Single nucleotide editing at the 11q23.3 adiposity risk locus by the prime editing technique

Ingrid Johansen Petersen

This thesis is submitted in partial fulfillment 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

Acknowledgements

The work presented in this thesis was conducted from August 2023 to June 2024 at the Hormone Laboratory, Department of Clinical Science at Haukeland University Hospital, through the Department of Biological Sciences at the University of Bergen.

First, I would like to thank all of my supervisors. A special thank you to my main supervisor, Jan-Inge Bjune, for always being available for questions, whether in person or online, day and night. Your patience in explaining difficult concepts and guiding me through the steep learning curve of a Master's degree has been invaluable. I would also like to thank my two co-supervisors, Simon Erling Nitter Dankel and Annika Sem Sippel Krill. Thank you to Simon for giving me this incredible opportunity, for always engaging in discussion, and for staying up on late nights to give me feedback. Thank you to Annika, for inspiring me to work harder and think deeper about all decisions I make in the lab, big or small. All three of you have been essential to my success.

Second, I would like to thank my partner in crime, Tiffany Chor Kiu Ngan. Thank you for the countless hours of collaboration, the countless laughs, and the countless cakes. You made the transition from dependent to independent in the lab so much easier, and it has been a joy to learn alongside you.

Third, I would like to thank the numerous incredible people that work at the Hormone Laboratory. This thesis would not have been possible without all the people I have gotten to know here. Through the course of the last year, I have learned that it could not be more true that research relies on collaboration. Thank you to Linn Skartveit and Margit Hildershavn Solsvik for always taking the time to answer my questions, and for pointing me in the right direction when I got lost on my way to find more pipette tips. Thank you to Pouda Panahandeh and Sayani Mukherjee for going out of your way to make sure I was prepared when I was unsure about what to do next, and for always providing interesting insight. Thank all of you for endless hours in the lab, endless questions, and endless discussions over lunch.

Lastly, I would like to thank my family and friends. More specifically, I would like to thank my partner Trevor for always trying your best to understand and for extending your love through all the long and hard days. Thank you to the rest of my family for being patient when I was busy, and for providing me with a safe space and warm dinners. Thank you to all the old friends I have known for ages and the new ones I have gained over the last year. Specifically, I would like to thank Hannah, Thea, and Katrin. Thank you for providing me with distraction when I needed it, and for cheering me on.

Bergen, June 2024
Ingrid Johansen Petersen

Contents

List of Abbreviations

Abbreviation	Explanation	Abbreviation	Explanation
AAV	Adeno-Associated Virus	SAT	Subcutaneous Adipose Tissue
AD-MSC	Adipose Derived Mesenchymal Stem Cells	SNP	Single Nucleotide Polymorphism
ANOVA	Two-Way Analysis of Variance	SNV	Single Nucleotide Variant
BAT	Brown Adipose Tissue	SSC-A	Side Scatter Area
bp	Base pair	tracrRNA	Trans-activating CRISPR RNA
BSA	Bovine Serum Albumin	TSS	Transcription Start Site
CFU	Colony Forming Unit	VAT	Visceral Adipose Tissue
CNV	Copy Number Variants	WAT	White Adipose Tissue
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats	WB	Western Blot
crRNA	CRISPR-RNA	WES	Whole-Exome Sequencing
Ct	Cycle threshold	WGS	Whole-genome sequencing
DSB	Double-Stranded Break	WT	Wild Type
E. coli	*Escherichia coli*		
eQTL	Expression Quantitative Trait Locus		
f.c.	Final concentration		
FSC-A	Forward Scatter Area		
FSC-H	Forward Scatter Height		
GFP	Green Fluorescent Protein		
GTEx	Genotype-Tissue Expression (GTEx) project		
GV	GoStix™ Value		
GWAS	Genome-Wide Association Studies		
HEK	Human Embryonic Kidney		
HDR	Homology Directed Repair		
IFU	Infectious Unit		
Kb	Kilobases		
kDa	Kilodalton		
LD	Linkage Disequilibrium		
MMR	Mismatch Repair		
MOI	Multiplicity Of Infection		
MSC	Mesenchymal Stem Cell		
nCas9	Cas9 nickase		
NHEJ	Non-Homologous End Joining		
PAM	Protospacer Adjacent Motif		
PE	Prime Editing		
pegRNA	Prime editing guide RNA		
RT	Reverse Transcriptase		
S. pyogenes	*Streptococcus pyogenes*		
sgRNA	Single guide RNA		

Summary

Unraveling the mysteries hidden within the small fraction of the genome that makes every human unique, is still ongoing. Most of these genetic variations require interactions with environmental factors to affect the risk of developing complex diseases, and uncovering their biological roles remains a big challenge. Genome-Wide Association Studies (GWAS) have been instrumental in identifying single nucleotide polymorphisms (SNPs) associated with diseases, although the causal variants often elude detection due to the complexity of the genome. Connecting the biology of SNPs to their statistically linked diseases can be achieved by gene editing technologies such as CRISPR/Cas9 and prime editing (PE).

Accumulation of visceral adipose tissue (VAT) – fat stored around the internal organs – has been linked to insulin resistance, metabolic disease, and cardiovascular disease. While GWASs have identified numerous SNPs associated with overall adiposity (overweight and obesity as measured by BMI), few have focused on the distribution of fat itself. A recently identified locus, 11q23.3, has been found to be associated with VAT mass. Specifically, the A-allele of the SNP rs1799993 within this locus is associated with an increase in VAT.

The current thesis is part of a larger project with the aim of understanding the biological mechanisms of this locus, with a special focus on the SNP rs1799993. Our lab has identified regions with enhancer activity within the locus, and the SNP rs1799993 was found to be located within a potential enhancer. Moreover, epigenetic data suggest this enhancer is active in adipose-derived mesenchymal stem cells. By utilizing prime editing in the mesenchymal stem cell line ASC52telo, the underlying mechanisms of rs1799993 can be uncovered. The overall aims of this thesis are therefore to first establish a proof-of-concept prime editing protocol in the easily transfected cell line HT1080, followed by developing a lentiviral transduction protocol for the PE components in the mesenchymal stem cell line ASC52telo with the prospect of using PE to generate ASC52telo clones homozygous for the risk and protective alleles of rs1799993.

Sanger sequencing was applied to genotype the HT1080 cell line, and it was found to be heterozygous for the SNP rs1799993. This meant that the edit needed to be performed in both directions to achieve homozygous cell lines. First, different transient transfection protocols were tested and assessed using qPCR or fluorescence imaging. The PE components (PEmax-P2A-GFP, pegRNAs, and mismatch repair inhibitor) were transfected using the optimal protocol, and GFP positive cells were sorted as single cells by FACS for single-cell colony expansion. Single-cell expansion yielded 162 colonies, 87 of which were genotyped by Sanger sequencing. One colony (~1%) was positively edited from heterozygous A/C to the homozygous C protective allele, suggesting that the PE protocol worked, at least with the designed pegRNAs encoding the C-allele.

Based on the successful edit in the HT1080 cells, PE of the ASC52telo proceeded. Lentivirus carrying the PEmax sequence and blasticidin resistance were created in the HEK293T cell line, and subsequent quantification of viral particles and infectious units were determined by a rapid antigen detection test and estimated by a functional viral titer test, respectively. ASC52telo cells were transduced with a multiplicity of infection of 2 and 5, and successfully transduced cells were selected using blasticidin. The lentiviral transduction of ASC52telo and selection of transduced cells were successful, but no single-cell colonies were expanded despite comprehensive effort. Therefore, several heterogeneous cell lines with stable expression of PEmax were established instead. Genomic integration and mRNA expression of PEmax was demonstrated by qPCR but could be validated by western blotting.

Finally, a pilot was conducted to introduce the remaining PE-components into the heterogenous cell lines expressing PEmax. The cells were sorted by FACS after nucleofection, but none of the cells survived. Therefore, more work should be done to further develop clonal expansion of this cell line, and to optimize nucleofection for improved viability. As a result, the SNP rs1799993 was not edited in the ASC52telo cell line. However, the groundwork has been laid for further the development of a PE protocol and clonal expansion, as well as utilizing the PEmax-expressing cells to edit rs1799993 and other genomic regions by differently designed pegRNAs. Once these protocols have been developed, any disease-associated SNP in these cells can in principle be edited, their downstream target genes be identified, and the affected biology be determined.

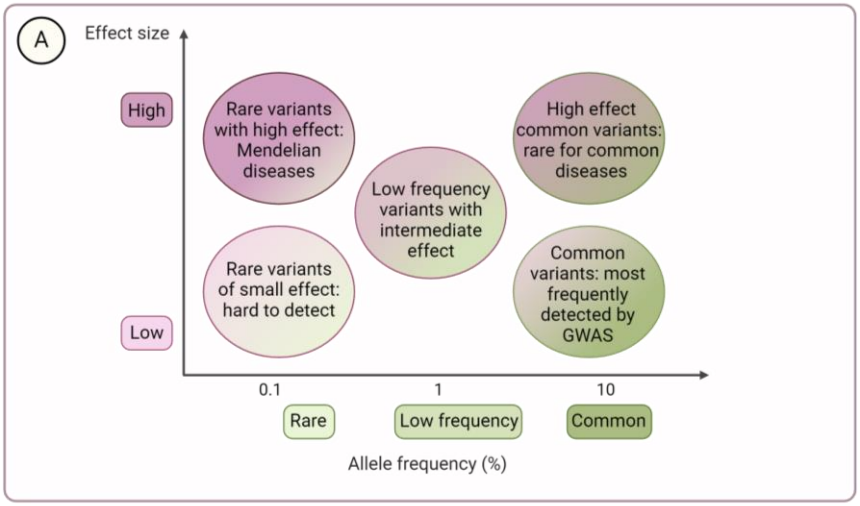
1. Introduction

1.1. Decoding the genome to uncover the causes of complex disease

Every human is unique, yet if any two genomes were subjected to comparison, they would be about 99.6% identical. The remaining 0.4% includes millions of variants among the 6 billion base pairs that constitute the human genome (NIH, 2023a). While most of the affected base pairs are found in larger structural rearrangements such as deletions, duplications, inversions, etc., the most common type of variation is the single nucleotide variation (SNV). In any given individual, about 5 million SNVs can be found (NIH, 2023a). SNVs that appear more frequently in the population than 1% are referred to as single nucleotide polymorphisms (SNPs). After the Human Genome Project was completed, and further technological developments in the early 21st century allowed for cheaper sequencing (Hood and Rowen, 2013, Auton et al., 2015), researchers were, for the first time in history, able to map associations between SNPs and different diseases and traits, in so-called genome-wide association studies (GWASs). To date, GWASs have identified more than 400 000 SNPs associated with more than 5000 diseases and traits (Sollis et al., 2022).

SNPs often occur in clusters in the genome, but can, in theory, occur at any nucleotide. Uncovering the specific causal SNPs that affect bodily functions, and how they do so, has however proven challenging, in large part because the vast majority of SNPs occur in non-coding regions (McVean et al., 2012, Barroso and McCarthy, 2019), as opposed to in coding regions, of the DNA (Freedman et al., 2011, Hood and Rowen, 2013). While the effect of SNPs in protein-coding sequences of DNA can be readily determined, the role of SNPs in non-coding DNA is far less clear. It is believed that trait-affecting non-coding SNPs will often reside in regulatory regions, affecting gene expression levels rather than gene/protein structure. However, enhancer regions are not completely mapped in all cell types, and the target gene(s) can be far away from the SNP, sometimes up to megabases away (Claussnitzer et al., 2015).

Coding SNPs that affect the gene product typically cause monogenic diseases, where a single mutation is sufficient to affect the phenotype (Bick et al., 2021). In these diseases, there is therefore a clear path between genotype and phenotype. However, monogenic disease-causing SNPs are rare, with a minor allele frequency of $< 0.5\%$ in the general population (Figure 1.1 A). Most SNPs do not cause disease by themselves but contribute to disease risk in combination with dozens to hundreds of other SNPs, each making a minor contribution. However, these SNPs are more common in the population and, along with environmental factors, contribute to a range of lifestyle related diseases such as obesity, heart disease and type-2 diabetes (Figure 1.1 B) (Barroso and McCarthy, 2019, NIH, 2023b). These diseases are termed “polygenic” or “complex” because they usually do not present a disease-associated phenotype until a given combination of variants and environmental factors causes it to penetrate (Lvovs et al., 2012).



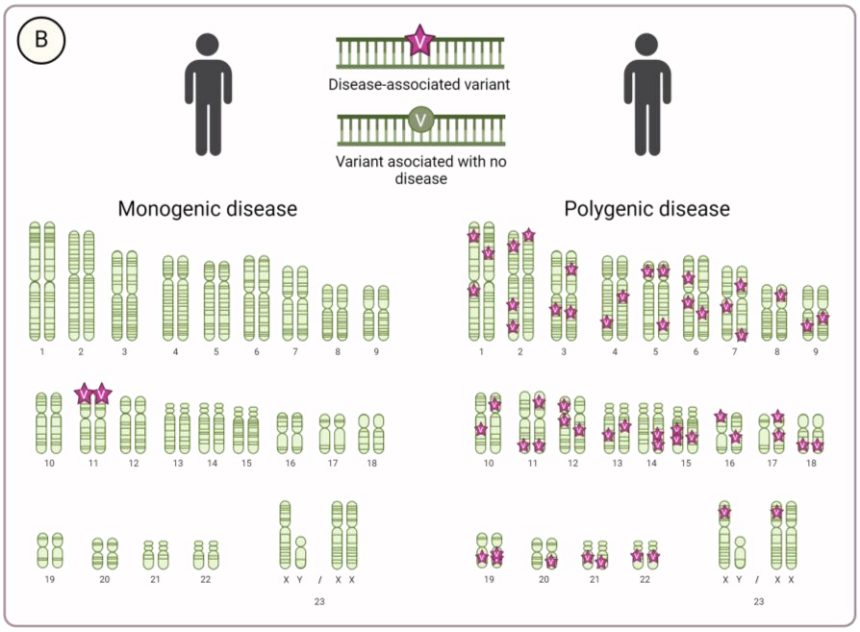


Figure 1.1: Effect sizes of variants discovered by GWASs and the genetic makeup of monogenic and polygenic disease. (A) The y-axis denotes the penetrance of the allele (penetrance of the phenotype), while the x-axis denotes the allele frequency in the population. Genome wide association studies (GWASs) are designed to detect genetic variants associated with disease. Most variants detected by a GWAS tend to have relatively low effects. The variants picked up by a GWAS are also the most common in the population, meaning that rare variants are usually missed. Rare variants with high effects are usually the cause of diseases with typical Mendelian inheritance patterns, which have a minor allele frequency of < 0.5%. Figure created in BioRender.com adapted from McCarthy et al. (2008). **(B)** Disease-associated variants, single nucleotide polymorphisms included, are denoted with a star. In monogenic diseases, as highlighted on the left in the illustration, the disease is caused by one variant in the genome. In the case of classic dominant inheritance, only one allele is needed for the disease to be inherited. In the case of recessive diseases, such as in the illustration, one copy on each chromosome is necessary. These diseases usually also follow Mendelian inheritance patterns, meaning that diseases can often be identified by mapping family history. Characteristics like these make monogenic diseases generally easier to characterize than polygenic diseases. For a phenotype to penetrate in the case of complex or polygenic diseases (right side of the figure), the combination of variants and environment play a central role. The variants also show different effect sizes, and the effects can vary in combination with lifestyle and other variants. This means that even if one possesses the 'correct' combination of variants, they might not cause a change of phenotype until the right environment is provided. Figure created with BioRender.com adapted from the National Human Genome Research Institute (NIH, 2023b).

While GWASs can provide clues about which SNPs may be involved in disease mechanisms, the presence of these SNPs are not always good predictors of future disease on their own. Still, identifying the molecular mechanisms behind disease-associated SNPs found in GWASs could provide important new insight into SNP-environment interactions in disease pathogenesis, potentially leading to new therapeutic targets for personalized prevention and treatment. However, the challenge remains: how can the genetic contributions of polygenic diseases be effectively mapped when environmental factors play such a big role, and how can the underlying molecular mechanisms be uncovered?

1.1.1. From sequencing a single genome to detecting patterns across many

The Human Genome Project and increased accessibility of effective sequencing techniques were a springboard for several big projects aiming to gain a better understanding of human genetic variation. One of these projects was the HapMap project (2002-2005), which aimed to determine common patterns of DNA sequence variation and make it publicly available, with the goal of providing a database for studying common, complex disorders (Gibbs et al., 2003, Altshuler et al., 2005). The project began when it was discovered that patterns of recombination in the genome could be used to our advantage when studying inheritance and disease. Certain areas on chromosomes appear to have a higher frequency of recombination (recombination hotspots), compared to other areas where recombination seems to be effectively absent (Gibbs et al., 2003, Wall and Pritchard, 2003, Altshuler et al., 2005). Thus, genetic variants that occur in the areas of no recombination are statistically linked, a concept known as linkage disequilibrium (LD) (Wall and Pritchard, 2003). Patterns of LD in the human genome can be described by LD-blocks, whereby neighboring blocks are separated by recombination hotspots (Daly et al., 2001, Gabriel et al., 2002, Wall and Pritchard, 2003). The collective variants residing in each block denote a haplotype. The main haplotype of each block can be labeled and characterized by specific tag SNPs, which provides an effective mechanism for screening haploblocks in large-scale association studies (Gabriel et al., 2002, Gibbs et al., 2003) (Figure 1.2). These observations provided the foundation for creating a haplotype map of the human genome, such as the HapMap project.

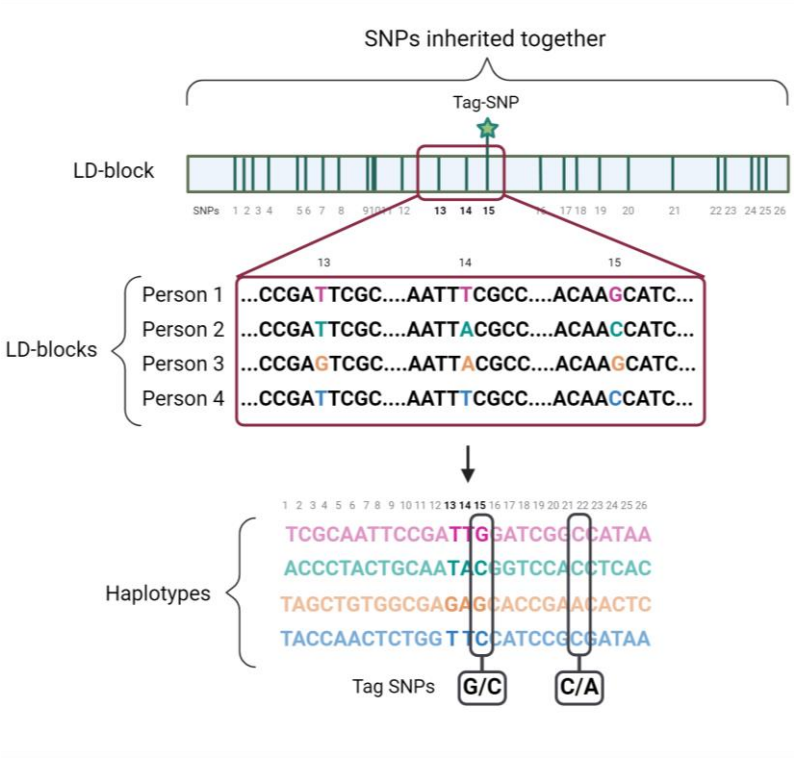


Figure 1.2: The role of tag-SNPs in LD-blocks and haplotypes. LD-blocks are characterized by several SNPs in linkage-disequilibrium. This means that throughout evolution, specific parts of the genome were inherited together more often than not. This distinct feature of inheritance can be used to our advantage when developing methods to analyze the human genome. In the case of the HapMap project, the SNPs within an LD-block are put together to form distinct haplotypes, which works as descriptors of variance. Certain SNPs can be chosen to represent these blocks, termed tag-SNPs. When performing large-scale association studies, the tag-SNPs alone can be analyzed, negating the need for investigating the entire genome. Figure created with BioRender.com adapted from the International HapMap Consortium (Gibbs et al., 2003).

