Expanding the CRISPR Toolbox for Chinese Hamster Ovary Cell Line Engineering

Julie la Cour Karottki, Karen

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Expanding the CRISPR Toolbox for Chinese Hamster Ovary Cell Line Engineering

Karen Julie la Cour Karottki

PhD Thesis
February 2019

DTU Biosustain
The Novo Nordisk Foundation Center for Biosustainability

Expanding the CRISPR Toolbox for Chinese Hamster Ovary Cell Line Engineering

Ph.D. Thesis

Karen Julie la Cour Karottki

Novo Nordisk Foundation Center for Biosustainability
The Technical University of Denmark
February 2019

Supervisors: Senior Researcher and Co-PI Helene Faustrup Kildegaard & Senior Researcher and Co-PI Lasse Ebdrup Pedersen
Co-Supervisors: Associate Professor Nathan E. Lewis & Assistant Professor Jae Seong Lee

"For the trailing white spaces… "

PREFACE

This PhD was conducted between the 1st of March 2016 and the 28th of February 2019. The work was carried out at the Novo Nordisk Foundation Center for Biosustainability (CFB), at the Technical University of Denmark (DTU). I was supervised by senior researcher and Co-PI Helene Faustrup Kildegaard from the beginning of my PhD until the end of July 2018 after which senior researcher and Co-PI Lasse Ebdrup Pedersen assumed these responsibilities. Throughout my PhD I was co-supervised by assistant professor Jae Seong Lee, Ajou University, Seoul, South Korea and associate professor Nathan E. Lewis, University of California San Diego, USA. The period from November 2017 to May 2018 was spent at the Systems Biology and Cell Engineering lab at University of California San Diego gaining a range of computational skills under the supervision of associate professor Nathan E. Lewis. The PhD was funded by the Novo Nordisk Foundation. The six months external research stay was partially funded by Danmark Amerika Fondet.

Karen Julie la Cour Karottki
28th February 2019
The Novo Nordisk Foundation Center for Biosustainability
Technical University of Denmark

ABSTRACT

Recombinant biotherapeutic proteins are used in the treatment of some of the most serious diseases, such as cancer and autoimmune diseases, affecting millions of people worldwide. In this field, Chinese hamster ovary (CHO) cells dominate as the expression host of choice. For decades CHO cells have delivered high quality products approved by regulatory agencies, yet the underlying mechanisms governing their behaviour are not well understood. To meet the increasing global demand from a growing and aging population it is imperative to be able to precisely manipulate cells to obtain desirable traits. Recent advances in unveiling the genetic context and developments in genetic engineering tools have unlocked a new era for understanding and changing the nature of CHO cells. The gene editing technology CRISPR has gained foothold as a powerful genetic engineering system in multiple mammalian systems and has proven useful for engineering CHO cells.

We here present the application of novel CRISPR techniques in CHO cells with the aim of expanding the tool box to enable more diverse CHO cell line engineering and target discovery. We focus on three methods: (1) CRISPR interference to repress transcription, (2) CRISPR activation to enhance transcription, and (3) large-scale CRISPR knockout screening to enable high-throughput novel target identification. We conclude that these three tools can be used individually or in conjunction to permit a multilevel exploration of the CHO phenotypic space, which should permit more rapid development of rationally engineered CHO cell lines for biotherapeutic protein production.

DANSK SAMMENFATNING

Rekombinante bioterapeutiske proteiner anvendes til behandling af nogle af de alvorligste sygdomme, såsom kræft og autoimmune sygdomme, der rammer millioner af mennesker globalt. Kinesisk hamster ovarie (CHO) celler er en af de foretrukne produktionsværter indenfor dette område. I årtier har CHO-celler leveret myndighedsgodkendte produkter af høj kvalitet, men de underliggende mekanismer, der styrer deres adfærd, er ikke velkendte. For at imødekomme den stigende globale efterspørgsel fra en voksende og aldrende befolkning er det nødvendigt at kunne manipulere cellerne præcist for at opnå de ønskede egenskaber. Nylige fremskridt indenfor kendskabet til CHO cellers genetiske kontekst og udvikling af gentekniske værktøjer har åbnet helt nye muligheder for at forstå og ændre CHO-cellernes natur. Genredigeringsmetoden CRISPR har vundet fodfæste som et kraftigt genteknisk system i flere pattedyr-systemer og har vist sig nyttigt til at modificere CHO-celler.

Her præsenterer vi anvendelsen af nye CRISPR-teknikker i CHO-celler med det formål at udvide værktøjskassen for CHO-cellelinje redigerings teknikker og for at finde frem til nye redigeringsmål. Vi fokuserer på tre metoder: (1) CRISPR interferens til at undertrykke transkription, (2) CRISPR aktivering for at øge transkription og (3) storskala CRISPR knockout screening for at muliggøre high-throughput identifikation af nye redigeringsmål. Vi konkluderer at disse tre værktøjer kan anvendes individuelt eller sammen med henblik på at muliggøre en udforskning af CHO-fænotypisk rum på flere niveauer, hvilket burde muliggøre hurtigere udvikling af rationelt udviklede CHO-cellelinier til forbedret bioterapeutisk proteinproduktion.

LIST OF MANUSCRIPTS

Articles and manuscripts included in thesis (published and unpublished):
* Denotes equal contribution

I. **CRISPR toolbox for mammalian cell engineering**
Daria Sergeeva, Karen Julie la Cour Karottki, Jae Seong Lee and Helene Faustrup Kildegaard, Cell Culture Engineering: Recombinant Protein Production (Advanced Biotechnology), Wiley-Blackwell Biotechnology Series (accepted for publication)

II. **Reduced Apoptosis in Chinese Hamster Ovary Cells via Optimized CRISPR Interference**
Kai Xiong*, Kim Fabiano Marquart*, Karen Julie la Cour Karottki*, Shangzhong Li, Isaac Shamie, Jae Seong Lee, Nan Cher Yeo, Alejandro Chavez, Gyun Min Lee, Nathan E. Lewis, Helene Faustrup Kildegaard (resubmitted after revision)

III. **CRISPR Activation for Transcriptional Reprogramming of Endogenous Genes in Chinese Hamster Ovary Cells**
Karen Julie la Cour Karottki*, Isaac Shamie*, Kai Xiong, Hooman Hefzi, Anders Holmgaard Hansen, Songyuan Li, Jae Seong Lee, Shangzhong Li, Samuel Roth, Sascha Duttke, Christopher Benner, Gyun Min Lee, Helene Faustrup Kildegaard, Nathan E. Lewis (manuscript in preparation)

IV. **Establishing a Pooled Metabolic CRISPR-Cas9 Knockout Screen in Suspension Chinese Hamster Ovary Cells**
Karen Julie la Cour Karottki, Hooman Hefzi, Songyuan Li, Lasse Ebdrup Pedersen, Philipp Spahn, Alex Thomas, Nachon C. Pedersen, Jae Seong Lee, Gyun Min Lee, Nathan E. Lewis, Helene Faustrup Kildegaard (manuscript in preparation)

Contributions to manuscripts not included in this thesis:

Application of CRISPR/Cas9 genome editing to improve recombinant protein production in CHO cells
Lise Marie Grav, Karen Julie la Cour Karottki, Jae Seong Lee, Helene Faustrup Kildegaard Heterologous Protein Production in CHO Cells, Springer. pp. 101-118. (Methods in Molecular Biology, Vol. 1603)

Multiplex genome editing eliminates the Warburg Effect without affecting oxidative metabolism or growth rate
Hooman Hefzi, Iván M. Monge, Soo Min Noh, Karen Julie la Cour Karottki, Marianne Decker, Johnny Arnsdorf, Stefan Kol, Nuša Pristovšek, Anders Holmgaard Hansen, Sara Petersen Bjorn, Karen Kathrine Brøndum, Elham Maria Javidi, Kristian Lund Jensen, Thomas Beuchert Kallehauge, Daniel Ley, Patrice Ménard, Helle Munck Petersen, Zulfiya Sukhova, Lars K. Nielsen, Gyun Min Lee, Helene Faustrup Kildegaard, Bjørn G. Voldborg, Nathan E. Lewis (manuscript in preparation)

ACKNOWLEDGEMENTS

Where to start…

First of all, I would like to thank my group of supervisors. It has been an ever-changing journey, but I have never doubted your support. Helene, thank you for introducing me to the world of CHO, CRISPR and the world where basic research tangents industrial application. Thank you for your never-ending enthusiasm for science. It's infectious. Lasse, thank you for taking up the role as main supervisor near the end where all things collide. Also thank you for always providing insights to the table on conversations about everything. I also want to thank Gyun Min for keeping everything in check at a distance and always being there to give insightful feedback. Jae, thank you for always providing a path forward for any problem I might have encountered and for continuing to check in for a chat about my progress even though you had changed job and moved across the world (also thank you for accepting our Skype calls from CHO retreat). And last but not least, thank you Nate for being a constant throughout my thesis. I'm not sure how you work but I'm grateful for your ability to quickly assess what I have presented to you and seeing a world of possibilities. I will always be appreciative of my stay with your group in San Diego.

It is invaluable to work in such a helpful and open group of researchers as in the CLED. It has been a pleasure working with every one of you over the years with all the challenges and joys moving location and changing personnel can bring. There has been a general air of support and exchange of knowledge that is rare to find. A warm thank you to all current and previous members. Special thanks to Song, Kai and Kim for joining the projects I have been working on and with it bringing an extra wave of energy and alternative approaches. Your collaboration has been priceless. Thank you Nachon for always providing a helping hand when needed and keeping the lab space in check. Thank you also to all previous and current members of CHO core - you have been extremely supportive by sharing your expertise. A special thank you to Sara for always being up for challenging the status quo. This also goes for the rest of the building. Thank you, Marta, for the chats and bike rides. I also want to extend a thank you to the service personnel and core labs of CFB for making the center run like a well-oiled machine.

During my PhD I was lucky enough to spend six months at UCSD, tapping into the world of computational biology. I want to thank the whole group for taking the time to guide a novice. Special thank you to Isaac for close collaboration and Shangzhong for keeping it cool even with the most frustrating data. Thank you, Austin, for your patience. Thank you,

Sharon, for advice on everything (the swim club is still alive). Special thanks to the whole group for every trip outside the lab - especially the food adventures - I never felt like a stranger.

And last, but not least, I want to thank my two rocks. I cannot have imagined doing this work without you. Lise, you have been there for me for the whole ride. You made me feel so welcome when I first joined the lab and have continued to do so throughout the years. It has been invaluable to have a sister in arms to go through this process with. Hooman, I'm not sure this, past year especially, would have been possible without you. Thank you for your endless positive attitude. Thank you for keeping me grounded when I was about to fly off to Venus and for reminding me of why we are here. There will be no future for trailing white spaces, I promise. I am grateful for both of your friendships and am looking forward to what the future brings.

A very special thank you to my friends and family for reminding me that there is a world outside of the lab, but also to support and take a genuine interest in what I was working with.

Julie.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

CAS	CRISPR-Associated
CHO	Chinese Hamster Ovary
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRISPRa	CRISPR activation
CRISPRi	CRISPR interference
crRNA	CRISPR RNA
DSB	Double Strand Break
gRNA	guide RNA
HDR	Homology Directed Repair
Indel	Insertion/Deletion
KRAB	Krüppel Associated Box
NCD	Non-Communicable Disease
NHEJ	Non-Homologous End Joining
PAM	Protospacer-Adjacent Motif
PTM	Post-Translational Modification
RNP	Ribonucleoprotein
RTA	R Transactivator
SAM	Synergistic Activation Mediator
TALEN	Transcription Activator-Like Effector Nuclease
T-PA	Tissue-Plasminogen Activator
tracrRNA	Trans-activating crRNA
TSS	Transcription Start Site
VPR	VP64-p65-rta
WHO	World Health Organization
ZFN	Zinc-Finger Nuclease

AIMS AND THESIS STRUCTURE

The aim of this thesis was to expand the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein (Cas) 9 toolbox for improved Chinese hamster ovary (CHO) cell line engineering, more specifically to explore the space surrounding large-scale pooled CRISPR/Cas9 screening in CHO cells. CRISPR mediated activation and repression of endogenous genes were implemented in CHO cells with the aim of utilizing them in conjunction with pooled CRISPR/Cas9 screening for increased versatility. Pooled CRISPR screening proves a promising tool for finding hitherto completely novel phenotype to genotype connections and unraveling the biological black box of CHO cells.

Chapter 1 presents an introduction to the manuscripts presented in the remainder of the thesis. This chapter has been created with short excerpts from Manuscript I. In **Chapter 2,** Manuscript II explores CRISPR interference (CRISPRi) for repression of endogenous genes in CHO cells and demonstrates the repression of genes involved in apoptosis. **Chapter 3** applies CRISPR activation (CRISPRa) techniques to CHO cells in order to force expression of previously unexpressed genes leading to altered glycosylation and amino acid essentiality. This work comprises part of Manuscript III. In **Chapter 4**, Manuscript IV presents a pooled CRISPR knockout screen used to identify novel targets to improve growth in media lacking glutamine. **Chapter 5** concludes the thesis by assessing the impact the new tools have on CHO cell engineering and perspectives of large-scale screening in CHO cells.

CHAPTER 1

INTRODUCTION

Contains short excerpts from Manuscript I

Biotherapeutic medicines are defined by the World Health Organization (WHO) as '*produced from biological sources such as living organisms rather than synthesised chemicals'* and are extremely important in today's treatment of diseases, especially non-communicable diseases (NCDs)--such as inflammation-related conditions, cancers, and diabetes--affecting millions of people worldwide[1,2]. Within cancer alone, in 2018 there were over 1.7 million new diagnoses in the United States[3] and more than 18 million new cases globally[4]. The rise in global NCD incidences is a result of a cocktail of causes largely related to the increased global socio-economic status, lengthening life-spans, a trend that is unlikely to stop anytime soon[1,5]. Great efforts in research and development of biotherapeutic medicines have provided optimism for treatment of previously untreatable diseases. However, the high costs associated with development and production have restricted the availability in developing countries, and put a substantial burden on the health care systems in higher income countries[6].

Biotherapeutics include diverse molecules such as clotting factors, cytokines, hormones, growth factors, and interferons, but are today largely dominated by monoclonal antibodies (mAbs)[7]; mAbs represent 53% of first time approved biopharmaceuticals between 2015 and July 2018[8]. The share of approvals for biotherapeutics has been more or less constant over the years 1995 to 2014; however, over the last four and a half years it has accelerated noticeably. From January 2014 to July 2018 biotherapeutics comprised 47% of all new drug approvals in the United States[8], a marked increase from previous years (21% from January 2006 to July 2010[9] and 26% from January 2010 to July 2014[10]) (Figure 1A). The global market for biotherapeutics is also expanding robustly, reaching US$188 billion in 2017[8]. The market for mAbs alone increased by 18.3% from 2016 to 2017 with sales exceeding US$98 billion. This has been steadily increasing from 2013 and is projected to further increase to US$130-200 billion by 2022[7] (Figure 1B).

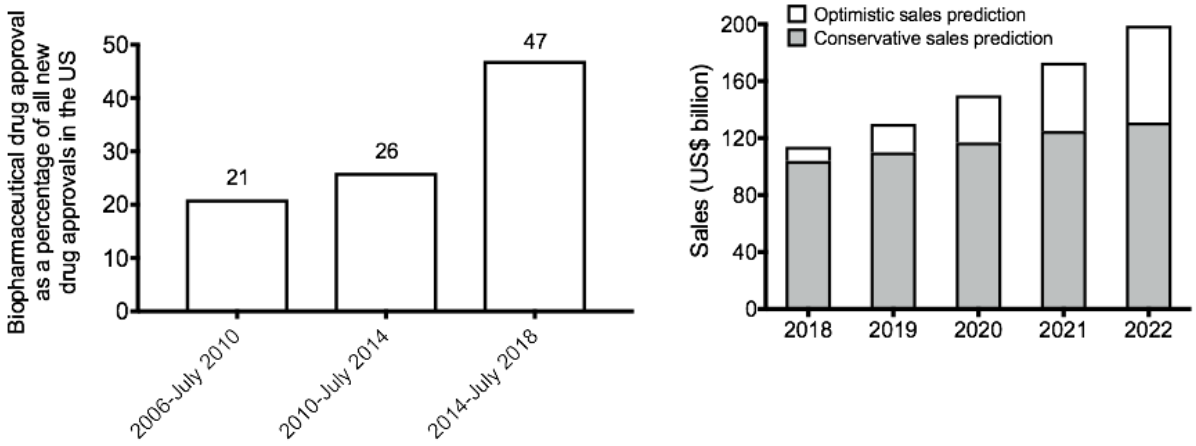


Figure 1 The biotherapeutic market

A) Overview of approvals of biopharmaceutical drugs as a percentage of all new drug approvals approved by US regulators (created with data from Walsh, 2018[8]) **B)** Prediction of mAb sales. The mAb market is expected to grow in the coming years (created with data from Grilo & Mantalaris, 2018[7]).

While at one point most biotherapeutics were extracted from tissues and bodily fluids, a very costly and complex affair, advances in recombinant DNA technology and cellular expression systems have played a central role in the explosive growth of the biotherapeutic field by driving down costs significantly. Of the expression hosts available, CHO cells have been the preferred host and the trend is continuing; 84% of mAbs approved by July 2018 were produced in CHO cells[8].

Chinese Hamster Ovary Cells

The Chinese hamster (*Cricetulus griseus*) was first brought into the laboratory in 1919 at Peking Union Medical College by Dr. E.T. Hsieh to use for typing pneumococcal bacteria, mainly because of the shortage of mice and the abundance of hamsters in the surrounding area. It soon proved that the hamsters were a useful model, as they could easily be infected to study a range of diseases. Moreover, it was found that the hamsters have few but large chromosomes, facilitating cytogenetic studies. They were gifted to Dr. Robert Briggs Watson

(picture 1B) who smuggled them out of China to the United States in 1948 and provided them to Victor Schwentker, a rodent breeder who was the first to successfully establish a colony outside of China. In 1951, George Yerganian established that the Chinese hamsters have a total of 22 chromosomes and he later took over breeding and distribution of Chinese hamsters to the scientific community after Schwentker ceased his sales in 1954[11] (Picture 1A and 1B).

Picture 1 History of CHO cells

A) Hamsters of one the original Chinese hamster litters from which todays colonies stem from. **B)** George Yerganian (right) and Robert Briggs Watson (left). Both images are sourced from Gottesman, 1985[11].

Cell lines derived from the Chinese hamster were first established by Dr. Theodore Puck in 1957. Already having successfully established clonal HeLa cell cultures, Puck eyed the potential of the Chinese hamster. He extracted primary cells from hamster ovarian tissue retrieved from Yerganian's colonies and succeeded in establishing a cell line from the fibroblast-like cells via apparent spontaneous immortalization while growing them for over 10 months[11,12]. The cells were widely distributed and extensively used to study mammalian genetics and mammalian cell physiology due to their fast growth, robustness, and ease of use[13]. Although not mentioned in Puck's original paper, CHO cells are believed to have a clonal origin supported by the fact that all CHO cell lines are deficient in proline synthesis[14]. The cells were adapted to grow in serum-free media in 1977[15] allowing for cell growth in defined media, an important trait for simplifying downstream processes and regulatory approval. Since

the original CHO cell was isolated, several cell lines with unique properties have emerged of which extensive reviews have been written[12,16]. Their heterogeneity is both displayed in their genetic make-up and their phenotypic characteristics. The most popular CHO cell lines in use include CHO-K1, CHO-DG44, CHO-S, and CHO-DXB11 (see Figure 2)[13,17].

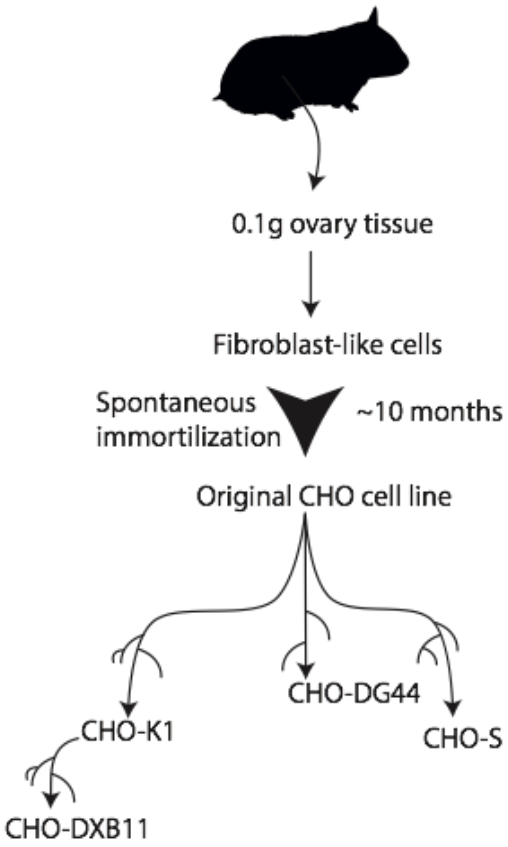


Figure 2 The origin of CHO cell lines

Ovary tissue was resected from a Chinese hamster (*Cricetulus griseus*) and fibroblast-like cells were extracted. These were grown in culture for over 10 months during which they ostensibly underwent spontaneous immortalization resulting in the original CHO cell line from which all current CHO cell lines originate.

Since the first recombinant protein produced in CHO cells, human tissue plasminogen activator (t-PA, Activase®, Genentech), was approved for clinical use in 1987[10] CHO cells have been the preferred workhorse for the industry, largely owing to their ability to perform human-like post transcriptional modifications (PTMs), fast growth, ease of manipulation, resistance to viruses infecting human cells and acceptance from regulatory agencies[18]. Although there has been a steady increase in production rates over the past three decades, major advances have centered around optimizing external factors such as expression vector design[19], expression vector selection and gene amplification[20,21], culture optimization[22] and selection of high producing clones[23]. While CHO cells have been used effectively for biotherapeutic protein production, it is likely that their ovarian cell origin--not a traditional 'professional secretor'--means that their internal machinery is not optimized for protein secretion. Rational engineering is an attractive path to improving these cells[24], however this has been nigh impossible until recently. Advances in both understanding the genomic basis of CHO and development of molecular engineering tools have ushered in a new era of rational data-driven and precise engineering.

The Genomic Basis of CHO Cells

For decades, the work in CHO has been carried out without an understanding of the specific molecular components, and their genetic basis, that underpin biotherapeutic protein production in this cell line. Advancements in sequencing technology and associated drop in cost quickly changed that. The first complete genome sequence of CHO-K1 was released in 2011[25] and since then several efforts have been carried out to sequence the genomes of the Chinese hamster and different CHO cell lines[17,26,27]. Recently an updated reference genome of the Chinese hamster has been published that is similar in quality to that of other model organisms such as mouse and rat[28], opening up new opportunities for CHO genome engineering.

Genome Engineering Tools

Currently, three major classes of programmable nucleases are in use for targeted genome engineering of virtually any site in the genome; zinc-finger nucleases (ZFNs), transcription activator-like Effector nucleases (TALENs) and the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein (Cas) system[29,30]. The CRISPR/Cas system is the most recently developed, and because of its relative ease, efficiency, and affordability it has quickly become a routine genome editing tool for many labs[29,30]. Its application to improve CHO cell line engineering will be the focus of this thesis.

CRISPR/Cas9 Genome Editing

CRISPR/Cas systems have evolved in prokaryotes as a defense mechanism against foreign genetic elements such as viruses[31]. The most investigated CRISPR/Cas system is composed of three main components: Cas9 endonuclease, CRISPR RNA (crRNA), and trans-activating crRNA (tracrRNA). In vivo, crRNAs are processed from CRISPR arrays which are clusters of repeat sequences interspaced by variable sequences (spacers) homologous to, and derived from, foreign genetic elements (protospacers). Transcribed crRNA and tracrRNA hybridize to a RNA duplex known as guide RNA (gRNA) that binds to the Cas9 protein and form an active ribonucleoprotein (RNP) complex. With the guidance of the gRNA, this complex searches for complementary sequences in foreign DNA and directs its cleavage by two nuclease domains of Cas9. Each nuclease domain cleaves one strand of DNA, hence Cas9 activity leads to a double-stranded break (DSB) of DNA[32].

The breakthrough in the CRISPR/Cas9 field happened when it was shown that the bacterial CRISPR/Cas9 system could be used for genome editing in eukaryotic cells, especially in mammalian cells[33,34]. To simplify the system, the crRNA:tracrRNA duplex was fused into a chimeric single gRNA. The gRNA binds to a target DNA through an approximately 20-nucleotide long region that is adjacent to the protospacer-adjacent motif (PAM), which is recognized by the PAM-interacting domain of Cas9. Thus, by customizing a 20-nt region of the gRNA to pair with the DNA sequence of interest, Cas9 can be targeted to any genomic

locus containing a PAM sequence, making it an easily programmable platform for genome editing. Since the first publications of CRISPR/Cas9 application in mammalian cells, diverse CRISPR-based tools have been developed for gene editing, gene regulation[35], epigenetic modification[36], chromatin remodeling[37], and genome imaging[38].

CRISPR/Cas9-Mediated Gene Knockout and Integration

Gene disruption was the first application of CRISPR/Cas9 in mammalian cells and remains the most used method for functional knockout of the target genes[33,34]. The easiest way to create a knockout cell line is to introduce two components into the cell: a gRNA specific to the target sequence and a Cas9 protein. DSBs induced by CRISPR/Cas9 are preferentially repaired by error-prone non-homologous end joining (NHEJ), leading to generation of insertion/deletion (indel) mutations. These mutations can cause a frameshift in the coding regions of genes that disrupts their proper translation and results in a functional knockout (Figure 3A). CRISPR/Cas9 can also be harnessed to perform site-specific integration of DNA by taking advantage of another central DNA repair mechanism, homology directed repair (HDR). Where NHEJ most-often leads to disruption of genes, HDR allows precise targeted gene integration. When a proper donor DNA template is presented, HDR of CRISPR/Cas9-introduced DSBs near the targeted genomic loci can lead to the error-free introduction of desired mutations or insertion of a desired gene[39] (Figure 3B).

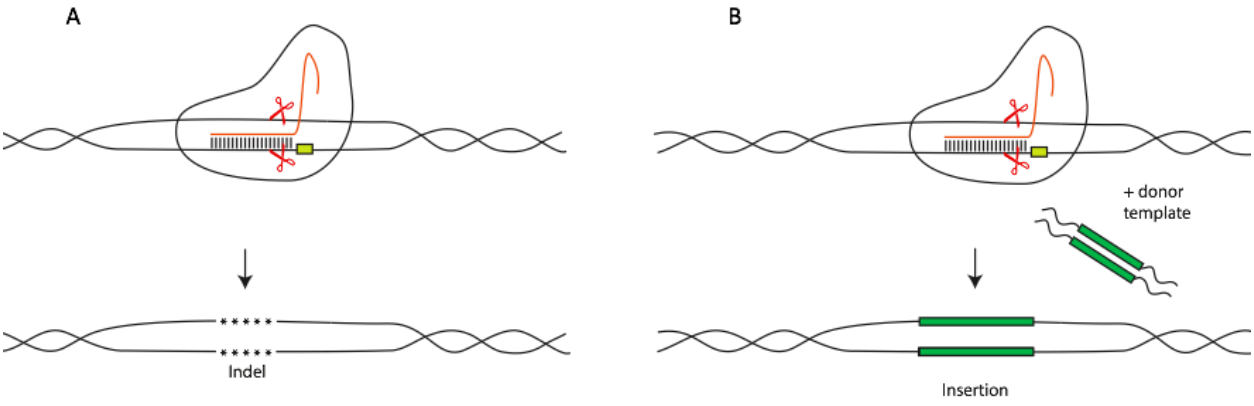


Figure 3 CRISPR mediated knockout and integration

A) CRISPR/Cas9 mediated knockout. A ~20bp gRNA (red) adjacent to a PAM site (yellow) guides the Cas9 to the site of interest and creates a double strand break. In most cases the non-homologous end joining repair pathway fixes the break and creates a mutation in the target region leading to a functional knockout. **B)** The double strand break can also be repaired by homology directed repair, which can lead to site-specific integration when a DNA donor template (green) with correct homology arms is provided.

CRISPR/Cas9-Mediated Genome Modification

With CRISPR gaining foothold as a powerful genome targeting method, many have started repurposing the system for site-specific purposes other than cleavage. The catalytically inactive version of Cas9 (dCas9) combined with different effector molecules has been used for various applications including transcriptional regulation[35], epigenetic modification of DNA or histones[36], and genome imaging[38] for tracking genomic structure and dynamics.

Transcriptional Regulation

The first use of CRISPR as a transcriptional regulator was performed in bacteria and relied on targeting the gRNA and dCas9 to the gene of interest to sterically hinder RNA polymerase from binding (Figure 4A). This led to considerable repression of specific transcription and the tool was coined CRISPR interference (CRISPRi)[35]. The effect in

eukaryotic cells, however, was moderate, likely because of the complexity of transcriptional regulation in these higher-order organisms. Follow up studies have shown that fusing dCas9 with repressing regulatory elements, for example, the repressive chromatin modifier domain Krüppel Associated Box (KRAB[40]), could yield improved transcriptional repression. Stable (as opposed to transient) expression of dCas9 further increased repression, regardless of whether the gRNA was transiently or stably expressed[41].

In line with the principles underpinning CRISPRi, it has been explored whether fusing dCas9 with activating regulatory elements could increase the expression of targeted genes, referred to as CRISPR activation (CRISPRa) (Figure 4B). Initial experiments have shown that fusing well-known transcription activators, such as the VP16 (a viral transactivator that forms an activation complex with host transcription factors[42]) or p65 (a transcription activating subunit of transcription factor NF-kappa-B[43]) activation domains to dCas9 and co-expressing with gRNA increases expression of targeted genes[41]. Several different combinations of gRNAs and activation effectors have since been tested across various cell types in various settings. A comparison of so-called second-generation activators including VP64 (4 copies of VP16), VPR (a combinatorial activator comprising VP64, p65, and the Epstein-Barr virus R transactivator (rta)[44]), SAM (synergistic activation mediator[45]), and SunTag[46], in human, mouse, and fly cell lines revealed that SAM, Suntag, and VPR were consistently superior at increasing expression across cell lines[47].

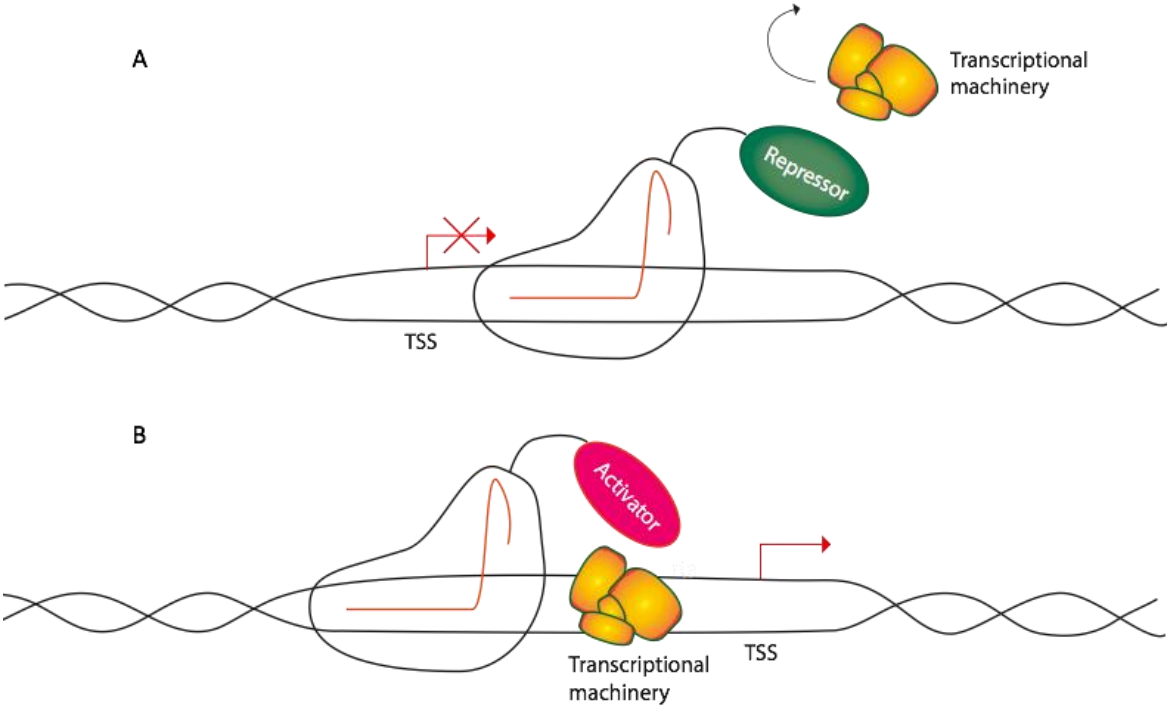


Figure 4: CRISPR mediated repression and activation

A catalytically inactive dCas9 is fused to a transcriptional repressor (green) or activator (magenta) and guided by the gRNA (red) to a region down or upstream of the transcription start site (TSS) where it **A**) represses or **B**) enhances transcription of the target gene.

CRISPRi and CRISPRa have since been applied to a variety of organisms with varied success rates. For better reproducibility of CRISPR-based regulation between genes, evidence suggests that careful design of the gRNA in relation to the transcription start site (TSS) as well as nucleosome occupancy of the target site is of great importance[47–49].

CRISPR/Cas9-Mediated Genome-Scale Screening

In addition to the aforementioned applications of CRISPR/Cas9 for gene editing and regulation, the technique and said tools can be extended to high-throughput forward genetic screens that can be used not only to interrogate gene function but also to identify novel targets

for treatment of disease or engineering industrial mammalian cells. In contrast to reverse genetic engineering, where known genes are modulated and the resulting phenotype is studied, the approach taken in forward genetic screening includes changing numerous genes in a pool of cells, applying a phenotypic selection and subsequently identifying which genetic changes are responsible for cellular behavior. Traditionally these screens have relied on random mutagenesis and isolating individuals with an interesting phenotype[50–53]. As there are many mutations from this process, identification of causality is a very lengthy and difficult process and represents one of the main weaknesses of this system. With the advancement of RNA interference (RNAi) techniques came a much-welcomed alternative approach to forward genetic screening. RNAi targeting specific mRNAs replaced random mutagenesis and the identification of causal genetic alterations was much simplified[54]. Many discoveries have been made using RNAi screens[55–58], however, the incomplete knockdown of targeted genes and substantial off-target effects have restrained the extent to which RNAi screens can be used[59]. The advancement of CRISPR technologies presents a novel approach to attempt to solve these issues and increase the versatility for the next generation of forward screening methods. Indeed, as early as 2014, genome-scale CRISPR knockout screens were applied to human cell lines[60,61], and alternative effectors such as CRISPRi and CRISPRa have since been included in the screening setup[62–64]. Today, results from large-scale CRISPR screens are being published at a fast-moving pace with varied purposes and designs[65–70].

CRISPR in CHO

The first demonstration of CRISPR/Cas9 activity in CHO cells was performed in our lab and published in 2014 when two genes involved in glycosylation (C1GALT1 specific chaperone 1 (*Cosmc*) and alpha-1,6-fucosyltransferase (*Fut8*)) were disrupted with indel frequencies up to 47% in a pool of cells[71]. Later, the enrichment of cells carrying GFP-Cas9 plasmid and multiplexing knockout strategies were developed[72]. Simultaneous introduction of three gRNA plasmids and fluorescent enrichment yielded a triplex indel frequency of 59% in *Fut8*, *Bax,* and *Bak* genes. The knockout cell lines showed improved resistance to apoptosis

and disrupted fucosylation activity, relevant for prolonged cell cultivation and production of non-fucosylated therapeutic proteins.

Since the successful application of CRISPR/Cas9-mediated gene editing, various CHO engineering strategies employing CRISPR/Cas9 have been developed in an attempt to improve product quality and to expedite cell line development for high productivity[73–78]. The successes lead us to investigate CRISPR tools that have not been applied in CHO cells yet, namely CRISPRi and CRISPRa of endogenous genes and pooled CRISPR knockout screens with the ultimate goal of providing the tools for performing versatile genome-wide CRISPR screens in CHO cells.

CHAPTER 2

CRISPR MEDIATED REPRESSION

Manuscript II

Reduced Apoptosis in Chinese Hamster Ovary Cells via Optimized CRISPR Interference

Kai Xiong[1*], Kim Fabiano Marquart[1,2*], Karen Julie la Cour Karottki[1*], Shangzhong Li[3,4,10], Isaac Shamie[3,10], Jae Seong Lee[5], Signe Gerling[1,6], Nan Cher Yeo[7], Alejandro Chavez[8], Gyun Min Lee[1,9], Nathan E. Lewis[3,4,10,§], Helene Faustrup Kildegaard[1,11,§]

[1]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Denmark.

[2]Integrated Biologics Profiling/Cell line development, Novartis, Basel, Switzerland (current address).

[3]Department of Pediatrics, University of California, San Diego.

[4]Department of Bioengineering, University of California, San Diego.

[5]Department of Molecular Science and Technology, Ajou University, Suwon, Republic of Korea.

[6]Novo Nordisk A/S, Hillerød, Denmark

[7]Department of Pharmacology and Toxicology, Precision Medicine Institute, University of Alabama, Birmingham, USA.

[8]Department of Pathology and Cell Biology, Columbia University College of Physicians and Surgeons, New York, New York, USA.

[9]Department of Biological Sciences, KAIST, Daejeon, Republic of Korea.

[10]The Novo Nordisk Foundation Center for Biosustainability, University of California, San Diego, United States.

[11]Novo Nordisk A/S, Måløv, Denmark (current address).

*These authors contributed equally to this publication.

§ Correspondence to: Helene Faustrup Kildegaard (hef@biosustain.dtu.dk) and Nathan E. Lewis (nlewisres@ucsd.edu).

Keywords: CHO, CRISPRi, apoptosis, mitochondrial membrane integrity, Caspase, *Bax*, *Bak*, *Casp3*.

Abstract

Chinese hamster ovary (CHO) cells are widely used for biopharmaceutical protein production. One challenge limiting CHO cell productivity is apoptosis stemming from cellular stress during protein production. Here we applied CRISPR interference (CRISPRi) to downregulate the endogenous expression of apoptotic genes *Bak*, *Bax*, and *Casp3* in CHO cells. In addition to reduced apoptosis, mitochondrial membrane integrity was improved and the caspase activity was reduced. Moreover, we optimized the CRISPRi system to enhance the gene repression efficiency in CHO cells by testing different repressor fusion types. An improved Cas9 repressor has been identified by applying C-terminal fusion of a bipartite repressor domain, KRAB–MeCP2, to nuclease-deficient Cas9. These results collectively demonstrate that CHO cells can be rescued from cell apoptosis by targeted gene repression using the CRISPRi system.

Chinese hamster ovary (CHO) cells are the most commonly used host cells for the production of therapeutic and diagnostic proteins[10]. During protein production, metabolic stresses in CHO cells can induce apoptosis and thus reduce the viability and productivity of cells[79]. Suppressing apoptosis in CHO cell culture is therefore a crucial challenge faced by academia and industry. Under cellular stress, pro-apoptotic proteins *Bak* (Bcl-2 homologous antagonist/killer) and *Bax* (Bcl-2 associated X) are involved in permeabilization of the mitochondrial outer membrane, thus leading to release of cytochrome c[80]. The released cytochrome c can then activate caspases (cysteine-dependent aspartyl-specific proteases) to cleave several hundred cellular proteins for further programmed apoptosis[80,81]. To overcome apoptosis in CHO cells based on this known pathway, apoptosis-resistance had been shown by gene inhibition or deletion of *Bak* and *Bax*[72,82,83] as well as by gene inhibition of *Casp3* and *Casp7*[84,85].

