CRISPR/Cas9-mediated epigenetic repression of an enhancer linked to visceral obesity

Tiffany Chor Kiu Ngan

This thesis is submitted in partial fulfilment of the requirements for the degree of Master of Science

Department of Biological Sciences

&

The Hormone Laboratory Research Group

Department of Clinical Science

University of Bergen

Norway

June 2024

Acknowledgments

Table of contents

Summary

Projects like the Human Genome Project, ENCODE, and RoadMap Epigenomics have enabled genome-wide association studies (GWAS) to map associations between genetic variants and human disease. However, most variants detected in GWAS are found in non-coding regions, which limits mechanistic and biological understanding of the associations. Many of these variants are thought to be situated in enhancers, affecting gene expression rather than gene structure. However, both enhancers themselves, as well as their target genes are incompletely mapped in most cells and tissues. Therefore, there is a great need for developing methods that can both detect enhancers in disease-associated loci and map their target genes.

Active enhancers are characterized by the specific histone marks H3K4me1 and H3K27ac. The former mark can be removed by the lysine-specific demethylase 1 (LSD1) enzyme, which leads to repression of the enhancer. By fusing LSD1 to a catalytically dead Cas9 (dCas9) protein and introduce it to a relevant cell type, targeted epigenetic repression of an enhancer can be achieved. Moreover, by providing a transcriptional repressor alongside LSD1, an even more potent repression of the enhancer's target gene(s) can be achieved.

Visceral obesity is associated with increased risk of metabolic disease and premature death, but the genes contributing to disease risk are largely unknown. A recent GWAS identified the 11q23.3 locus to be associated with visceral obesity, but the associated variants are located in non-coding regions. Previous work by our lab has identified a predicted causal variant, rs1799993, and found it to be situated in an enhancer that is active in adipose-derived mesenchymal stem cells (AdMSC).

The overall aim of this thesis is therefore to establish a method of epigenetic repression in AdMSC to identify the target genes of the 11q23.3 enhancer associated with visceral obesity.

To this end, a dox-inducible lentiviral LSD1-dCas9 system coupled with constitutive lentiviral sgRNA(MS2)-MCP-KRAB expression was used in this thesis. Three different lentiviral constructs harboring Tet-on-3G-BSD, LSD1-dCas9-mCherry and sgRNA(MS2)-MCP-KRAB-zsGreen, respectively were generated and successively introduced into the AdMSC cell line ASC52telo after validation of the plasmids and viruses. Cells transduced with viruses encoding Tet-on-3G-BSD were selected using blasticidin, and cells transduced with LSD1-dCas9 and the sgRNA system were selected using FACS. Integration of lentiviruses into the genome of

transduced cells and expression of the encoded constructs was assessed by qPCR of gDNA and cDNA, respectively. Protein levels of transduced constructs were analyzed by WB.

The lentiviral plasmids were sequenced and found to contain the correct insert, except for a previously used Tet-On-3G construct that was found to be incorrect. Only correct plasmids were used for further experiments in this thesis. The plasmids were further functionally validated by transient overexpression in HT1080 cells with and without doxycycline, and found to express Tet-On-3G and LSD1-dCas9. A clear Tet-On-3G dependent dox-inducible effect was demonstrated for dCas9.

Lentiviruses harboring the respective constructs were next successfully generated and found to contain viral concentrations of 4-9 x 10^{10} physical particles pr ml. A functional titer test revealed that about 1/1000 physical particles were infectious. ASC52telo cells were transduced with virus harboring Tet-On-3G-BSD using different multiplicity of infections (MOIs) and after BSD selection, the best MOI was found to be 15. The transduced cells were not able to grow as single colonies, thus a heterogenous population of Tet-On-3G expressing cells was expanded and successively transduced with lentivirus harboring LSD1-dCas9-mCherry. Only 2-8% of the cells were mCherry positive, but these populations were sorted, expanded and used in the subsequent transduction with virus containing sgRNA and KRAB. The latter transduction was more successful with 16-18% efficiency as determined by FACS. Triple-transduced cells, as well as single- and double-transduced controls, were sorted and successfully expanded. Intriguingly, mCherry expression was detected independently of Tet-On-3G. While expression of the Tet-On-3G and sgRNA-KRAB constructs were validated by qPCR, no clear expression of LSD1-dCAs9 could be demonstrated by qPCR or WB in the transduced cells, despite the positive mCherry signal in these cells.

In conclusion, the establishment of an inducible CRISPR-Cas9 enhancer repression system in AdMSCs was partially established, but it remains to be determined whether the essential component LSD1-dCas9 is correctly expressed or not. Thus, repression of the 11q23.3 enhancer could not be performed, and consequently the target genes of the enhancer were not identified.

List of Abbreviations

AdMSCs	Adipose-derived mesenchymal stem cells
BSD	Blasticidin
Cas9	CRISPR-associated protein 9
CREs	Cis-regulatory elements
CRISPR	Clustered regularly interspaced palindromic repeats
CoREST	Corepressor to the REST (RE1 silencing transcription factor/neural restrictive silencing factor)
DBS	Double-stranded break
dCas9	catalytic deactivated Cas9
Dox	Doxycycline
enCRISPR/Cas9i	enhancer targeting CRISPR/Cas9-mediated inactivation
gRNA	guide RNA
GWAS	Genome-wide association study
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HMT	Histone methyltransferase
KRAB	Krüppel associated box
KDM	Lysine demethylase
LSD1	Lysine-specific demethylase 1
MCP	MS2 coat protein
o/n	overnight
RT	Room temperature
sgRNA	Single guide RNA
SNPs	single-nucleotide polymorphisms
Tet	Tetracycline
TetO	Tet operator
TetR	Tet repressor
TF	Transcription Factor
TRE	Tet response element
VAT	Visceral adipose tissue

Introduction

1.1. Background

The genome is the foundation of all life. The human genome is 99.9% identical between any two individuals; thus, the remaining 0.1% gives us individual variation. Over the past three decades, a series of scientific projects have been carried out with the aim of uncovering this DNA variation which confers individuality. The sequence of the human genome was first mapped by the Human Genome Project, completed in 2002 (Hood and Rowen, 2013). Following this achievement, the Encyclopedia of DNA Elements (ENCODE) project was conducted to map out the functional elements of the genome, including areas involved in transcription, gene regulatory regions/areas of transcription factor association, chromatin structure, and histone modification (Dunham et al., 2012). The RoadMap Epigenomics project provided an even more detailed resolution of the epigenetic landscape of gene regulatory elements across the genome and for a much larger number of different tissues and cell types (Kundaje et al., 2015). These findings have allowed us to ascribe biochemical functions to 80% of the genome, which before this was especially challenging beyond the well-studied protein-coding regions.

The Human Genome Project took us into a new era of research into the genetics of disease. By knowing the sequence of the human genome, we could identify genetic variants associated with disease in a systematic way through genome-wide association studies (GWAS), which have now been performed for more than 1000 different human traits (Aguet et al., 2019). Strikingly, the vast majority of the trait-associated genetic variants in these GWAS are found in the non-coding DNA where the functions are largely unknown (Claussnitzer et al., 2020, Gaulton et al., 2023). Many, if not most, of these single-nucleotide variants are likely to affect gene regulatory mechanisms, such as by affecting transcription factor binding to the DNA. Thus, disease phenotypes can often be linked to a particular cell type or transcription factor (TF) (Dunham et al., 2012).

However, the single-nucleotide variants, also known as single nucleotide polymorphisms (SNPs), typically reported in GWAS, are usually just representatives of multiple SNPs co-inherited together. Thus, the actual causal SNPs affecting transcription factor binding are usually unknown (Barroso and McCarthy, 2019, Claussnitzer et al., 2020). Thus, despite the comprehensive information of human genetic variation associated with disease now available,

only a tiny fraction of this information has been translated into biological understanding (Gaulton et al., 2023, Claussnitzer et al., 2020).

Most non-communicable human complex diseases are characterized by genetic risk factors, which can be represented as alterations in gene expression patterns, including upregulation, downregulation, or abnormal expression of wild-type or mutant genes. Unlike monogenic diseases, which are caused by harmful coding mutations in single genes, characterized by loss of function (Tsai et al., 2022), complex diseases are influenced by hundreds of genetic variants (SNPs) which each have a small impact (Barroso and McCarthy, 2019). As mentioned above, these disease risk variants associated with complex traits and diseases are primarily located in non-coding regulatory regions of genes. This underscores the importance of characterizing the function of these non-coding regions and mapping their target genes.

The ENCODE and RoadMap Epigenomics projects have made a significant advance towards characterizing such non-coding regions by identifying enhancers in different cells and tissues. In addition, these newly identified elements exhibit a correlation with disease-associated SNPs, aiding in the interpretation of such genetic variations (Dunham et al., 2012) . In recent years, several studies have conducted disease-specific prioritization of genes at the level of integrating (GWAS) summary statistics data with functional genomics information, such as gene expression patterns and gene networks (Dey et al., 2022). These projects provided new insights and understanding of the human genome and the regulation of our genes, which is an invaluable resource for further research.

1.2. Principles of gene regulation

From a biological standpoint, every cell within an organism carries identical genetic material; however, the cells are differentiated into distinct specialized cell types, which differ in function and morphology. This intricate process underscores the beauty of genetic regulation. One of the components of gene regulation is Cis-regulatory elements (CREs), which constitute one layer of control over cell diversity. CREs are DNA components within the non-coding sequences, which consist of promoters, enhancers, silencers, insulators, and locus control regions. Among these elements, enhancers play a crucial role in regulating tissue-specific and temporally dependent gene expression (Chatterjee et al., 2011).

Another important component or layer of gene regulation is chromatin/histone modifications. Chromatin is the state where DNA is packaged within the cell. When the DNA is tightly packed, the genes and CREs are not accessible for gene expression. However, chromatin is not a rigid structure but rather a malleable DNA scaffold that responds to external cues (Mansisidor and Risca, 2022). Thus, changes in chromatin play key roles in gene regulation. There are two main types of chromatin changes: histone modifications and ATP-dependent chromatin remodeling (Vignali et al., 2000). The focus of this thesis is on histone modifications and the ATP-dependent chromatin remodeling will therefore not be further discussed. Histone modifications play essential roles in the majority of biological processes that manipulate and regulate DNA expression (Bannister and Kouzarides, 2011). There are many forms of histone modifications; in this thesis, we are going to focus on histone acetylation and histone methylation.

Histone acetylation is the acetylation of lysine residues in the N-terminal histone tails and is regulated by two opposing families of enzymes: histone acetyltransferases and histone deacetylases. Histone acetyltransferase (HAT) employs acetyl CoA as a cofactor and catalyzes the transfer of an acetyl group to the ε-amino group of the lysine side chains, neutralizing the lysine's positive charge. This leads to the weakening of the interaction between histones and the DNA, leading to the opening of the chromatin. Histone deacetylases (HDAC) remove the lysine acetylation, thus restoring the positive charge of the lysine. This stabilizes the local chromatin structure, making it inaccessible for transcription and thus acting as transcriptional repression (Bannister and Kouzarides, 2011).

Histone methylation, on the other hand, occurs primarily in the side chains of lysine (and arginine), and is performed by histone methyltransferase (HMT). Unlike acetylation, this modification does not alter the charge of the histone protein (Miller and Grant, 2013). The HMT can transfer up to three methyl groups to the ε-amino group of a lysine residue, resulting in mono-, di-, or tri-methylation, respectively, which adds a level of complexity (Martinez-Gamero et al., 2021). Most of the HMT enzymes methylate lysine within the N-terminal tails. These enzymes tend to be relatively specific, meaning that certain histone lysine methyltransferases only work on specific parts of the histone. (Bannister and Kouzarides, 2011). The position and methylation state of the histone tails make up the histone code, where specific marks or combinations of marks give different gene regulatory signals. For instance, trimethylation of lysine 4 on histone 3 (H3K4me3) is predominantly found at promoters of actively transcribed genes, whereas H3K27me3 is primarily present in repressed regions,

covering the gene body and flanking regions. H3K4me1 alone signifies primed enhancers, while its presence alongside H3K27me3 or H3K27ac signals poised or active enhancers, respectively (Martinez-Gamero et al., 2021).
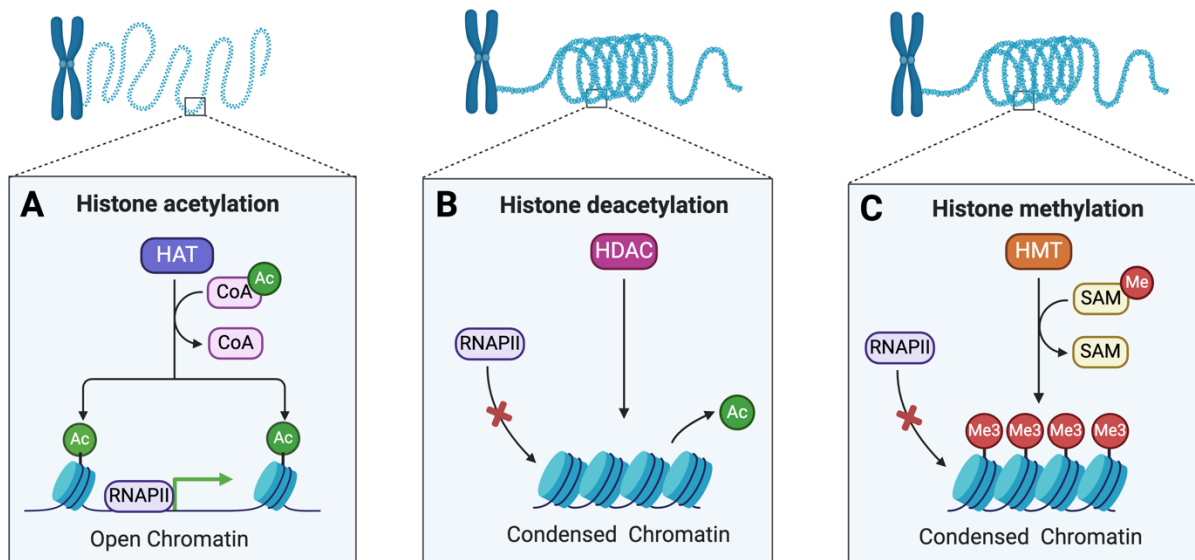
Figure 1.2.1: Histone methylation, deacetylation, and acetylation. This illustration demonstrates chromatin accessibility. When the chromatin is condensed, the DNA is tightly packed around the nucleosomes, and gene transcription is unavailable since RNA polymerase II (RNAPII) cannot bind the DNA. On the other hand, when the chromatin is open, the DNA is accessible for the RNAPII so that transcription can be performed. The state of the chromatin can be affected by several histone modifications. A: Histone acetylation involves the recruitment of histone acetyltransferase (HAT), which uses CoA as a cofactor to transfer the acetyl group to the ε-amino group of the histone tail, resulting in an open chromatin structure. B: Histone deacetylation involves the removal of the acetyl group by histone deacetylase (HDAC), leading to a more compact chromatin structure. RNAPII cannot access the DNA, and gene expression is repressed C: Histone methylation is controlled by histone methyltransferase (HMT), which adds methyl groups to the histone tail. Different HMTs work on different lysine on the histone. HMT uses S-adenosylmethionine (SAM), a universal methyl donor, to transferer methyl groups on the histone tail, creating mono-, di- and trimethylation, depending on the methyl groups transferred, respectively. This illustrates one of the examples of histone methylation, whereas trimethylation of lysine 4 on histone 3 (H3K4me3), resulting in the tight winding of the DNA to the nucleosome. The figure was made with Biorender.

Until two decades ago, histone methylation was thought to be a stable modification, which was only removed upon histone exchange or during DNA replication. This perspective shifted upon discovering that lysine-specific demethylase 1 (LSD1) facilitates the demethylation of H3K4me1 and H3K4me2. Consequently, this indicates that histone lysine methylation can undergo dynamic regulation at specific genes through the recruitment of methyltransferases and demethylases, similar to acetylation (Højfeldt et al., 2013). Yet, this pathway is not a direct reversal of methylation (Bannister and Kouzarides, 2011). LSD1 operates as a flavin-adenine dinucleotide (FAD)-dependent amine oxidase, playing a crucial role in demethylating histone H3K4 mono- and di-methylation. This enzymatic activity results in transcriptional repression

and activation. Notably, LSD1 does not remove trimethyl groups (Martinez-Gamero et al., 2021). LSD1 is further described in section 1.4.3. Since the discovery of LSD1, multiple other types of histone lysine demethylases (KDMs) have been identified (Qu et al., 2023), but these will not be further described in this thesis.

The histone modifications indicate that the genome is highly complicated, and the expression of genes is not as straightforward as it would seem, with all the mechanisms that control it; this is often referred to as the epigenome and epigenetic modifications. The epigenome represents a heritable layer of information not directly encoded in the DNA sequence of the genome but rather in the chemical modifications alongside transcriptional factors, operating as regulations of genome activity. This implies that epigenetic information comprises covalent chemical alterations, such as post-translational modifications of histone proteins and cytosine base methylation, which modify the structure and physiochemical properties of DNA and/or DNA-bound histones. Epigenetic modifications are inherently dynamic and are regulated by enzymes to transfer or remove the modifications. The combinations of modifications determine the epigenetic state of a genomic region, although the precise mechanisms underlying this code remain largely unclear (Brocken et al., 2018).

Taken together, non-coding regions of the genome harbor regulatory regions, such as enhancers, that can be identified by mapping specific histone modifications. Thus, epigenetic data can be used to narrow down disease-associated SNPs to those in active enhancers in relevant cell types.

1.3. Enhancers

The two most common regulatory elements are promoters and enhancers. Promoters are defined as DNA sequences that initiate transcription from a transcription start site. During transcription, RNA polymerase II (RNAP II) assembles at the promoter with the assistance of general transcription factors (TFs) (Kadonaga, 2012). While enhancers are segments of non-coding DNA elements that regulate gene transcription by the interaction with TFs (Carullo and Day, 2019). A key characteristic of the enhancer's ability to serve as integrated TF binding platforms, typically spanning a length of 200-500 bp (Calo and Wysocka, 2013). Enhancers contain clusters of TF motifs (Carullo and Day, 2019), enabling them to increase transcription by binding to a promoter (Spitz and Furlong, 2012). Collectively, Promoters and enhancer are crucial in transcription initiation, as promoters initiate RNA synthesis, while enhancers stimulate promoter activity of the target genes (Nguyen et al., 2016).

Enhancers are regulated by epigenetic modifications of the histone tails of the surrounding chromatin. Active enhancers exhibit a high density of H3K27 acetylation (H3K27ac) markers, which opens the chromatin and enables TF-binding (Carullo and Day, 2019) (Figure 1.3.1 B). In general, the activation of enhancers requires the presence of multiple TFs and sequence-dependent effectors of signaling pathways, which ensures the incorporation of intrinsic and extrinsic environmental signals at these elements. The ability of TFs to initiate transcription on chromatin templates relies on the recruitment of coactivator proteins (Calo and Wysocka, 2013). Conversely, inactive enhancers are highly methylated, resulting in a compacted chromatin structure, preventing TF binding (Carullo and Day, 2019) (Figure 1.3.1 A)
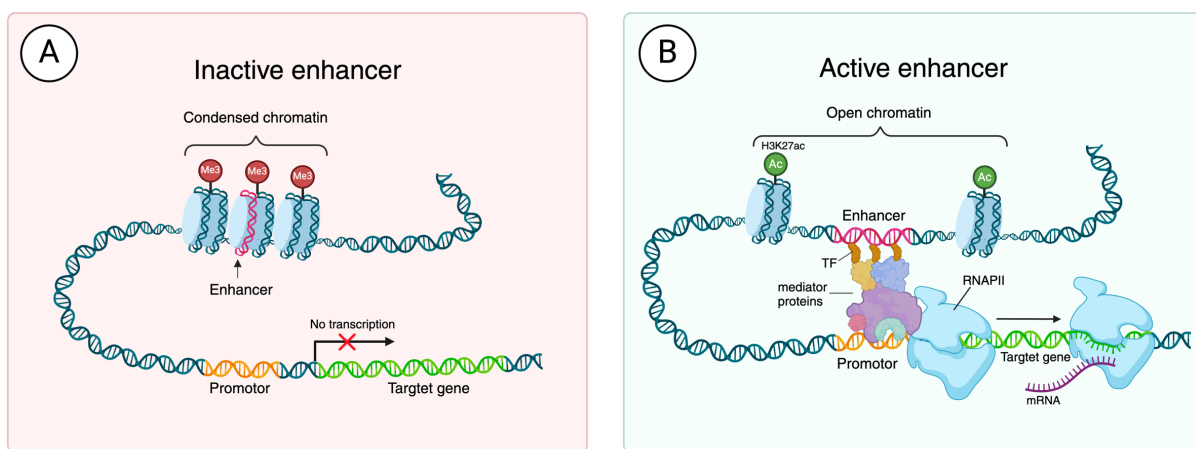
Figure 1.3.1: The epigenetic modulation of active versus inactive enhancers. In instances where enhancers are situated within compacted chromatin regions, they remain inaccessible for transcription factors (TFs) to bind, thus failing to facilitate the recruitment of RNA polymerase II (RNAPII) for transcription initiation. Conversely, enhancers located within open chromatin structures permit TF binding, as well as interaction with mediator proteins, thereby facilitating RNAPII binding and subsequent mRNA transcription.

Identifying enhancers and what genes they express is often challenging because, unlike promoters, enhancers can regulate genes that lie a great distance from the transcription site, and are orientation-independent, meaning the enhancer can be both upstream and downstream from the gene they regulate. In addition, enhancers are cell-type specific (Li et al., 2020, Banerji et al., 1981). The chromatin states are consistent across cell types, in contrast to histone modification, which are chemical changes to the histone protein; may vary between cell types (Heintzman et al., 2009). This may explain the various enhancer activities for different cell types as to why specific enhancers can be active in one and inactive in another. Since enhancer activity is linked to the presence and binding of TFs and chromatin accessibility. Consequently, the same enhancer can modulate different genes in different cell types (Ong and Corces, 2011, Panigrahi and O'Malley, 2021)

The majority of enhancers are still unidentified, and their target genes remain unknown (Carullo and Day, 2019, Lewis et al., 2019). So, identifying SNPs that impact this mode of gene regulation is thought to be one of the main impediments to linking GWAS association to mechanistic function. As mentioned, enhancer elements act in a cell-specific manner, governing tissue-specific gene expression (Carullo and Day, 2019). However, given that the DNA sequence remains constant across cell types, it does not provide insight into the spatial or temporal activity of the enhancer.

A considerable segment of the non-coding genome is dedicated to enhancers. Current estimates suggest that vertebrate genomes might encode anywhere from tens of thousands to potentially millions of active enhancers (Gray et al., 2015, Li et al., 2016). Therefore, it is not unexpected that most disease-associated SNPs are found in non-coding regions of the genome. A growing number of these SNPs have been linked to the functional activity of enhancers (Carullo and Day, 2019, Davidson et al., 2011, Voisin et al., 2015). In the recent ENCODE project, it was mapped out and annotated 400 000 putative human enhancers in different sets of human cell lines (Dunham et al., 2012, Calo and Wysocka, 2013). It is estimated that the human genome harbors over a million enhancer elements (Dunham et al., 2012).

One of the strategies to investigate predicted enhancers and their target genes involves epigenetic modulation, which entails either the activation or repression of the enhancer, followed by the evaluation of the expression levels of the proximal genes.

1.4. Epigenome editing – strategies to manipulate gene expression

1.4.1. CRISPR/Cas9

The continual development of technologies facilitates new techniques for altering the genome. Today, there are many methods of genome editing and gene engineering, where DNA can be inserted, removed, or modified in living cells. One of the most accurate methods of genome editing in living cells, and which is utilized in many applied disciplines, is the CRISPR/Cas9 system (Mengstie and Wondimu, 2021).

The CRISPR/Cas9 system is a gene-editing tool consisting of two components: a guide RNA (gRNA) to match the DNA region of interest, and a Cas9 endonuclease that binds the gRNA and thereby is directed to the desired DNA target. CRISPR/Cas9 was initially discovered within

the genome of the bacterial species *Streptococcus pyogenes*, where it serves as a defense mechanism against bacteriophages (Redman et al., 2016). It was discovered that bacteria stored 20-nucleotide fragments of previously encountered viral DNA in their genomes, referred to as "spacers", serving as a genetic memory of the viral infection. The CRISPR/Cas9 defense mechanism can attack when the same virus invades again by using the stored spacer sequence to recognize and destroy the invading viral DNA (Mengstie and Wondimu, 2021). It does this by producing a RNA molecule containing a copy of the spacer sequence which becomes associated with Cas9 and then guides Cas9 to the target DNA sequence, which is referred to as the "protospacer" and is identical to the spacer sequence of the RNA. To prevent the system from cleaving the bacterium's own DNA, spacers are only produced from protospacers that have a particular three-nucleotide sequence (NGG) directly downstream, called the protospacer-adjacent motif (PAM), and the Cas9/gRNA complex must bind to the PAM and protospacer before cleavage can take place (Mojica et al., 2009).

Since its discovery, CRISPR/Cas9 has evolved into a revolutionary tool for precise editing of the genome in mammalian cells. CRISPR/Cas9 genome editing involves three steps: recognition, cleavage, and repair (Figure 1.4.1). The Cas9/gRNA complex recognizes the target sequence through base pairing between the sgRNA spacer sequence and the complementary strand of the protospacer sequence in the target DNA. Cas9 has two nuclease domains that then induce a double-stranded break (DBS) at a site within the protospacer, three base pairs upstream of the PAM sequence. Once cleavage has taken place, the Cas9/sgRNA complex dissociates, and the cell uses endogenous systems to repair the break. This occurs through either non-homologous end joining (NHEJ) or homology-directed repair (HDR) pathways. NHEJ facilitates the DBS by joining DNA fragments enzymatically without exogenous homologous DNA and is active in all cell cycle phases. This method is an error-prone mechanism leading to mutation because of small random insertion or deletion at the cleavage site, leading to the generation of frameshift mutation or premature stop codon, and is, therefore, useful for disabling genes. HDR, on the other hand, can be used to make more precise edits, such as gene insertion or replacement. This method requires a donor DNA template containing the sequence to be inserted, which must be designed with homologous ends that match the sequences flanking the cut site of the target DNA, and thus the cell will incorporate the sequence at the site (Mengstie and Wondimu, 2021, Janik et al., 2020) (Figure 1.4.1).
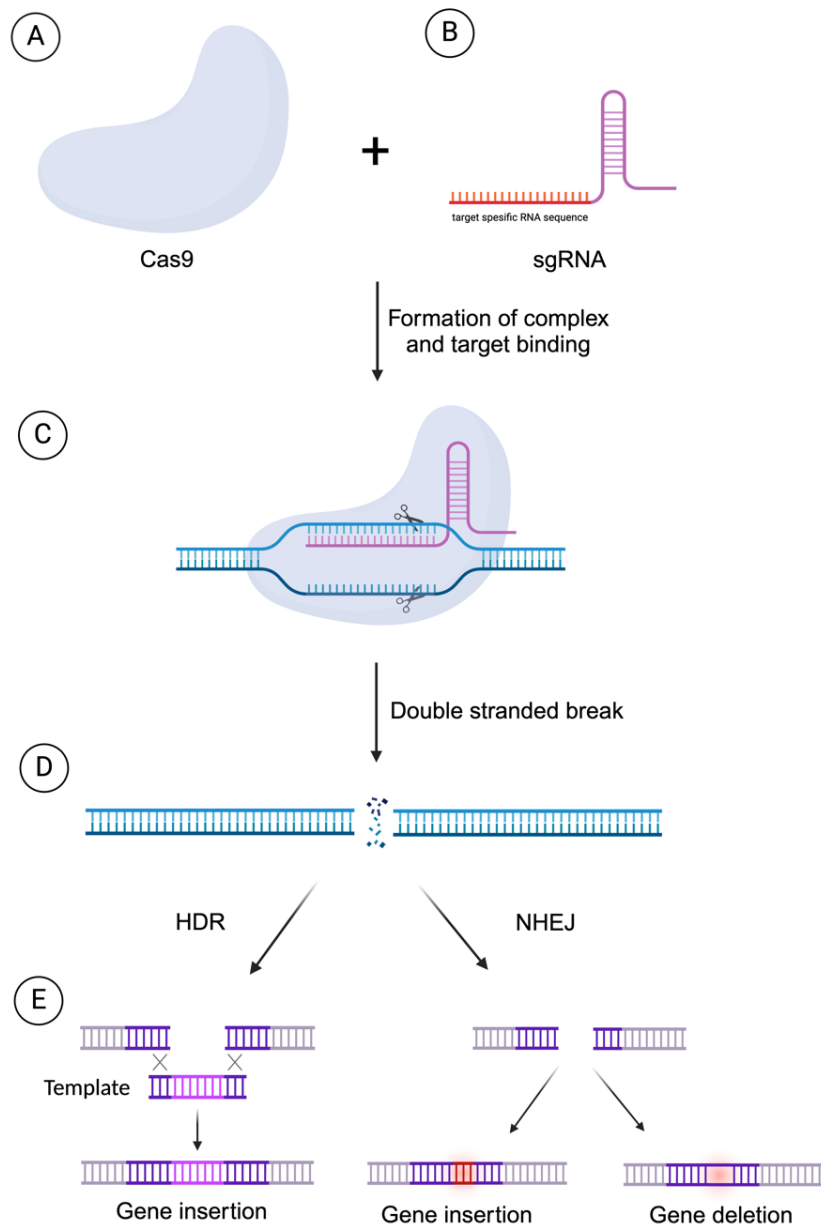
Figure 1.4.1: Schematic overview of the CRISPR/Cas9 system. The CRISPR/Cas9 system consists of two parts: the Cas9 protein (A) and the single guide RNA (sgRNA) (B). The Cas9 protein binds to the sgRNA and forms the complex, which binds to the specific DNA sequence coded by the sgRNA (C). The Cas9 cleaves both strands, three bases upstream of the protospacer-adjacent motif (PAM) sequence, creating a double-stranded break (DSB) (D). After the break, the cell tries to repair the DNA, either by homology-directed repair (HDR) pathway (E), if an appropriate donor is present, or by non-homologous end joining (NHEJ) (E). HDR repair allows accurate genome editing at the target site with a precise template, while NHEJ introduces short insertions or deletions. The figure was made in Biorender and adapted from Janik et al. (2020).