After the HapMap project was completed, the 1000 Genomes project set out to provide a thorough description of genetic variation in humans by applying whole genome sequencing to multiple individuals from multiple populations (McVean et al., 2012). In 2015 they reported having completed sequencing of 2504 individuals from 26 populations, resulting in the identification of 84.7 million SNPs, which encompasses > 99% of all SNP variants with a population frequency of > 1% (Auton et al., 2015).

Building on these findings, the challenge has been to move past the need to focus on only a small fraction of the genome at a time, and instead achieve a multidimensional approach able to match genome-wide variants with functional analyses. Decoding and cataloguing functional activity have therefore been an important steppingstone, with the ENCODE and Roadmap Epigenomics projects being at the forefront. These projects have generated maps of chromatin accessibility, three-dimensional structures in the genome, histone modifications, and other regulatory annotations across hundreds of tissues and cell types (ENCODE, 2012, Kundaje et al., 2015). Results from these large, collaborative, and publicly available projects make the possibility of decoding the genome, and thus finding the underlying causes of complex disease, closer than ever before.

1.1.2. Genome-Wide Association Studies

The premise of a GWAS is to test numerous genetic variations across many genomes to statistically link variants (most commonly SNPs) with specific diseases or traits. The data produced from the HapMap project,

the 1000 Genomes project, and other projects like them, supported the conduction of the first GWASs. These were performed on a much smaller scale than today, but some of them still managed to uncover strong associations across many diseases. However, the vast majority of variants uncovered had more subtle effects than expected (Claussnitzer et al., 2020). The increased access to sequencing techniques we have now makes it possible to build larger databases across populations. Larger data sets increase the chances to discover rare and low-frequency SNPs compared to previous GWASs, and thus getting closer to identifying more causal SNPs (Lange et al., 2014, Walter et al., 2015, Collins et al., 2021).

Causal variants have been difficult to identify due to the high correlation between adjacent variants in LD. While the extensive disequilibrium in the genome is what makes GWASs possible through LD-blocks and tag-SNPs, it also directly causes a lot of the complexity in interpreting GWAS results. SNPs that occur in strong LD with actual causal variants therefore tend to present similar statistical significance. This, combined with SNPs occurring mostly in non-coding regions (assumed regulatory regions), makes linking the causal variants to a specific gene even harder (Figure 1.3). The obesity-associated *FTO* locus is a good example of this, as the variants within the locus are located in the introns of the *FTO* gene, but despite being named after the *FTO* gene, the haploblock has been shown to be positioned within an enhancer that regulates expression *IRX3* and *IRX5* up to 1.5 million bases away (Smemo et al., 2014, Claussnitzer et al., 2015, Bjune et al., 2023) Nevertheless, GWASs serve as a valuable tool for guiding researchers toward specific disease-associated genomic loci, warranting further exploration to understand SNPs in a phenotypic, genomic, and biological context.

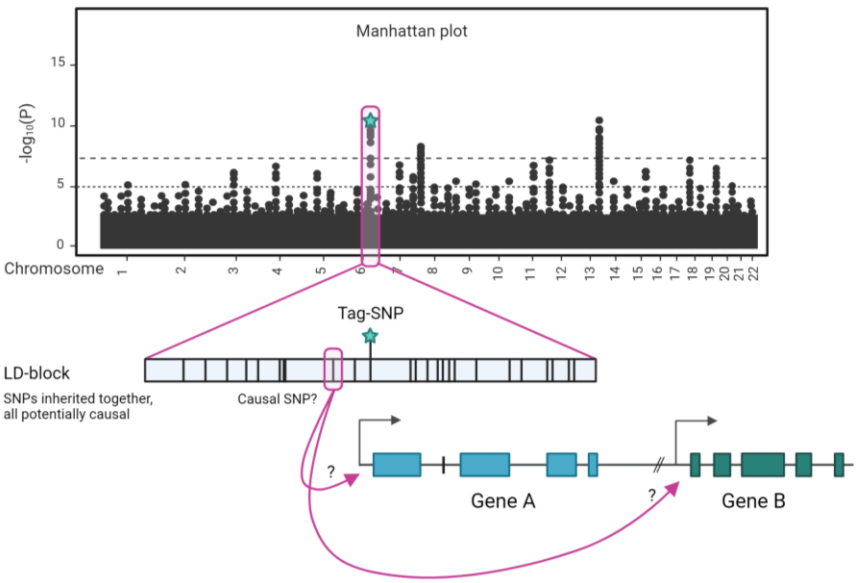


Figure 1.3: The relationship between SNPs and their connected genes. Results from genome wide association studies (GWASs) are often presented in the form of Manhattan plots, where the statistical significance of each variant's association to the phenotype measured as -$\log_{10}$ transformed P-values (y-axis) plotted against their position in the genome (x-axis). Each dot represents a variant in the form of a single-nucleotide polymorphism (SNP). Tag-SNPs are used to represent a region of the genome with high linkage disequilibrium (LD), termed LD-blocks. The combination of SNPs within an LD-block represents a specific haplotype. When differences in LD patterns occur between populations, SNPs can lose association with each other and cause changes in the haplotype. This is important to keep in mind when choosing tag-SNPs for GWASs or other association studies. The use of tag-SNPs negates the need to genotype every SNP in the genome and

reduces both cost and time. As tag-SNPs represent several other SNPs as well, there is a possibility of synthetic associations. This means that as a phenotype is found to be associated with a haplotype, some SNPs will be false positives. Another important consideration is that the causal SNP does not necessarily lie within a coding region, or even link molecularly to the closest gene. Results from a GWAS will not provide any information on whether a specific SNP is linked to the closest gene (illustrated as 'gene A') or to one that lies further away ('gene B'). GWAS results will also not provide any information about which cell type(s) or tissue(s) the SNP is linked to. Performing further analysis and experiments to confirm which SNPs are truly causal and the mechanisms behind them is therefore important. Figure created with BioRender.com adapted from Jan-Inge Bjune, unpublished.

1.1.3. GWAS signals in regulatory regions

As mentioned above, most disease-associated GWAS signals occur outside of protein-coding genes and are presumed to lie within regulatory regions (Freedman et al., 2011, Welter et al., 2014). A regulatory region is a broad term for regulatory elements that are involved in controlling gene expression through interaction with transcription factors (TFs). Regulatory elements include promoters, enhancers, silencers, and insulators. The interactions between regulatory regions, TFs, cofactors, and specific epigenetic variations such as histone modifications, DNA methylation, and DNA looping establish complex molecular networks that give cells their specific identities. Promoters lie within 1-2 kilobases (kb) of a gene's transcription start site and contain short motifs that aid to assemble the transcription initiation complex (Doane and Elemento, 2017). Enhancers are regions located distally to promoters and are often needed to achieve more than minimal transcription. Enhancers interact with site-specific TFs that promote or inhibit transcription (Shlyueva et al., 2014). As opposed to protein-coding sequences, the sequence codes of enhancers are poorly understood. Enhancers and TF binding sites can therefore not be inferred from DNA sequence alone.

Enhancer activity is also often restricted to specific cell types and tissues, to certain developmental time points, or in response to various environmental conditions (Pennacchio et al., 2013). This is controlled by the chromatin status and epigenetic markers of the histone tails in the different cell types. Open chromatin is associated with active transcription, which is initiated by acetylation of the histone tails. A high density of methylation, however, is associated with closed chromatin and inactive transcription (Jaenisch and Bird, 2003). When an enhancer is active, activator proteins promote transcription by binding to distal control elements in the enhancer (Figure 1.4, step 1). This is thought to facilitate DNA looping by recruitment of proteins that cause the DNA to bend (Figure 1.4, step 2). This brings the activator proteins, along with the enhancer, closer to the promoter. The activator proteins bound to the enhancer, together with RNA polymerase and other co-factors, then form the transcription initiation complex (Figure 1.4, step 3) (Shlyueva et al., 2014, Doane and Elemento, 2017).

When SNPs are located within enhancers, they may increase or decrease enhancer function, either by strengthening protein binding affinity or weakening it. As mentioned above, the *FTO* locus contains SNPs that are shown to affect gene regulation of genes further away. For instance, the rs1421085 (T→C) risk variant alters the binding motif of the ARID5B repressor, leading to an upregulation of the distal genes *IRX3* and *IRX5* specifically in preadipocytes, promoting a pro-obesity phenotype (Claussnitzer et al., 2015).

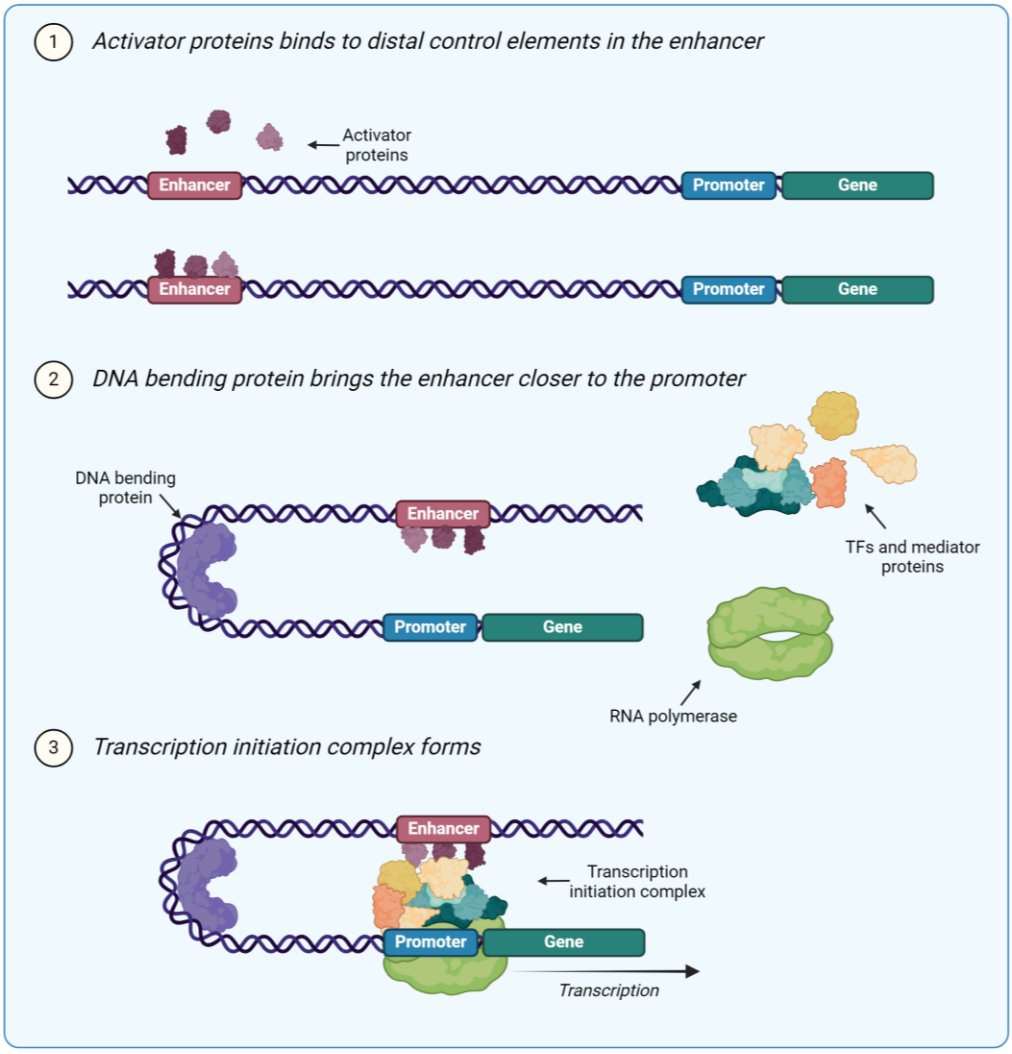


Figure 1.4: Enhancer activation. (1) Enhancers contain sequences that act as binding sites for specific transcription factors (TFs) that facilitate or inhibit transcription. When transcription is initiated, activator proteins bind to the distal regulatory elements within the enhancer. **(2)** These proteins recruit other TFs and co-factors that bend the DNA, causing the enhancer and activator proteins to come into close proximity of the promoter. **(3)** The recruited TFs and other transcription mediator proteins then assemble along with RNA polymerase to form the transcription initiation complex, which initiates transcription of the gene. Figure created with BioRender.com adapted from Shlyueva et al. (2014).

1.1.4. Fine-mapping of likely causal variants

Many variants within a genetic locus may show a statistically significant association with a trait due to LD with the actual causal variant. This creates clusters of associated variants, making it difficult to narrow down the true casual variants. Fine-mapping analysis is a computational method that narrows down which variants are most likely to be causal, by analyzing patterns of LD and statistical associations (Uffelmann et al., 2021). The simplest fine-mapping analysis is conditional association analysis, which adjusts for the effect of the variant with the strongest association within the region of LD. This process is performed iteratively until no more further significant associations are found, also known as stepwise conditional analysis (Uffelmann et al., 2021). More advanced fine-mapping analyses use Bayesian statistical models, which are powerful, but become more complex and the statistical power decreases when multiple associations are present within a region (Wang et al., 2020).

Reporter assays can be conducted as an alternative, or in addition, to statistical fine mapping analyses to narrow down causal variants. Enhancer function, for instance, can be validated through reporter assays, where a candidate enhancer sequence is linked to a minimal promoter followed by a reporter gene (GFP, luciferase, etc.). The reporter vectors are then introduced into cultured cell lines or organisms, and the expression of the reporter gene is measured. If the presence of the sequence preceding the promoter changes the reporter gene expression, it strongly indicates enhancer activity (Inoue and Ahituv, 2015). A region of multiple associations can therefore be divided into smaller tiles, the tiles cloned into several vectors, and the enhancer activity of the tiles can be compared to each other. Each SNP within the tiles can then be mutated to the alternative allele, and the tiles can be reassessed by another reporter assay to look for changes in the reporter gene expression. (Claussnitzer et al., 2015, Samuelsen, 2021, Mirza, 2022). Thus, several associations can be associated with enhancer activity, and narrowed down to more or less likely causal.

1.1.5. Narrowing down relevant cell models and putative target genes

Once the likely causal variant has been narrowed down (Figure 1.5 steps 1-2), the task of uncovering molecular mechanisms commences. In this context, it is imperative to consider the specific cellular environment in which the variant operates. For instance, the preadipocyte-specific effect of rs1421085 in the *FTO* locus was due to its position in an enhancer that was only active in this cell type (Claussnitzer et al., 2015). Similarly, many enhancers are only active in specific cells and tissues, meaning that SNPs that reside in these regions may only have an effect in those specific cells. Thus, mapping the chromatin status across different cell types and tissues surrounding the SNP of interest is critical to narrow down the place of action of the SNP.

Chromatin status in various cell types has been provided by large projects such as ENCODE and the Roadmap Epigenomic projects and is crucial in predicting which cell types disease-associated signals are active in.

ENCODE uses histone modification markers to assign assumed functions to their respective regions, which includes predicting regulatory elements such as enhancers. Their data shows that global patterns of modifications are highly variable across cell types, which changes in accordance with transcriptional activity, with open chromatin being associated with active transcription as opposed to closed chromatin (ENCODE, 2012). Thus, information of different histone modifications and other epigenetic markers can be used in a systematic way to assign functional properties to genomic regions. Additionally, the Genotype-Tissue Expression (GTEx) project characterizes transcriptomes in a variety of tissues and cell types, and identifies associations between SNPs and gene expression, known as expression quantitative trait loci (eQTLs) (Farh et al., 2015, Aguet et al., 2017). While eQTL analyses can indicate loci associated with molecular, cellular, or organismal phenotypes, variants in LD with the causal variant will likely show a statistical association with the trait, and eQTLs therefore fall short of determining the true causal variant. Moreover, eQTLs usually only consider genes close to each SNP and will miss long range regulation.

Chromosome conformation capture (3C) techniques can detect the long-range interactions that eQTLs miss. The technique involves treating cells in a way that causes crosslinking of DNA and observing pairs of interacting loci. Several techniques have been developed that build on each other and offer more involved analyses of interactions compared to the one-vs-one approach of the original 3C technique. These include 4C, as a one-vs-all approach, and T2C, which offers a many-vs-all approach, and can also be adapted to serve as an all-vs-all approach (Dekker et al., 2002, Zhao et al., 2006). eQTL and chromosome conformation capture techniques can therefore be used together to provide comprehensive maps of genomic interactions and chromatin states across different cell types, which point researchers in the right direction when selecting the most fitting cell types in which to study the SNP of interest (Figure 1.5 step 3). Ultimately, to conclusively determine which variants are truly causal, experimental testing through genome editing and subsequently observing the effects on target genes in the relevant cell types is necessary (Figure 1.5 step 4).