The CRISPR technology has emerged as an efficient method for precise genetic and translational modification. A simplified CRISPR system generally includes two components: a guide RNA (gRNA) and a Cas9 protein. The gRNA confers target specificity via DNA-RNA base-pairing and the gRNA-Cas9 complex forms to mediate DNA cleavage at the target site[33,34]. Moreover, a catalytically inactive version of Cas9 (dCas9) can be generated by mutagenesis to remove the nuclease activity of Cas9. RNA-guided dCas9 can be applied to mediate gene repression by binding to the region proximate to the transcription start site (TSS) of a target gene thus blocking transcription initiation and RNA polymerase elongation[35]. CRISPRi has previously been shown to repress a transgene *Dhfr* (coding dihydrofolate reductase) in CHO cells to enhance protein production[78]; however, the endogenous gene repression using CRISPRi in CHO cells has not yet been demonstrated.

In this study, we use CRISPRi to downregulate the endogenous expression of apoptotic genes, including *Bak*, *Bax,* and *Casp3* (Figure 1a). To test the effect of CRISPRi of on apoptotic genes in CHO cells, an N-terminal KRAB (Krüppel-associated box, transcription repressor) domain was fused to a CHO codon-optimized dCas9 to generate a recombinant KRAB-dCas9 repressor and three gRNA vectors were designed to target each apoptotic gene. First, in a pilot experiment (see Supplemental Figure S1), KRAB-dCas9 and the individual gRNAs were co-

transfected into CHO-S cells to test all gRNAs individually or in combination. For each test, we measured the transcriptional fold change of each target gene by quantitative reverse-transcription PCR (qRT-PCR). Then, the temporal effect of CRISPRi of *Bak* and *Bax* was investigated. Collecting samples 24 hours after applying CRISPRi demonstrated the most efficient repression compared to 48 and 72 hours (Supplemental Figure S2). Based on these data, the best gRNA (gRNA3-Bak, gRNA3-Bax, gRNA1-Casp3) or a mix of all three gRNA for each gene was applied in following experiments and sampling was performed 24 hours after transfection. As shown in Figure 1b, we further confirmed that significant gene repression of *Bak* and *Bax* mediated by CRISPRi was achieved by transfecting the gRNA mix and KRAB-dCas9; however, when we again transfected KRAB-dCas9 and each best gRNA from our pilot experiment, the CRISPRi-mediated downregulation was only evident for *Bax* (Figure 1b). In order to improve the repression efficiency of CRISPRi, we established a KRAB-dCas9-CHO cell line displaying stable and high expression of KRAB-dCas9 by precisely integrating a KRAB-dCas9 expression cassette into a transcriptionally active site (Figure 1c). This integration site was previously selected as an active genome site surrounded by highly expressing genes based on our previous study[86]. In these established KRAB-dCas9-CHO cells, CRISPRi-mediated gene repression of *Bak*, *Bax,* and *Casp3* was achieved by transfecting the best gRNA. Moreover, all apoptotic genes have successfully been repressed when transfected with their respective gRNA mixes (Figure 1d). Taken together, using the combination of a gRNA mix and constitutive KRAB-dCas9 in CRISPRi displayed a more powerful repression, consistent with previous literature[87]. In addition, the potential of using CRISPRi for simultaneous inhibition of multiple genes (*Bak*+*Bax*) was also investigated in KRAB-dCas9-CHO cells; however, only *Bax* showed significant repression by CRISPRi of *Bak*+*Bax* combinations (Supplemental Figure S3). Correspondingly, the a modest (but not significant) increase of viable cell density (VCD) was observed after the CRISPRi treatment (Supplemental Figure S4).

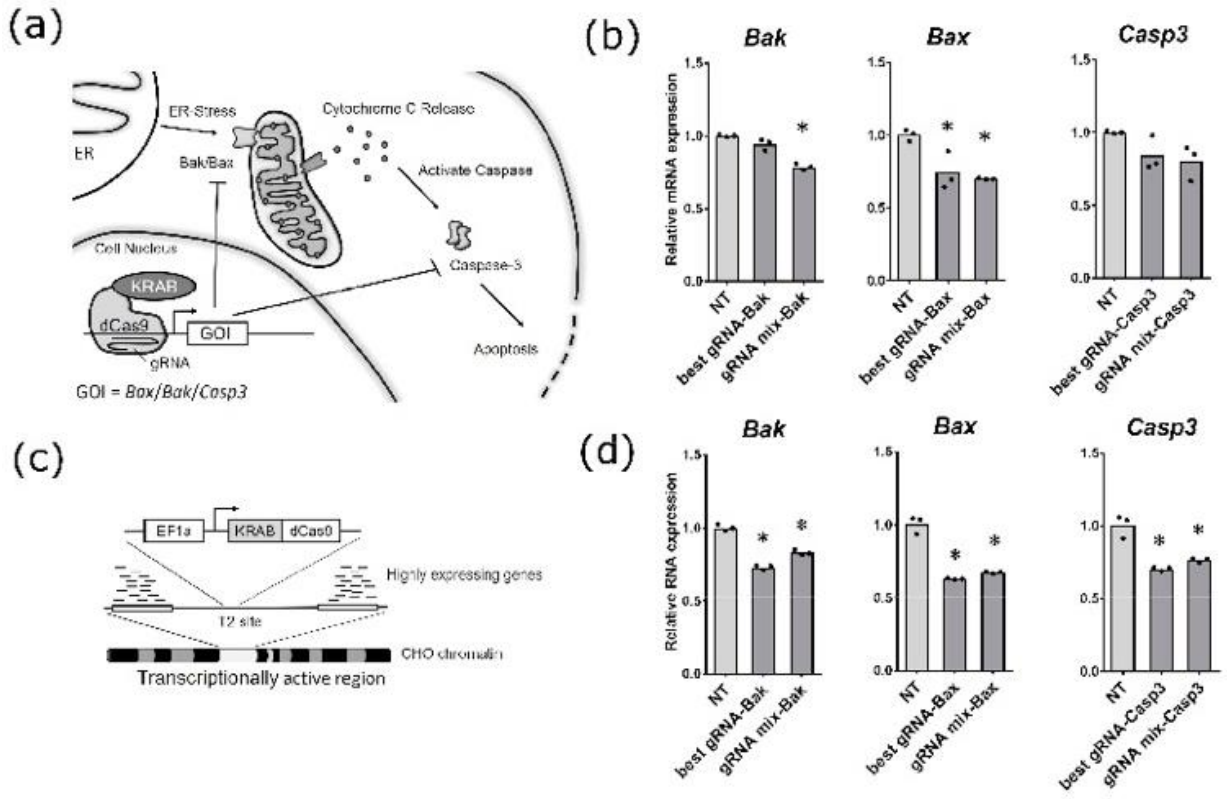


Figure 1 Characterization of CRISPRi induced repression of apoptotic genes in CHO cells

(a) Schematic illustration of the application of CRISPRi in disrupting the apoptosis pathway in CHO cells. (b) Transcriptional repression of *Bak*, *Bax,* and *Casp3* via transient expression of gRNA and KRAB-dCas9 in CHO cell measured by qRT-PCR. (c) Schematic illustration of the KRAB-dCas9-CHO cell line with precise integration of the KRAB-dCas9 expression cassette. (d) Transcriptional repression of *Bak*, *Bax,* and *Casp3* via transient expression of gRNA in KRAB-dCas9-CHO cells measured by qRT-PCR. In (b) and (d), The best gRNA or the mixture of three designed gRNAs for each gene were tested in this study. mRNA was harvested 24h after transfection. NT indicates transfection with scrambled non-target gRNA and KRAB-dCas9 as a negative control. Targeted gene expression fold was shown as mean of the biological triplicates (n=3, * indicates $p<0.05$ vs NT) which were normalized to *Fkbp1a* and *Gnb1* as housekeeping genes. Each transfection replicate was measured with technical triplicates of qRT-PCR. Dot represents the mean of technical triplicates in qRT-PCR.

To investigate if CRISPRi repression of apoptotic genes can reduce the apoptosis in CHO, we transfected gRNAs targeting apoptotic genes in KRAB-dCas9-CHO cells. After 24 hours, we induced apoptosis by using Actinomycin D (ActD). 24 hours post induction, a mitochondrial potential assay was performed. In apoptotic cells, depolarized mitochondria lose the ability to aggregate JC-1 (5',6',6'-tetrachloro-1,1',3,3'- tetraethylbenzimidazolyl- carbocyanine iodide) monomers. Therefore, the ratio of JC-1 aggregates to monomers is an indicator of mitochondrial membrane integrity (Figure 2a). CCCP (Carbonyl cyanide m-chlorophenyl hydrazone), a mitochondrial membrane disruptor, was used as a positive control to fully depolarize mitochondria. As the results show, CRISPRi of *Bak* and *Bax* rescued the decreasing ratio of JC-1 aggregates to monomers after ActD induction compared with the negative control transfected with non-targeting gRNA (Figure 2b), indicating CRISPRi of apoptotic genes inhibits the depolarization of mitochondria. In addition, a caspase activity assay was performed by incubating cells with a DEVD peptide linked fluorescein (FITC). Caspase 3/7 in apoptotic cells can mediate the cleavage of the peptide and release the FITC to stain the cell DNA (Figure 2c). As expected, ActD induction enhances the caspase activity by increasing the percentage of FITC positive cells from basal levels of around 10% to around 90%; however, CRISPRi of *Casp3* decreases caspase activity in apoptotic cells by reducing FITC positive cells by ~10% (Figure 2d). In summary, CRISPRi can reduce the apoptotic phenotype in CHO cells. Along with the reduced apoptotic phenotype (Figure 2), there is an increase of VCD, albeit not significant (Supplemental Figure S5).

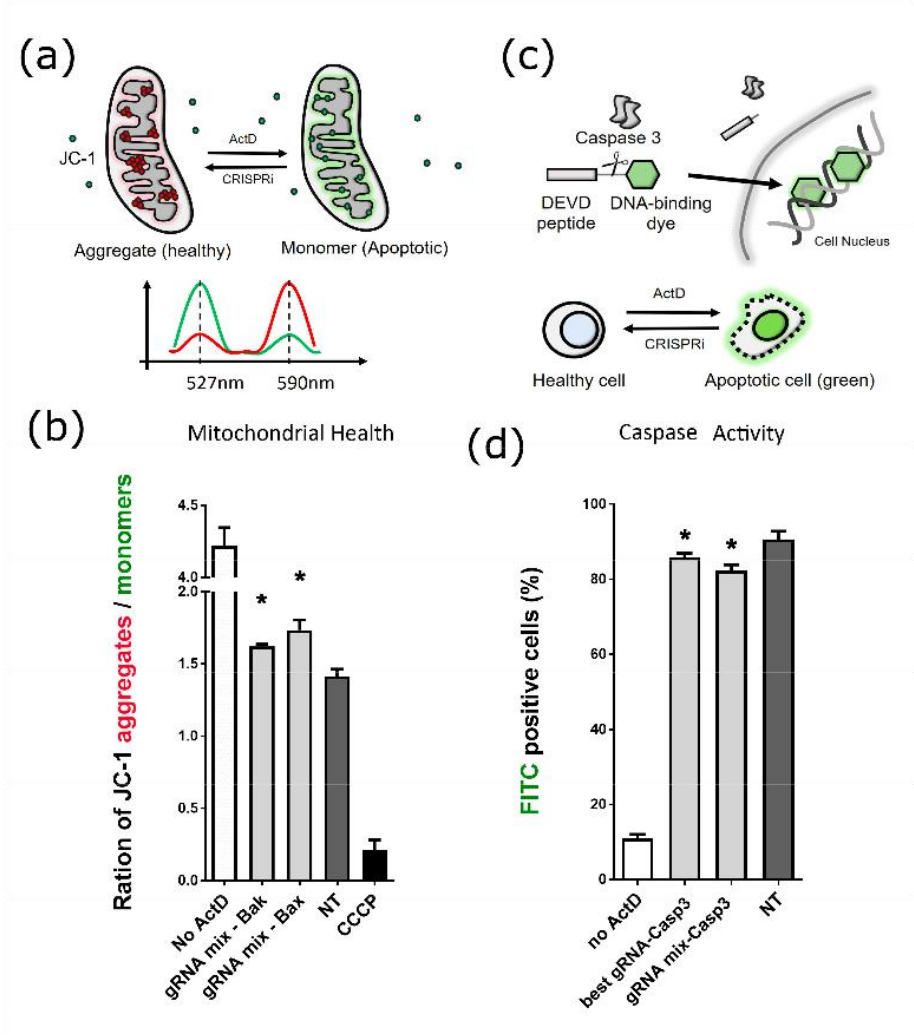


Figure 2 Reduced apoptotic phenotype via CRISPRi

CHO apoptosis was induced by ActD 24 hours after transfection of gRNA in KRAB-dCas9-CHO cells. Another 24 hours after induction, cells were harvested and analyzed for changes in the apoptotic phenotype. (a) Schematic illustration of determination of mitochondrial membrane integrity. (b) Mitochondrial membrane integrity after CRISPRi of apoptotic genes in KRAB-dCas9-CHO cells. The mixture of three gRNAs for each gene was tested in this study. (c) Schematic illustration of caspase activity analysis. (d) Caspase activity after CRISPRi of apoptotic genes in KRAB-dCas9-CHO cells. The best gRNA or the mixture of three gRNAs for *Casp3* were tested in this study. NT represents transfection with scrambled non-target gRNA. CCCP represents carbonyl cyanide m-chlorophenyl hydrazone, a mitochondrial membrane disruptor to fully depolarize mitochondria. Each treatment was performed in biological triplicates (n=3, * indicates $p<0.05$ vs NT).

With the significant decrease in apoptosis and only modest increases in VCD, we wondered if the CRISPRi system could be enhanced. Although previous studies have shown that the N-terminal KRAB fusion works more effectively than a C-terminal fusion for CRISPRi in human cells[41,63], the terminal-effects of KRAB fusion in CHO cells remained unknown. To address this, another dCas9 repression construct was designed by fusing a C-terminal KRAB to dCas9 to obtain a dCas9-KRAB construct (Figure 3a). As the results revealed by qRT-PCR (Figure 3b), C-terminal KRAB in dCas9-KRAB constructs demonstrated more efficient repression compared to KRAB-dCas9 (indicated with # in Figure 3b), which is, interestingly, in contrast to the conclusion in human cells.

We further evaluated the impact of using a second repressor domain to further enhance the suppression of apoptosis and increase VCD. Thus, we fused the C-terminal MeCP2 transcriptional repression domain (TRD)[88,89] to dCas9-KRAB to generate a dCas9-KRAB-MeCP2 construct (Figure 3a). As predicted, the dCas9-KRAB-MeCP2 construct showed enhanced repression efficiency in 5 out of 8 cases compared to KRAB-dCas9 (indicated with # in Figure 3b) and in 4 out of 8 cases compared to dCas9-KRAB (indicated with § in Figure 3b). We thus find that all three tested CRISPRi systems could mediate the target gene repression in most cases (21 out of 24, indicated with *), albeit with different efficacies. Moreover, while the standard CRISPRi experiments only showed a modest increase in VCD (Supplementary Figure S4), the VCD after CRISPRi improved significantly when using the optimized dCas9-KRAB and dCas9-KRAB-MeCP2 constructs (Figure 3c). Notably, use of a C-terminal KRAB and fusing a MeCP2 repressor to dCas9 constructs can improve the gene repression efficiency of CRISPRi in CHO cells.

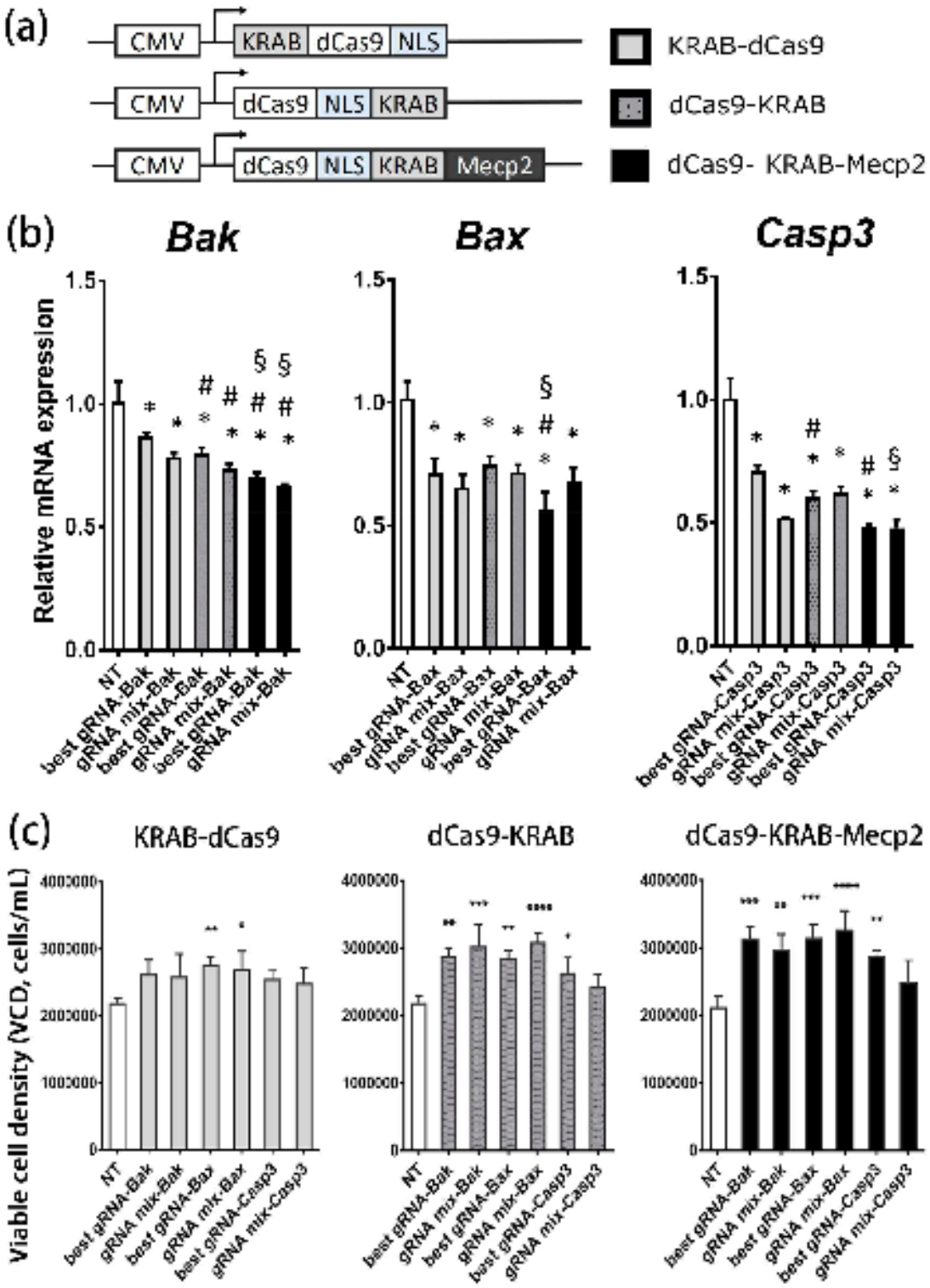

(a)
CMV
KRAB dCas9 NLS
KRAB-dCas9
CMV
dCas9 NLS KRAB
dCas9-KRAB
CMV
dCas9 NLS KRAB Mecp2
dCas9- KRAB-Mecp2
(b)
Bak
Bax
Casp3
Relative mRNA expression
1.5
1.0
0.5
0.0
NT
best gRNA-Bak
gRNA mix-Bak
best gRNA-Bax
gRNA mix-Bax
best gRNA-Casp3
gRNA mix-Casp3
(c)
KRAB-dCas9
dCas9-KRAB
dCas9-KRAB-Mecp2
Viable cell density (VCD, cells/mL)
4000000
3000000
2000000
1000000
0

Figure 3 CRISPRi optimization in CHO cells

(a) Schematic comparison of CRISPRi vectors. NLS, nuclear localization signal; CMV, cytomegalovirus promoter; MeCP2, gene element that encodes the transcriptional repression domain of methyl CpG binding protein 2. (b) Transcriptional repression of *Bak*, *Bax* ,and *Casp3* via the three CRISPRi systems we tested was measured by qRT-PCR. We tested the best gRNA (gRNA3 for *Bax*, gRNA3 for *Bak*, and gRNA1 for *Casp3*) and the mixture of designed gRNAs for each gene. gRNA and dCas9-repressors were transiently co-transfected and the mRNA samples was harvested 24 hours after transfection. NT represents transfection with scrambled non-target gRNA and dCas9-KRAB-MeCP2. Targeted gene expression fold change was shown as mean of the biological triplicates (n=3, * indicates $p<0.05$ vs NT, # indicates $p<0.05$ vs KRAB-dCas9, § indicates $p<0.05$ vs dCas9-KRAB). Fold change was normalized to *Fkbp1a* and *Gnb1* as housekeeping genes. Each biological replicate was measured with technical triplicates of qRT-PCR. Error bar represents the standard deviation of transfection replicates. (c) The viable cell density of the groups in (b) was measured by flow cytometry 24 hours after transfection (n=3, * indicates $p<0.05$ vs NT, ** indicates $p<0.01$ vs NT, *** indicates $p<0.001$ vs NT, **** indicates $p<0.0001$).

In summary, we found that CRISPRi can be successfully applied to repress endogenous expression of apoptotic genes and rescue CHO cells from cell stress induced apoptosis. Thus, the CRISPRi system joins other tools available for gene knockdowns, such as RNAi. However, CRISPRi provides unique advantages, including potentially lower off-target effects than RNAi[90] and advantages in that CRISPRi suppresses transcription, while RNAi induces RNA degradation. These tools, however, will be complementary and enable pooled loss-of-function screens to identify genes associated with desired or undesirable traits, thereby guiding targeted cell line engineering. Thus, this platform enables gene function dissection and provides a new horizon for CHO cell engineering.

gRNA Design, Vector Construction and Generation of a KRAB-dCas9-CHO Cell Line

gRNAs were designed to mediate CRISPRi of *Bak*, *Bax,* and *Casp3* (see Supplemental Table S1 for target sequences and Supplemental Table S2 for oligos used for cloning). gRNAs were combined with either transient delivery of KRAB-dCas9, dCas9-KRAB, and dCas9-KRAB-MeCP2 or delivered to cells stably expressing KRAB-dCas9 (Supplemental Figure S6). Detailed materials and methods for gRNA design, vector construction and generation of a KRAB-dCas9-CHO cell line are elaborated in supplementary information.

Cell Culture, Transfection and Apoptosis Induction

CHO-S cells (Thermo Fisher Scientific) were maintained in CD CHO medium supplemented with 8 mM L-Glutamine (Thermo Fisher Scientific) and 2 µL/mL Anti-Clumping reagent (AC, Life Technologies). Cells were cultivated in 125 mL Erlenmeyer shake flasks (Corning Inc., Acton, MA), incubated in a humidified incubator at 37˚C, 5% CO_2 at 120 rpm and passaged every 2-3 days. One day before transfection, cells were seeded into the culture medium without AC. On the day of transfection, cells were transfected using a total amount of 3.75 µg of DNA packaged with 3.75 µL of FreeStyle Max transfection reagent (Thermo Fisher Scientific) per well of a 6-well plate (BD Biosciences) in 3 mL culture medium without AC at the adjusted density of 1.0×10^6 viable cells/mL (details about ratio of gRNA and dCas9 repressor for transfection in Supplementary Table S3). All transfections were performed in biological duplicates or triplicates. To induce apoptosis, 10 µg/mL of ActD (Sigma Aldrich) was added into the culture of KRAB-dCas9-CHO cells at 24 hours after transfection with or without gRNA. 24 hours post ActD induction, cells were harvested for detection of mitochondrial membrane potential and caspase activity as described below.

Detection of Mitochondrial Membrane Potential

For each sample, 1.0×10^6 cells were pelleted and resuspended in 1 mL warm phosphate-buffered saline (PBS, Sigma-Aldrich). For the apoptosis control, 1 µL of 50 nM

CCCP was added and incubated for 5 minutes at 37 °C. 10 μL of 200 μM JC-1 was added to each sample and the cells were incubated for 30 minutes at 37 °C, 5% CO_2. After incubation, cells were washed and resuspended in PBS and measured via MACSQuant® VYB flow cytometry (Miltenyi Biotec, Bergisch Gladbach, Germany) using 488 nm excitation and 529 nm (JC-1 monomers) + 590 nm (JC-1 aggregates) emission. The ratio of JC-1 aggregates to monomers was calculated as a measure of mitochondrial membrane integrity.

Caspase Activity Assay

Caspase activity was detected by using CellEvent™ Caspase-3/7 Flow Cytometry Assay Kit (Life Technologies). For each sample, 1×10^6 cells were pelleted and resuspended in 1 mL warm PBS and mixed with CellEvent™ Caspase 3/7 detection reagent to a final concentration of 500 nM. After 25 minutes of incubation at 37 °C and 5 % CO_2, the SYTOX® AADvanced™ reagent was added at a final concentration of 100 μM. After an additional incubation of 5 min at 37 °C and 5 % CO_2, the percentage of FITC positive cells was measured via MACSQuant® VYB flow cytometry.

Quantitative Detection of Gene Transcription

After CRISPRi, transcription level changes of each target gene were detected by qRT-PCR. The primers and probes for the qRT-PCR were designed using PrimerQuest™ (Integrated DNA Technologies, Coralville, IA) and listed in Supplemental Table S4. For each group 10^6 cells were spun at 200xg, supernatant was removed, and total RNA was extracted using the RNeasy Plus Mini Kit (Qiagen, Venlo, Netherlands) according to the manufacturer's instructions, followed by the quality and quantity measurement on the NanoDrop spectrophotometer 2000 (Thermo Fisher Scientific). Isolated RNA was reverse transcribed into cDNA using the Maxima First strand kit (Thermo Fisher Scientific), followed by qRT-PCR using Taqman assay in the QuantStudio 5 instrument (Applied Biosystems, Waltham, USA). Each sample had 3 technical replicates during quantification. Transcription levels of target genes were normalized to those of *Fkbp1a* and *Gnb1* and calculated using the $2^{-\Delta\Delta Ct}$ method[91].

Viable Cell Density and Viability Measurement

Generally, viable cell density (VCD) and viability were monitored using the NucleoCounter NC-200 Cell Counter (ChemoMetec, Denmark). Particularly, necrosis assay was performed in Figure 3. 100 µL cell suspension was pelleted and resuspended in 100 µL of staining solution composed of CD CHO cell culture medium supplemented with 5 µg/mL propidium iodide (Carl Roth, Karlsruhe, Germany). Cells were protected from light and incubated for 25 min at RT. Fluorescence assisted cell analysis was performed using a MACSQuant® VYB benchtop flow cytometer (Miltenyi Biotec, Bergisch Gladbach, Germany) equipped with a violet (405 nm), blue (488 nm) and yellow (561 nm) excitation laser. For each sample, a total volume of 25 µL (approximately 20-30 k cells) was analyzed at a flow rate of 100 µL per min. Propidium iodide positives were considered as necrotic.

Statistical Analysis

Statistical analysis was performed using Microsoft Excel and GraphPad Prism. Unless otherwise stated, an unpaired two-tailed t-test was applied to determine statistical significance between independent sets of samples. The one-way analysis of variance (ANOVA) followed by a Dunnett's multiple comparison post-test was conducted for comparing more than two data sets and for many-to-one comparisons. To compare several means to each other, the Tukey's multiple comparisons test was applied. Significance was assumed for p-values less than 0.05.

Acknowledgements

The authors thank Nachon Charanyanonda Petersen and Karen Kathrine Brøndum for assistance with the FACS and Mads Valdemar Anderson for computational support. We acknowledge support from the Novo Nordisk Foundation (NNF10CC1016517 and NNF16OC0021638) and support to NEL from NIGMS (R35 GM119850). AC was funded by the Burroughs Wellcome Career Award for Medical Scientist. The authors declare no competing financial interests.

CHAPTER 3

CRISPR MEDIATED ACTIVATION

Manuscript III

CRISPR Activation for Transcriptional Reprogramming of Endogenous Genes in Chinese Hamster Ovary Cells

Karen Julie la Cour Karottki[1]*, Isaac Shamie[2]*, Kai Xiong[1], Hooman Hefzi[2], Anders Holmgaard Hansen[1], Songyan Li[1], Jae Seong Lee[1], Shangzhong Li[2], Samuel Roth[2,4], Sascha Duttke[4], Christopher Benner[4], Gyun Min Lee[3], Helene Faustrup Kildegaard[1], Nathan E. Lewis[2]

(1) The Novo Nordisk Foundation Center For Biosustainability, Technical University Of Denmark, Denmark

(2) The Novo Nordisk Foundation Center For Biosustainability At The University Of California, San Diego, USA

(3)Department Of Biological Sciences, Kaist, 291 Daehak-Ro, Yuseong-Gu, Daejeon 305-701, Republic Of Korea

(4)Department Medicine, University of California, San Diego, USA

*These authors contributed equally to this publication.

Key words: CRISPR, activation, CHO, transcription start site, *Mgat3*, *St6gal1*, *Aldh18a1*

Abstract

Chinese hamster ovary (CHO) cells are the preferred workhorse for the biopharmaceutical industry. In recent years CRISPR/Cas9 has proven a powerful tool for generating targeted gene perturbations in CHO cells. Here we expand the CRISPR engineering toolbox with CRISPR activation (CRISPRa) to increase transcription of endogenous genes. We designed gRNAs to target *Mgat3, Aldh18a1,* and *St6gal1*. We successfully increased transcription of all genes and verified the effect of *Mgat3* and *St6gal1* overexpression on a functional level by detecting that the appropriate glycan structures were produced. This work demonstrates the applicability of CRISPRa for targeted alterations of CHO cells toward desired phenotypic traits.

Mammalian genomes encode over 20,000 genes but only a fraction of these are expressed in any given tissue[92]. This is true for Chinese hamster ovary (CHO) cells as well; across ~1600 transcriptomic datasets nearly half of all genes are never expressed (unpublished data). Within production of recombinant therapeutic proteins, CHO cells are the most widely used expression host[8]. Despite the ubiquity of their use, the recombinant proteins being produced in CHO cells are not perfect matches to their human equivalent, largely due to differences in glycosylation. For example, the bisecting N-Acetylglucosamine attached by Mgat3 (Mannosyl (Beta-1,4-)-Glycoprotein Beta-1,4-N- Acetylglucosaminyltransferase) and the alpha 2,6 sialic acid attached by St6gal1 (St6 beta galactoside alpha-2,6-sialyltransferase 1) are common on many human glycoproteins[73,93], but are not found on proteins expressed in CHO cells. Recent advances in gene editing tools and genomic understanding have eased our ability to manipulate CHO cells in order to overcome some of these shortcomings.

The clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated protein 9 (Cas9) system has gained foothold as an easy and cost-effective tool for manipulating the genome and has demonstrated to work efficiently for generating CHO knockout cell lines[72,73,94,95]. The CRISPR/Cas9 system does not limit itself to gene disruption, however; by mutating the nuclease domains of the native Cas9 endonuclease to a catalytically inactive Cas9--referred to as dead Cas9 (dCas9)--and fusing it with various domains, the system has been repurposed from inflicting DNA cleavage to delivering effector molecules to sites of interest. Fusing dCas9 to transcription activators, a tool known as CRISPR activation (CRISPRa), and guiding it to a region upstream of the transcription start site (TSS) has been shown to be a powerful tool to increase transcription of endogenous genes across different eukaryotic organisms[41,47,96,97]. A comparative study across different organisms revealed that fusing dCas9 to the activators SAM, Suntag, and VPR increased expression most effectively[47]. Further work on accurately predicting the activity of gRNAs established that the position of the gRNA relative to the TSS was one of the most important predictors of transcriptional activation[49].

This study is the first, to our knowledge, to apply CRISPRa to increase transcription of genes in CHO cells. We targeted *Mgat3* and *St6gal1* due to their importance in generating

human-like glycoforms as well as *Aldh18a1*, a gene responsible for the hallmark proline auxotrophy of CHO cell lines[98] using gRNAs designed against computationally predicted TSSs as well as an alternative TSS dataset built off of experimental data.

Results

CRISPRa Targeting NCBI CHO-K1 Transcription Start Sites Increases Gene Expression

One of the central factors for successful activation using CRISPRa is the gRNA design. For well-known organisms, like human or mouse, online predictive tools exist[99,100]; however, there is no such tool for CHO cells because of limited genomic resources. We defined the TSS to be the first base pair of the target gene using the CHO-K1 annotation available on NCBI (from here on referred to as NCBI TSS). We designed the gRNAs--based on studies in other mammalian cells--at distances to the TSS that lead to a predicted high CRISPRa activity (25 to 550 bp upstream)[49,63] (Figure 1A). We chose to design gRNAs against three industrially relevant genes (*Mgat3* and *St6gal1* impacting glycosylation, *Aldh18a1* being responsible for hallmark proline auxotrophy of CHO cells) that had low expression in CHO-S transcriptomic data[24] (Supplemental Table S1). For our activator construct, we used VPR since it is the simplest construct that showed high levels of activation in previous studies[47] (Figure 1B). Finally, *Mgat3* and *St6gal1* activation was carried out in an engineered cell line designed to produce primarily bianntenary, nonsialylated, nonfucosylated glycans so as to make detection of altered glycans easier (see Methods).

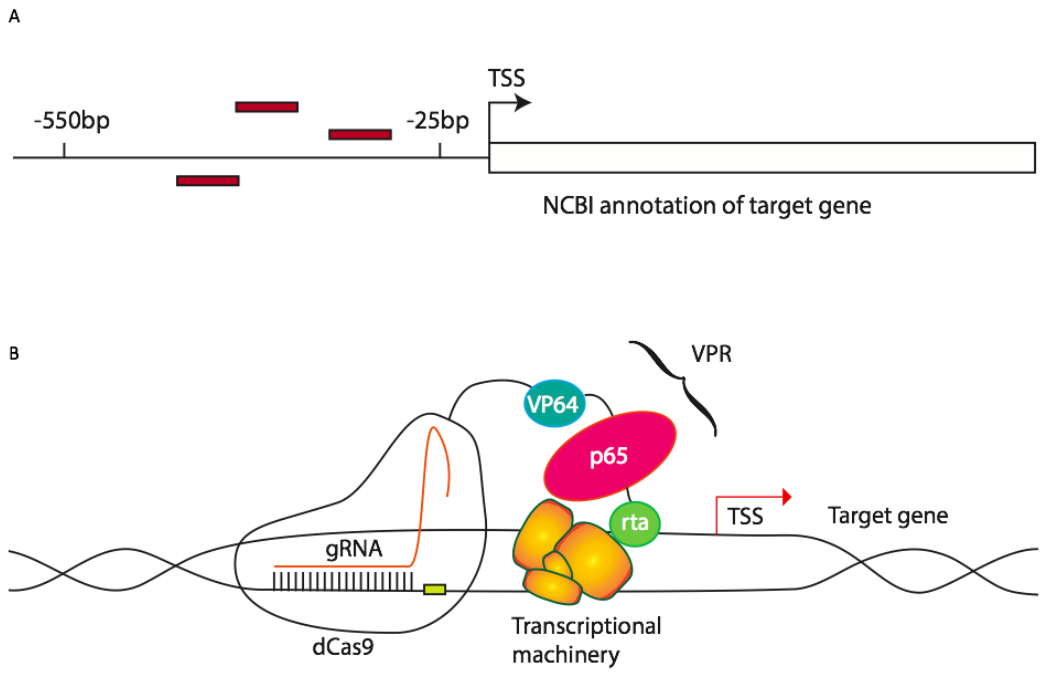


Figure 1 Components for CRISPRa

A) gRNA design using the CHO-K1 annotation available on NCBI. The TSS was set to be the starting base pair of the annotated target genes. Three gRNAs (red rectangles) were designed between 25 bp and 550 bp upstream of the TSS. B) A plasmid encoding VPR-dCas9 was transfected along with either one of the gRNA plasmids or a mix of all of the gRNA plasmids. The gRNAs guide the VPR-dCas9 to a region upstream of the TSS where the VPR increases expression of the target gene.

Using gRNAs designed against the NCBI TSS we were able to increase the transcription of all target genes (Figure 2). RNA expression for *Mgat3* and *St6gal1* was increased noticeably--up to 1500 times--whereas for *Aldh18a1* the results were modest (a maximum of ~4-fold activation was observed using gRNA 1). We observed clear differences between the individual gRNAs and expression level. For example, gRNA 1 against *Mgat3* and gRNA 2 against *St6gal1* failed to increase RNA expression. For *Mgat3* and *St6gal1,* transfecting with a mix of the 3 gRNAs showed the most potent activation, corroborating findings in the literature[47,101]. Given the varied results we decided to investigate whether we could improve activation by designing gRNAs against experimentally found TSSs.

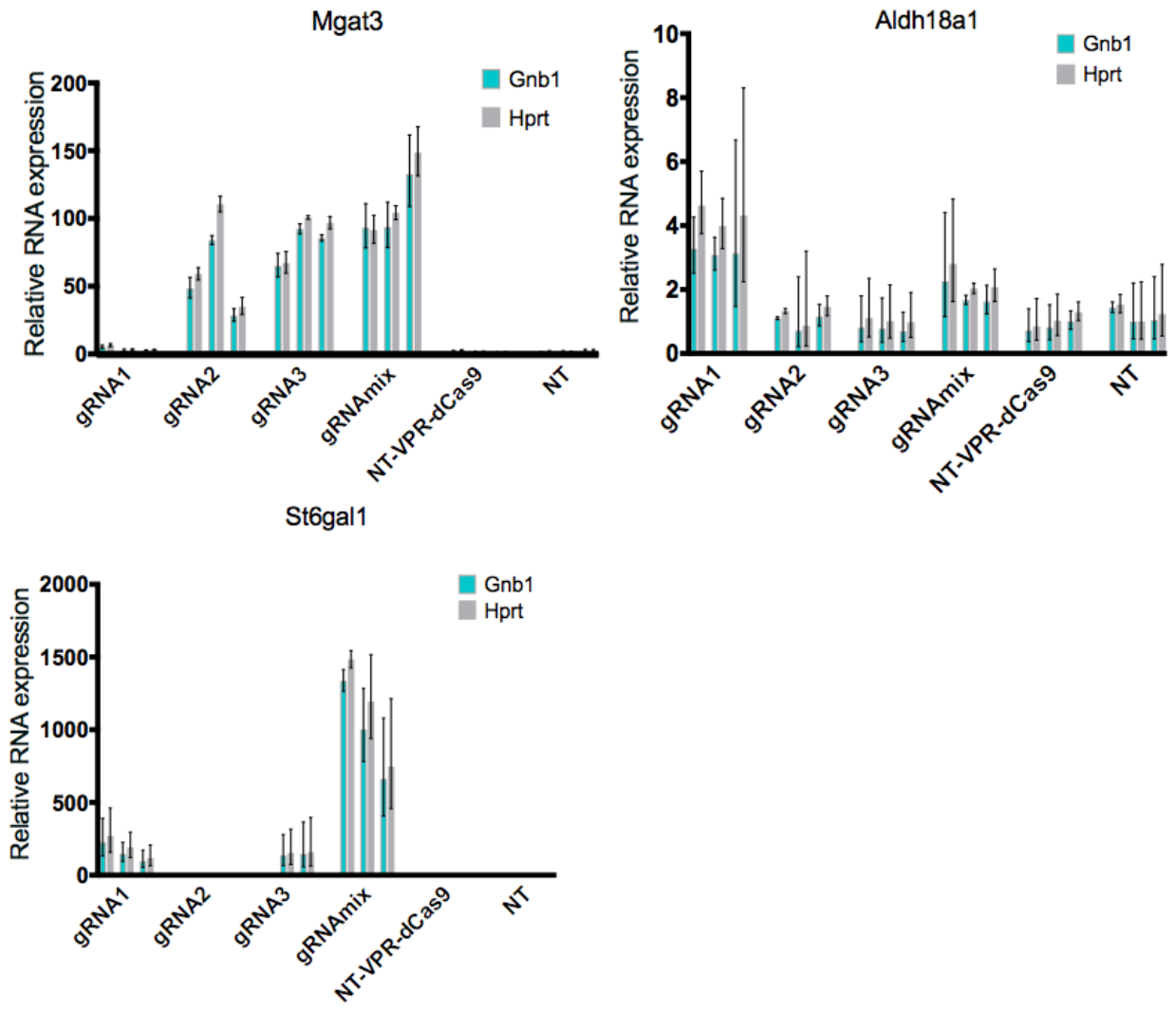


Figure 2 CRISPR induced activation of endogenous genes in CHO cells using gRNAs designed against the NCBI annotation

Activation of *Mgat3, Aldh18a1* and *St6gal1* via transient expression of gRNA and VPR-dCas9 was measured by qRT-PCR. CHO cells were transfected with gRNAs 1-3 individually or a mix of the 3 gRNAs in three biological replicates. mRNA was harvested 48h after transfection. For controls, cells were transfected with non-targeting gRNA and VPR-dCas9 (NT-VPR-dCas9) or non-targeting gRNA (NT). Relative RNA expression is shown as mean of the technical triplicates normalized to housekeeping genes *Gnb1* (turquoise bars) or *Hprt* (grey bars). Relative expression is calculated with respect to the NT biological replicate with the largest ΔCt.