One remarkable feature of the CRISPR-Cas9 system is its possibility of customization. The gRNA can be designed to target any 20 nt, called the protospacer if it is upstream from the PAM sequence. The design of the protospacer is important as it guides the Cas9 to the target site; inaccurate design can lead to off-target results (Motoche-Monar et al., 2023). The gRNA also

includes a scaffold sequence, which is necessary for the binding of Cas9 Felt (Rainha et al., 2020(Rainha et al., 2020, Janik et al., 2020).

The major advantages of CRISPR/Cas9 compared to other techniques are that it is easier to use, cheaper, and more precise, minimizing unwanted gene edits at off-target sites (Janik et al., 2020) However, there are still many challenges in applying CRISPR, such as off-target effects, which are still a major concern. Furthermore, the efficiency of on-targets is low, especially when comparing HDR to NHEJ (Yip, 2020).

To address some of the limitations of the CRISPR/Cas9 system, several variations of the Cas9 endonuclease have been engineered. High-fidelity Cas9 nucleases have been engineered for higher efficiency and specificity, resulting in fewer off-target effects (Allemailem et al., 2023). Other types of the Cas9 nuclease with different properties have also been engineered, such as Cas9 nickase (nCas9), where one of the nuclease domains has been deactivated, meaning it only cuts the DNA on one strand. This variant is used in prime editing to precisely substitute specific base pairs in the genome with a reduced number of unintended errors because it does not rely on DSBs or utilize donor DNA templates (Anzalone et al., 2019, Janik et al., 2020). Another example is the catalytically deactivated/dead Cas9 (dCas9), which refers to the deactivation of the cleavage domains, meaning it does not cut either strand; it only targets and binds the DNA with high precision and specifically. dCas9 is often used in the recruitment of epigenetic modifiers at specific genomic locations to achieve specific modifications (Maeder et al., 2013).

The CRISPR/Cas9 complex can be delivered into cells by three primary methods: using physical, chemical, or viral vectors. Physical methods include electroporation, where electrical pulses are applied to the cells and disturb the cell membrane, which increases the temporary permeability of the membrane, enabling the delivery of the cargo of Cas9 and gRNA into the cells (Prasanna and Panda, 1997, Yip, 2020). Chemical methods involve lipid and polymer-based nanoparticles (Mengstie and Wondimu, 2021). Lipid nanoparticles package negatively charged nucleic acid within a positively charged lipid bilayer, which can be taken up by the cells via endocytosis (Duan et al., 2021). Polymeric nanoparticles functions similarly to lipid nanoparticles, but a polymer is used as the carrier for the cargo (Duan et al., 2021). The Viral vectors, which include adenoviral vectors, adeno-associated viruses (AAV), and lentivirus vectors, are particularly adept at in vivo CRISPR/Cas9 delivery. Viral vectors are often preferred for this purpose due to their superior delivery efficiency relative to physical or chemical methods (Mengstie and Wondimu, 2021). AAVs are the most used vector for transient

gene delivery (Daya and Berns, 2008, Yip, 2020). On the other hand, lentiviral vectors integrate the cargo nucleic acid into the genome of the host cell; the drawback is random integration (Yip, 2020). Still, a major challenge is the large size of the CRISP/Cas9 system; one solution is to pack the Cas9 and the sgRNA in separate viruses (Mengstie and Wondimu, 2021).

1.4.2. CRISPR/Cas9-mediated activation/inactivation

Recent advancements have revolutionized the modulation of endogenous gene expression by adapting the CRISPR/Cas9 system without altering the genetic composition of the DNA. This modification involves coupling the dCas9 to either activator or repressor domains, enabling targeted modifications in gene transcription (Allemailem et al., 2023, Mengstie and Wondimu, 2021) The CRISPR/Cas9-mediated activation/inactivation operates as a dual-effector system to manipulate a target enhancer by epigenetically altering the chromatin structure and by modulating interacting promoters. This involves the recruitment of transcriptional regulators on the specific site to promote or repress transcription. This approach relies on a Cas9-fusion protein to precisely target a specific genomic region of interest. In this context, dCas9 is fused with a chromatin remodeling domain, such as the H3K27 acetyltransferase p300 for enhancer activation or the H3K27 demethylase LSD1 for inactivation (Li et al., 2020). The second effector is associated with the system via the singe-guide RNA (sgRNA). The sgRNA is designed to incorporate two MS2 hairpins, a structure recognized by the MS2 coat proteins (MCP). For enhancer activation, the MCP is fused with a transcriptional activator VP64 domain, while for inactivation, the MCP is fused with the transcriptional repressor KRAB. In this thesis, the focus is on the inactivation of enhancers (Figure 1.4.2).
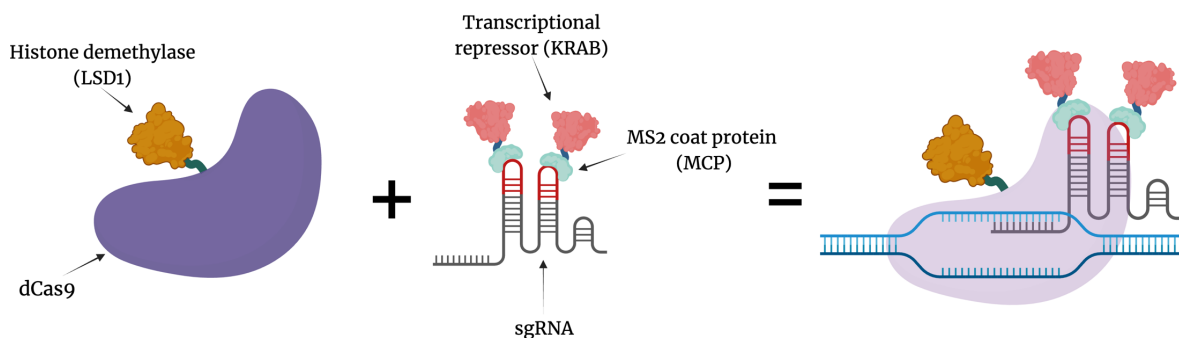
Figure 1.4.2: Schematic of the enhancer targeting CRISPR/Cas9-mediated inactivation: The system has two components. The first one is a catalytic deactivated Cas9 (dCas9) is fused with a histone demethylase LSD1. The sgRNA is bound to the MCP and is fused with the transcriptional repressor KRAB. The two components will form the complex and attach to the DNA, located by the specific target sequence of the spacer of the design sgRNA. Since the Cas9 endonuclease is catalytic deactivated, no DNA strand is cut. Figure made in Biorender

The Krüppel-associated box (KRAB) domain is a transcriptional repressor domain and part of zinc finger proteins. Zinc finger proteins are the most prominent family of transcriptional regulators. Through zin finger motifs, they bind to DNA recruit corepressors complexes via the KRAB domain and act as potent transcriptional repressor (Ying et al., 2015). KRAB is associated with heterochromatin formation (Li et al., 2020).

Lysine-specific histone demethylase 1 (LSD1), also referred to as KMD1a, was the first protein described to exhibit histone demethylase activity and contribute significantly to various essential functions in mammalian biology (Maiques-Diaz and Somervaille, 2016). LSD1 specifically catalyzes the demethylation of mono- and di-methylated H3K4 (Nagasaka et al., 2019). It is primarily recognized for its role in regulating lysine methylation status in both histone and non-histone proteins. LSD1 facilitates the demethylation of histone H3K4me1 and H3K4me2, thus leading to transcriptional repression, partly through the modulation of enhancer activity (Pinter et al., 2021). LSD1 can join protein complexes to influence chromatin structure. One of these complexes is the CoREST transcription complex along with HDACs and proteins. Although LSD1 can demethylate histones or peptides alone, it requires the LSD1-CoREST complex formation for activity within the nucleosome, which stabilizes LSD1 on its target region and mediates the demethylation of the surrounding nucleosomes. The demethylation of HeK4me1/me2 results in transcription repression because of the chromatin closed structure. (Martinez-Gamero et al., 2021). Consequently, LSD1 plays a crucial role in the epigenetic control of gene expression (Figure 1.4.3).
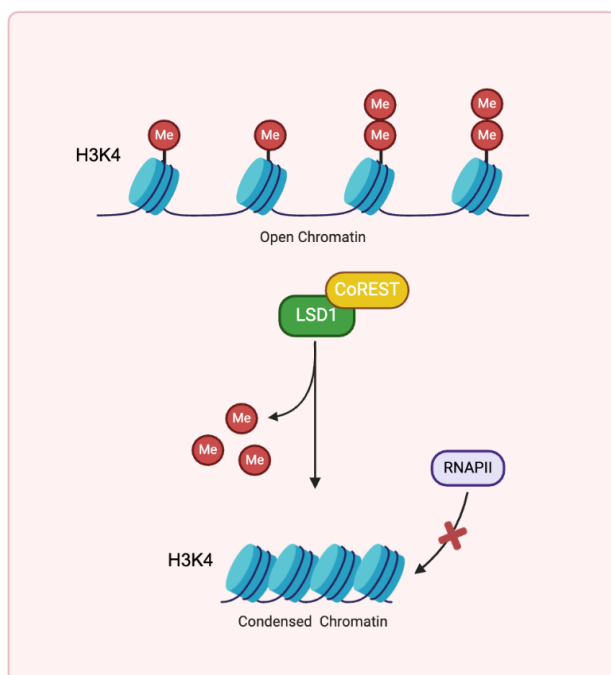
Figure 1.4.3: Transcriptional repression with LDS1 by histone demethylation. LSD1 interacts with CoREST or other protein complexes and catalysis H3K4 me1 and H3K4me2 demethylation, resulting in condensed chromatin structure and transcriptional repression of genes. Figure illustrated in Biorender, inspires by (Martinez-Gamero et al., 2021)

With the components outlined above, the enCRISPR/Cas9i system is assembled. Similar to conventional CRISPR/Cas9 systems, the LSD1-dCas9 fusion protein forms a complex with the single-guide RNA (sgRNA), directing it to the vicinity of the target binding site within the enhancer. Upon binding, the LSD1 enzyme fused with Cas9 initiates demethylation of H3K4 by engaging in complexes with other proteins, resulting in a condensed chromatin structure surrounding the enhancer and subsequent inactivation. Another repression mechanism involves KRAB, which is attached to the sgRNA. Together with zinc-finger motifs, KRAB binds to the DNA, serving as transcriptional repressors for the target genes of the enhancer. (Figure 1.4.4)
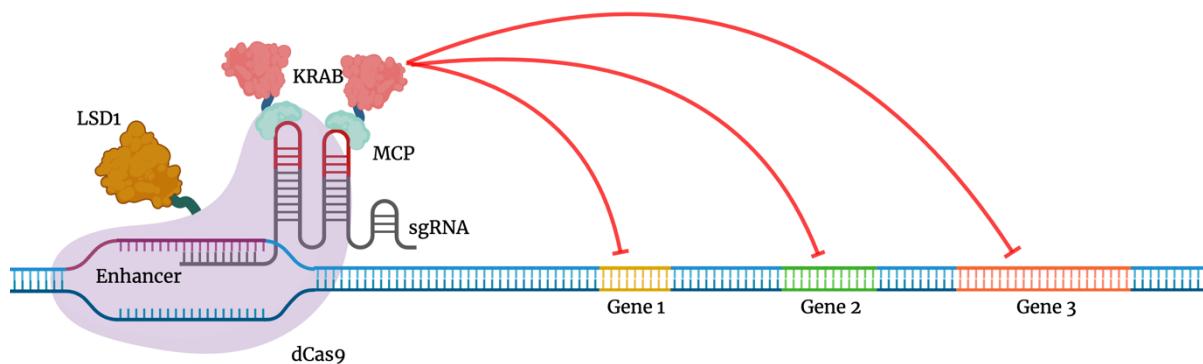
Figure 1.4.4: the CRISPR/Cas9 mediated enhancer inactivation in action. The sgRNA, featuring a specific design spacer, directs the dCas9 to the binding site on/near the enhancer. Upon binding, the histone demethylase LSD1 contributes to enhancer repression by demethylation of H3K4, making the enhancer unavailable in the closed-up chromatin. Additionally, the transcriptional repressor KRAB acts to suppress the transcription of genes targeted by the enhancer. It is noteworthy that enhancers can exert an effect on multiple genes. Figure illustrated in Biorender, adapted from Li et al. (2020)

1.4.3. Inducible enhancer CRISPR/Cas9 inactivation

Enhancer activity is often restricted not only to certain cells and tissues, but also to specific developmental stages (Li et al., 2020, Banerji et al., 1981). To experimentally test when an enhancer is active during cellular differentiation, precise control over the timing of epigenetic repression is essential, which can be achieved using an inducible system. This allows for turning on the inactivation at specific time points during differentiation.

The Tet system is one of the most widely used inducible mammalian expression systems available (Loew et al., 2010). It is used to regulate the activity of genes in eukaryotic cells (Das et al., 2016). Tet-On-3G is a modified reverse tetracycline-controlled transactivator (rtTA) protein for the Tet-on expression system (Ying et al., 2015). The Tet-On-3G protein consists of

tetracycline (-responsive) repressor (TetR) proteins (Das et al., 2016). The Tet-on system is based on its sensitivity to doxycycline (dox), a tetracycline analog, and its ability to bind tightly to a promoter with a Tet operator (TetO), also known as Tet response element (TRE) (Ying et al., 2015). This means that the Tet-On-3G rtTA protein, along with a gene of interest under the control of a TRE3G promoter (P_{TRE3G}), will exhibit elevated expression levels of the gene of interest only in the presence of dox, as shown in Figure 1.4.5 .
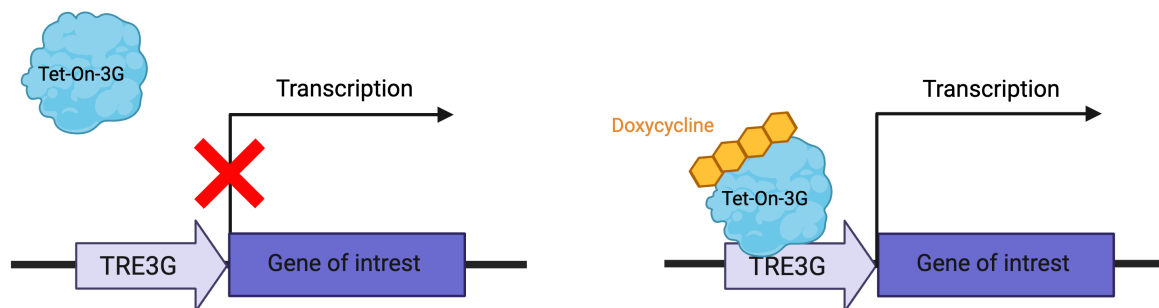
Figure 1.4.5: an overview of the Tet-On-3G system. The transcription of the gene of interest is placed downstream of the TRE3G promoter, can only be processed when Tet-On-3G and doxycycline are present. Tet-On-3G only binds to the promoter when doxycycline is bound to it. This means that the Tet-On-3G inducible system regulates the gene expression of TRE3G promoter-controlled genes with Tet-On-3G and the presence of Doxycycline. Figure illustrated in Biorender.

1.5. The 11q23.3 locus associated with visceral adiposity

Over the past century, there has been a significant rise in the global prevalence of obesity and overweight, resulting in increasing health issues across the world (Ezzati et al., 2017). Visceral adipose tissue (VAT) is fat stored around the internal organs and has been shown to be more harmful compared to subcutaneous fat, which is stored directly under the skin (Kwok et al., 2016, Shuster et al., 2012). Specifically, VAT is associated with an increased risk of metabolic diseases, such as type 2 diabetes, cardiovascular diseases, and cancer, as well as all-cause mortality (Brown et al., 2017, Fox et al., 2007, Vega et al., 2006). And yet, the contribution of genetics to VAT and its disease-related effects is mainly unexplored due to the lack of advanced imaging tools to measure VAT accurately (Karlsson et al., 2019).. However, in 2019, a GWA Study was published where the authors developed a computational model to predict VAT mass based on information available through the UK Biobank (UKBB). With data from 400 000

individuals, the study identified over 200 nonoverlapping loci associated with visceral obesity (Karlsson et al., 2019).

One of the novel VAT-mass-associated GWAS signals was found in locus 11q23.3, which contains three genes; *HYOU1, VPS11* and *HMBS*. However, the reported (tag) SNP and the 38 SNPs co-inherited together (any of which could be causal) in a so-called LD-block, are situated in non-coding regions (Samuelsen, 2021). Thus, the target gene(s) and downstream mechanisms are still unknown, although *HMBS* has been proposed as one possible target gene (Karlsson et al., 2019). Specifically, Karlsson et al. (2019) found the SNP rs1799993 to have a modest but significant effect on promoter activity of a short isoform of *HMBS*. However, since this isoform is only expressed in blood (Chretien et al., 1988, The UniProt Consortium, 2022), it is not likely to affect VAT mass. Instead, since rs1799993 is situated in the first intron of the long isoform of *HMBS*, our research group hypothesizes that it may be situated in an enhancer rather than a promoter. Thus, both the long isoform of *HMBS*, as well as the other nearby genes could be regulated by such an enhancer. Indeed, eQTL data show that expression of both the *HMBS* and *VPS11* genes are affected by the genotype of rs1799993 in both visceral and subcutaneous adipose tissues (Samuelsen, 2021), suggesting that both of these genes are likely regulated by the enhancer.

Moreover, since enhancers have been shown to regulate genes hundred thousand to more than a million base pairs away (Claussnitzer et al., 2015, Ragvin et al., 2010, Schoenfelder and Fraser, 2019), approximately 50 genes could be within reach of an enhancer in the 11q23.3 locus (Samuelsen, 2021). Therefore, to discover the biological mechanisms underlying the association between variants in 11q23.3 and visceral obesity, there is a need for determining whether rs1799993 is indeed situated in an enhancer and whether there are other enhancers harboring any of the other co-inherited SNPs. Furthermore, experimentally testing what gene(s) are in fact controlled by these enhancers is required to unravel the underlying mechanisms. This thesis is part of a larger project aiming to address these issues.

Previous work in the lab has determined that the LD block in 11q23.3 that is associated with visceral obesity consists of 38 SNPs in a 30 kb region. Moreover, Roadmap epigenetic data revealed the presence of three predicted enhancers in adipose-derived mesenchymal stem cells in this region (Samuelsen, 2021). To investigate whether these enhancers were functional, genomic regions (named tiles in this thesis) corresponding to these enhancers were cloned from patients with homozygous risk or protective variants in the 11q23.3 locus. The tiles were cloned

into a luciferase reporter plasmid vector containing a minimal promoter and tested for genotype-dependent enhancer activity (Mirza, 2022, Samuelsen, 2021). Through this approach, the LD block was narrowed down to a single 1.8 kb tile (Tile 8a) containing only five SNPs. These five SNPs were further analyzed by using site-directed mutagenesis of the reporter plasmid *in vitro*, followed by luciferase analysis, and it was found that each SNPs individually contributed to an increase in enhancer activity when altered from the VAT mass-associated protective variant to the risk allele, suggesting that all five are potential causal SNPs (Mirza, 2022).

Importantly, one of these five potential causal SNP in Tile 8a was rs1799993, demonstrating that this SNP, which Karlsson previously found to affect promoter activity as described above, is indeed also situated in an enhancer and affecting it activity. Thus, the next step in uncovering the mechanisms underlying the association between rs1799993 and visceral obesity is to experimentally identify the target gene(s) of this enhancer using CRISPR/Cas9-mediated epigenetic repression.

1.6. Aims and objectives

This thesis is part of a larger project aiming to map the mechanisms underlying the association between gene variants in the 11q23.3 locus and their associated risk of increased visceral obesity. A subset of these variants, including rs1799993, has been found to reside in an enhancer and the risk alleles have been found to increase the enhancer activity *in vitro*. Epigenetic data suggest that the enhancer is active in mesenchymal stem cells derived from fat (AdMSCs), but it is not known which gene(s) are regulated by this enhancer.

The main aim of this thesis is to establish an inducible CRISPR/Cas9 mediated epigenetic repression system in AdMSCs and use it to repress the enhancer in 11q23.3 to identify its target gene(s).

Specific objectives include:

1. Test the functionality of the system's plasmids through transient transfection.
2. Produce lentivirus carrying Tet-On-3G, LSD1-dCas9, and sgRNA-KRAB constructs, respectively.
3. Sequentially transduce AdMSCs with the different lentiviruses and expand positive clones.
4. Treat triple-induced AdMSC cells with doxycycline at various stages of adipocyte differentiation and measure gene expression to identify the enhancer's target gene.

2. Materials

Table 2.1: Plasmids

Plasmid Name	Size (kb)	Supplier	Catalog nr.	Description
Lenti_**sgRNA**(MS2)_MCP-**KRAB-IRES**-zsGreen1	12.1	Addgene	138460	lenti sgRNA cloning backbone with MS2 loops and EF1a-MCP-KRAB. Ampicillin resistance
pCMV6-XL4/5/6-**empty**	4.7	OriGene	PCMV6-XL4	An empty vector used as control plasmid.
pCMVR8.74	11.9	Addgene	22036	2nd generation lentiviral packaging plasmid with TAT.
pHR_TRE3G-**LSD1-dCas9**-P2A-mCherry	16.4	Addgene	138462	2nd Generation Lentiviral vector. Expresses an N-terminal LSD1-dCas9 fusion protein and mCherry. Plasmid sent in bacteria as agar stab. Ampicillin resistance
pLVX-**Tet3G**_blasticidin	8.6	Addgene	128061	Lentiviral vector encoding the Tet-On-3G transactivator protein, which allows for inducible gene expression. Ampicillin resistance.
pSFFV-**Tet-On-3G_BFP**	10.5	Bo Huang, forwarded by Jian Xu		Expresses Tet-on 3G from the EF1Alpha promoter.
VSV-G	Unknown	A kind gift from Pouda Panahandeh (Strømland)		Envelope Plasmid A, for lentiviral production. Compatible with both 2nd and 3rd generation lentiviral transfer vectors.

Table 2.2: Plasmid Purification Components

Component	Supplier	Catalog nr.
Ampicillin. Sodium salt	CALBIOCHEM	171254
Bacto™ Agar	BD Biosciences	214050
Difco™ LB (Luria-Bertani) Broth Miller	BD Biosciences	244620
Glycerol	Merk	104094

Table 2.3: Cell Lines

Cell line	Supplier	Catalog nr.	Description
ASC52telo (ASC52)	ATCC	SCRC-4000	An hTERT immortalized adipose-derived Mesenchymal stem cell line (AdMSC), isolated from the adipose tissue of a white female.
HT-1080	ATCC	CCL-121	An epithelial human cell line derived from connective tissue from a Fibrosarcoma patient. Diploid and easy to transfect
HEK293T	Gifted from Nils Halberg, University of Bergen, Norway.	-	Human embryonic kidney cells and express the simian virus 40 large T antigen, promoting a high expression of protein (Iuchi et al., 2020). Easily transfectable

Table 2.4: Cell culture components

Component	Supplier	Catalog nr.
Accutase	Merk Millipore	SCR005
Amphotericin B	Sigma-Aldrich	A2942
CoolCell™ LX Cell Freezing Container	Corning®	CLS432001
Dimethyl Sulfoxide (DMSO)	EMD Millipore	317275
DMEM High Glucose	EuroClone	ECM0728L
Doxycycline	Sigma Aldrich	D5207
Dulbecco's Phosphate Buffered Saline (PBS)	Sigma-Aldrich	D5652-10X1L
EGF	Cell signalling technologies	72528S
Fetal Bovine Serum (FBS)	Gibco™	10270-106
Fetal Bovine Serum, Tet system approved, USDA-approved regions, OneShot™ format	Gibco™	A4736301
FGF 2 BASIC	MERK	SRP4037
G418	InvivoGen®	108321-42-2
Gentamicin	Sigma-Aldrich	1405-41-0

MEM		
Mesenchymal Stem Cell Basal Medium	ATCC®	PCS-500-030
Mesenchymal Stem Cell Growth Kit for Adipose and Umbilical-derived MSCs – Low serum	ATCC®	PCS-500-040
Nunc™ Cell Culture Dishes, Ø 10 cm	Thermo Scientific	168381
Nunc™ MicroWellTM 96-Well, Nunclon Delta-Treated, Flat- Bottom Microplate	Thermo Scientific	167008
Penicillin and Streptomycin (PEST)	LONZA	17-745E
Trypan blue stain (0.4%)	Invitrogen™	T10282
Trypsin-EDTA (0.25 %), with Phenol Red	Gibco™	25200056

Table 2.5: Components for Virus Production and Transfection reagents

Component	Supplier	Catalog nr.
Bovine serum Albumin (BSA)	Sigma-Aldrich	A7030
RetroNectin® Recombinant Human Fibronectin Fragment	Takara bio	T100A
Lenti-X™ GoStix™ Plus	Takara bio	631280
Lenti-X™ Concentrator	Takara bio	631232
TansIT-LT1	Mirus	MIR 2305
OptiMEM	Gibco™	31985062
Lipofectamine Stem Cell	Invitrogen	STEM00001

Table 2.6: Primers for sequencing and PCR

NAME	Sequence (5' -> 3')	Binds to	Application
AK61_dCas9_fwd3	TCACCGACGAGTACAAGGTG	dCas9	qPCR of RNA and sequencing
AK62_dCas9_rev3	CTGTTTCTCCGCTGTCGAAC	dCas9	qPCR of RNA and sequencing
AK63_dCas9_fwd4	TCACCGACGAGTACAAGGTGC	dCas9	qPCR of RNA and sequencing
AK64_dCas9_rev4	CGCCGATCAGGTTCTTCTTGAT	dCas9	qPCR of RNA and sequencing

AK69_Tet3G_fwd	TAAAGTGCATCTCGGCACCC	Tet-On-3G	qPCR of RNA and gDNA, and sequencing
AK70_Tet3G_rev	TGTTCCTCCAATACGCAGCC	Tet-On-3G	qPCR of RNA and gDNA and sequencing
MMP9 upstream fwd	CTTCAGAGCCAGGCAGTTCT	MMP9	gDNA reference position for qPCR
MMP9 upstream rev	AGCCTCTCGTTTCATCCTCA	MMP9	gDNA reference position for qPCR
PRLPO rev	CAATCTGCAGACAGACACTGG	PRLPO	mRNA (cDNA) reference for qPCR
RPLPO fwd	TCTACAACCCTGAAGTGCTTGAT	PRLPO	mRNA (cDNA) reference for qPCR
RPS13 fwd	TTCACCGTTTGGCTCGATA	RPS13	mRNA (cDNA) reference for qPCR
RPS13 rev	AAACAATCATTTTATTGCTTGAGTACA	RPS13	mRNA (cDNA) reference for qPCR
TN1_zsGreen_fwd	GCGAATCACTCTCATCTTTGGG	zsGreen1	qPCR of RNA
TN2_zsGreen_rev	CTTCAAGAACTCCTGCCCCG	zsGreen1	qPCR of RNA
TN3_MCP_fwd	ACGCCGTAGAACTTGGACTC	MCP	qPCR of RNA
TN4_MCP_rev	GTCGCCCCAAGCAACTTC	MCP	qPCR of RNA
TN5_TRE3GV_fwd	GCTCTGACGAACGCTACAGG	TRE3G promoter	qPCR of RNA
TN6_TRE3GV_rev	GTATGTCGAGGTAGGCGTGT	TRE3G promoter	qPCR of RNA
16q12.2_ rev	AAATACAGTCAGGTAGGTTTCGT	Genomic region at position 16q12.2,	gDNA reference position for qPCR
16q12.2_fwd	TTTTCCTTCTCTCCTTCTGCCC	Genomic region at position 16q12.2, forward	gDNA reference position for qPCR
141F_VP1.5_fwd	GGACTTTCCAAAATGTCG	pCMV6-XL plasmid	Sequencing

Table 2.7: Spacer for enhancer targeted CRISPR/Cas9-mediated inactivation

NAME	Sequence (5' -> 3')	PAM	Distance to the SNP rs1799993 (nts)
Spacer-01	GAAACCAGAGGCAGCACTCT	AGG	9
Spacer-02b	GTCCATGTTGGTGAAACCAG	AGG	21
Spacer-04a	ACTGCATGCATTAAGGACTG	AGG	173
Spacer-10b	GTAGTCTTATGGAACATCTG	AGG	42

Spacers design by Krill (2023)
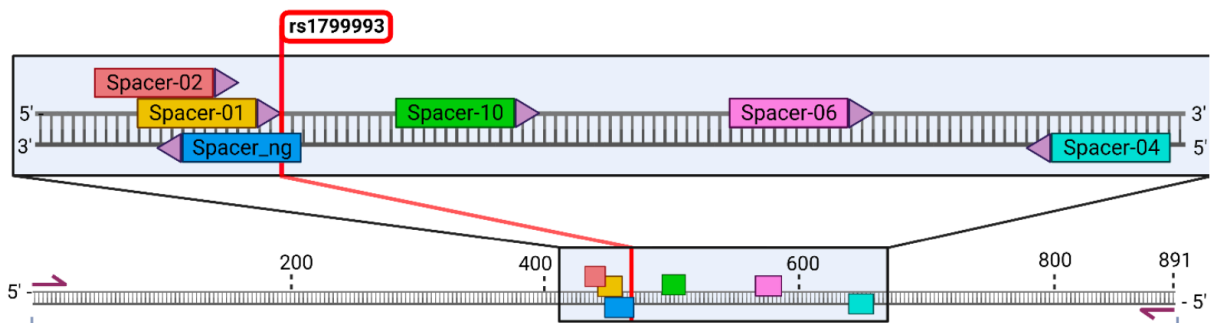
Figure 2.7.1: an overview of the location of the spacers relative to the SNP rs1799993 in the locus 11q.23.3. Spacers sequences were designed by Krill (2023) surrounding the causal SNP. Those were amplified to the gRNA plasmid Lenti_sgRNA(MS2)_MCP-KRAB-IRES-zsGreen1 for enhancer-targeting CRISPR/Cas9 inactivation system. Spacer_ng was designed for another system, and Spacer-06 was not available. Figure is in courtesy of Krill (2023).