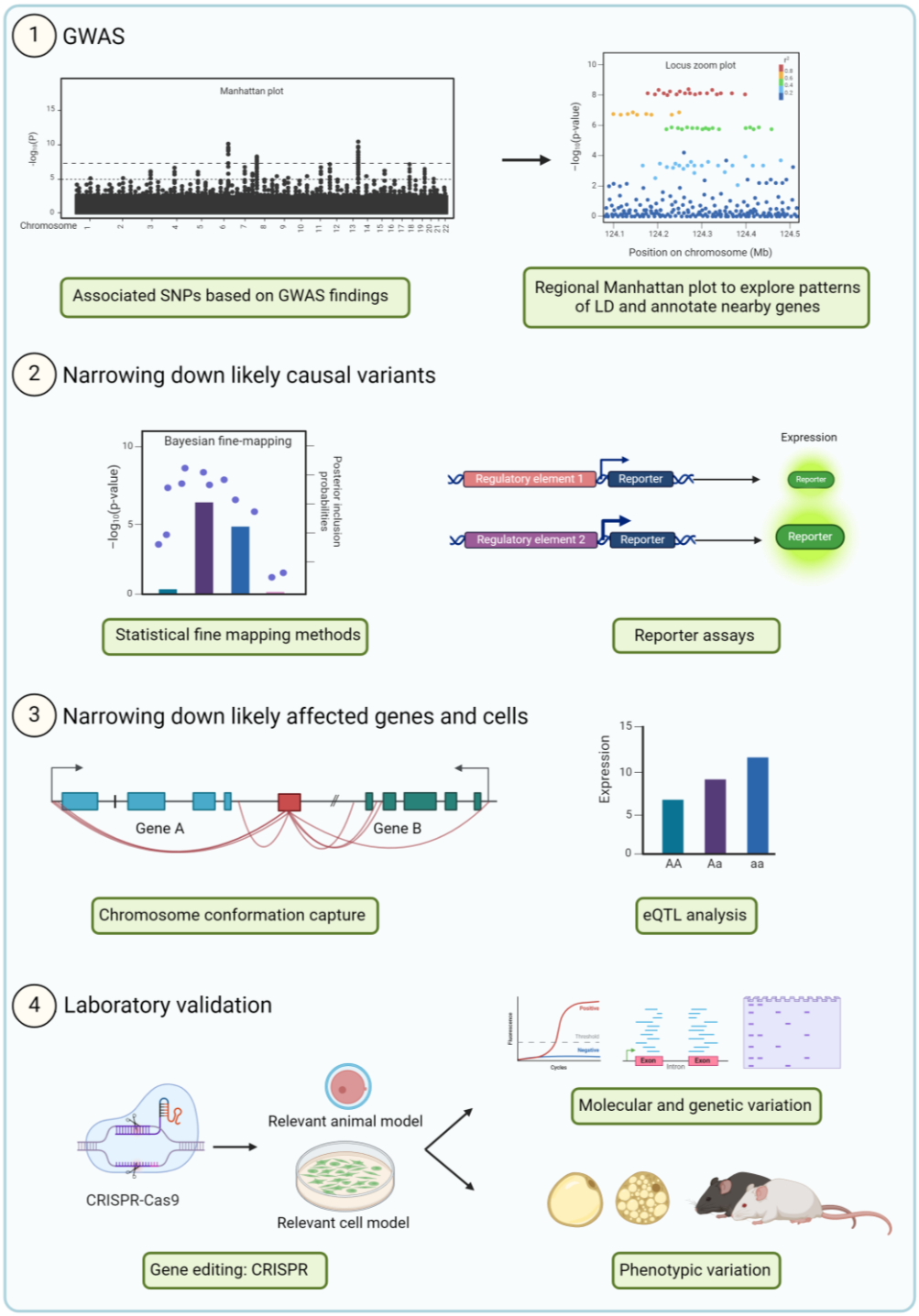


Figure 1.5: Example of typical workflow from initial GWAS to uncovering variant function. **(1)** To understand the links between diseases and genomic variation, Genome Wide Association Studies (GWASs) can be conducted. To achieve this, a GWAS is conducted on a large population before a list of single-nucleotide polymorphisms (SNPs) associated with the disease is created. SNPs are often visualized in Manhattan plots to show their genomic positions and strength of association. Regional Manhattan plots are then created to study the individual values of the associations in the region, to study the patterns of linkage disequilibrium, and annotate nearby genes. **(2)** Because GWASs only create a statistical overview of the variants that may be causal, many methods exist to narrow the list down to more or less likely causal.

Statistical fine-mapping is applied to identify areas that are likely to include the causal variant. Reporter assays (such as Luciferase assays) can be conducted to measure whether an allele is associated with gene expression, for instance to discover potential enhancers. **(3)** Chromosome conformation capture can be utilized to predict target genes and used together with eQTL analysis to narrow down the tissues/cells they are active in. **(4)** While all the previous steps are necessary for narrowing down the search of the causal variant and its place of action, further laboratory validation is necessary to verify the findings and unravel the biological and mechanistic action. This can for example be done using genome editing such as CRISPR in the cell models or tissues of animals that were predicted to be the action site of the SNP. Phenotypic and molecular variation can then be measured and linked to the SNP in question. Figure created in BioRender.com adapted from Rao et al. (2021).

1.2. Strategies to reveal molecular mechanisms by gene editing

Once the associated SNPs are limited to a few candidates and the likely cells in which they act are identified, the next step in interpreting GWAS signals involves testing these associations for causality through experimental validation. One way to gain more insight into the underlying biology is through genome editing of the candidate causal SNP(s) in the relevant cell model to identify affected gene(s) and biological pathways. However, editing a single base pair in a cell model can be challenging. This is where the CRISPR/Cas-system revolutionized the world of gene editing.

1.2.1. Genome editing strategies: CRISPR/Cas

The discovery of clustered regularly interspaced short palindromic repeats (CRISPR) in procaryotes dates back to the 1980s, long before the underlying mechanism was understood (Ishino et al., 1987). These repeats are interspersed by unique spacers, and in 2007 it was demonstrated that these unique sequences derive from phages and other foreign invaders. CRISPR-associated genes (Cas genes) encode naturally occuring endonucleases, which, when coupled with spacer-derived RNA (tracrRNA and crRNA), target and induce double stranded breaks (DSBs) in foreign DNA (Barrangou et al., 2007). Essentially, the CRISPR/Cas-system functions as an adaptive immune-system for bacteria and archaea.

In vitro reconstruction of the CRISPR/Cas system from *Streptococcus pyogenes* (*S. pyogenes*) has enabled its programming to target specific sequences in other species. More specifically, by pairing an engineered version of the tracrRNA and crRNA along with the Cas9 endonuclease allows for precise DNA targeting (Figure 1.6). This is possible by constructing the crRNA to complement a target DNA region and linking it to the tracrRNA. This creates a single guide RNA (sgRNA) (Jinek et al., 2012). Placing genes encoding the sgRNA and the Cas9 on plasmids then enables transfection into the cells of interest. Cas9 recognizes the target site through a protospacer adjacent motif (PAM) next to the target sequence. The Cas9-sgRNA complex binds to the target DNA via the sgRNA's complementary guide sequence, resulting in a DSB mediated by Cas9's RuvC and HNH nuclease domains. The cell's repair mechanism then either rejoins the broken DNA ends through non-

homologous end joining (NHEJ) or homology-directed repair (HDR) (Figure 1.6) (Hille et al., 2018, Anzalone et al., 2020). As a result, the target is disrupted.

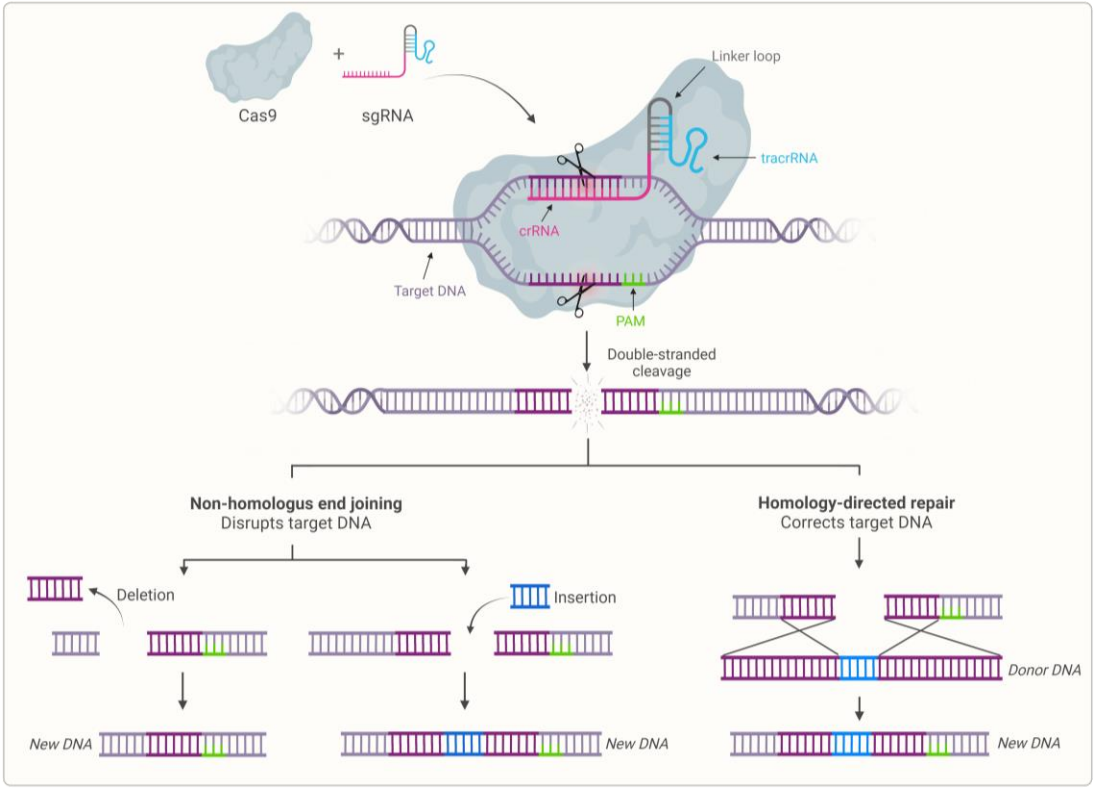


Figure 1.6: The mechanism of the CRISPR/Cas9 gene editing system. The illustration depicts the CRISPR/Cas9 complex, complete with the sgRNA bound to the target DNA. Together, this system can accurately target regions in the DNA for genome editing. The sgRNA is composed of the tracrRNA and the crRNA fused by a linker loop, where the crRNA is engineered to be complementary to the target DNA. After the crRNA sequence binds to the DNA, Cas9 recognizes the PAM sequence next to the target and performs double stranded cleavage by the HNH and RuvC endonuclease domains. The cleaved DNA sequence is repaired by the cell's own repair mechanism, either non-homologous end joining (NHEJ) or by homology-directed repair (HDR). The NHEJ pathway causes the target DNA to be disrupted either by deletion or insertion of DNA, while the HDR pathway uses a donor DNA template to repair the break. Figure created with BioRender.com adapted from Anzalone et al. (2020).

Compared to earlier methods, the CRISPR/Cas9 system offers a quicker, cheaper, and more straightforward approach, requiring only the engineering of a short RNA molecule. In recent years, various improvements and adaptations have enhanced the system's versatility. Some of these adaptations even enable precise editing of single nucleotides by relying on HDR and providing a template containing the single mutation of interest (Li et al., 2022). However, traditional CRISPR/Cas editing strategies remain inefficient due to their reliance on DSBs in the DNA, which are related to large deletions, complex rearrangements, and apoptotic responses in the cells (Haapaniemi et al., 2018, Ihry et al., 2018, Kosicki et al., 2018).

1.2.2. Genome editing strategies: Prime editing

To overcome the limitations of traditional CRISPR/Cas9, base editing offers a solution by chemically altering single nucleotides without cleaving the DNA backbone. This approach utilizes a deaminase enzyme to perform the four transition mutations (C→T, T→C, A→G, G→A). It cannot, however, perform the eight transversion mutations (C→A, A→C, C→G, G→C, G→T, T→G, A→C, C→A) (Rees and Liu, 2018, Anzalone et al., 2019). Furthermore, no method has been reported to be able to perform targeted deletions or insertions without DSBs, apart from prime editing (PE).

PE is a CRISPR/Cas9-derived editing strategy that uses a Cas9 variant with an inactivated HNH domain, termed Cas9 nickase (nCas9) fused to a reverse transcriptase (RT) (Figure 1.7 step 1). The system utilizes a special prime editing guide RNA (pegRNA) that both specifies the target site and encodes the desired edit (Anzalone et al., 2019). The pegRNA guides the nCas9 to perform a nick at the target site, which exposes a 3' hydroxyl group and primes the strand for reverse transcription. The extension on the pegRNA that encodes the edit is then transcribed directly into the target site (Figure 1.7 step 2). This creates two exposed DNA strands: an edited 3' flap and an unedited 5' flap (Figure 1.7 step 3). Although re-incorporating the 5' flap is thermodynamically favored due to the mismatch in the 3' flap, cellular repair mechanisms tend to remove the 5' flap as it is the preferred target of structure-specific endonucleases (Liu et al., 2004, Keijzers et al., 2015). However, the mismatch in the 3' flap still needs to be repaired by the cells own repair mechanism, such as mismatch repair (MMR), which means that whether the edit is incorporated or reverted back to its original sequence is random (Anzalone et al., 2019).

There are several PE systems, each containing slight variations from each other. PE1 consists of a nCas9 fused to a wild type (WT) RT. Mutation of the RT to increase editing efficiency created the PE2 system. The system is called PE3 if an sgRNA is included (in addition to the pegRNA) that guides the nCas9 to nick the unedited strand as well, as the presence of a nick on the complementary strand can cause the edited strand to be favored during DNA repair. To minimize the presence of unwanted simultaneous nicks, nicking the complementary strand should happen after edit-conversion. This can be achieved by designing an sgRNA with a spacer that specifically binds the strand after it has been edited. This system is denoted PE3b, and results in a decrease of the amount of indels compared to PE3 (Anzalone et al., 2019). Other improvements have also been made to the system in recent years. Studies involving CRISPR interference uncovered that specific DNA MMR genes greatly inhibit PE efficiency by reverting the edited strand back to its original sequence, and subsequently the introduction of a dominant negative MMR protein (MLH1) gave rise to two new PE systems: PE4 (PE2 including MLH1) and PE5 (PE3 including MLH1) (Yan et al., 2020, Chen et al., 2021, Chen and Liu, 2023). In addition, further mutation of the PE fusion protein resulted in the development of an enzyme with improved editing efficiencies, named PEmax (Chen et al., 2021).

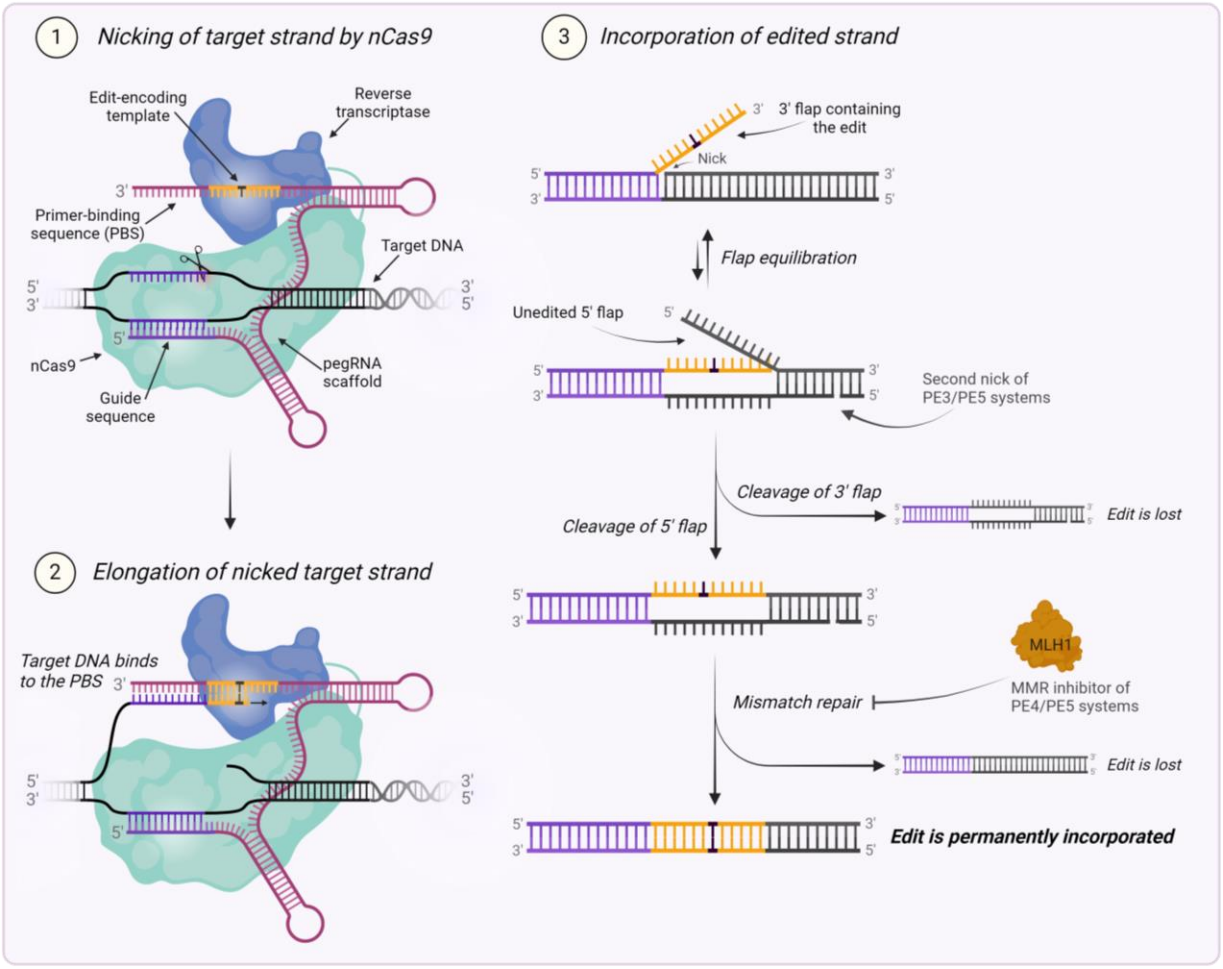


Figure 1.7: The mechanism of the prime editing (PE) systems. (1) The goal of the PE system is to edit a target in the genome from one sequence to another. For this to work, the PE complex consists of a Cas9 nickase (nCas9) and reverse transcriptase (RT) fusion protein complexed with a pegRNA. The pegRNA contains the primer binding sequence (PBS), the edit encoding template, the guide sequence (spacer), and the scaffold that binds everything together. These components enable the targeting of the correct genomic region and provides a template for the RT to transcribe the edit. The DNA is targeted by the spacer, which binds to the strand complementary to the target. This enables the nCas9 to create a nick on the target strand, resulting in a free 3' end. **(2)** The PBS hybridizes to the 3' target strand, and the RT transcribes the edit from the 3' end, extending it. **(3)** The transcribed strand results in an edited 3' flap, which overlaps the strand on the other side of the nick (unedited 5' strand). These flaps will equilibrate, which often will lead to 5' strand excision due to being favorized by structure-specific endonucleases. The 3' flap still contains a mismatch that needs to be repaired by the cells own repair mechanisms, such as DNA mismatch repair (MMR). Whether the edit is incorporated or reverted to its original sequence therefore happens at random. The PE3 and PE5 systems therefore include the presence of an sgRNA that induces a nick on the complementary strand, which will bias the cell's repair mechanism into incorporating the edit. To further favor the edited strand, MMR is inhibited in the PE4 and PE5 systems by inclusion of an MMR inhibitor, which inhibits removal of the edited strand. The edit will eventually be permanently incorporated into the genome. Figure created with BioRender.com adapted from Anzalone et al. (2019), Chen and Liu (2023).

The most common method for introducing the PE system into the target cell is to encode the components on vector plasmids for transient transfection. This results in transient expression of the PE components, reducing the chances of off-target effects. However, transient transfection can be challenging in some cell types, due to the length of the PE coding sequence resulting in a very large plasmid (Krill, 2023). To overcome this hindrance, the system can be adapted for lentiviral delivery, where the virus packages and delivers the prime editor sequence to the cell where it will be permanently incorporated into the genome. A lentiviral approach creates a higher chance of off-target effects since the PE protein will be constitutively expressed, but the risk can be reduced by transiently expressing the remaining components of the system (i.e. pegRNA).