CRISPRa Targeting Experimentally Identified Transcription Start Sites is Mixed in Increasing Gene Expression.

With the aim of improving the gRNA design, we first experimentally determined the TSSs in a mix of hamster tissues (see Supplementary Methods). Furthermore, as some regions upstream of genes contain gaps in the NCBI CHO-K1 assembly (example shown in Supplemental Figure S2), a more complete genome assembly for CHO would be ideal for gRNA design. While the newly published Chinese hamster genome is the most complete Chinese hamster genome assembly published to date, closing many of the sequence gaps of the previous assembly[28] (example shown in Supplemental Figure S2), CHO cells have accumulated a considerable number of mutations in the course of their use[17]. By correcting the Chinese hamster genome for genetic variants present in CHO-S (see Methods), we created a genome sequence representing the most accurate picture of the CHO-S genome available, from here on referred to as CHOoptPICR. We mapped the experimentally measured TSS sites back to the CHOoptPICR (see Methods) and carried out gRNA and experimental design in the same manner as with the NCBI TSSs (Figure 3).

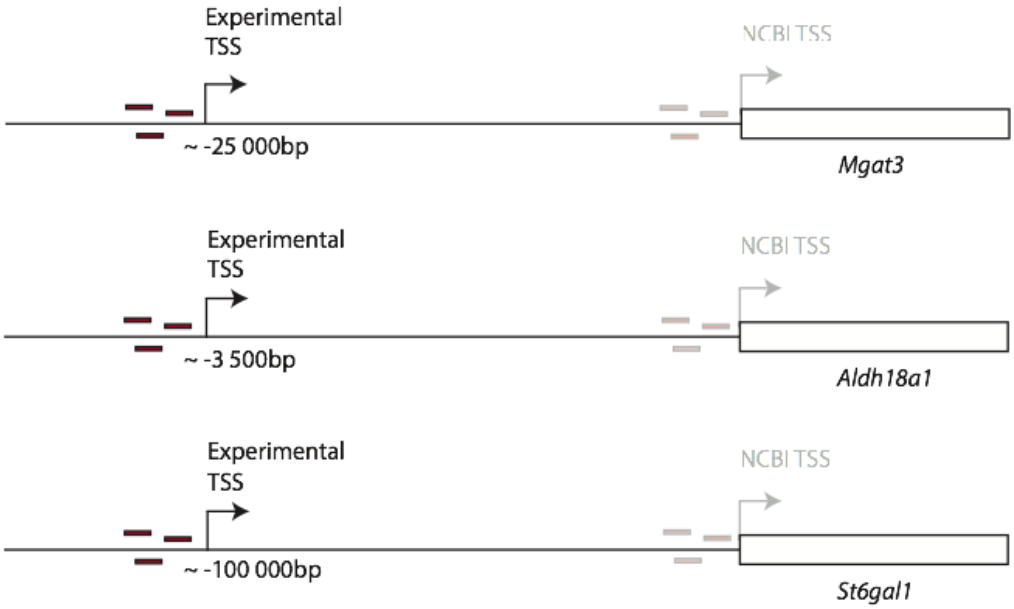


Figure 3 gRNA design against experimental TSSs

TSSs were defined based on experimental data and mapped to CHOoptPICR. gRNAs were designed in the region 25 bp to 550 bp upstream of the experimental TSS. The experimental TSSs for the different target genes were located at different distances upstream of the NCBI TSSs (faded); *Mgat3,* 25kb; *Aldh18a1,* 3.5kb; *St6gal1,* 100kb.

By using gRNAs against experimentally defined TSSs, we succeeded in increasing gene expression with some of the gRNAs (Figure 4). All three gRNAs targeting *Mgat3* showed equal or greater increase of expression compared to gRNAs using the NCBI TSSs. For *Aldh18a1* there were no noticeable changes in expression levels, largely similar to what was observed using the gRNAs designed against the NCBI TSSs (with the exception of gRNA 1). The gRNAs targeting *St6gal1* showed varied results, and did not achieve the increase in expression levels as seen with the gRNAs against the NCBI TSS annotation.

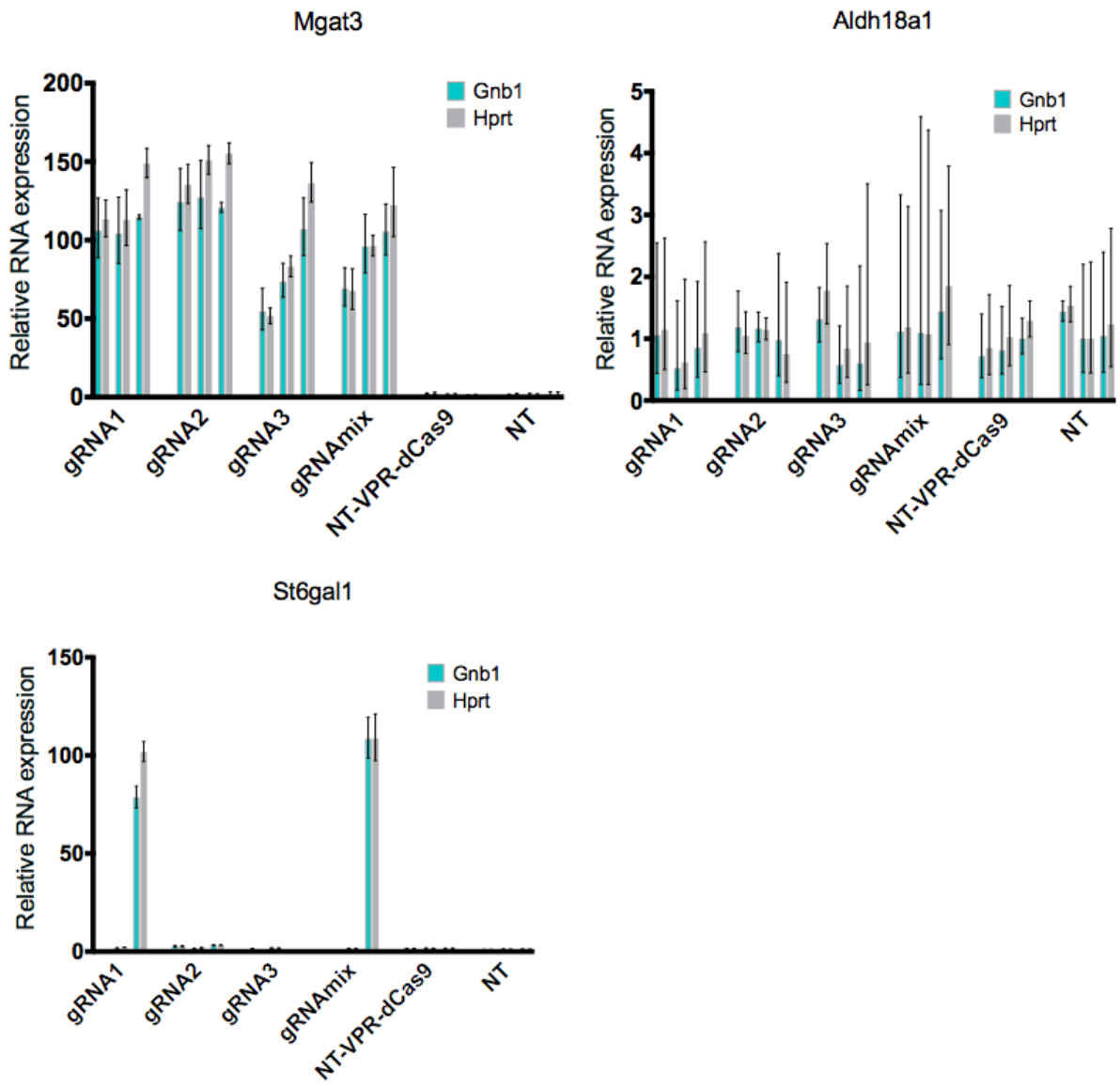


Figure 4 CRISPR induced activation of endogenous genes in CHO cells using gRNAs designed against experimentally discovered TSSs

Activation of *Mgat3, Aldh18a1* and *St6gal1* via transient expression of gRNA and VPR-dCas9 was measured by qRT-PCR. CHO cells were transfected with gRNAs 1-3 individually or a mix of the 3 in three biological replicates. The mRNA was harvested 48h after transfection. For controls, cells were transfected with non-targeting gRNA and VPR-dCas9 (NT-VPR-dCas9) or non-targeting gRNA (NT). Relative RNA expression is shown as mean of the technical triplicates normalized to housekeeping genes *Gnb1* (turquoise bars) or *Hprt* (grey bars). Relative expression is calculated with respect to the NT biological replicate with the largest ΔCt.

Activation of Mgat3 and St6gal1 is Observed on the Glycan Level

To confirm that the effect seen transcriptionally has a functional impact, the best performing gRNA for each target and TSS source--as well as the non-targeting (NT) gRNA--had the glycans on secreted proteins quantified (see Methods). In the NT gRNA samples no glycans indicating Mgat3 (bisecting GlcNAc) or St6gal1 (alpha-2,6 sialic acid) activity were detected. Glycans containing bisecting GlcNAcs were detected in both the mix of gRNAs designed against the NCBI TSS and for gRNA 2 designed against the CHOoptPICR TSS for *Mgat3*, albeit at low levels. A more drastic increase in alpha-2,6 sialylation was observed in the gRNA mix designed against the NCBI TSS for *St6gal1*, while no such glycans were observed in the gRNA mix against the CHOoptPICR TSS (Figure 5), in agreement with the inconsistent (seen only 1 of 3 biological replicates) activation observed with the latter.

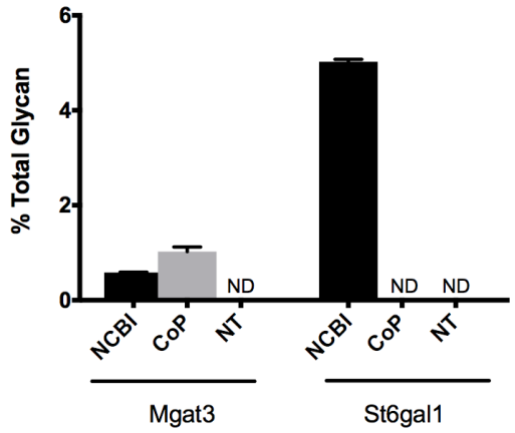


Figure 5 Glycan analysis of secreted proteins during CRISPR induced activation of *Mgat3* and *St6gal1*

Fraction of the total N-glycans on secreted proteins that contain bisecting GlcNAc (increased Mgat3 activity) or any number (mono- or bi-) or alpha-2,6 linked sialic acids (increased St6gal1 activity). Shown are the results from the gRNA showing the largest activation transcriptionally from the NCBI defined TSS (NCBI, see Figure 2) and the experimental TSS in CHOoptPICR (CoP, see Figure 3), as well as the non-targeting gRNA treated samples (NT). ND: not detected.

Concluding Remarks

In this study we draw on CRISPRa studies successfully carried out in other mammalian organisms to expand the CRISPR toolbox for CHO cell line engineering. We succeeded in increasing transcription of our target genes by transiently transfecting dCas9 fused with activator domain VPR and gRNAs targeting the region directly upstream of the TSS. Most striking was the activation of *Mgat3* transcription, which succeeded with two different TSSs 25000 bp apart. We confirmed the activation from both TSSs on a functional level by observing the associated glycan structure. While we succeeded in increasing transcription of *St6gal1* up to 1500-fold using a TSS predicted from the NCBI CHO-K1 annotation, results were mixed when targeting the experimentally determined TSS (~100000 bp upstream of the NCBI TSS). Activation--when observed--was lower, and the response was inconsistent between biological replicates for gRNAs displaying activation. It is possible that the called TSS is actually an enhancer RNA or other non-coding RNA rather than the true TSS. We confirmed our observations on a functional level as we only observed the associated glycan structure using gRNAs against the NCBI CHO-K1 annotation. *Aldh18a1* showed a low level of activation (~4-fold) for 1 gRNA designed against the NCBI TSS and no measurable activation for all other gRNAs (against both TSSs). With only 3500 bp between the tested TSSs, it is unclear why this gene was difficult to activate; many factors, including chromatin state (for dCas9 and gRNA accessibility) could play a role. In summary, we showed that CRISPRa can successfully be applied to increase gene expression in CHO cells, however there is still more research to be done on how we can obtain a more consistent activation pattern across all genes and TSSs in CHO cells.

Construction of a CHO-Optimized Genome Assembly

Variants were called using GATK[102] using Illumina DNA-sequencing reads from CHO-S against a draft version of the latest *C. griseus* genome (PICR)[28]. The resulting VCF file was used to correct CHO-S specific mutations in the PICR assembly using the FastaAlternateReferenceMaker function of GATK, generating the CHOoptPICR assembly used for gRNA design.

gRNA Design

For gRNAs designed against NCBI CHO-K1 annotation, a 1000 bp region upstream of the 1st basepair of the gene was obtained from the CHO-K1 assembly on NCBI. This region was analyzed using an inhouse gRNA prediction tool to rank all possible gRNA sequences in the region based on off-targets. For gRNAs designed against experimentally found TSSs (see Supplementary Methods) the 1000 bp region upstream of the most highly ranked TSS in a draft version of the PICR assembly[28] was blasted against CHOoptPICR to find the corresponding region when correcting for CHO specific sequence variants. These regions were then analyzed using the same prediction tool. For both TSS sources, the gRNAs were further filtered for a proximity of 550 bp-25 bp upstream of the TSS. Target sequences and gRNA oligos are listed in Supplemental Table S2 and Table S3.

Vector Construction

gRNA vectors were constructed using Uracil-Specific Excision Reagent (USER) friendly cloning as previously described[71]. The CHO codon-optimized dCas9 cassette was generated from a synthesized CHO codon-optimized wild-type Cas9 (GeneArt, Thermo Fisher Scientific) by mutagenesis of Cas9 to introduce the D10A and H840A mutations using Q5 Site-Directed mutagenesis kit (New England Biolabs). The CHO codon-optimized dCas9 was then fused to a VPR domain cloned from a AAV_NLS-SaCas9-NLS-VPR vetor (Addgene plasmid #68496). The construct will from here on be referred to as VPR_dCas9. The plasmid

construct and sequence are listed in Supplemental Figure S1. All plasmids were purified using NucleoBond Xtra Midi EF (Macherey-Nagel) according to manufacturer's protocol and verified by Sanger sequencing.

Cell Culture and Transfection

CHO-S cells (Thermo Fisher Scientific) and a CHO-S cell with knockouts in *Mgat4a,4b* and *5*, *St3gal3,4* and *6*, *B3gnt2*, *Sppl3,* and *Fut8* (from here on referred to as CHO-C1) were maintained in CD-CHO medium supplemented with 8mM L-Glutamine (Thermo Fisher Scientific) and 1 µl/mL Anti-Clumping (AC) reagent (Life Technologies) unless otherwise specified. The cells were cultivated in 125mL Erlenmeyer shake flasks (Corning Inc.) in a humidified incubator at 37˚C, 5% CO_2 at 120 RPM and passaged every 2-3 days. Viable cell density (VCD) and viability were monitored using the Nucleocounter NC-200 Cell Counter (ChemoMetec). One day prior to transfection 0.6 x 10^6 cells/mL cells were spun down at 200 x g and resuspended in 6-well plates (BD Biosciences) in 3 mL culture medium without AC per well. Cells were transfected using a total of 3.75 µg DNA and 3.75 µl FreeStyle™ Max transfection reagent (Thermo Fisher Scientific) together with OptiPRO SFM medium (Life Technologies) according to manufacturer's protocol. The ratio of DNA used from respective plasmids is listed in Supplemental Table S5. Expression plasmids for gRNAs targeting *Aldh18a1* were transfected into CHO-S cells and for gRNAs against *Mgat3* and *St6gal1* were transfected into CHO-C1 cells. In addition, non-targeting gRNA with and without dCas9 were transfected for each cell line as controls. All transfections were done in biological triplicates.

RNA Isolation, cDNA Generation and Quantitative Reverse-Transcription PCR

qRT-PCR was carried out to determine the relative expression of target genes after CRISPRa using pre-designed TaqMan Gene Expression Assays (St6gal1: Cg04423241_g1, Aldh18a1: Cg04498748_m1, and Hprt: Cg04448435_m1) (Thermo Fisher Scientific) or custom assays (Mgat3 and Gnb1--see Supplemental Table S5). Two days after transfection, 500 µl of cells were spun at 200 x g, the supernatant was removed, and the pellet was resuspended in 300 µl TRIzol™ Reagent (Thermo Fisher Scientific). RNA was isolated using

96 well Direct-zol RNA kits (Zymo Research) following manufacturer's protocol, followed by quality and quantity measurement on the NanoDrop spectrophotometer 2000 (Thermo Fisher Scientific). Isolated RNA was reverse transcribed into cDNA using using the qScript Flex cDNA synthesis kit (QuantaBio), followed by qRT-PCR with *Hprt* and *Gnb1* as housekeeping genes following manufacturer's protocol using the QuantStudio 5 instrument (Applied Biosystems). Amplification was performed under the following conditions: 50°C for 2 min, 95°C for 10 min; 45x: 95°C for 15 s, 60°C for 1 min. Three technical replicates were performed for each sample. Transcript levels were normalized to *Hprt* or *Gnb1* and relative expression levels were calculated using using the $2^{-\Delta\Delta Ct}$ method[91]. For *St6gal1* no amplification was detected in the NT samples, so the Ct value was arbitrarily set to 46 (1 cycle after the run was terminated).

Glycan Analysis

After centrifugation to remove the cells and cell debris, the supernatant (~3 mL) was up-concentrated using Amicon® Ultra-4 Centrifugal Filter Units following manufacturer' instructions. Secretome proteins were fluorescently labeled with GlycoWorks RapiFluor-MS N-Glycan Kit (Waters, Milford, MA) according to the manufacturer's protocol. N-linked glycan analysis was performed by LC-MS system using a Thermo Ultimate 3000 HPLC with the fluorescence detector coupled on-line to a Thermo Velos Pro Iontrap MS and was run in positive mode. Separation gradient was 30% to 43% buffer. The amount of N-glycan was measured by integrating the areas under the normalized fluorescence spectrum peaks with Thermo Xcalibur software (Thermo Fisher Scientific, Waltham, MA) giving the normalized, relative amount of the glycans.

Acknowledgements

The authors thank Helle Munck Petersen and Sanne Schoffelen for help with glyco labeling. The authors acknowledge support from the Novo Nordisk Foundation (NNF10CC1016517 and NNF16OC0021638)

CHAPTER 4

LARGE SCALE POOLED CRISPR SCREENING

Manuscript IV

Establishing a Pooled Metabolic CRISPR-Cas9 Knockout Screen in Suspension Chinese Hamster Ovary Cells

Karen Julie la Cour Karottki[1], Hooman Hefzi[2], Songyuan Li[1], Lasse Ebdrup Pedersen[1], Philipp Spahn[2], Alex Thomas[2], Jae Seong Lee[1], Gyun Min Lee[3], Nathan E. Lewis[2], Helene Faustrup Kildegaard[1]

(1) The Novo Nordisk Foundation Center For Biosustainability, Technical University Of Denmark, Denmark

(2) The Novo Nordisk Foundation Center For Biosustainability At The University Of California, San Diego, USA

(3)Department Of Biological Sciences, Kaist, 291 Daehak-Ro, Yuseong-Gu, Daejeon 305-701, Republic Of Korea

Keywords: CHO, CRISPR pooled screen, glutamine, metabolism

Abstract

Chinese hamster ovary (CHO) cells are the most common mammalian cell line used for producing bio-therapeutic proteins and serve as the expression system of choice for the best-selling biologics. Although these cells are widely used, the genetic bases underlying desirable phenotypes remain difficult to elucidate. CRISPR/Cas9-mediated genome engineering has proven effective in CHO cells in smaller formats, yet rational target identification on a high throughput level remains a bottleneck. CRISPR-Cas9 pooled screens represent a powerful tool for discovery of novel genotype to phenotype relations. Here we describe the first published CHO-specific CRISPR/Cas9 knockout screening platform in CHO cells, enabling high throughput target discovery under industrially relevant selection pressures. We have designed a library comprising ~16,000 gRNAs against ~2500 metabolic targets using the CHO-K1 genome and the genome scale model of CHO cell metabolism. The library was used to generate a pool of cells each expressing a single gRNA. As a proof of concept, we subjected the pool of cells to glutamine selection to identify genes that influence cell growth under glutamine deprivation. The screen successfully identified expected targets and revealed novel genes that enhanced growth in media lacking glutamine.

Chinese hamster ovary (CHO) cells are the most commonly used mammalian cells for biotherapeutic protein production and serve as the expression system of choice for the best-selling biologics[8]. Consequently, improving product quality and decreasing manufacturing costs in CHO cells is of great interest to the biopharmaceutical industry. Since their first use in the late 1980s, final product titer from CHO cells has improved 50-fold, largely through bioprocess optimization[103]. Although effective, these empirical approaches are highly variable and demand extensive labor, time and resources. However, the general lack of understanding of the underlying molecular processes and unique CHO cell idiosyncrasies have slowed down the development of rationally designed engineering methods. While the recent release of CHO cell genome sequences[17,25,26] and improved systems biology approaches[24,104] have laid the groundwork for a new era of targeted CHO cell line development, the question of the best way to discover and engineer targets remains open.

Although it has been possible to knock out genes in CHO cells via zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs)[105,106], their complexity, and difficulty of use and cost has made anything beyond single gene deletions nearly impossible. With more than 20,000 genes in the CHO genome—many of which have an unknown impact on the cell—an efficient, high-throughput method to characterize genetic perturbations is necessary to explore the phenotypic space. Although RNA interference (RNAi) screening initially showed promise via high-throughput investigation of gene knockdowns[107], the inability to achieve full knockout, a significant amount of off-target effects and inconsistent results have limited their use[108]. The recent development of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) technology has drastically transformed the outlook of CHO cell line engineering, allowing for easy, robust, and inexpensive manipulation. Briefly, a 19-20 nucleotide RNA molecule (the guideRNA (gRNA)) guides a Cas9 nuclease to the complementary genomic sequence and induces a double strand break, which is often repaired imperfectly (e.g., with frameshift causing insertions or deletions)[33,109,110]. Indeed, CRISPR/Cas9 mediated genome engineering of CHO cells has already been established in smaller formats for targeted cell design[71] and permits easy multiplex genome editing[72]. The adaptation of CRISPR to large-

scale pooled screening presents a novel approach for forward genetic screens and functions much like the RNAi screens, except by employing the CRISPR/Cas9 system rather than RNAi, avoiding many of the pitfalls present in RNAi screening[111]. This method has been established in several cell lines and organisms, mainly mice and human, increasing the robustness for the next generation of forward genetic screening methods[60,111–114].

With the intent of generating a platform for gaining insight into CHO cell metabolism, we present the first large-scale CRISPR-Cas9 knockout screen performed in CHO cells. We generated a CHO-specific gRNA library targeting genes involved in CHO cell metabolism and deployed the CRISPR-Cas9 knockout screening method against an industrially relevant selection pressure, glutamine deprivation. The results were confirmed by generating a knockout cell line of one of the top targets of the screen and characterizing its response to the presence and absence of glutamine.

Results

Establishing a CRISPR Knockout Library in CHO Cells

We first generated CHO-S cell lines constitutively expressing GFP-Cas9 (CHO-S^{Cas9}) via puromycin selection followed by single cell sorting and expansion for use with gRNA library transduction. The functionality of Cas9 in the clonal cell lines was assessed by transfecting CHO-S^{Cas9} with gRNAs for *Mgat1* and validating the cleavage efficiency by indel analysis of the target region (Supplementary Table S1). As a second test, Cas9-linked GFP expression was evaluated by analyzing fluorescence of the clonal cell lines (Supplementary Figure S1). We picked a clone positive for GFP expression and efficient Cas9 cleavage for further use. To build the CRISPR knockout library, we designed a large CHO-specific gRNA library containing multiple gRNAs against the metabolic genes of CHO. Genes in the screen included genes from the genome scale metabolic model of CHO[24], genes with metabolism-associated GO terms and transcription factors associated with the aforementioned metabolic genes. This resulted in a final library consisting of 15,654 gRNAs against 2,599 genes (1,765 genes from the model, 782 from GO terms and 52 transcription factors). CHO-S^{Cas9} cells were

transduced with the gRNA library at low MOI (Supplementary Methods) to ensure only a single gRNA integration event per cell, generating a CHO CRISPR knockout library for use in pooled screening (overview figure in Supplemental Figure S2).

Glutamine Screening

Glutamine is a key metabolite for cell function and thus an important media component for animal cell culture[115]. However, glutamine catabolism necessarily produces ammonia, a toxic byproduct that negatively impacts cell growth, production, and product quality[116–119]. Therefore, it is of interest to identify engineering strategies that permit improved cell behavior in glutamine free conditions. We thus screened the CHO CRISPR knockout library cells for growth in media with and without glutamine for fourteen days. The cells were passaged every three days (growth profile in Supplementary Figure S3) and 30 x 10^6 cells were collected at the beginning and the end of the screen for analysis.

The gRNA Library is Well Represented Prior to Selection

In CRISPR screening it is important to verify that all gRNAs are present at the initiation of the screen to ensure that all possible gene knockouts are assessed during selection. We therefore sampled the cells just prior to the initiation of glutamine deprivation (T0) and sequenced the gRNAs present in the starting cell pool. In all samples, median normalized gRNA sequencing depth was greater than 40 counts per million (CPM) (Figure 1A). Furthermore, we quantified the number of gRNAs/gene that were present at >10 CPM to assess whether most genes had multiple targeting gRNAs available. More than 75% of genes were represented by more than 80% of their respective gRNAs (Figure 1B). Thus, the majority of the library was well represented before the CRISPR knockout library was subjected to screening.

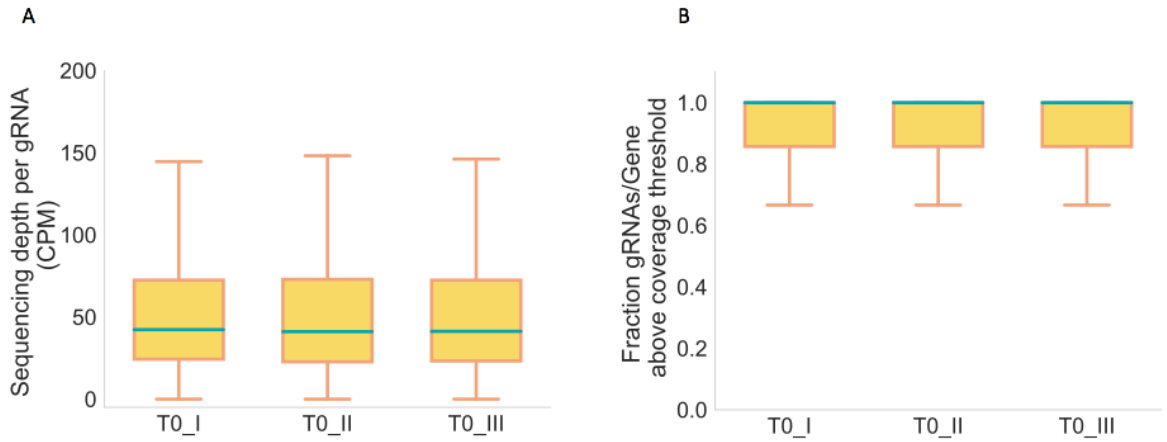


Figure 1 Screen verification

A) Read depth. Read count per gRNA in three replicate experiments (I, II, III) at time point 0 (T0) normalized to total read count in counts per million (CPM). Outliers are excluded. The median normalized gRNA sequencing depth was greater than 40 CPM for all samples B) Genes with a good global gRNA coverage. Fraction of gRNA/genes covered by a CPM of 10 or more. More than 75% of genes are represented by more than 80% of their respective gRNAs

Glutamine Screening Revealed Expected and Novel Targets

To see which gRNAs had an effect on CHO cell growth in glutamine free media, we analyzed the enrichment and depletion of gRNAs between endpoint samples grown for fourteen days in media with and without glutamine. As expected, the absence of glutamine does not display a very strong selection pressure (Supplementary Figure S4), consistent with the ability of CHO cells to express low levels of endogenous glutamine synthetase and thus survive via endogenous glutamine biosynthesis albeit at lower growth rates. Comparison of results for each replicate revealed conserved gene targets with beneficial or detrimental effects on growth in glutamine depleted media (Figure 2B). Comfortingly, growing the CRISPR knockout library cells in media without glutamine resulted in significant depletion of gRNAs against *Glul* (glutamine synthetase), consistent with its role as the enzyme responsible for *de novo* glutamine synthesis. Similarly, significant enrichment of *Gls* (glutaminase) gRNAs was observed, protecting the intracellular glutamine pool from undesirable catabolism. We continued to investigate the candidate showing the strongest and most consistent gRNA enrichment, Target X.

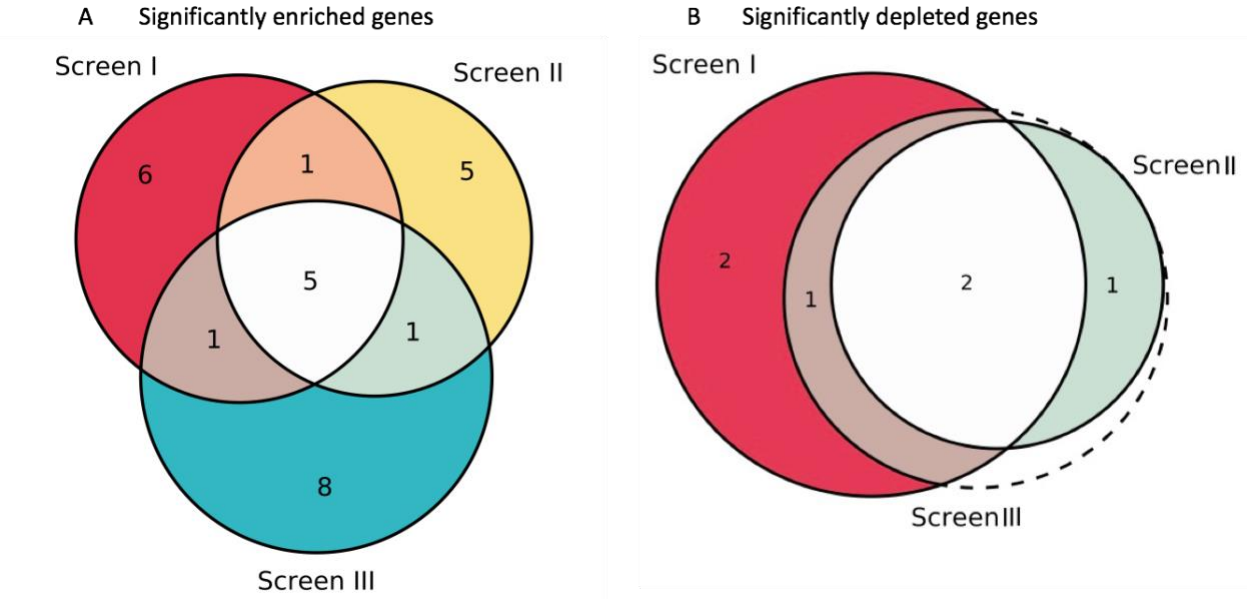


Figure 2 Significantly enriched and depleted genes in three replicates of glutamine screening

A) Venn diagram of significantly enriched genes. Genes were ranked using the aRRA gene ranking method. Five significantly enriched genes overlap between the three screen replicates (Screen I-III) shown in white. B) Venn diagram of significantly depleted genes. Genes were ranked using the aRRA gene ranking method. Five significantly depleted genes overlap between the three replicates (Screen I-III) shown in white.

Disruption of Target X is Conditionally Beneficial Dependent on Glutamine Presence

To validate the screen, we generated clonal Target X knockout cell lines using CRISPR/Cas9 and assessed their growth in media with and without glutamine. In accordance with the screening results, Target X knockout cells show improved growth in glutamine free media (Figure 3A) but also depressed growth in glutamine-containing media compared to control cells (Figure 3B).

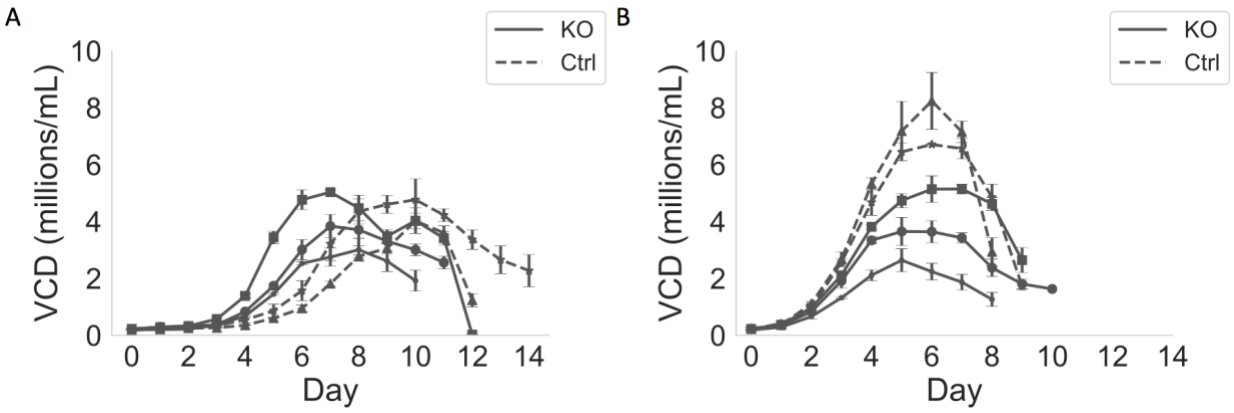


Figure 3 Screen Target X growth curves

Growth curves for three Target X knockout and two control cell lines grown in three replicates in media without glutamine (A) and supplemented with glutamine (B). Viable cell density (VCD) was measured every day over a period of 14 days.

Interestingly, nutrient utilization along the glutamine/glutamate axis is altered dependent on genotype and glutamine presence as well. In media containing glutamine, Target X knockout cells tend to accumulate extracellular glutamate (Figure 4A) while also consuming less glutamine (Figure 4B) before ceasing growth. In media lacking glutamine, however, this increase in glutamate is not observed (Figure 4A). It is possible that the Target X knockout in some way mitigates 'glutamine addiction'[120] via rewiring of metabolic flux that proves advantageous in glutamine depleted conditions (perhaps by leading to retention of glutamate for glutamine biosynthesis) but comparatively detrimental in glutamine replete conditions by impeding glutamine utilization. To explore the relationship between Target X and glutamine metabolism, we further analyzed knockout and control cell lines and compared their transcriptomic profile when grown in media with and without glutamine.

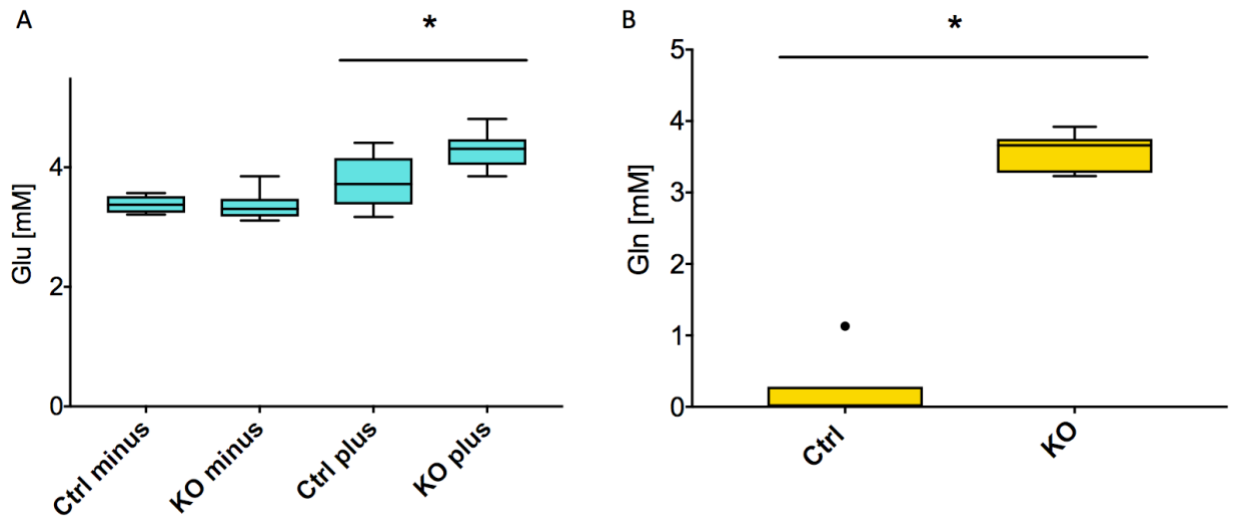


Figure 4 Impact of knockout and media condition on glutamate and glutamine metabolism

The maximum glutamate (A) and minimum glutamine (B) levels observed during the growth phase of culture (from seeding until the maximum VCD is reached) for control (Ctrl) and knockout (KO) clones in medium containing (plus) or lacking (minus) glutamine. Outliers are indicated with •. * indicates a statistically significant difference (α=0.05) as calculated by a two-tailed Welch's t-test with the Dunn–Šidák correction for multiple comparisons.