Table 2.8: Components for FACS

Component	Supplier	Catalog nr.
96 F microwell plate w/out lid flat bottom	Thermo Scientific	269620
5 ml Polystyrene round-bottom tube with cell strainer cap	Corning	352235
Automatic Setup beads	SONY	LE-B3001
Sorting Chip 100 µm	SONY	LE-C3210

Table 2.9: Kits

Kits	Supplier	Catalog nr.	Application
BigDye™ Terminator v3.1 Cycle Sequencing Kit	Applied Biosystems™	4337455	Sequencing of DNA
High-Capacity cDNA Reverse Transcription Kit	Applied Biosystems™	43-688-14	cDNA synthesis
HiSpeed Plasmid Maxi Kit	Qiagen	12663	Purification of plasmid
RNA/DNA/protein purification plus kits	NORGEN biotek	47700	Purification of RNA/DNA
RNeasy Mini Kit	Qiagen	74104	RNA purification
SYBR™ Green PCR Master Mix	Applied Biosystems™	4309155	qPCR

Table 2.10: Antibodies

Antibody	Supplier	Catalog nr.	Host species	Description
Anti-beta Actin antibody	abcam	ab6276	Mouse	Monoclonal antibody (used as control)
Anti-CRISPR-Cas9	abcam	ab191468	Mouse	Monoclonal antibody
Anti-KDM1/LSD1	abcam	ab17721	Rabbit	Polyclonal antibody
Anti-TetR a	abcam	ab302642	Mouse	Monoclonal TETR2 antibody
HRP Goat Anti-Mouse Ig	BD Biosciences	554002	Goat	Polyclonal antibody (Secondary antibody)
Goat anti-Rabbit IgG HRP	Invitrogen	31460	Goat	Polyclonal antibody (Secondary antibody)
HMBS polyclonal	Invitrogen	PA5-102482	Rabbit	Polyclonal antibody

Table 2.11: Components for protein quantification and Western blotting

Component	Supplier	Catalog nr.
4 x SDS Sample Buffer	VWR	70607-3
96 F microwell plate w/out lid flat bottom	Thermo Scientific	269620
Bovine serum Albumin BSA	Sigma-Aldrich	A7030
DC™ protein Assay	Bio-Rad	5000112
Dulbecco's Phosphate Buffered Saline (PBS)	Sigma-Aldrich	D5652-10X1L
eSpCas9 protein	Sigma-Aldrich	ESPCAS9PRO
MagicMark™ XP Western ProteinStandard	Invitrogen	10610856
Ponceau S	MP Biomedical	190644
Precision Plus Protein™ Kaladidoscope™ Prestained protein standards	Bio-Rad	161-0375
Restore™ Western Blot Stipping buffer	Thermo Scientific	21059
SuperSignal™ West Femto Maximum Sensitivity Substrate	Thermo Scientific	34096
SurePAGE™, Bis-Tris, 10x8, 4-20%, 15 wells	GenScript	M00657
Trans-Blot Transfer Pack, 0.2 ¼ Midi, nitrocellulose	Bio-Rad	170-4159
Tris-MOPS-SDS Running buffer powder	GenScript	M00138
Tween 20	Bio-Rad	1706531

Table 2.12: Instruments

Instrument	Supplier	Application
Countess II Automated Cell Counter	Invitrogen	Estimation of cell number
GeneAmp® PCR System 9700	Applied Biosystems	PCR: sequencing and cDNA synthesis
LightCycler 480	Roche	qPCR: quantification of RNA
NanoDrop® ND-1000 Spectrophotometer	Thermo Scientific	Quantification of nucleic acid concentration
Nikon TE2000	Nikon	Fluorescence microscope for cell imaging
QIAcube	QIAGEN	Purification of RNA
QIAxpert	QIAGEN	Quantification of nucleic acid concentration
SONY SH800 Cell Sorter	SONY	FACS, florescent sorting of cells

SpectraMax plus 384 microplate reader	Molecular Devices	Quantification of protein concentration
Trans-Blot® Turbo™ Transfer System	Bio-Rad	Western blot, transfer of protein from gel to membrane
Universal Hood II Gel Doc XR system w/ ChemiDoc XRS system	Bio-Rad	Western blot, membrane imaging
Nikon Ecpilse TS100 Microscope	Nikon	Observing and imaging cell culture

Table 2.13: Online tools

Name	Purpose	URL
Beacon Designer	Evaluating primers	https://www.premierbiosoft.com/qOligo/Oligo.jsp?PID=1
Biorender	Create scientific illustrations	https://www.biorender.com
BLAT	Sequence search in UCSC genome browser	https://genome.ucsc.edu/cgi-bin/hgBlat
NCBI Primer-BLAST	Designing primers	https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?GROUP_TARGET=on
OligoEvaluator	Evaluation of primers	http://www.oligoevaluator.com/
UCSC Genome Browser	Genome browser	https://genome.ucsc.edu/

Table 2.14: Software

Name	Purpose
FlowJo v.10.10	Analyses of FACS data
Graphpad Prism v.10.2.2	Make graph and statistical analyses
Inkscape v.1.3	Annotate figures
TakaRa GoStix™ Plus	Reading of the GoStix cassettes

3. Methods

3.1. Prepping and purification of plasmids

Several plasmids were used in this thesis (Table 2.1), and most of these were already available in the lab. However, two of the plasmids (pCMV8.74 and pLVX-Tet3G_blasticidin) were recently purchased from Addgene and required prepping and purification prior to use. These plasmids arrived pre-transformed in E. coli in agar stabs. A small portion of each stab was inoculated onto an agar plate with Ampicillin [100 μg/mL] and incubated at 37°C o/n. The day after, a single colony was picked and incubated o/n at 37°C in 5 ml LB medium with shaking at 225 rpm. The culture was then added to 150 ml LB medium and again incubated o/n at 37°C whilst shaking at 225 rpm. A small part of the culture was made into glycerol stock by adding 20% glycerol (f.c) and transferred to cryotubes, flash frozen in nitrogen, and then stored at -80°C long long-term storage.

The plasmids in the rest of the culture were immediately purified using the Qiagen HiSpeed kit according to the manufacturer's protocol. In short, the culture was first centrifuged at 6000 rpm for 15 min. at 4°C to pellet the bacteria. The pellet was resuspended in P1 buffer, mixed with P2 buffer by gentle inversion, and incubated for 5 min, then mixed with P3 buffer as before, transferred to the QiaFilter cartridge, and incubated for 10 min. The HiSpeed tip was equilibrated with QBT buffer. The lysate was filtered through the HiSpeed tip, the HiSpeed tip was washed with QC buffer, and the DNA was eluted with QF buffer through the HiSpeed tips. Isopropanol was added to the eluted DNA and incubated for 5 min. The eluate was filtered through a syringe attached to a QIAprecipitator, and 70% ethanol was also filtered to wash the DNA. To elute the DNA, the QIAprecipitator was reattached to a new syringe, 500 μl TE buffer was added, and constant pressure was applied. The eluate was filtered through the QIAprecipitator a second time for the optimal yield. The plasmid DNA concentration was measured with NanoDrop.

3.2. Sequencing of plasmids

Plasmids were sequenced to verify if the plasmid contained the insert. BigDye™ Terminator v3.1 Cycle Sequencing Kit was used by mixing 1X Sequencing buffer, 10% (v/v) Big Dye Ready Reaction Mix (enzyme), 3.2 pmol primers (fwd or rev), 200 ng DNA template with a total reaction volume of 10 μl. for each template, fwd and rev primers were run separately, with two replicates of each. The reactions were run in the thermocycler for 30 sec as initial

denaturation at 96℃, then 35 cycles of denaturation for 10 sec at 96℃, annealing for 5 sec at 50℃, and elongation for 4 min. at 60℃. Following completion of the run, the samples were delivered to the Core facility at the Department of Medical Genetics, Haukeland University Hospital. The returned result was viewed and analyzed with SnapGene viewer v7.2.

3.3. Cell culture

The human ASC52telo adipose-derived mesenchymal stem cell line was the main cell line of this thesis and the target for lentiviral-mediated CRISPR-Cas9 enhancer inactivation. HT-1080 cells were used for transient transfection of lentiviral cargo plasmids to test the function of each plasmid prior to use in lentiviral production. Finally, HEK293T cells were used for transient transfection of lentiviral packaging and cargo plasmids to produce lentivirus.

3.3.1. Cultivation of cells

The different cell lines were cultivated in cell-specific medium (Table 3.3.1) in 10 cm culture-grade dishes and maintained in a humidified cell incubator at 37 ℃ with 5 % CO_2. The growth medium was changed every 3-4 days. The cells were subcultured when reaching 70-80% confluence. To dissociate the cells, ASC52telo and HT-1080 cells were enzymatically treated with accutase or trypsin, respectively, while the HEK293T cells could be physically detached by washing over the plate with the medium. The cells were subcultured by transferring a subset of the detached cells to a new culture dish with the appropriate volume of growth medium. Usually, the cells were counted before seeding, with seeding densities of 5000 cells/cm^2 for ASC52telo cells and 1500-2500 cells/cm^2 for HT-1080 cells. These seeding densities usually translated to 1:4-1:5 and 1:10 dilutions, respectively, and sometimes the cells were seeded at these dilutions without counting. HEK293T cells were subcultured at 1:10 or 1:20 dilutions. After seeding, the cells were sometimes placed in RT for 20 min to allow the cells to attach to the dish without clumping together, preventing uneven distribution of cells in the dish. Cells were generally maintained for a maximum of 10 passages, with the notable exception of transduced cells, which underwent additional passages in the process of generating new clonal cell lines.

Table 3.3.1. Composition of cell-specific growth medium

Cell line	Basal medium	Supplements[a]
ASC52	MSC Basal Medium	MSC Growth Kit: - 2% (v/v) FBS, - 5 ng/ml rhFGF basic - 5 ng/ml rhFGF acidic, - 5 ng/ml EGF - 2.4 mM L-Alanyl-L-Glutamine 0.25 µg/ml Amphotericin B 1% (v/v) PEST[b] 100 µg/ml Gentamicin[c] 200 µg/ml G418
HT-1080	MEM	10% (v/v) FBS 1% (v/v) PEST[b] L-Glutamine
HEK293T	DMEM High Glucose with L-Glutamine	10% (v/v) FBS 1% (v/v) PEST[b]

[a] final concentration
[b] Penicillin/streptomycin
[c] added in the last half of the project

3.3.2. Counting cells

Cells were counted on the Invitrogen Countess 2 Automated Cell Counter. Briefly, cell suspensions were centrifuged for 3 min at 300 g and resuspended in 1 ml medium. A small amount was used for counting by mixing with an equal amount of trypan blue. Each cell suspension was counted twice, and the average number of live cells was used in calculations.

3.3.3. Freezing cells

To freeze cells, the cells were harvested, counted, and resuspended in growth medium containing 10% (v/v) DMSO at a cell concentration of 250,000-500,000cells/mL. 1 mL aliquots were transferred to 1.5 mL cryotubes and placed in a CoolCell Container at -80°C o/n for gradual freezing before transferring to the liquid nitrogen tank for long-term storage.

3.3.4. Thawing cells

Cryotubes with cells were thawed by removing them from the liquid nitrogen tank, quickly thawed at 37°C (max. 2 min.), and immediately diluted 1:5 in medium, followed by

centrifugation at 300 g for 3 min to remove most of the DMSO. The cell pellet was resuspended in regular growth medium, and the cell suspension transferred to a 10 cm culture dish, and cultured as described above.

3.4. Transient Overexpression

Before making lentiviruses, transient overexpression of the lentiviral cargo plasmids was performed in the HT-1080 cell line to test the functionality of the plasmids. The HT-1080 cell line was chosen because it is much easier to transfect compared to the ASC52telo cells

The cells were seeded out one day before transfection at a cell density of about 13 000 per cm^2 in 24-well plates in tetracycline-free medium so the cells were about 80% confluent on the day of transfection. The experiment was performed by transfecting cells with 0.2 µg each of lentiviral cargo plasmids harboring Tet3G-BSD and/or LSD1-dCas9-mCherry. The pCVM6-empty plasmid was used as negative control and to keep DNA amounts constant at 0.4 µg total. Two different transfection reagents were tested, Lipofectamine Stem and TransIT-LT1, both with a 1:3 ratio of DNA and transfection reagent. For each reaction, the plasmid DNA was first mixed with 50 µl Opti-MEM before addition of the respective transfection reagents. The transfection solutions were mixed gently, incubated at RT for 20 min, and added dropwise into each well containing 450 µL culture medium and HT-1080 cells. After transfection, the cells were incubated at 37℃ for two days. With the Nikon TE2000 fluorescent microscopy, the mCherry expression was assessed. The filter TRITC with excitation 528-552 nm and emission 577-523 nm was used for this purpose. Then RNA was harvested by washing the cells with 500 µL ice-cold PBS and then lysing the cells with 350 µl RTL buffer from the RNeasy mini Kit, followed by pipetting up and down and scraping. The lysed samples were then immediately placed on ice and either directly purified using the RNeasy Mini kit on the QIAcube or stored at -80℃ for later purification.

3.5. Analysis of gene expression

3.5.1. RNA Purification

To isolate the RNA from lysed cells, the RNA was purified by the QIAcube with the RNeasy Mini Kit. If the samples were frozen, the samples were thawed at 55℃ for a maximum of 5 min. The samples were loaded into the QIAcube along with the components of the RNAeasy Mini Kit, and the QIAcube protocol was followed as instructed, all steps in the purification procedure are fully automated. When the samples were done, they were put directly on the ice to keep the integrity of the RNA. Then, the concentration of RNA was measured with the QIAxpert and continued to qPCR and further analysis.

3.5.2. cDNA synthesis

To run qPCR, purified RNA must be made into cDNA, and this is done by cDNA synthesis by reverse transcription of the RNA template, with the help of polymerase chain reaction (PCR). The reaction was performed with the AB High-Capacity cDNA Reverse Transcription Kit with an input of 100 ng. One reaction contains 1X RT buffer, 1X dNTP mix, 1X RT Random Primers, and 2.5 U MultiScribe Reverse Transcriptase, with a total volume of 20 µl. The reactions were run in the thermocycler for 10 min. at 25℃, 120 min. at 37℃, 5 min at 85℃ and hold at 4℃. The cDNA is diluted 1:5 with PCR-graded water and stored in 4℃ for short-term and -20℃ for long-term storage. The cDNA was continued to qPCR for analyses of the RNA expression.

3.5.3. Primer design

Primers were needed for genomic analysis of the gene of interest. Primers were designed with the NCBI primer-BLAST online tool, using the default setting with the additional parameters of Maximum Template Mis-priming and Maximum Self and Pair Complementarity set as low as possible. The primers were checked for melting (T_m) and prediction for primer-dimers and secondary structures with the OligoEvaluator online tool by Sigma-Aldrich and Beacon Designer by Premier Biosoft. Primers were also cross-checked for appearance in other parts of the human genome with BLAT (UCSC browser). For the plasmid-specific primers, the sequence from the plasmid map is used as the query, while for primers for target genomes, the RefSeq ID (found in the UCSC browser) for the gene was used. The primers are ordered from Sigma-Aldrich.

3.5.4. qPCR

Quantitative PCR (qPCR) was performed to check if the gene of interest is present in the cell line. This is a PCR-based technique that couples the amplification of a target DNA sequence

(with specific primers) with the quantification of the concentration of the DNA. qPCR was performed on the transfected HT-1080 to check the plasmid expression of transient overexpression and on the transduced ASC52telo to check the integration of the genes of interest. All qPCRs followed the same step with different inputs and primers.

One reaction in a 384-well plate had a total volume of 10 µl, which contained 1X SYBR™ Green PCR Master Mix, 0.4 µM of forward and reverse primer, and 1.25 µl of the synthesized cDNA diluted (1:5 with PCR-graded water) was template. With the LightCycler 480, the following was run: denaturation at 95°C for 5 min, followed by 45 cycles of amplification at 95°C for 10 sec., 60°C for 15 sec., and 72°C for 10 sec., then melting curve with 95°C for 5 sec., 65°C for 1 min., and then the temperature increases until temperature reaches 97°C at a ramp rate of 0.11, then cooling at 40°C for 10 s. The data from the melting curve was calculated. The amplification data was analyzed with 2nd derivative quantification, the Cp values were exported, further calculations were performed in Microsoft Excel, and the graph was plotted in GraphPad Prism v.10.21.

3.6. Lentiviral production

Three viruses were made to incorporate the whole Tet-on inducible, CRISPR/Cas9-mediated inactivation system into the genome of the ASC52telo (Figure 3.6.1). The viruses were made separately and transduced one after the other when positive clones were selected.
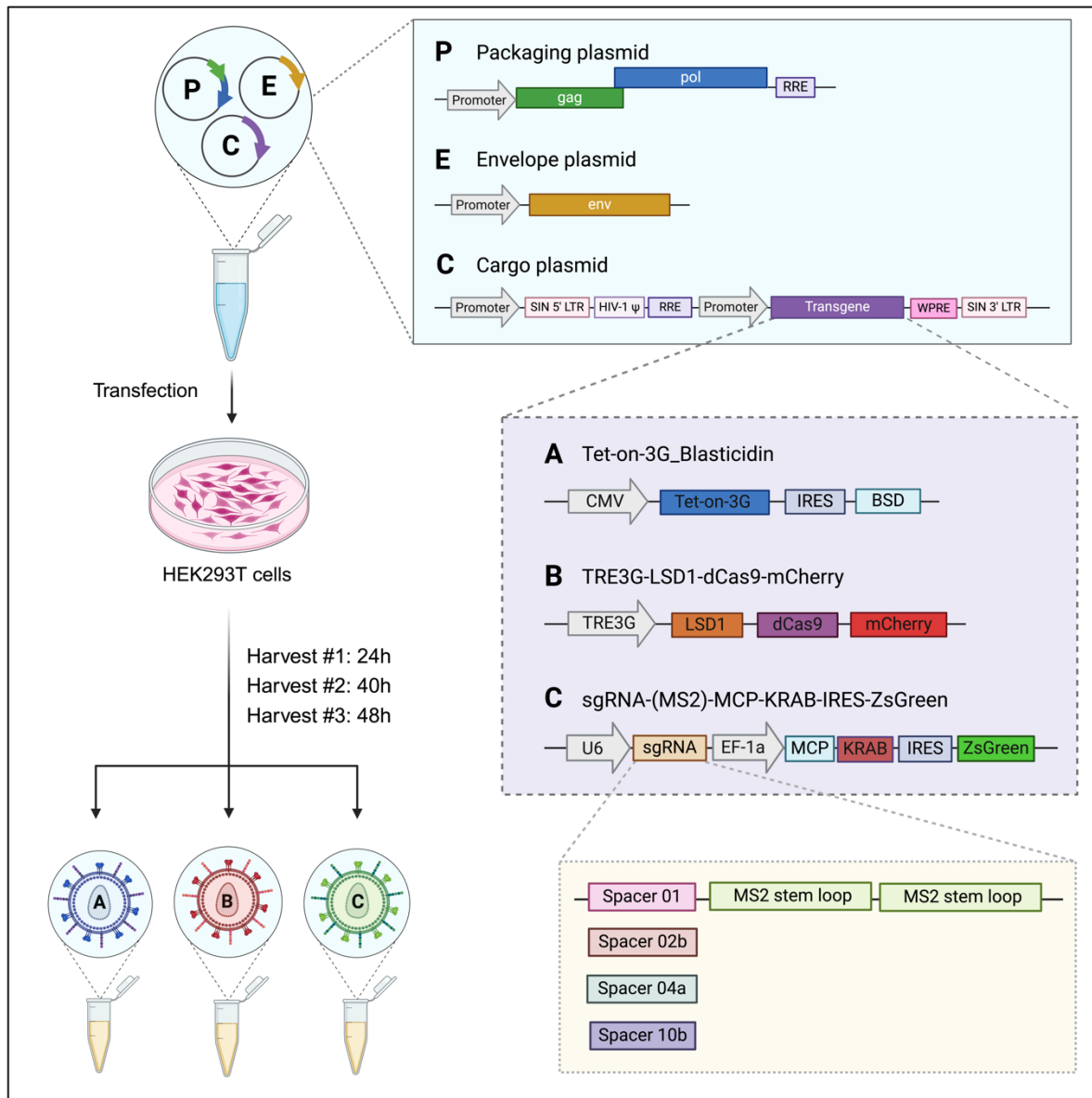
Figure 3.6.1: an overview of the lentiviral packaging in the HEK293T cell line. The lentiviral vectors were produced in the packaging cell line HEK293T by transfection with the 2nd generation lentivirus packaging plasmid, envelope plasmid, and the cargo plasmids with the desired cargo (A), (B), or (C). The 2nd generation lentiviral system has the lentiviral HIV-1 genes divided into two plasmids, the packaging plasmid and the envelope plasmid, as a precaution not to produce replication-competent viral vectors. The packaging plasmid contains the gag-pol genes and the Rev Response Element (RRE). Gag encodes group-specific antigen, a polyprotein that contains the core structural proteins making the viral capsule, which is necessary to assemble the viral particles. Meanwhile, pol encodes the proteins responsible for the synthesis and integration of viral DNA into the host genome. The RRE allows for rev-dependent mRNA export from the nucleus to the cytoplasm. The envelope plasmid encodes the surface proteins; in this thesis, the VSV-G envelope gene encodes the vesicular stomatitis

virus G glycoprotein. The transfer plasmid contains the desired sequences that are to be integrated into the host genome, in addition to the necessary elements for the incorporation, which are psi (Ψ) packaging signal, RRE, and the 5' and 3' LTRs (long terminal repeats) which facilitate the integration into the genome by integrase. Three versions of the transferer plasmids contain the different components of the inducible CRISPR/Cas9 inactivation system. Plasmid A contains the genes for Tet3G with the blasticidin-resistant gene. Plasmid B encodes for the histone demethylase LSD1 and the dCas9 protein along with the fluorescence marker mCherry. Plasmid C has the genes for the sgRNA, which has four different variants (illustrated in the figure, see spacer sequences in Table 2.7) in four different plasmids; these are paired with MS2 stem-loops, resulting in sgRNA with specific spacers with two MS2 stem-loops. This plasmid also encodes for MCP (MS2 coat proteins) and the transcription repressor KRAB. The three (packaging, envelope, and transferer) plasmids were transfected into the HEK293T cell line with TransIT-LT1 Transfection Reagent, and harvest of the virus particles was performed at 24h, 40h, and 48h. Each transfer plasmid was packaged in a separate reaction, making three separate viruses. Figure made in Biorender, with inspiration taken from Krill (2023)

3.6.1.   Preparation of HEK293T cells for transfection

HEK293T cells were seeded to reach 80-90% confluency on the transfection day in an antibiotic-free medium (DMEM, 4.5 g/L Glucose, Glutamine, 10% Tetracycline-free FBS (3.1 million cells in a 3x15 cm dish)

3.6.2.   Transfection of HEK293T cells

The envelope and packaging plasmid were transfected into the packaging cell line, HEK293T, and the transfer plasmid containing gene interest to produce viral particles. HEK293T cells were seeded three days prior to transfection in 15-cm dishes and had 80-90% confluency on the day of transfection. Transfection was carried out using TransIT-LT1 reagent at a 1:2 DNA-to-reagent ratio, following the manufacturer's protocol. Each transfection contained 22.5 µg of cargo plasmid containing the gene of interest ((A) Tet3G_BSD, (B) LSD1-dCas9 or (C) sgRNA-KRAB) 14.6 µg pCMVR8.74 (2nd generation packaging plasmid with gag, POL, REV, TAT) and 7.9 µg VSV-G (envelope protein) (Table 2.1 for more details), a total of 45 µg plasmid per transfection. The total transfection volume was a total of 3 ml with the addition of Opti-MEM, this was dropwise added to the 15 cm plate of HEK293T cells. Three separate transfections were conducted to produce viruses containing plasmids A, B, and C. Plasmid C had four different versions as a spacer was inserted into the plasmid to enable sgRNA binding to different sites near the putative enhancer. All four versions of plasmid C were transfected together, totaling 22.5 µg of cargo plasmid, to generate a virus with pooled spacers. The cell medium (without antibiotics) was changed immediately before transfection and again 16 hours post-transfection.

3.6.3. Harvest

Viral particles were harvested by collecting the medium from the cells. The first harvest was performed eight hours after the last medium changed (after transfection), the second harvest was 24 hours, and the third harvest was 32 hours after the medium changed. The viral supernatant was kept on ice while collecting, centrifuged at 500 g for 10 min to remove cell debris, and then kept at 4°C for short storage. The harvests were filtered with a non-pyogenic filter (to avoid destroying viral particles) and later combined.

3.6.4. Up-concentration of viral stocks

Since the lentivirus supernatant is collected from the packaging cell line, the virus produced in the HEK293T medium is incompatible with transducing the target cell line. The Lenti-X Concentrator was used to concentrate the viral stock. The Lenti-X^{TM} Concentrator Protocol-at-a-Glance (PT4421-2) from the manufacturer was performed. In brief, the protocol included combining the 1 volume of Lenti-X Concentrator with three volumes of filtered viral supernatant, incubation for at least 30 min at 4 °C, centrifugation at 1500 x g for 45 min, and resuspended in 1/100th of the original volume in Opti-MEM or antibiotic-free ASC medium. The concentrated viral was aliquoted and frozen at -80°C.

3.6.5. Quantification of physical viral titer

To quantify the virus stock, Lenti-X^{TM} GoStix Plus was used to detect the presence of lentiviral p24 capsid protein. The Lenti-X™ GoStix™ Plus Protocol-At-A-Glance from the manufacturer was followed. In short, 20 µl viral supernatant (diluted if necessary) and 80 µl of chase buffer were added to the GoStix cassette, incubated for 10 min, and with the Lenti-X GoStix Plus app on a smartphone, the virus titer given as GoStix value (GV), which is equal to ng/ml p24 protein, by comparing the intensities of the test and the control bands. Quantification was performed on both filtered, unconcentrated, and concentrated viral stock; a freeze test on the concentrated virus was also performed to see if freezing would degrade the virus. The physical viral particle (PP) is calculated from the obtained GV. 1 ng p24 contains 1×10^7 PP (Sigma-Aldrich, n.d.)

## 3.7. Transduction of Lentivirus in ASC52 cells

3.7.1. RetroNectin

RetroNectin is a transduction reagent that enhances lentiviral-mediated gene transfer. Some infection reagents, e.g., Polybrene, are toxic to mammalian stem cells (Malach et al., 2023). RetroNectin is a recombinant human fibronectin fragment that features three domains: the cell-binding domain, the heparin-binding domain, and the CS-1 sequence. The product enhances the efficiency of retroviral gene transduction by facilitating the co-localization of target cells and virions (TakaraBio, n.d).

Wells were coated with RetroNectin according to the manufacturer's protocol (Takara Bio Inc). Each well was coated with 5µg/cm^2 RetroNectin (20 µg/ml) and incubated for two hours at RT. RetroNectin solution was then removed and blocked with an appropriate volume of sterile 2% bovine serum albumin (BSA) in PBS for 30 min in RT. The wells were washed with PBS and then ready for seeding or could be stored at 4°C when sealed with parafilm for up to one week.