When choosing a genome editing system, various factors must be considered, including the presence of a PAM-site near the target, cell type, type of edit, delivery method, and downstream applications. Flexibility in targeting is offered by different Cas9 variants recognizing alternative PAM-sites, as well as variations in size and amino acid composition (Edraki et al., 2019). Ultimately, the choice of editing system comes down to specific experimental requirements. Taken together, the benefits of the PE system make it particularly attractive for editing SNPs and therefore a valuable tool in the experimental approach to validate putative causal SNPs discovered by GWASs.

1.3. A genetic locus associated with visceral obesity

The present thesis focuses on investigating a genetic variant associated with visceral obesity and unraveling the biological mechanisms. Visceral adipose tissue (VAT) – deposits of fat surrounding the internal organs – has been linked to an increased risk of cardiovascular and metabolic disease, as well as increased cancer-specific mortality. Subcutaneous fat storage, on the other hand, is considered neutral or even beneficial (Koster et al., 2015, Karlsson et al., 2019). Visceral obesity falls under the category of complex polygenic diseases and is therefore a suitable candidate for GWASs. While multiple GWASs have identified more than 1000 SNPs associated with adiposity in general as measured by BMI (Loos and Yeo, 2022), very few studies have addressed the genetic associations with fat distribution. Obtaining this information is vital since visceral fat storage, as opposed to high BMI, is particularly associated with metabolic disease. However, in order to properly quantify visceral fat, imaging modalities such as MRI or DEXA must be performed, which are profoundly more costly and time consuming than measuring BMI. A steppingstone in this area was made with a 2019 GWAS conducted by Karlsson et al. (2019), which identified 102 novel loci linked specifically to visceral adiposity.

One notable adipogenesis-associated locus uncovered by Karlsson et al. (2019) is situated at the chromosomal location 11q23.2, harboring the *HMBS*, *VPS11* and *HYOU1* genes. Within this region, which contains 38 SNPs in LD (Samuelsen, 2021), eQTL signals have been observed in VAT for the *HMBS* and *VPS11* genes, as well

as for the lncRNA *RP11-110I1.14*, indicating potential gene regulatory mechanisms underlying the association. One SNP in particular, rs1799993, is situated near the proximal promoter of a short isoform of *HMBS* (Figure 1.8), and luciferase analyses found the (A) risk-allele to slightly increase promoter activity *in vitro* compared to the protective (C) allele. Moreover, the risk allele was associated with an increased expression of the *HMBS* gene (isoform not specified), and an increase in VAT mass (Karlsson et al., 2019). However, because the short *HMBS* gene is only expressed in blood (Chretien et al., 1988, UniProt, 2022), altered expression of this isoform is less likely to contribute to fat storage. The promoter of the short isoform lies within the first intron of the long *HMBS* isoform, which is ubiquitously expressed, and it is therefore likely that the region also harbors an enhancer element that is active in other cell types. Indeed, ENCODE data supports the presence of an enhancer overlapping rs1799993 (Figure 1.8). This enhancer may regulate the expression of the long *HMBS* isoform or other distant genes, and it may be this function, rather than the promoter function, that is altered by the risk allele of rs1799993 and contributing to visceral obesity.

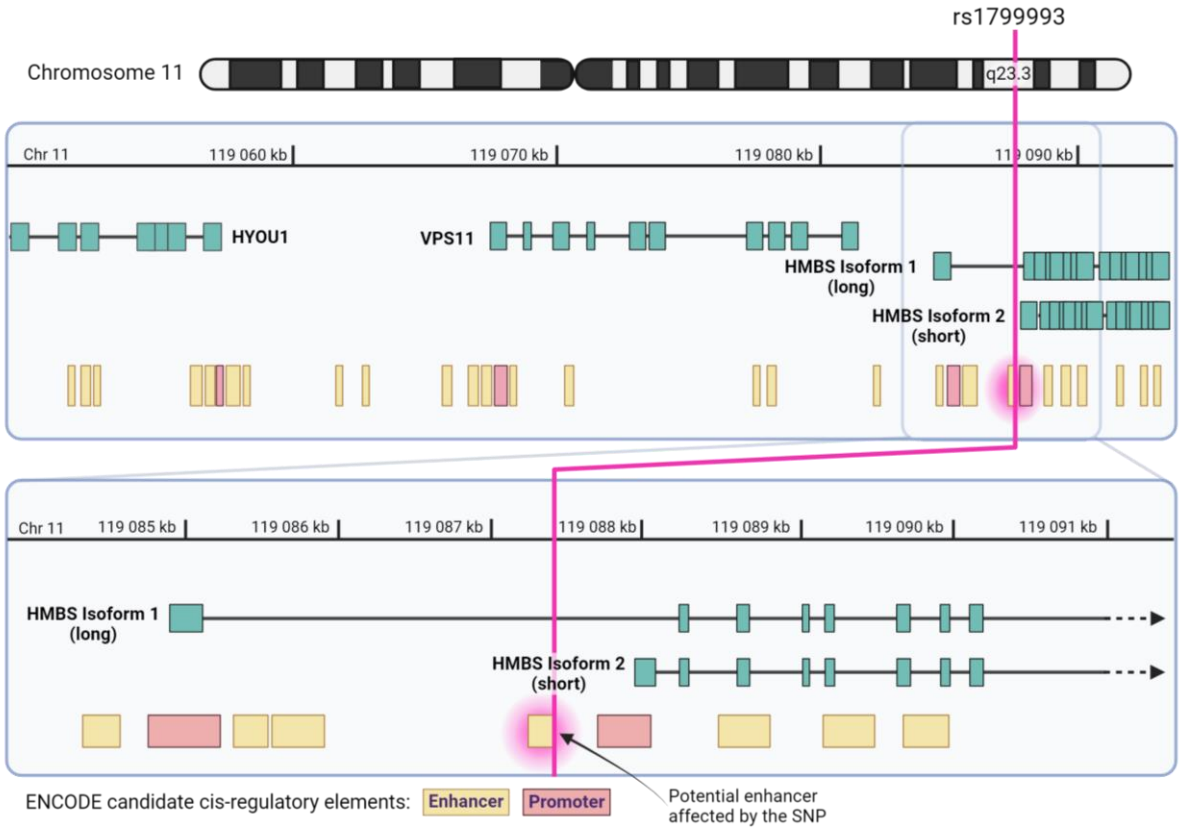


Figure 1.8: Illustration of the genomic location 11q23.3 and the SNP rs1799993. Exons are denoted by green boxes, introns by black lines in between. Yellow and red boxes denote regions associated with enhancer or promoter activity (respectively), according to data from ENCODE candidate cis-regulatory elements. 38 single nucleotide polymorphisms (SNPs) in linkage disequilibrium have been discovered within the chromosomal region 11q23.3, which harbors the genes *HMBS*, *VPS11*, and *HYOU1*. One SNP, rs1799993 (pink line), is located near the promoter of the short isoform of *HMBS*, and within the first intron of the long isoform of *HMBS*. The risk allele for the SNP (A) was found to increase promoter

activity *in vitro* compared to the protective allele (C). The risk allele was also associated with an increase in expression of *HMBS* (unspecified isoform), and an increase in visceral adipose tissue. The short isoform is, however, only expressed in blood, and it is therefore likely that the promoter overlaps a region with enhancer activity that is active in other cell types. Indeed, ENCODE candidate cis-regulatory elements shows a potential enhancer region implicated by the SNP which is highlighted in pink. Illustration created with BioRender.com, adapted from Krill (2023), information gathered from UCSC Genome Browser on Human (GRCh38/hg38).

1.3.1. The biology of fat cells

As the SNP rs1799993 is associated with obesity and adipose tissue, adipocytes are an intuitive choice as a model system to study its genotypic effect. Adipose tissue is a loose connective tissue primarily consisting of adipocytes derived from mesenchymal stem cells (MSCs). The process by which precursor cells mature and differentiate into adipocytes is called adipogenesis. It is regulated by various factors, with the TF peroxisome proliferator activated receptor gamma, PPARγ, acting as the master regulator (Rosen and Spiegelman, 2014, Richard et al., 2020). Additionally, steroids play a significant role in the regulation of adipogenesis, as adipocytes express several steroid hormone receptors.

White adipose tissue (WAT), brown adipose tissue (BAT), and beige adipose tissue are three of the main components of adipose tissue (Richard et al., 2020). The differences between them give rise to specialized tissue functions. For example, WAT plays a crucial role in energy storage, endocrine signaling, and insulin sensitivity, while BAT is associated with non-shivering heat production for maintaining body temperature and contributes to fat oxidation and diet-induced thermogenesis (Richard et al., 2020). BAT and WAT cells also differ in shape, size, and intracellular structure. White adipocytes typically contain a single large lipid droplet that occupies the majority of the intracellular space, allowing for extensive lipid storage (Richard et al., 2020). In contrast, brown adipocytes contain multiple smaller lipid droplets surrounded by iron-rich mitochondria, which are considered responsible for BAT's thermogenic activity (Mota de Sá et al., 2017, Richard et al., 2020). Beige adipocytes exhibit characteristics of both white and brown adipocytes. These cells represent a distinct type of thermogenic fat cells that start out as white adipocytes that can undergo 'beiging' or 'browning' in response to stimuli like exercise, diet, or pharmacological agents (Wang et al., 2019, Kaisanlahti and Glumoff, 2019, Mika et al., 2019). This process may give protective effects against obesity-related metabolic dysfunction (Richard et al., 2020).

The majority of WAT is categorized as either subcutaneous adipose tissue (SAT) or VAT. Accumulation of VAT has been linked to insulin resistance, metabolic disease, and cardiovascular disease, in contrast to SAT accumulation, which may be protective against metabolic syndrome (Koster et al., 2015, Zhang et al., 2015, Smith et al., 2019). Observations like these lead to the distinction between metabolically healthy and

metabolically unhealthy obesity. Metabolically healthy obese individuals are characterized with high fat mass, high SAT mass, low VAT mass, high insulin sensitivity, low levels of liver fat, and normal blood pressure (Smith et al., 2019), suggesting that the location of adipose tissue may be a strong predictor of metabolic health risks rather than looking at overall obesity (Harvey et al., 2020, Richard et al., 2020). Consequently, studying the genetics and biology of VAT accumulation becomes imperative for understanding obesity-related diseases.

1.3.2. Previous findings and implications for future investigations

Building upon the discoveries made by Karlsson et al. (2019), our lab hypothesized that the SNPs within the 11q23.3 locus associated with VAT mass are affecting enhancer activity rather than promoter activity. Our approach aimed to identify genetic regions within the locus exhibiting cis-acting enhancer activity, predict the causal variant(s), and assess the effect of genetic manipulation *in vitro.* To narrow down putative causal SNPs, ENCODE epigenetic data was used to identify predicted enhancers, and these regions were cloned from human genomic DNA harboring the risk or protective haplotypes for the LD-block at 11q23.3 into luciferase reporters and tested for enhancer activity. Several regions showed enhancer activity, but only one, which contains rs1799993 and five other SNPs, showed genotype-dependent enhancer activity when the two haplotypes were compared (Samuelsen, 2021, Mirza, 2022).

When mutated individually from protective to risk genotype, four of the SNPs, including rs1799993, each resulted in a significantly increased gene regulatory activity in the luciferase assay (Samuelsen, 2021, Mirza, 2022). While *in silico* screens identified predicted transcription factor binding sites overlapping all five SNPs, only rs1799993 had a predicted genotypic impact on binding affinity. Specifically, RARB, RARG and RXRB were all found to bind the protective, but not the risk variant. Moreover, the disruption of the RARB binding site was predicted to potentially disrupt a retinol-RARB anti-adipogenic signaling pathway, revealing a potential mechanism that could lead to increased VAT storage in risk carriers (Mirza, 2022). Thus, with a plausible upstream biological pathway connected to the associated phenotype, we considered rs1799993 as the best candidate for functional validation by genome editing.

To obtain a biologically meaningful result, the genome editing should be performed in a cell type where the SNP, and the enhancer in which it resides, are active. The ENCODE epigenetic data used to identify the enhancer signal surrounding rs1799993 also revealed that this enhancer was specifically active in adipose-derived MSCs (AD-MSC), but not in mature adipocytes (Mirza, 2022). Thus, a human AD-MSC cell line (ASC52telo), which is heterozygous for rs1799993, was selected for genome editing by PE. Spacers for pegRNAs targeting rs1799993 to edit the SNP from risk allele (A) to protective allele (C) and vice versa were been designed by *in silico* analyses and validated for on-target efficiency *in vitro* (Samuelsen, 2021, Krill, 2023). However, despite extensive testing of several types of transfection methods (including electroporation, which normally leads to superior transfection efficiencies (Chong et al., 2021), the ASC52telo cells did not

take up the large (~10.6 Kb) PE plasmids (Krill, 2023). Moreover, while another pilot PE experiment in a preadipocyte cell line (A41WAT) did result in some clones taking up the plasmids, only four of the clones survived and none of them were edited (Samuelsen, 2021). Thus, due to difficulties in transfecting human adipose-derived stem-cells and preadipocytes, the performance of the PE method itself, including the designed pegRNAs, has not been validated in cell culture in our lab. Consequently, the current thesis is dedicated to address these challenges and advance the development of a PE protocol in MSCs using the improved PEmax enzyme.

Aims

This thesis is part of a larger project aiming to dissect the biological mechanisms underlying the association between genetic variants in the 11q23.3 locus and visceral obesity. A key step in achieving this aim is genome editing of the predicted causal SNP, rs1799993, in AD-MSCs. Prime editing is a CRISPR/Cas9 derived technique particularly suited to single base pair genomic editing, but the method has not been optimized for use in AD-MSCs, a cell type that is very hard to transfect.

The overall aims of this thesis are to validate the prime editing method in an easy-to-transfect human diploid cell (HT1080) line and to optimize it in adipose-derived MSCs.

Specific objectives include:

- Genotype rs1799993 in the HT1080 cell line in preparation for proof-of-concept genome editing.
- Establish a transfection protocol for PE4 in the HT1080 cells.
- Optimize clonal expansion, FACS, and high-throughput screening of HT1080 mutants.
- Validate performance of the designed pegRNAs in PE targeting rs1799993.
- Adapt a lentiviral delivery approach of PEmax in the mesenchymal stem cell line ASC52telo.
- Optimize clonal expansion, FACS, and high-throughput screening of ASC52telo mutants.
- Establish a cell line of ASC52telo that stably expresses PEmax.
- Transiently introduce pegRNAs to edit rs1799993 from heterozygous to homozygous risk and protective alleles.

2. Materials

Table 2.1 | Online software

Resource	Purpose	URL
Benchling Molecular Biology	Visualization and annotation of DNA sequences	https://www.benchling.com/
Human BLAT Search	Evaluation of primer design by finding sequence similarities in the genome.	https://genome-euro.ucsc.edu/cgi-bin/hgBlat
BioRender	Creating scientific illustrations	www.BioRender.com
NCBI Primer-BLAST	Designing primers	https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi
OligoEvaluator™	Evaluating primer design	http://www.oligoevaluator.com/LoginServlet

Table 2.2 | Software

Product	Supplier	Application/description
Fiji ImageJ	Open source	Image processing
DNA Baser Sequence assembly software	Heracle BioSoft S.R.L.	Visualizing DNA sequences
FlowJoTM v.10.10.0	Dickinson and Company	Handling and visualization of FACS data
IncuCyte 2021C	Sartorius	Live-cell imaging and analysis
Prism	GraphPad	Handling and visualization of data
SnapGene® software	Dotmatics	Visualization of DNA sequences and chromatograms

Table 2.3 | Instruments

Product	Supplier	Application
QIACube	QIAGEN	Nucleic acid purification
QIAxpert	QIAGEN	Quantification of nucleic acids
Countess® II	Life Technologies	Counting of live and dead cells in culture
GeneAmp® PCR System 9700	Applied Biosystems	PCR, sequencing
LightCycler® 480 II	Roche	Quantitative real-time PCR (qPCR)
NanoDrop® ND-1000 Spectrophotometer	Thermo Scientific	Measurement of nucleic acid concentrations
Nikon Eclipse TS100	Nikon	Cell monitoring and imaging

Nikon TE2000 Eclipse	Nikon	Monitoring of fluorescence
Nikon Eclipse TI2	Nikon	Cell imaging and monitoring of fluorescence
Sartorius IncuCyte S3 Live-Cell Analysis systems	Sartorious	Live cell monitoring and imaging
SONY SH800S Cell Sorter	SONY Biotechnology Inc.	Flow cytometry and Fluorescence Activated Cell Sorting (FACS)
Gel Doc™ EZ Imager	Bio-Rad	Agarose gel electrophoresis imaging
Gel Doc™ XR System	Bio-Rad	Western Blot imaging
Trans-Blot® Turbo™ Transfer System	Bio-Rad	Protein to membrane transfer for Western Blot
PowerPac™ High Current Power Supply	Bio-Rad	SDS-PAGE protein separation
SPECTRAMax® *PLUS*^{384} Microplate Reader	Molecular Devices	Measurement of protein concentrations

Table 2.4 | Plasmids

Plasmid name	Size (bp)	Supplier	Application
pEF1a-hMLH1NTD-NLS	5109	Addgene (#174826)	PE, increases editing efficiency.
pLenti-PE2max-BSD	15 050	Addgene (#191102)	PE, lentiviral cargo plasmid of PEmax
pCMV-PEmax-P2A-GFP	10 578	Addgene (#180020)	PE, transient expression of PEmax
pU6-pegRNA-rs1799993-Peg1-A	2259 + spacer and extension (table 2.5)	Based on Addgene #132777, cloned by Nina T. Samuelsen	PE, plasmids encoding different pegRNA sequences. (Table 2.5) Transient expression.
pU6-pegRNA-rs1799993-Peg2-A		Based on Addgene #132777, cloned by Annika S. S. Krill	
pU6-pegRNA-rs1799993-Peg3-A		Based on Addgene #132777, cloned by Annika S. S. Krill	
pU6-pegRNA-rs1799993-Peg1-C		Based on Addgene #132777, cloned by Nina T. Samuelsen	
pU6-pegRNA-rs1799993-Peg2-C		Based on Addgene #132777, cloned by Annika S. S. Krill	

pU6-pegRNA-rs1799993-Peg3-C		Based on Addgene #132777, cloned by Annika S. S. Krill	
pCMVR8.74	11 904	Addgene, #22036	2nd generation lentiviral packaging plasmid expressing GAG, POL, REV, and TAT.
VSV-G	5 551	Plasmid gifted from Pouda Panahandeh	2nd and 3rd generation lentiviral plasmid. Encoding envelope protein.
pMax_GFP	3 527	Addgene, #177825	Fluorescent marker, monitoring transfection efficiencies, positive control in FACS.
pCMV6-XL5-empty	4 707	Origene, #PCMV6XL5	Empty vector, used as control

Table 2.5 | pegRNA spacers and extensions

pegRNA name	Spacer sequence (5'->3')	Extension sequence* (5'->3')
Peg1-A	GAAACCAGAGGCAGCACTCT	TTCGTtCACCTAGA*GTGCTGCCTCTGGTT*
Peg1-C		TTCGTgCACCTAGA*GTGCTGCCTCTGGTT*
Peg2-A	GAAACCAGAGGCAGCACTCT	TAAAGTTCGTtCACCTAGA*GTGCTGCCTC*
Peg2-C		TAAAGTTCGTgCACCTAGA*GTGCTGCCTC*
Peg3-A	TTCCATGTTGGTGAAACCAG	TTCGTtCACCTAGAGTGCTGCCTCTG*GTTTCACCAACA*
Peg3-C		TTCGTgCACCTAGAGTGCTGCCTCTG*GTTTCACCAACA*