Transcriptomic Analysis Implicates Attenuation of the Interferon Response with Improved Growth in Glutamine Free Media

Gene set enrichment analysis of a comparison between the control cells in mid-exponential in media containing or lacking glutamine showed significant enrichment in numerous interferon associated gene sets (data not shown), which was surprising given their canonical association with viral infection. However, as one of the interferon stimulated responses leads to increased RNAse activity[121], this could be a reason for retarded growth in media lacking glutamine. Interestingly, when looking at how sets of interferon responsive genes behave in different comparisons of varying cell types (Target X knockout or control) and/or media condition (with or without glutamine), this signature of differential expression of interferon associated genes was primarily observed for control cells when grown in media lacking glutamine (Figure 5A). Specifically, this concerted change in expression of interferon

responsive genes does not appear when looking at Target X knockout vs. control cells in -glutamine conditions or Target X knockout cells in -glutamine vs +glutamine. For Target X knockout in -glutamine vs. control cells in +glutamine, while about half of the interferon responsive genes that were differentially expressed in the +glutamine vs -glutamine comparison for control cells were also differentially expressed, the amplitude of the differential expression was significantly (almost 2-fold) lower (Figure 5B)). It is also noteworthy that the increase in RNAse L expression is exclusive to the control cells when grown in media lacking glutamine. The directionality of this relationship is uncertain, specifically it is unclear if glutamine deprivation induces the interferon response leading to a decrease in growth or if glutamine deprivation leads to a decrease in growth that induces the interferon response due to a misinterpretation of biochemical cues. However, while the mechanisms are unclear, the effect is striking.

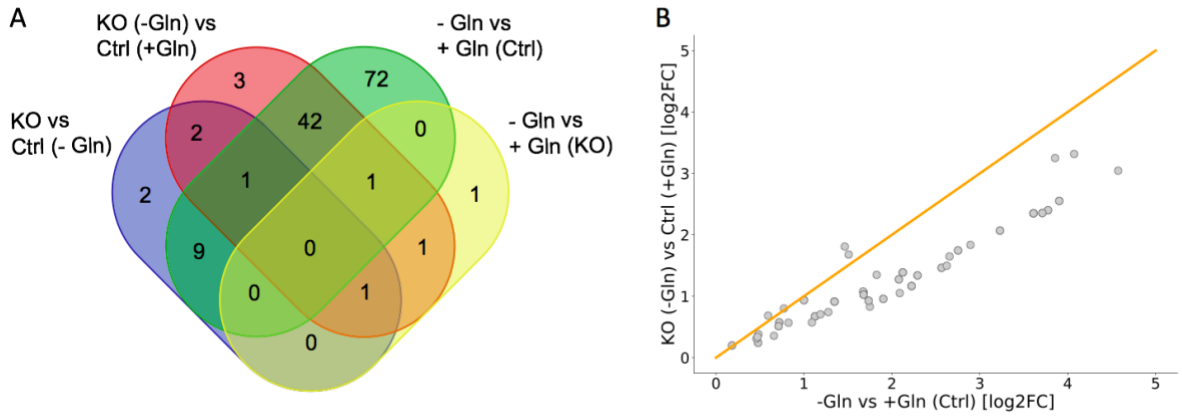


Figure 5 Differential expression of genes involved in the interferon response

A) Overlap between interferon response genes (see Methods) that are differentially expressed in various comparisons between clones of different genotype or growth conditions. B) Fold changes in differentially expressed interferon response genes (n=42, Figure 5A) that are in common between the -glutamine (-Gln) vs. +glutamine (+Gln) control (Ctrl) cell and -Gln knockout (KO) vs. +Gln control cell comparisons.

Discussion

As CHO cells are the primary workhorse for the production of biopharmaceuticals, significant time and effort has been invested towards producing optimal cell lines for growth, high protein titer, and good protein quality. Here, we contribute with a high-throughput approach to identify novel targets for CHO cell line engineering. The objective was bipartite; first to establish a CHO-specific metabolic CRISPR-Cas9 knockout screening platform in CHO cells and second to use this platform to explore the nature of CHO cell metabolism using an industrially relevant screening setup.

Glutamine is one of the major nutrients taken up by mammalian cells and plays an important role as an energy source in *in vitro* culture[115,122]. The fast consumption of glutamine results in accumulated ammonia in the medium, which inhibits cell growth, reduces productivity, and alters glycosylation patterns on heterologously expressed proteins[117–119]. While growth on glutamine-free media is possible, a significant decrease in growth rate is almost always observed[123]. It is therefore of interest to investigate genetic alterations that elicit a positive growth response to media lacking glutamine. We found several genes whose knockout resulted in a growth benefit in media without glutamine. Unsurprisingly, one of these genes was glutaminase, which codes for the primary glutamine-catabolizing enzyme. The remaining targets found were novel with respect to their protective role in glutamine depletion in CHO cells and their roles in a biological context are a topic for further investigation. We chose to follow up on Target X, a novel target that showed the most marked enrichment of gRNAs in cells grown under glutamine depleted conditions. Our results suggest that the interruption of Target X inhibits glutamate catabolism in some way. We observed depressed growth of Target X knockout cells in glutamine containing media alongside glutamate accumulation in the media and cessation of growth prior to glutamine depletion. As glutamate is a major source of TCA cycle intermediates[124], the secretion of glutamate is consistent with the observed reduced growth rate. Conversely, in glutamine free media, Target X knockout cells exhibited improved growth compared to the wild type cells without differences in glutamate secretion. We postulate that the inhibition of glutamate catabolism leads to retention of glutamate and increases its availability for conversion to glutamine, leading to better growth.

Unfortunately, an ironclad explanation for the mechanism underlying the glutamine dependent effect of Target X knockout was not readily apparent from transcriptomic analysis of the different cell lines/growth conditions. We did find a striking correlation between the apparent activation of an apparent interferon response-like signature for cells grown in media lacking glutamine that was largely attenuated for Target X knockout cells. However, more extensive studies are needed to draw definitive conclusions about the role Target X plays in glutamine metabolism. However, while the exact mechanism remains to be elucidated, the considerable increase in growth of Target X knockout cells in glutamine-free media is strongly reproducible and its inhibition could benefit applications wherein glutamine-depleted media is preferred.

High-throughput CRISPR-Cas9 screening presents a novel approach to conduct forward genetic engineering and can provide an abundance of knowledge in the study of genotype to phenotype relationships. Over recent years CRISPR-Cas9 screens have been applied to a variety of mammalian cell types to study biological function[60,61,112]. Since the publication of initial CRISPR-Cas9 screens, extensive reviews and comprehensive method articles have been published[59,64,125]. We show here that CRISPR screening techniques can be applied to the industrially relevant CHO cell line. This approach enables a wide array of studies in CHO cells by applying different screening conditions or exploiting the existing variations of the Cas protein such as catalytically inactive Cas9 coupled to transcriptional activators and repressors for activation or repression screens as has already shown potential in other mammalian cells[63,64,126–129]. With continuous advances in CRISPR screen design and comprehensive annotation of the CHO cell genome these types of screens will enable a new era of targeted engineering to improve CHO cell phenotypes.

Methods

Plasmid Design and Construction

The GFP_2A_Cas9 plasmid was constructed as previously described[72]. A Cas9 expression vector for generation of a Cas9 expressing CHO cell line (from here on be referred to as CHO-S^{Cas9}), was constructed by cloning the 2A peptide-linked Cas9 ORF from the GFP_2A_Cas9 expression vector[72] into a pcDNA™3.1(+) vector (Thermo Fisher Scientific) between the HindIII and BamHI sites. The construct will from here on be referred to as pcCas9. gRNA vectors were constructed using Uracil-Specific Excision Reagent (USER) friendly cloning as previously described[71]. Plasmids were purified using NucleoBond Xtra Midi EF (Macherey-Nagel) according to manufacturer's protocol. Target sequences and gRNA oligos are listed in Supplementary Table S2.

Cell Culture

CHO-S wild type cells from Life Technologies were cultivated in CD-CHO medium (Thermo Fisher Scientific) supplemented with 8 mM L-Glutamine and 2 µL/mL Anti-Clumping Agent (AC) (Thermo Fisher Scientific) in a humidified incubator at 37 °C, 5 % CO_2 at 120 RPM shake in sterile Corning® Erlenmeyer culture flasks (Sigma-Aldrich) unless otherwise stated. Viable cell density (VCD) was measured using the NucleoCounter® NC-200™ (Chemometec) utilizing fluorescent dyes acridine orange and 4',6-diamidino-2-phenylindole (DAPI) for the detection of total and dead cells. Cells were seeded at 0.3 x 10^6 cells/mL every three days or 0.5 x 10^6 cells every two days.

Transfection and Cell Line Generation

For all transfections, CHO-S wild type cells at a concentration of 1 x 10^6 cells/mL in a six well plate (BD Biosciences) in AC free media were transfected with a total of 3.75 µg DNA using FreeStyle™ MAX reagent together with OptiPRO SFM medium (Life Technologies) according to the manufacturer's instructions. For generation of CHO-S^{Cas9}, CHO-S wild type cells were transfected with pcCas9. Stable cell pools were generated by seeding transfected

cells at 0.2 x 10^6 cells/mL in 3 mL selection media containing 500 μg/mL G418 (Sigma-Aldrich) in CELLSTAR® 6 well Advanced TC plates (Greiner Bio-one) two days post transfection. Medium was changed every four days during selection. After two weeks of selection, cells were detached and adapted to grow in suspension. Single cell sorting of the stable cell pools was performed to generate clonal CHO-S^{Cas9} cell lines using a BD FACSJazz cell sorter (BD Biosciences) upon gating for GFP positive cell population as described previously[72]. The clonal cell lines were analysed by Celigo Cell Imaging Cytometer (Nexcelom Bioscience) based on the green fluorescence level using the mask (blue fluorescence representing individual cells stained with NucBlue™ Live ReadyProbes™ Reagent; Thermo Fisher Scientific) + target 1 (green fluorescence) application. For generating knockout cell lines of screen targets, CHO-S wild type cells were transfected with GFP_2A_Cas9 and appropriate gRNA expression vectors at a DNA ratio of 1:1 (w:w). Two days after transfection cells were single cell sorted using a FACSJazz (BD Bioscience), gating for GFP positive cell population as described previously[72]. Indels in targeted genes were verified by NGS as described previously[72]. Primers are listed in Supplementary Table S2. Three clones with a confirmed indel and two control clones without indels were and expanded to 30 mL media before they were frozen down at 1 x 10^7 cells per vial in spent CD-CHO medium with 5 % DMSO (Sigma-Aldrich).

Characterizing CHO-S^{Cas9} Functionality

To characterize Cas9 functionality we transfected clonal CHO-S^{Cas9} cells with a vector expressing gRNA against *Mgat1* and verified indel generation on a pool level by Next Generation Sequencing (NGS) as described previously[130] (using gRNA oligo primers MGAT1_gRNA_fwd and MGAT1_gRNA_rev and NGS primers MGAT1_miseq_fwd and MGAT1_miseq_rev listed in Supplementary Table S2). To analyze GFP expression, clonal cells were seeded in wells of a 96-well optical-bottom microplate (Greiner Bio-One) and identified GFP positive cells on the Celigo Cell Imaging Cytometer (Nexcelom Bioscience) using the green fluorescence channel. GFP negative gating was set on the basis of fluorescence emitted from CHO-S wild type cells.

Library Design and Construction

For design of the metabolic gRNA library, a list of metabolic genes was extracted from the CHO metabolic network reconstruction[24] along with a list of genes with metabolic GO terms in CHO and associated transcription factors. The gRNA templates were computationally designed using CRISPy (http://crispy.biosustain.dtu.dk/), resulting in a gRNA library with a minimum of 5 gRNAs per gene. The oligo library was synthesized by CustomArray. Full-length oligonucleotides were amplified by PCR using KAPA Hifi (Kapa Biosystems), size selected on a 2% agarose gel and purified with a QIAquick Gel Extraction Kit (Qiagen) as per manufacturer's protocol. The gRNA-LGP vector (Addgene #52963) was digested using BsmBI (New England BioLabs) (4 μg gRNA-LGP vector, 5 μL buffer 3.1, 5 μL 10 x BSA, 3 μL BsmBI and H_2O up to 50 μL were mixed and incubated at 55˚C for 3 hours). Subsequently, 2 μL of calf intestinal alkaline phosphatase (New England BioLabs) was added to the digested vector and the mix was incubated at 37˚C for 30 minutes before it was purified with a QIAquick PCR Purification Kit (Qiagen) as per manufacturer's protocol. To assemble the gRNAs into the vector a 20 μL Gibson ligation reaction (New England BioLabs) was carried out (25 ng linearized vector, 10 ng purified insert, 10 μL 2 x Gibson Assembly Master Mix (New England BioLabs) and up to 20 μL H_2O were mixed and incubated at 50˚C for 1 hour). The assembled vector was purified using QIAquick PCR purification (Qiagen) and transformed into chemically competent *E. coli* (Invitrogen). Transformed bacteria were plated onto LB-carbenicillin plates for overnight incubation at 37˚C, and plasmid DNA was purified using a HiSpeed Plasmid Maxi Kit (Qiagen).

Lentiviral Packaging

To produce the lentivirus, HEK293T cells were cultivated in DMEM supplemented with 10% Fetal Bovine Serum (FBS). One day prior to transfection, cells were seeded in a 15-cm tissue culture plate at a density suitable for reaching 70-80% confluency at time of transfection. Culture medium was replaced with prewarmed DMEM containing 10% FBS. 36 μL Lipofectamine 3000 (Life Technologies) was diluted in 1.2 mL OptiMEM (Life

Technologies) and in a separate tube 48 μL P3000 reagent, 12 μg pCMV (Addgene #12263), 3 μg pMD2.G (Addgene #12259) and 9 μg lentiviral vector were diluted in 1.2 mL OptiMEM. The solutions were incubated for 5 minutes at room temperature, mixed, incubated for another 30 minutes before they were added dropwise to the HEK293T cells. 48 hours and 72 hours after transfection the viral particles were concentrated using Centricon Plus-20 Centrifugal ultrafilters (100 kDa pore size), aliquoted and stored at -80˚C.

Puromycin Kill Curve

To determine the concentration of puromycin to be used to select the CHO library cells for gRNA insertion, a puromycin kill curve for CHO cells was determined. CHO-S wild type cells at a concentration of 1 x 10^6 cells/mL in media containing various amounts of puromycin (0, 0.25, 0.5, 1, 2, 3, 4, 5, 6, 7, 8 and 10 μg/mL). Cell viability and VCD was monitored over 7 days and based on halted growth and complete cell death of wild type cells 10 μg/mL was used for further experiments (Supplementary Figure S5).

Determining Multiplicity of Infection in CHO Cells

CHO-S wild type were seeded at 0.3 x 10^6 cells/mL in 1 mL media in 5 wells of a 12 well plate (BD Biosciences). Four wells were transduced with 4, 10, 20 and 40μL virus/well along with 8 μg/mL Polybrene (Sigma-Aldrich). Cells in the remaining well were left non-transduced as a negative control. After 24 hours, the cells were washed in PBS (Sigma-Aldrich) by centrifugation at 200 x g, resuspended in media and seeded in a new 12 well plate. After 24 hours cells were expanded to 3 mL media in wells of 6 well plates (BD Biosciences). Selection for cells containing the gRNA insert was initiated by adding 10 μg/mL puromycin (Thermo Fisher Scientific) to each well. Non-transduced control cells were monitored for complete cell death, consonant with finalised selection. After selection, the transduced cells were single cell sorted using a BD FACSJazz cell sorter (BD Biosciences) into Corning® 384 well plates (Sigma-Aldrich) containing 30 μL media without AC supplemented with 1% antibiotic-antimycotic (Gibco), and 1.5% HEPES buffer (Gibco). After 14 days of static incubation, viable cells were transferred to an MD96F Falcon™ plate (Thermo Fisher Scientific). Once

confluent, 50 uL cell suspension from each well was transferred to a MicroAmp Fast 96 well reaction plate (Thermo Fisher Scientific). The plate was centrifuged at 1000 x g for 10 minutes, then the supernatant was removed via rapid inversion. Genomic DNA (gDNA) was extracted by resuspending the pellet in 20 μL QuickExtract™ DNA Extraction Solution (Epicentre) and processing the plate in the thermocycler (65°C for 15 minutes followed by 95°C for 5 minutes). gRNA integration was verified by NGS using a modified version of the Illumina 16S Metagenomic Sequencing Library Preparation as described previously[72] (using primers lib_miseq_fwd, lib_miseq_rev, OMA1_fwd and OMA1_rev listed in Supplementary Table S2).

Transducing CHO-S^{Cas9} with Library Virus

CHO-S^{Cas9} cells were seeded at 0.3 x 10^6 cells/mL in 1 mL media in 26 wells of 12 well plates (BD Biosciences). In 25 of the wells, cells were transduced with 4 μL library virus/well along with 8 μg/mL Polybrene (Sigma-Aldrich) aiming for an MOI at 0.3-0.4 (Supplementary Results). Cells in the remaining well were left non-transduced as a negative control. After 24 hours, the cells were washed in PBS (Sigma-Aldrich) by centrifugation at 200 x g, resuspended in media and seeded in a new 12 well plate. After 24 hours, cells were expanded to 3 mL media in wells of 6 well plates (BD Biosciences). Selection for cells containing the gRNA insert was initiated by adding 10 μg/mL puromycin (Thermo Fisher Scientific) to each well (see puromycin kill curve in Supplementary Figure S5). Non-transduced control cells were monitored for complete cell death, consonant with finalised selection. The cells were washed and passed twice before they were expanded to attain enough cells to create a cell bank. Cells were frozen down at 1 x 10^7 cells per vial in spent CD-CHO medium with 5 % DMSO (Sigma-Aldrich) and will from here on be referred to as CHO-S^{Cas9} library cells.

Screening and DNA Extraction

CHO-S^{Cas9} library cells were thawed in 30 mL media and expanded to 60 mL before starting the screen. On day 0 (T0) 1.5 x 10^7 cells were spun down at 200 x g and resuspended in 60 mL appropriate screening media. The cells were grown for 14 days (passed to 0.25 x 10^6

cells/mL every third day). 30 x 10^6 cells were collected at T0 and on day 14 (T14). The pellets were stored at -80˚C until further use. gDNA extraction of all 30 x 10^6 cells was carried out using GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific) following the manufacturer's protocol. gDNA was eluted in 100 µL preheated elution buffer from the purification kit and incubated for 10 minutes before final centrifugation for maximum gDNA recovery.

Preparation for Next Generation Sequencing

50 µL PCR reactions with 3 µg input gDNA per reaction were run using Phusion® Hot Start II High-Fidelity DNA Polymerase (Thermo Fisher Scientific) (95°C for 4 min; 30 times: 98°C for 45s, 60°C for 30 s, 72°C for 1 min; 72°C for 7 min) using primers flanking the gRNA insert containing overhang sequenced compatible with Illumina Nextera XT indexing and 8 random nucleotides to increase the diversity of the sequences (LIB_8xN_NGS_FWD and LIB_8xN_NGS_REV listed in Supplemental Table S2). Double size selection was performed using Agencourt AMPure XP beads (Beckman Coulter) to exclude primer dimers and genomic DNA. The amplicons were indexed using Nextera XT Index Kit v2 (Illumina) sequence adapters using KAPA HiFi HotStart ReadyMix (KAPA Biosystems) (95°C for 3 min; 8 times: 95°C for 30s, 55°C for 30 s, 72°C for 30 s; 72°C for 5 min) and subjected to a second round of bead-based size exclusion. The resulting library was quantified with Qubit® using the dsDNA HS Assay Kit (Thermo Fisher Scientific) and the fragment size was determined using a 2100 Bioanalyzer Instrument (Agilent) before running the samples on a NextSeq 500 sequencer (Illumina).

Analysis

Raw FASTQ files were uploaded to PinAPL-PY (http://pinapl-py.ucsd.edu/[131]) along with a custom file containing the sequences for all gRNAs contained in the library. Top candidates for enriched and for depleted gRNAs were ranked by an adjusted robust rank aggregation (aRRA) method[132] and filtered for significance, compared between the replicates

and used for verification of the screen. The screen was analyzed using default parameters set by PinAPL-PY.

Batch Culture of Screen Target X

Screen target knockout cell lines were seeded at 0.3 x 10^6 cells/mL in 90 mL CD-CHO media with and without glutamine supplemented with 1µl/mL AC in 250 mL Corning® Erlenmeyer culture flasks (Sigma-Aldrich). Cell viability and density were measured every day for a maximum of fourteen days.

RNA Extraction for RNA-Seq

During batch culture 2 x 10^6 cells were harvested in mid-exponential phase. Cells were spun down at 200 x g, supernatant was discarded and the cell pellet was resuspended in 600 µL TRIzol™ Reagent (Thermo Fisher Scientific). RNA was isolated using Direct-zol™ RNA MiniPrep Plus (Zymo Research) following manufacturer's protocol. RNA concentration was measured with Qubit fluorometric analysis (Life Technologies) and RNA quality was determined with Agilent 2100 bioanalyzer and Fragment analyzer automated CE system (Advanced Analytical Technologies, Inc.).

Library Preparation and RNA-Seq

The RNA samples were processed by the NGS lab at the Novo Nordisk Foundation Center for Biosustainability (Technical University of Denmark). The samples were prepared with Illumina's TruSeq Stranded mRNA sample preparation kit according to manufacturer's instructions, pooled and sequenced on a NextSeq 500 machine (Illumina) using the NextSeq High Output Kit v2, for 75 cycles of single-end reads for an average of 10 million reads per sample.

RNA-Seq Analysis

Sequencing reads were aligned against the NCBI CHO-K1 genome (GCF_000223135.1) using STAR[133]. Aligned reads were quantified using HTSeq[134] to obtain counts for annotated genes. Differential gene expression was determined by DESeq2[135] following their standard workflow. Genes with fewer than 10 counts across all samples were excluded from analysis. Comparisons were made for Target X knockout vs control cells in media containing and lacking glutamine as well as Target X knockout or control cells in media lacking glutamine vs control cells in media containing glutamine. CHO gene IDs were mapped to human genes and analyzed using GSEA[136]. Three specific gene sets, Hallmark Interferon Alpha Response, Hallmark Interferon Gamma Response, and Browne Interferon Responsive Genes[137] were examined on the individual gene level.

Acknowledgements

The authors wish to thank Nachon Charanyanonda Petersen for assistance with FACS and Anna Koza, Alexandra Hoffmeyer, and Pannipa Pornpitapong for assistance with NGS. This work was supported by the Novo Nordisk Foundation (NNF10CC1016517 and NNF16OC0021638).

CHAPTER 5

CONCLUDING REMARKS AND FUTURE DIRECTIONS

In this thesis we draw on research carried out in other mammalian cells and explore novel CRISPR tools with the overarching aim of expanding the CRISPR toolbox for genetic engineering of CHO cells. We deployed CRISPRi to repress the expression of apoptosis related genes and CRISPRa for increasing the expression of endogenous genes. We further carried out a CHO-specific pooled knockout CRISPR screen, targeting genes involved in metabolism to identify novel genes involved in glutamine metabolism. We have shown that the three tools can work independently, laying the groundwork for genome-scale CRISPR screening in CHO cells.

In **Chapter II**, we show that it is possible to repress endogenous genes in CHO cells using CRISPRi. We tested different inhibitor Cas9 constructs and found dCas9 fused with KRAB-MeCP2 to be most effective, in line with studies in other mammalian cells[89]. CRISPRi's closest competitor is RNAi. While RNAi has a longer history as an engineering tool, there are several advantages to CRISPRi. It presents minimal off-target effects[63,138], which is a considerable issue for RNAi[139]. Furthermore, by acting on the DNA level rather than the mRNA level, CRISPRi allows the manipulation of regulatory elements and other nuclear-localized components, such as long noncoding RNAs[128] and miRNAs[41].

We further applied CRISPRa to CHO cells, and succeeded in increasing transcription of endogenous genes as shown in **Chapter III**, albeit with different success for the different genes and the different TSSs tested. CRISPRa represents a competitor to overexpressing genes using cDNA with several advantages. For example, CRISPRa allows you to increase expression of genes with the closest to natural post-transcriptional processing in their native genomic content and is especially attractive for genes that are large or difficult to clone[140]. In addition, CRISPRa allows you to easily activate multiple genes simultaneously. One of the main challenges going forward is fine-tuning the TSS annotations and CHO genome assembly onto which the gRNAs are designed such that we can get a robust activation across all target sites. As mentioned in **Chapter I**, aside from gRNA proximity to the TSS the second-most predictive factor for successful activation is nucleosome position[49] and incorporation of this data will further optimize gRNA design.

Both CRISPRi and CRISPRa are valuable tools in their own right, but they are especially useful when combined with pooled CRISPR screening. In our final chapter, **Chapter IV**, we showed that it is possible to carry out a large-scale CRISPR knockout screen in CHO cells. We believe that for successful identification of novel targets it is imperative to have designed a library consisting of CHO-specific gRNAs. We showed that when using our library and screening in glutamine free media, we can get both expected and novel targets that we can verify.

These three studies open the possibility of using CRISPR screening to investigate the underlying genetic mechanisms of CHO cells under specific conditions. The gRNA design, Cas operators and screening methods can easily be adjusted to fit the purpose of the screen, whether it be genome-wide searches or smaller focused studies.

The gRNA library can be designed to target all annotated genes in the CHO genome such that we can perform true genome-wide screens. However, there is a trade-off to increasing the number of gRNAs--the handling of a larger library is labour intensive and the results can become more noisy and difficult to interpret. Depending on the purpose you can customize your library, for example, gRNAs can be designed to target only expressed genes when performing a CRISPR knockout screen, or only non-expressed genes when performing a CRISPRa screen. By designing the library with constructs containing two gRNAs, it is also possible to perform combinatorial screens to study genetic interactions[141–143].

Ongoing research into new Cas orthologues, both by mining nature and driving mutagenesis of desired features, have gained traction in the wake of CRISPR/Cas9. These may confer advantages to future screening experiments, expanding the target range by having less stringent requirements for the PAM sequence and decreasing the off-target effects[144]. Cas variance also makes it possible to perform screening of cells with simultaneous gene knockout and activation by designing the library with dual gRNAs that independently works with different Cas operators[145–147]. Lastly, second generation Cas proteins with clear ownership (such as Cpf1[148]) or free from licence (such as Mad7[149]) present attractive alternatives to circumvent the complicated patent landscape surrounding Cas9[150].

Careful choice of a selection that will bring out desired biological properties is essential. For example, flow cytometry assisted selection of protein expression[151,152], viability screening such as the one we carried out in **Chapter IV** or drug resistance screens[153,154]. One way to look for traits that are attractive to the industry is to screen for cells containing gRNAs that confer high productivity, as has been seen in HEK cells[151]; by fluorescently labeling the product, and sorting for cells with high fluorescence, gene perturbations conferring high productivity can be identified. Another option that is of industrial interest is to screen for cells that grow in media containing toxic by-products--already performed in bacteria[155]--such as lactate or ammonia and see if there are any gRNAs that provide superior growth under these conditions. While perhaps not as relevant for CHO engineering purposes, creative screening conditions have been used to study issues ranging from viral infection[156], responsiveness to immunotherapy[157] to cancer metastasis[67] in other cell lines, illustrating the broad utility of large-scale CRISPR screens.

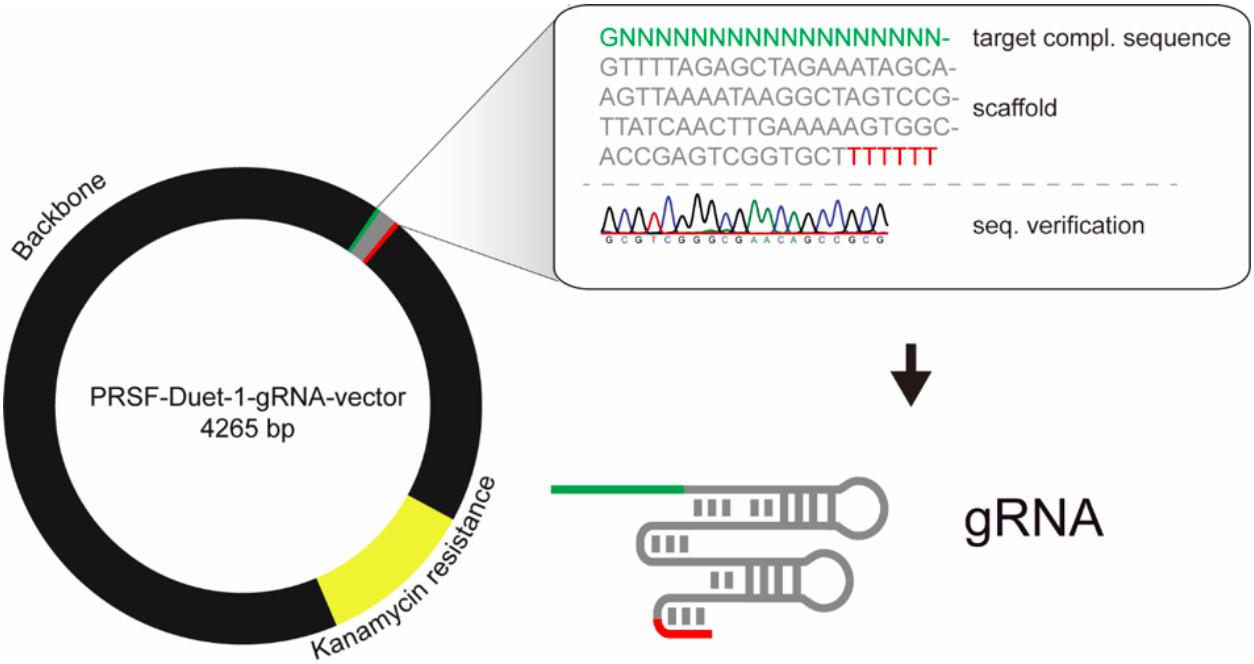
Only a few years ago, the number of genes that had been perturbed in CHO (i.e., overexpressed or inhibited/knocked out) numbered fewer than a hundred[16]. While CRISPR screens can be labour intensive, their impact is undeniable, especially for cells from non-model organisms, which lack much of the information that has accumulated for other organisms (e.g., yeast, *E. coli*, or mouse). We now can simultaneously investigate perturbations of all genes of the CHO genome, which should greatly accelerate the rate at which genotype-phenotype relationships can be discovered, enabling improved biotherapeutic protein production.

APPENDIX

Supplementary Information for Chapter II

gRNA Design and gRNA Vector Construction

gRNAs were designed to target the region immediately downstream of the gene start position from a draft version of the latest hamster genome for *Bak*, *Bax* and *Casp3*[28]. Three gRNA vectors were designed for each target gene using CRISPy algorithm[71] and target sequences are listed in Supplemental Table S1. Oligonucleotides (Supplemental Table S2) were designed for Uracil-Specific Excision Reagent (USER) friendly cloning technology[158] containing the target binding sequence of the gRNA were synthesized, annealed, cloned and transformed in *E. coli* MACH1 for amplification and plasmid purification. The U6 promoter, the targeting region and the gRNA scaffold were verified by Sanger sequencing. The gRNA expression plasmid map and vector sequence is indicated as below:



GGGGAATTGTGAGCGGATAACAATTCCCCTGTAGAAATAATTTTGTTTAACTTTAATAAGGAG
ATATACCATGGGCAGCAGCCATCACCATCATCACCACAGCCAGGATCCGAATTCGAGCTCGGT
ACTGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGGTACCAAGGTC
GGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTT
AGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGT
AGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCAT
ATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGA

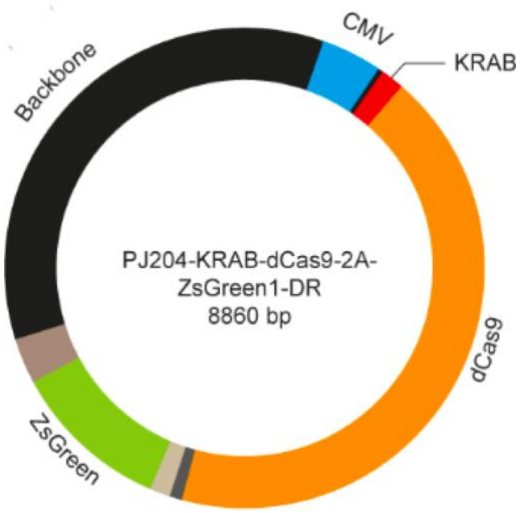
AACACCNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTTCTAGACCCAGCTTT
CTTGTACAAAGTTGGCATTATGCGGCCGCATAATGCTTAAGTCGAACAGAAAGTAATCGTATT
GTACACGGCCGCATAATCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACA
ATTCCCCATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGGCAGATCTCAATT
GGATATCGGCCGGCCACGCGATCGCTGACGTCGGTACCCTCGAGTCTGGTAAAGAAACCGCT
GCTGCGAAATTTGAACGCCAGCACATGGACTCGTCTACTAGCGCAGCTTAATTAACCTAGGCT
GCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGG
TTTTTTGCTGAAACCTCAGGCATTTGAGAAGCACACGGTCACACTGCTTCCGGTAGTCAATAA
ACCGGTAAACCAGCAATAGACATAAGCGGCTATTTAACGACCCTGCCCTGAACCGACGACAAG
CTGACGACCGGGTCTCCGCAAGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGT
TTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATCGA
GCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCG
TTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATC
GGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAG
GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTAT
GCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATC
AACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGGTCGCTGTTAAA
AGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAA
TATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGGGGATCGCAGT
GGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATA
AATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGC
CATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAATCGATAGATTGTCGCACCTG
ATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAA
TCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATACTCTTCCTTTTTCAATATTAT
TGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAAT
AAACAAATAGGCATGCAGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTACGCTCGGTCGTT
CGACTGCGGCGAGCGGTGTCAGCTCACTCAAAAGCGGTAATACGGTTATCCACAGAATCAGG
GGATAAAGCCGGAAAGAACATGTGAGCAAAAAGCAAAGCACCGGAAGAAGCCAACGCCGCA
GGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGCCAGAG
GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC
GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGT
GGCGCTTTCTCATAGCTCACGCTGTTGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT
GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCT
TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCATTGGTAACTGATTTAG
AGGACTTTGTCTTGAAGTTATGCACCTGTTAAGGCTAAACTGAAAGAACAGATTTTGGTGAG
TGCGGTCCTCCAACCCACTTACCTTGGTTCAAAGAGTTGGTAGCTCAGCGAACCTTGAGAAAA
CCACCGTTGGTAGCGGTGGTTTTTCTTTATTTATGAGATGATGAATCAATCGGTCTATCAAGT
CAACGAACAGCTATTCCGTTACTCTAGATTTCAGTGCAATTTATCTCTTCAAATGTAGCACCTG
AAGTCAGCCCCATACGATATAAGTTGTAATTCTCATGTTAGTCATGCCCCGCGCCCACCGGAA
GGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTG
AGCTAACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGC
CAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGG
GTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGA
GAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGT

GGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATGT
CCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCG
TTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTTGCATGGTTTGTTGAAAA
CCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGA
TATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAG
CGCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATG
GGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACAT
TAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATGATCAGC
CCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCG
TTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGAC
AATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGT
TTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCC
ACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTG
ATAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCT
GAATTGACTCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGT
GTCCGGGATCTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAATTAATACGACTCACTA
TA

dCas9 Repressor Vector Construction

The CHO codon-optimized dCas9 cassette was generated from a synthesized CHO codon-optimized wild-type Cas9 (GeneArt, Thermo Fisher Scientific) by mutagenesis of Cas9 to introduce the D10A and H840A mutations using Q5 Site-Directed mutagenesis kit (New England Biolabs). This CHO codon-optimized dCas9 was fused to a KRAB domain at N- or C- terminal to construct a KRAB-dCas9 or dCas9-KRAB expression vector. Additionally the TRD of MeCP2 (cloned from Addgene plasmid #110821) was fused to C- terminal of dCas9-KRAB to generate the vector dCas9-KRAB-MeCP2. All plasmid cloning works were performed by using the USER friendly cloning technology[158]. All functional DNA elements of the final vectors were verified by Sanger sequencing (Eurofins Genomics GmbH, Ebersberg, Germany). The plasmid maps and sequences of KRAB-dCas9, dCas9-KRAB and dCas9-KRAB-MeCP2 were shown as below. Blue letters indicate CMV promoter, while red, orange and violet ones indicate KRAB, dCas9 and MeCP2 elements respectively.