3.7.2. Transduction

Cells were seeded out with about optimal cell seeding confluency (ASC52telo 5000 cells/cm^2) in an antibiotic-free medium 24h before transduction. The desired amount of virus diluted in an antibiotic-free medium was added, and after 6 hours, the viral supernatant was removed, and the cells were washed with PBS twice before the normal antibiotic-containing medium was added. The cells underwent a 48h period of dormancy before initiating the selection phase, induced by the introduction of blasticidin (BSD) or doxycycline (dox). A schematic overview is shown in Figure 3.7.1.
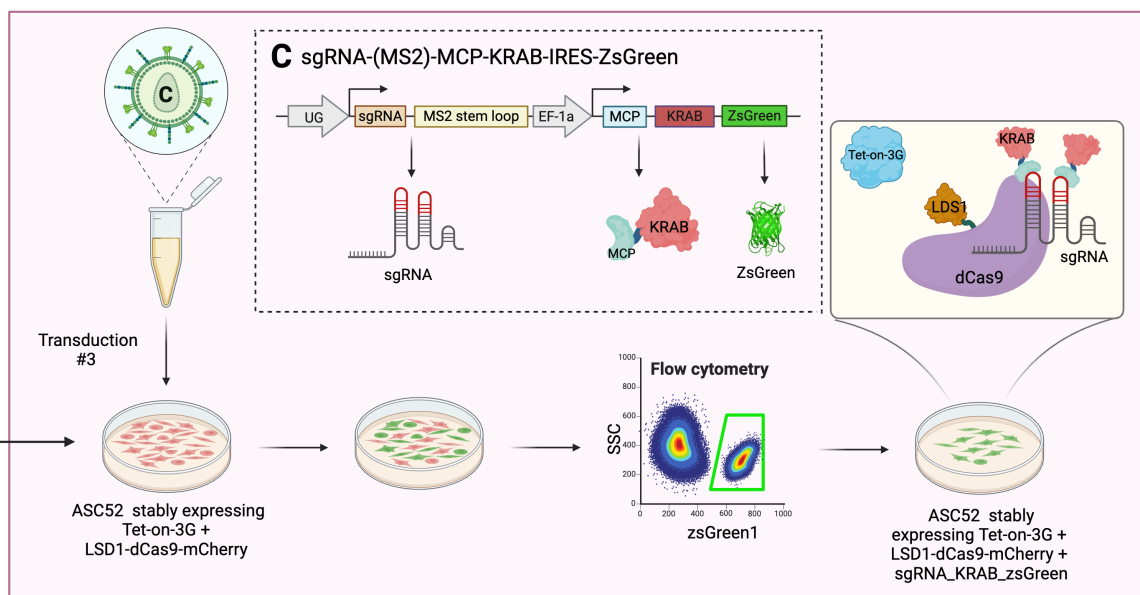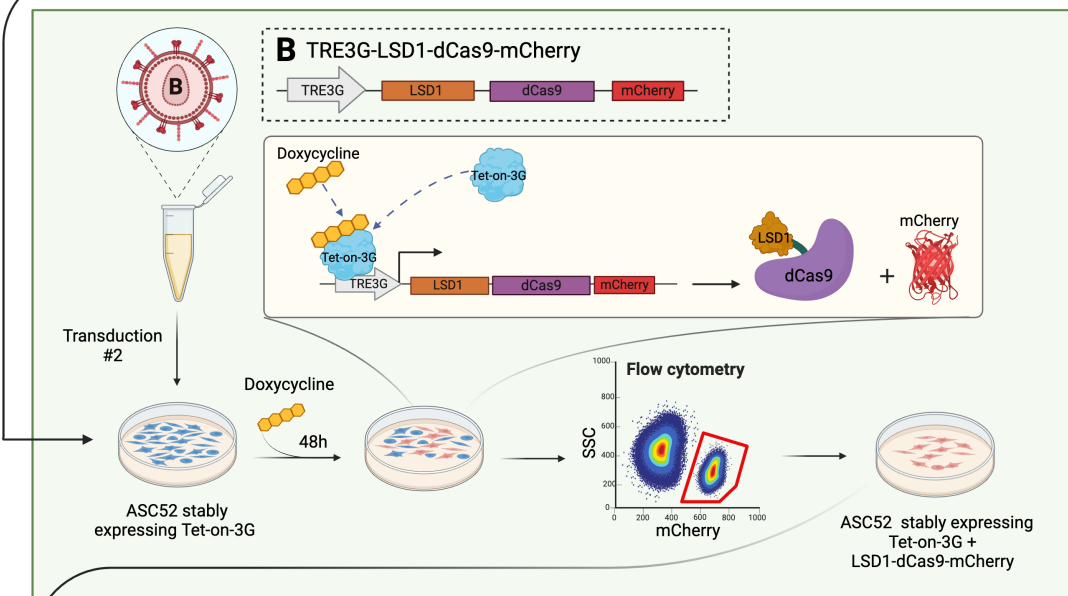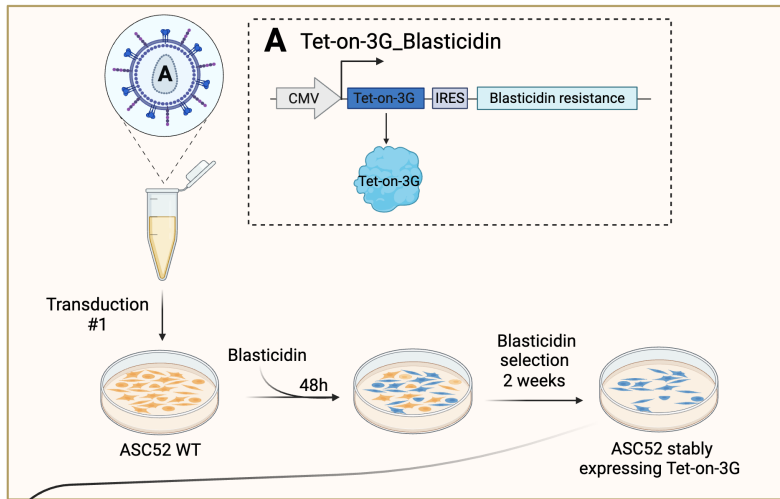
Figure 3.7.1: The sequence of transductions and the assembly of the Tet-on 3G inducible enCRISPR/Cas9i complex system. Firstly, virus A, comprising the Tet3G component, was transduced into WT ACS52 cells. Following successful transduction, Blasticidin (BSD) was administered after a 48-hour period, initiating a 14-day selection phase with a dosage of 20 µg/ml. Upon successful selection of positive clones, the second transduction was conducted on these cells. The second virus contained the dCas9 construct fused with LSD1. After 48h post-transduction, doxycycline (dox) was introduced to induce the transcription of dCas9, LSD1, and mCherry, as these genes are positioned downstream of the TRE3G promoter, which is activated exclusively by Tet-On-3G in the presence of dox. Upon achieving stable gene expression and sufficient cell numbers, fluorescence-activated cell sorting (FACS) was employed to sort cells expressing mCherry. Only cells expressing mCherry, and subsequently dCas9 and LSD1, were retained and expanded. Subsequently, the cells, now expressing both Tet3G and LSD1-dCas9, were transduced with virus C, containing the sgRNA and KRAB, with zsGreen serving as the selection marker. The transduced cells underwent FACS and were sorted for zsGreen. At this stage, triple-transduced cells expressing all components of the inducible enhancer CRISPR/Cas9 inactivation system were obtained. Figure made in Biorender, adapted from Krill (2023)

3.7.3. Blasticidin Selection and Functional Titer Test of ASC52 Transduce with Tet3G-BSD

Knowing the viral titer is important for evaluating the quality of the viral production and how much virus is needed for a successful infection. The Lenti-X™ GoStix Plus only predicted the physical titer, which measures the total viral particles, both functional and non-functional. A functional titer test was performed to find out how many viral particles are actually able to infect cells. This was done by making a serial dilution of viral supernatant. Two series of dilutions, 1:5 and 1:10 fold serial dilutions with eight folds, were chosen. The cells were treated the same as transduction, 48h post-transduction, 20 µg/ml blasticidin (BSD) was added for the antibiotic selection of the cells transduced with the Tet3G_BSD virus. Selecting only ASC52telo_Tet3G_BSD positive cells. A kill-well seeded with ASC52 WT (not transduced) received the same treatment. The titer test was performed over 14 days, with regular changes of fresh medium with BSD every 3-4 days and washing with PBS to remove dead cells and debris. Counting of the surviving cells on day the day all cells in the kill well were dead. Cells were observed in the Nikon Eclipse TS100 microscope.

The functional titer was estimated using the following equation (Addgene, 2023). The number of surviving cells was used as an indication of the titer by multiplying with the total cells transduced and the dilution factor, all dividing by the total volume of the virus supernatant. The functional titer is given in infectious units (IFU) per. ml. With this knowledge, the Multiplicity of infection (MOI) can be calculated and used for transfection of ASC52.

Equation 3.7.1

$$T = \frac{N * F * D}{V_T}$$

T = Titer [IFU/mL]

N = number of cells transducted

F = Fraction of cells surrived, $\frac{number\ of\ surriving\ cells}{N}$

D = dilution factor

V_T = Transduction Volume [ml]

3.8. FACS – Fluorescence-activated cell sorting of transduced ASC52cells

The transduced cells with viruses B and C included fluorescence markers, mCherry, and zsGreen1, respectively. Flow cytometry was used to sort the cells for fluorescence-positive, more accurately Fluorescence-activated cell sorting (FACS) with the Sony SH800 Cell Sorter to detect fluorescence.

Transduced ASC52telo cells were detached by accutase and centrifuged at 300 x g for 3 min., and the cell pellet was resuspended with 500 µl FACS buffer and put on ice until sorting. The Sony SH800 Cell Sorter was set up and calibrated as instructed with the help of Flow personnel. To sort the cells, a series of gates must be set to ensure the correct sorting of the target cells. The first gate removed debris by setting the x-axis for the forward scatter area and the y-axis for the side scatter area (FSC-A vs. SSC-A). The second gate was set to eliminate doublets and only sort single cells, done with forward scatter area vs. forward scatter height (FSC-A vs. FSC-H). The third and fourth gates were set to sort the mCherry-positive and mCherry-negative cells. mCherry is detected by the FL3 (617/50) filter since mCherry has an excitation peak at 587 nm and an emission peak at 610 nm. Gates were set with SSC-A vs. FL3-A filter. First, a negative control was run for a few minutes to set the mCherry-negative gate, and then a sample with mCherry to set the gates for the positive cells. Then, the machine was ready to sort ASC52 cells that have integrated LSD1-dCas9_mCherry into its genome. The samples were first run for a few minutes to check the gates were set correctly; small changes were made to adjust for the different samples to ensure precise gating for all the samples. Then, the sorting could start by loading the collection, the plates with medium, and selecting the ultra-purity mode and the

number of cells needed in each well for optimal seeding density. Each sorting was also recorded by clicking the record button to save the data for later analysis.

For sorting the cells with sgRNA-KRAB_zsGreen1 integrated, the FL1 (525/59) filter was used, which has an excitation peak at 492 nm and an emission peak at 506 nm. The same gates as above were set; the mCherry gates were gated because these cells are supposed to be triple transduced and will have both mCherry and zsGreen1. To set the zsGreen1-negative/positive gates, the graph was chosen to show SSC-A vs. FL1-A filter. A last graph was made, which plotted the FL1-A filter vs. the FL3-A filter, which showed zsGreen1 and mCherry in the same graph. Then the following gates were made fluorescence negative, just mCherry positive, just zsGreen1 positive, and double positive (both mCherry and zsGreen1 positive). So, the double positive could be sorted. The procedure was the same as the mCherry integrated cells, but the cells were sorted into 15 ml tubes with some medium and then seeded at optimal seeding density for the number of cells sorted.

In addition to the mCherry-positive and zsGreen1-positive cells, a series of controls were also sorted with different combinations of viruses. For the samples with no fluorescence marker, all single cells were then sorted. The flow cytometry data that was recorded was analyzed with FlowJo^{TM} Software v10.10.

3.9. Isolation of single clones

For the ASC52_Tet3G cell line, single-cell colonies were attempted to establish cell lines with successfully integrated Tet-On-3G, which are clones, meaning that the genomic composition of the cells in the colonies is identical. This will give more control down the pipeline of the experiment, knowing that all the cells express the same amount of Tet-On-3G.

The first attempt was to seed the cells thinly on a cell culture dish so that the cells were far from each other. Approximately 2000 cells were seeded into a 10-cm culture dish, during the BSD selection, regular changes of media containing 20 µg/ml BSD during the whole selection were used, and 10 µg/ml BSD for maintenance. Given the considerable spacing between seeded cells, isolated cells were identifiable and monitored individually. Cells adhering at an appropriate distance from neighboring cells were marked and observed periodically using a light

microscope to track colony formation from single cells. Upon successful colony formation, colonies were meticulously transferred to a 96-well plate by gentle scraping with a cell scraper and collection using a pipette under microscopic.

The second attempt was by FACS. The Sony SH800 Cell Sorter was utilized to get one single cell in each well of a 96-well plate. The same gating methods as in section 3.8 were used as mentioned in section 3.8, only for the non-debris and single cells; since these cells did not have any fluorescence marker, the machine sorted all single cells directly into 96-well plates. One well was set with 100 cells to help monitor cell division on the single cells and find focus with the microscope. Since the cells are isolated in separate wells, growth factors are needed for cell division, EGF (1 ng/ml (v/v) f.c.) and FGF (10 ng/ml (v/v) f.c.) were added to standard ASC52 medium.

3.10. Validation of transduced cells

To validate the integration of the lentiviral transduction, genomic DNA (gDNA), mRNA and protein were extracted from the transduced cells. The gDNA was used to validate the integration of the components of the CRISPR/Cas9-system. mRNA was assessed to validate the expression of the integrated genes in ASC52 cells and their translation into proteins. Protein extraction was performed to verify the presence of the proteins, thereby confirming the functionality of the CRISPR/Cas9 inactivation system.

3.10.1. Isolation of gDNA, mRNA and proteins from transduced cells

The genius of the NORGEN RNA/DNA/protein purification plus kit is that it can isolate RNA, DNA, and protein from the same sample. The kits were used to extract RNA and gDNA from the same cell sample. Single-transduced, double-transduced, and triple-transduced ASC52telo were lysed to test the gene expression on RNA and genomic DNA (gDNA) levels. The cells were lysed by washing with the appropriate amount of ice-cold PBS, then added 300 µl SKP buffer and incubated for 5 min in RT with gentle swirling, then transferred to Eppendorf tubes and placed on ice to avoid the degradation of RNA. The samples were stored at -80°C or purified immediately.

The samples were purified after the manufacturer's instructions. If the samples were frozen, they were thawed at 55°C for a maximum of five min. The gDNA was purified by filtering the lysate through the gDNA purification column with centrifugation at 8,000 RPM for one minute. The flowthrough was kept on ice for RNA purification while the column was reassembled into a new collection tube. Washing solution A was added to the column and centrifuged at 6,000 RPM for one minute, discarding the flowthrough; this was done twice, then centrifuged empty at 13,400 RPM for two min. to dry the resin. At last, 50 µl elution buffer F was added and incubated for 2 min in RT and then centrifuged for 2 min. at 2,000 RPM, followed by one minute. at 13,400 RPM.

The RNA was then isolated by spinning the RNA flowthrough, which was set aside from the gDNA extraction, with ethanol (for every 100 µl flowthrough, 60 µl ethanol was added) through an RNA/Protein Purification Column for 2 min. at 6,000 RMP. The flowthrough is kept if protein purification is needed. Then, the samples were followed by the same washing steps as gDNA, but this time, they were washed with trice. Then 50 µl of elution Solution A was added to the column to elute the gDNA, then centrifuged for 2 min. at 2,000 RPM, followed by one minute. at 13,400 RPM. The elute was then put through the same column and spun again for a higher yield. The column was kept for protein purification if needed.
The concentration of RNA and gDNA were measured by NanoDrop Spectrometer and continued to qPCR for genomic analysis.

The protein purification was done as follows: mix the protein flowthrough retained from the RNA purification with an equal amount of molecular-graded water, and 8% (v/v) Binding Buffer A of the volume of flowthrough was added. The mixture was centrifuged in the same column as the RNA purification for 2 min at 8,000 RPM, discarding the flowthrough. Wash Solution C was added to the column and centrifuged like last time. Then, 9.3 µl of Protein Neutralizer was added to an Elution tube, which the column was transferred to, and 100 µl of Elution Buffer C was added to the column and centrifuged for 2 min. at 8,000 RMP to elute the bound proteins. Proteins were then stored at -80°C for long-term storage.

3.10.2. Detection of integrated virus in the ASC52 gDNA

To detect integrated virus sequences, qPCR was conducted targeting the respective lentiviral cargo plasmid within the isolated gDNA. The ASC52_Tet3G cell line was examined using primers specific to Tet3G (refer to Table 2.6), along with the MMP9 and IRX3 enhancer for

gDNA normalization. The qPCR protocol followed the methodology outlined in section 3.5.2 Notably, the results were expressed relative to the genomic region in human cells rather than employing a housekeeping gene in reverse transcription qPCR.

3.10.3. RNA Expression from integrated viruses

To assess the expression of integrated viral vectors, cDNA synthesis and qPCR analyses were conducted following established protocols (sections 3.5.2-3.5.4). Utilizing the same target and control primers as those used in transient overexpression experiments (refer to Table 2.6) qPCR was performed.

3.10.4. Protein expression from integrated viruses

The protein extracted from the triple transduced cells and their respective controls were examined for protein expression. This analysis was carried out using Western blotting techniques outlined in subsequent sections (3.11)

3.11. Quantification of protein

Prior to proceeding with protein quantification, it is imperative to determine the concentration of the protein. protein was measured with the SpectraMax plus 384 microplate reader alongside a BSA 2-fold serial dilution (2 µg/µl —0.125 µg/µl) to make the BSA standard curve. The DC^{TM} protein Assay kit was used to quantify the protein. The kit consisted of reagents A, B, and S. Reagents A and S were mixed 1:50, respectively. For each sample/standard, 5x reagent A+S and 40x reagent B were loaded into a flat-bottom 96-well plate with three blank controls. Mixed by pipetting up and down, and shake on 80 rpm for 15 min in RT. Then, the plate was read on the plate reader with an absorbance of 750 nm. The absorbance data was transferred to Microsoft Excel to make the standard curve and calculate the protein concentration. Then, a Western blot was performed to check if the protein of interest was present.

3.11.1. Western blot

Western blot analysis (WB) was conducted to ascertain the presence of the anticipated proteins. This analytical procedure was executed using the isolated protein extracted from the triple-transduced cells and its various controls. Western blots consist of several steps: SDS-page, blotting, and staining.

SDS-PAGE

SDS-page was performed to separate the proteins according to their molecular size. Sodium dodecyl sulfate (SDS) denatures the protein and makes it linear. The samples were diluted to the same concentration according to the sample with the lowest concentration, mixed with 1X SDS loading buffer, and heated at 95°C for 5 min. The samples (6.2 µg) were loaded into the premade SurePage Tris-Bis gels alongside 5 µl kaleidoscope protein ladder, and the positive Cas9 control (0.2 µg). The gels ran at 200 V for 45 min.

Semi-dry blotting (transfer to membrane)

When the proteins are separated, they are transferred to a membrane using the Trans-Blot® Turbo™ Transfer System with the Trans-Blot Transfer Pack, following the manufacturer's instructions, but the transferred time was doubled, to ensure the all the sizes of proteins are successfully transferred. The protein transfer was checked with ponceau S staining (0.1% (w/v) ponceau powder dissolved in 5% (v/v) pure acetic acid). Appropriate amount of staining solution was applied to the membrane for approximately 5 min, until the red bands started to develop. The membrane was washed with PBST to destain the red staining.

Antibody staining

First, the membrane was blocked with 5% dried skimmed milk (w/v) in PBST (0.001 % tween 20 in 1x PBS) whilst shaking at 30 rpm for 1 hr in RT or o/n at 4°C to avoid unspecific binding to the primary antibody. Then, the membrane was washed 2x5 sec. and 2x5 min. with PBST. The primary antibody solution (see Table 3.11.1) was added and incubated whilst shaking at 30 rpm for 1h in RT or o/n at 4°C. Afterward, the membrane was washed 2x5 sec and 1x15 min before staining with the secondary antibody (Table 3.11.1) on the shaker for 30 min. at RT. Then, washed again for 2x5 sec., 1x15 min., and 4x5 min. Following, the membrane was soaked in Femto detection solution (1:1:1 ratio of peroxide buffer, luminol, and MilliQ) for 2 min in RT in the dark. Then, placed in between two pieces of transparent paper for imaging. The membrane was imaged with the Universal Hood II Gel Doc system. A picture of the ladder was first taken manually and exposed at 0.002. Then, live acquire was used to image the proteins. The membrane was put back in PBST.

If more antibodies need to be run on the same membrane, the membrane can be stripped and stained again. Then, the following is performed: the membrane is washed for 5 min in BPST, and then stripping solution is added and incubated for 8 min in RT on a shaker covered in aluminum foil. Then, it is washed again for 2x5 sec. and 1x15 min. Then, the membrane can be

blocked and stained with a new primary antibody. The membrane is stored by drying it in filter paper and then in aluminum foil at 4°C.

Table 3.11.1: Antibody Solutions for Staining

Buffer	Content
Primary antibody solution	3 % BSA in PBST Dilution of Primary antibody: - Anti-KDM1/LSD1 1:5000 - Anti-TetR 1:1000 - Anti-CRISPR-Cas9 1:1000 - HMBS 1:1000 - Anti-beta Actin 1:5000.
Secondary antibody solution	3 % BSA in PBST Dilution of Secondary antibody: - Anti-mouse 1:7500 - Anti-rabbit 1:10 000

4. Results

4.1. Sequencing of plasmids

The plasmids required for the inducible LSD-dCas9 inactivation system were sequenced to verify the presence of the target genes within their sequences. The plasmid pSFFV-Tet-On-3G-BFP, which previously failed to produce BFP (Krill, 2023), was found to not contain the correct sequence of Tet-On-3G, but was contaminated with dCas9 (Figure 4.1.1.A). In contrast, the replacement plasmid, Tet3G_BSD, which was prepped in the current work, was found to contain the correct sequence (Figure 4.1.1.B.) Moreover, the plasmid LSD1-dCas9-mCherry, as well as pCMVR8.74 were also found to contain the correct sequences (Figure 4.1.1.C-D).
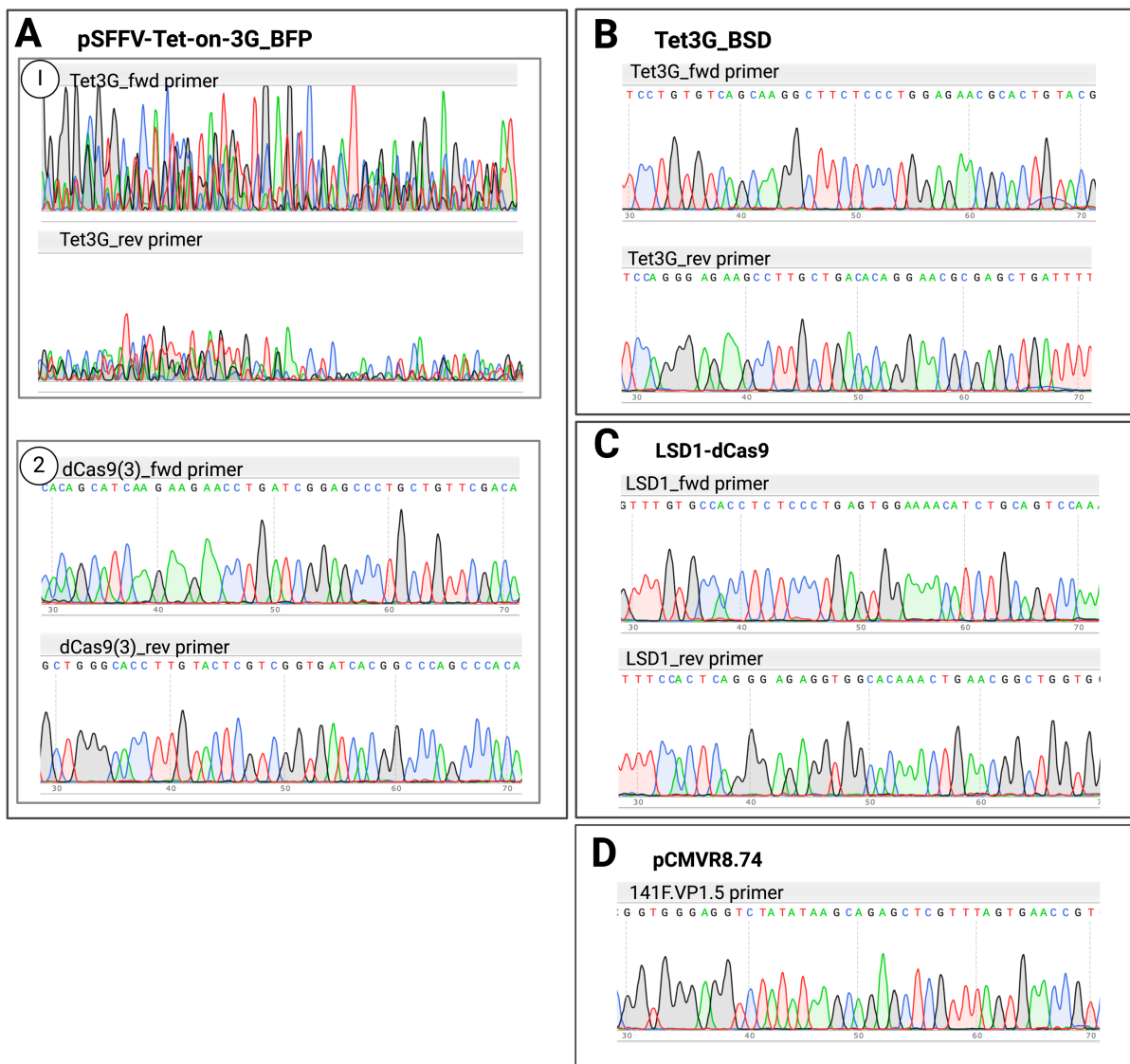
4.2. Transient overexpression of lentiviral cargo vectors in HT-1080 cells

To test the functionality of the inducible Tet-On-3G LSD1-dCas9 inactivation system *in cellulo*, the lentiviral cargo plasmids harboring Tet-On-3G_BSD and LSD1-dCas9-mCherry were co-transfected into HT-1080 for transient overexpression. Following transfection, the cells were cultured with or without doxycycline treatment in tetracycline-free medium to test for activation of the TRE3G promoter by the dox-inducible Tet-On 3G transactivator protein (Figure 4.2.1). 48h after transfection, fluorescence microscopy revealed the dox-induced mCherry signal when the Tet3G and LSD1-dCas9-mCherry plasmids were co-transfected into the HT-1080 cells (Figure 4.2.1). The experiment included several control conditions. Co-transfection of pCMV6-empty and LSD1-dCas9 without dox (Figure 4.2.1 A) and with dox (Figure 4.2.1 B) were conducted as negative controls for Tet-On-3G, as the LSD1-dCas9-mCherry that is regulated by the TRE3G promoter should not be induced in its absence. However, several cells in these negative controls were mCherry positive, suggesting a leaky TRE3G promoter or the presence of an alternative promoter. Cells transfected with Tet-On-3G but cultured in the absence of doxycycline did not increase the number of mCherry positive cells (Figure 4.2.1 C), while transfection with Tet-On-3G in the presence of doxycycline resulted in a considerably higher number of mCherry-expressing cells (Figure 4.2.1 D). Taken together, these data suggest the expression of LSD1-dCas9 could indeed be induced by Tet-On-3G and doxycycline.
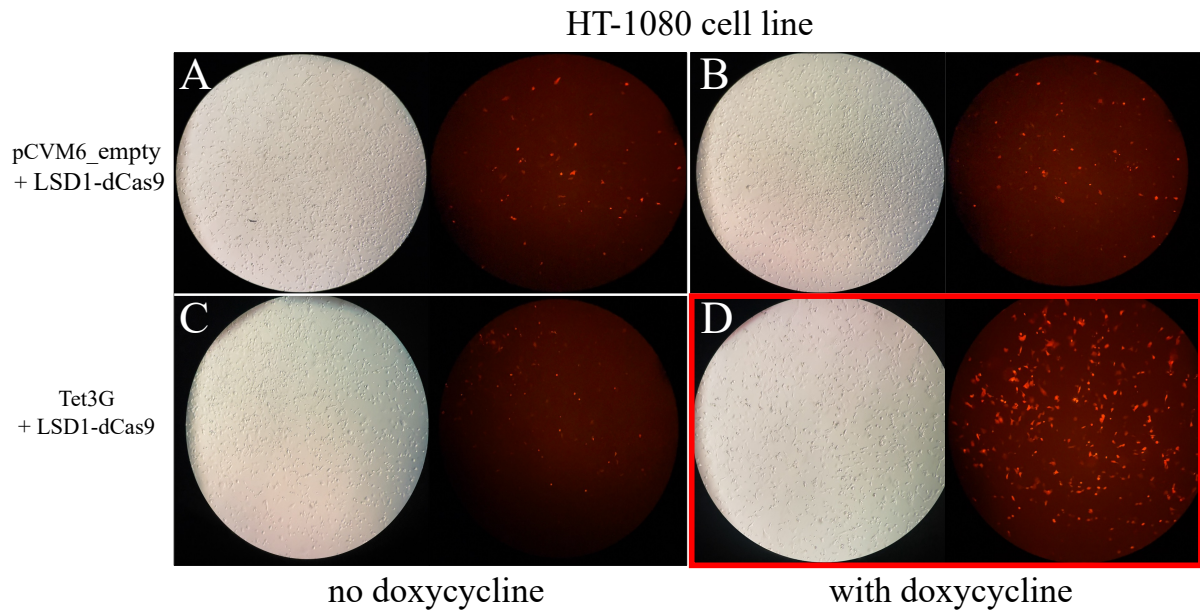
Figure 4.2.1: Validation of mCherry signal of transient overexpression of Tet3G and LSD1-dCas9 in HT-1080 cells. The plasmids Tet3G and LSD1-dCas9 were co-transfected with a total of 0.4 ng plasmid DNA. As a control, the pCMV6-empty (pCMV6-4/5/6XL) was used. Doxycycline (dox) was added to the medium during transfection since the LSD1-dCas9 plasmid was constructed with the construct sequence downstream from the TRE3G promoter, which is only activated with Tet-On-3G in the presence of dox. As a control, cells transfected with the same combination of plasmids were also cultured without dox. After 48 hours post-transfection, the cells were examined for fluorescence under a microscope to confirm successful transfection, as LSD1-dCas9 contains an mCherry fluorescence marker. The cells were assessed under the Nikon TE2000 to check for the expression of mCherry, red fluorescence, with the TRITC filter. The red square (D) is the only cell line meeting all the criteria for LSD1- and dCas9-expression, as well as fluorescent marker mCherry