**Italic= Primer binding sequence*
Underlined = Edit-encoding reverse complementary template
Red = Complementary template for the (A) risk allele
Blue = Complementary template for the (C) protective allele

Table 2.6 | Patient samples

Name	Description	Genotype	Ethical approval number
V586	gDNA	A/A	From the Western Norway Obesity Biobank (WNOB), approved by REK 2010/502 and REK 2018/2020
V602	gDNA	C/C	

Table 2.7 | Cell lines

Cell line	Description	Ethical approval number	Supplier (cat.no)
ASC52telo	Mesenchymal stem cells derived from adipose tissue. Exhibits fibroblast-like morphology. hTERT immortalized.	EK589/2011, EK200/2005	SCRC-41000
HEK293T	Epithelial-like cells derived from human embryonic kidney (HEK) cells.	-	Gift from Nils Henrik Halberg, University of Bergen, Norway.
HT1080	Epithelial cells derived from fibrosarcoma connective tissue.	-	CCL121

Table 2.8 | Cell culture components

Component	Supplier	Cat.no
Accutase	MerckMillipore	SCR005
Amphotericin B solution	Sigma-Aldrich	1397-89-3
Blasticidin	InvivoGen	Ant-bl
CoolCell® Container	Corning	432000
DMEM High Glucose	VWR	392-0416
Dimethyl Sulfoxide (DMSO)	EMD Millipore	317275
EMEM	Lonza™	12-662F
Fetal Bovine Serum (FBS), Qualified	Gibco™	10270-106
FGF-2 human	Merck	SRP4037
G418	InvivoGen®	108321-42-2
Gentamicin solution	Sigma-Aldrich	1405-41-0
Human EGF	Cell Signaling Technology	72528S
Low Temperature Freezing Vials	VWR	479-1208
MEM	Gibco™	10370-021
Mesenchymal Stem Cell Basal Medium	ATCC®	PCS-500-030
Mesenchymal Stem Cell Growth Kit	ATCC®	PCS-500-040
Nunc™ Cell Culture Dishes	Thermo Scientific	168381
Nunc™ Cell-Culture Treated Multidishes	Thermo Scientific	140675
Nunc™ MicroWell™ 96-Well, Nunclon Delta-Treated	Thermo Scientific	167008
Trypan Blue Stain (0.4%)	Invitrogen™	T10282

Opti-MEM™	Gibco™	31985-062
Pen/Strep Amphotericin B	Lonza™	17-745E
Penicillin and Streptomycin (PEST)	Sigma-Aldrich	P0781

Table 2.9 | Transfection reagents

Reagent	Supplier	Cat.no
Lipofectamine Stem Transfection Reagent	Invitrogen™	STEM00008
P1 Primary Cell 4D-Nucleofector™ X Kit	Lonza	V4XP-1032
TransIT®-LT1 Transfection Reagent	Mirus	MIR2300

Table 2.10 | Buffers and solutions

Buffer/solution	Composition*	Application
1x PBS	0.2 g/l KCl, 0.24 g/l KH2PO4, 8 g/l NaCl, 1.44 g/l Na2HPO4, pH adjusted with HCl to 7.4	Cell culture and Western Blot
PBST	1x PBS, 0.1% (v/v) Tween	Western Blot
1x TAE	2 M Tris, 1 M EDTA-Na_2, 1 M Acetic acid	Agarose gel electrophoresis
FACS buffer	1x PBS, 0.1% BSA (v/v), 2 mM EDTA	FACS and flow cytometry
LB-medium	2.5% (w/v) BD Difco Dehydrated Culture media, pH adjusted with NaOH to 7.5	Bacterial culture
RIPA WB lysis buffer	1x RIPA (Millipore), 1x Roche protease inhibitors, 5 mM EDTA, 0.1 µM Na orthovanadate	Cell lysis for Western blot

** All concentrations given as final concentrations*

Table 2.11 | Primers for PCR and sequencing

Primer name (ID)	Sequence (5' → 3')	Binds to	Application
AK63_dCas9_fwd3	TCACCGACGAGTACAAGGTGC	*Cas9*, forward	qPCR and sequencing, targets Cas9
AK64_dCas9_rev3	CGCCGATCAGGTTCTTCTTGAT	*Cas9*, reverse	qPCR and sequencing, targets Cas9
IP11_PEmax_F	GTGATGAAGCAGCTGAAGCG	*PEmax (nCas9)*, forward	qPCR, targets nCas9
IP12_PEmax_R	AGGATTGTCTTGCCGGACTG	*PEmax (nCas9)*, reverse	qPCR, targets nCas9
128F_human_RPS13	TTCACCGTTTGGCTCGATA	*RPS13*, forward	Reference gene for qPCR

129R_human_ RPS13 R	AAACAATCATTTTATTGCTTGAGTACA	*RPS13*, reverse	Reference gene for qPCR
RPLP0 F	TCTACAACCCTGAAGTGCTTGAT	*RPLP0*, forward	Reference gene for qPCR
RPLP0 R	CAATCTGCAGACAGACACTGG	*RPLP0*, reverse	Reference gene for qPCR
16q12.2_F	TTTTCCTTCTCTCCTTCTGCCC	Genomic region at position 16q12.2, forward	gDNA reference position for qPCR
16q12.2_R	AAATACAGTCAGGTAGGTTTCGT	Genomic region at position 16q12.2, reverse	gDNA reference position for qPCR
643F_HMBS_ rs1799993	TGTCTGATGTAGAAAGAAATGGGA	*HMBS* and rs1799993, forward	Genotyping and sequencing of rs1799993 in HT1080 cells
644R_HMBS_ rs1799993	CTGGGGAACCTGTGCTGAG	*HMBS* and rs1799993, reverse	Genotyping and sequencing of rs1799993 in HT1080 cells
728F_rs1799993_ C_prot	AGAGGCAGCACTCTAGGTGC	*rs1799993_C* forward	qPCR screening of mutants
729R_rs1799993_ U	AAAGGTTGGACTCAGGTTTCTTCT	*rs1799993_A/C* reverse	qPCR screening of mutants
730F_rs1799993_ A_risk	AGAGGCAGCACTCTAGGTGA	*rs1799993_A* forward	qPCR screening of mutants
141F_VP1.5_ pCMV6-XL	GGACTTTCCAAAATGTCG	pCMV plasmid, forward	Sequencing
04R_BGH_ pCR3.1	TAGAAGGCACAGTCGAGG	*MLH1*, reverse	Sequencing

Table 2.12 | Components and kits for DNA and RNA methods

Component	Supplier	Cat.no	Application
BigDye™ Terminator v.3.1 enzyme	Applied Biosystems™	4337455	Sequencing
BigDye™ Terminator v.3.1 buffer	Applied Biosystems™	4337455	Sequencing
Chelex 100	Bio-Rad	1421253	gDNA extraction

GenElute™ Mammalian Genomic DNA Miniprep kit	Sigma-Aldrich	G1N70	gDNA extraction
GeneRuler 1 kb Plus DNA Ladder	Thermo Fischer	SM1331	Agarose gel electrophoresis
High-Capacity cDNA Reverse Transcription Kit	Applied Biosystems™	4368814	cDNA synthesis
LightCycler® 480 SYBR Green I Master	Roche	04887352001	qPCR
RNA/DNA/Protein Purification Plus	Norgen Biotek	47700	DNA and RNA purification
RNeasy Mini QIAcube Kit	QIAGEN	74116	RNA purification
Buffer RLT	QIAGEN	79216	Cell lysis
HiSpeed Plasmid Maxi Kit	QIAGEN	12663	Plasmid purification
Platinum™ SuperFi™ II PCR	Invitrogen™	12368010	PCR
Illustra™ ExoProStar 1-step kit	Cytiva	US77701	PCR cleanup
SeaKem® LE agarose	Lonza	50004	Agarose gel electrophoresis
GelRed™	VWR	730-2958	Agarose gel electrophoresis
RNase A	QIAGEN	19101	Chelex 100 gDNA extraction
Proteinase K	QIAGEN	19131	Chelex 100 gDNA extraction

Table 2.13 | Components and kits for protein methods

Component	Supplier	Cat.no	Application
Bovine Serum Albumin	Sigma-Aldrich	SLCJ4594	Blocking of RetroNectin coating and Western Blot
DC™ Protein Assay Kit II	Bio-Rad	500-0112	Protein quantification
10x RIPA lysis buffer	Millipore	20-188	Protein extraction
96F Microwell plate	Thermo Scientific	269620	Protein quantification
cOmplete™, Protease Inhibitor Cocktail	Roche	04693116001	Protein extraction
Ponceau S	MP Biomedical	190644	Protein visualization
Precision Plus Protein™ Kaleidoscope™ Prestained Protein Standards	Bio-Rad	1610375	SDS-PAGE
Restore™ Western Blot Stripping buffer	Thermo Scientific	21059	Western blot membrane stripping

RNA/DNA/Protein Purification Plus	Norgen Biotek	47700	Protein extraction
SDS Sample Buffer	Sigma-Aldrich	70607-3	SDS-PAGE
SuperSignal™ West Femto Maximum Sensitivity Substrate	Thermo Scientific	34096	Western blot detection
SurePAGE™, Bis-Tris 4-20% gel	GenScript	M00657	Western blot
Trans-Blot Turbo Transfer Pack	Bio-Rad	1704159	Western blot member transfer
Tris-MOPS-SDS Running Buffer	GenScript	M00138	SDS-page running buffer

Table 2.14 | Antibodies for Western Blot

Component	Supplier	Cat.no	Application
Anti-beta Actin	Abcam	ab6276	Primary antibody targeting beta actin
Anti-CRISPR-Cas9 antibody	Abcam	ab191468	Primary antibody targeting Cas9
HRP Goat Anti-Mouse Ig	BD Biosciences	554002	Secondary antibody

Table 2.15 | Lentiviral transduction components

Reagent	Supplier	Cat.no	Application
Lenti-X™ Concentrator	Takara Bio	631232	Concentration of lentiviral particles
Lenti-X™ GoStix™ Plus	Takara Bio	631280	Quantification of lentiviral particles
RetroNectin	Takara Bio	T100A	Enhancement of lentiviral transduction

Table 2.16 | FACS components

Component	Supplier	Cat.no	Application
5 ml Polystyrene Round-Bottom Tube with 35 µm Cell-Strainer Cap	Corning	352235	Straining of cells prior to sorting
100 µm Sorting Chip	SONY	LE-C3210	Cell sorting
Automatic Setup Beads	SONY	LE-B3001	Calibration of SONY SH800S Cell Sorter

3. Methods

3.1. Ethics and approvals

The work in this project implicates the use of human biological material, which requires ethical considerations. The project has been approved by the Norwegian Regional Ethics Committee (REK). This includes the use of genomic DNA samples from the Western Norway Obesity Biobank (WNOB), approved by REK2010/502 and REK 2018/2020. All patients provided written informed consent. The ASC52telo immortalized cell line was developed and commercialized with the approval of the Ethics Committee of the University of Vienna, (EK589/2011) and Ethikkommission der Landes Oberösterreich, (EK200/2005).

3.2. Cell culture and cell culture methods

3.2.1. Cell lines

The present thesis utilized three cell lines: ASC52telo, HEK293T, and HT1080. HT1080 and ASC52telo were the targets for prime editing (PE) and were utilized in the majority of experiments. The HEK293T was specifically used for production of lentivirus.

ASC52telo is a human hTERT-immortalized adipose-derived MSC line that displays fibroblast-like morphology. This cell line has the potential to be differentiated into mature adipocytes in addition to osteoblasts and chondrocytes. HT1080 cells, on the other hand, are immortalized human epithelial cells derived from connective tissue of a patient with fibrosarcoma. This cell line is mostly diploid and known to be a suitable transfection host (ATCC, 2023), making it ideal for proof-of-principle genome editing. HEK293T is an immortalized human embryonic kidney cell line. It is known for being robust, fast growing, and easy to transfect. The presence of the oncogenic SV40 large T-antigen is also thought to promote plasmid-mediated gene expression and suppress antiviral defense mechanisms, and thus make HEK293T an ideal cell line for virus production (Ferreira et al., 2020). It is, however, not suitable for genome editing due to its high ploidy and unstable karyotype (Lin et al., 2014).

3.2.2. Cultures and passaging

The cells were cultured in cell-line specific growth medium (Table 3.1) in various culture-grade dishes, but most often in 10 or 15 cm dishes. Growth medium was changed every 2-3 days until cell confluency reached 70-90%, when they were subsequently passaged. Passaging was performed by either gentle manual detachment (HEK293T) or incubation with either trypsin (HT1080) or Accutase (ASC52telo) followed by centrifugation at 300 x g for 3 minutes, resuspension in fresh culture medium and seeding into a new culture dish. The cells were usually counted in order to seed an exact number but were sometimes instead split at an appropriate dilution for approximate cell number. To count the cells, a 20 µL aliquot of cell suspension was diluted 1:2

with Trypan Blue and counted using the Invitrogen Countess II Automated Cell Counter. All counts were performed in duplicates, and the mean was calculated based on the number of live cells measured. The percentage of live cells was used to assess the health of the cell culture.

The ASC52telo cells were seeded at a density of 5000 cells/cm^2 according to ATCC's recommendations. The HT1080 cells were sub-cultured at a ratio of between 1:4 to 1:8 also according to ATCC's recommendations. Cell lines were maintained for a maximum of 10 passages before being taken down, the exception being cells used for PE. HEK293T cells were sub-cultured at a ratio of between 1:10 to 1:20. All cell cultures were incubated at 37°C and with 5% CO_2.

Table 3.1: Cell culture medium

Cell line	Basal media	Additional supplements*
ASC52telo	Mesenchymal Stem Cell Basal Medium	MSC supplement (2% FBS, 5 ng/mL rh FGF basic, 5 ng/mL rh FGF acidic, 5 ng/mL rh EGF) L-Alanyl-L-Glutamine (2.4 mM) Amphonectrin B (250 ng/ml) G418 (200 ng/ml) PEST (1%, v/v)
HEK293T	DMEM High Glucose w/ L-Glutamine w/ Sodium Pyruvate	FBS (10%, v/v) PEST (1%, v/v)
HT1080	EMEM	FBS (10%, v/v) L-Glutamine (1%, v/v) PEST (1%, v/v)

** All concentrations given as final concentrations.*

3.2.3. Thawing cells

Cryopreserved cell lines were briefly thawed at 37°C and immediately diluted 1:10 in fresh culture medium. The diluted cell suspension was then centrifuged for 3 minutes at 300 x g to remove residual DMSO. The pellet was then resuspended in fresh medium, seeded in a suitable cell culture dish and cultured as described in section 3.2.1.

3.2.4. Freezing cells

Cells were cryopreserved in cell culture medium containing 10% DMSO at a concentration of 150 000 – 500 000 cells/ml in 1 ml cryotubes. The cryotubes were gradually cooled down and frozen over night at -80°C in a CoolCell Container before transfer to liquid nitrogen for long-term storage.

3.3. Amplification and purification of plasmids

3.3.1. Plasmids

The key plasmids utilized in this thesis include two PE plasmids (Figure 3.1 A-B), one suitable for transient transfection methods and another serving as a lentiviral cargo plasmid. Both PE plasmids encode the optimized version of nCas9 fused with reverse transcriptase (PEmax), but carry separate selection markers, specifically green fluorescent protein (GFP) and a blasticidin resistance gene (BSD), respectively. The lentiviral version additionally contains sequences required for packaging into lentiviral particles and integration into the host genome (Figure 3.1 B).

To direct the PEmax fusion proteins to the correct genomic locus and specify the edit, pegRNAs must be co-delivered with the PE plasmids. A pool of different pegRNA sequences were used in this thesis (Table 2.5), based on the same backbone plasmid (Figure 3.1 C) An MMR-inhibitor-encoding (MLH1) plasmid designed to improve editing efficiencies was included for the PE4 system (Figure 3.1 D). Furthermore, a GFP-encoding plasmid (Figure 3.1 E) was co-transfected with the pU6-pegRNA plasmids (no selection marker) to serve as a proxy for positive transfection.

Plasmids for lentiviral production include the 2nd generation PE lentiviral cargo plasmid (Figure 3.1 A), a 2nd generation lentiviral packaging plasmid expressing the key lentiviral genes (GAG, POL, REV, and TAT), and a lentiviral packaging plasmid encoding the VSV-G envelope protein, which is compatible with both 2nd and 3rd generation cargo plasmids (Table 2.4).

An empty vector used as control in experiments, pCMV6-XL5-empty, was also included in this thesis.

