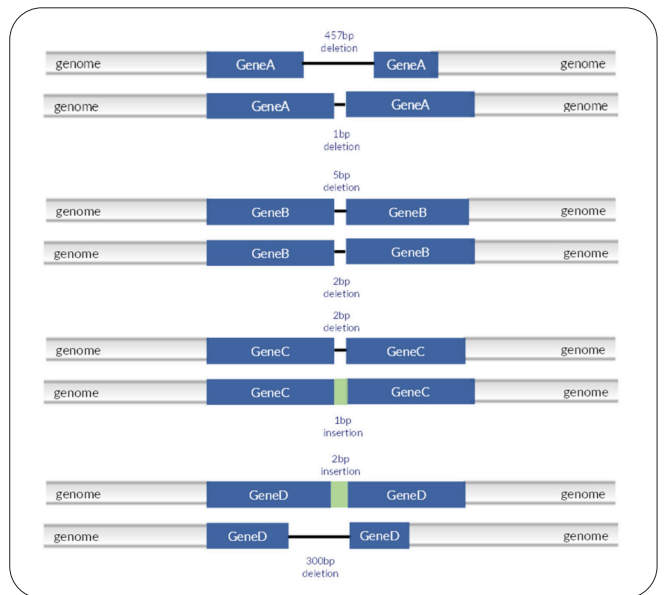


Figure 5. Schematic representation of the gene editing events detected in each gene analyzed for sample 1. Inversion in the locus B, containing a large inversion.

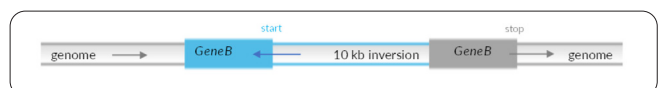


Figure 6. Schematic representation of the genomic inversion in the locus B, containing a large inversion. Data not shown corresponding to Sample 2.

## Discussion

Here we have demonstrated the successful application of Revvity's established gene editing pipeline (including proprietary transfection methodology and in-house bioinformatic tools) to perform a simultaneous multiplexed knock-out of four different target genes in Revvity's CHOSOURCE GS KO cells. Furthermore, we have shown how the use of multiplexed gene editing can be used to detect potential synergistic effects that might not be apparent following one-at-a-time gene editing approaches. Mass spectrometry analysis of different edited clones not only validated the absence of the impurities in those clones where successful edit of all alleles had been accomplished, but also identified the absence of a fifth protein impurity unrelated to the other four knocked out genes (Table 4), supporting the existence of such synergistic interplay between them.

After amplicon-based sequence screening to identify clones with all 4 target genes knocked out, a Sanger sequencing screen was performed to validate the NGS data, and the presence of out-of-frame indels was confirmed. However, amplicon-based sequencing did not allow a full understanding of all the different outcomes resulting from the genetic manipulation. In addition, it is known that CRISPR editing may introduce abnormalities in chromosomes due to editing by DNA double-strand breaks, leading to genomic instability,<sup>16,17,18</sup> which could have a negative impact in the desired performance of the cell hosts. TLA, was therefore used to further characterize all editing outcomes in the selected KO clones, with a particular focus in the detection of such potential rearrangement events.

The application of TLA in two selected clones confirmed and expanded the initial findings by amplicon-based NGS and Sanger sequencing. Specifically, TLA could detect two different editing outcomes per loci in each sample, while the amplicon-based approaches could only detect the most frequent event (see comparison between methods in table 5). In the case of Gene A and Gene D, these minor frequency alleles were larger deletions ( $\geq 300$ bp) that would have evaded ampliconbased detection due to complete removal of the primer binding sites. The other small deletions detected at lower frequencies per loci were still relatively frequent (48% or 47%) suggesting important limitations in the sensitivity of the amplicon-based methods compared to TLA in the detection of the complete range of gene editing outcomes.

Importantly, TLA confirmed the absence of complex chromosomal rearrangements between the simultaneously targeted loci, further underscoring the robustness of the multiplexed gene editing approach used. However, it is important to mention that one of the analyzed samples did reveal the presence of a chromosomal rearrangement, which was also missed by the amplicon-based screening, that does not appear to be related to CRISPR-Cas9 editing. More specifically, TLA detected an inversion in Gene B and a chromosomal rearrangement between a region in Gene C and an unannotated region of the CHO genome in one specific scaffold (data not shown). These events appear to be unrelated to a specific off-target effect of the Cas9, which again highlights the importance of comprehensive genetic QC when characterizing gene editing outcomes and selecting clones.

It is also worth noting that TLA analysis helped resolve all (genetic) events occurring at each locus, by providing exact breakpoint sequences at single nucleotide resolution and generating comprehensive coverage across a range of 40-90 kb surrounding the loci of interest. Furthermore, a single TLA assay generated high coverage ( $>1000x$ ) of all 4 loci, thus proving to be a time-effective and comprehensive genetic QC approach for characterization of gene editing outcomes.