To confirm the results from the microscopy, RNA was harvested from these cells as well as from non-transfected HT-1080 WT cells, followed by cDNA synthesis and analysis with qPCR. Target-specific primers designed for analyzing the lentiviral cargo plasmids were used, including *Tet-On-3G*, *LSD1*, and *dCas9*, as well as primers targeting the reference genes *RPS13* and *RPLPO* for normalization of the results (Table 2.6). There was a strong and significant difference in *Tet-On-3G* mRNA expression between control cell lines (non-transfected or transfected with pCMV6-empty) compared to those transfected with Tet3G_BSD, regardless of dox presence (Figure 4.2.2 A). In line with the microscopy observations, there was a noticeable dox-independent expression of LSD1-dCas9-mCherry in the absence of Tet-On 3G (Figure 4.2.2 B-C). Moreover, the cells transfected with both Tet3G_BSD and LSD1-dCas9-mCherry showed significantly higher dCas9 (but not LSD1) mRNA expression, but only in the presence of dox (Figure 4.2.2 C). Taken together, these data confirm that the plasmids encode the correct genes and that LSD1-dCas9-mCherry expression could be induced, despite high Tet-On-3G

53

independent background. Hypothesizing that the transient overexpression of LSD1-dCas9 may have exploited an additional promoter in the vector that may not be packaged into lentivirus, lentiviral production for plasmid delivery was performed next.
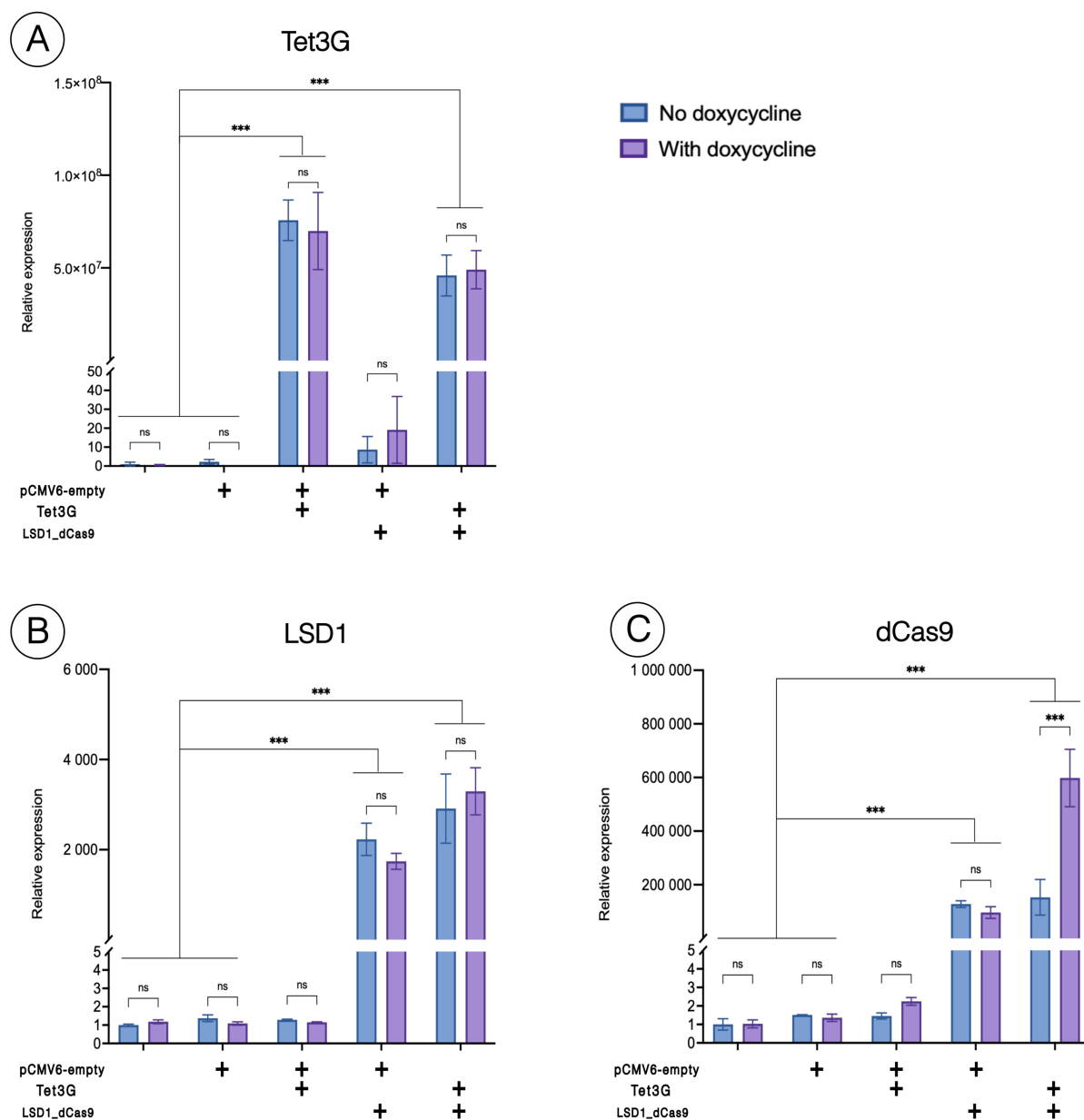
Figure 4.2.2: Transient overexpression of Tet3G and LSD1-dCas9 in HT-1080 cells measured by qPCR. This experiment involved co-transfection of the lentiviral cargo plasmids Tet3G and LSD1-dCas9 into HT-1080 cells with 0.4 4g DNA input. The pCMV6-empty plasmid served as the control plasmid. A no-plasmid control was also included. The cells with the same conditions were transfected in a doxycycline-containing medium and in a tetracycline-free medium. For the induction of the inducible Tet-On-3G system. RNA extracted from transfected HT-1080 cells with different combinations of the plasmids. The data are presented as grouped datasets with the same transfections performed in the presence and absence of dox. The panels show the mRNA expression of Tet3G (A), LSD1 (B) and dCas9 (C). The y-axis represents the relative expression normalized to the mean of the *RPS13*

and *RPLPO* reference genes. The data are shown as mean ± SD (n = 3). Statistical analysis was performed utilizing the two-way analysis of variance (ANOVA) with Holm-Šídák's multiple comparisons test. Significance levels are denoted with *** indicating $p_{adj} < 0.001$. Figures were made with GraphPad Prism v10, as well as statistical analysis. The figures were annotated with Inkscape. Abbreviations: LSD1: Lysine specific demethylase, Tet3G: Tet-On-3G, ns: non-significant.

4.3. Lentiviral production

With the plasmids verified to express Tet-On-3G and LSD1-dCas9, which are fundamental for the inducible enCRISPR/Cas9i system, lentiviruses harboring these components could be made. Three different viruses needed to be made in order for the assembly of the whole inducible enCRISPR/Cas9i system to be integrated into the target cell line (Figure 3.7.1).The viruses contained Tet3G (Virus A), LSD1, and LSD1-dCas9 (Virus B), and sgRNA and KRAB (Virus C), respectively; each of the viruses were produced individually (Figure 3.6.1). Lentiviral particles were produced with the HEK293T packaging cell line and harvested at 24h, 40h, and 48h post-transfection. To confirm the generation of viral particles and estimate viral titer, the harvested samples were assessed using the instant lentiviral titer test, Lenti-X™ GoStix Plus. The LentiX GoStix test measures the concentration of the viral capsid protein p24, consequently indicating the presence of virus. The test results were quantified using the accompanying app from the manufacturer. As an example, Figure 4.3.1 shows the quantification of lentivirus containing LSD1-dCas9 (Virus B) from harvest #1, which resulted in a p24 concentration (GV value) of 390 ng/ml for undiluted lentiviral supernatant before concentration.
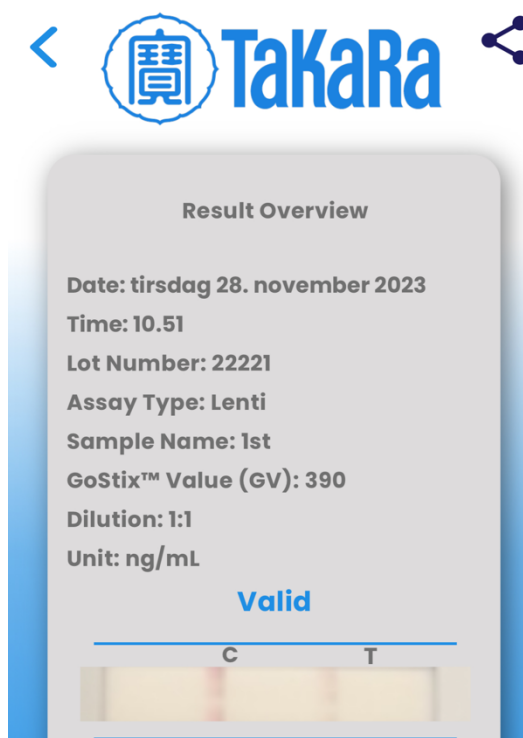
Figure 4.3.1: Quantification of lentiviral titer with LentiX GoStix. Viral titer estimation was conducted by measuring the concentration of the viral capsid protein p24. A small sample of the viral supernatant was added to a test cassette, developing two lines upon the presence of the p24 protein. The test band's (T) intensity indicated the concentration relative to the control (C) line. Quantitative assessment was facilitated by the accompanying app, which provided the GosStix Value (GV) equivalent to the p24 concentration (ng/ml). The result is of virus B harvest #1, with a 1:1 dilution and GV of 390.

Following quantification, the harvested samples were pooled and concentrated, followed by re-measuring the p24 concentration (GV). The GV obtained was translated to the number of physical viral particles (PP). These calculations are based on the fact that 1 ng p24 contains about 10 million PP (Sigma-Aldrich, n.d.). Virus A (Tet-On 3G) was found to have 4.08×10^{10} PP/ml (Table 4.3.1), Virus B (LSD1-dCas9) had a yield of 9.24×10^{10} PP/ml (Table 4.3.2) and Virus C (sgRNA and KRAB) had 6.87×10^{10} PP/ml (Table 4.3.3). A freeze-test of the concentrated viral stocks showed that the freezing did not negatively impact the viral stock; on the contrary, a higher GV for all viruses was found following freeze-thaw (Table 4.3.1-Table 4.3.3).

Table 4.3.1: Quantification of the viral titer for Virus A with Tet-On 3G.

Virus A: Tet3G		GV [ng/ml]	Total amount of p24 [ng]	Total volume [ml]	Recovery	PP/ml
unconcentrated virus stock	Harvest #1	10	420	42	-	-
	Harvest #2	249	10 458	42	-	-
	Harvest #3	64	2 688	42	-	-
concentrated virus stock	Pooled harvest	4 078	5 152	1.26	38%	4.08×10^{10}
Freeze test of concentrated virus stock	Pooled harvest	4 125	5 198	1.26	38%	4.13×10^{10}

GV: GoStix Value, Recovery of viruses after concentration: GV (concentrated)/(GV(unconcentrated), PP: physical particles, IFU: infectious units.

Table 4.3.2: Quantification of the viral titer for Virus B with dCas9_LDS1.

Virus B: LSD1-dCas9		GV [ng/ml]	Total amount of p24 [ng]	Total volume [ml]	Recovery	PP/ml
unconcentrated virus stock	Harvest #1	390	1 6380	42	-	-
	Harvest #2	297*	8 019**	42	-	-
	Harvest #3	477	20 034	42	-	-
concentrated virus stock	Pooled harvest	9 244	11 092	1.20	25%**	9.24×10^{10}
Freeze test of concentrated virus stock	Pooled harvest	16 230	19 476	1.20	44%	$16.2. \times 10^{10}$

GV: GoStix Value, Recovery of viruses after concentration: GV (concentrated)/(GV(unconcentrated), PP: physical particles, IFU: infectious units. *Suspected wrong titer estimation from GoStix; first test with 1:1 dilution was out of range, and this result was following 1:5 dilution. **continued calculation error.

Table 4.3.3: Quantification of the viral titer for Virus B with sgRNA-KRAB.

| Virus C: sgRNA-KRAB		GV [ng/ml]	Total amount of p24 [ng]	Total volume [ml]	Recovery	PP/ml
unconcentrated virus stock	Harvest #1-3	223	20 070	90	-	-
concentrated virus stock	Pooled harvest	6867	5 150	0.90	30%	6.87×10^{10}
Freeze test of concentrated virus stock	Pooled harvest	7988	5 991	0.90	30%	8.00×10^{10}

GV: GoStix Value, Recovery of viruses after concentration: GV (concentrated)/(GV(unconcentrated), PP: physical particles, IFU: infectious units

4.4. Estimation of functional titer

The physical titer obtained from the GoStix instant lentiviral titer test provides a quantitative measure of viral particles based on the detection of viral proteins, specifically the p24 capsid proteins. However, this measurement may not accurately reflect the infectivity or functional capability of the virus. The physical titer quantifies all p24 proteins present, including those not incorporated into intact viral particles, such as free p24 and defective viral particles. Consequently, titers determined solely by p24 quantification tend to be overestimated. To assess the functional titer, a method involving the transduction of the Tet3G_BSD virus into WT ASC52 cells via serial dilution (5-fold and 10-fold with eight folds) followed by microscopic monitoring was employed. Since the Tet3G_BSD virus carries the BSD resistance gene, a selection process involving 14 days of BSD exposure was conducted. Subsequently, the surviving cells were counted on the tenth day, as all the cells in the kill well had died, and the functional titer was calculated based on the observed cell counts (Figure 4.4.1). No or very few surviving cells were observed for transduction of 1:1 dilution of the virus. Dilutions 1:5-1:3125 and $1:10-1:10^5$ were too many cells to manually count. Therefore, the counting was conducted on the last two dilutions of each series. The 1:15 625 dilution had 12 cells left in the well, the 1:78 125 dilution left one surviving cell, the $1:10^6$ dilution gave 3 surviving cells and the highest dilution of $1:10^7$ had no cells.

Funtional titer test Virus A (Tet3G_BSD)

5-fold serial dilutions	10-fold serial dilutions
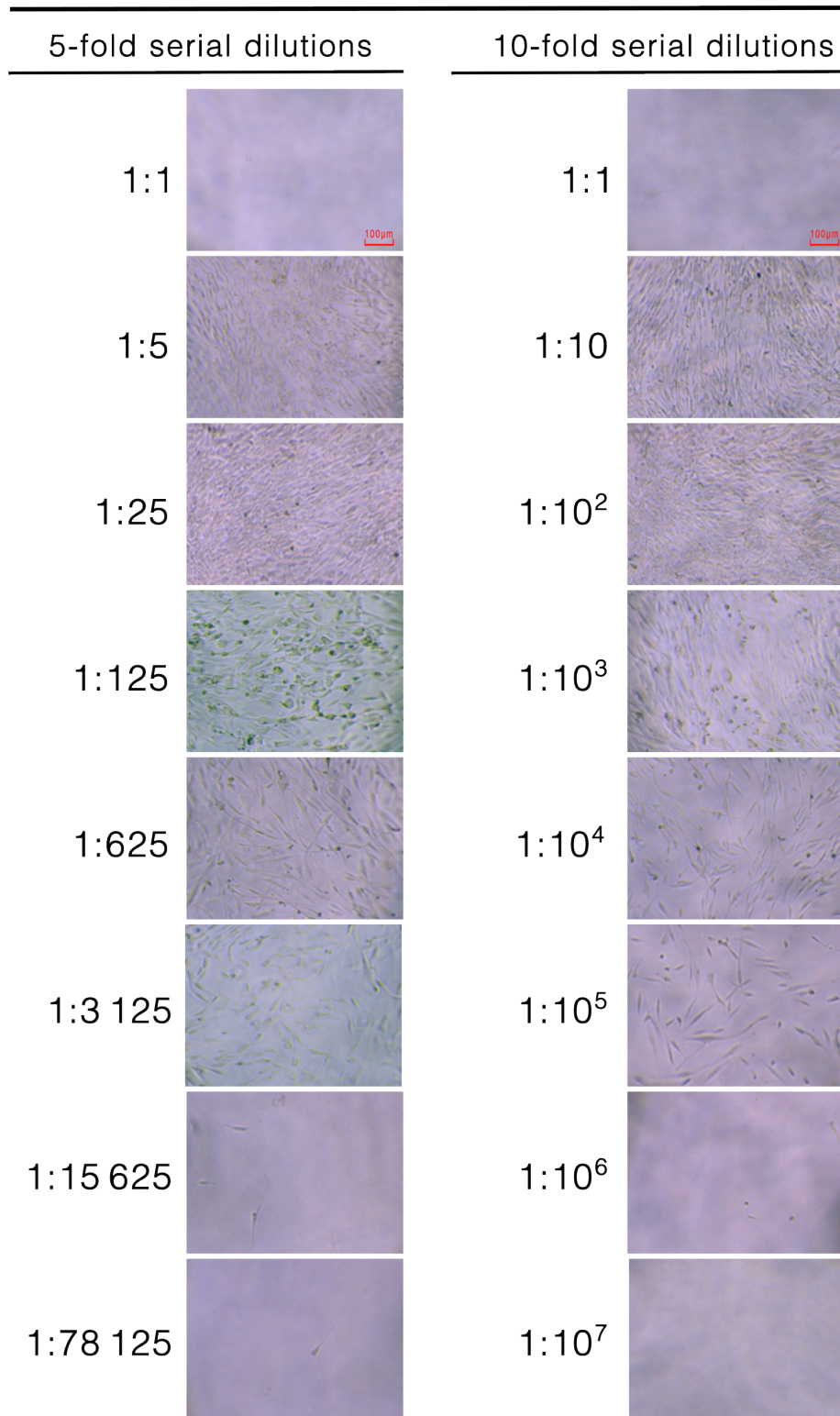
Figure 4.4.1: Functional titer test of Virus A (Tet3G_BSD) with 5- and 10-fold serial dilution. A functional titer test was conducted by transducing ASC52 WT cells with a 5- and 10-fold serial dilution of the lentiviral stock of virus containing Tet3G_BSD. Blasticidin was added 48h-post-transduction, and cells were counted when the kill well was dead. So, the whole blasticidin selection lasted 10 days. Cells were monitored during the whole selection with Nikon Eclipse TS100 Microscope. Figure made in Inkscape.

Using Equation 3.7.1, the functional titer of Virus A containing Tet3G-BSD was estimated based on the number of surviving cells following viral dilution transduction. For viral dilution of 1:15 625, with 12 total surviving cells, the functional titer was 4.58 x10^6 IFU/ml. Similarly, for viral dilution of 1:10^7 with 3 surviving cells, the functional titer was determined to be 7.50 x 10^7 IFU/ml. The mean of the two calculated titers is 3.98 x10^7 IFU/ml. This indicates that the proportion of total PP that are infectious units (IFU/PP) is 9.76 x 10^{-4} for Virus A (Tet3G_BSD), which was used as a standard for all downstream lentiviral transduction of ASC52. Then, functional titer was calculated to be 9.02 x 10^7 IFU/ml for Virus B (LSD1-dCas9), and for Virus C (sgRNA-KRAB) it was 6.70 x 10^7 IFU/ml, based on the PP/ml for each virus. These are the functional titer used to calculate MOI for each viral transduction.

4.5. Transduction of lentiviral Tet3G-BSD in ASC52 cells

The next step to achieve epigenetic modulation with an inducible enCRISPR/Cas9i system was to transduce ASC52 WT cells with a virus containing Tet3G-BSD (Virus A) (

Figure 3.7.*1*). ASC52telo cells were transduced as described in section 3.4 with MOIs 80, 15, 3, and 0.6. After 48h, the cells were split into two parts: one intended for single-cell colony selection and the other designated for the expansion of a heterogeneous mix of all transfected cells. Then the cells moved on to BSD selection to select cells with positive integrated Tet-On-3G.

4.5.1. Blasticidin selection

Laboratory-established findings indicated that a concentration of 20 µg/ml of BSD for ASC52 cells effectively eradicated all cells during a 10-day period, thus, this concentration was used for the selection. Two days after the transduction of Virus A, BSD was added to the culture medium to select positive clones with Tet3G-BSD successfully integrated into the ASC52 genome, as the BSD resistance sequence is also integrated with the transduction of the plasmid. A non-transduced kill-well was treated in the same way. The selection ended when all the cells in the kill-well had died, which was 10 days. All MOIs survived the selection and were used for further analyses. MOI 15 has the most surviving cells and grows fastest, while MOI 0.6 had the lowest number of surviving cells and the slowest growth. The establishment of a library of ASC52 cells stably expressing Tet-On-3G had been achieved, referred to henceforth as

ASC52_Tet3G. Subsequently, to preserve antibiotic resistance, half of the optimal concentration was utilized for the ongoing maintenance of the cell lines.

4.5.2. Single-cell clonal selection of transduced cells

With the generation of heterogeneous ASC52_Tet3G cells, two methods were implemented to establish uniform cell lines featuring consistent expression levels of Tet3G. This approach aimed to attain colonies harboring identical genomic compositions, given that lentiviral DNA integration occurs randomly and may result in multiple integrations within the genome. The BSD selection was performed in parallel with the isolation of single-cell colonies.

The initial approach to form single-cell colonies entailed sparsely seeding cells onto a 10 cm culture plate, identifying individual cells, monitoring their proliferation, and subsequently expanding each colony independently. This process proved to be difficult and labor-intensive. Despite the presence of multiple cells on the same culture dish allowing cell-cell communication, single cells failed to develop into colonies in the absence of very adjacent cells and instead naturally migrated towards neighboring cells to initiate heterogeneous colonies. As depicted in Figure 4.5.1, all 160 monitored clones eventually died. This was consistent across all MOIs of ASC52_Tet3G. Thus, this experiment failed to yield homogenous single-cell populations.
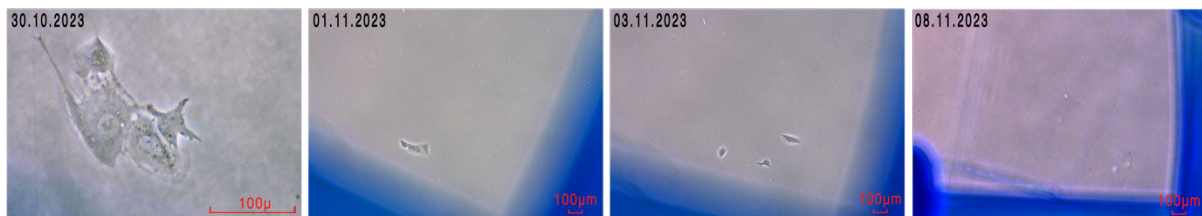
Figure 4.5.1: Example of a single cell dividing when seeded sparsely on a 10 cm plate. ASC52_Tet3G MOI 15 were seeded sparsely on a culture plate, and cells located at an appropriate distance from neighboring cells were marked and monitored with microscopy. This cell started divided four days after seeding and became a colony of three cells, but shortly after the whole colony died. Cells were monitored during the whole selection with Nikon Eclipse TS100 Microscope. Figure made in Inkscape.

The second method was FACS of the heterogeneous mixtures with MOI 80 and 15. The cells were sorted as single cells directly into 96-well plates to facilitate the expansion of single-cell colonies. However, this attempt encountered similar challenges regarding cell viability, either during the sorting process or subsequent culturing.

In summary, none of these methods yielded colonies from one single cell. Simultaneously, however, heterogeneous mixes of the different MOIs were expanded. These heterogeneous

mixes were pools of all ASC52_Tet3G cells transduced with the same MOI. In contrast to single cells or sparsely seeded cells, these heterogeneous mixes were cultured at the optimal seeding density of 5000 cells/cm^2, which resulted in excellent proliferation. Hence, heterogeneous mixes of ASC52_Tet3G from MOI 80, 15, 3 and 0.6 were successfully expanded and used for further testing and downstream transductions.

4.6. Validation of Tet3G integration and expression

The heterogeneous populations of ASC52_Tet3G cells, subjected to varying MOIs, underwent lysis to extract gDNA and RNA, which were analyzed with qPCR to determine the most effective MOI for viral transduction and to identify the optimal conditions for subsequent transductions. qPCR was run on ASC52_Tet3G with MOI 80, 15, 3 and 0.6, with the gene-specific primer for *Tet-On-3G* (Table 2.6) for both gDNA and mRNA.

The qPCR results of the extracted gDNA clearly demonstrated the presence of Tet-On-3G in the genome of all MOIs (4.6.1 A), with a corresponding strong mRNA expression of Tet-On-3G (4.6.1 B). Significant expression was found in almost all MOIs for gDNA and mRNA when compared to ASC52 WT (non-transduced), except for mRNA from MOI 3 and 0.3 (4.6.1 B). When the relative RNA expression was divided by relative gRNA expression, it was clear that only MOI 15 had a significant difference when compared to the rest (4.6.1 C). In conclusion, the Tet-On-3G sequence was successfully transduced into ASC52 cells, with MOI 15 showing the strongest expression from the lowest amount of integrated virus and was thus selected for subsequent transductions. Henceforth, this particular ASC52 _Tet3G variant is exclusively referenced as ASC52_Tet3G.
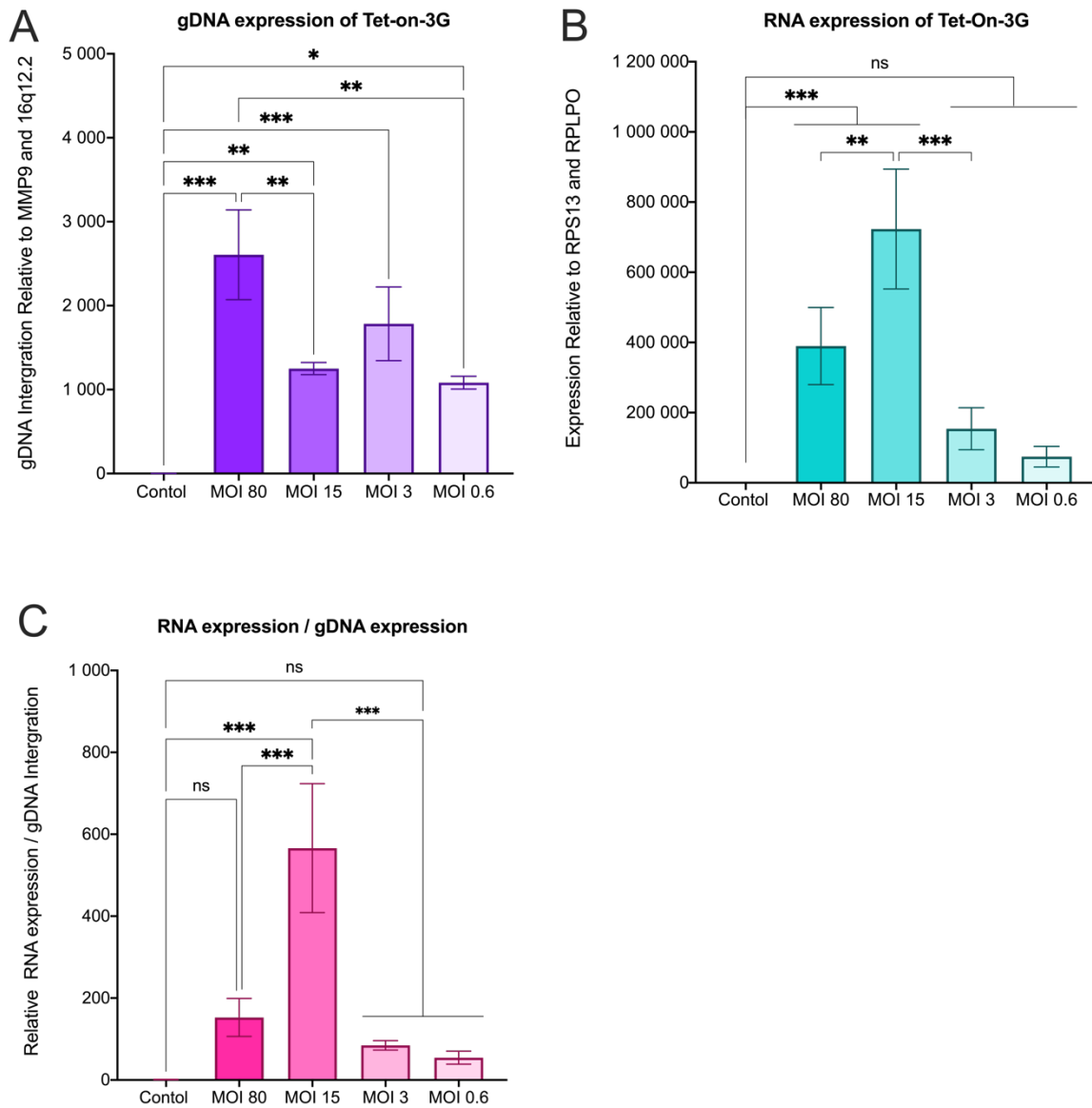
Figure 4.6.1: Validation of integration of Tet3G in ASC52 cells with qPCR. The gDNA and RNA were extracted from the heterogenous pool of ASC52_Tet3G transduced with MO1 80, 15, 3, and 0.6. The qPCR with the gDNA (A), RNA (B) was run with primers to detect Tet-On-3G, the expression is relative to the control genes, MMP9 + 16q12.2 and RPS13 + RPLPO, respectively. When the relative RNA expression is divided with the relative gDNA integration (C), it visualizes the RNA expression per integration of Tet-On-3G. The data are shown as mean ± SD (n = 3). Graphing and statical analysis were performed with GraphPad Prism v10.