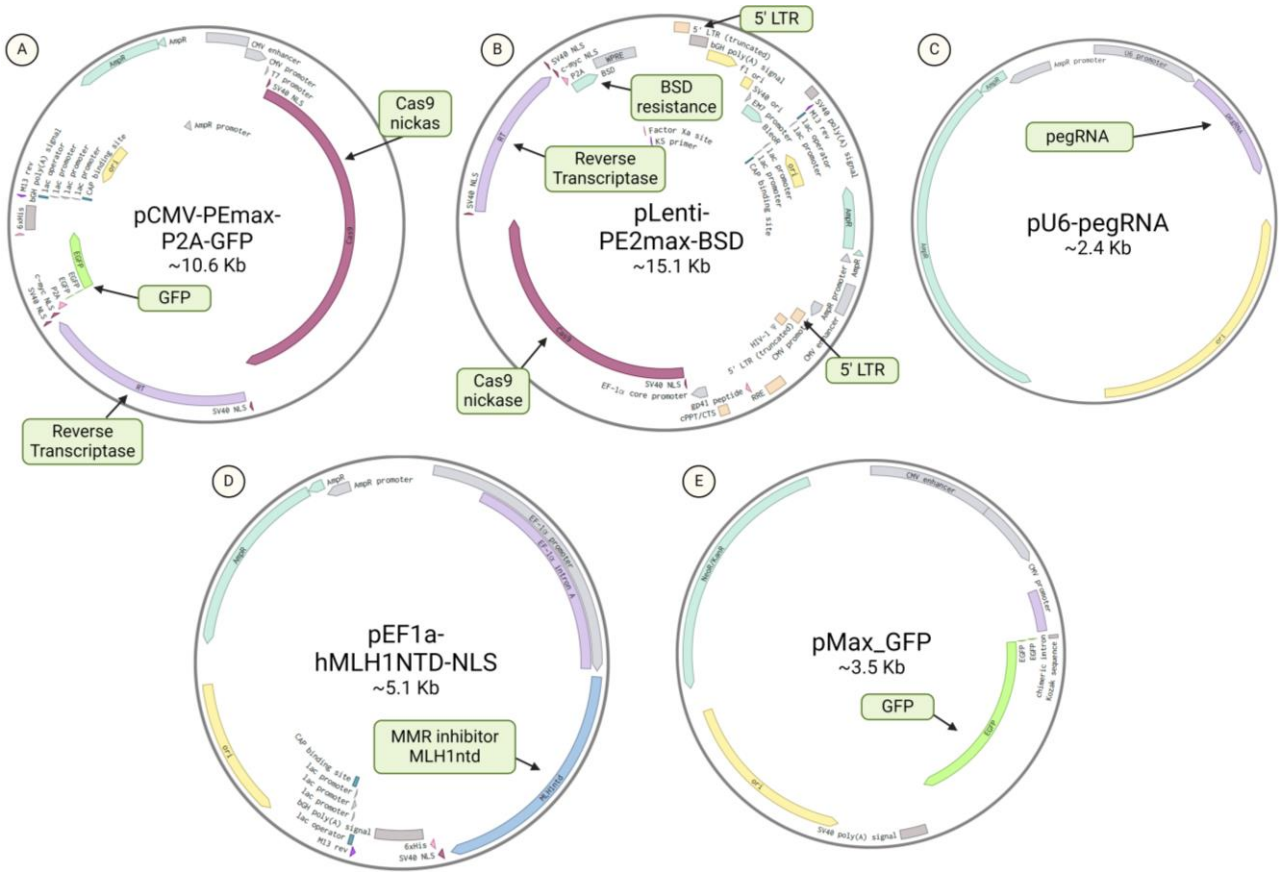


Figure 3.1: Plasmids maps of the plasmids utilized for PE. (A) PE plasmid encoding PEmax (Cas9 nickase + RT fusion protein) utilized in proof-of-concept genome editing of HT1080. Transfected by TransIT-LT1 along with plasmids (C) and (E). The plasmid contains GFP, making it possible to screen successfully transfected cells by imaging methods and FACS. **(B)** PE plasmid encoding PEmax (Cas9 nickase + RT fusion protein) utilized in lentiviral transduction. The plasmid was transfected into HEK293T cells along with the lentiviral packaging and envelope plasmids (not pictured, see Table 2.4) The area between the 5'-LTRs on the plasmid becomes packaged as RNA by the virus, which then becomes deposited as DNA permanently in the genome of the transduced cell. The plasmid contains a BSD resistance gene, making it possible to select successfully transduced cells by antibiotic selection. **(C)** PE pegRNA plasmid with pegRNA sequence inserted. Purple annotated area shows were pegRNA sequences for encoding A or C were previously cloned into the plasmids by Mirza (2022) and Krill (2023). See Table 2.5 for specific sequences. pegRNA plasmids were transfected by nucleofection into PE lentiviral transduced MSC. **(D)** Plasmid encoding the MMR inhibitor MLH1ntd of the PE4 and PE5 systems. Inhibition of MMR increases editing efficiency (Chen et al., 2021). **(E)** Plasmid encoding GFP, transfected by nucleofection into PE lentiviral transduced MSC, making selection by FACS possible for GFP-positive cells. This plasmid was also used as a positive control for GFP-positive cells. The combination of these plasmids in PE amounts to the PE4 system. Plasmid maps are collected from the manufacturer, analyzed in Benchling.com, and the figure is created with BioRender.com.

3.3.2. Amplification of plasmids in E. coli

Cultures of *E. coli* were used for plasmid amplification. *E. coli* in agar stabs pre-transformed with the plasmids of interest were purchased from the manufacturer and streaked on agar plates containing ampicillin (100 µg/ml) and left to grow over night at 37°C. The resulting colonies were assessed and single high-quality colonies (no satellite colonies, far away from other colonies) were chosen for further culture expansion. Mini-cultures were prepared by inoculating 5 ml LB-medium containing ampicillin (75 µg/ml) with the chosen colonies. The mini-cultures were incubated at 37°C over night while shaking (225 rpm).

Maxi-cultures were prepared by transferring the entire volume of the mini-culture to 150 ml of LB medium containing ampicillin (75 µg/ml) followed by incubation over night at 37°C while shaking (225 rpm). The bacteria were harvested by centrifugation at 5600 x g for 20 minutes at 4°C, and the supernatant discarded. While awaiting purification, the pellets were stored at -20°C. The bacteria were maintained at -80°C in 15% glycerol (v/v) stocks.

3.3.3. Maxi-prep plasmid purification

In order to purify the plasmids from the pelleted bacteria, the cells were lysed, and released plasmids purified using the QIAGEN HiSpeed Maxi Kit according to manufacturer's instructions. Plasmids were eluted in 500 µl TE buffer. Plasmid concentrations and purities were measured with the NanoDrop.

3.4. RNA and DNA methods

3.4.1. RNA extraction

When extracting RNA alone from cell cultures, the RNeasy Mini kit was utilized as per manufacturer's instructions. When the cells reached 80-90% confluency in 24-well or 12-well plates, they were lysed with 350 µl RLT buffer, scraped to aid lysis, before they were stored at -80°C unless RNA purification proceeded immediately. If the samples were frozen, they were thawed for 2 minutes at 55°C. The purification itself was done using the RNeasy Mini program on the QIACube. RNA concentrations were then measured on the QIAxpert or by NanoDrop.

3.4.2. gDNA extraction

To quickly harvest gDNA from a large number of samples, Chelex 100 chelating resin beads were utilized. The cells were grown in 96-well or 24-well plates, before a 5% Chelex solution was prepared with 0.2 mg/ml (f.c.) Proteinase K and 50 µg/ml (f.c.) RNase A. The cells were washed in PBS before 50 µl or 100 µl (respectively) of 5% Chelex solution was added to each well. To aid cell lysis the cells were also gently scraped

by the pipette tip while in solution. The cells were then set to incubate at room temperature for 2 minutes before a 4-hour incubation period at 56°C while shaking. Following incubation, the plate was briefly spun down to collect evaporation before the lysate was transferred to PCR tubes. The tubes were spun down at 600 x g for 2 minutes to pellet the Chelex beads, and the supernatant was transferred to new PCR plates. The resulting crude lysate containing single-stranded gDNA was stored at -20°C if not used directly.

To obtain gDNA of higher quality, silica-based spin column technology was used. When extracting gDNA alone from cell cultures, the GenElute™ Mammalian Genomic DNA Miniprep Kit was utilized according to manufacturer's instructions. The cells were harvested from 24- or 12-well plates at 80-90% confluency by trypsin, pelleted by centrifugation for 5 minutes at 300 x g, and resuspended in 200 µl Resuspension solution. The cells were then lysed by addition of 20 µl Proteinase K and 200 µl Lysis Solution C before a 10-minute incubation at 70°C. After column purification, the gDNA was eluted in 100 µl Elution Solution. Samples of gDNA was stored at -20°C if not used directly.

The concentrations of gDNA were measured on the QIAxpert or Nanodrop

3.4.3. Co-extraction of RNA and gDNA

To simultaneously extract both RNA and gDNA from the same sample, the Norgen RNA/DNA/Protein extraction kit was utilized according to manufacturer's instructions (details on protein described in section 3.6.1). The cells were lysed by 300 µl Buffer SKP, incubated at room temperature for 5 minutes with gentle manual swirling, before storage at -80°C until RNA and gDNA extraction proceeded. The RNA was eluted in 50 µl Elution Buffer A whereas the gDNA was eluted in 50-100 µl Elution Buffer F. The concentrations of RNA and gDNA were measured on the QIAxpert or Nanodrop. RNA was stored at -80°C and gDNA was stored at -20°C if not used directly.

3.4.4. cDNA synthesis

In order to quantify mRNA, cDNA synthesis was performed using the Applied Biosystems High-Capacity cDNA Reverse Transcriptase kit on 100-350 ng RNA extracted the Norgen RNA/DNA/Protein extraction kit or the RNeasy QIAcube Kit. In each reaction, 1x (f.c.) Rxn buffer, 4 mM (f.c.) dNTP mix, 1x (f.c.) RT Random Primers and 2.5 U/µl (f.c.) MultiScribe™ Reverse Transcriptase was used, with the final volume being 20 µl. The Reaction was run on the thermocycler for one cycle of 10 minutes at 25°C for annealing, followed by 120 minutes at 37°C for elongation, and finally 5 minutes at 85°C for enzyme deactivation before cooling down to 4°C.

3.4.5. Real-time quantitative PCR (qPCR)

To measure mRNA expression, qPCR was performed on 1:5 diluted cDNA using 1x (f.c.) LightCycler® 480 SYBR Green I Master, 0.4 µM (f.c.) forward primer (table 2.9), 0.4 µM (f.c.) reverse primer, and 1.25 µl template per reaction, with a total volume of 10 µl. For gDNA, the reactions were run similarly, using either 11 ng purified (1.25 µl) template or 4 µl crude Chelex extract with unknown concentration per reaction.

The reactions were run on the LightCycler® 480 II with an amplification profile of one cycle of 5 minutes at 95°C for denaturation, 45 cycles of amplification with 10 seconds at 95°C for denaturation, 15 seconds at 60°C for annealing, and 10 seconds at 72°C for extension. Additionally melting curve analysis was performed by denaturation at 95°C for 5 seconds, followed by cooling at 65°C for 1 minute, and then gradually increasing the temperature up to 97°C while recording at 0.1°C increments. qPCR results were analyzed using the ΔΔCt method.

3.4.6. Polymerase chain reaction (PCR) amplification of gDNA

Polymerase chain reactions (PCRs) amplification of gDNA for downstream sequencing was performed using 10 µl Platinum™ SuperFi™ Green II PCR master mix, 0.5 µM each of forward and reverse primers (f.c.), and 100 ng gDNA template. Each reaction had a final volume of 20 µl and were run in the thermocycler for one cycle of 30 seconds at 98°C for initial denaturation, followed by 30 cycles of amplification with 10 seconds at 98°C for denaturation, 10 seconds at 60°C for annealing, and 30 seconds at 72°C for elongation, before a final elongation cycle of 5 minutes at 72°C and cooldown at 4°C.

3.4.7. Agarose gel electrophoresis

To verify clean amplification by PCR, agarose gel electrophoresis was run on PCR products using a 1.5% (w/v) agarose gel with 1:10 000 (v/v) GelRed and 1x TAE buffer. Each well was loaded with 5 µl Superfi Green PCR product (containing loading dye), and 2 µl GeneRuler 1kb plus DNA ladder. The gel was then run at 100 V for 1 hour. Pictures were taken using the GelDoc EZ Imager.

3.4.8. Enzymatic clean-up

To prepare PCR products for sequencing and remove unincorporated dNTPs and primers, the Illustra™ ExoProStar 1-step kit was used as per manufacturer's instructions with 5 µl PCR product mixed with 2 µl reagent. The reaction was then set to incubate at 37°C for 15 minutes followed by 80°C for another 15 minutes. Concentrations were measured using the Nanodrop.

3.4.9. Sequencing

Sanger sequencing of plasmids and PCR products were performed using the BigDye™ sequencing kit (v. 3.1). 1x (f.c) Sequencing Buffer, 10% (v/v) BigDye Terminator Ready Reaction Mix, and 0.32 µM (f.c.) primer

was added to 200 ng plasmid or 20 ng PCR product to a final volume of 10 µl. Each sample were run with two replicates for both forward and reverse primers. The samples were run in the thermocycler for one cycle of 30 seconds at 96°C for initial denaturation, followed by 35 cycles of amplification with 10 seconds at 96°C for denaturation, 5 seconds at 50°C for annealing, and 4 minutes at 60°C for elongation. The samples were then analyzed at the Genomics Core Facility at the department of Medical Genetics at Haukeland University Hospital. Sequencing results were analyzed using Benchling.

3.4.10. Primer design

Primers for PCR, qPCR, and sequencing were designed using the NCBI online tool Primer-BLAST and all primers were double checked for matches in the Human (hg38) BLAT search engine, as well as quality checked for primer dimers and additional secondary structure predictions by the OligoEvaluator™ online tool by Sigma-Aldrich. Primers for qPCR of cDNA were designed to achieve a PCR product of between 70-150 bp, spanning exon-exon junction (not applicable for primers for introduced genes like PEmax), and max self-complementarity adjusted from 8 to 3.

3.5. Methods of transient DNA delivery

3.5.1. Chemical- and lipid-based transfection methods

Two transfection reagents were initially tested in the HT1080 cells to determine which reagent resulted in higher transfection efficiency, short-term cell viability, and to check whether the lentiviral PEmax encoding plasmid expressed PEmax as expected. Which reagent would be utilized in later experiments was also assessed.

The reagents that were tested were the lipid-based Lipofectamine Stem Cell reagent and the chemical-based TransIT-LT1 reagent. For each reagent, 25 000 HT1080 cells were seeded per well in a 24-well plate to reach 50-60% confluency at the time of transfection. A 1:3 ratio of plasmid DNA (pLenti-PE2max-BSD and pCMV6-XL5-empty) and reagent was used, with a total of 400 ng DNA per reaction. Both the plasmid DNA and the reagents were suspended in Opti-MEM. After mixing DNA and reagent, they were incubated at room temperature for 20 minutes, before 50 µl of the transfection mix was added to each well. The cells were supplied with fresh medium prior to transfection. Post transfection, the cells were incubated for 48 hours before assessing the cell viability by visual assessment in the microscope and transfection efficiencies by mRNA expression in the cells as measured by qPCR. A follow-up experiment was conducted where the DNA concentration was adjusted to 250 ng for the cells transfected with Lipofectamine Stem Cell reagent, while maintaining the 1:3 ratio of DNA to reagent to reduce cell death.

In downstream experiments, TransIT-LT1 was selected as the primary transfection reagent for both transfecting the PE-system into the HT1080 cells and for virus production in HEK293T cells. A separate test

on different ratios of TransIT-LT1 reagent and DNA was run on the HT1080 cell line prior to the main genome editing using PE, this is described in detail in section 3.8.2.

3.5.2. Nucleofection

Based on transfection tests conducted by Krill (2023), which demonstrated that nucleoporation was the best method for delivering small plasmids to the ASC52telo cells, nucleofection was chosen as the method for transient delivery of plasmids encoding pegRNAs, pMax_GFP and the MMR inhibitor MLH1 in ASC52telo. A pre-optimized kit and protocol for mesenchymal stem cells (P1 Primary Cell 4D-Nucleofector X Kit) was used according to manufacturer's instructions. Briefly, the cells were harvested, resuspended in culture medium and counted. A total of 50 000 cells per reaction was transferred to new tubes, pelleted, resuspended in Nucleofector Solution with supplement (1:4.5 ratio) and immediately transferred to the nucleofection chamber. A total of 800 ng DNA was then added per 50 000 cells, corresponding to 300 ng of pEF1a-hMLH1NTD-NLS, 300 ng pegRNA-A or -C, and 200 ng pMax_GFP.

The cells were run on the P1 program with EW-104 pulse condition. Following the completion of the program, the cells were rested at room temperature for 10 minutes before mixed in pre-warmed medium and seeded at a density 100 000 cells per well in 6-well plates. See section 3.9.7 for more details.

3.6. Protein methods

3.6.1. Protein harvest

Protein was harvested from cells at 80-90% confluency in 6-well plates using either the Norgen kit (according to manufacturer's instructions, see section 3.4.3) eluted in 100 µl Elution Buffer C, or RIPA Western Blot (WB) lysis buffer.

The cells to be treated with RIPA WB lysis buffer were transferred to ice and washed with ice-cold PBS. 600 µl PBS was added to each well, and the cells were scraped with cell scraper in the PBS before the cell suspension was transferred to Eppendorf tubes. The cells were then pelleted by centrifugation at 2500 x g at 4°C for 3 minutes before resuspension in 100 µl RIPA WB lysis buffer. The samples were then vortexed and kept on ice for 45 minutes for complete lysis.

Both the Norgen and RIPA WB protein samples were stored at -80°C if not used directly.

3.6.2. Measuring protein concentrations

A Bovine Serum Albumin (BSA) standard curve was made to calculate protein concentrations. Sample lysates were thawed on ice and centrifuged at 4°C for 7 minutes at 13 000 x g to pellet debris. BSA was serial diluted

in either Norgen Elution Buffer C or in RIPA WB lysis buffer, starting with a 2 µg/µl stock concentration. A 2-fold serial dilution was made with 5 samples in total, as well as a negative control.

The DC™ Protein Assay reagents were added according to manufacturer's instruction along with 5 µl of BSA standard or sample to each well of a blank 96-well plate to perform a colorimetric assay of the protein concentration. The samples were then incubated at room temperature for 15 minutes while shaking. The absorbance at 750 nm of each well was measured using the SPECTRAMax® *PLUS*^{384} Microplate and a standard curve was then made based on the known BSA concentrations. Finally, the sample concentrations were calculated using the equation obtained from the resulting standard curve.

3.6.3. SDS-PAGE and Western Blot (WB)

SDS-PAGE was performed on the protein samples extracted using RIPA WB lysis buffer or the Norgen kit using SurePAGE™ 4-20% gels. The samples were prepared using 1x (f.c.) SDS Sample Buffer. 5 µl Kaleidoscope™ Prestained Protein Standard ladder was loaded in one well. 200 ng recombinant Cas9 was loaded as a positive control in one well. An equal amount of protein (10-20 µg depending on the experiment) was loaded in each well, and the gel was run in 1x Tris-MOPS-SDS Running Buffer at 200 V for 30 minutes. The proteins were transferred to a nitrocellulose membrane by the Trans-Blot Turbo Transfer System using 2.5 A and 25 V for up to 20 minutes, or until all large proteins on the ladder were observed to be transferred. The membrane was then blocked in 5% (w/v, f.c.) nonfat dried milk dissolved in 1x PBST for 1 hour at room temperature while shaking, before being thoroughly washed in PBST.

Prior to antibody staining, the transfer of proteins was confirmed by Ponceau staining. A Ponceau solution was prepared at a concentration of 0.1% (w/v, f.c.) dissolved in 5% (v/v, f.c.) pure acetic acid. The membranes were left in the stain for approximately 5 minutes before being thoroughly washed in PBST.

Primary antibodies included were anti-Cas9 and anti-beta-actin. They were diluted 1:1000 and 1:5000 (respectively) in 3% BSA (w/v, f.c.) in PBST. The membranes were incubated in the primary antibody solution over night at 4°C while shaking. Before the secondary antibody solution was added the membrane was thoroughly washed in PBST. HRP Goat Anti-Mouse secondary antibody was diluted 1:7500 in 3% BSA in PBST as well and incubated on the membrane for 30 minutes at room temperature while shaking. The membrane was then once again washed with PBST prior to the addition of detection solution. The detection solution was made by mixing a 1:1:1 solution of SuperSignal West Femto Luminol, SuperSignal West Femto Peroxide, and MilliQ H_2O, and before it was left to incubate on the membrane for 2 minutes in room temperature. The resulting light produced by the HRP-conjugates was detected by the Gel Doc XR System.

After detection, the membrane was stripped with Restore Western Blot Stripping buffer for 8 minutes at room temperature while shaking. The primary antibodies were stored for later use at 4°C in 0.02% (w/v) NaN_3.