KRAB-dCas9



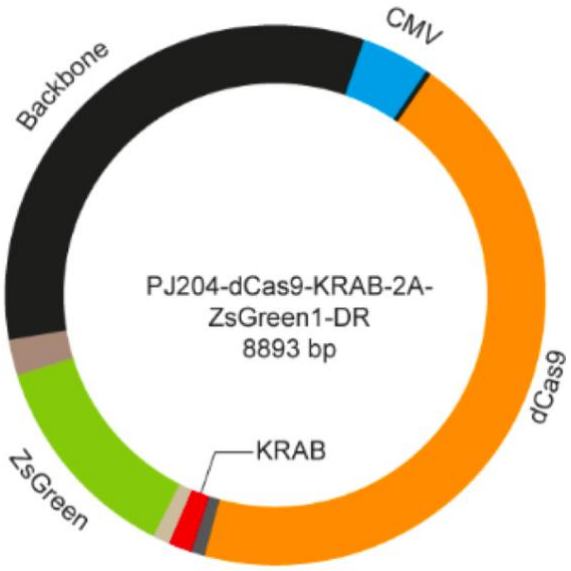
AGTCGGTGTGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAG
TTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGAC
CGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC
AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGC
ATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT
CGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACT
CACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATC
AACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGT
GTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTG
GCTTATCGAAATAAGCAGCGTCGCCACCATGGACGCGAAATCACTTACGGCATGGTCGAGAA
CACTGGTTACGTTCAAGGACGTGTTTGTGGACTTTACACGTGAGGAGTGGAAATTGCTGGAT
ACTGCGCAACAAATTGTGTATCGAAATGTCATGCTTGAGAATTACAAGAACCTCGTCAGTCTC
GGATACCAGTTGACGAAACCGGATGTGATCCTTAGGCTCGAAAAGGGGGAAGAACCTAAAG
CCTATGACAAGAAATACTCCATCGGCCTTGCCATCGGGACCAACAGCGTGGGTTGGGCCGTG
ATCACTGACGAATACAAAGTGCCATCGAAAAAGTTTAAGGTGCTGGGCAATACTGACAGACA
TAGCATCAAGAAAAACCTGATTGGAGCCCTCCTGTTCGACTCCGGGGAAACCGCCGAAGCCA
CTCGCCTGAAACGGACCGCACGCCGGAGGTACACTCGCCGGAAGAATAGAATCTGCTACCTT
CAAGAGATCTTTAGCAACGAGATGGCCAAAGTGGACGATTCGTTCTTCCACCGGCTGGAGGA
GTCATTCCTGGTCGAAGAGGACAAAAAGCATGAGAGACATCCGATCTTCGGTAACATTGTGG
ATGAGGTCGCCTACCACGAGAAGTACCCCACTATCTACCATTTGCGCAAGAAGTTGGTCGACT
CGACTGATAAGGCCGACCTCAGACTCATCTACCTCGCTTTGGCACACATGATCAAGTTTAGGG
GTCACTTCCTGATTGAAGGGGATCTGAACCCAGACAACTCGGATGTGGATAAACTGTTTATCC
AGCTGGTCCAGACTTACAATCAGCTCTTTGAGGAGAACCCGATCAATGCCTCCGGCGTGGAT
GCTAAGGCAATCCTGTCCGCACGCTTGTCAAAGAGCCGGCGGCTCGAAAATCTGATCGCACA
GCTTCCGGGAGAAAAGAAAAACGGACTGTTTGGAAACCTTATCGCCCTCTCGCTGGGACTCA
CTCCTAACTTCAAATCCAATTTTGATTTGGCCGAGGATGCCAAGCTGCAGCTGTCGAAGGACA
CCTACGACGATGACCTCGACAATCTTCTGGCCCAGATCGGGGACCAGTACGCCGATCTGTTCC
TCGCGGCCAAAAACCTGTCAGACGCAATCTTGCTGAGCGATATCTTGAGGGTCAATACCGAA

ATCACTAAGGCTCCATTGTCAGCATCGATGATCAAGAGGTACGATGAACACCATCAGGACCTG
ACCCTGCTCAAGGCACTGGTGCGGCAGCAACTGCCGGAAAAGTACAAGGAGATCTTCTTCGA
TCAAAGCAAGAATGGGTACGCAGGGTACATCGATGGTGGAGCATCACAAGAAGAGTTCTACA
AATTCATCAAGCCCATCCTTGAAAAGATGGACGGGACGGAAGAACTGCTGGTGAAGCTGAAT
CGGGAAGATCTGCTGCGGAAGCAGCGCACCTTCGACAATGGTTCCATCCCACACCAAATCCAT
CTCGGCGAACTGCACGCGATCCTTCGCCGGCAGGAAGATTTCTACCCGTTCTTGAAGGATAAC
AGAGAAAAGATCGAGAAAATCCTGACCTTTAGAATCCCGTACTACGTGGGCCCGTTGGCTCG
CGGAAACTCAAGATTCGCCTGGATGACTAGAAAATCCGAAGAAACGATTACCCCGTGGAACT
TTGAAGAGGTCGTCGACAAGGGAGCCTCGGCACAGTCGTTCATCGAACGGATGACCAATTTC
GACAAGAACCTCCCCAACGAAAAGGTGCTCCCGAAACACTCGCTGCTGTATGAGTACTTCACG
GTCTACAACGAGCTGACCAAGGTGAAGTACGTGACCGAAGGAATGCGGAAGCCCGCTTTTCT
GTCCGGTGAACAGAAAAAGGCCATCGTGGACCTCCTCTTCAAGACTAACAGAAAGGTGACCG
TGAAACAGCTTAAAGAGGACTACTTCAAGAAGATCGAGTGTTTTGACTCCGTGGAAATCTCG
GGAGTGGAGGATAGGTTTAACGCTTCGCTGGGGACCTACCATGACCTTCTCAAGATCATCAA
GGATAAGGATTTCCTGGACAACGAAGAAAACGAGGACATCCTTGAGGACATCGTGTTGACCT
TGACCCTGTTTGAGGATCGGGAAATGATCGAGGAGAGACTGAAAACTTACGCGCATCTTTTC
GACGACAAAGTGATGAAGCAGCTGAAAAGAAGAAGATACACTGGCTGGGGACGGCTGTCGA
GGAAACTGATTAACGGAATTAGGGATAAACAAAGCGGAAAGACTATCCTCGATTTTCTCAAG
TCGGACGGCTTCGCCAATCGGAACTTTATGCAGCTCATCCACGACGATAGCCTGACCTTCAAG
GAAGATATCCAAAAAGCCCAAGTGTCCGGCCAGGGAGATAGCTTGCACGAGCACATTGCTAA
CCTGGCCGGTTCACCAGCCATTAAGAAGGGAATCCTCCAAACCGTGAAGGTGGTCGACGAAC
TCGTCAAGGTGATGGGCCGCCATAAGCCAGAGAACATCGTGATCGAGATGGCTCGCGAAAAT
CAGACTACTCAGAAGGGTCAAAAGAACTCCCGCGAGCGCATGAAGCGCATCGAAGAAGGAAT
TAAGGAGCTGGGCTCACAAATCCTCAAAGAACATCCTGTCGAGAACACCCAACTTCAGAACGA
GAAACTGTACCTCTACTATCTCCAAAATGGCCGGGACATGTACGTCGATCAGGAATTGGATAT
CAACCGCCTCTCAGACTACGACGTGGATGCTATCGTGCCGCAGTCGTTTCTCAAAGATGATAG
CATCGACAACAAGGTCCTGACGAGGTCGGATAAGAACCGCGGGAAATCCGACAATGTGCCTT
CGGAAGAGGTGGTGAAAAAGATGAAGAACTACTGGAGGCAATTGCTTAATGCCAAACTGAT
CACCCAGCGCAAATTCGACAACCTTACTAAGGCCGAGCGCGGAGGACTGTCCGAACTCGATA
AGGCGGGGTTCATCAAAAGACAATTGGTCGAAACTCGGCAAATTACCAAGCATGTCGCTCAG
ATCCTCGACTCGCGCATGAACACTAAGTATGATGAGAACGACAAGCTCATTCGCGAAGTGAA
AGTGATTACCCTCAAATCAAAGCTGGTGTCGGACTTCAGAAAAGACTTTCAATTCTACAAGGT
GCGGGAGATTAACAACTACCACCACGCGCACGACGCCTACCTCAATGCAGTGGTGGGAACCG
CCCTCATCAAGAAGTATCCAAAGCTCGAAAGCGAGTTCGTGTACGGGGACTATAAAGTGTAC
GATGTGCGGAAAATGATCGCAAAGAGCGAGCAAGAGATCGGCAAAGCAACGGCCAAATACT
TCTTCTACAGCAACATTATGAACTTTTTCAAGACCGAGATTACCCTGGCGAACGGAGAAATCC
GGAAGCGGCCTTTGATCGAAACGAATGGGGAAACTGGAGAAATCGTGTGGGACAAGGGCA
GAGACTTTGCGACCGTGAGAAAGGTGCTGTCCATGCCACAAGTCAACATCGTGAAGAAAACT
GAAGTGCAGACTGGAGGATTTTCCAAGGAGTCAATCCTTCCGAAGCGCAACAGCGACAAGCT
GATCGCCAGGAAGAAGGATTGGGACCCCAAGAAGTACGGTGGTTTTGATTCACCTACTGTCG
CTTACAGCGTGCTGGTGGTGGCCAAGGTGGAAAAGGGGAAGTCAAAGAAATTGAAGTCGGT
GAAGGAATTGCTGGGGATTACTATCATGGAGAGGAGCAGCTTCGAAAAGAATCCCATCGACT
TCTTGGAGGCCAAGGGATACAAAGAGGTGAAGAAAGACCTGATCATTAAGCTGCCGAAGTAC
TCCTTGTTCGAACTGGAAAACGGCAGAAAGCGGATGCTGGCCTCCGCCGGAGAGCTGCAGA
AGGGTAACGAGCTGGCTCTGCCCAGCAAATACGTGAATTTCCTCTACCTGGCCTCCCATTACG

AGAAGCTCAAGGGAAGCCCGGAGGATAATGAGCAAAAACAACTGTTCGTCGAGCAGCACAA
GCACTACCTCGACGAGATCATCGAACAAATCTCCGAGTTCTCGAAGCGGGTGATTCTGGCCG
ACGCAAACCTTGATAAAGTCCTCAGCGCCTACAACAAGCACCGCGACAAACCGATCAGAGAAC
AAGCTGAGAACATCATCCACCTGTTCACGCTGACTAACTTGGGAGCGCCTGCCGCCTTCAAGT
ACTTCGACACCACTATCGATCGGAAACGGTACACTTCCACTAAGGAAGTGCTGGACGCAACCC
TGATCCATCAGTCCATCACCGGACTTTATGAGACTCGGATCGATCTCAGCCAGCTTGGCGGTG
ATTCAAGAGCTGACCCGAAGAAGAAGCGCAAGGTCAGACGTCATGGCAGTGGAGAGGGCAG
AGGAAGTCTGCTAACATGCGGTGACGTCGAGGAGAATCCTGGCCCAATGGCCCAGTCCAAGC
ACGGCCTGACCAAGGAGATGACCATGAAGTACCGCATGGAGGGCTGCGTGGACGGCCACAA
GTTCGTGATCACCGGCGAGGGCATCGGCTACCCCTTCAAGGGCAAGCAGGCCATCAACCTGT
GCGTGGTGGAGGGCGGCCCCTTGCCCTTCGCCGAGGACATCTTGTCCGCCGCCTTCATGTAC
GGCAACCGCGTGTTCACCGAGTACCCCCAGGACATCGTCGACTACTTCAAGAACTCCTGCCCC
GCCGGCTACACCTGGGACCGCTCCTTCCTGTTCGAGGACGGCGCCGTGTGCATCTGCAACGC
CGACATCACCGTGAGCGTGGAGGAGAACTGCATGTACCACGAGTCCAAGTTCTACGGCGTGA
ACTTCCCCGCCGACGGCCCCGTGATGAAGAAGATGACCGACAACTGGGAGCCCTCCTGCGAG
AAGATCATCCCCGTGCCCAAGCAGGGCATCTTGAAGGGCGACGTGAGCATGTACCTGCTGCT
GAAGGACGGTGGCCGCTTGCGCTGCCAGTTCGACACCGTGTACAAGGCCAAGTCCGTGCCCC
GCAAGATGCCCGACTGGCACTTCATCCAGCACAAGCTGACCCGCGAGGACCGCAGCGACGCC
AAGAACCAGAAGTGGCACCTGACCGAGCACGCCATCGCCTCCGGCTCCGCCTTGCCCAAGCTT
CCGCGGAGCCATGGCTTCCCGCCGGCGGTGGCGGCGCAGGATGATGGCACGCTGCCCATGT
CTTGTGCCCAGGAGAGCGGGATGGACCGTCACCCTGCAGCCTGTGCTTCTGCTAGGATCAAT
GTGTAGACACAGTCTCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCC
TTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATC
GCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGG
GGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGACTTGCG
TAGTGAGTCGAATAAGGGCGACACAAAATTTATTCTAAATGCATAATAAATACTGATAACATC
TTATAGTTTGTATTATATTTTGTATTATCGTTGACATGTATAATTTTGATATCAAAAACTGATTT
TCCCTTTATTATTTTCGAGATTTATTTTCTTAATTCTCTTTAACAAACTAGAAATATTGTATATA
CAAAAAATCATAAATAATAGATGAATAGTTTAATTATAGGTGTTCATCAATCGAAAAAGCAAC
GTATCTTATTTAAAGTGCGTTGCTTTTTTCTCATTTATAAGGTTAAATAATTCTCATATATCAA
GCAAAGTGACAGGCGCCCTTAAATATTCTGACAAATGCTCTTTCCCTAAACTCCCCCCATAAAA
AAACCCGCCGAAGCGGGTTTTTACGTTATTTGCGGATTAACGATTACTCGTTATCAGAACCGC
CCAGGGGGCCCGAGCTTAAGACTGGCCGTCGTTTTACAACACAGAAAGAGTTTGTAGAAACG
CAAAAAGGCCATCCGTCAGGGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTCCCTACTCTCGC
CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAG
CTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG
TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA
TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC
CGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTC
CGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTC
ATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTG
CACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACC
CGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGG
TATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGGCTAACTACGGCTACACTAGAAGAAC
AGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTG

ATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGC
GCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG
AACGACGCGCGCGTAACTCACGTTAAGGGATTTTGGTCATGAGCTTGCGCCGTCCCGTCAAG
TCAGCGTAATGCTCTGCTTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCT
ATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTT
ACCATCTGGCCCCAGCGCTGCGATGATACCGCGAGAACCACGCTCACCGGCTCCGGATTTATC
AGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCT
CCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGC
GCAACGTTGTTGCCATCGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCAT
TCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGG
TTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG
TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGG
TGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGG
CGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAAC
GTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCA
CTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAAC
AGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATA
TTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT
TTGAATGTATTTAGAAAAATAAACAAATAGGGGTCAGTGTTACAACCAATTAACCAATTCTGA
ACATTATCGCGAGCCCATTTATACCTGAATATGGCTCATAACACCCCTTGTTTGCCTGGCGGC
AGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGA
TGGTAGTGTGGGGACTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAA
GGCTCAGTCGAAAGACTGGGCCTTTCGCCCGGGCTAATTATGGGGTGTCGCCCTTATTCGAC
TC

dCas9-KRAB



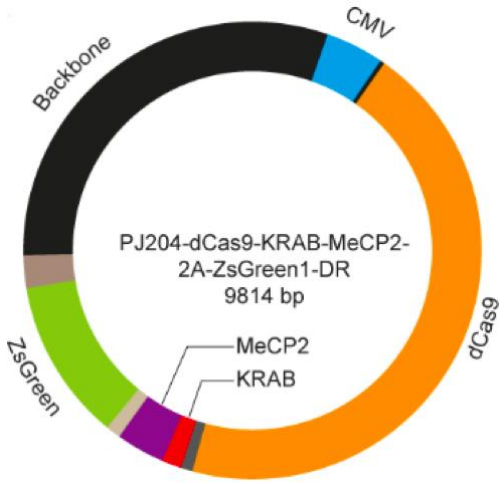
AGTCGGTGTGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAG
TTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGAC
CGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC
AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGC
ATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT
CGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACT
CACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATC
AACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGT
GTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTG
GCTTATCGAAATAAGCAGCGTCGCCACCATGGACAAGAAATACTCCATCGGCCTTGCCATCGG
GACCAACAGCGTGGGTTGGGCCGTGATCACTGACGAATACAAAGTGCCATCGAAAAAGTTTA
AGGTGCTGGGCAATACTGACAGACATAGCATCAAGAAAAACCTGATTGGAGCCCTCCTGTTC
GACTCCGGGGAAACCGCCGAAGCCACTCGCCTGAAACGGACCGCACGCCGGAGGTACACTCG
CCGGAAGAATAGAATCTGCTACCTTCAAGAGATCTTTAGCAACGAGATGGCCAAAGTGGACG
ATTCGTTCTTCCACCGGCTGGAGGAGTCATTCCTGGTCGAAGAGGACAAAAAGCATGAGAGA
CATCCGATCTTCGGTAACATTGTGGATGAGGTCGCCTACCACGAGAAGTACCCCACTATCTAC
CATTTGCGCAAGAAGTTGGTCGACTCGACTGATAAGGCCGACCTCAGACTCATCTACCTCGCT
TTGGCACACATGATCAAGTTTAGGGGTCACTTCCTGATTGAAGGGGATCTGAACCCAGACAA
CTCGGATGTGGATAAACTGTTTATCCAGCTGGTCCAGACTTACAATCAGCTCTTTGAGGAGAA
CCCGATCAATGCCTCCGGCGTGGATGCTAAGGCAATCCTGTCCGCACGCTTGTCAAAGAGCC
GGCGGCTCGAAAATCTGATCGCACAGCTTCCGGGAGAAAAGAAAAACGGACTGTTTGGAAAC
CTTATCGCCCTCTCGCTGGGACTCACTCCTAACTTCAAATCCAATTTTGATTTGGCCGAGGATG
CCAAGCTGCAGCTGTCGAAGGACACCTACGACGATGACCTCGACAATCTTCTGGCCCAGATCG
GGGACCAGTACGCCGATCTGTTCCTCGCGGCCAAAAACCTGTCAGACGCAATCTTGCTGAGC
GATATCTTGAGGGTCAATACCGAAATCACTAAGGCTCCATTGTCAGCATCGATGATCAAGAG
GTACGATGAACACCATCAGGACCTGACCCTGCTCAAGGCACTGGTGCGGCAGCAACTGCCGG

AAAAGTACAAGGAGATCTTCTTCGATCAAAGCAAGAATGGGTACGCAGGGTACATCGATGGT
GGAGCATCACAAGAAGAGTTCTACAAATTCATCAAGCCCATCCTTGAAAAGATGGACGGGAC
GGAAGAACTGCTGGTGAAGCTGAATCGGGAAGATCTGCTGCGGAAGCAGCGCACCTTCGAC
AATGGTTCCATCCCACACCAAATCCATCTCGGCGAACTGCACGCGATCCTTCGCCGGCAGGAA
GATTTCTACCCGTTCTTGAAGGATAACAGAGAAAAGATCGAGAAAATCCTGACCTTTAGAATC
CCGTACTACGTGGGCCCGTTGGCTCGCGGAAACTCAAGATTCGCCTGGATGACTAGAAAATC
CGAAGAAACGATTACCCCGTGGAACTTTGAAGAGGTCGTCGACAAGGGAGCCTCGGCACAGT
CGTTCATCGAACGGATGACCAATTTCGACAAGAACCTCCCCAACGAAAAGGTGCTCCCGAAAC
ACTCGCTGCTGTATGAGTACTTCACGGTCTACAACGAGCTGACCAAGGTGAAGTACGTGACC
GAAGGAATGCGGAAGCCCGCTTTTCTGTCCGGTGAACAGAAAAAGGCCATCGTGGACCTCCT
CTTCAAGACTAACAGAAAGGTGACCGTGAAACAGCTTAAAGAGGACTACTTCAAGAAGATCG
AGTGTTTTGACTCCGTGGAAATCTCGGGAGTGGAGGATAGGTTTAACGCTTCGCTGGGGACC
TACCATGACCTTCTCAAGATCATCAAGGATAAGGATTTCCTGGACAACGAAGAAAACGAGGA
CATCCTTGAGGACATCGTGTTGACCTTGACCCTGTTTGAGGATCGGGAAATGATCGAGGAGA
GACTGAAAACTTACGCGCATCTTTTCGACGACAAAGTGATGAAGCAGCTGAAAAGAAGAAGA
TACACTGGCTGGGGACGGCTGTCGAGGAAACTGATTAACGGAATTAGGGATAAACAAAGCG
GAAAGACTATCCTCGATTTTCTCAAGTCGGACGGCTTCGCCAATCGGAACTTTATGCAGCTCA
TCCACGACGATAGCCTGACCTTCAAGGAAGATATCCAAAAAGCCCAAGTGTCCGGCCAGGGA
GATAGCTTGCACGAGCACATTGCTAACCTGGCCGGTTCACCAGCCATTAAGAAGGGAATCCTC
CAAACCGTGAAGGTGGTCGACGAACTCGTCAAGGTGATGGGCCGCCATAAGCCAGAGAACA
TCGTGATCGAGATGGCTCGCGAAAATCAGACTACTCAGAAGGGTCAAAAGAACTCCCGCGAG
CGCATGAAGCGCATCGAAGAAGGAATTAAGGAGCTGGGCTCACAAATCCTCAAAGAACATCC
TGTCGAGAACACCCAACTTCAGAACGAGAAACTGTACCTCTACTATCTCCAAAATGGCCGGGA
CATGTACGTCGATCAGGAATTGGATATCAACCGCCTCTCAGACTACGACGTGGATGCTATCGT
GCCGCAGTCGTTTCTCAAAGATGATAGCATCGACAACAAGGTCCTGACGAGGTCGGATAAGA
ACCGCGGGAAATCCGACAATGTGCCTTCGGAAGAGGTGGTGAAAAAGATGAAGAACTACTG
GAGGCAATTGCTTAATGCCAAACTGATCACCCAGCGCAAATTCGACAACCTTACTAAGGCCGA
GCGCGGAGGACTGTCCGAACTCGATAAGGCGGGGTTCATCAAAAGACAATTGGTCGAAACTC
GGCAAATTACCAAGCATGTCGCTCAGATCCTCGACTCGCGCATGAACACTAAGTATGATGAGA
ACGACAAGCTCATTCGCGAAGTGAAAGTGATTACCCTCAAATCAAAGCTGGTGTCGGACTTCA
GAAAAGACTTTCAATTCTACAAGGTGCGGGAGATTAACAACTACCACCACGCGCACGACGCCT
ACCTCAATGCAGTGGTGGGAACCGCCCTCATCAAGAAGTATCCAAAGCTCGAAAGCGAGTTC
GTGTACGGGGACTATAAAGTGTACGATGTGCGGAAAATGATCGCAAAGAGCGAGCAAGAGA
TCGGCAAAGCAACGGCCAAATACTTCTTCTACAGCAACATTATGAACTTTTTCAAGACCGAGA
TTACCCTGGCGAACGGAGAAATCCGGAAGCGGCCTTTGATCGAAACGAATGGGGAAACTGG
AGAAATCGTGTGGGACAAGGGCAGAGACTTTGCGACCGTGAGAAAGGTGCTGTCCATGCCA
CAAGTCAACATCGTGAAGAAAACTGAAGTGCAGACTGGAGGATTTTCCAAGGAGTCAATCCT
TCCGAAGCGCAACAGCGACAAGCTGATCGCCAGGAAGAAGGATTGGGACCCCAAGAAGTAC
GGTGGTTTTGATTCACCTACTGTCGCTTACAGCGTGCTGGTGGTGGCCAAGGTGGAAAAGG
GGAAGTCAAAGAAATTGAAGTCGGTGAAGGAATTGCTGGGGATTACTATCATGGAGAGGAG
CAGCTTCGAAAAGAATCCCATCGACTTCTTGGAGGCCAAGGGATACAAAGAGGTGAAGAAAG
ACCTGATCATTAAGCTGCCGAAGTACTCCTTGTTCGAACTGGAAAACGGCAGAAAGCGGATG
CTGGCCTCCGCCGGAGAGCTGCAGAAGGGTAACGAGCTGGCTCTGCCCAGCAAATACGTGA
ATTTCCTCTACCTGGCCTCCCATTACGAGAAGCTCAAGGGAAGCCCGGAGGATAATGAGCAA
AAACAACTGTTCGTCGAGCAGCACAAGCACTACCTCGACGAGATCATCGAACAAATCTCCGAG

TTCTCGAAGCGGGTGATTCTGGCCGACGCAAACCTTGATAAAGTCCTCAGCGCCTACAACAAG
CACCGCGACAAACCGATCAGAGAACAAGCTGAGAACATCATCCACCTGTTCACGCTGACTAAC
TTGGGAGCGCCTGCCGCCTTCAAGTACTTCGACACCACTATCGATCGGAAACGGTACACTTCC
ACTAAGGAAGTGCTGGACGCAACCCTGATCCATCAGTCCATCACCGGACTTTATGAGACTCGG
ATCGATCTCAGCCAGCTTGGCGGTGATTCAAGAGCTGACCCGAAGAAGAAGCGCAAGGTCAG
TGGTGGAGGAAGTGGCGGGTCAGGGTCGATGGACGCGAAATCACTTACGGCATGGTCGAG
AACACTGGTTACGTTCAAGGACGTGTTTGTGGACTTTACACGTGAGGAGTGGAAATTGCTGG
ATACTGCGCAACAAATTGTGTATCGAAATGTCATGCTTGAGAATTACAAGAACCTCGTCAGTC
TCGGATACCAGTTGACGAAACCGGATGTGATCCTTAGGCTCGAAAAGGGGGAAGAACCTTG
GCTGGTAAGACGTCATGGCAGTGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTGACGTC
GAGGAGAATCCTGGCCCAATGGCCCAGTCCAAGCACGGCCTGACCAAGGAGATGACCATGAA
GTACCGCATGGAGGGCTGCGTGGACGGCCACAAGTTCGTGATCACCGGCGAGGGCATCGGC
TACCCCTTCAAGGGCAAGCAGGCCATCAACCTGTGCGTGGTGGAGGGCGGCCCCTTGCCCTT
CGCCGAGGACATCTTGTCCGCCGCCTTCATGTACGGCAACCGCGTGTTCACCGAGTACCCCCA
GGACATCGTCGACTACTTCAAGAACTCCTGCCCCGCCGGCTACACCTGGGACCGCTCCTTCCT
GTTCGAGGACGGCGCCGTGTGCATCTGCAACGCCGACATCACCGTGAGCGTGGAGGAGAAC
TGCATGTACCACGAGTCCAAGTTCTACGGCGTGAACTTCCCCGCCGACGGCCCCGTGATGAA
GAAGATGACCGACAACTGGGAGCCCTCCTGCGAGAAGATCATCCCCGTGCCCAAGCAGGGCA
TCTTGAAGGGCGACGTGAGCATGTACCTGCTGCTGAAGGACGGTGGCCGCTTGCGCTGCCA
GTTCGACACCGTGTACAAGGCCAAGTCCGTGCCCCGCAAGATGCCCGACTGGCACTTCATCCA
GCACAAGCTGACCCGCGAGGACCGCAGCGACGCCAAGAACCAGAAGTGGCACCTGACCGAG
CACGCCATCGCCTCCGGCTCCGCCTTGCCCAAGCTTCCGCGGAGCCATGGCTTCCCGCCGGCG
GTGGCGGCGCAGGATGATGGCACGCTGCCCATGTCTTGTGCCCAGGAGAGCGGGATGGACC
GTCACCCTGCAGCCTGTGCTTCTGCTAGGATCAATGTGTAGACACAGTCTCTGTGCCTTCTAG
TTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCC
ACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
TGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCAT
GCTGGGGATGCGGTGGGCTCTATGGACTTGCGTAGTGAGTCGAATAAGGGCGACACAAAAT
TTATTCTAAATGCATAATAAATACTGATAACATCTTATAGTTTGTATTATATTTTGTATTATCGT
TGACATGTATAATTTTGATATCAAAAACTGATTTTCCCTTTATTATTTTCGAGATTTATTTTCTT
AATTCTCTTTAACAAACTAGAAATATTGTATATACAAAAAATCATAAATAATAGATGAATAGTT
TAATTATAGGTGTTCATCAATCGAAAAAGCAACGTATCTTATTTAAAGTGCGTTGCTTTTTTCT
CATTTATAAGGTTAAATAATTCTCATATATCAAGCAAAGTGACAGGCGCCCTTAAATATTCTGA
CAAATGCTCTTTCCCTAAACTCCCCCCATAAAAAAACCCGCCGAAGCGGGTTTTTACGTTATTT
GCGGATTAACGATTACTCGTTATCAGAACCGCCCAGGGGGCCCGAGCTTAAGACTGGCCGTC
GTTTTACAACACAGAAAGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGGGCCTTCTGC
TTAGTTTGATGCCTGGCAGTTCCCTACTCTCGCCTTCCGCTTCCTCGCTCACTGACTCGCTGCG
CTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCA
CAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA
CCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAA
AAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTC
CCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGC
CTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGT
GTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCG
CCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAG

CAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA
GTGGTGGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGC
CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGC
GGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCC
TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGACGCGCGCGTAACTCACGTTAAGG
GATTTTGGTCATGAGCTTGCGCCGTCCCGTCAAGTCAGCGTAATGCTCTGCTTTTACCAATGC
TTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCC
CCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGCGCTGCGATGATA
CCGCGAGAACCACGCTCACCGGCTCCGGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC
CGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGA
AGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATCGCTACAGGCAT
CGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG
AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGT
CAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGA
ATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCAC
ATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGA
TCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATC
TTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGG
GAATAAGGGCGACACGGAAATGTTGAATACTCATATTCTTCCTTTTTCAATATTATTGAAGCA
TTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAA
TAGGGGTCAGTGTTACAACCAATTAACCAATTCTGAACATTATCGCGAGCCCATTTATACCTG
AATATGGCTCATAACACCCCTTGTTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCC
ATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGACTCCCCATGCGA
GAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCG
CCCGGGCTAATTATGGGGTGTCGCCCTTATTCGACTC

dCas9-KRAB-MeCP2



AGTCGGTGTGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAG
TTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGAC
CGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC
AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGC
ATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT
CGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACT
CACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATC
AACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGT
GTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTG
GCTTATCGAAATAAGCAGCGTCGCCACCATGGACAAGAAATACTCCATCGGCCTTGCCATCGG
GACCAACAGCGTGGGTTGGGCCGTGATCACTGACGAATACAAAGTGCCATCGAAAAAGTTTA
AGGTGCTGGGCAATACTGACAGACATAGCATCAAGAAAAACCTGATTGGAGCCCTCCTGTTC
GACTCCGGGGAAACCGCCGAAGCCACTCGCCTGAAACGGACCGCACGCCGGAGGTACACTCG
CCGGAAGAATAGAATCTGCTACCTTCAAGAGATCTTTAGCAACGAGATGGCCAAAGTGGACG
ATTCGTTCTTCCACCGGCTGGAGGAGTCATTCCTGGTCGAAGAGGACAAAAAGCATGAGAGA
CATCCGATCTTCGGTAACATTGTGGATGAGGTCGCCTACCACGAGAAGTACCCCACTATCTAC
CATTTGCGCAAGAAGTTGGTCGACTCGACTGATAAGGCCGACCTCAGACTCATCTACCTCGCT
TTGGCACACATGATCAAGTTTAGGGGTCACTTCCTGATTGAAGGGGATCTGAACCCAGACAA
CTCGGATGTGGATAAACTGTTTATCCAGCTGGTCCAGACTTACAATCAGCTCTTTGAGGAGAA
CCCGATCAATGCCTCCGGCGTGGATGCTAAGGCAATCCTGTCCGCACGCTTGTCAAAGAGCC
GGCGGCTCGAAAATCTGATCGCACAGCTTCCGGGAGAAAAGAAAAACGGACTGTTTGGAAAC
CTTATCGCCCTCTCGCTGGGACTCACTCCTAACTTCAAATCCAATTTTGATTTGGCCGAGGATG
CCAAGCTGCAGCTGTCGAAGGACACCTACGACGATGACCTCGACAATCTTCTGGCCCAGATCG
GGGACCAGTACGCCGATCTGTTCCTCGCGGCCAAAAACCTGTCAGACGCAATCTTGCTGAGC
GATATCTTGAGGGTCAATACCGAAATCACTAAGGCTCCATTGTCAGCATCGATGATCAAGAG
GTACGATGAACACCATCAGGACCTGACCCTGCTCAAGGCACTGGTGCGGCAGCAACTGCCGG
AAAAGTACAAGGAGATCTTCTTCGATCAAAGCAAGAATGGGTACGCAGGGTACATCGATGGT
GGAGCATCACAAGAAGAGTTCTACAAATTCATCAAGCCCATCCTTGAAAAGATGGACGGGAC

GGAAGAACTGCTGGTGAAGCTGAATCGGGAAGATCTGCTGCGGAAGCAGCGCACCTTCGAC
AATGGTTCCATCCCACACCAAATCCATCTCGGCGAACTGCACGCGATCCTTCGCCGGCAGGAA
GATTTCTACCCGTTCTTGAAGGATAACAGAGAAAAGATCGAGAAAATCCTGACCTTTAGAATC
CCGTACTACGTGGGCCCGTTGGCTCGCGGAAACTCAAGATTCGCCTGGATGACTAGAAAATC
CGAAGAAACGATTACCCCGTGGAACTTTGAAGAGGTCGTCGACAAGGGAGCCTCGGCACAGT
CGTTCATCGAACGGATGACCAATTTCGACAAGAACCTCCCCAACGAAAAGGTGCTCCCGAAAC
ACTCGCTGCTGTATGAGTACTTCACGGTCTACAACGAGCTGACCAAGGTGAAGTACGTGACC
GAAGGAATGCGGAAGCCCGCTTTTCTGTCCGGTGAACAGAAAAAGGCCATCGTGGACCTCCT
CTTCAAGACTAACAGAAAGGTGACCGTGAAACAGCTTAAAGAGGACTACTTCAAGAAGATCG
AGTGTTTTGACTCCGTGGAAATCTCGGGAGTGGAGGATAGGTTTAACGCTTCGCTGGGGACC
TACCATGACCTTCTCAAGATCATCAAGGATAAGGATTTCCTGGACAACGAAGAAAACGAGGA
CATCCTTGAGGACATCGTGTTGACCTTGACCCTGTTTGAGGATCGGGAAATGATCGAGGAGA
GACTGAAAACTTACGCGCATCTTTTCGACGACAAAGTGATGAAGCAGCTGAAAAGAAGAAGA
TACACTGGCTGGGGACGGCTGTCGAGGAAACTGATTAACGGAATTAGGGATAAACAAAGCG
GAAAGACTATCCTCGATTTTCTCAAGTCGGACGGCTTCGCCAATCGGAACTTTATGCAGCTCA
TCCACGACGATAGCCTGACCTTCAAGGAAGATATCCAAAAAGCCCAAGTGTCCGGCCAGGGA
GATAGCTTGCACGAGCACATTGCTAACCTGGCCGGTTCACCAGCCATTAAGAAGGGAATCCTC
CAAACCGTGAAGGTGGTCGACGAACTCGTCAAGGTGATGGGCCGCCATAAGCCAGAGAACA
TCGTGATCGAGATGGCTCGCGAAAATCAGACTACTCAGAAGGGTCAAAAGAACTCCCGCGAG
CGCATGAAGCGCATCGAAGAAGGAATTAAGGAGCTGGGCTCACAAATCCTCAAAGAACATCC
TGTCGAGAACACCCAACTTCAGAACGAGAAACTGTACCTCTACTATCTCCAAAATGGCCGGGA
CATGTACGTCGATCAGGAATTGGATATCAACCGCCTCTCAGACTACGACGTGGATGCTATCGT
GCCGCAGTCGTTTCTCAAAGATGATAGCATCGACAACAAGGTCCTGACGAGGTCGGATAAGA
ACCGCGGGAAATCCGACAATGTGCCTTCGGAAGAGGTGGTGAAAAAGATGAAGAACTACTG
GAGGCAATTGCTTAATGCCAAACTGATCACCCAGCGCAAATTCGACAACCTTACTAAGGCCGA
GCGCGGAGGACTGTCCGAACTCGATAAGGCGGGGTTCATCAAAAGACAATTGGTCGAAACTC
GGCAAATTACCAAGCATGTCGCTCAGATCCTCGACTCGCGCATGAACACTAAGTATGATGAGA
ACGACAAGCTCATTCGCGAAGTGAAAGTGATTACCCTCAAATCAAAGCTGGTGTCGGACTTCA
GAAAAGACTTTCAATTCTACAAGGTGCGGGAGATTAACAACTACCACCACGCGCACGACGCCT
ACCTCAATGCAGTGGTGGGAACCGCCCTCATCAAGAAGTATCCAAAGCTCGAAAGCGAGTTC
GTGTACGGGGACTATAAAGTGTACGATGTGCGGAAAATGATCGCAAAGAGCGAGCAAGAGA
TCGGCAAAGCAACGGCCAAATACTTCTTCTACAGCAACATTATGAACTTTTTCAAGACCGAGA
TTACCCTGGCGAACGGAGAAATCCGGAAGCGGCCTTTGATCGAAACGAATGGGGAAACTGG
AGAAATCGTGTGGGACAAGGGCAGAGACTTTGCGACCGTGAGAAAGGTGCTGTCCATGCCA
CAAGTCAACATCGTGAAGAAAACTGAAGTGCAGACTGGAGGATTTTCCAAGGAGTCAATCCT
TCCGAAGCGCAACAGCGACAAGCTGATCGCCAGGAAGAAGGATTGGGACCCCAAGAAGTAC
GGTGGTTTTGATTCACCTACTGTCGCTTACAGCGTGCTGGTGGTGGCCAAGGTGGAAAAGG
GGAAGTCAAAGAAATTGAAGTCGGTGAAGGAATTGCTGGGGATTACTATCATGGAGAGGAG
CAGCTTCGAAAAGAATCCCATCGACTTCTTGGAGGCCAAGGGATACAAAGAGGTGAAGAAAG
ACCTGATCATTAAGCTGCCGAAGTACTCCTTGTTCGAACTGGAAAACGGCAGAAAGCGGATG
CTGGCCTCCGCCGGAGAGCTGCAGAAGGGTAACGAGCTGGCTCTGCCCAGCAAATACGTGA
ATTTCCTCTACCTGGCCTCCCATTACGAGAAGCTCAAGGGAAGCCCGGAGGATAATGAGCAA
AAACAACTGTTCGTCGAGCAGCACAAGCACTACCTCGACGAGATCATCGAACAAATCTCCGAG
TTCTCGAAGCGGGTGATTCTGGCCGACGCAAACCTTGATAAAGTCCTCAGCGCCTACAACAAG
CACCGCGACAAACCGATCAGAGAACAAGCTGAGAACATCATCCACCTGTTCACGCTGACTAAC

TTGGGAGCGCCTGCCGCCTTCAAGTACTTCGACACCACTATCGATCGGAAACGGTACACTTCC
ACTAAGGAAGTGCTGGACGCAACCCTGATCCATCAGTCCATCACCGGACTTTATGAGACTCGG
ATCGATCTCAGCCAGCTTGGCGGTGATTCAAGAGCTGACCCGAAGAAGAAGCGCAAGGTCAG
TGGTGGAGGAAGTGGCGGGTCAGGGTCGATGGACGCGAAATCACTTACGGCATGGTCGAG
AACACTGGTTACGTTCAAGGACGTGTTTGTGGACTTTACACGTGAGGAGTGGAAATTGCTGG
ATACTGCGCAACAAATTGTGTATCGAAATGTCATGCTTGAGAATTACAAGAACCTCGTCAGTC
TCGGATACCAGTTGACGAAACCGGATGTGATCCTTAGGCTCGAAAAGGGGGAAGAACCTTG
GCTGGTATCGGGAGGTGGTTCGGGTGGCTCTGGATCAAGCCCAAAGAAGAAACGGAAGGT
GGAAGCCTCAGTGCAGGTGAAAAGGGTGCTGGAAAAATCCCCCGGCAAACTCCTCGTGAAG
ATGCCCTTCCAGGCTTCCCCTGGCGGAAAAGGTGAAGGGGGTGGCGCAACCACATCTGCCCA
GGTCATGGTCATCAAGCGACCTGGAAGGAAAAGAAAGGCCGAGGCTGACCCTCAGGCCATTC
CAAAGAAACGGGGACGCAAGCCAGGGTCCGTGGTCGCAGCTGCAGCAGCTGAGGCTAAGAA
AAAGGCAGTGAAGGAAAGCTCCATCCGCAGTGTGCAGGAGACTGTCCTGCCCATCAAGAAGA
GGAAGACTAGGGAGACCGTGTCCATCGAGGTCAAAGAAGTGGTCAAGCCCCTGCTCGTGTCC
ACCCTGGGCGAAAAATCTGGAAAGGGGCTCAAAACATGCAAGTCACCTGGACGGAAAAGCA
AGGAGTCTAGTCCAAAGGGGCGCTCAAGCTCCGCTTCTAGTCCCCCTAAAAAGGAACACCATC
ACCATCACCATCACGCCGAGTCTCCTAAGGCTCCTATGCCACTGCTCCCACCACCTCCACCACC
TGAGCCACAGTCAAGCGAAGACCCCATCAGCCCACCCGAGCCTCAGGATCTGTCCTCTAGTAT
TTGCAAAGAGGAAAAGATGCCCAGAGCAGGCAGCCTGGAGAGTGATGGCTGTCCAAAAGAA
CCCGCCAAGACCCAGCCTATGGTGGCAGCCGCTGCAACTACCACCACAACCACAACTACCACA
GTGGCCGAAAAATACAAGCATCGCGGCGAGGGCGAACGAAAGGACATTGTGTCAAGCTCCA
TGCCCAGACCTAACCGGGAGGAACCAGTCGATAGTAGGACACCCGTGACTGAGAGAGTCTCA
AGACGTCATGGCAGTGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTGACGTCGAGGAG
AATCCTGGCCCAATGGCCCAGTCCAAGCACGGCCTGACCAAGGAGATGACCATGAAGTACCG
CATGGAGGGCTGCGTGGACGGCCACAAGTTCGTGATCACCGGCGAGGGCATCGGCTACCCC
TTCAAGGGCAAGCAGGCCATCAACCTGTGCGTGGTGGAGGGCGGCCCCTTGCCCTTCGCCGA
GGACATCTTGTCCGCCGCCTTCATGTACGGCAACCGCGTGTTCACCGAGTACCCCCAGGACAT
CGTCGACTACTTCAAGAACTCCTGCCCCGCCGGCTACACCTGGGACCGCTCCTTCCTGTTCGA
GGACGGCGCCGTGTGCATCTGCAACGCCGACATCACCGTGAGCGTGGAGGAGAACTGCATG
TACCACGAGTCCAAGTTCTACGGCGTGAACTTCCCCGCCGACGGCCCCGTGATGAAGAAGAT
GACCGACAACTGGGAGCCCTCCTGCGAGAAGATCATCCCCGTGCCCAAGCAGGGCATCTTGA
AGGGCGACGTGAGCATGTACCTGCTGCTGAAGGACGGTGGCCGCTTGCGCTGCCAGTTCGA
CACCGTGTACAAGGCCAAGTCCGTGCCCCGCAAGATGCCCGACTGGCACTTCATCCAGCACAA
GCTGACCCGCGAGGACCGCAGCGACGCCAAGAACCAGAAGTGGCACCTGACCGAGCACGCC
ATCGCCTCCGGCTCCGCCTTGCCCAAGCTTCCGCGGAGCCATGGCTTCCCGCCGGCGGTGGC
GGCGCAGGATGATGGCACGCTGCCCATGTCTTGTGCCCAGGAGAGCGGGATGGACCGTCAC
CCTGCAGCCTGTGCTTCTGCTAGGATCAATGTGTAGACACAGTCTCTGTGCCTTCTAGTTGCC
AGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGT
CCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGG
GGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTG
GGGATGCGGTGGGCTCTATGGACTTGCGTAGTGAGTCGAATAAGGGCGACACAAAATTTAT
TCTAAATGCATAATAAATACTGATAACATCTTATAGTTTGTATTATATTTTGTATTATCGTTGA
CATGTATAATTTTGATATCAAAAACTGATTTTCCCTTTATTATTTTCGAGATTTATTTTCTTAAT
TCTCTTTAACAAACTAGAAATATTGTATATACAAAAAATCATAAATAATAGATGAATAGTTTAA
TTATAGGTGTTCATCAATCGAAAAAGCAACGTATCTTATTTAAAGTGCGTTGCTTTTTTCTCAT