## 4.7. Transduction of lentiviral LSD1-dCas9-mCherry in ASC52 cells

Having identified the best-performing ASC52_Tet3G population, the next transduction (Virus B) could be conducted, following the pipeline in (Figure 3.7.1). This transduction introduced dCas9 fused with LSD1, the first part of the enCRISPR/Cas9 system, as well as the co-expression of mCherry. The LSD1-dCas9-mCherry construct is under the control of the TRE3G promoter, inducible by Tet-On-3G in the presence of dox. Briefly, a pool of ASC52_Tet3G MOI 15 cells, alongside WT control cells, were transduced with Virus B, containing LSD1-dCas9-mCherry with MOI 35. Two days after transduction, the cells received fresh medium with or without dox. Of note, no tetracycline-free version of the ASC52 medium was available, so dox was added to the regular medium. After 24h of dox treatment, fluorescent microscopy was performed to analyze the cells for mCherry expression (Figure 4.7.1). As expected, control cells with no transduction of LSD1-dCas9-mCherry showed no expression of mCherry (Figure 4.7.1 A-B), while cells transduced with Tet3G_BSD and LSD1-dCas9-mCherry did exhibit expression of mCherry (Figure 4.7.1 C-D). Interestingly, the presence of dox in the medium did not seem to affect the expression of mCherry. WT cells transduced with LSD1-dCas9, but not Tet-On-3G, also displayed no mCherry expression (Figure 4.7.1E-F).
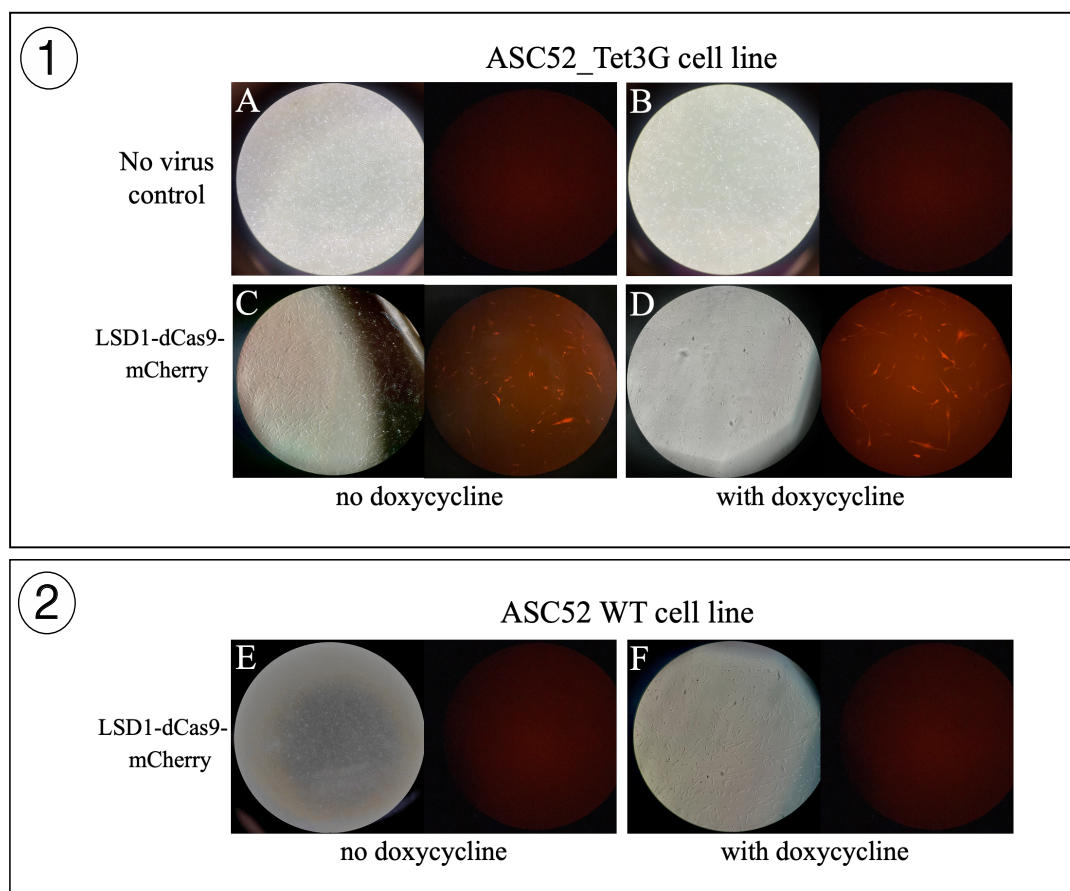
Figure 4.7.1: Assessment of mCherry-expression in ASC52_Tet3G transduced with LSD1-dCas9-mCherry cells. The ASC52_Tet3G MOI 15 cell line (1) and ASC52 WT (serving as control) (2) underwent transduction with LSD1_dCas9 MO1 35. Culturing conditions involved the presence and the absence of doxycycline (dox) because of the nature of the inducible Tet-On-3G system. The expression of LSD1, dCas9, and mCherry is regulated by the TRE3G promotor; it is exclusively activated in combination with the Tet-On-3G transactivator and dox. Dox was administered 48h post-transduction and viewed in the Nikon TE2000 with the TRITC 24h post-doc treatment. Figure made in Inkscape.

Subsequently, the ASC52_Tet3G cells and ASC52 WT cells transduced with LSD1-dCas9-mCherry were sorted using FACS to validate the impressions from the microscopy and to isolate doubly transduced, mCherry-expressing cells (Figure 4.7.2). Of all the ASC52_Tet3G cells transduced with LSD1-dCas9-mCherry and treated with dox, only about 3000 cells were mCherry-positive and sorted, constituting only 2.11% of the population. In contrast, approximately 11 000 mCherry positive cells were detected and sorted from the same cell line without dox, accounting for 8.4% of the cells (Table 4.7.1)

Intriguingly, and in contrast to the microscope observations, the ASC52 WT control cell line transduced with only LSD1-dCas9-mCherry also exhibited mCherry-positive cells despite Tet-On-3G lacking in these cells (Figure 4.7.1). In these cells, 10.4% of the cells cultured with dox were mCherry-positive, and 14.3% without dox, totaling around 60 000 cells which were isolated and pooled. The ASC52 WT control line which was not transduced with any virus was mCherry-negative as expected, these cells were also processed through FACS and sorted from the mCherry negative gate. These control cell lines, having undergone the same sorting conditions, would continue to serve as true controls for comparison with the mCherry-expressing cell lines. Thus, four cell lines were sorted in this FACS experiment (Table 4.7.1). These cell lines were seeded in bulk, successfully expanded as heterogeneous populations, and used for downstream transductions and testing.
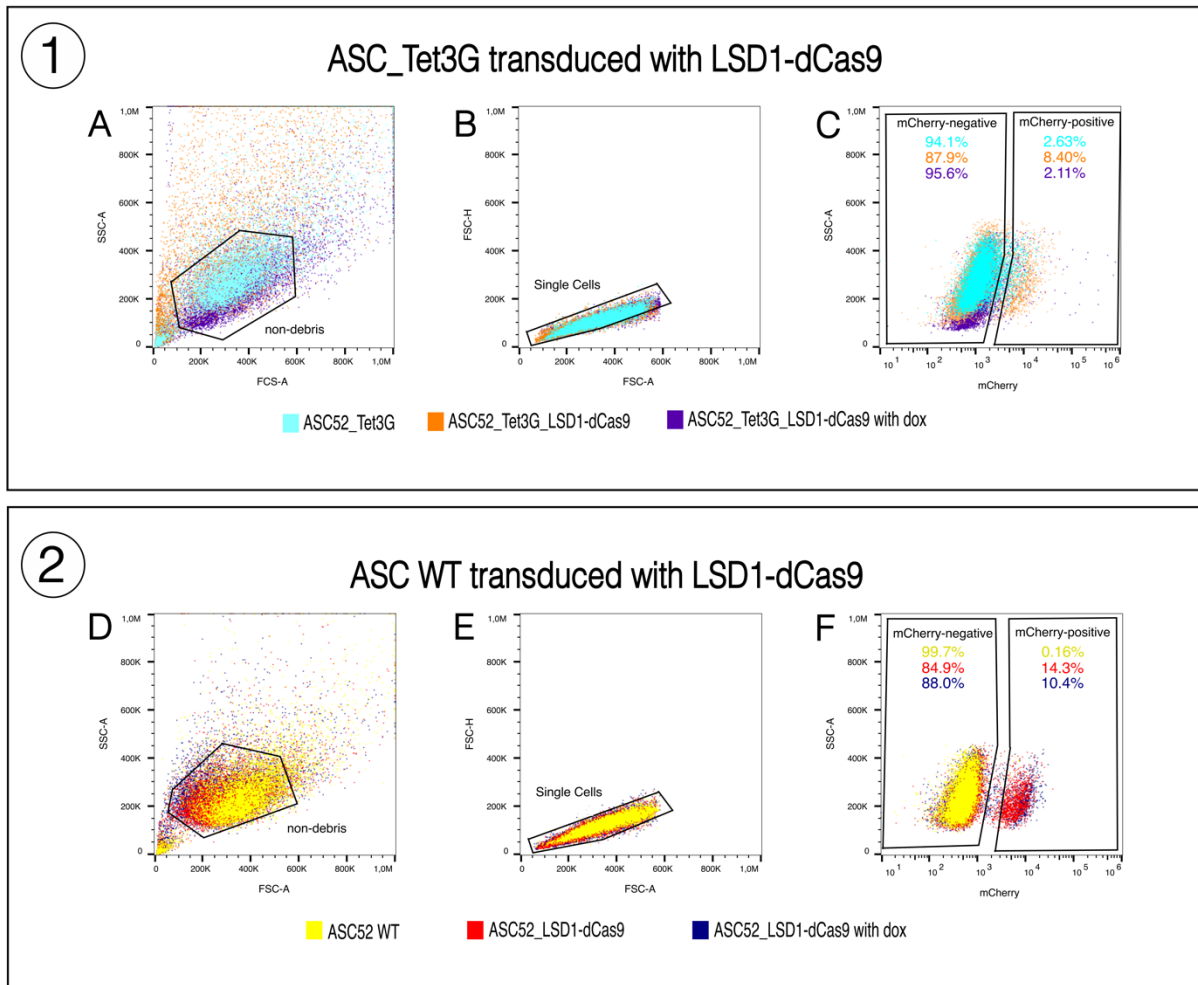
Figure 4.7.2: FACS of ASC52_Tet3G and ASC52 WT transduced with LSD1-dCas9-mCherry. ASC52_Tet3G cells (1) were transduced with LSD1-dCas9-mCherry to achieve double transduced cells, and ASC52 WT cells (2) were transduced as a control. A and D graph all events and the gating gates. B and E graph non-debris and gates for single cells, filtering out all the duplet cells. C and F show the graph of the single cells and gates for mCherry negative and mCherry positive cells. The samples were sorted with the SONY SH800 Cell Sorter. Data was analyzed in FlowJo™ Software v10.10 and annotated in Inkscape.

Table 4.7.1: FACS results for ASC_LSD1-dCas9

Sample	Annotation	Non-debris (Freq. of parent) [%]	Single cells (Freq. of parent) [%]	mCherry negative (Freq. of parent) [%]	mCherry positive (Freq. of parent) [%]
ASC52 WT	WT control	78.5	78.3	99.7	0.16
ASC52_LSD1-dCas9	Tet3G negative control	88.5	87,8	84.4	14.3
ASC52_LSD1-dCas9 with Dox	Tet3G negative control	85.6	84.9	88.0	10.4
ASC52_Tet3G	LSD1-dCas9 negative control	66.2	61.1	94.1	2.63
ASC52_Tet3G_LSD1-dCas9	Double transduced.	62.8	60.3	87.9	8.40
ASC52_Tet3G_LSD1-dCas9 with Dox	Double transduced. dox	70.0	66.8	95.6	2.11

Table 4.7.2: overview of cell lines sorted for mCherry

Cell line number	Cell line name	Tet-On-3G	LSD1_dCas9	Gate sorted
1	ASC52 WT	-	-	mCherry neg
2	ASC52_Tet3G	+	-	mCherry neg
3	ASC52_LSD1-dCas9	-	+	mCherry pos*
4	ASC52_Tet3G_LSD1-dCas9	+	+	mCherry pos

* Designed to be mCherry negative control, but was positive.

4.8. Transduction of lentiviral sgRNA-KRAB-zsGreen in ASC52 cells

Following the successful expansion of heterogeneous populations of ASC52_Tet3G_LSD1-dCas9, the next step was to transduce the cell line with the last virus (Virus C) containing sgRNA-KRAB-zsGreen1 (Figure 3.7.1) to establish a target cell line with stable expression of the complete inducible enCRISPR/Cas9i system. ASC52_Tet3G_LSD1-dCas9 cells, along with various control cell lines, ASC52_Tet3G, ASC52_LSD1-dCas9 and ASC52 WT (Table 4.7.2) were transduced with Virus C at MOI 15 and 24. The lentiviral sgRNA-KRAB cargo plasmid has zsGreen1 as a fluorescence marker. Cells expressing both zsGreen1 and mCherry were sorted using FACS (Figure 4.8.1 and Table 4.8.1) to isolate those successfully transduced with sgRNA-KRAB-zsGreen1 and to confirm the presence of both LSD1-dCas9-mCherry and sgRNA-KRAB-zsGreen1, thereby creating the target cell line.

Of all ASC52_Tet3G_LSD1-dCas9 cells transduced with sgRNA-KRAB-zsGreen MOI 15, around 27 000 cells were positive for both zsGreen and mCherry, constituting 16.0% of the total single cells. Using MOI 24, ca. 56 000 zsGreen and mCherry positive cells were sorted, constituting 17.5% of the single cell population (Figure 4.8.1 E and Table 4.8.1).
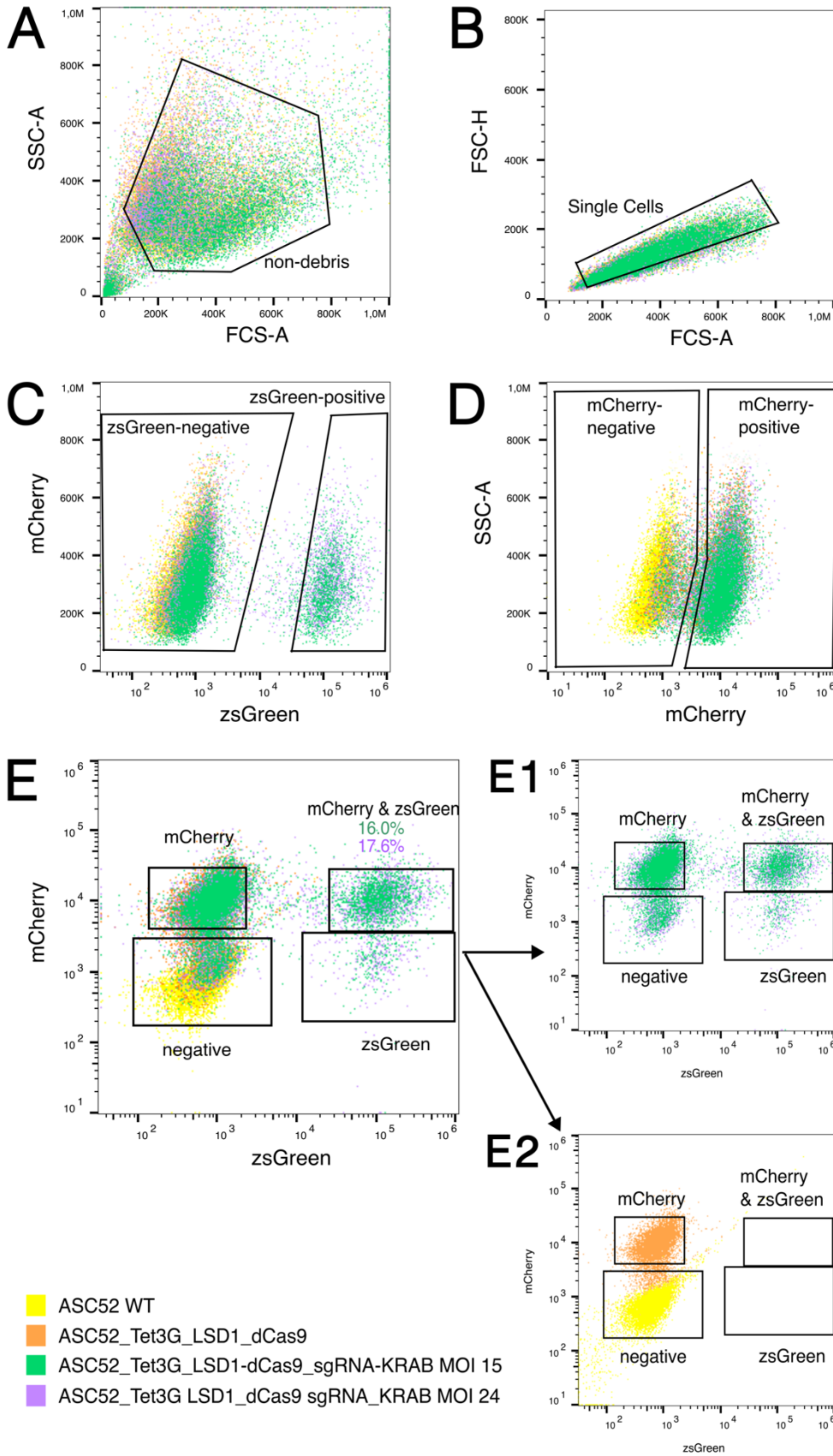
Table 4.8.1 FACS result of ASC52_Tet3G_LSD1-dCas9_sgRNA-KRAB

| Sample	Non-debris (Freq. of parent) [%]	Single cells (Freq. of parent) [%]	mCherry negative (Freq. of parent) [%]	mCherry positive (Freq. of parent) [%]	zsGreen negative (Freq. of parent) [%]	zsGreen positive (Freq. of parent) [%]	zsGreen & mCherry positive (Freq. of parent) [%]
ASC52 WT	85.1	89.9	99.9	0.073	99.9	0	0
ASC52_Tet3G_ LSD1-dCas9	77.1	84.8	153	79.4	99.9	0	0
ASC52_Tet3G_ LSD1dCas9_ sgRNA-KRAB MOI 15	73.5	90.0	10.6	84.8	79.4	18.2	16.0
ASC52_Tet3G_ LSD1dCas9_ sgRNA-KRAB MOI 24	78.9	83.3	11.1	84.4	77.7	19.7	17.5

Freq. of parent: frequency of parent, parent as in previous gate.

ACS52_Tet3G_LSD1-dCas9 cells transduced with sgRNA-KRAB

Figure 4.8.1: FACS of ASC52_Tet3G_LSD1-dCas9 transduced with sgRNA-KRAB-zsGreen, with control cell lines. ASC52_Tet3G-LSD1-dCas9 cells were transduced with sgRNA-KRAB-zsGreen MOI 15 (green) and MOI 24 (purple), and sgRNA-KRAB was validated with FACS and double positive cells for zsGreen and mCherry was isolated. ASC52_Tet3G-LSD1 cells and ASC52 WT cells non-transduced (orange and yellow, respectively) were also run through FACS. A) graphs all events and the gating gates. B) graphs non-debris and gates for single cells, filtering out all the duplet cells. C) graphs all single cells and gates for zsGreen negative and positive cells. D) shows the graph of the single cells and gates for mCherry negative and mCherry positive cells. The samples were sorted with the SONY SH800 Cell Sorter. Data was analyzed in FlowJo™ Software v10.10 and annotated in Inkscape.
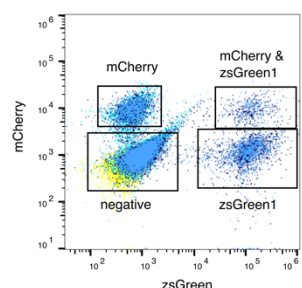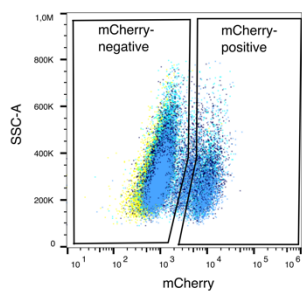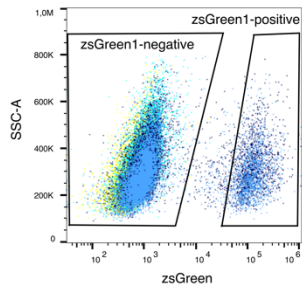
Transduction of Virus C (sgRNA-KRAB-zsGreen) was also performed on the control cell lines, ASC52 WT, ASC52_Tet3G, and ASC52_LSD1-dCas9 (Table 4.7.2). To verify the functionality of Virus C without interference from previously transduction cells, ASC52 WT cells were transduced and sorted with FACS for zsGreen fluorescence expression, yielding 9.92% and 12.4% zsGreen-positive cells of all single cells of the transduced cells with MOI 15 and MOI 24, respectively (Figure 4.8.2 and Table 4.8.2) This cell line served as the double-negative control, being negative for both zsGreen and mCherry. The ASC52_Tet3G cell line was transduced with Virus C at MOI 15 and MOI 24 and sorted for zsGreen, resulting in 15.7% and 15.8% zsGreen-positive cells, respectively (Figure 4.8.2 and Table 4.8.2), thus serving as a LSD1-negative control. Despite the cell line exhibiting about 22-24% mCherry positive cells and ~3% double positive cells, for this cell line to act as the control as expected, only the zsGreen positive and the mCherry negative gates were sorted. The last control cell line was transduction of ASC52_LSD1-dCas9, sorted for both zsGreen and mCherry, and yielded 10.8% and 12.8% zsGreen+mCherry-positive cells for transduction with MOI 15 and MOI 24, respectively (Figure 4.8.2 and Table 4.8.2). This cell line acted as the Tet-On-3G negative control. Including all the controls, 8 cell lines were generated in total (Table 4.8.3).

ACS52 WT cells transduced with sgRNA-KRAB
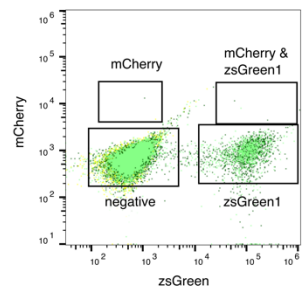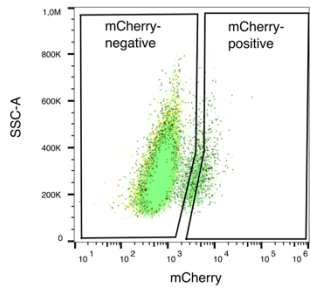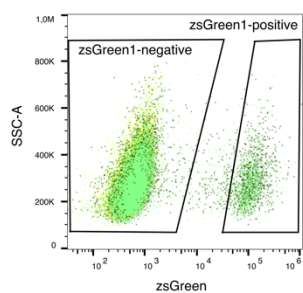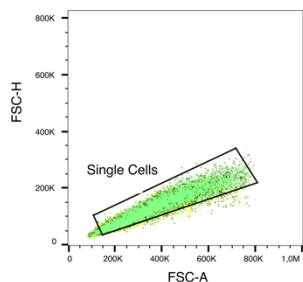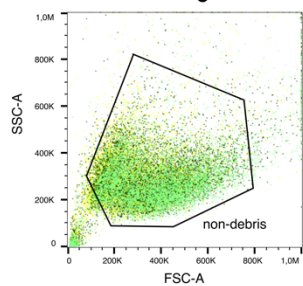
ACS52_Tet3G cells transduced with sgRNA-KRAB

ACS52_LSD1-dCas9 cells transduced with sgRNA-KRAB

ASC52 WT
ASC52_sgRNA-KRAB MOI 15
ASC52_sgRNA-KRAB MOI 24

ASC52 WT
ASC52 _Tet3G
ASC52 _Tet3G_sgRNA_KRAB MOI 15
ASC52 _Tet3G_sgRNA_KRAB MOI 24
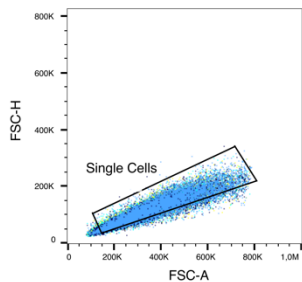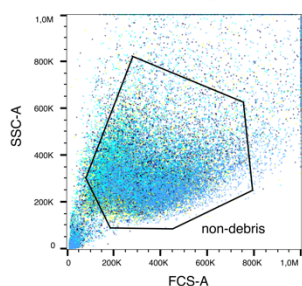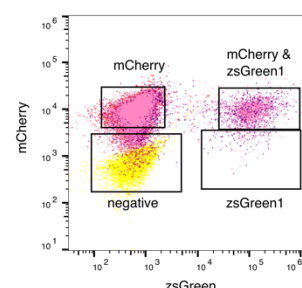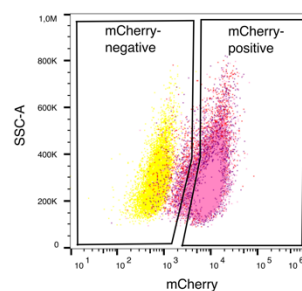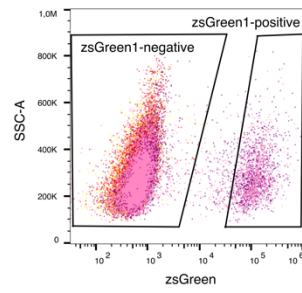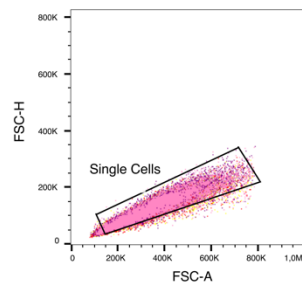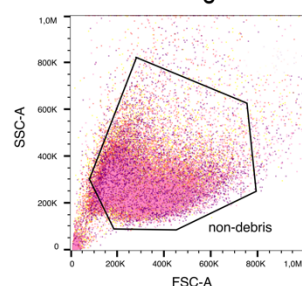
ASC52 WT
ASC52_LSD1-dCas9
ASC52_LSD1-dCas9_sgRNA-KRAB MOI 15
ASC52_LSD1-dCas9_sgRNA-KRAB MOI 24

70

Figure 4.8.2: FACS results of the control cell lines of the triple-transduced ASC52 cell line. Along with the transduction of the triple transduced cell line, which harbors the entire inducible enCRISPR/Cas9i system, various cells were obtained as controls. Transduction of sgRNA-KRAB-zsGreen1 with both MOI 15 and 24 was performed on ASC52 WT (yellow) ASC52_Tet3G (turquoise) and AS52_LSD1(Red), establishing the control cell lines of ASC52_sgRNA-KRAB (light and dark green), ASC25_Tet3G _sgRNA-KRAB (light and dark blue) ASC52_LSD1-dCas9_sgRNA-KRAB (pink and purple). The samples were sorted with the SONY SH800 Cell Sorter. Data was analyzed in FlowJoTM Software v10.10 and annotated in Inkscape.

Table 4.8.2 : FACS results for the control lines transduced with sgRNA-KRAB-zsGreen

Sample	Non-debris (Freq. of parent) [%]	Single cells (Freq. of parent) [%]	mCherry negative (Freq. of parent) [%]	mCherry positive (Freq. of parent) [%]	zsGreen negative (Freq. of parent) [%]	zsGreen positive (Freq. of parent) [%]	zsGreen & mCherry positive (Freq. of parent) [%]
ASC52 WT	85.1	89.9	99.9	0.07	99.9	0	0
ASC52_sgRNA-KRAB MOI 15	83.1	95.3	91.6	4.13	88.6	9.92	0.02
ASC52_sgRNA-KRAB MOI 24	84.9	93.7	89.7	4.89	85.9	12.4	5.22
ASC52_Tet3G	75.2	83.6	79.6	18.2	99.9	0	0
ASC52_Tet3G_ sgRNA-KRAB MOI 15	76.6	86.6	69.9	23.9	82.4	15.7	3.13
ASC52_ Tet3G_ sgRNA-KRAB MOI 24	78.7	83.9	70.8	22.5	82.0	15.8	2.94
ASC52_LSD1-dCas9	84.2	87.2	5.42	87.8	99.9	0	0
ASC52_LSD1-dCas9_sgRNA-KRAB MOI 15	86.9	94.2	4.89	88.6	87.2	11.1	10.8
ASC52_LSD1-dCas9_sgRNA-KRAB MOI 24	87.3	92.0	5.01	88.1	85.6	12.6	12.8

Freq. of parent: frequency of parent, parent as in previous gate

Table 4.8.3: overview of cell lines sorted for zsGreen and mCherry

Cell line number	Cell line name	Tet-On-3G	LSD1-dCas9	sgRNA-KRAB	Gate sorted
1	ASC52 WT	-	-	-	all neg
2	ASC52_Tet3G	+	-	-	all neg
3	ASC52_LSD1-dCas9	-	+	-	mCherry pos*
4	ASC52_Tet3G_ LSD1-dCas9	+	+	-	mCherry pos
5	ASC52_sgRNA-KRAB	-	-	+	zsGreen pos
6	ASC52_Tet3G_sgRNA-KRAB	+	-	+	zsGreen pos
7	ASC52_LSD1-dCas9_ sgRNA-KRAB	-	+	+	mCherry*& zsGreen pos
8	ASC52_Tet3G_LSD1-dCas9_sgRNA-KRAB	+	+	+	mCherry & zsGreen pos

* Designed to be mCherry negative control, but was positive.