3.7. Fluorescence Activated Cell Sorting (FACS)

To select for successfully transfected/nucleofected cells, a GFP fluorescent marker was included in the transfections, either on the same plasmid as the gene of interest (pCMV-PEmax-P2A-GFP), or co-transfected on a separate plasmid (pMax_GFP). FACS was performed using the SONY SH800 Cell Sorter according to manufacturer's instructions. Prior to sorting, the cells were seeded in 6-well plates after/during treatment and sorted no more than two days after treatment to ensure high GFP expression. On the day of sorting, the cells were washed with PBS, detached from the plate using Accutase or trypsin-EDTA (depending on cell line) and spun down at 300 x g for 3 minutes, before resuspension in 500 µl FACS buffer. The cells were filtered through a 100 µm cell strainer and stored on ice prior to sorting. To excite GFP-positive cells, the 488 nm laser was used, and emitted fluorescence was detected using the FL1 (525/50) bandpass filter. Negative and positive controls were run through prior to sorting to set the florescent gates (Figure 3.2). After validating the presence of cells in each sample, the flow rate was reduced to a minimum (setting 1 or 2) before sorting. The cells were sorted using the Ultra-Purity mode either as single cells in 96-well plates or as a pooled cell population in tubes containing cell culture medium and immediately transferred to 37ºC in a CO_2 incubator. Data was analyzed with FlowJo v.10.10.0.

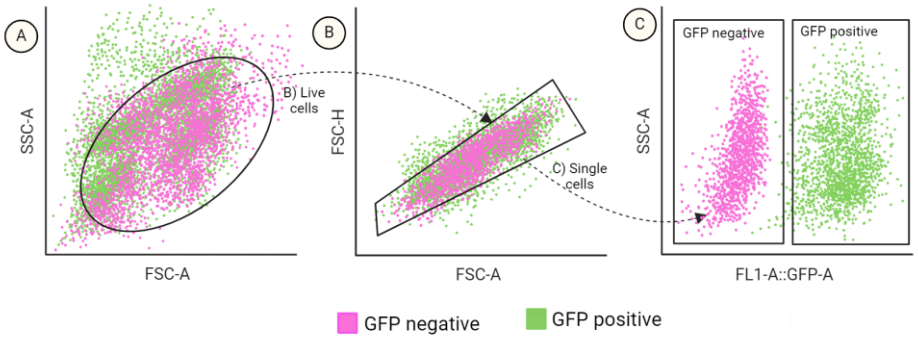


Figure 3.2: The following gates were used to sort GFP positive cells: **(A)** The first gate (live cells) was set on forward scatter area (FSC-A) against side scatter area (SSC-A) to get an estimation of the main population of live cells, assess granularity and to filter out debris. **(B)** The live cell population was shown in a new plot as FSC-A against forward scatter height (FSC-H) and the second gate (single cells) was set diagonally to exclude doublets. **(C)** The single cells were shown in a new plot as FL1-A (GFP) against SSC-A, and GFP negative and GFP positive gates were set according to positive and negative fluorescent controls. Figure created with BioRender.com.

3.8. Proof-of-concept prime editing in the HT1080 cell line

To validate that the PE system works, and verify the functionality of the specifically designed pegRNAs, a proof-of-concept experiment was conducted in a cell line that is easy to transfect and fast-growing, specifically HT1080. The cells were transfected with the components corresponding to the PE4 system (PE2 including MMR inhibitor MLH1) (Figure 3.3 step 1) , sorted by FACS (Figure 3.3 step 2), and single-cell colonies were then expanded and genotyped (Figure 3.3 steps 3-6). Single-cell colony expansion is necessary to achieve homozygous cell lines for the risk and protective alleles.

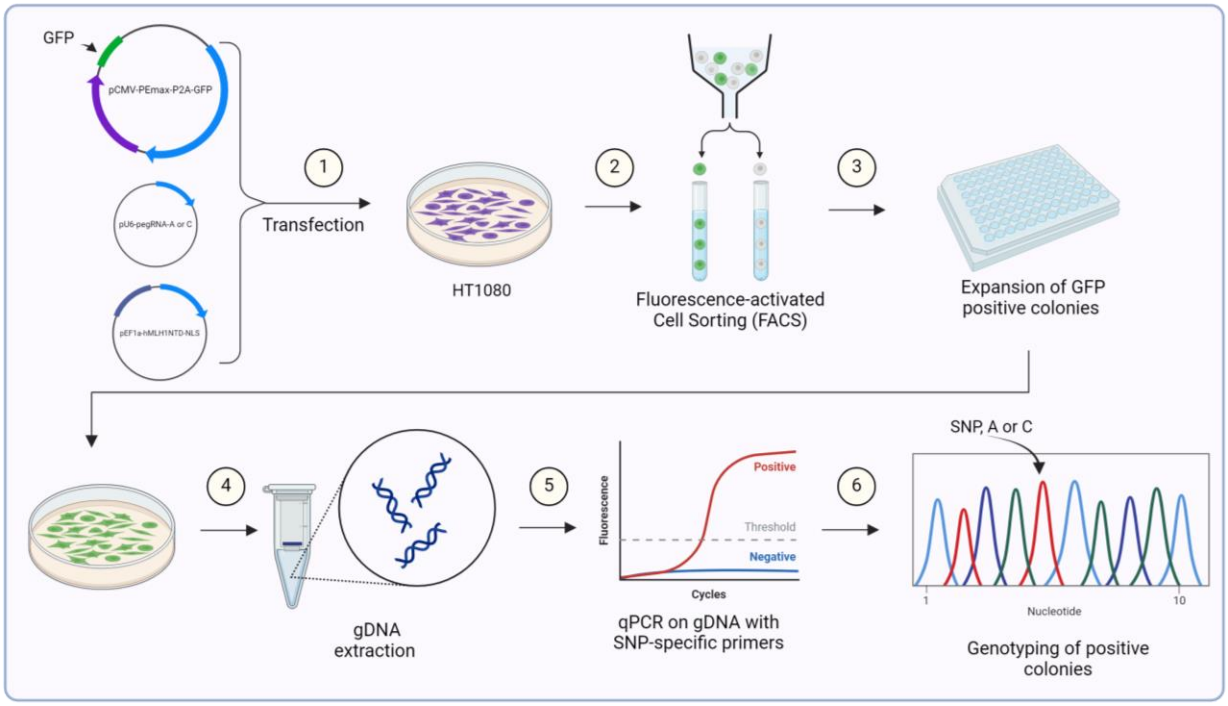


Figure 3.3: Overview of the workflow for the proof-of-concept prime editing (PE) in the HT1080 cell line. (1) To achieve an edit of the single nucleotide polymorphism (SNP) rs1799993, the PE system was introduced into the cell line by transfection. **(2)** The PEmax-plasmid contains a GFP sequence, meaning that the successfully transfected cells were selected by FACS. **(3)** Single-cell colony expansion of the GFP-positive cells was then necessary to achieve homozygous cell lines for the SNP. **(4)** To screen the cells for a successful edit, genomic DNA from the expanded colonies was extracted and analyzed by **(5)** qPCR with SNP-specific primers. **(6)** Validation of the edit was then performed by sequencing on genomic DNA. Figure created with BioRender.com.

3.8.1. Genotyping of HT1080 cells

To be able to edit the HT1080 cells, knowing which alleles of the SNP rs1799993 are present is crucial. 20 000 HT1080 cells were grown in each well of a 12-well plate until > 80% confluent. The cells were then lysed and gDNA was isolated using the GenElute Mammalian Genomic DNA Miniprep kit as described in section 3.4.2. The region surrounding rs1799993 was amplified by PCR using 100 ng gDNA per reaction and the Platinum™ SuperFi™ II PCR kit as described in section 3.4.6, using the 643F_HMBS and 644R_HMBS primers before the size of the product was confirmed by agarose gel electrophoresis (section 3.4.7). The PCR product was cleaned up using the Illustra™ ExoProStar 1-step kit as described in section 3.4.8, before sequencing by BigDye as described in section 3.4.9 using the same primers as for PCR. Sequencing results were analyzed in Benchling by the MAFFT auto alignment function to align the sequencing results with the reference sequence.

3.8.2. Optimizing transfection of the PE-plasmid in HT1080 cells

Using the TransIT-LT1 transfection reagent, a transfection experiment was conducted using the pCMV-PEmax-P2A-GFP and pMax_GFP plasmids to determine the optimal concentrations of plasmid of interest and reagent for highest number of successfully transfected cells. As the PEmax plasmid is larger than the pMax_GFP plasmid, optimalization of transfection was necessary.

Initially, 3 200 HT1080 cells were seeded into each well of a 96-well plate 24 hours pre transfection. The ratios of DNA to reagent tested were 1:1, 1:2, 1:3, and 1:4, with varying amounts of DNA (0 ng, 25 ng, 50 ng, and 100 ng) for each ratio, with three parallels in total for each combination. Each plasmid was transfected individually, with pMax_GFP acting as a positive control. The transfection was otherwise performed as described in section 3.5.1. GFP expression was monitored using the IncuCyte S3 Live-Cell Analysis System at 0, 24, 48, 72, and 96 hours post transfection. Five pictures were taken of each well at each time point at 10 x magnification with detection of GFP. After the monitoring was completed, an analysis was run to identify the number of GFP positive objects in each well.

3.8.3. Transient transfection of the complete PE system in HT1080 cells

The best combination of plasmid and transfection reagent volumes, as identified in the 96-well high throughput screening, was used as a starting point to scale the reaction up to 6-well plates. Accordingly, a DNA to reagent ratio of 1:4 and 2.5 µg plasmid DNA were selected. One day before transfection, 96 000 HT1080 cells were seeded into each well of a 6-well plate and incubated for 24 hours to reach (50-60% confluency the next day). Each reaction included 1.5 µg of pCMV-PEmax-P2A-GFP, 0.5 µg of pegRNA-A or -C, and 0.5 µg of either pEF1a-hMLH1ntd-NLS (PE4) or pCMV6-XL5 (PE2). A positive control of transfection included 1.5 µg pMax_GFP with 1 µg pCMV6-XL5 and a negative PE control included 1.5 µg pCMV-PEmax-P2A-GFP and 1 µg pCMV6-XL5 (no pegRNA).

GFP expression was assessed using the Nikon Eclipse Ti2 24 hours post transfection. 48 hours post transfection, the GFP positive cells were sorted by FACS as described in section 3.73.7. Cells from each reaction were both sorted as single cells in one 96-well plate and as a pool for expansion on larger cell culture plates to form colonies.

3.8.4. Expansion of single-cell colonies from FACS-sorted HT1080 cells

The cells that were sorted as single cells in 96-well plates were expanded until 80% confluency, before being split 1:3 and copied into three new 96-well plates. The cultures that were pooled were kept as carry-on cultures in 6-well plates and 10 cm plates, before between 1000 and 2000 cells were seeded in 15 cm plates and monitored every day for a up to two weeks for colony formation. The colonies that formed were gently

removed by cell scraper or pipette tip and transferred to 96-well plates by pipette. Colonies that survived were later expanded, split 1:3 and copied into three 96-well plates once they reached 70% confluency. Backups of the colonies were frozen using the freezing cells technique described in section 3.2.4, except the cells were suspended in 10% DMSO directly in the wells. They were then wrapped in a thick layer of paper towels and aluminum foil, before storage at -80^{o}C.

3.8.5. Screening of edited HT1080 colonies using qPCR with SNP-specific primers

Successfully expanded colonies were analyzed by qPCR and SNP-specific primers to indicate whether the edit had been successful. The SNP-specific primers were tested on 6 ng gDNA from previously genotyped homozygous patient samples, and their efficiency was estimated based on gene copy number (GCN) described in more detail in section 3.9.6. To screen the colonies using these primers, they were expanded to 24-well plates before gDNA was harvested by Chelex 100 as described in section 3.4.2. A total of 4 µl crude Chelex lysate diluted 1:5 was then used in the qPCR reaction, and the Ct-values were compared between the A and C specific primers to give an indication of homozygosity or heterozygosity. Antilog values of the difference between the Ct values ($2^{-\Delta Ct}$) represented A and scaled to represent C by negative one over the antilog value ($-1/2^{-\Delta Ct}$).

3.8.6. Genotype validation of edited HT1080 colonies

The colonies that were validated by qPCR and SNP-specific primers and deemed possibly homozygous for the SNP were expanded and sequenced as described in section 3.8.1. The colonies that were selected for sequencing (including unedited controls) were thawed by addition of pre-warmed medium directly in the wells of the 96-well plate, before transferal directly to 1 ml medium in 24-well plates to dilute the DMSO. The colonies were then expanded to 12-well plates, before gDNA was isolated by the GenElute Mammalian Genomic DNA Miniprep kit as described in section 3.4.2.

The area surrounding the SNP was subsequently amplified by PCR (section 3.4.6), and the size of the product was validated by agarose gel electrophoresis (section 3.4.7). Once validated, the product was cleaned up by the Illustra^{TM} ExoProStar 1-step kit and sequenced by BigDye as described in section 3.4.9. The sequencing results were analyzed in Benchling.

3.9. Prime editing in ASC52telo cells using lentiviral delivery

Having tested the PE concept in HT1080 cells (section 3.8), editing of the target cell line ASC52telo was attempted. While nucleoporation is suitable to transfect this cell line with small plasmids, this method is not able to deliver the large plasmid encoding the PE fusion protein and GFP (~10.6 Kb) (Krill, 2023). Instead, lentiviral transduction to stably integrate the PE construct into the genome was tested as an alternative approach (Figure 3.4 steps 1-2). Once a cell line of ASC52telo that stably expresses the PEmax construct has been

constructed, the small plasmids encoding pegRNAs can be transiently introduced by nucleofection, and single-cell colonies can be expanded and subsequently screened for successful edit by qPCR (Figure 3.4 steps 6-9). Once edited, the phenotype of cells homozygous for the risk and protective allele at rs1799993 (A and C, respectively can be analyzed by qPCR, WB, and other methods both before, during, and after differentiation into mature adipocytes to uncover biological processes affected by the alleles. The ASC52telo cells were attempted edited by the PE4 system (PE2 with MLH1).

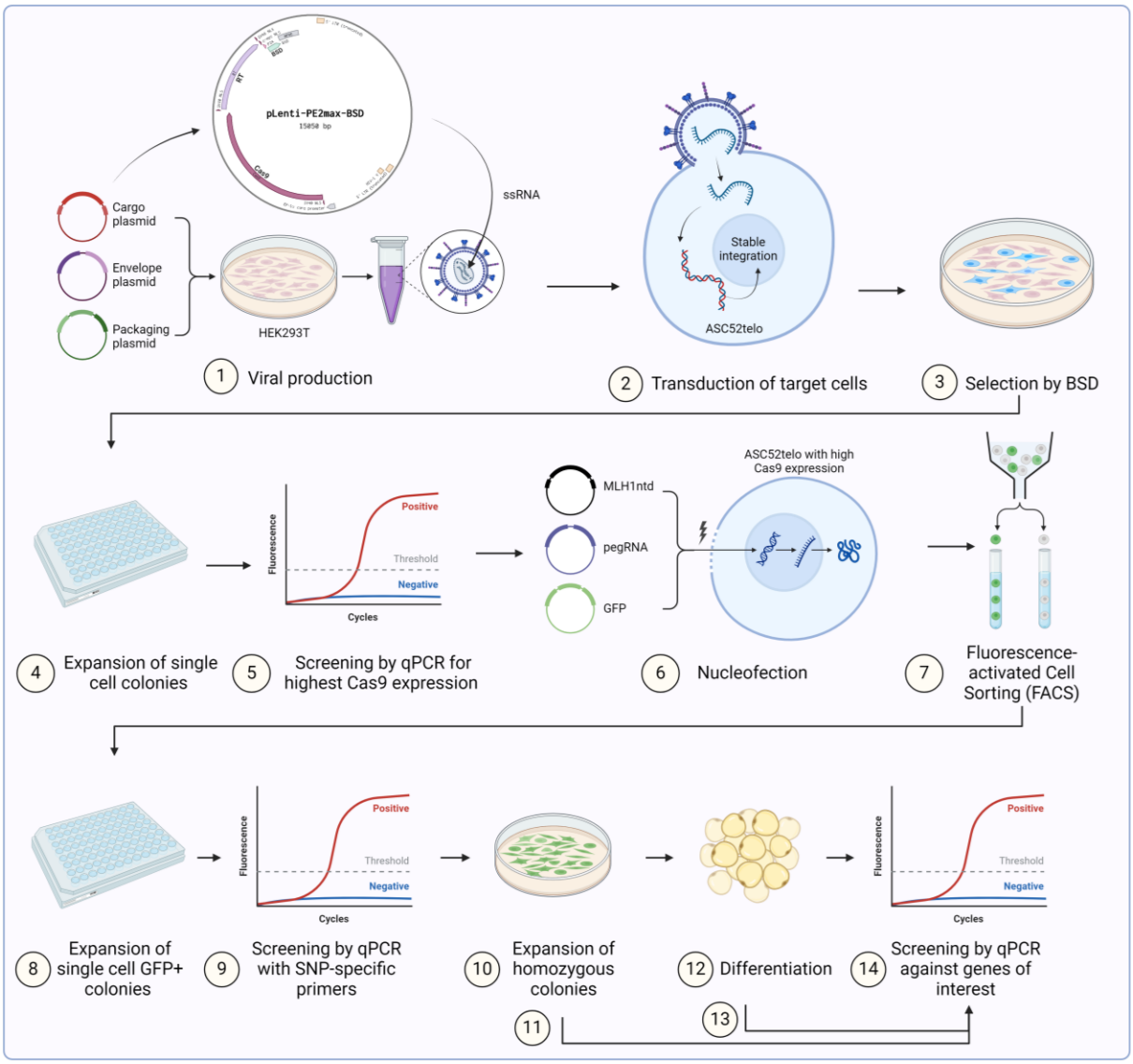


Figure 3.4: Workflow of editing ASC52telo by lentiviral transduction of the PE-system. (1) As the ASC52telo cell line is hard to transfect, the prime editing (PE) plasmid was packaged as single stranded RNA in lentivirus by HEK293T cells. The plasmid carries the PEmax enzyme and resistance to blasticidin (BSD). **(2)** ASC52telo was transduced by the virus, and **(3)** selection of successfully transduced cells was done by BSD. **(4)** Expanding single-cell colonies and **(5)** screening by qPCR was then necessary to single out cells with the highest expression of PEmax. **(6)** Nucleofection to introduce the remaining plasmids of the system, pegRNAs encoding either the A or C allele, a mismatch repair inhibitor

(MLH1), as well as a GFP-encoding plasmid to screen transfected cell, before **(7)** screening by FACS for GFP positive cells. **(8)** Expanding GFP positive cells as single-cell colonies is necessary to achieve homozygous cell lines, **(9)** which can be screened by qPCR with SNP-specific primers and by Sanger sequencing. **(10)** Expansion of homozygous colonies with the genotype of A or C. **(11)** Screening by qPCR against genes of interest before differentiation. **(12)** Differentiation of edited colonies. **(13)** Screening by qPCR against genes of interest during differentiation. **(14)** Screening by qPCR against genes of interest after differentiation. Figure created with BioRender.com.

3.9.1. Viral production in HEK293T cells

To produce lentivirus that incorporates PEmax and resistance to BSD as cargo, plasmids that carry specific lentiviral components can be transfected into a production cell line (most commonly HEK293T cells) (Merten et al., 2016). The cargo plasmid (pLenti-PE2max-BSD) encodes the cargo-genes of interest flanked by two LTR-sequences, and the psi packaging signal. The area between the LTR's become packaged into the virus as ssRNA. The packaging plasmid (pCMVR8.74) encodes the key lentiviral genes GAG, POL, TAT, and REV, and the envelope plasmid (VSV-G) encodes the envelope protein ENV. By splitting the viral components across multiple plasmids and only incorporating specific genes, the system ensures no uncontrolled viral replication. This results in the production of virus encoding the gene of interest that is released into the medium of the production cell line. The virus can then be harvested and applied to the target cell line.