In summary, we have shown that using Revvity's expertise, which includes well-defined pipelines and tools to isolate, identify and characterize edited KO clones, we were able to develop a multiplexed gene targeting strategy where four different loci are targeted simultaneously to generate a host cell line with reduced impurity levels. We were also able to further validate these clones by using TLA-based solutions to identify all the genetic events occurring due to gene editing in these particular clones. In addition to the different genetic validation performed, we have further confirmed the successful inactivation of the 4 target genes in the KO cells plus the elimination of an additional impurity as confirmed by mass spectrometry.



## Conclusion

Simultaneous gene editing using CRISPR-Cas9 is now part of the established capabilities offered by Revvity. The robust pipeline incorporates stringent genetic QC steps to ensure comprehensive characterization of the generated cell lines, including TLA analysis of top selected clones. TLA allows the determination of all gene editing outcomes in both alleles in one single experiment while also providing large coverage (40-90kb) into the surrounding genome. The outcomes reported by TLA are not limited to small deletions and insertions, but also large deletions, inversion, and chromosomal translocations, which are usually missed by more conventional amplicon-based methods.

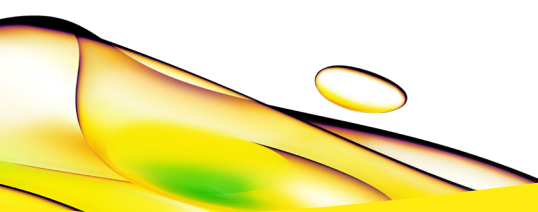
## References

1. Fischer, S., Handrick, R. & Otte, K. The art of CHO cell engineering: A comprehensive retrospect and future perspectives. *Biotechnol. Adv.* 33, 1878-1896 (2015).
2. Cost, G. J. et al. BAK and BAX deletion using zinc-finger nucleases yields apoptosis-resistant CHO cells. *Biotechnol. Bioeng.* 105, 330-340 (2010).
3. Cristea, S. et al. In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration. *Biotechnol. Bioeng.* 110, 871-880 (2013).
4. Malphettes, L. et al. Highly efficient deletion of FUT8 in CHO cell lines using zinc-finger nucleases yields cells that produce completely nonfucosylated antibodies. *Biotechnol. Bioeng.* 106, 774-783 (2010).
5. Yamane-Ohnuki, N. & Satoh, M. Production of therapeutic antibodies with controlled fucosylation. *mAbs* 1, 230-236 (2009).
6. Yamane-Ohnuki, N. et al. Establishment of FUT8 knockout Chinese hamster ovary cells: An ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity. *Biotechnol. Bioeng.* 87, 614-622 (2004).
7. Dovgan, T. et al. Targeted CHO cell engineering approaches can reduce HCP-related enzymatic degradation and improve mAb product quality. *Biotechnol. Bioeng.* 118, 3821-3831 (2021).
8. Ronda, C. et al. Accelerating genome editing in CHO cells using CRISPR Cas9 and CRISPy, a web-based target finding tool. *Biotechnol. Bioeng.* 111, 1604-1616 (2014).
9. Cullot, G. et al. CRISPR-Cas9 genome editing induces megabasescale chromosomal truncations. *Nat. Commun.* 10, 1136 (2019).
10. Liu, M. et al. Global detection of DNA repair outcomes induced by CRISPR-Cas9. *Nucleic Acids Res.* 49, 8732-8742 (2021).
11. Dhiman, H., Campbell, M., Melcher, M., Smith, K. D. & Borth, N. Predicting favorable landing pads for targeted integrations in Chinese hamster ovary cell lines by learning stability characteristics from random transgene integrations. *Comput. Struct. Biotechnol. J.* 18, 3632-3648 (2020).
12. Feary, M., Moffat, M. A., Casperson, G. F., Allen, M. J. & Young, R. J. CHOK1SV GS-KO SSI expression system: A combination of the Fer1L4 locus and glutamine synthetase selection. *Biotechnol. Prog.* 37, e3137 (2021).
13. Stadermann, A. et al. Structural analysis of random transgene integration in CHO manufacturing cell lines by targeted sequencing. *Biotechnol. Bioeng.* 119, 868-880 (2022).

14. de Vree, P. J. P. et al. Targeted sequencing by proximity ligation for comprehensive variant detection and local haplotyping. *Nat. Biotechnol.* 32, 1019–1025 (2014).
15. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* 26, 589–595 (2010).
16. Kosicki, M., Tomberg, K. & Bradley, A. Erratum: Repair of doublestrand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* 36, 899 (2018).
17. Zuccaro, M. V et al. Allele-Specific Chromosome Removal after Cas9 Cleavage in Human Embryos. *Cell* 183, 1650–1664.e15 (2020).
18. Cullot, G. et al. CRISPR–Cas9 genome editing induces megabasescale chromosomal truncations. *Nat. Commun.* 10, 1136 (2019).

## Affiliations

1. Revvity Discovery, 8100 Cambridge Research Park, Waterbeach, Cambridge, CB25 9TL, United Kingdom.
2. Cergentis B.V., Yalelaan 62, 3584 CM Utrecht, The Netherlands.



revvity