4.9. Validation of LSD1-dCas9 and sgRNA-MCP-KRAB-zsGreen integration in ASC52 cells

By successfully isolating zsGreen and mCherry from the triple transduced cell line ASC52_Tet3G_LSD1-dCas9_sgRNA-KRAB, which harbors the complete inducible enCRISPR/Cas9i system, the next step was to confirm that the cells have integrated the system's genes, and that expression of these genes is occurring within the cell. RNA was harvested from said cell line and the various control cell lines (Tabel 4.8.3) and analyzed with qPCR. Primers specific to the system's components were used, including Tet3G, TRE3G, LSD1, dCas9, MCP, and zsGreen (Table 2.6). Even though Tet-On-3G was validated before, it was verified again to ensure that the Tet-On-3G transactivator was still present and had not been lost.

The results showed that all cell lines that were transduced with Virus A (Tet3G), did express the sequence for the Tet-On-3G transactivator, although this finding was only statistically significant in cells co-transduced with LSD1-dCas9 (Figure 4.9.1 A). For transduction of Virus B (TRE3G-LSD1-dCas9-mCherry), the cell lines were analyzed with three different primer pairs targeting TRE3G (expressed part of promoter), LSD1, and dCas9. While the TRE3G primers showed a significant expression of the TRE3G promoter from the TRE3G-

LSD1-dCas9-mCherry construct in the expected cell lines, the dCas9 primer yielded no results from the qPCR, and the LSD1 primers only showed a negligible overexpression in cells transduced with LSD1-dCas9 (Figure 4.9.1 C).

On the other hand, there is no doubt that transduction of Virus C (sgRNA-MCP-KRAB-zsGreen) was successful, as expression of both MCP and zsGreen in these cell lines were found to be significant compared to all the control cell lines (Figure 4.9.1 D-E).

Taken together with the previous data, the results shows that Tet3G most likely was successfully introduced into the genome of the transduced cells and that those cells are expressing the Tet-On-3G transactivator. The same can be said for the sgRNA-KRAB construct. However, expression of the LSD1-dCas9 construct is uncertain, as neither dCas9 or LSD1 primers detected its presence in any cell line, even though the presence of mCherry was confirmed with FACS and the expression of the TRE3G promoter was verified with qPCR.

Virus transducted ASC52 cell lines
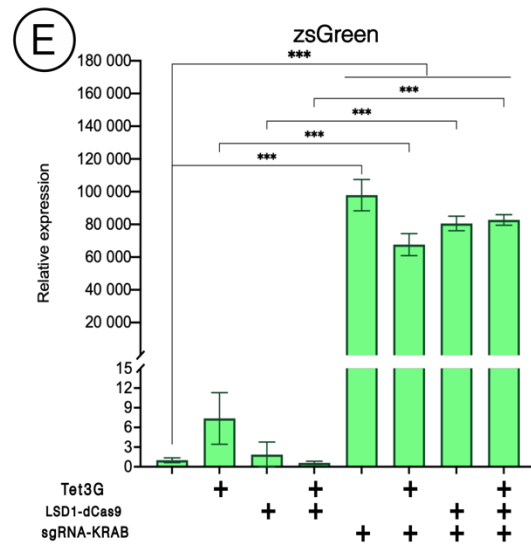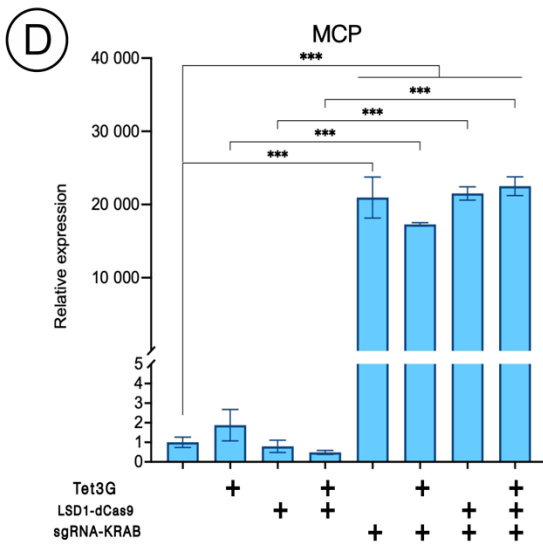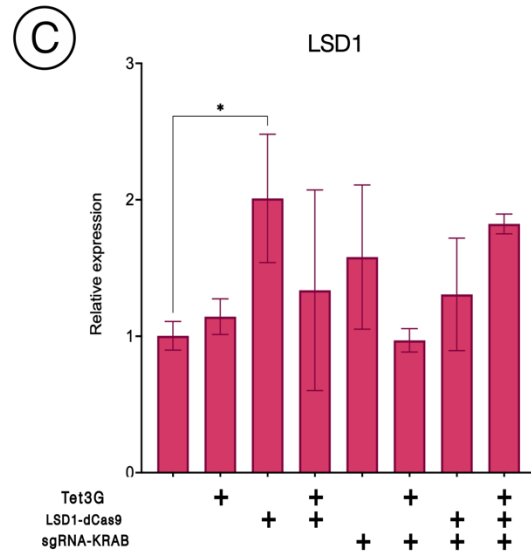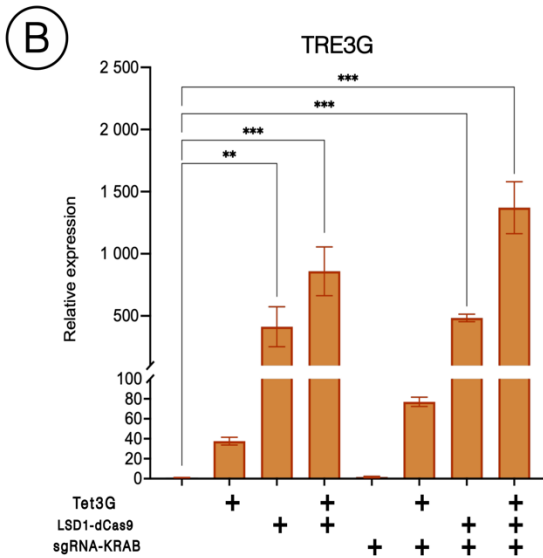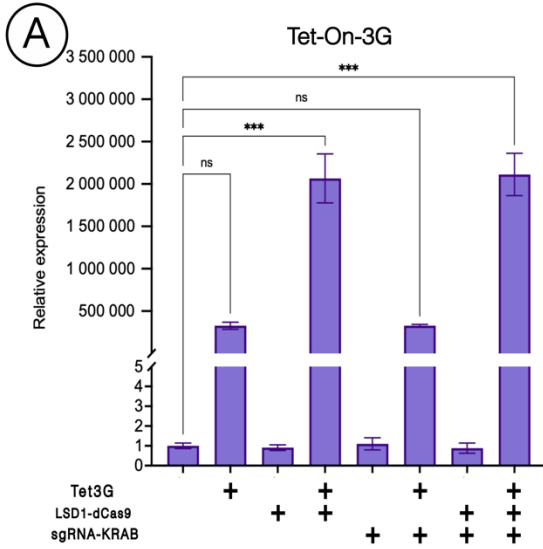
Figure 4.9.1: Validation of lentiviral transduction of ASC52 cells. Initially, ASC52 WT was transduced with virus A (Tet3G-Blasticidin), followed by a selection of positive clones using blasticidin. Then, these cells were transduced with virus B (LSD1-dCas9-mCherry), and mCherry-positive cells were sorted. Finally, the cells were transduced with Virus C (sgRNA-KRAB-zsGreen) and sorted based on the fluorescent marker zsGreen. Alongside the target cell line ASC52_Tet3G-LSD1-dCas9_sgRNA-KRAB, which harbors the complete enCRISPR/Cas9i system, several control cell lines were generated with varying combinations of viral transductions. These controls include ASC52_Tet3G, ASC52_LSD1-dCas9, ASC52_Tet3G-LSD1-dCas9, ASC52_sgRNA-KRAB, ASC52_Tet3G-sgRNA-KRAB, and ASC52_LSD1-dCas9-sgRNA-KRAB. qPCR was performed to analyze the components of the enCRISPR/Cas9i construct using specific primers for the genes Tet3G (A), TRE3G (B), LSD1 (C), MCP (D), and KRAB (E), with the housekeeping genes RPS13 and RPLPO for normalization. The y-axis represents the relative expression normalized to the reference gene, which was RPS13 for all except Tet3G, where RPLPO. The data are shown as mean ± SD (n = 3) Statistical analyses were conducted using one-way analysis of variance (ANOVA), with significance indicated as follows ns: non-significant, * P<0.033, ** P<0.002, P< 0.001. The figure was made with GraphPad Prism and annotated using Inkscape.

4.10. WB targeting the enCRISPR/Cas9i and HMBS proteins

Because the qPCR analysis of the generated cell lines was not conclusive regarding LSD1-dCas9 expression, proteins were harvested from the cells and western blotting was conducted using antibodies targeting LSD1, Cas9, Tetracycline repressor (TetR), and HMBS, the suspected target of the enCRISPR/Cas9i-mediated repressed enhancer, along with beta-actin as a loading control. Of note, no antibody targeting KRAB or MCP was available. Before staining with antibodies, Ponceau S staining was performed to verify the transfer of the proteins from the gel to the membrane. While proteins up to 75 kDa were well transferred, proteins above this size were less efficiently transferred (Figure 4.10.1).
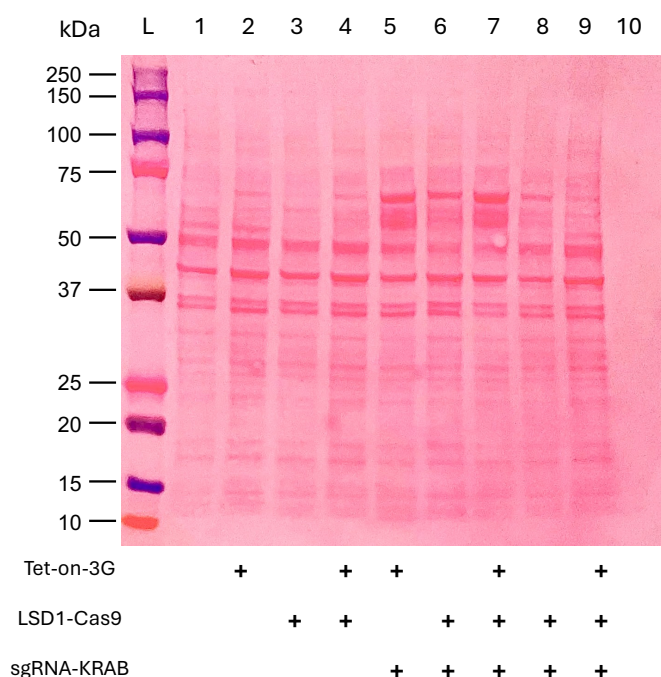
Figure 4.10.1: Verification of protein transfer from the SDS-gel to the membrane with Ponceau S staining. Lane L is the kaleidoscope protein ladder used to estimate protein size. Lane 1 contains the ASC52 WT for control. Lanes 2-9 contain the transduced ASC52 cell lines of Tet3G, LSD1-dCas9, and sgRNA-KRAB in different combinations, as indicated below the figure. Lanes 6 and 8 are transduced with the same set of proteins, only with different MOI of sgRNA-KRAB, 15, and 24, respectively. The same with lanes 7 and 9; these two have all the lentiviral cargo plasmid incorporated. Lane 10 is the Cas9 protein positive control.

By staining the membrane with the TetR antibody, the presence of Tet-On-3G can be confirmed, indicated by a band at approximately 23.4 kDa. However, no specific band corresponding to overexpression of Tet-On-3G could be detected (Figure 4.10.2 A). The second transduction involved the LSD1-dCas9 cargo plasmid. The LSD1 protein is estimated to be 112.8 kDa, and the dCas9 is 158.2 kDa. In this construct, the dCas9 is a fusion protein with LSD1, resulting in an expected band size of 271 kDa when the membrane is stained with Anti-LSD1 or Anti-Cas9. The anti-LSD1 antibody revealed bands with the size of LSD1 alone, corresponding to endogenous LSD1 proteins. However, no bands corresponding to the LSD1-dCas9 fusion protein were detected (Figure 4.10.2 A). In line with this finding, no band was obtained for the transduced cells using anti-Cas9, while a recombinant purified Cas9 control sample resulted in a strong band of expected (lower) size (Figure 4.10.2 B). The β-actin loading control verified the presence of proteins from the transduced cells and demonstrated equal loading between the samples. Finally, the HMBS antibody did not reveal any band of the expected molecular weight of 21 kDa, but rather showed a band at ~ 18 kDa. However, there was no suppression of this band in the triple-transduced cells compared to the controls.
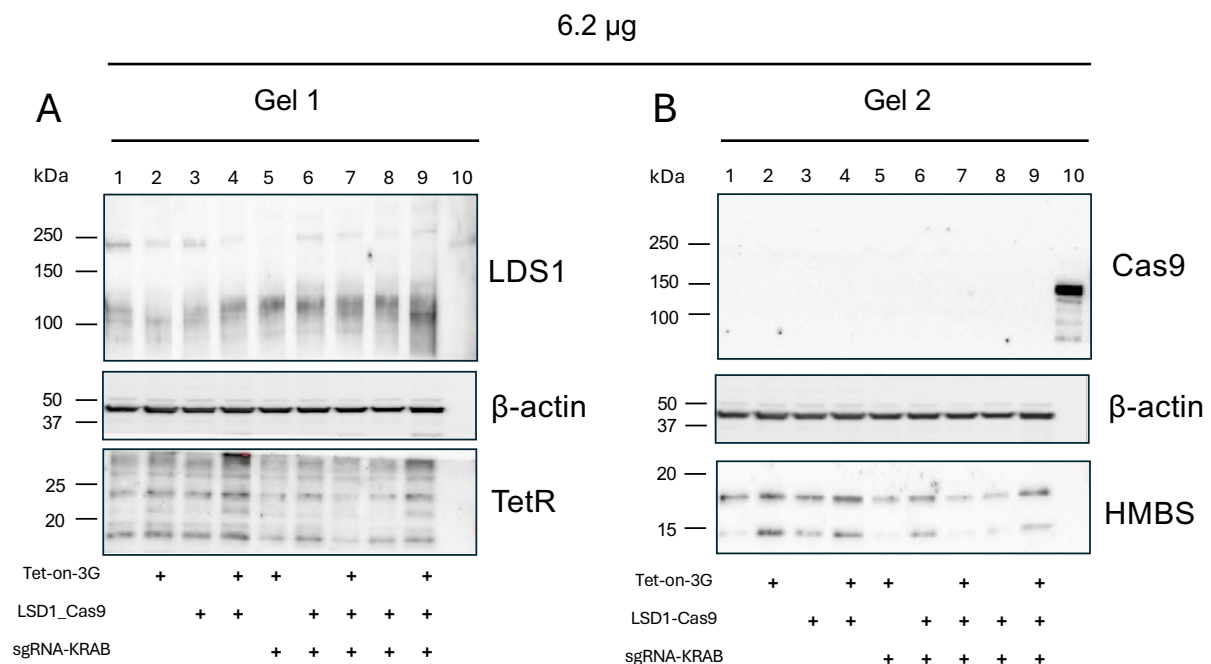
Figure 4.10.2: Western blot for triple transduced cells and its various controls. The ASC52 cells were transduced with MOI X Tet3G, Moi Y LSD1-dCas9 and Moi X sgRNA-KRAB in different combinations, creating many control cell lines and the target cell, which harbors the whole enCRISPR/Cas9i-system. The proteins were purified with the NORGEN RNA/DNA/protein kit. The SDS-Page was run with 6.2 μg protein and ran with 200V for 30 min. Lane 1 is the ASC52 WT, as control. Lanes 2-9 are the target samples with triple transduced cell lines and their respective controls. Lane 2 is the ASC52_Tet3G. Lane 3 is the ASC52_LSD1. Lane 4 is the double-transduced cell line ASC52_Tet3G_LSD1-dCas9. Lanes 5-7 harbors the same cell line as lanes 2-4, but they are

also transduced with sgRNA-KRAB with MOI W. Lanes 8-9 have the same makeup as lanes 6 and 7 only with transduced with MOI Y sgRNA-KRAB. Lane 10 is a Cas9 recombinant purified protein positive control. The membrane was stained with the antibodies Anti-LSD1/KMD1 (A), Anti-TetR (C), Anti_Cas9 (D), and Anti-HMBS (F), and Anti-β-actin (B, E) as an internal control. A protein ladder kaleidoscope was used to estimate the protein size.

5. Discussions

Although significant progress has been made during the last decade in mapping out the functional elements in our DNA, it remains a major challenge in genomics research to identify enhancers and their target genes (Dunham et al., 2012, Gray et al., 2015). This has a direct negative impact on translating results from the rapidly growing number of GWAS being performed since >90% of tag-SNPs reported from these studies are found in non-coding regions, with a large proportion believed to be situated in thus far unidentified enhancers (Peña-Martínez and Rodríguez-Martínez, 2024, Gray et al., 2015). Therefore, translating the association signals identified in GWAS into causal biological mechanisms has become a major challenge and a bottleneck in understanding genetic susceptibility to disease (Barroso and McCarthy, 2019).

For example, although general obesity is known to be highly inheritable (Elks et al., 2012, Maes et al., 1997) with more than 1000 loci associated with BMI (Loos and Yeo, 2022), only a few hundred loci have been associated with *visceral* obesity (Karlsson et al., 2019), which is the pathogenic form of obesity. So far, no biological mechanisms have been established for these associated loci, so there is a great need to functionally interpret these associations.

To address the limitations of GWAS, using visceral obesity as a model disease, our lab has identified a novel enhancer in one of the associated loci (11q23.3,) which harbors a set of five SNPs which all individually affect the enhancer activity *in vitro* (Krill, 2023, Mirza, 2022, Samuelsen, 2021). To link these SNPs to the target gene(s) of the enhancer, this thesis aimed to establish a CRISPR/Cas9-based epigenetic editing method to repress the enhancer in adipose-derived mesenchymal stem cells, the cell type where the enhancer has been found to be active. By analyzing global gene expression in the cells with and without repression of the enhancer, the target genes can be identified.

Epigenetic data suggest the enhancer is only active in undifferentiated mesenchymal stem cells and not during differentiation towards mature fat cells (Krill, 2023, Mirza, 2022, Samuelsen, 2021). To experimentally prove this, the enhancer repression should be inducible so that repression could be limited to various stages of differentiation, including the undifferentiated state. Therefore, a lentiviral-based, dox-inducible CRISPR/Cas9-mediated enhancer repression system described by Li et al. (2020) was selected and adjusted to target our enhancer of interest in a commercial mesenchymal stem cell line (ASC52telo).

This system is complicated, consisting of multiple components that have not been optimized for use with the ASC52telo cells. In this thesis, many, but not all, of these components were successfully introduced into these cells. In the following sections, key issues encountered are discussed, including plasmid quality control, the inducible system, single-cell clonal isolation, and the transduction efficiency of the lentiviruses.

5.1 Quality control of plasmids

A previous attempt at establishing the CRISPR/Cas9 mediated epigenetic repression of the 11q23.3 enhancer in ASC52telo cells in the lab failed, as neither Tet-On-3G nor LSD1-dCas9 were found to be expressed after a pilot lentiviral transduction (Krill, 2023). Therefore, in the present thesis, an emphasis has been placed on testing each component of the system to identify (Gray et al., 2015) the likely cause of previous problems and find solutions to overcome them. In particular, the Tet-On-3G-BFP construct and the lentiviral production/transduction protocols were the main suspects and the starting point of the troubleshooting.

Sequencing of the plasmids revealed that the Tet-On-3G-BFP plasmid previously used did not contain the right sequence, which is likely one major explanation for the previous failed experiment. Therefore, a different lentiviral Tet3G construct was obtained, sequenced and experimentally tested by transient overexpression before making lentiviruses. The previously used LSD1-dCas9-mCherry construct was also tested in the same way. Sequencing revealed that the new Tet3G plasmid was correct and validated the previously sequenced LSD1-dCas9-mCherry construct. Accordingly, overexpression in HT1080 cells revealed the presence of both Tet3G, LSD1, and dCas9 as analyzed by qPCR and fluorescence microscopy (Figure 4.2.1-4.2.2), thereby confirming the plasmids were correct and functional.

5.2. Issues with the inducible system and doxycycline

The LSD1-dCas9 plasmid is designed to be a part of an inducible system controlled by the TRE3G promoter, which is activated by the Tet-On-3G transactivator protein in the presence of doxycycline. The TRE3G promoter is positioned upstream of *LSD1* and *dCas9* (as well as *mCherry*). RNA polymerase binds to the promoter in the presence of Tet-On-3G and dox, resulting in the transcription of the downstream genes. Because doxycycline is derived from tetracycline, the latter can also activate the Tet-On-3G activator. This is a potential problem

since the FBS added to the cell culture medium is usually derived from countries like Brazil or USA, where cattle are often given high doses of antibiotics (Martin et al., 2015). Thus, regular FBS often contains sufficient levels of tetracycline to cause background activity in Tetracycline-based inducible systems (ThermoFisher, n.d.). Therefore, special FBS with certified absence of tetracycline was used in the medium of the HT1080 during the transient overexpression.

Dox was expected to only affect the expression of LSD1-dCas9-mCherry. This was indeed observed in the transient overexpression through increased mCherry expression, as detected by fluorescent microscopy (Figure 4.2.1), and dCas9 expression, as measured by qPCR in dox-treated cells transfected with both Tet3G and LSD-dCas9-mCherry (Figure 4.2.2). Dox treatment resulted in four-fold higher expression levels of dCas9 compared to vehicle and no other sample responded to dox. Intriguingly, however, LSD1 expression was not increased in response to dox. This was highly unexpected since LSD1 and dCas9 are part of the same fusion protein. One possible explanation might be a technical error in cDNA synthesis or qPCR since both dCas9 and mCherry responded to dox.

Moreover, the surprisingly high expression of LSD1 and dCas9 in the absence of Tet-On-3G (and thereby also regardless of dox) raises questions about the construct's design. The whole point of inducible LSD1-dCas9 expression is to have a low basal expression in the absence of (activated) Tet-On-3G, but these data suggest either a very leaky construct where the tetracycline-responsive TRE3G promoter also responds to endogenous transcription factors in the HT1080 cells, or the presence of an additional promoter that is active when the lentiviral construct is transiently expressed but not packaged into lentiviral particles. The latter was found to not be true since a similar result was obtained in lentivirus-transduced cells (Figure 4.7.2). Alternatively, contamination with Tet-On-3G could be imagined. The qPCR results show a marginal expression of Tet-On-3G where LSD1-dCas9 was supposed to be overexpressed alone, but this is more likely the effect of primer dimers being detected in the qPCR reaction rather than actual cross-contamination, especially when considering how much higher the Tet-On-3G levels were in the Tet-On-3G positive samples and the fact that LSD1-dCas9 was roughly equal in Tet-On-3G positive and negative samples.

The ASC52 cells were cultured in a standard ASC52 medium containing low serum (2% FBS). However, this medium was not tetracycline-approved for Tet-systems, and may contain trace amounts of tetracycline that can activate the Tet-on system. Because the FBS was included in a stem cell supplement containing several other components, it could not simply be replaced with

the Tet-approved FBS used for the HT1080 cells. This likely resulted in high background, as evidenced by FACS analysis of ASC52_Tet3G cells transduced with LSD1-dCas9-mCherry (Figure 4.7.2). Of note, ASC52 WT transduced with LSD1-dCas9-mCherry also exhibited mCherry-positive cells despite the absence of both Tet-On-3G and dox (Figure 4.7.2). This demonstrates the inherent leakiness of the Tet-On-3G system, a phenomenon also observed in the HT-1080 transient overexpression experiments, as mentioned above.

The introduction of dox did not increase the number of mCherry-positive cells, indicating that tetracycline may be present and thus activate the Tet-on system. On the contrary, of all cell lines, the added dox yielded fewer positive clones after transduction, suggesting an adverse effect of dox on the cells. During the cultivation of the various transduced ASC52 cell lines, those cultivated in the presence of dox exhibit slower growth compared to the same cell line in normal medium without dox, as observed in the fluorescence microscopy images. Likewise, FACS analysis showed that cell lines treated with dox had fewer mCherry-positive cells compared to their non-dox counterparts.

Tetracyclines function by inhibiting bacterial protein synthesis and growth through binding to the bacterial 30S ribosomal subunit, thereby preventing aminoacyl tRNA from attaching to the ribosome (Chopra and Roberts, 2001, Das et al., 2016). Tetracyclines, as an antibiotic, exhibit great activity on a wide range of bacteria (Chopra and Roberts, 2001) because tetracyclines are predominantly absorbed by bacterial cells and not by eukaryotic cells; their effect is mainly considered specific to bacteria (Sanchez et al., 2020). This assumption led to the use of tetracyclines in eukaryotic experimental gene expression systems, such as the Tet-On-system for regulation transcription (Sanchez et al., 2020). Nowadays, dox is preferred in Tet systems due to its greater stability compared to tetracycline (Honnorat-Benabbou et al., 2001)

One concern regarding the use of dox in mammalian cell lines is its potential to disrupt mitochondrial function and affect characteristics like cell proliferation (De Boeck and Verfaillie, 2021). Ahler et al. (2013) discovered that dox alters the metabolism and proliferation of human cell lines. Metabolic pathway enrichment analysis revealed significant enrichment of pathways such as oxidative phosphorylation and glycolysis in dox-treated cell lines, with robust changes in gene expression in the presence of dox (Ahler et al., 2013). Fife et al. (1998) showed cell death with a 10 µg/ml dosage of dox and a three-fold reduction of proliferation when the dosage is halved (no other dosage tested) for human prostate adenocarcinoma cells. Similarly, Ahler et al. (2013) found that 1 µg/ml significantly reduces proliferation in most of the human

cell lines tested, while no significant finding was related for 100 ng/ml dox-treated cells. Moullan et al. (2015) reported that tetracycline disturbs the mitochondrial function of eukaryotic models. Collectively, these studies highlight the potential drawbacks of using dox in cell cultivation. In our Tet-On-3G inducible system, 1 µg/ml dox was used following the protocol of Li et al. (2020). One improvement may be to minimize the dose of dox; Zhou et al. (2006) have found evidence that 100 ng/ml dox can successfully induce a Tet-system-based overexpression (De Boeck and Verfaillie, 2021)

5.3. Issues with clonal selection

Despite our methods for clonal selection of positive transduced ASC52 cells being unsuccessful, these tested techniques have proven effective in other cell lines. Another method not tested was limiting dilution, which establishes clonal populations by highly diluted cell suspensions to achieve a density of less than cone cell per aliquot (Underwood and Bean, 1988). However, this method has its own limitations, as it does not guarantee that colonies originate from a single cell due to the difficulty of ensuring that each well contains only one cell due to the challenge of identifying single cells under a microscope (Ye et al., 2021)

However, the primary issue with our single-cell isolation is not the technicality of the method but rather the inherent growth characteristics of transduced ASC52 cells. These cells do not thrive in isolation, often migrating towards neighboring cells when seeded sparsely. Single cells isolated by FACS in separate wells did not form colonies even when growth factors were added. This issue was compounded by the fact that FACS-based single-cell isolation typically relies on cells expressing fluorescent markers for sorting (Basu et al., 2010). Our heterogeneous population of ASC52_ Tet3G that we attempted this with did not contain fluorescent markers, and no viability dyes were used, raising questions about the accuracy and viability of cells sorted into wells.

One potential solution could be the adaptation of an indirect co-culture system using transwells. Traditionally, this method is used to cultivate two different cell types in the same culture medium but separated by a membrane, allowing communication via secretory factors, and cell signaling molecules are the main focus (Rasouli and Safari, 2024)). Seeding a single cell in the top layer of the transwell and the bottom layer with ASC52 cells at optimal density would be a possible design. This approach aims to leverage the signaling from the ASC52 cell population, encouraging the isolated single cells for cell proliferation and colony formation.

5.4. Transduction efficiency of produced lentiviruses and FACS

Another important point that needs to be highlighted is the low transduction efficiency of the LSD1-dCas9 virus when transduced in ASC52_Tet3G cells, resulting in only 2.1 % and 8.4 % positive mCherry-expressing cells with and without dox, respectively (Table 4.7.1). Meanwhile, the transduction of the same virus in ASC52 WT cells yielded 10.4% and 14.3% mCherry-positive cells, respectively (Table 4.7.1). In contrast, the transduction of virus containing sgRNA-KRAB-zsGreen1 had better transduction efficiencies, ranging from 10-20% (Table 4.8.1-4.8.2) for all cell lines transduced. Interestingly, here, the double transduced cell lines with Tet3G and LSD1-dCas9 yielded the highest levels of transduction efficacy of 18.2% and 19.7% for MOI 15 and 25, respectively.