TTATAAGGTTAAATAATTCTCATATATCAAGCAAAGTGACAGGCGCCCTTAAATATTCTGACAA
ATGCTCTTTCCCTAAACTCCCCCCATAAAAAAACCCGCCGAAGCGGGTTTTTACGTTATTTGCG
GATTAACGATTACTCGTTATCAGAACCGCCCAGGGGGCCCGAGCTTAAGACTGGCCGTCGTT
TTACAACACAGAAAGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGGGCCTTCTGCTTA
GTTTGATGCCTGGCAGTTCCCTACTCTCGCCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTC
GGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAG
AATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG
TAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAA
TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCC
CTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTT
TCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTA
GGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT
ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGC
CACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG
TGGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT
ACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGG
TTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGAT
CTTTTCTACGGGGTCTGACGCTCAGTGGAACGACGCGCGCGTAACTCACGTTAAGGGATTTT
GGTCATGAGCTTGCGCCGTCCCGTCAAGTCAGCGTAATGCTCTGCTTTTACCAATGCTTAATC
AGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCG
TGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGCGCTGCGATGATACCGCGA
GAACCACGCTCACCGGCTCCGGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCG
CAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG
AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATCGCTACAGGCATCGTGGT
GTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTAC
ATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAG
TAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCAT
GCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG
TATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCA
GAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTAC
CGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTAC
TTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAA
GGGCGACACGGAAATGTTGAATACTCATATTCTTCCTTTTTCAATATTATTGAAGCATTTATCA
GGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGG
TCAGTGTTACAACCAATTAACCAATTCTGAACATTATCGCGAGCCCATTTATACCTGAATATGG
CTCATAACACCCCTTGTTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGA
ACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGACTCCCCATGCGAGAGTAGG
GAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGCCCGGGC
TAATTATGGGGTGTCGCCCTTATTCGACTC

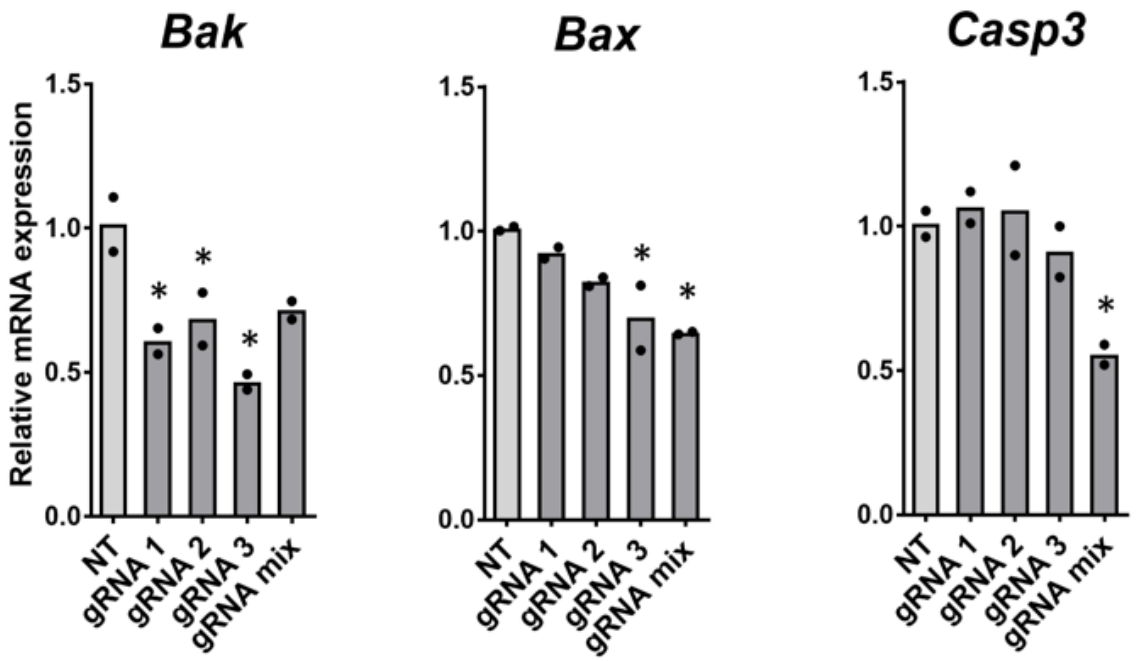
Generation of a KRAB-dCas9-CHO Cell Line

To establish a cell platform for stably expressing KRAB-dCas9, KRAB-dCas9-CHO cell lines were derived from a parent cell line containing a recombinase-mediated cassette exchange (RMCE) landing pad. The parent cell line was made by CRISPR-mediated homology directed targeted integration of CHO-S cells (Life Technologies, Carlsbad, USA) as previously described[86,130]. The parent cell line contains a mCherry coding sequence flanked by a loxP sequence at the 5 end and a lox2272 sequence at the 3 end (pEF1-loxP-mCherry- lox2272-BGHpA), and the 5 and 3 homology arms target at a transcriptionally active site within a non-coding region. Promoterless and polyAless RMCE vectors were constructed by assembly of PCR fragments containing KRAB-dCas9 region that was flanked by loxP and lox2272 sequences. The parent cell line was transfected with RMCE KRAB-dCas9 plasmids and Cre-recombinase vector in 3:1 ratio to exchange mCherry coding sequence with KRAB-dCas9. For Cre recombinase expression, PSF-CMV-CRE recombinase vector (OGS591, Sigma-Aldrich) was used. Transfected cell pool was passaged two times after transfection. After 7 days, flow cytometry revealed that 0.5% of the cells in transfected cell pool were changed from mCherry positive to negative. mCherry-negative single cell sorting was performed by FACS using a BD FACSJazz cell sorter (BD Biosciences). After 14 days, the monoclonal KRAB-dCas9-CHO cell line was established after verification of proper integration and confirmation of single copy of KRAB-dCas9 (Supplemental Figure S4).

Copy Number Analysis Using qPCR

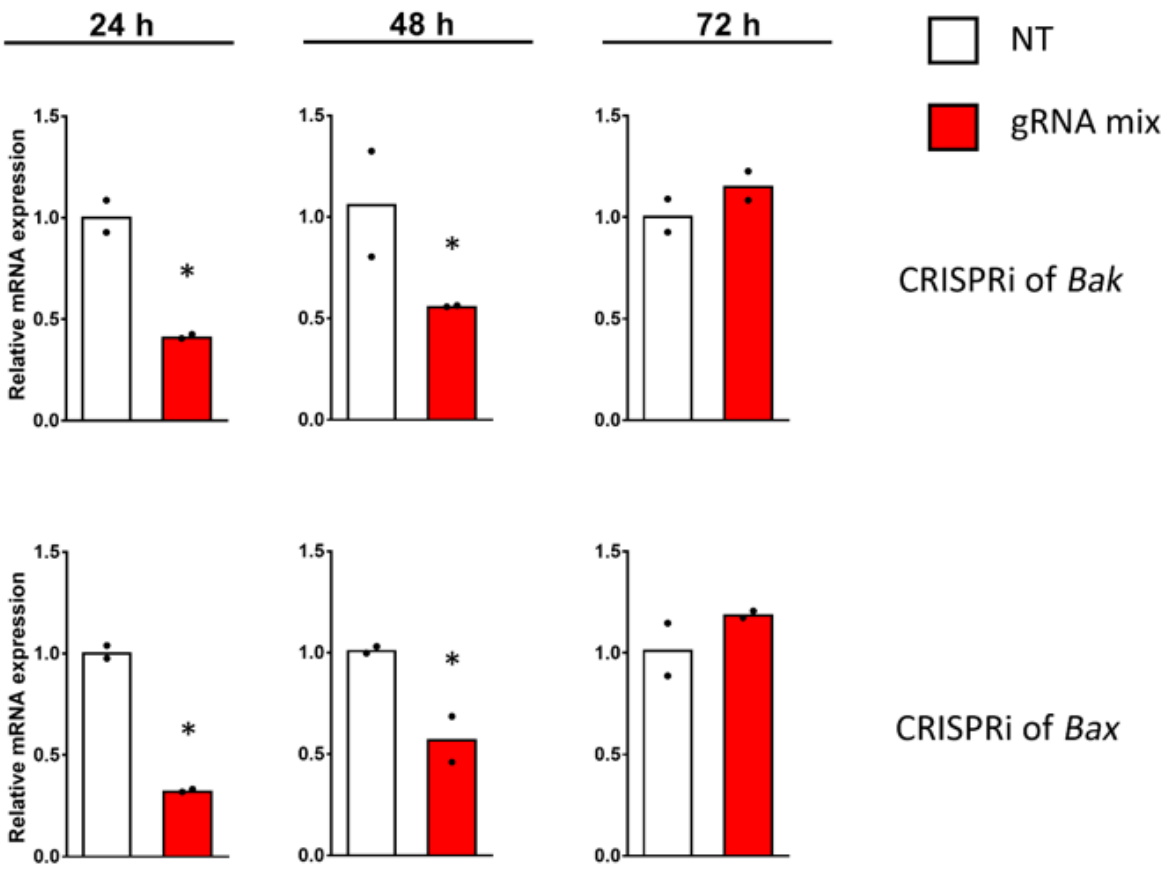
Relative copy numbers of the integrated dCas9 gene was determined by qPCR. Genomic DNA was extracted using the GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific) according to manufacturer's instruction. qPCR was run on a QuantStudio 5 Real-Time PCR System (Agilent Technologies). Amplification was performed under the following conditions: 50°C for 2 min, 95°C for 10 min; 40x: 95°C for 15 s, 60°C for 1 min. Copy numbers of dCas9 were determined with C1GALT1C1 (COSMC) as an internal control gene for normalization using TaqManTM Gene Expression Master Mix and custom-made TaqManTM probes for all three genes (Thermo Fisher Scientific). Primers and probes are

listed in the Supplementary Table S4 and were validated by melting curve analysis and primer efficiency test. A delta-delta threshold cycle (ΔΔCt) method was applied to calculate the copy number of dCas9 integrated using previously generated (d)Cas9-CHO cell lines with a single-copy of (d)Cas9 integrated in the identical site (single-copy calibrators) . Each experiment was performed in technical triplicates using Genomic DNA from CHO-S cells as a negative control.



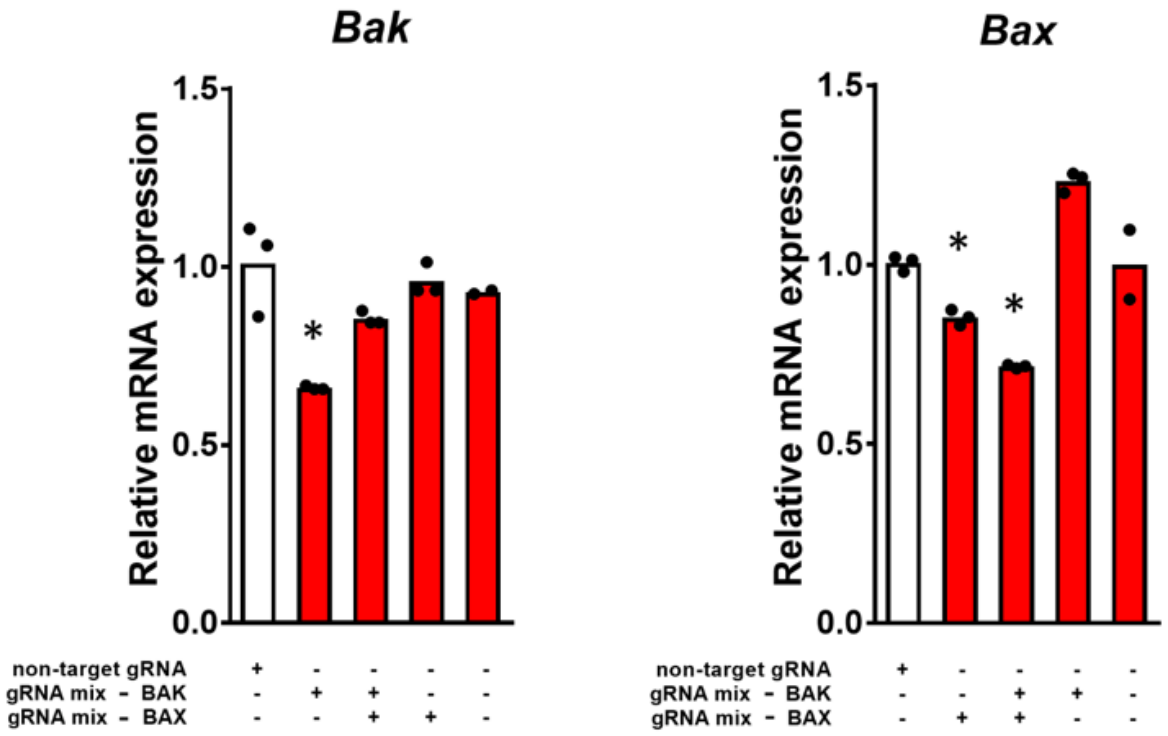
Supplemental Figure S1: Verification of gRNAs Used in this Study.

Transcriptional repression of *Bak*, *Bax* and *Casp3* via transient expression of gRNA and KRAB-dCas9 in CHO-S cell measured by qRT-PCR. Three individual gRNA or the mixture of three gRNAs for each gene were tested in this study. mRNA was harvested 24h after transfection. NT indicates transfection with scrambled non-target gRNA as a negative control. Targeted gene expression fold change shown is the mean of the individual transfection duplicates. *Fkbp1a* and *Gnb1* was used as housekeeping genes to normalizing the fold change. Transfection was performed in biological duplicates (n=2, * indicates $p<0.05$ vs NT) and each transfection repeat was measured with three technical repeats of qRT-PCR. Dot represents the mean of technical triplicates in qRT-PCR.



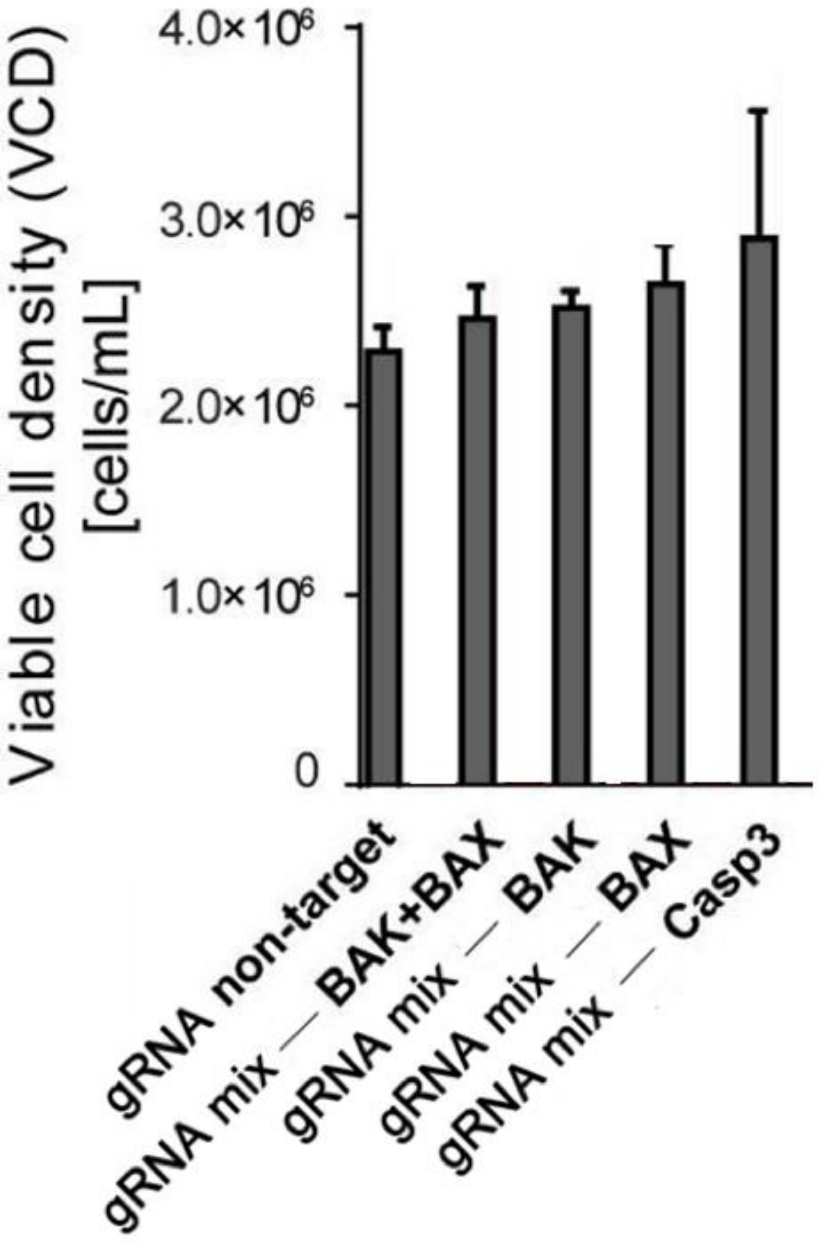
Supplemental Figure S2: Time Course of CRISPRi in CHO Cells.

Transcriptional repression of *Bak* and *Bax* measured by qRT-PCR. mRNA was harvested at 24h, 48h and 72h after transfection of KRAB-dCas9 and respective gRNA mix. NT indicates transfection with scrambled non-target gRNA as a negative control. Targeted gene expression fold was shown as mean of the biological duplicates (n=2, * indicates $p<0.05$ vs NT) which were normalized to *Fkbp1a* and *Gnb1* as housekeeping genes. Each biological repeat was measured with three technical repeats of qRT-PCR. Dot represents the mean of technical triplicates in qRT-PCR.



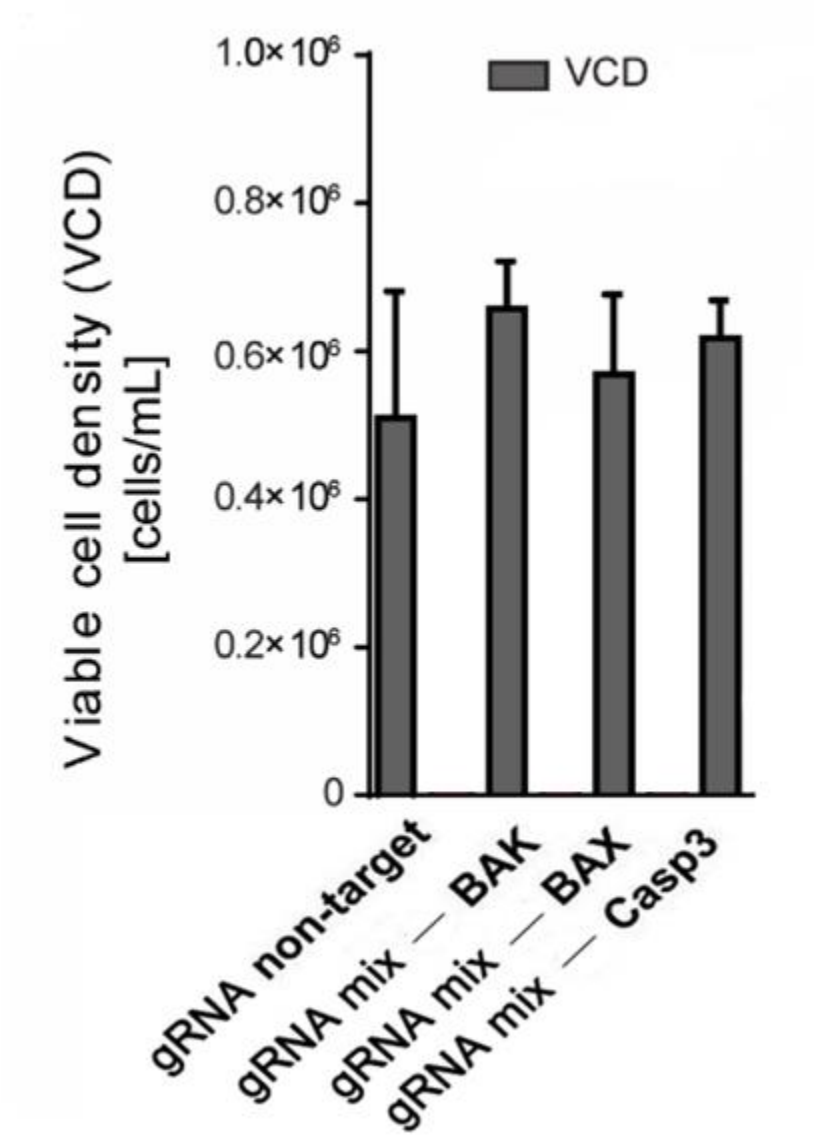
Supplemental Figure S3: Simultaneous CRISPRi of Gene Duplex in CHO.

Transcriptional repression of *Bak* and *Bax* via transient expression of gRNA in KRAB-dCas9-CHO cells measured by qRT-PCR. Simultaneous CRISPRi of *Bak*+*Bax* was performed. mRNA was harvested 24h after transfection. NT indicates transfection with scrambled non-target gRNA as a negative control. Targeted gene expression fold was shown as mean of the biological duplicates (n= 3, * indicates $p<0.05$ vs NT) which were normalized to *Fkbp1a* and *Gnb1* as housekeeping genes. Each transfection repeat was measured in technical triplicates of qRT-PCR. Dot represents the mean of technical triplicates in qRT-PCR.



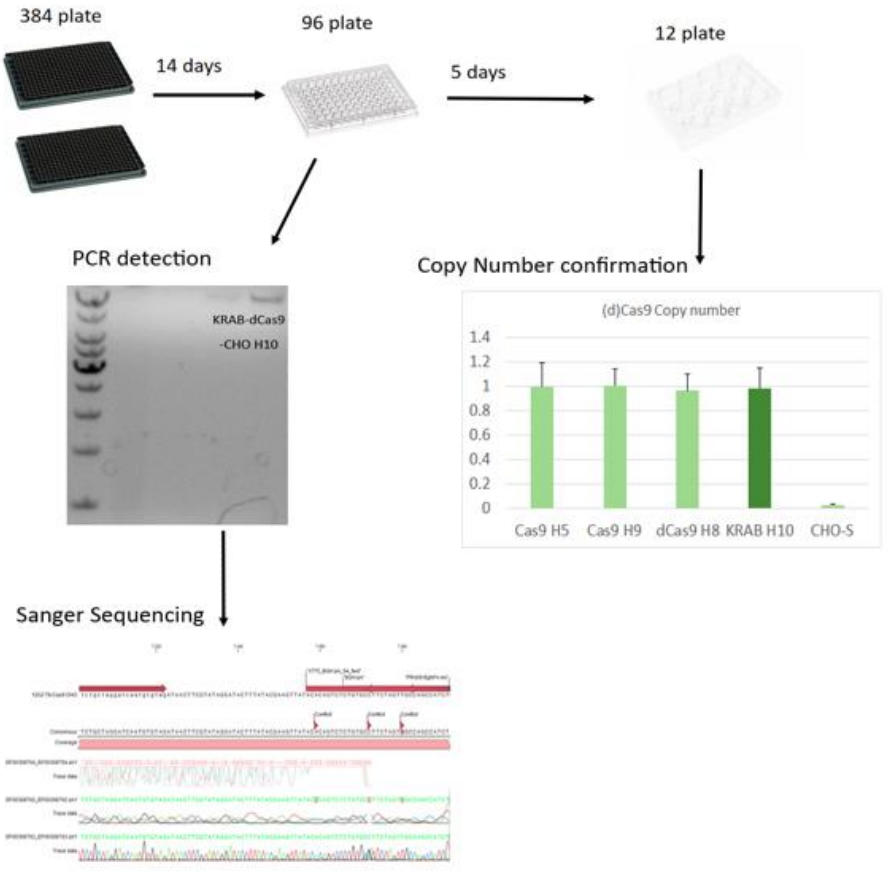
Supplemental Figure S4: Viable Cell Density of CRISPRi in CHO

Viable cell density of CHO cells 24 h post transfection of indicated gRNAs in KRAB-dCas9-CHO cell lines. The mixture of gRNAs for each gene (or gene combination) were tested in this study.



Supplemental Figure S5: Viable Cell Density of CRISPRi in CHO

Viable cell density of CHO cells 24 h post induction of AcD in Figure 2. Transfection of indicated gRNAs in KRAB-dCas9-CHO cell lines was performed 24 h before AcD induction The mixture of gRNAs for each gene were tested in this study. Each treatment was performed in biological triplicates (n=3, * indicates $p<0.05$ vs NT).



Supplemental Figure S6: Process of KRAB-dCas9-CHO Establishment and Verification.

The parent cell line was transfected with RMCE KRAB-dCas9 plasmids and Cre-recombinase and cultured for 7 days followed by flow cytometry analysis. Single cell sorting was performed by FACS plating into 384-well plates. After 14 days, the monoclonal KRAB-dCas9-CHO cell line was verified by PCR and following Sanger sequencing to screen for the positive clone of proper integration. A positive clone was selected for one copy number confirmation of dCas9 measured by qPCR. In copy number measurement, CHO-S cells was used as a negative control and other (d)Cas9-CHO lines previously verified with one copy number were used as calibrators.

Supplemental Table S1: gRNA Target Sequence and Location. T=template strand; NT= Non-Template Strand

gRNA	Target Sequence (5'-3', PAM indicated in bold font)	Strand	Distance downstream from TSS (bp)
Bak gRNA1	TCCTTAGCGGGTTCCCGACG**GGG**	T	235
Bak gRNA2	GCCCCGTCGGGAACCCGCTA**AGG**	NT	291
Bak gRNA3	CCACAGCCGGGCAGGTACGC**TGG**	NT	207
Bax gRNA1	CCCGGGGCGCACCCGGCGAG**AGG**	NT	61
Bax gRNA2	ACATCCGCGTGCGCGCCCCC**GGG**	NT	259
Bax gRNA3	GGCAGGCCCGGGCTTGTCGC**TGG**	NT	280
Casp3 gRNA1	CTGGAGCACGGGCCGCGCCG**AGG**	NT	381
Casp3 gRNA2	GGATGAGCGTGGCGGTCGAG**TGG**	NT	495
Casp3 gRNA3	CCGCCTTTTGCTCCTCGGCG**CGG**	T	134

Supplemental Table S2. Oligos used for gRNA Vector Cloning.

Oligos	Sequence(5'-3')
Bak_1_gRNAfwd	GGAAAGGACGAAACACCGCCTTAGCGGGTTCCCGACGGTTTTAGAGCTAGAAAT
Bak_1_gRNArev	CTAAAACCGTCGGGAACCCGCTAAGGCGGTGTTTCGTCCTTTCCACAAGATAT
Bak_2_gRNAfwd	GGAAAGGACGAAACACCGCCCCGTCGGGAACCCGCTAGTTTTAGAGCTAGAAAT

Bak_2_gRNArev	CTAAAACTAGCGGGTTCCCGACGGGGCGGTGTTTCGTCCTTTCCACAAGATAT
Bak_3_gRNAfwd	GGAAAGGACGAAACACCGCACAGCCGGGCAGGTACGCGTTTTAGAGCTAGAAAT
Bak_3_gRNArev	CTAAAACGCGTACCTGCCCGGCTGTGCGGTGTTTCGTCCTTTCCACAAGATAT
Bax_1_gRNAfwd	GGAAAGGACGAAACACCGCCGGGGCGCACCCGGCGAGGTTTTAGAGCTAGAAAT
Bax_1_gRNArev	CTAAAACCTCGCCGGGTGCGCCCCGGCGGTGTTTCGTCCTTTCCACAAGATAT
Bax_2_gRNAfwd	GGAAAGGACGAAACACCGCATCCGCGTGCGCGCCCCCGTTTTAGAGCTAGAAAT
Bax_2_gRNArev	CTAAAACGGGGGCGCGCACGCGGATGCGGTGTTTCGTCCTTTCCACAAGATAT
Bax_3_gRNAfwd	GGAAAGGACGAAACACCGGCAGGCCCGGGCTTGTCGCGTTTTAGAGCTAGAAAT
Bax_3_gRNArev	CTAAAACGCGACAAGCCCGGGCCTGCCGGTGTTTCGTCCTTTCCACAAGATAT
Casp3_1_gRNAfwd	GGAAAGGACGAAACACCGTGGAGCACGGGCCGCGCCGGTTTTAGAGCTAGAAAT
Casp3_1_gRNArev	CTAAAACCGGCGCGGCCCGTGCTCCACGGTGTTTCGTCCTTTCCACAAGATAT
Casp3_2_gRNAfwd	GGAAAGGACGAAACACCGGATGAGCGTGGCGGTCGAGGTTTTAGAGCTAGAAAT
Casp3_2_gRNArev	CTAAAACCTCGACCGCCACGCTCATCCGGTGTTTCGTCCTTTCCACAAGATAT
Casp3_3_gRNAfwd	GGAAAGGACGAAACACCGCGCCTTTTGCTCCTCGGCGGTTTTAGAGCTAGAAAT
Casp3_3_gRNArev	CTAAAACCGCCGAGGAGCAAAAGGCGCGGTGTTTCGTCCTTTCCACAAGATAT

Supplemental Table S3. The amount of gRNA and dCas9 repressor used for transfection. Total amount of 3.75 μg of DNA were used for each well of transfection.

Groups	gRNAs	dCas9 repressors
Transfection of gRNA(s) into KRAB-dCas9-CHO cell	3.75 μg (1.25 μg for each gRNA when applying three gRNA mix)	/
Co-transfection of gRNA(s) and dCas9 repressor into CHO-S cell	1.875 μg (0.625 μg for each gRNA when applying three gRNA mix)	1.875 μg

Supplemental Table S4. Primers and Probes used for qRT-PCR and qPCR

Target Gene	Type	Sequence	Reporter-Quencher pair
Bak	Primer forward	GAGGAGGTCTTTCGAAGCTATG	-
Bak	Primer reverse	AAGATGCTGTTGGGTTCTAGG	-
Bax	Primer forward	ATCCGTCTACCAAGAAGTTGAG	-
Bax	Primer reverse	GGGAGTTAGTATCCACATTAGCA	-
Casp3	Primer forward	GCCATGGTGAAGAAGGAGTAA	-
Casp3	Primer reverse	CCAGTTAGACTTCGGCAATAGT	-

Fkbp1a	Primer forward	CTCTCGGGACAGAAACAAGC	-
Fkbp1a	Primer reverse	GACCTACACTCATCTGGGCTAC	-
Gnb1	Primer forward	CCATATGTTTCTTTCCCAATGGC	-
Gnb1	Primer reverse	AAGTCGTCGTACCCAGCAAG	-
Cas9	Primer forward	CCGCCCTCATCAAGAAGTATC	-
Cas9	Primer reverse	AGCTCGAAAGCGAGTTCGTGTACG	-
COSMC	Primer forward	GCAGCCTTTCTATCTAGGACAC	-
COSMC	Primer reverse	CCACCTTGTTCAGGACACTT	-
Bak	Probe	CCACCATCAGGAACAAGAGACCCA	FAM-MGB
Bax	Probe	AGGCGAATTGGAGATGAGCTGGAC	FAM-MGB
Casp3	Probe	ACAGATGGGCCTGTTGACCTGAAA	FAM-MGB
Fkbp1a	Probe	ATGCTAGGCAAGCAGGAGGTGATC	VIC-MGB
Gnb1	Probe	ACTGGTTCAGACGATGCTACGTGC	ABY-QSY
Cas9	Probe	CGATCTCTTGCTCGCTCTTT	FAM-MGB
COSMC	Probe	AGTGACAGCCATATTGGAACAGCATCC	VIC-MGB

Supplementary Information for Chapter III

Supplementary Methods

Methods for Identifying Experimental TSS

RNA Processing

Sequence data for 5'-GRO-seq[159], GRO-seq[160],and a variant of Start-seq[161,162] were quality controlled using FastQC, and cutadapt[163] was used to trim adapter sequences and low quality bases from the reads. Reads were aligned to a draft version of the recently published updated Chinese Hamster genome[28] using STAR[133]. Reads mapping to multiple locations were removed from analysis.

Peak Calling

To call TSS peaks, the Homer version 4.10 5GRO-Seq pipeline was used. Briefly, aligned reads for TSS samples and control samples were estimated to have a fragment size of 1 basepair (bp). Counts, or tags, were normalized to a million mapped reads, or CPM. Regions of the genome are then scanned at a fixed width of 150 bps and clusters regions that have maximum density of tags. As clusters are found, the regions immediately adjacent are excluded to ensure there are no piggyback peaks feed off the signal of large peaks. Peaks must be greater than 2x the peak width apart from on another. This continues until all tags have been assigned to clusters. After all clusters have been found, a tag threshold is established to correct for the fact that we may expect to see clusters simply by random chance. These are modelled as a Poisson distribution and the expected number of counts is determined. A False Discovery Rate of 0.001 is used for multiple hypothesis correction. Importantly, in experiments where the cap is enriched, efficiency is not perfect, and additional reads tend to occur in high-expressing genes. To correct for this, we use control samples, GRO-Seq and START-input for GRO-Cap 9 and START, respectively. These experiments do not enrich for the 5 cap, and

thus will be found along the gene body. We enforce our peaks to be more than 2-fold enriched compared to the controls.

Merging Samples

Samples were merged using Homer's mergePeaks command[164]. Briefly, if samples have peaks that overlap at all, they are combined into one, where the Start position is min(Start_1,Start_2) and the End is max(End_1,End_2). Additionally, when merging the samples peak expression in the same tissue, the average CPM is used.

mRNA TSS Calling

Since both TSS methods enrich for nascent transcripts, many peaks end up in enhancer regions, where enhancer RNA (eRNA) is transcribed. Therefore, in order to annotate protein-coding TSSs, a distance threshold from the annotated gene was enforced. Ultimately, we used a distance of -1kb to +1kb from the initial reported TSS. We used a draft version of an expanded annotation of the Chinese hamster[165] that incorporates multiple omic data types and identifies additional putative isoforms for many genes. Additionally, intron enhancers are known to occur, and so any intron peaks were filtered.

Peak Naming and Ranking

Peaks identified at the permissive threshold were ranked by the total number of tags supporting each and then sequentially numbered (for example, p1@GFAP corresponds to the promoter of GFAP which has the highest tag support). This is a similar approach to how the FANTOM Consortium reported their TSSs in mouse and the human. Additionally, if there is no peak for that gene, the original RefSeq location is used, and the name is numbered as 0 (for example p0@RRP7A). In order to rank peaks of the same transcript and gene, we looked to take into account both the number of samples and the expression value itself. Todo this, CPM was first log transformed, and the median value of all samples with a peak was multiplied by the number of samples that had that peak. This gives where T is the set of all tissues, is the

CPM at genomic region i, and is a counts threshold, here taken to be 0. Peaks are then ranked according to their value:

$$s_i = |\{x : LCPM_{x,i} > \theta\}| * \text{median}(\{LCPM_{t,i} : t \in T, LCPM_{t,i} > \theta\})$$

Where s_i is the activation value for that peak in region i, LCPMx,i is the log CPM in tissue X in genomic region i, and theta is the threshold, currently taken as 0 (of course these peaks are significant, as described before). The first term on the right-hand side is the number of tissues that have a peak in this region, and the second term is the median value amongst those tissues.

Peak Location

When different samples are merged together, sometimes peaks do not land on the same base-pair, either due to noise or having multiple peaks within the 150 base-pair window used to call peaks. Although all of these are nearby each other, we opt to report the location from one of the samples to avoid any nucleotide shifting. When reporting the final location we use one of two options: if the peak was seen in CHO, we report the CHO location, otherwise, we report the location of the peak in the sample which had the most CPM.