All MOI calculations were based on the functional titer test of the transduction of Tet3G-BSD in ASC52 WT cells with blasticidin selection, where manual counting of the surviving cells was performed at the end using by a light microscope. This method has accuracy limitations due to potential human error and the difficulty in distinguishing live cells from dead cells and debris. Therefore, the estimation of the functional titer is a rough estimation. One improvement could have been using the IncuCyte® Live-Cell Analysis System to monitor the cell throughout the selection process and counting with the calculated confluency of the cells after the selection period. Staining with live/dead markers detectable by the system would also enhance accuracy. Conversely, a titer test with a virus containing LSD1-dCas9 or sgRNA-KRAB-zsGreen, which includes fluorescent markers, would simplify the process. Flow cytometry, like FACS, could be employed in this case. The virus stock would be diluted and transduced as before, and FACS would be used to count both transduced and non-transduced cells, yielding more accurate results. This would have made the calculation more accurate, equation 3.8.1. could also have been used here as it is based on virus transduction with fluorescent markers. In the functional titer test, the number of transduced cells was estimated under the assumption that the cells doubled in one day, as they were seeded one day prior to transduction. While an attempt was made to count the cells, since the test was conducted in a 96-well plate, the number of cells was too little to give an inaccurate result with the Countess.

The transduction of the lentivirus containing the LSD1-dCas9-mCherry was only performed for one MOI based on the titer test from the previous virus. This is not optimal as different lentivirus transduction efficacy and a titer test on each virus should have been performed to find the

optimal MOI for each virus. However, one MOI was selected to keep the total number of generated cell lines to a manageable level.

Of note, the FACS analysis of the LSD1-dCas9 negative control cell line (ASC52_ Tet3G) revealed a small portion of mCherry positive cells of 2.63%, which is comparable to the transduction of ASC52_Tet3G cells with LSD1-dCas9 (2.1% in the presence of dox). Of note, most of the cells in these gates were found along the left-hand side of the gate (Figure 4.7.2). These results raise questions about the accuracy of the mCherry negative gate. The mCherry positive gate was established using the double-transduced ASC52_Tet3G_LSD1-dCas9 cell line, while the negative gate was based on ASC52 WT cells. Given that ASC52_Tet3G cells without transduction also yielded mCherry-positive signals, it is possible that the gate was slightly incorrectly set, leading to false-positive sorting in other mCherry-positive cell lines.

Additionally, the gating strategy for non-debris cells differed between ASC52_WT and ASC52_Tet3G cells transduced with LSD1-dCas9 due to inherent differences in their characteristics and light scatter properties as detected by FACS. This variation in gating could explain the discrepancies and potential inaccuracies. Considering the possibility of false-positive mCherry signals in the transduced ASC52_Tet3G_LSD1-dCas9 cells, the actual transduction efficiency of the virus may be closer to 5%. Which is about less than half of the transduction efficiency of the sgRNA-KRAB, where the successful integration of *MCP* and *zsGreen* was detected with qPCR of RNA with an indication of very high overexpression (Figure 4.9.1)

This raises the question about the presence of LSD1-dCas9 in the genome. qPCR analysis, using specifically designed primers (Table 2.6), failed to detect any dCas9 expression. While endogenous LSD1 expression was detected, showing that the primers worked, no overexpression of LSD1 was observed (Figure 4.9.1). Likewise, no LSD1-dCas9 fusion protein was detected by WB in the same samples by antibodies targeting either protein. However, the WB findings may be explained by incomplete transfer to the membrane, as discussed below. Moreover, the mCherry tag of the LSD1-dCas9 construct was expressed in 88% of the FACS-enriched cells (Table 4.8.2). In addition, the qPCR analysis of the construct's TRE3G promoter also yielded significant overexpression for cell lines with the transduction of LSD1-dCas9. Although it is puzzling that the promoter is transcribed, the primers bind very close to the open reading frame and, therefore, likely downstream of the transcription start site. Thus, whether

LSD1-dCas9 is truly present or not in cells supposed to have it has not been unambiguously determined.

An additional piece of evidence would be to analyze the gDNA harvested from the triple-transduced cell line with qPCR, like that which was done for the ASC52_Tet3G, to search for integration of LSD1 and dCas9. However, due to time restrictions, this analysis was not performed in this thesis. This analysis can truly confirm the presence (or absence) of the LSD1-dCas9 integration in the genome of ASC52 transduced cell lines. Currently, it remains unclear whether the issue lies in the integration or expression of the genes.

5.5. Troubleshooting of Western Blotting

A series of issues with the WB resulted in many unspecific bands or no band at all, complicating the verification of the protein of interest and its presence. The Ponceau S staining of the membrane revealed poor transfer of the larger protein located on the top of the membrane (Figure 4.10.1). Meanwhile, the proteins in the middle and bottom were transferred well. The Trans-Blot® Turbo™ Transfer System is not optimal for all proteins (Bass et al., 2017); in this case, it struggles with proteins of higher molecular weight. Even after doubling the transfer time from the manufacturer's protocol and achieving complete ladder transfer, staining indicated otherwise. Bands of larger molecular weight require longer transfer time compared to smaller proteins (Liu et al., 2023). The traditional wet blotting method could potentially yield better results for larger proteins. Wet transfer involves creating a "sandwich" with cassette backing, pad, blotting paper, gel, and membrane. While this method is time-consuming, it allows for customizable transfer times and buffer compositions specific to the protein of interest. The buffer composition significantly affects transfer efficiency, particularly for high molecular weight proteins (Bass et al., 2017). The LSD1-dCas9 fusion protein, for instance, is a very large protein with an expected molecular weight of 270 kDa. On the other hand, prioritizing the transfer of the larger proteins, the transfer of the proteins with low molecular weight may be lost. This poses a challenge since TetR and HMBS proteins of small molecular weight, at approximately 23 kDa and 21 kDa, respectively.

On discussion point the method of lysis of the transduced ASC52 cells. The samples were harvested with NORGEN DNA/RNA/protein purification plus kit, whereas DNA, RNA and protein is extracted from the same sample. This kit allows for the simultaneous extraction of DNA, RNA, and protein from the same sample, ensuring that all analyses are performed on

identical samples with minimal variation. However, despite the advantages of this approach, there may be some drawbacks. While the DNA and RNA yields from this purification method are of high quality, the same may not be true for the protein yields. Results from our lab suggest that the NORGEN-harvested proteins differ from those obtained using RIPA lysis, as indicated by Ponceau S staining. An experiment conducted on HT-1080 cells integrated with Cas9 nickase compared the two lysis methods. When stained with the anti-Cas9 antibody, RIPA lysis demonstrated a significantly higher expression of Cas9 compared to NORGEN. Therefore, while the NORGEN method is useful for DNA and RNA extraction, it may not be the best choice for protein extraction.

Additionally, the WB was conducted without positive controls for the different antibodies targeting the protein of interest, except for Cas9. Due to the presence of numerous nonspecific bands in the antibody staining for LSD1 and TetR (Figure 4.10.2), it was impossible to distinguish the protein of interest from other false positive bands without a positive control. A positive control could have been created by lysing cells transiently overexpressing the cargo plasmid in HT-1080 cells and loading these lysates in parallel with the protein lysate from the transduced ASC52 cells.

However, the WB revealed (Figure 4.10.2) nice bands for the control protein β-actin and HMBS. The two bands observed for HMBS likely represent its two different isoforms (long and short). There was no suppression of HMBS in the samples where cells were triple-transduced, which was expected since no LSD1-dCas9 was detected by WB or qPCR.

6. Conclusion

The establishment of an inducible CRISPR/Cas9 enhancer repression system in the AdMSC cell line ASC52telo was partially achieved in this thesis. However, a successful integration and expression of the LSD1-dCas9 fusion protein could not be clearly demonstrated. Thus, epigenetic inactivation of the 11q23.3 enhancer associated with visceral adipose mass was not achieved, preventing identification of the target gene(s) of the enhancer.

However, this thesis has established a foundation for the continued development of the enhancer targeting CRISPR/Cas9 inactivation method, as several of the previous challenges have been overcome:

- A previously used lentiviral cargo plasmid was found to be incorrect.
- New lentiviral cargo and packaging plasmids were validated to contain correct inserts.
- Three lentiviruses carrying Tet-On-3G, LSD1-dCas9, and sgRNA-KRAB constructs, respectively, were successfully made.
- ASC52telo cells were successively transduced with the three lentiviruses, and heterogeneous pools of positively transduced cells were sorted and expanded.
- The presence of Tet-On-3G and sgRNA-KRAB was verified in triple-transduced cells.

7. Future perspective

The leakiness of the Tet-On-3G inducible system needs to be resolved. Also, the fact that the dCas9 construct fused with LSD1 was not confirmed in the target cell line requires future attention. The natural next step is to analyze the gDNA harvest from the triple-transduced cells and investigate the integration of *LDS1* and *dCas9*. Alternatively, repeating the transduction of LSD1-dCas9 in the cells validated to express the gRNA and KRAB could be performed using different amounts of MOI. This can be done in both Tet3G positive and negative cells, however, due to the LSD1-dCas9 construct's apparent independence of Tet3G, it may be easiest to abandon the inducible strategy only use Tet3G negative cells Additionally, a method of establishing single-cell clones of the transduced cells should be further explored.

8. References

Addgene. 2023. *Fluorescence Titering Assay for Lentivirus* [Online]. Available: https://www.addgene.org/protocols/fluorescence-titering-assay/ [Accessed 22.04 2024].

Aguet, F., Barbeira, A. N., Bonazzola, R., Brown, A., Castel, S. E., Jo, B., Kasela, S., Kim-Hellmuth, S., Liang, Y., Oliva, M., et al. 2019. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *bioRxiv*, 787903.

Ahler, E., Sullivan, W. J., Cass, A., Braas, D., York, A. G., Bensinger, S. J., Graeber, T. G. & Christofk, H. R. 2013. Doxycycline alters metabolism and proliferation of human cell lines. *PLoS One,* 8, e64561.

Allemailem, K. S., Almatroodi, S. A., Almatroudi, A., Alrumaihi, F., Al Abdulmonem, W., Al-Megrin, W. A. I., Aljamaan, A. N., Rahmani, A. H. & Khan, A. A. 2023. Recent Advances in Genome-Editing Technology with CRISPR/Cas9 Variants and Stimuli-Responsive Targeting Approaches within Tumor Cells: A Future Perspective of Cancer Management. *Int J Mol Sci,* 24.

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., Chen, P. J., Wilson, C., Newby, G. A., Raguram, A., et al. 2019. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature,* 576, 149-157.

Banerji, J., Rusconi, S. & Schaffner, W. 1981. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. *Cell,* 27, 299-308.

Bannister, A. J. & Kouzarides, T. 2011. Regulation of chromatin by histone modifications. *Cell Research,* 21, 381-395.

Barroso, I. & McCarthy, M. I. 2019. The Genetic Basis of Metabolic Disease. *Cell,* 177, 146-161.

Bass, J. J., Wilkinson, D. J., Rankin, D., Phillips, B. E., Szewczyk, N. J., Smith, K. & Atherton, P. J. 2017. An overview of technical considerations for Western blotting applications to physiological research. *Scand J Med Sci Sports,* 27, 4-25.

Basu, S., Campbell, H. M., Dittel, B. N. & Ray, A. 2010. Purification of specific cell population by fluorescence activated cell sorting (FACS). *J Vis Exp.*

Brocken, D. J. W., Tark-Dame, M. & Dame, R. T. 2018. dCas9: A Versatile Tool for Epigenome Editing. *Current Issues in Molecular Biology*, 15-32.

Brown, J. C., Harhay, M. O. & Harhay, M. N. 2017. Anthropometrically-predicted visceral adipose tissue and mortality among men and women in the third national health and nutrition examination survey (NHANES III). *Am J Hum Biol,* 29.

Calo, E. & Wysocka, J. 2013. Modification of enhancer chromatin: what, how, and why? *Mol Cell,* 49, 825-37.

Carullo, N. V. N. & Day, J. J. 2019. Genomic Enhancers in Brain Health and Disease. *Genes,* 10, 43.

Chatterjee, S., Bourque, G. & Lufkin, T. 2011. Conserved and non-conserved enhancers direct tissue specific transcription in ancient germ layer specific developmental control genes. *BMC Developmental Biology,* 11, 63.

Chopra, I. & Roberts, M. 2001. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiol Mol Biol Rev,* 65, 232-60 ; second page, table of contents.

Chretien, S., Dubart, A., Beaupain, D., Raich, N., Grandchamp, B., Rosa, J., Goossens, M. & Romeo, P. H. 1988. Alternative transcription and splicing of the human porphobilinogen deaminase gene result either in tissue-specific or in housekeeping expression. *Proc Natl Acad Sci U S A,* 85, 6-10.

Claussnitzer, M., Cho, J. H., Collins, R., Cox, N. J., Dermitzakis, E. T., Hurles, M. E., Kathiresan, S., Kenny, E. E., Lindgren, C. M., MacArthur, D. G., et al. 2020. A brief history of human disease genetics. *Nature,* 577, 179-189.

Claussnitzer, M., Dankel, S. N., Kim, K. H., Quon, G., Meuleman, W., Haugen, C., Glunk, V., Sousa, I. S., Beaudry, J. L., Puviindran, V., et al. 2015. FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *N Engl J Med,* 373**,** 895-907.

Das, A. T., Tenenbaum, L. & Berkhout, B. 2016. Tet-On Systems For Doxycycline-inducible Gene Expression. *Current Gene Therapy,* 16**,** 156-167.

Davidson, S., Lear, M., Shanley, L., Hing, B., Baizan-Edge, A., Herwig, A., Quinn, J. P., Breen, G., McGuffin, P., Starkey, A., et al. 2011. Differential activity by polymorphic variants of a remote enhancer that supports galanin expression in the hypothalamus and amygdala: implications for obesity, depression and alcoholism. *Neuropsychopharmacology,* 36**,** 2211-21.

Daya, S. & Berns, K. I. 2008. Gene therapy using adeno-associated virus vectors. *Clin Microbiol Rev,* 21**,** 583-93.

De Boeck, J. & Verfaillie, C. 2021. Doxycycline inducible overexpression systems: how to induce your gene of interest without inducing misinterpretations. *Mol Biol Cell,* 32**,** 1517-1522.

Dey, K. K., Gazal, S., van de Geijn, B., Kim, S. S., Nasser, J., Engreitz, J. M. & Price, A. L. 2022. SNP-to-gene linking strategies reveal contributions of enhancer-related and candidate master-regulator genes to autoimmune disease. *Cell Genom,* 2.

Duan, L., Ouyang, K., Xu, X., Xu, L., Wen, C., Zhou, X., Qin, Z., Xu, Z., Sun, W. & Liang, Y. 2021. Nanoparticle Delivery of CRISPR/Cas9 for Genome Editing. *Front Genet,* 12**,** 673286.

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., Epstein, C. B., Frietze, S., Harrow, J., Kaul, R., et al. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature,* 489**,** 57-74.

Elks, C. E., den Hoed, M., Zhao, J. H., Sharp, S. J., Wareham, N. J., Loos, R. J. & Ong, K. K. 2012. Variability in the heritability of body mass index: a systematic review and meta-regression. *Front Endocrinol (Lausanne),* 3**,** 29.

Ezzati, M., Bentham, J., Di Cesare, M., Bilano, V., Bixby, H., Zhou, B., Stevens, G., Riley, L., Taddei, C., Hajifathalian, K., et al. 2017. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults.

Fife, R. S., Sledge, G. W., Roth, B. J. & Proctor, C. 1998. Effects of doxycycline on human prostate cancer cells in vitro. *Cancer Letters,* 127**,** 37-41.

Fox, C. S., Massaro, J. M., Hoffmann, U., Pou, K. M., Maurovich-Horvat, P., Liu, C. Y., Vasan, R. S., Murabito, J. M., Meigs, J. B., Cupples, L. A., et al. 2007. Abdominal visceral and subcutaneous adipose tissue compartments: association with metabolic risk factors in the Framingham Heart Study. *Circulation,* 116**,** 39-48.

Gaulton, K. J., Preissl, S. & Ren, B. 2023. Interpreting non-coding disease-associated human variants using single-cell epigenomics. *Nat Rev Genet,* 24**,** 516-534.

Gray, J. M., Kim, T. K., West, A. E., Nord, A. S., Markenscoff-Papadimitriou, E. & Lomvardas, S. 2015. Genomic Views of Transcriptional Enhancers: Essential Determinants of Cellular Identity and Activity-Dependent Responses in the CNS. *J Neurosci,* 35**,** 13819-26.

Heintzman, N. D., Hon, G. C., Hawkins, R. D., Kheradpour, P., Stark, A., Harp, L. F., Ye, Z., Lee, L. K., Stuart, R. K., Ching, C. W., et al. 2009. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature,* 459**,** 108-12.

Honnorat-Benabbou, V. C., Lebugle, A. A., Sallek, B. & Duffaut-Lagarrigue, D. 2001. Stability study of tetracyclines with respect to their use in slow release systems. *Journal of Materials Science: Materials in Medicine,* 12**,** 107-110.

Hood, L. & Rowen, L. 2013. The Human Genome Project: big science transforms biology and medicine. *Genome Medicine,* 5**,** 79.

Højfeldt, J. W., Agger, K. & Helin, K. 2013. Histone lysine demethylases as targets for anticancer therapy. *Nature Reviews Drug Discovery,* 12**,** 917-930.

Iuchi, K., Oya, K., Hosoya, K., Sasaki, K., Sakurada, Y., Nakano, T. & Hisatomi, H. 2020. Different morphologies of human embryonic kidney 293T cells in various types of culture dishes. *Cytotechnology,* 72**,** 131-140.

Janik, E., Niemcewicz, M., Ceremuga, M., Krzowski, L., Saluk-Bijak, J. & Bijak, M. 2020. Various Aspects of a Gene Editing System—CRISPR–Cas9. *International Journal of Molecular Sciences,* 21**,** 9604.

Kadonaga, J. T. 2012. Perspectives on the RNA polymerase II core promoter. *Wiley Interdiscip Rev Dev Biol,* 1**,** 40-51.

Karlsson, T., Rask-Andersen, M., Pan, G., Höglund, J., Wadelius, C., Ek, W. E. & Johansson, Å. 2019. Contribution of genetics to visceral adiposity and its relation to cardiovascular and metabolic disease. *Nat Med,* 25**,** 1390-1395.

Krill, A. S. S. 2023. *Advanced CRISPR-Cas9 techniques for modulation of non-coding disease-associated genetic variants.* Master of Science, The University of Bergen.

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M. J., et al. 2015. Integrative analysis of 111 reference human epigenomes. *Nature,* 518**,** 317-330.

Kwok, K. H. M., Lam, K. S. L. & Xu, A. 2016. Heterogeneity of white adipose tissue: molecular basis and clinical implications. *Experimental & Molecular Medicine,* 48**,** e215-e215.

Lewis, M. W., Li, S. & Franco, H. L. 2019. Transcriptional control by enhancers and enhancer RNAs. *Transcription,* 10**,** 171-186.

Li, K., Liu, Y., Cao, H., Zhang, Y., Gu, Z., Liu, X., Yu, A., Kaphle, P., Dickerson, K. E., Ni, M., et al. 2020. Interrogation of enhancer function by enhancer-targeting CRISPR epigenetic editing. *Nature Communications,* 11**,** 485.

Li, W., Notani, D. & Rosenfeld, M. G. 2016. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nature Reviews Genetics,* 17**,** 207-223.

Liu, D., Wu, H., Cui, S. & Zhao, Q. 2023. Comprehensive Optimization of Western Blotting. *Gels,* 9.

Loew, R., Heinz, N., Hampf, M., Bujard, H. & Gossen, M. 2010. Improved Tet-responsive promoters with minimized background expression. *BMC Biotechnology,* 10**,** 81.

Loos, R. J. F. & Yeo, G. S. H. 2022. The genetics of obesity: from discovery to biology. *Nat Rev Genet,* 23**,** 120-133.

Maeder, M. L., Linder, S. J., Cascio, V. M., Fu, Y., Ho, Q. H. & Joung, J. K. 2013. CRISPR RNA-guided activation of endogenous human genes. *Nat Methods,* 10**,** 977-9.

Maes, H. H. M., Neale, M. C. & Eaves, L. J. 1997. Genetic and Environmental Factors in Relative Body Weight and Human Adiposity. *Behavior Genetics,* 27**,** 325-351.

Maiques-Diaz, A. & Somervaille, T. C. 2016. LSD1: biologic roles and therapeutic targeting. *Epigenomics,* 8**,** 1103-16.

Malach, P., Kay, C., Tinworth, C., Patel, F., Joosse, B., Wade, J., Rosa do Carmo, M., Donovan, B., Brugman, M., Montiel-Equihua, C., et al. 2023. Identification of a small molecule for enhancing lentiviral transduction of T cells. *Mol Ther Methods Clin Dev,* 31**,** 101113.

Mansisidor, A. R. & Risca, V. I. 2022. Chromatin accessibility: methods, mechanisms, and biological insights. *Nucleus,* 13**,** 236-276.

Martin, M. J., Thottathil, S. E. & Newman, T. B. 2015. Antibiotics Overuse in Animal Agriculture: A Call to Action for Health Care Providers. *Am J Public Health,* 105**,** 2409-10.

Martinez-Gamero, C., Malla, S. & Aguilo, F. 2021. LSD1: Expanding Functions in Stem Cells and Differentiation. *Cells,* 10**,** 3252.

Mengstie, M. A. & Wondimu, B. Z. 2021. Mechanism and Applications of CRISPR/Cas-9-Mediated Genome Editing. *Biologics: Targets and Therapy,* Volume 15**,** 353-361.

Miller, J. L. & Grant, P. A. 2013. The role of DNA methylation and histone modifications in transcriptional regulation in humans. *Subcell Biochem,* 61**,** 289-317.

Mirza, S. S. 2022. *Genotype dependent gene regulation in visceral adiposity: a study of the 11q23.3 locus.* Master of Science, The University of Bergen.

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology (Reading),* 155**,** 733-740.

Motoche-Monar, C., Ordoñez, J. E., Chang, O. & Gonzales-Zubiate, F. A. 2023. gRNA Design: How Its Evolution Impacted on CRISPR/Cas9 Systems Refinement. *Biomolecules,* 13.

Moullan, N., Mouchiroud, L., Wang, X., Ryu, D., Williams, E. G., Mottis, A., Jovaisaite, V., Frochaux, M. V., Quiros, P. M., Deplancke, B., et al. 2015. Tetracyclines Disturb Mitochondrial Function across Eukaryotic Models: A Call for Caution in Biomedical Research. *Cell Rep,* 10**,** 1681-1691.

Nagasaka, M., Tsuzuki, K., Ozeki, Y., Tokugawa, M., Ohoka, N., Inoue, Y. & Hayashi, H. 2019. Lysine-Specific Demethylase 1 (LSD1/KDM1A) Is a Novel Target Gene of c-Myc. *Biol Pharm Bull,* 42**,** 481-488.

Nguyen, T. A., Jones, R. D., Snavely, A. R., Pfenning, A. R., Kirchner, R., Hemberg, M. & Gray, J. M. 2016. High-throughput functional comparison of promoter and enhancer activities. *Genome Res,* 26**,** 1023-33.

Ong, C.-T. & Corces, V. G. 2011. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nature Reviews Genetics,* 12**,** 283-293.

Panigrahi, A. & O'Malley, B. W. 2021. Mechanisms of enhancer action: the known and the unknown. *Genome Biology,* 22**,** 108.

Peña-Martínez, E. G. & Rodríguez-Martínez, J. A. 2024. Decoding Non-coding Variants: Recent Approaches to Studying Their Role in Gene Regulation and Human Diseases. *Front Biosci (Schol Ed),* 16**,** 4.

Pinter, S., Knodel, F., Choudalakis, M., Schnee, P., Kroll, C., Fuchs, M., Broehm, A., Weirich, S., Roth, M., Eisler, S. A., et al. 2021. A functional LSD1 coregulator screen reveals a novel transcriptional regulatory cascade connecting R-loop homeostasis with epigenetic regulation. *Nucleic Acids Res,* 49**,** 4350-4370.

Prasanna, G. L. & Panda, T. 1997. Electroporation: basic principles, practical considerations and applications in molecular biology. *Bioprocess Engineering,* 16**,** 261-264.

Qu, L., Yin, T., Zhao, Y., Lv, W., Liu, Z., Chen, C., Liu, K., Shan, S., Zhou, R., Li, X., et al. 2023. Histone demethylases in the regulation of immunity and inflammation. *Cell Death Discovery,* 9**,** 188.

Ragvin, A., Moro, E., Fredman, D., Navratilova, P., Drivenes, Ø., Engström, P. G., Alonso, M. E., de la Calle Mustienes, E., Gómez Skarmeta, J. L., Tavares, M. J., et al. 2010. Long-range gene regulation links genomic type 2 diabetes and obesity risk regions to HHEX, SOX4, and IRX3. *Proc Natl Acad Sci U S A,* 107**,** 775-80.

Rainha, J., Rodrigues, J. L. & Rodrigues, L. R. 2020. CRISPR-Cas9: A Powerful Tool to Efficiently Engineer Saccharomyces cerevisiae. *Life (Basel),* 11.

Rasouli, M. & Safari, F. 2024. Principles of Indirect Co-culture Method Using Transwell. *Methods Mol Biol.*

Redman, M., King, A., Watson, C. & King, D. 2016. What is CRISPR/Cas9? *Arch Dis Child Educ Pract Ed,* 101**,** 213-5.

Samuelsen, N. T. 2021. *Functional dissection of a genetic locus for visceral fat mass (11q23.3).* Master of Science, The University of Bergen   .

Sanchez, G., Linde, S. C. & Coolon, J. D. 2020. Genome-wide effect of tetracycline, doxycycline and 4-epidoxycycline on gene expression in Saccharomyces cerevisiae. *Yeast,* 37**,** 389-396.

Schoenfelder, S. & Fraser, P. 2019. Long-range enhancer-promoter contacts in gene expression control. *Nat Rev Genet,* 20**,** 437-455.

Shuster, A., Patlas, M., Pinthus, J. H. & Mourtzakis, M. 2012. The clinical importance of visceral adiposity: a critical review of methods for visceral adipose tissue analysis. *Br J Radiol,* 85**,** 1-10.

Sigma-Aldrich. n.d. *Successful Transduction Using Lentivirus* [Online]. Available: https://www.sigmaaldrich.com/NO/en/technical-documents/technical-article/genomics/advanced-gene-editing/successful-transduction-lentivirus [Accessed May 2024].

Spitz, F. & Furlong, E. E. M. 2012. Transcription factors: from enhancer binding to developmental control. *Nature Reviews Genetics,* 13**,** 613-626.

TakaraBio. n.d. *RetroNectin* [Online]. Available: https://www.takarabio.com/products/gene-function/t-cell-transduction-and-culture/retronectin-elisa-kit [Accessed desember 2023].

The UniProt Consortium 2022. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research,* 51**,** D523-D531.

ThermoFisher. n.d. *Fetal Bovine Serum, Tet system approved, US origin* [Online]. Available: https://www.thermofisher.com/order/catalog/product/A4736201 [Accessed May 2024].

Tsai, H.-C., Pietrobon, V., Peng, M., Wang, S., Zhao, L., Marincola, F. M. & Cai, Q. 2022. Current strategies employed in the manipulation of gene expression for clinical purposes. *Journal of Translational Medicine,* 20**,** 535.

Underwood, P. A. & Bean, P. A. 1988. Hazards of the limiting-dilution method of cloning hybridomas. *J Immunol Methods,* 107**,** 119-28.

Vega, G. L., Adams-Huet, B., Peshock, R., Willett, D., Shah, B. & Grundy, S. M. 2006. Influence of body fat content and distribution on variation in metabolic risk. *J Clin Endocrinol Metab,* 91**,** 4459-66.

Vignali, M., Hassan, A. H., Neely, K. E. & Workman, J. L. 2000. ATP-dependent chromatin-remodeling complexes. *Mol Cell Biol,* 20**,** 1899-910.

Voisin, S., Almén, M. S., Zheleznyakova, G. Y., Lundberg, L., Zarei, S., Castillo, S., Eriksson, F. E., Nilsson, E. K., Blüher, M., Böttcher, Y., et al. 2015. Many obesity-associated SNPs strongly associate with DNA methylation changes at proximal promoters and enhancers. *Genome Med,* 7**,** 103.

Ye, M., Wilhelm, M., Gentschev, I. & Szalay, A. 2021. A Modified Limiting Dilution Method for Monoclonal Stable Cell Line Selection Using a Real-Time Fluorescence Imaging System: A Practical Workflow and Advanced Applications. *Methods Protoc,* 4.

Ying, Y., Yang, X., Zhao, K., Mao, J., Kuang, Y., Wang, Z., Sun, R. & Fei, J. 2015. The Krüppel-associated box repressor domain induces reversible and irreversible regulation of endogenous mouse genes by mediating different chromatin states. *Nucleic Acids Research,* 43**,** 1549-1561.

Yip, B. H. 2020. Recent Advances in CRISPR/Cas9 Delivery Strategies. *Biomolecules,* 10**,** 839.

Zhou, X., Vink, M., Klaver, B., Berkhout, B. & Das, A. T. 2006. Optimization of the Tet-On system for regulated gene expression through viral evolution. *Gene Ther,* 13**,** 1382-90.