Supplementary Tables and Figures

Supplementary Table S1: TPM Expression Summary from in-house CHO-S Transcriptomic Samples

Gene	**Aldh18a1**	**Mgat3**	**St6gal1**
Cell type (# of samples)	CHO-S (40)	CHO-S (40)	CHO-S (40)
minimum:	0.019488	0.0	0.0
25th percentile:	0.100172725	0.0	0.0
50th percentile:	0.195765	0.0	0.0
75th percentile:	0.23229675	0.0	0.0
maximum:	0.443554	0.0486695	0.0

Supplementary Table S2: gRNA Target Sequences and Oligos

Target	Assembly	gRNA#	Target sequence (5'-3', PAM indicated in bold font)
St6gal1	NCBI	gRNA_1	CATGGTGCATAGACTAAGCG**GGG**
St6gal1	NCBI	gRNA_2	GGAAAGATGCATAAACGTTC**AGG**
St6gal1	NCBI	gRNA_3	ACAGAGGACTTCCGAGAATC**TGG**
St6gal1	CHOoptPICR	gRNA_1	CTGGTGAAAACGCTCGTACA**AGG**
St6gal1	CHOoptPICR	gRNA_2	CAGAGAAAAGTTACTGATCG**CGG**
St6gal1	CHOoptPICR	gRNA_3	CCTCTACTGTCCCGAAGTAT**TGG**
Aldh18a1	NCBI	gRNA_1	ACAAAACTGAAATGTCGGCT**TGG**

Aldh18a1	NCBI	gRNA_2	TAGAGAGTCCTTAACAGGAC**AGG**
Aldh18a1	NCBI	gRNA_3	CAGCAGGCTGCTATGAAGGT**TGG**
Aldh18a1	CHOoptPICR	gRNA_1	GTCCGACTCGAGCTACGCCG**CGG**
Aldh18a1	CHOoptPICR	gRNA_2	AAGCGCCTGCGCCGTTGCCG**CGG**
Aldh18a1	CHOoptPICR	gRNA_3	GACCCGTCGCCTAACGTCAG**CGG**
Mgat3	NCBI	gRNA_1	AGGTGTGGTTACCACGCCCA**GGG**
Mgat3	NCBI	gRNA_2	TGTGATGGGGTCTCGGGTAT**GGG**
Mgat3	NCBI	gRNA_3	AACAGGTCAGGGCAAGCCGG**AGG**
Mgat3	CHOoptPICR	gRNA_1	CCTAGCTCTAGAAGCCGTCT**TGG**
Mgat3	CHOoptPICR	gRNA_2	CATATCAAACTCCCACTAGC**AGG**
Mgat3	CHOoptPICR	gRNA_3	ATAAAGTCCAAGCATGCTAG**AGG**

Supplementary Table S3: Oligos Used for gRNA Vector Cloning

Target	Assembly	gRNA#	Sequence
St6gal1	NCBI	gRNA_1_fwd	GGAAAGGACGAAACACCGATGGTGCATAGACTAAGCGGTTTT AGAGCTAGAAAT
St6gal1	NCBI	gRNA_1_rev	CTAAAACCGCTTAGTCTATGCACCATCGGTGTTTCGTCCTTTCC ACAAGATAT
St6gal1	NCBI	gRNA_2_fwd	GGAAAGGACGAAACACCGGAAAGATGCATAAACGTTCGTTTT AGAGCTAGAAAT
St6gal1	NCBI	gRNA_2_rev	CTAAAACGAACGTTTATGCATCTTTCCGGTGTTTCGTCCTTTCC ACAAGATAT

St6gal1	NCBI	gRNA_3_fwd	GGAAAGGACGAAACACCGCAGAGGACTTCCGAGAATCGTTTT AGAGCTAGAAAT
St6gal1	NCBI	gRNA_3_rev	CTAAAACGATTCTCGGAAGTCCTCTGCGGTGTTTCGTCCTTTCC ACAAGATAT
Aldh18a1	NCBI	gRNA_1_fwd	GGAAAGGACGAAACACCGCAAAACTGAAATGTCGGCTGTTTT AGAGCTAGAAAT
Aldh18a1	NCBI	gRNA_1_rev	CTAAAACAGCCGACATTTCAGTTTTGCGGTGTTTCGTCCTTTCC ACAAGATAT
Aldh18a1	NCBI	gRNA_2_fwd	GGAAAGGACGAAACACCGAGAGAGTCCTTAACAGGACGTTTT AGAGCTAGAAAT
Aldh18a1	NCBI	gRNA_2_rev	CTAAAACGTCCTGTTAAGGACTCTCTCGGTGTTTCGTCCTTTCC ACAAGATAT
Aldh18a1	NCBI	gRNA_3_fwd	GGAAAGGACGAAACACCGAGCAGGCTGCTATGAAGGTGTTTT AGAGCTAGAAAT
Aldh18a1	NCBI	gRNA_3_rev	CTAAAACACCTTCATAGCAGCCTGCTCGGTGTTTCGTCCTTTCC ACAAGATAT
Mgat3	NCBI	gRNA_1_fwd	GGAAAGGACGAAACACCGGGTGTGGTTACCACGCCCAGTTTT AGAGCTAGAAAT
Mgat3	NCBI	gRNA_1_rev	CTAAAACTGGGCGTGGTAACCACACCCGGTGTTTCGTCCTTTC CACAAGATAT
Mgat3	NCBI	gRNA_2_fwd	GGAAAGGACGAAACACCGGTGATGGGGTCTCGGGTATGTTTT AGAGCTAGAAAT
Mgat3	NCBI	gRNA_2_rev	CTAAAACATACCCGAGACCCCATCACCGGTGTTTCGTCCTTTCC ACAAGATAT
Mgat3	NCBI	gRNA_3_fwd	GGAAAGGACGAAACACCGACAGGTCAGGGCAAGCCGGGTTTT AGAGCTAGAAAT
Mgat3	NCBI	gRNA_3_rev	CTAAAACCCGGCTTGCCCTGACCTGTCGGTGTTTCGTCCTTTCC ACAAGATAT
St6gal1	CHOoptPICR	gRNA_1_fwd	GGAAAGGACGAAACACCGTGGTGAAAACGCTCGTACAGTTTT AGAGCTAGAAAT
St6gal1	CHOoptPICR	gRNA_1_rev	CTAAAACTGTACGAGCGTTTTCACCACGGTGTTTCGTCCTTTCC ACAAGATAT
St6gal1	CHOoptPICR	gRNA_2_fwd	GGAAAGGACGAAACACCGAGAGAAAAGTTACTGATCGGTTTT AGAGCTAGAAAT
St6gal1	CHOoptPICR	gRNA_2_rev	CTAAAACCGATCAGTAACTTTTCTCTCGGTGTTTCGTCCTTTCC ACAAGATAT
St6gal1	CHOoptPICR	gRNA_3_fwd	GGAAAGGACGAAACACCGCTCTACTGTCCCGAAGTATGTTTTA GAGCTAGAAAT

St6gal1	CHOoptPICR	gRNA_3_rev	CTAAAACATACTTCGGGACAGTAGAGCGGTGTTTCGTCCTTTC CACAAGATAT
Aldh18a1	CHOoptPICR	gRNA_1_fwd	GGAAAGGACGAAACACCGTCCGACTCGAGCTACGCCGGTTTTA GAGCTAGAAAT
Aldh18a1	CHOoptPICR	gRNA_1_rev	CTAAAACCGGCGTAGCTCGAGTCGGACGGTGTTTCGTCCTTTC CACAAGATAT
Aldh18a1	CHOoptPICR	gRNA_2_fwd	GGAAAGGACGAAACACCGAGCGCCTGCGCCGTTGCCGGTTTT AGAGCTAGAAAT
Aldh18a1	CHOoptPICR	gRNA_2_rev	CTAAAACCGGCAACGGCGCAGGCGCTCGGTGTTTCGTCCTTTC CACAAGATAT
Aldh18a1	CHOoptPICR	gRNA_3_fwd	GGAAAGGACGAAACACCGACCCGTCGCCTAACGTCAGGTTTTA GAGCTAGAAAT
Aldh18a1	CHOoptPICR	gRNA_3_rev	CTAAAACCTGACGTTAGGCGACGGGTCGGTGTTTCGTCCTTTC CACAAGATAT
Mgat3	CHOoptPICR	gRNA_1_fwd	GGAAAGGACGAAACACCGCTAGCTCTAGAAGCCGTCTGTTTTA GAGCTAGAAAT
Mgat3	CHOoptPICR	gRNA_1_rev	CTAAAACAGACGGCTTCTAGAGCTAGCGGTGTTTCGTCCTTTC CACAAGATAT
Mgat3	CHOoptPICR	gRNA_2_fwd	GGAAAGGACGAAACACCGATATCAAACTCCCACTAGCGTTTTA GAGCTAGAAAT
Mgat3	CHOoptPICR	gRNA_2_rev	CTAAAACGCTAGTGGGAGTTTGATATCGGTGTTTCGTCCTTTC CACAAGATAT
Mgat3	CHOoptPICR	gRNA_3_fwd	GGAAAGGACGAAACACCGTAAAGTCCAAGCATGCTAGGTTTT AGAGCTAGAAAT
Mgat3	CHOoptPICR	gRNA_3_rev	CTAAAACCTAGCATGCTTGGACTTTACGGTGTTTCGTCCTTTCC ACAAGATAT

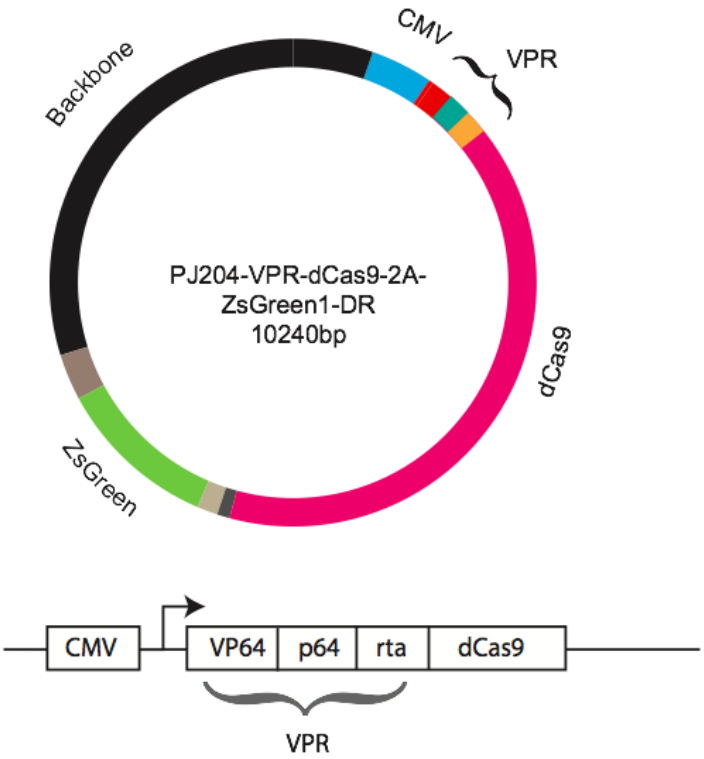
Supplementary Table S4: qRT-PCR Custom Assays

Target	Fwd primer	Rev primer	Probe
Mgat3	GAGACGCTACAAGCTCTTTCT	GGGAAGGTGACATAGGAT AAGG	TGTAAGAAGGAGATGA GGCACAGGC
Gnb1	CCATATGTTTCTTTCCCAATGGC	AAGTCGTCGTACCCAGCA AG	ACTGGTTCAGACGATGC TACGTGC

Supplementary Table S5: The amount of gRNA and VPR-dCas9 used for transfection. A total amount of 3.75 μg DNA was used for each transfection.

Transfection	gRNA	VPR-dCas9
Co-transfection of a single gRNA and VPR-dCas9	1.875 μg	1.875 μg
Co-transfection of a mix of 3 gRNAs and VPR-dCas9	0.625 μg/gRNA	1.875 μg

Supplementary Figure S1: VPR-dCas9 Construct and Sequence



VPR-dCas9 sequence (CMV-VP64-p65-rta-dCas9)

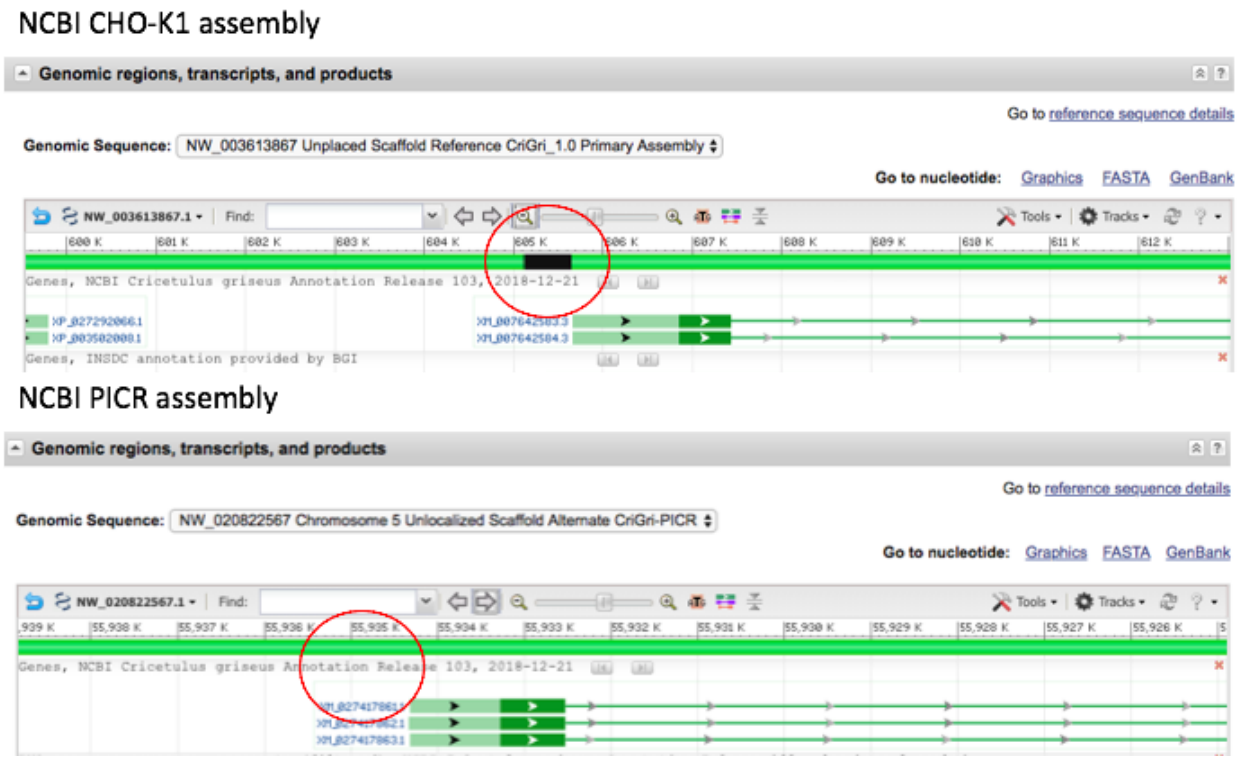
AGTCGGTGTGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAG
TTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGAC
CGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC
AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGC
ATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT
CGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACT
CACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATC
AACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGT
GTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTG
GCTTATCGAAATAAGCAGCGTCGCCACCATGGAGGCCAGCGGTTCCGGACGGGCTGACGCAT
TGGACGATTTTGATCTGGATATGCTGGGAAGTGACGCCCTCGATGATTTTGACCTTGACATG
CTTGGTTCGGATGCCCTTGATGACTTTGACCTCGACATGCTCGGCAGTGACGCCCTTGATGAT
TTCGACCTGGACATGCTGATTAACTCTAGAAGTTCCGGATCTCCGAAAAAGAAACGCAAAGTT
GGTAGCCAGTACCTGCCCGACACCGACGACCGGCACCGGATCGAGGAAAAGCGGAAGCGGA
CCTACGAGACATTCAAGAGCATCATGAAGAAGTCCCCCTTCAGCGGCCCCACCGACCCTAGAC
CTCCACCTAGAAGAATCGCCGTGCCCAGCAGATCCAGCGCCAGCGTGCCAAAACCTGCCCCCC
AGCCTTACCCCTTCACCAGCAGCCTGAGCACCATCAACTACGACGAGTTCCCTACCATGGTGTT
CCCCAGCGGCCAGATCTCTCAGGCCTCTGCTCTGGCTCCAGCCCCTCCTCAGGTGCTGCCTCA

GGCTCCTGCTCCTGCACCAGCTCCAGCCATGGTGTCTGCACTGGCTCAGGCACCAGCACCCGT
GCCTGTGCTGGCTCCTGGACCTCCACAGGCTGTGGCTCCACCAGCCCCTAAACCTACACAGGC
CGGCGAGGGCACACTGTCTGAAGCTCTGCTGCAGCTGCAGTTCGACGACGAGGATCTGGGA
GCCCTGCTGGGAAACAGCACCGATCCTGCCGTGTTCACCGACCTGGCCAGCGTGGACAACAG
CGAGTTCCAGCAGCTGCTGAACCAGGGCATCCCTGTGGCCCCTCACACCACCGAGCCCATGCT
GATGGAATACCCCGAGGCCATCACCCGGCTCGTGACAGGCGCTCAGAGGCCTCCTGATCCAG
CTCCTGCCCCTCTGGGAGCACCAGGCCTGCCTAATGGACTGCTGTCTGGCGACGAGGACTTC
AGCTCTATCGCCGATATGGATTTCTCAGCCTTGCTGGGCTCTGGCAGCGGCAGCCGGGATTC
CAGGGAAGGGATGTTTTTGCCGAAGCCTGAGGCCGGCTCCGCTATTAGTGACGTGTTTGAG
GGCCGCGAGGTGTGCCAGCCAAAACGAATCCGGCCATTTCATCCTCCAGGAAGTCCATGGGC
CAACCGCCCACTCCCCGCCAGCCTCGCACCAACACCAACCGGTCCAGTACATGAGCCAGTCGG
GTCACTGACCCCGGCACCAGTCCCTCAGCCACTGGATCCAGCGCCCGCAGTGACTCCCGAGGC
CAGTCACCTGTTGGAGGATCCCGATGAAGAGACGAGCCAGGCTGTCAAAGCCCTTCGGGAG
ATGGCCGATACTGTGATTCCCCAGAAGGAAGAGGCTGCAATCTGTGGCCAAATGGACCTTTC
CCATCCGCCCCCAAGGGGCCATCTGGATGAGCTGACAACCACACTTGAGTCCATGACCGAGG
ATCTGAACCTGGACTCACCCCTGACCCCGGAATTGAACGAGATTCTGGATACCTTCCTGAACG
ACGAGTGCCTCTTGCATGCCATGCATATCAGCACAGGACTGTCCATCTTCGACACATCTCTGTT
TAAAGCCTATGACAAGAAATACTCCATCGGCCTTGCCATCGGGACCAACAGCGTGGGTTGGG
CCGTGATCACTGACGAATACAAAGTGCCATCGAAAAAGTTTAAGGTGCTGGGCAATACTGAC
AGACATAGCATCAAGAAAAACCTGATTGGAGCCCTCCTGTTCGACTCCGGGGAAACCGCCGA
AGCCACTCGCCTGAAACGGACCGCACGCCGGAGGTACACTCGCCGGAAGAATAGAATCTGCT
ACCTTCAAGAGATCTTTAGCAACGAGATGGCCAAAGTGGACGATTCGTTCTTCCACCGGCTG
GAGGAGTCATTCCTGGTCGAAGAGGACAAAAAGCATGAGAGACATCCGATCTTCGGTAACAT
TGTGGATGAGGTCGCCTACCACGAGAAGTACCCCACTATCTACCATTTGCGCAAGAAGTTGG
TCGACTCGACTGATAAGGCCGACCTCAGACTCATCTACCTCGCTTTGGCACACATGATCAAGT
TTAGGGGTCACTTCCTGATTGAAGGGGATCTGAACCCAGACAACTCGGATGTGGATAAACTG
TTTATCCAGCTGGTCCAGACTTACAATCAGCTCTTTGAGGAGAACCCGATCAATGCCTCCGGC
GTGGATGCTAAGGCAATCCTGTCCGCACGCTTGTCAAAGAGCCGGCGGCTCGAAAATCTGAT
CGCACAGCTTCCGGGAGAAAAGAAAAACGGACTGTTTGGAAACCTTATCGCCCTCTCGCTGG
GACTCACTCCTAACTTCAAATCCAATTTTGATTTGGCCGAGGATGCCAAGCTGCAGCTGTCGA
AGGACACCTACGACGATGACCTCGACAATCTTCTGGCCCAGATCGGGGACCAGTACGCCGAT
CTGTTCCTCGCGGCCAAAAACCTGTCAGACGCAATCTTGCTGAGCGATATCTTGAGGGTCAAT
ACCGAAATCACTAAGGCTCCATTGTCAGCATCGATGATCAAGAGGTACGATGAACACCATCAG
GACCTGACCCTGCTCAAGGCACTGGTGCGGCAGCAACTGCCGGAAAAGTACAAGGAGATCTT
CTTCGATCAAAGCAAGAATGGGTACGCAGGGTACATCGATGGTGGAGCATCACAAGAAGAG
TTCTACAAATTCATCAAGCCCATCCTTGAAAAGATGGACGGGACGGAAGAACTGCTGGTGAA
GCTGAATCGGGAAGATCTGCTGCGGAAGCAGCGCACCTTCGACAATGGTTCCATCCCACACC
AAATCCATCTCGGCGAACTGCACGCGATCCTTCGCCGGCAGGAAGATTTCTACCCGTTCTTGA
AGGATAACAGAGAAAAGATCGAGAAAATCCTGACCTTTAGAATCCCGTACTACGTGGGCCCG
TTGGCTCGCGGAAACTCAAGATTCGCCTGGATGACTAGAAAATCCGAAGAAACGATTACCCC
GTGGAACTTTGAAGAGGTCGTCGACAAGGGAGCCTCGGCACAGTCGTTCATCGAACGGATG
ACCAATTTCGACAAGAACCTCCCCAACGAAAAGGTGCTCCCGAAACACTCGCTGCTGTATGAG
TACTTCACGGTCTACAACGAGCTGACCAAGGTGAAGTACGTGACCGAAGGAATGCGGAAGCC
CGCTTTTCTGTCCGGTGAACAGAAAAAGGCCATCGTGGACCTCCTCTTCAAGACTAACAGAAA
GGTGACCGTGAAACAGCTTAAAGAGGACTACTTCAAGAAGATCGAGTGTTTTGACTCCGTGG
AAATCTCGGGAGTGGAGGATAGGTTTAACGCTTCGCTGGGGACCTACCATGACCTTCTCAAG
ATCATCAAGGATAAGGATTTCCTGGACAACGAAGAAAACGAGGACATCCTTGAGGACATCGT
GTTGACCTTGACCCTGTTTGAGGATCGGGAAATGATCGAGGAGAGACTGAAAACTTACGCGC
ATCTTTTCGACGACAAAGTGATGAAGCAGCTGAAAAGAAGAAGATACACTGGCTGGGGACG
GCTGTCGAGGAAACTGATTAACGGAATTAGGGATAAACAAAGCGGAAAGACTATCCTCGATT
TTCTCAAGTCGGACGGCTTCGCCAATCGGAACTTTATGCAGCTCATCCACGACGATAGCCTGA
CCTTCAAGGAAGATATCCAAAAAGCCCAAGTGTCCGGCCAGGGAGATAGCTTGCACGAGCAC

ATTGCTAACCTGGCCGGTTCACCAGCCATTAAGAAGGGAATCCTCCAAACCGTGAAGGTGGT
CGACGAACTCGTCAAGGTGATGGGCCGCCATAAGCCAGAGAACATCGTGATCGAGATGGCTC
GCGAAAATCAGACTACTCAGAAGGGTCAAAAGAACTCCCGCGAGCGCATGAAGCGCATCGAA
GAAGGAATTAAGGAGCTGGGCTCACAAATCCTCAAAGAACATCCTGTCGAGAACACCCAACTT
CAGAACGAGAAACTGTACCTCTACTATCTCCAAAATGGCCGGGACATGTACGTCGATCAGGA
ATTGGATATCAACCGCCTCTCAGACTACGACGTGGATGCTATCGTGCCGCAGTCGTTTCTCAA
AGATGATAGCATCGACAACAAGGTCCTGACGAGGTCGGATAAGAACCGCGGGAAATCCGAC
AATGTGCCTTCGGAAGAGGTGGTGAAAAAGATGAAGAACTACTGGAGGCAATTGCTTAATG
CCAAACTGATCACCCAGCGCAAATTCGACAACCTTACTAAGGCCGAGCGCGGAGGACTGTCC
GAACTCGATAAGGCGGGGTTCATCAAAAGACAATTGGTCGAAACTCGGCAAATTACCAAGCA
TGTCGCTCAGATCCTCGACTCGCGCATGAACACTAAGTATGATGAGAACGACAAGCTCATTCG
CGAAGTGAAAGTGATTACCCTCAAATCAAAGCTGGTGTCGGACTTCAGAAAAGACTTTCAATT
CTACAAGGTGCGGGAGATTAACAACTACCACCACGCGCACGACGCCTACCTCAATGCAGTGG
TGGGAACCGCCCTCATCAAGAAGTATCCAAAGCTCGAAAGCGAGTTCGTGTACGGGGACTAT
AAAGTGTACGATGTGCGGAAAATGATCGCAAAGAGCGAGCAAGAGATCGGCAAAGCAACGG
CCAAATACTTCTTCTACAGCAACATTATGAACTTTTTCAAGACCGAGATTACCCTGGCGAACGG
AGAAATCCGGAAGCGGCCTTTGATCGAAACGAATGGGGAAACTGGAGAAATCGTGTGGGAC
AAGGGCAGAGACTTTGCGACCGTGAGAAAGGTGCTGTCCATGCCACAAGTCAACATCGTGAA
GAAAACTGAAGTGCAGACTGGAGGATTTTCCAAGGAGTCAATCCTTCCGAAGCGCAACAGCG
ACAAGCTGATCGCCAGGAAGAAGGATTGGGACCCCAAGAAGTACGGTGGTTTTGATTCACCT
ACTGTCGCTTACAGCGTGCTGGTGGTGGCCAAGGTGGAAAAGGGGAAGTCAAAGAAATTGA
AGTCGGTGAAGGAATTGCTGGGGATTACTATCATGGAGAGGAGCAGCTTCGAAAAGAATCC
CATCGACTTCTTGGAGGCCAAGGGATACAAAGAGGTGAAGAAAGACCTGATCATTAAGCTGC
CGAAGTACTCCTTGTTCGAACTGGAAAACGGCAGAAAGCGGATGCTGGCCTCCGCCGGAGA
GCTGCAGAAGGGTAACGAGCTGGCTCTGCCCAGCAAATACGTGAATTTCCTCTACCTGGCCTC
CCATTACGAGAAGCTCAAGGGAAGCCCGGAGGATAATGAGCAAAAACAACTGTTCGTCGAGC
AGCACAAGCACTACCTCGACGAGATCATCGAACAAATCTCCGAGTTCTCGAAGCGGGTGATTC
TGGCCGACGCAAACCTTGATAAAGTCCTCAGCGCCTACAACAAGCACCGCGACAAACCGATCA
GAGAACAAGCTGAGAACATCATCCACCTGTTCACGCTGACTAACTTGGGAGCGCCTGCCGCCT
TCAAGTACTTCGACACCACTATCGATCGGAAACGGTACACTTCCACTAAGGAAGTGCTGGACG
CAACCCTGATCCATCAGTCCATCACCGGACTTTATGAGACTCGGATCGATCTCAGCCAGCTTG
GCGGTGATTCAAGAGCTGACCCGAAGAAGAAGCGCAAGGTCAGACGTCATGGCAGTGGAGA
GGGCAGAGGAAGTCTGCTAACATGCGGTGACGTCGAGGAGAATCCTGGCCCAATGGCCCAG
TCCAAGCACGGCCTGACCAAGGAGATGACCATGAAGTACCGCATGGAGGGCTGCGTGGACG
GCCACAAGTTCGTGATCACCGGCGAGGGCATCGGCTACCCCTTCAAGGGCAAGCAGGCCATC
AACCTGTGCGTGGTGGAGGGCGGCCCCTTGCCCTTCGCCGAGGACATCTTGTCCGCCGCCTT
CATGTACGGCAACCGCGTGTTCACCGAGTACCCCCAGGACATCGTCGACTACTTCAAGAACTC
CTGCCCCGCCGGCTACACCTGGGACCGCTCCTTCCTGTTCGAGGACGGCGCCGTGTGCATCT
GCAACGCCGACATCACCGTGAGCGTGGAGGAGAACTGCATGTACCACGAGTCCAAGTTCTAC
GGCGTGAACTTCCCCGCCGACGGCCCCGTGATGAAGAAGATGACCGACAACTGGGAGCCCTC
CTGCGAGAAGATCATCCCCGTGCCCAAGCAGGGCATCTTGAAGGGCGACGTGAGCATGTACC
TGCTGCTGAAGGACGGTGGCCGCTTGCGCTGCCAGTTCGACACCGTGTACAAGGCCAAGTCC
GTGCCCCGCAAGATGCCCGACTGGCACTTCATCCAGCACAAGCTGACCCGCGAGGACCGCAG
CGACGCCAAGAACCAGAAGTGGCACCTGACCGAGCACGCCATCGCCTCCGGCTCCGCCTTGC
CCAAGCTTCCGCGGAGCCATGGCTTCCCGCCGGCGGTGGCGGCGCAGGATGATGGCACGCT
GCCCATGTCTTGTGCCCAGGAGAGCGGGATGGACCGTCACCCTGCAGCCTGTGCTTCTGCTA
GGATCAATGTGTAGACACAGTCTCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTC
CCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAA
ATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACA
GCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGG
ACTTGCGTAGTGAGTCGAATAAGGGCGACACAAAATTTATTCTAAATGCATAATAAATACTGA
TAACATCTTATAGTTTGTATTATATTTTGTATTATCGTTGACATGTATAATTTTGATATCAAAAA

CTGATTTTCCCTTTATTATTTTCGAGATTTATTTTCTTAATTCTCTTTAACAAACTAGAAATATT
GTATATACAAAAAATCATAAATAATAGATGAATAGTTTAATTATAGGTGTTCATCAATCGAAA
AAGCAACGTATCTTATTTAAAGTGCGTTGCTTTTTTCTCATTTATAAGGTTAAATAATTCTCAT
ATATCAAGCAAAGTGACAGGCGCCCTTAAATATTCTGACAAATGCTCTTTCCCTAAACTCCCCC
CATAAAAAAACCCGCCGAAGCGGGTTTTTACGTTATTTGCGGATTAACGATTACTCGTTATCA
GAACCGCCCAGGGGGCCCGAGCTTAAGACTGGCCGTCGTTTTACAACACAGAAAGAGTTTGT
AGAAACGCAAAAAGGCCATCCGTCAGGGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTCCCTA
CTCTCGCCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCG
GTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAA
GAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCG
TTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGG
CGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT
CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCG
CTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGC
TGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG
TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAG
AGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGGCTAACTACGGCTACACT
AGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG
TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCA
GATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGC
TCAGTGGAACGACGCGCGCGTAACTCACGTTAAGGGATTTTGGTCATGAGCTTGCGCCGTCC
CGTCAAGTCAGCGTAATGCTCTGCTTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCG
ATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGG
AGGGCTTACCATCTGGCCCCAGCGCTGCGATGATACCGCGAGAACCACGCTCACCGGCTCCG
GATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTT
ATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAA
TAGTTTGCGCAACGTTGTTGCCATCGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTAT
GGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAA
AAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATC
ACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCT
GTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTC
TTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT
TGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGAT
GTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGA
GCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAA
TACTCATATTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGG
ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTCAGTGTTACAACCAATTAACC
AATTCTGAACATTATCGCGAGCCCATTTATACCTGAATATGGCTCATAACACCCCTTGTTTGCC
TGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTA
GCGCCGATGGTAGTGTGGGGACTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAA
AACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGCCCGGGCTAATTATGGGGTGTCGCCCTT
ATTCGACTC

Supplementary Figure S2: *Bcl2* is an example for which the region upstream of the gene has an N filled gap (black region) in the NCBI CHO-K1 assembly and no gap in the PICR assembly (red circles). Screenshots from NCBI.



Supplementary Information for Chapter IV

Supplementary Results

Determining Multiplicity of Infection in Suspension CHO Cells

When creating a CRISPR-Cas9 library using lentiviruses, the initial number of infections per cell, and thus number of gRNA per cell, follows a Poisson distribution[166]. That distribution depends on the ratio of viruses to cells. This is known as the "multiplicity of infection" or MOI. In CRISPR-Cas9 screens it is important to transduce at a low MOI to decrease the number gRNA integration events per cell. If a cell carries multiple different gRNAs the genotype to phenotype relation becomes complex to analyze. This confounder is alleviated by having multiple different gRNAs targeting each gene. However, a low MOI also increases the number of cells needed for 100% library coverage which can be a practical limitation. Most CRISPR-Cas9 screens report transduction with an MOI of 0.2-0.5[45,60,64,112,126,167–176]. Though a reason is not always given, one can speculate that the reason is that at an MOI of ~0.3, the Poisson distribution probability mass function (Poisson pmf) predicts that the fraction of cells having more than 1 gRNA should be:

$$p(k) = \frac{\lambda^k e^{-\lambda}}{k!} \Rightarrow p(k > 1) = 1 - p(k = 0) - p(k = 1) = 1 - e^{-\lambda} - \lambda e^{-\lambda}$$

where p(k) is the fraction of cells with k infections and lambda is the MOI. So with MOI = 0.3, the fraction of cells with more than 1 infection, p(k>1), will most likely be ~3.7%, which presumably is then considered sufficiently low to prevent analysis issues. However previous studies often seem to lack any description as to how they actually obtain this MOI and/or measure it[49,61,112,113,167,168,171–173,175–180]. Furthermore, if they obtain/measure this MOI before antibiotic selection against uninfected cells (wild type cells), then the fraction of cells with more than 1 insertion in the final library will be much higher than 3.7% because all the uninfected cells are ideally gone. With the uninfected cells (k=0) gone, the formula changes to:

$$p_{after}(k) = \frac{\left(\frac{\lambda^k e^{-\lambda}}{k!}\right)}{p(k>0)} \Rightarrow \frac{\left(\frac{\lambda^k e^{-\lambda}}{k!}\right)}{1-e^{-\lambda}} \Rightarrow \frac{\lambda^k}{(e^{\lambda}-1)k!}$$

where p_{after} is the fraction of cells with k infections **after** antibiotic selection and k is larger than 0. So if you had an MOI of 0.3 before selection, then the fraction of cells with more than 1 infection after antibiotic selection, $p_{after}(k>1)$, would most likely be ~14%. We were only able to identify 1 other study that explicitly states this[67] and only 1 study that have selected an MOI that would lead to $p_{after}(k>1)$ being ~2.5%[113].

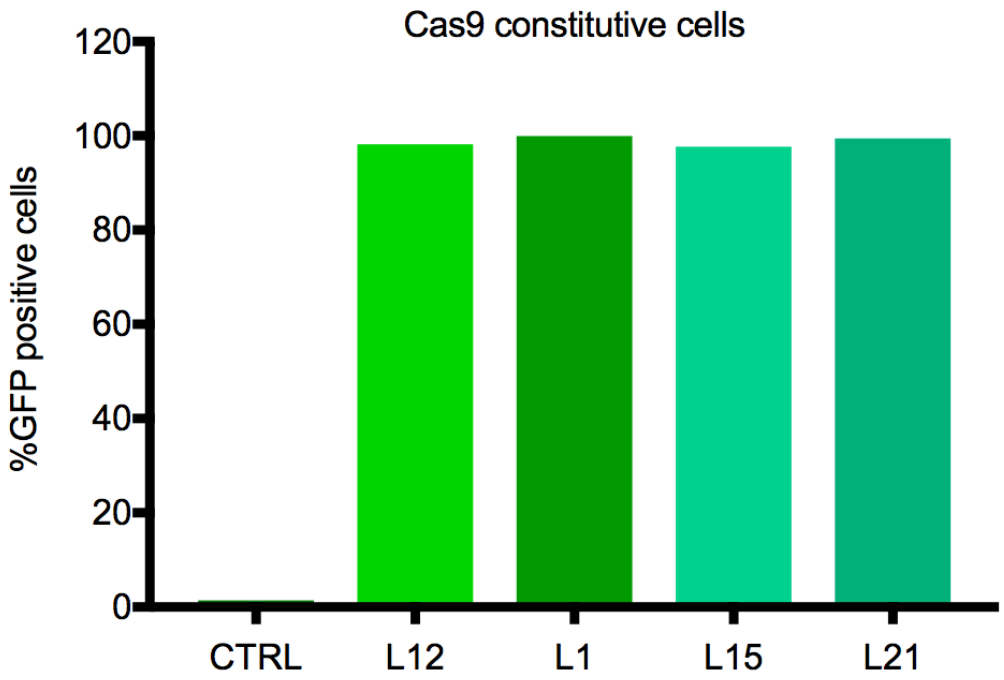
In this study we decided to aim for the lowest MOI possible where we could practically handle the number of cells needed for ~100% coverage of the library. Most of the previously referenced CRISPR screen studies estimate MOI by infecting a cell population with their lentiviral library and then counting the number of cells surviving with vs without an antibiotic selection. In such a study you are essentially counting the number of uninfected cells and assuming that the number of gRNAs per cell will otherwise follow the Poisson distribution. We speculate that it is possible that cells with multiple infections, and thus multiple antibiotic resistance genes, might survive better and thus perturb the assumed distribution. For this reason, we decided to actually count the number of different gRNAs per cell after antibiotic selection.

We set up an experiment identical to the process used in the actual library creation but with four different volumes of the viral library, 4, 10, 20 and 40 μL. After antibiotic selection, we single cell sorted the four resulting libraries, harvested gDNA from each clonal population and sequenced the gRNA insert and a control gene, OMA1. This allowed us to count how many different gRNA we had in each cell. We could then directly count how many percent of the cells had more than 1 gRNA and fitting this data to the $p_{after}(k)$ function, we could estimate lamda (pre-selection MOI) and estimate confidence intervals:

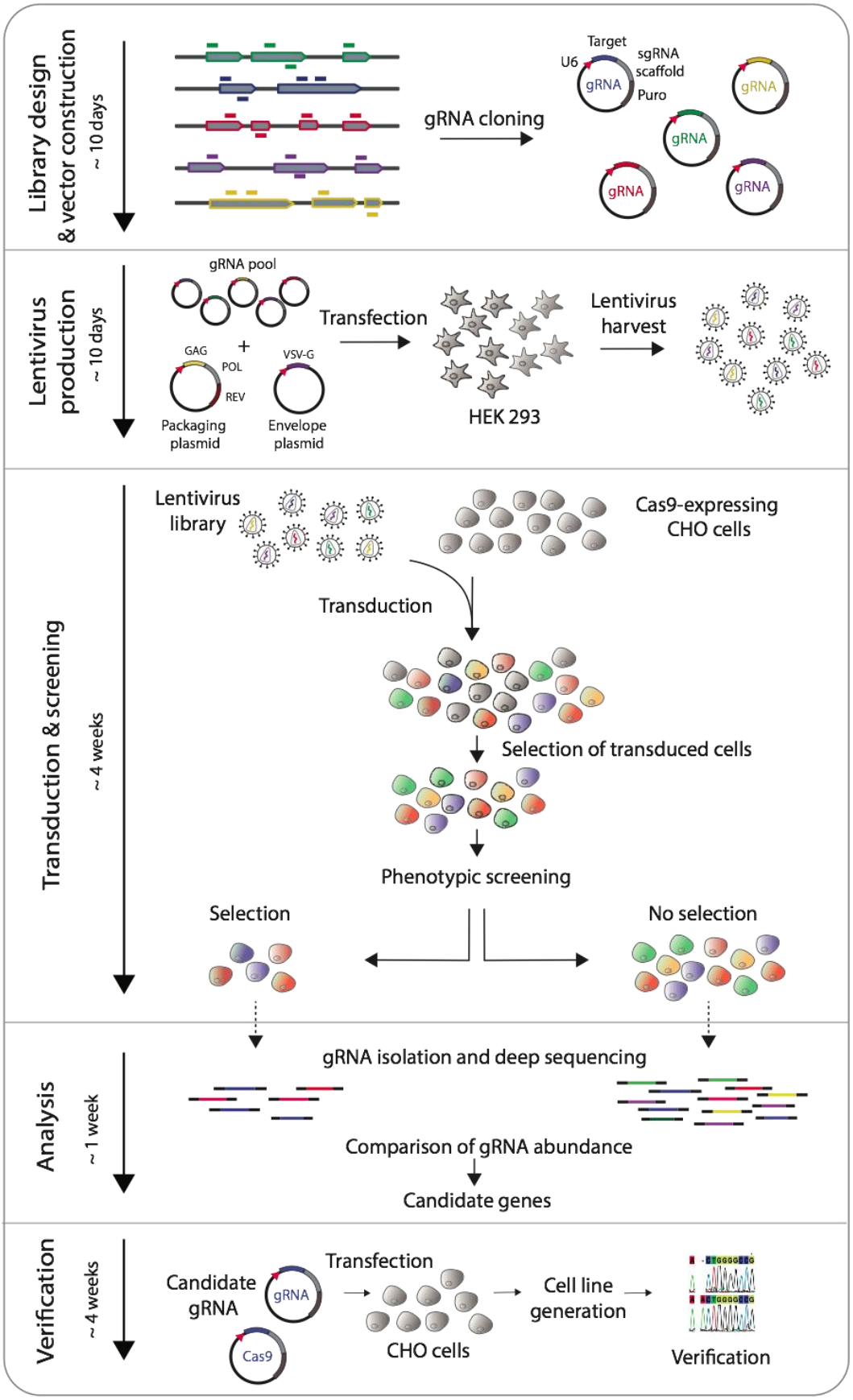
μL virus	Percent having more than 1 gRNA (lower/upper 95% conf interval)	Fitted lamda (moi) (lower/upper 95% conf interval)
4	11.8% [7.3%, 17.0%]	0.25 [0.15, 0.36]
10	30.1% [22.9%, 37.3%]	0.68 [0.5, 0.87]
20	35.3% [28.2%, 42.2%]	0.81 [0.63, 1.01]
40	56.8% [49.4%, 63.8%]	1.50 [1.24, 1.78]

Based on above results, we chose to use 4 μL virus per 0.6 x 10^6 cells/mL for transduction.

Supplementary Figures

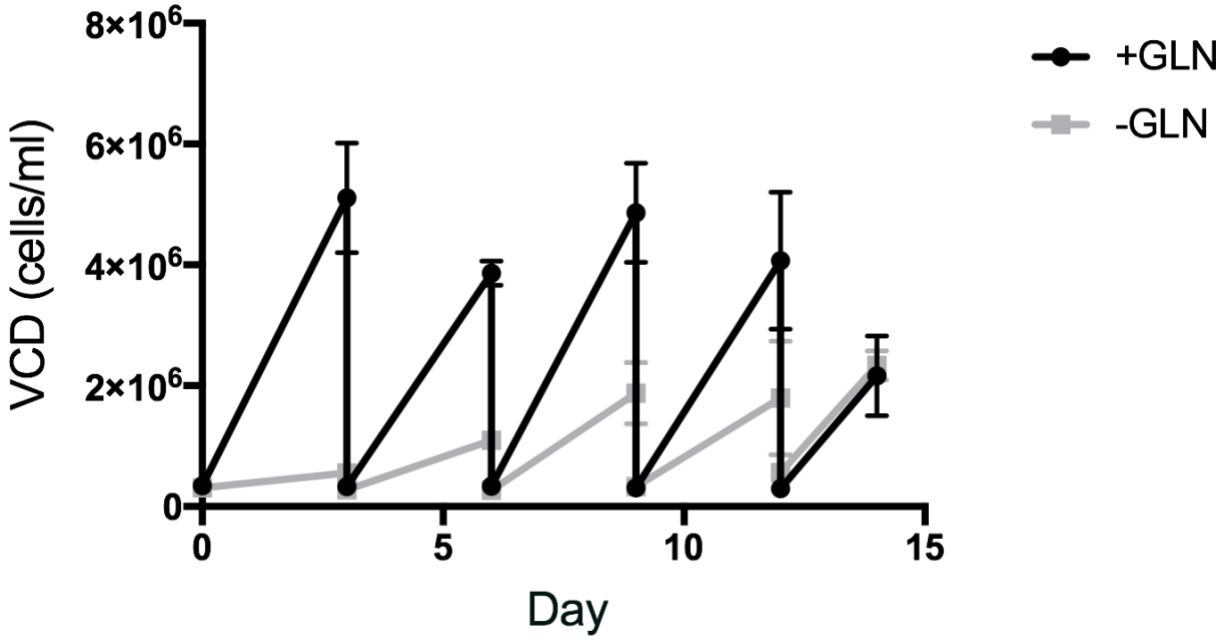


Supplementary Figure S1: CHO-S^{Cas9} Validation by GFP Fluorescence Analysis.
CHO-S^{Cas9} cells were seeded in wells of a 96 well plate and GFP fluorescence intensity was measured using Celigo. All 5 clones tested positive for GFP expression against CHO-S WT (CTRL). CHO-S^{Cas9} clone L12 was picked for use in the CRISPR KO screens.


Library design & vector construction
~ 10 days
gRNA cloning
Target
U6
gRNA
sgRNA scaffold
Puro
Lentivirus production
~ 10 days
gRNA pool
GAG
POL
REV
Packaging plasmid
VSV-G
Envelope plasmid
Transfection
HEK 293
Lentivirus harvest
Transduction & screening
~ 4 weeks
Lentivirus library
Cas9-expressing CHO cells
Transduction
Selection of transduced cells
Phenotypic screening
Selection
No selection
Analysis
~ 1 week
gRNA isolation and deep sequencing
Comparison of gRNA abundance
Candidate genes
Verification
~ 4 weeks
Candidate gRNA
Cas9
Transfection
CHO cells
Cell line generation
Verification

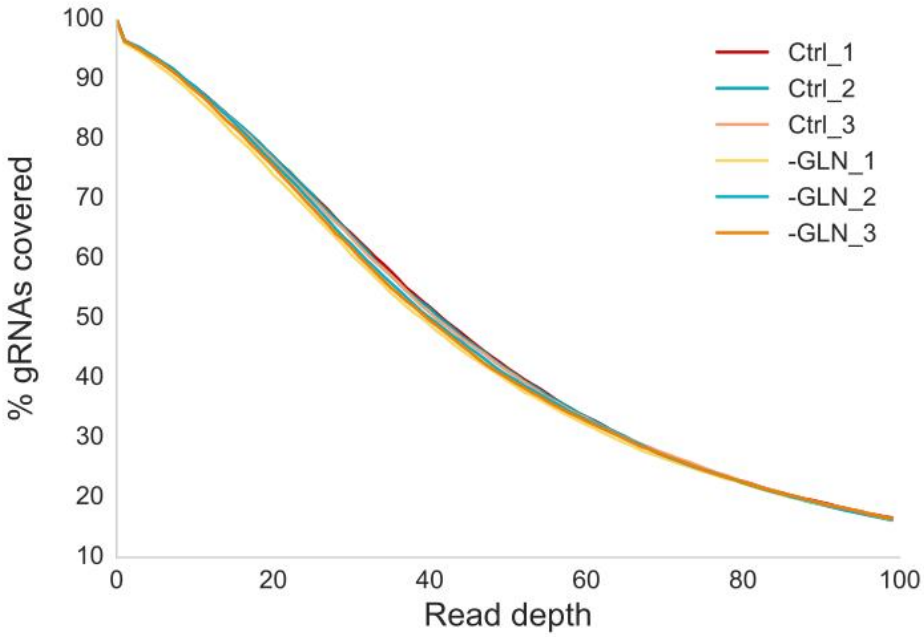
Supplementary Figure S2 Method Outline.

gRNAs are computationally designed to target the genes of interest, then synthesized and cloned into gRNA scaffold containing vectors. HEK cells are transfected with packaging vectors and gRNA vectors to generate a pool of viruses containing all the gRNA vectors. After harvest, the pooled library is used to transduce Cas9-expressing CHO cells at a low MOI to ensure a single integration event per cell. Cells positive for gRNA integration are selected for with antibiotics before undergoing a phenotypic screen. Genomic DNA is extracted from the collected cells and gRNA presence is compared between samples. Enriched or depleted gRNAs are ranked and candidate genes are phenotypically validated.



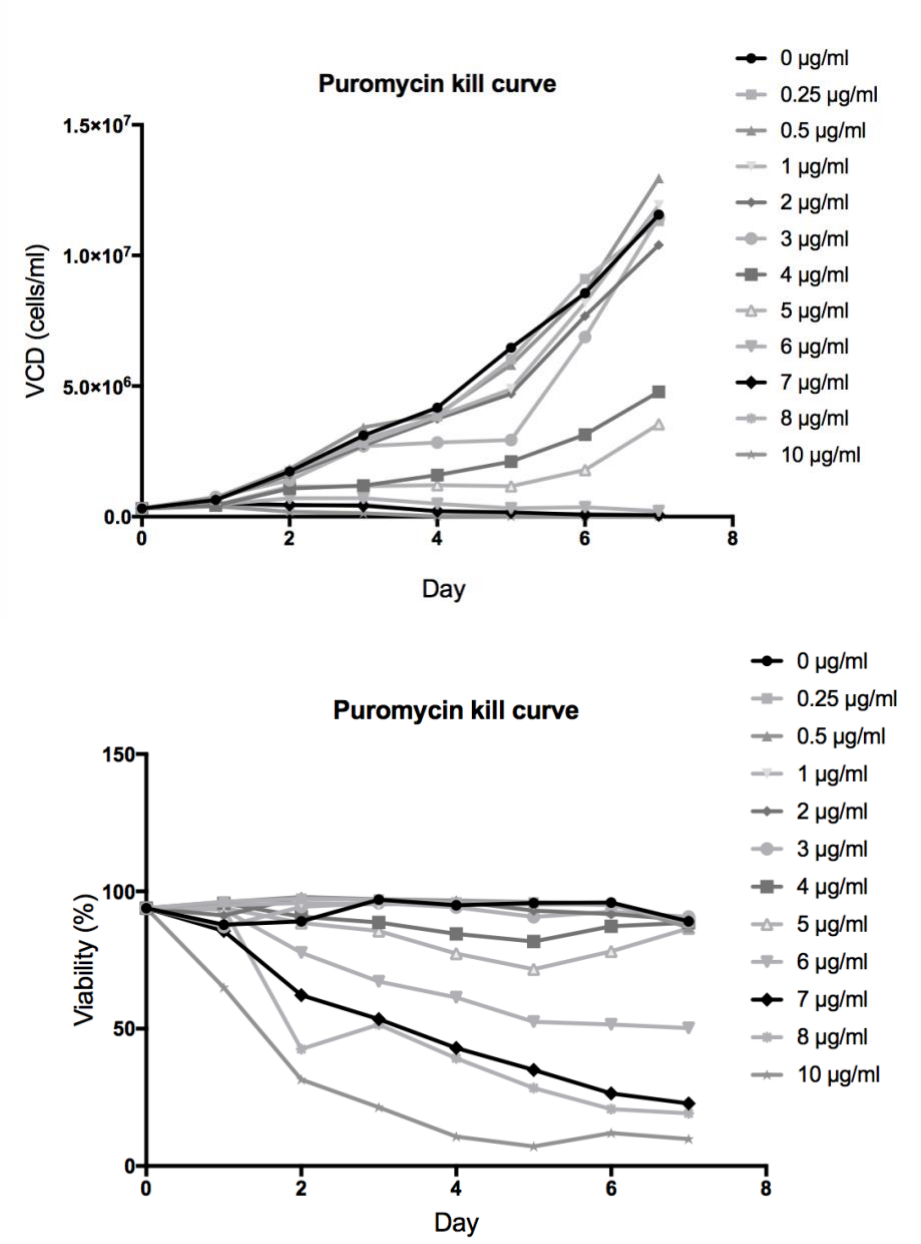
Supplementary Figure S3: **Viable Cell Density of the CHO^{Cas9} CRISPR Knockout Screen.**

CHO^{Cas9} CRISPR KO screens were grown in three replicates in media supplemented with glutamine (Gln) (black) and in media without Gln (grey) over 14 days. Cells were passed and VCD was measured every 3 days.



Supplementary Figure S4: Selection During Screening.

Percentage of gRNAs are plotted against read depth. There is little difference between cells grown for 14 days (T14) grown in media with glutamine (Ctrl) and without glutamine (-Gln) in agreement with weak selection.



Supplementary Figure S5: **Puromycin Kill Curve**

CHO-S cells were treated with varying concentrations of puromycin (spanning 0 – 10µg/ml). Viable cell density (VCD) and viability was measured across 7 days in culture. We continued to work with a concentration of 10µg/ml for sufficient stagnant VCD and decrease in viability.

Supplementary Tables

Supplementary Table S1: CHO-S^{Cas9} Validation by Indel Analysis.

CHO-S^{Cas9} cells were transfected with gRNA targeting the Mgat1 gene and sampled for indel analysis at a pool level. The above indel scores are of 5 different clonal CHO-S^{Cas9} cell lines from two different experiments. Cas9 showed ~60-70% cutting efficiency for all clonal cell lines.

Sample	Target	%Indel
CHO-S^{Cas9}_L1_mGat1	Mgat1	66,81%
CHO-S^{Cas9}_L15_mGat1	Mgat1	69,16%
CHO-S^{Cas9}_L21_mGat1	Mgat1	59,12%
CHO-S^{Cas9}_positive control	Mgat1	69,19%
CHO-S^{Cas9}_negative control	Mgat1	0,44%

Sample	Target	%Indel
CHO-S^{Cas9}_L12_mGat1	Mgat1	51,93%
CHO-S^{Cas9}_positive control	Mgat1	67,63%
CHO-S^{Cas9}_negative control	Mgat1	0,23%

Supplementary Table S2: Primers and Oligos

Primer name	Sequence (Green: NGS compatible overhangs)
MGAT1_gRNA_fwd	GGAAAGGACGAAACACCGTGGAGTTGGAGCGGCAGCGGTTTTAGAGCTAGAAAT
MGAT1_gRNA_rev	CTAAAACCGCTGCCGCTCCAACTCCACGGTGTTTCGTCCTTTCCACAAGATAT
MGAT1_miseq_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTTCTGGACACGCCCAGC
MGAT1_miseq_fwd	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCCACGGTGGGCACTTT
Lib_miseq_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATCTTGTGGAAAGGACGAAACAC
Lib_miseq_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCTCGGTGCCACTTTTTCAAGTT
OMA1_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTCCCATGAAGCAGCATAAGCA
OMA1_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGAGATCCCTGGGACATCCT
LIB_8xN_NGS_FWD	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNNNNGGCTTTATATATCTTGTGGAAAGGACGAAACACC
LIB_8xN_NGS_REV	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNNNNNNCCGACTCGGTGCCACTTTTTCAA
Target_X_gRNAfwd	XXXX
Target_X_gRNArev	XXXX
Target_X_Miseqfwd	XXXX
Target_X_Miseqrev	XXXX

REFERENCES

1. World Health Organization. *Global action plan for the prevention and control of noncommunicable diseases 2013-2020*. (2013).
2. World Health Organization. Biotherapeutics. *World Health Organization | Biotherapeutics* Available at: https://www.who.int/medicines/access/biotherapeutics/en/. (Accessed: 27th February 2019)
3. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2018. *CA Cancer J. Clin.* **68**, 7–30 (2018).
4. The Lancet. GLOBOCAN 2018: counting the toll of cancer. *Lancet* **392**, 985 (2018).
5. Rosling, H., Rosling, O. & Rönnlund, A. R. *Factfulness*. (Lindhardt og Ringhof, 2018).
6. Blackstone, E. A. & Fuhr, J. P., Jr. Biopharmaceuticals: the economic equation. *Biotechnol. Healthc.* **4**, 41–45 (2007).
7. Grilo, A. L. & Mantalaris, A. The Increasingly Human and Profitable Monoclonal Antibody Market. *Trends Biotechnol.* **37**, 9–16 (2019).
8. Walsh, G. Biopharmaceutical benchmarks 2018. *Nat. Biotechnol.* **36**, 1136–1145 (2018).
9. Walsh, G. Biopharmaceutical benchmarks 2010. *Nat. Biotechnol.* **28**, 917–924 (2010).
10. Walsh, G. Biopharmaceutical benchmarks 2014. *Nat. Biotechnol.* **32**, 992–1000 (2014).
11. Gottesman, M. M. *Molecular Cell Genetics*. (Wiley, 1985).
12. Wurm, F. CHO Quasispecies—Implications for Manufacturing Processes. *Processes* **1**, 296–311 (2013).
13. Wurm, F. & Wurm, M. Cloning of CHO Cells, Productivity and Genetic Stability—A Discussion. *Processes* **5**, 20 (2017).
14. Wurm, F. M. & Hacker, D. First CHO genome. *Nat. Biotechnol.* **29**, 718–720 (2011).
15. Hamilton, W. G. & Ham, R. G. Clonal growth of chinese hamster cell lines in protein-free media. *In Vitro* **13**, 537–547 (1977).
16. Fischer, S., Handrick, R. & Otte, K. The art of CHO cell engineering: A comprehensive retrospect and future perspectives. *Biotechnol. Adv.* **33**, 1878–1896 (2015).
17. Lewis, N. E. *et al.* Genomic landscapes of Chinese hamster ovary cell lines as revealed by the Cricetulus griseus draft genome. *Nat. Biotechnol.* **31**, 759–765 (2013).
18. Kim, J. Y., Kim, Y.-G. & Lee, G. M. CHO cells in biotechnology for production of recombinant

proteins: current state and further potential. *Appl. Microbiol. Biotechnol.* **93**, 917–930 (2012).

19. Almo, S. C. & Love, J. D. Better and faster: improvements and optimization for mammalian recombinant protein production. *Curr. Opin. Struct. Biol.* **26**, 39–43 (2014).
20. Kaufman, R. J. & Sharp, P. A. Amplification and expression of sequences cotransfected with a modular dihydrofolate reductase complementary dna gene. *J. Mol. Biol.* **159**, 601–621 (1982).
21. Bebbington, C. R. *et al.* High-level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker. *Biotechnology* **10**, 169–175 (1992).
22. Li, F., Vijayasankaran, N., Shen, A. (yijuan), Kiss, R. & Amanullah, A. Cell culture processes for monoclonal antibody production. *MAbs* **2**, 466–479 (2010).
23. Priola, J. J. *et al.* High-throughput screening and selection of mammalian cells for enhanced protein production. *Biotechnol. J.* **11**, 853–865 (2016).
24. Hefzi, H. *et al.* A Consensus Genome-scale Reconstruction of Chinese Hamster Ovary Cell Metabolism. *Cell Syst* **3**, 434–443.e8 (2016).
25. Xu, X. *et al.* The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat. Biotechnol.* **29**, 735–741 (2011).
26. Brinkrolf, K. *et al.* Chinese hamster genome sequenced from sorted chromosomes. *Nat. Biotechnol.* **31**, 694–695 (2013).
27. Kaas, C. S., Kristensen, C., Betenbaugh, M. J. & Andersen, M. R. Sequencing the CHO DXB11 genome reveals regional variations in genomic stability and haploidy. *BMC Genomics* **16**, 160 (2015).
28. Rupp, O. *et al.* A reference genome of the Chinese hamster based on a hybrid assembly strategy. *Biotechnol. Bioeng.* **115**, 2087–2100 (2018).
29. Carroll, D. Genome Editing: Past, Present, and Future. *Yale J. Biol. Med.* **90**, 653–659 (2017).
30. Qi, X. *et al.* The applications of CRISPR screen in functional genomics. *Brief. Funct. Genomics* **16**, 34–37 (2017).
31. Wright, A. V., Nuñez, J. K. & Doudna, J. A. Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. *Cell* **164**, 29–44 (2016).

32. Wang, H., La Russa, M. & Qi, L. S. CRISPR/Cas9 in Genome Editing and Beyond. *Annu. Rev. Biochem.* **85**, 227–264 (2016).

33. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).

34. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).

35. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).

36. Liu, X. S. *et al.* Editing DNA Methylation in the Mammalian Genome. *Cell* **167**, 233–247.e17 (2016).

37. Hilton, I. B. *et al.* Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* **33**, 510–517 (2015).

38. Chen, B. *et al.* Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491 (2013).

39. Liu, Z., Liang, Y., Ang, E. L. & Zhao, H. A New Era of Genome Integration-Simply Cut and Paste! *ACS Synth. Biol.* **6**, 601–609 (2017).

40. Margolin, J. F. *et al.* Krüppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 4509–4513 (1994).

41. Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013).

42. Wysocka, J. & Herr, W. The herpes simplex virus VP16-induced complex: the makings of a regulatory switch. *Trends Biochem. Sci.* **28**, 294–304 (2003).

43. Schmitz, M. L. & Baeuerle, P. A. The p65 subunit is responsible for the strong transcription activating potential of NF-kappa B. *EMBO J.* **10**, 3805–3817 (1991).

44. Hardwick, J. M., Tse, L., Applegren, N., Nicholas, J. & Veliuona, M. A. The Epstein-Barr virus R transactivator (Rta) contains a complex, potent activation domain with properties different from those of VP16. *J. Virol.* **66**, 5500–5508 (1992).

45. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588 (2015).

46. Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S. & Vale, R. D. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**, 635–646 (2014).

47. Chavez, A. *et al.* Comparison of Cas9 activators in multiple species. *Nat. Methods* **13**, 563–567 (2016).

48. Horlbeck, M. A. *et al.* Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *Elife* **5**, (2016).

49. Horlbeck, M. A. *et al.* Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *Elife* **5**, (2016).

50. Shuman, H. A. & Silhavy, T. J. The art and design of genetic screens: Escherichia coli. *Nat. Rev. Genet.* **4**, 419–431 (2003).

51. Casselton, L. & Zolan, M. The art and design of genetic screens: filamentous fungi. *Nat. Rev. Genet.* **3**, 683–697 (2002).

52. Jorgensen, E. M. & Mango, S. E. The art and design of genetic screens: caenorhabditis elegans. *Nat. Rev. Genet.* **3**, 356–369 (2002).

53. St Johnston, D. The art and design of genetic screens: Drosophila melanogaster. *Nat. Rev. Genet.* **3**, 176–188 (2002).

54. Boutros, M. & Ahringer, J. The art and design of genetic screens: RNA interference. *Nat. Rev. Genet.* **9**, 554–566 (2008).

55. Chou, Y.-C. *et al.* Variations in genome-wide RNAi screens: lessons from influenza research. *J. Clin. Bioinforma.* **5**, 2 (2015).

56. Goh, G. Y. & Bard, F. A. RNAi screens for genes involved in Golgi glycosylation. *Methods Mol. Biol.* **1270**, 411–426 (2015).

57. Galeev, R., Karlsson, C., Baudet, A. & Larsson, J. Forward RNAi Screens in Human Hematopoietic Stem Cells. *Methods Mol. Biol.* **1622**, 29–50 (2017).

58. Karlsson, C., Larsson, J. & Baudet, A. Forward RNAi screens in human stem cells. *Methods Mol. Biol.* **650**, 29–43 (2010).

59. Schuster, A. *et al.* RNAi/CRISPR Screens: from a Pool to a Valid Hit. *Trends Biotechnol.* **37**, 38–55 (2019).

60. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87

(2014).

61. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science* **343**, 80–84 (2013).

62. Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using CRISPR-Cas9. *Nat. Rev. Genet.* **16**, 299–311 (2015).

63. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647–661 (2014).

64. Joung, J. *et al.* Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* **12**, 828–863 (2017).

65. Park, R. J. *et al.* A genome-wide CRISPR screen identifies a restricted set of HIV host dependency factors. *Nat. Genet.* **49**, 193–203 (2017).

66. Korkmaz, G. *et al.* Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nat. Biotechnol.* **34**, 192–198 (2016).

67. Chen, S. *et al.* Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis. *Cell* **160**, 1246–1260 (2015).

68. Rauscher, B., Heigwer, F., Breinig, M., Winter, J. & Boutros, M. GenomeCRISPR - a database for high-throughput CRISPR/Cas9 screens. *Nucleic Acids Res.* **45**, D679–D686 (2017).

69. Hart, T. *et al.* High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* **163**, 1515–1526 (2015).

70. Cowley, G. S. *et al.* Parallel genome-scale loss of function screens in 216 cancer cell lines for the identification of context-specific genetic dependencies. *Sci Data* **1**, 140035 (2014).

71. 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).

72. Grav, L. M. *et al.* One-step generation of triple knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment. *Biotechnol. J.* **10**, 1446–1456 (2015).

73. Amann, T. *et al.* Glyco-engineered CHO cell lines producing alpha-1-antitrypsin and C1 esterase

inhibitor with fully humanized N-glycosylation profiles. *Metab. Eng.* **52**, 143–152 (2018).

74. Hefzi, H. & Lewis, N. Mammalian cells devoid of lactate dehydrogenase activity. *World Patent* (2017).
75. Marx, N. *et al.* CRISPR-Based Targeted Epigenetic Editing Enables Gene Expression Modulation of the Silenced Beta-Galactoside Alpha-2,6-Sialyltransferase 1 in CHO Cells. *Biotechnol. J.* **13**, 1700217 (2018).
76. Grav, L. M. *et al.* Minimizing Clonal Variation during Mammalian Cell Line Engineering for Improved Systems Biology Data Generation. *ACS Synth. Biol.* **7**, 2148–2159 (2018).
77. Schmieder, V. *et al.* Enhanced Genome Editing Tools For Multi-Gene Deletion Knock-Out Approaches Using Paired CRISPR sgRNAs in CHO Cells. *Biotechnol. J.* **13**, e1700211 (2018).
78. Shen, C.-C., Sung, L.-Y., Lin, S.-Y., Lin, M.-W. & Hu, Y.-C. Enhancing Protein Production Yield from Chinese Hamster Ovary Cells by CRISPR Interference. *ACS Synth. Biol.* **6**, 1509–1519 (2017).
79. Goswami, J., Sinskey, A. J., Steller, H., Stephanopoulos, G. N. & Wang, D. I. C. Apoptosis in batch cultures of Chinese Hamster Ovary cells. *Biotechnol. Bioeng.* **62**, 632–640 (1999).
80. Dewson, G. & Kluck, R. M. Mechanisms by which Bak and Bax permeabilise mitochondria during apoptosis. *J. Cell Sci.* **122**, 2801–2808 (2009).
81. Slee, E. A., Adrain, C. & Martin, S. J. Executioner caspase-3, -6, and -7 perform distinct, non-redundant roles during the demolition phase of apoptosis. *J. Biol. Chem.* **276**, 7320–7326 (2001).
82. Cost, G. J. *et al.* BAK and BAX deletion using zinc-finger nucleases yields apoptosis-resistant CHO cells. *Biotechnol. Bioeng.* **105**, 330–340 (2010).
83. Lim, S. F. *et al.* RNAi suppression of Bax and Bak enhances viability in fed-batch cultures of CHO cells. *Metab. Eng.* **8**, 509–522 (2006).
84. Kim, N. S. & Lee, G. M. Inhibition of sodium butyrate-induced apoptosis in recombinant Chinese hamster ovary cells by constitutively expressing antisense RNA of caspase-3. *Biotechnol. Bioeng.* **78**, 217–228 (2002).
85. Sung, Y. H., Lee, J. S., Park, S. H., Koo, J. & Lee, G. M. Influence of co-down-regulation of caspase-3 and caspase-7 by siRNAs on sodium butyrate-induced apoptotic cell death of Chinese hamster ovary cells producing thrombopoietin. *Metab. Eng.* **9**, 452–464 (2007).

86. Petersen, S. D. *et al.* Modular 5'-UTR hexamers for context-independent tuning of protein expression in eukaryotes. *Nucleic Acids Res.* (2018). doi:10.1093/nar/gky734

87. Long, L. *et al.* Regulation of transcriptionally active genes via the catalytically inactive Cas9 in C. elegans and D. rerio. *Cell Res.* **25**, 638–641 (2015).

88. Nan, X., Campoy, F. J. & Bird, A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88**, 471–481 (1997).

89. Yeo, N. C. *et al.* An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nat. Methods* **15**, 611–616 (2018).

90. Boettcher, M. & McManus, M. T. Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR. *Mol. Cell* **58**, 575–585 (2015).

91. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–408 (2001).

92. Uhlén, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).

93. Popp, O. *et al.* Development of a pre-glycoengineered CHO-K1 host cell line for the expression of antibodies with enhanced Fc mediated effector function. *MAbs* **10**, 290–303 (2018).

94. Lee, J. S., Grav, L. M., Pedersen, L. E., Lee, G. M. & Kildegaard, H. F. Accelerated homology-directed targeted integration of transgenes in Chinese hamster ovary cells via CRISPR/Cas9 and fluorescent enrichment. *Biotechnol. Bioeng.* **113**, 2518–2523 (2016).

95. Grav, L. M., la Cour Karottki, K. J., Lee, J. S. & Kildegaard, H. F. Application of CRISPR/Cas9 Genome Editing to Improve Recombinant Protein Production in CHO Cells. in *Methods in Molecular Biology* 101–118 (2017).

96. Lin, S., Ewen-Campen, B., Ni, X., Housden, B. E. & Perrimon, N. In Vivo Transcriptional Activation Using CRISPR/Cas9 in Drosophila. *Genetics* **201**, 433–442 (2015).

97. Chavez, A. *et al.* Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* **12**, 326–328 (2015).

98. Valle, D., Downing, S. J., Harris, S. C. & Phang, J. M. Proline biosynthesis: multiple defects in Chinese

hamster ovary cells. *Biochem. Biophys. Res. Commun.* **53**, 1130–1136 (1973).

99. Liu, H. *et al.* CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation: Fig. 1. *Bioinformatics* **31**, 3676–3678 (2015).
100. GPP Web Portal - Home. Available at: https://portals.broadinstitute.org/gpp/public/. (Accessed: 14th February 2019)
101. Cheng, A. W. *et al.* Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.* **23**, 1163–1171 (2013).
102. Van der Auwera, G. A. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, 11.10.1–33 (2013).
103. Jayapal, KP, Wlaschin, KF, Hu, WS & Yap, MGS. Recombinant protein therapeutics from CHO Cells - 20 years and counting. *Chem. Eng. Prog.* **103**, 40–47 (2007).
104. Gutierrez, J. M. & Lewis, N. E. Optimizing eukaryotic cell hosts for protein production through systems biotechnology and genome-scale modeling. *Biotechnol. J.* **10**, 939–949 (2015).
105. Sakuma, T. *et al.* Homologous Recombination-Independent Large Gene Cassette Knock-in in CHO Cells Using TALEN and MMEJ-Directed Donor Plasmids. *Int. J. Mol. Sci.* **16**, 23849–23866 (2015).
106. Santiago, Y. *et al.* Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 5809–5814 (2008).
107. Cullen, L. M. & Arndt, G. M. Genome-wide screening for gene function using RNAi in mammalian cells. *Immunol. Cell Biol.* **83**, 217–223 (2005).
108. Kaelin, W. G., Jr. Molecular biology. Use and abuse of RNAi to study mammalian gene function. *Science* **337**, 421–422 (2012).
109. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
110. Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).
111. Hart, T., Brown, K. R., Sircoulomb, F., Rottapel, R. & Moffat, J. Measuring error rates in genomic

perturbation screens: gold standards for human functional genomics. *Mol. Syst. Biol.* **10**, 733 (2014).

112. Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M. D. C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat. Biotechnol.* **32**, 267–273 (2014).
113. Zhou, Y. *et al.* High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature* **509**, 487–491 (2014).
114. Bassett, A. R., Kong, L. & Liu, J.-L. A genome-wide CRISPR library for high-throughput genetic screening in Drosophila cells. *J. Genet. Genomics* **42**, 301–309 (2015).
115. Yao, T. & Asayama, Y. Animal-cell culture media: History, characteristics, and current issues. *Reprod. Med. Biol.* **16**, 99–117 (2017).
116. Borys, M. C., Linzer, D. I. & Papoutsakis, E. T. Ammonia affects the glycosylation patterns of recombinant mouse placental lactogen-I by chinese hamster ovary cells in a pH-dependent manner. *Biotechnol. Bioeng.* **43**, 505–514 (1994).
117. Thorens, B. & Vassalli, P. Chloroquine and ammonium chloride prevent terminal glycosylation of immunoglobulins in plasma cells without affecting secretion. *Nature* **321**, 618–620 (1986).
118. Yang, M. & Butler, M. Effects of ammonia on CHO cell growth, erythropoietin production, and glycosylation. *Biotechnol. Bioeng.* **68**, 370–380 (2000).
119. Yang, M. & Butler, M. Effect of ammonia on the glycosylation of human recombinant erythropoietin in culture. *Biotechnol. Prog.* **16**, 751–759 (2000).
120. Wise, D. R. & Thompson, C. B. Glutamine addiction: a new therapeutic target in cancer. *Trends Biochem. Sci.* **35**, 427–433 (2010).
121. Fritsch, S. D. & Weichhart, T. Effects of Interferons and Viruses on Metabolism. *Front. Immunol.* **7**, (2016).
122. Newsholme, P. *et al.* Glutamine and glutamate as vital metabolites. *Braz. J. Med. Biol. Res.* **36**, 153–163 (2003).
123. Altamirano, C., Paredes, C., Cairo, J. J. & Godia, F. Improvement of CHO Cell Culture Medium

Formulation: Simultaneous Substitution of Glucose and Glutamine. *Biotechnol. Prog.* **16**, 69–75 (2000).

124. Ahn, W. S. & Antoniewicz, M. R. Parallel labeling experiments with [1,2-(13)C]glucose and [U-(13)C]glutamine provide new insights into CHO cell metabolism. *Metab. Eng.* **15**, 34–47 (2013).
125. Doench, J. G. Am I ready for CRISPR? A user's guide to genetic screens. *Nat. Rev. Genet.* **19**, 67–80 (2017).
126. Joung, J. *et al.* Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature* **548**, 343–346 (2017).
127. Heaton, B. E. *et al.* A CRISPR Activation Screen Identifies a Pan-avian Influenza Virus Inhibitory Host Factor. *Cell Rep.* **20**, 1503–1512 (2017).
128. Liu, S. J. *et al.* CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science* **355**, (2017).
129. Rosenbluh, J. *et al.* Complementary information derived from CRISPR Cas9 mediated gene deletion and suppression. *Nat. Commun.* **8**, 15403 (2017).
130. Lee, J. S., Kallehauge, T. B., Pedersen, L. E. & Kildegaard, H. F. Site-specific integration in CHO cells mediated by CRISPR/Cas9 and homology-directed DNA repair pathway. *Sci. Rep.* **5**, 8572 (2015).
131. Spahn, P. N. *et al.* PinAPL-Py: A comprehensive web-application for the analysis of CRISPR/Cas9 screens. *Sci. Rep.* **7**, 15854 (2017).
132. Li, W. *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).
133. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
134. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
135. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
136. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).

137. Browne, E. P., Wing, B., Coleman, D. & Shenk, T. Altered cellular mRNA levels in human cytomegalovirus-infected fibroblasts: viral block to the accumulation of antiviral mRNAs. *J. Virol.* **75**, 12319–12330 (2001).

138. Smith, I. *et al.* Evaluation of RNAi and CRISPR technologies by large-scale gene expression profiling in the Connectivity Map. *PLoS Biol.* **15**, e2003213 (2017).

139. Franceschini, A. *et al.* Specific inhibition of diverse pathogens in human cells by synthetic microRNA-like oligonucleotides inferred from RNAi screens. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4548–4553 (2014).

140. La Russa, M. F. & Qi, L. S. The New State of the Art: Cas9 for Gene Activation and Repression. *Mol. Cell. Biol.* **35**, 3800–3809 (2015).

141. Shen, J. P. *et al.* Combinatorial CRISPR–Cas9 screens for de novo mapping of genetic interactions. *Nat. Methods* **14**, 573–576 (2017).

142. Du, D. *et al.* Genetic interaction mapping in mammalian cells using CRISPR interference. *Nat. Methods* **14**, 577–580 (2017).

143. Han, K. *et al.* Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. *Nat. Biotechnol.* **35**, 463–474 (2017).

144. Cebrian-Serrano, A. & Davies, B. CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. *Mamm. Genome* **28**, 247–261 (2017).

145. Kweon, J. *et al.* Fusion guide RNAs for orthogonal gene manipulation with Cas9 and Cpf1. *Nat. Commun.* **8**, 1723 (2017).

146. Najm, F. J. *et al.* Orthologous CRISPR-Cas9 enzymes for combinatorial genetic screens. *Nat. Biotechnol.* **36**, 179–189 (2018).

147. Boettcher, M. *et al.* Dual gene activation and knockout screen reveals directional dependencies in genetic networks. *Nat. Biotechnol.* **36**, 170–178 (2018).

148. Zetsche, B. *et al.* Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).

149. Introducing MAD7 | Inscripta. *Inscripta* Available at: https://www.inscripta.com/madzymes/.

(Accessed: 24th February 2019)

150. Sherkow, J. S. The CRISPR Patent Landscape: Past, Present, and Future. *The CRISPR Journal* **1**, 5–9 (2018).

151. Sharon, D. M. & Kamen, A. A genome-wide CRISPR screen to generate high-yield cell lines for pandemic influenza vaccine production. (2018).

152. Parnas, O. *et al.* A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks. *Cell* **162**, 675–686 (2015).

153. Pettitt, S. J. *et al.* Genome-wide and high-density CRISPR-Cas9 screens identify point mutations in PARP1 causing PARP inhibitor resistance. *Nat. Commun.* **9**, 1849 (2018).

154. Ruiz, S. *et al.* A Genome-wide CRISPR Screen Identifies CDC25A as a Determinant of Sensitivity to ATR Inhibitors. *Mol. Cell* **62**, 307–313 (2016).

155. Wang, T. *et al.* Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. *Nat. Commun.* **9**, 2475 (2018).

156. Zhang, R. *et al.* A CRISPR screen defines a signal peptide processing pathway required by flaviviruses. *Nature* **535**, 164–168 (2016).

157. Patel, S. J. *et al.* Identification of essential genes for cancer immunotherapy. *Nature* **548**, 537–542 (2017).

158. Geu-Flores, F., Nour-Eldin, H. H., Nielsen, M. T. & Halkier, B. A. USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res.* **35**, e55 (2007).

159. Lam, M. T. Y. *et al.* Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature* **498**, 511–515 (2013).

160. Core, L. J., Waterfall, J. J. & Lis, J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845–1848 (2008).

161. Nechaev, S. *et al.* Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in Drosophila. *Science* **327**, 335–338 (2010).

162. Scruggs, B. S. *et al.* Bidirectional Transcription Arises from Two Distinct Hubs of Transcription Factor Binding and Active Chromatin. *Mol. Cell* **58**, 1101–1112 (2015).

163. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).

164. Homer Software and Data Download. Available at: http://homer.ucsd.edu/homer/ngs/tss/index.html. (Accessed: 18th February 2019)

165. Li, S. *et al.* Proteogenomic annotation of the Chinese hamster reveals extensive novel translation events and endogenous retroviral elements. (2018). doi:10.1101/468181

166. Ellis, E. L. & Delbrück, M. THE GROWTH OF BACTERIOPHAGE. *J. Gen. Physiol.* **22**, 365–384 (1939).

167. Evers, B. *et al.* CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat. Biotechnol.* **34**, 631–633 (2016).

168. Han, X. *et al.* Microfluidic Cell Deformability Assay for Rapid and Efficient Kinase Screening with the CRISPR-Cas9 System. *Angew. Chem. Int. Ed Engl.* **55**, 8561–8565 (2016).

169. Klann, T. S. *et al.* CRISPR-Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. *Nat. Biotechnol.* **35**, 561–568 (2017).

170. Liu, N. *et al.* Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. *Nature* **553**, 228–232 (2018).

171. Manguso, R. T. *et al.* In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target. *Nature* **547**, 413–418 (2017).

172. Munoz, D. M. *et al.* CRISPR Screens Provide a Comprehensive Assessment of Cancer Vulnerabilities but Generate False-Positive Hits for Highly Amplified Genomic Regions. *Cancer Discov.* **6**, 900–913 (2016).

173. Wong, A. S. L. *et al.* Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 2544–2549 (2016).

174. Wu, Y. *et al.* A genome-scale CRISPR-Cas9 screening method for protein stability reveals novel regulators of Cdc25A. *Cell Discov* **2**, 16014 (2016).

175. Zhong, C. *et al.* CRISPR-Cas9-Mediated Genetic Screening in Mice with Haploid Embryonic Stem Cells

Carrying a Guide RNA Library. *Cell Stem Cell* **17**, 221–232 (2015).

176. Zhu, S. *et al.* Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. *Nat. Biotechnol.* **34**, 1279–1286 (2016).
177. Cross, B. C. S. *et al.* Increasing the performance of pooled CRISPR–Cas9 drop-out screening. *Sci. Rep.* **6**, (2016).
178. Kiessling, M. K. *et al.* Identification of oncogenic driver mutations by genome-wide CRISPR-Cas9 dropout screening. *BMC Genomics* **17**, (2016).
179. Shi, J. *et al.* Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* **33**, 661–667 (2015).
180. Zhu, S., Zhou, Y. & Wei, W. Genome-Wide CRISPR/Cas9 Screening for High-Throughput Functional Genomics in Human Cells. *Methods Mol. Biol.* **1656**, 175–181 (2017).