Prior to transfection, the HEK293T cells were expanded in 15 cm dishes to reach 85-90% confluence. The cells were cultured in PEST-free medium to improve transfection efficiency. TransIT-LI1 (described in section 3.5.1) was chosen as the transfection reagent for the HEK293T cells. A total of 22.5 µg cargo plasmid (pLenti-PE2max-BSD), 14.6 µg packaging plasmid (pCMVR8.74), and 7.9 µg envelope plasmid (VSV-G) were mixed with 3 ml Opti-MEM for a total of 45 ug DNA and a DNA:TransIT-LT1 ratio of 1:2 pr 15 cm dish. Each transfection mix was scaled up to accommodate 3 x 15 cm dishes. The transfection mix was incubated for 20 minutes, and 3 ml was added dropwise to each dish. The cells were then incubated at 37°C and 5% CO_2 over-night. The ratio between the plasmids was based on a protocol developed and optimized by the Trono laboratory of Virology and Genetics (https://rb.gy/pzi0qe).

After the 16-hour incubation period, the medium was replaced with fresh medium (still without antibiotics). The first viral harvest was conducted 8 hours after the last medium change (24 hours post transfection) by collecting the medium and replacing it with fresh medium. Additional viral harvests were performed at 36 and 44 hours post transfection. The collected virus-containing medium was centrifuged at 500 x g for 10 minutes to pellet any cell debris and the supernatant was filtered using a 0.45 µm filter. The clarified supernatant was then stored at 4°C until all three harvests had been completed.

To concentrate the harvests the Lenti-X™ Concentrator was used according to manufacturer's instructions, with 1 volume concentrator added per 3 volumes of clarified lentiviral supernatant before a 30-minute incubation at 4°C. The mixture was then centrifuged at 1 500 x g for 45 minutes at 4°C. After removing the supernatant, the pellet was resuspended in Opti-MEM corresponding to 1:100 of the original volume of the harvests. The harvests were then aliquoted and stored at -80°C until transduction.

The harvested samples were quantified using Lenti-X™ GoStix™ before concentration, after concentration, and after a freeze-thaw cycle, following the manufacturer's instructions. The GoStix™ measures the amount of viral capsid protein (p24) in GV (GoStix Value) corresponding to ng/ml p24. This provides an estimation of viral particle concentration, as there is an approximate 1 x 10^7 physical viral particles (PP) in each ng of p24 protein measured (Sigma Aldrich - Merck, n.d.).

3.9.2. Determination of blasticidin dose on ASC52telo

As the lentiviral cargo plasmid (pLenti-PE2max-BSD) encodes resistance to blasticidin (BSD) to be used as selection of successfully transduced cells, it was necessary to find the concentration of BSD in which any non-transduced cells would not survive. To find the optimal concentration of BSD for a 7-14-day selection period, an antibiotics titration test was conducted. ASC52telo cells were seeded out to a density of 5000 cells/cm^2 and exposed to five different concentrations of BSD (1, 2.5, 5, 10, and 20 µg/ml), alongside a negative control (kill well). The medium was replaced after 3 days with a fresh dose of BSD. After 7 days, cell images were captured the cells using the Sartorius IncuCyte S3, and an analysis was run based on the confluency of each well. The concentration of BSD that resulted in the highest cell death was chosen for selection of the transduced cells.

3.9.3. Determination of functional viral titer

Physical viral particles do not necessarily correspond to viral particles capable of infection (infectious units; IFUs). Therefore, an estimate of IFU needed to be decided experimentally. Using the results given by the GoStix™, a titer assay was conducted on ASC52telo to estimate the number of IFUs in the concentrated harvests. To enhance transduction efficiency, RetroNectin was used to coat the wells of a 96-well plate before cell seeding. RetroNectin was applied at a concentration of 5 µg/cm^2 and incubated for 2 hours at room temperature. Subsequently, the coating solution was removed, and the wells were blocked with a 2% BSA solution (diluted in PBS) for 30 minutes. A control experiment was also conducted without the RetroNectin coating.

The ASC52telo cells were seeded in the RetroNectin-coated 96-well plates at a density of 4000 cells/well. Antibiotics-free medium was used for this step. The day after seeding, a 5-fold serial dilution of concentrated virus harvest consisting of 5 points (spanning 1:1 to 1:625) was prepared, before the dilutions were added to

the cells. A negative control not receiving virus was also included (kill well). The cells were then set to incubate for 6 hours at 37°C and 5% CO_2 before a medium change was performed, this time containing antibiotics.

After the cells had gone through a recovery-period of approximately 48 hours, the medium was changed to contain 20 µg/ml BSD, which corresponds to the results from the antibiotics titer test performed on the same cell line prior. The medium was changed every 3-4 days, each time resupplying with the same concentration of BSD. After 7 and 11 days of BSD-selection, images were taken using the Nikon Eclipse TS100 Microscope. The number of colonizing-forming units (CFU) was counted on day 11, and the corresponding IFU/ml was calculated based on dilution factor and GV. The number of CFU times the dilution factor of the titer test corresponds to the number of units capable of infection (IFU) in the viral sample. The number of IFU/PP could then be calculated by IFU/ml and the estimated PP extrapolated from the GV measurement. This provided the foundation for testing different multiplicities of infection (MOIs), representing the IFU-to-cell ratio, for the transduction.

3.9.4. Lentiviral transduction of ASC52telo

Prior to transduction, the wells of a 96-well plate were coated with 5 µg/cm^2 RetroNectin as previously described, before 2000 ASC52telo cells were seeded in each well. No antibiotics were used in this step. The cells were transduced the morning after seeding and based on the results from the functional titer test, MOIs of 0.05, 0.1, 1, 2, and 5 were chosen to infect the cells. A kill well was also included. After a 6-hour incubation period at 37°C and 5% CO_2, a medium change was performed, this time to include antibiotics. Two days post transduction the cells were moved to larger plates and wells, and BSD were included at 20 µg/ml to initiate the selection process.

3.9.5. Optimizing selection and expansion of single-cell colonies

The goal of the selection process was to isolate successfully transduced single cells that could give rise to colonies, and subsequently expand and screen them for expression of PEmax. Transduced cells were expanded and seeded sparsely in 6-well plates or 10 cm dishes under 20 µg/ml BSD selection for 14 days to ensure non-transduced cells were dead and to promote colony formation while maintaining cell-to-cell communication. Areas growing too dense were re-plated to ensure that colonies would arise from a single cell. The BSD concentration was reduced to 10 µg/ml (maintenance concentration) for three days, before the BSD was completely removed to promote the growth of surviving single colonies. However, 10 µg/ml BSD was later re-introduced into the medium to prevent growth of cells without the construct.

A second transduction was performed where two days post transduction the cells were loosened using Accutase, and half of the cells were moved to 10 cm dishes, while the other half were moved to 24-well plates

to be expanded for preliminary RNA and gDNA extraction. The cells were exposed to 20 µg/ml BSD for 15 days in total, before reduction to 10 µg/ml. These cells were not taken off BSD at all.

Attempts at achieving single-cell colonies were also done by seeding the cells at a density too high to form single-cell colonies, before isolating single cells by scraping away surrounding cells. Cells that surrounded single-cell colonies were scraped away approximately every other day, or when needed. Once the colonies had expanded to > 20 cells, they were gently removed by cell scraper and pipette to 96-well plates or 24-well plates.

FACS was also tested (section 3.7) to seed a pool of transduced ASC52telo cells as single cells in 96-well plates. To promote the growth of single cells, FGF (1 ng/ml, v/v, f.c.) and EGF (10 ng/ml, v/v, f.c.) was added to the growth medium. Conditioned medium from culture dishes with actively dividing WT ASC52telo cells was also added to single cells in an attempt to stimulate colony growth.

Some cells formed colonies without being monitored. These ambiguous and heterogeneous colonies were expanded and cryopreserved (section 3.2.4) for analysis and later attempts at achieving true single-cell colonies.

3.9.6. Calculation of gene copy number (GCN) in transduced cells

In order to analyze the expression of PEmax and identify the number of constructs taken up by the cells post transduction, RNA and gDNA were extracted from the heterogeneous colonies of transduced ASC52telo cells. The cells were grown in 12-well plates until 80-90% confluent, and RNA and gDNA was extracted using the Norgen RNA/DNA/Protein extraction kit (section 3.4.3). gDNA was analyzed by qPCR as described in section 3.4.5. Calculation of gene copy number (GCN) of PEmax was calculated based on the equation described by Doleschall et al. (2022), where the relationship between Ct-value and GCN is described by:

Eq. 1
$$C_t(target\ gene) - C_t(reference\ gene) = -\left(\frac{log_2(GCN)}{log_2(2)}\right) - 1$$

Which, solved for GCN, becomes

Eq. 2
$$GCN = 2^{-C_t(target\ gene)+C_t(reference\ gene)+1}.$$

3.9.7. Delivery of guide RNA plasmids and sorting of GFP positive cells

ASC52telo has previously been genotyped and found to be heterozygous for rs1799993 (Krill, 2023). It was therefore necessary to perform editing in both directions to achieve homozygous cells for both alleles. Nucleofection was chosen as method of delivery of pegRNAs in ASC52telo based on previous transfection tests (Krill, 2023). While the selection and expansion of single-cell colonies was still in process, a heterogeneous mix of MOI 2 was expanded for a preliminary nucleofection, to see whether or not the edit was possible in the PEmax-transduced cells. DNA was added to the cells in the following amounts per reaction for both A and C: 300 ng pEF1a-hMLH1NTD-NLS, 300 ng pU6-pegRNA-A or -C, and 200 ng pMax_GFP for a total of 800 ng DNA per 50 000 cells. Three controls were included, one control of an equal amount of cells that were not nucleoporated, one control of nucleoporated cells without DNA, and one control of nucleoporated cells without pegRNAs. Each reaction was run in duplicates, amounting to 100 000 cells per reaction which were seeded together in one well of a 6-well plate. Following nucleofection, the plates were placed in the incubator overnight before fluorescence was checked the next day using the Nikon TE2000 Eclipse.

Two days post nucleofection, the cells were sorted by FACS as single cells as described in section 3.7. The cells were seeded directly into 96-well plates (1 cell/well) containing medium with added FGF (1 ng/ml, v/v, f.c.) and EGF (10 ng/ml, v/v, f.c.).

4. Results

4.1. Sequencing of plasmids and transient overexpression

4.1.1. Sequencing of plasmids

The majority of the plasmids utilized in the project arrived in *E. coli* agar stabs and underwent subsequent amplification and purification. Sanger sequencing was therefore applied to the samples to verify that the plasmids contained the expected sequences. Each resulting sequence was analyzed in Benchling and aligned with the reference sequence (provided by the manufacturer) by using the MAFFT Auto multi-alignment algorithm. The chromatograms were of high quality with clean peaks and a mean coverage of 7-15%, showing that all the plasmids corresponded to the reference sequence (Figure 4.1 and Table 4.1).

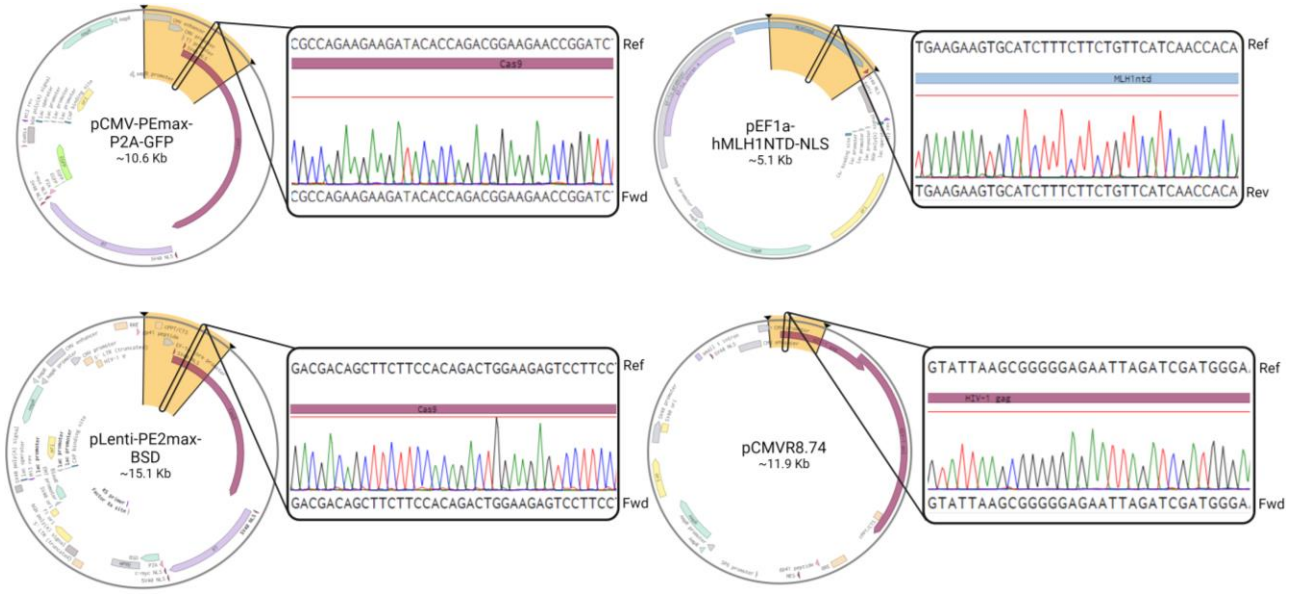


Figure 4.1: Mean coverage of Sanger sequencing performed on plasmids purified from E. coli. Each plasmid map showcases the area that the sequencing covered, marked in orange. A snippet of the chromatograms is showcased for each sequencing aligned with the reference sequence obtained from the manufacturer. The sequencing was performed using Sanger sequencing on 200 ng plasmids purified from *E. coli* and eluted in TE buffer.

Table 4.1 | MAFFT Auto alignment between Sanger sequencing results and plasmid reference sequence

Plasmid	Mean pairwise identity (%)	Mean coverage (%)
pCMV-PEmax-P2A-GFP	98.79	15.4
pLenti-PE2max-BSD	99.63	10.8
pEF1a-hMLH1NTD-NLS	99.47	16.3
pCMVR8.74	99.83	7.1

4.1.2. Transient overexpression of PE plasmid in HT1080

To confirm that the lentiviral PE plasmid was functional, a transient overexpression of the plasmid was performed in the HT1080 cell line using TransIT-LT1 and Lipofectamine Stem Cell transfection reagents with 0.4 µg DNA pr reaction for 25 000 cells in a 24-well plate. Expression of the PEmax enzyme was analyzed by qPCR 48 hours after transfection as described in section 3.4.5. The results were analyzed using the ΔΔCt method to calculate expression of PEmax relative to the reference gene *RPLP0*. The purpose was to assess whether the PE plasmids were functional and capable of high expression of PEmax when analyzed by qPCR. Consequently, the cells were lysed shortly two days after transfection, and long-term toxicity of the treatment was therefore not evaluated.

The cells transfected with Lipofectamine Stem Cell reagent exhibited a higher amount of cell death in response to the treatment than the cells transfected with TransIT-LT1 (data not shown). A second experiment was therefore run with the amount of DNA adjusted to 0.25 µg DNA for Lipofectamine Stem Cell, while maintaining 0.40 µg DNA for TransIT-LT1. Both treatments resulted in a strong overexpression of the PE enzyme (Figure 4.2), but cell viability was still higher with the TransIT-LT1 treatment (data no shown).

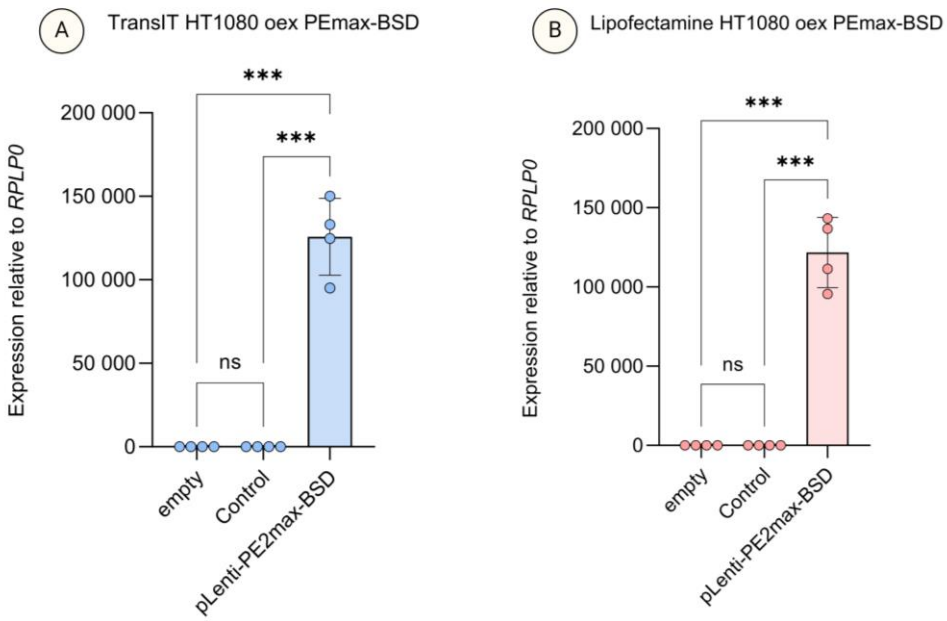


Figure 4.2: Transient overexpression of PEmax from the pLenti-PE2max-BSD plasmid in HT1080 cells. Two transfection reagents were tested, **(A)** TransIT-LT1, and **(B)** Lipofectamine Stem Cell. The ratio of DNA to reagent was 1:3, with 0.4 µg and 0.25 µg plasmid DNA (respectively), for 25 000 cells in a 24-well plate. The expression was analyzed by qPCR on the Roche LightCycler® 480 II and calculated using ΔΔCt with *RPLP0* as reference gene. The y-axis denotes expression of PEmax relative to *RPLP0*, and the x-axis denotes the experimental conditions with n = 4 biological replicates. The data are shown as mean ± SD. ***p_{adj}, < 0.001; ns, non-significant. The results were analyzed in GraphPad Prism using the ordinary one-way analysis of variance (ANOVA) along with the Holm-Šídák's multiple comparisons test. Figure created in GraphPad Prism.

4.2. Proof-of-concept prime editing in the HT1080 cell line

4.2.1. Genotyping of HT1080 cells

To be able to edit the SNP rs1799993 in the HT1080 cell line, the genotype needed to be known. In order to sequence the area around the SNP, gDNA was extracted (section 3.4.2) and the region was amplified by PCR using primers flanking the SNP (primers for *HMBS*, Table 2.1). The PCR product was analyzed using agarose gel electrophoresis and a strong band of expected size of about 900 bp was observed (Figure 4.3). A few, very weak unspecified bands could also be observed at ~1200 bp and ~2500 bp but were considered too weak to interfere with sequencing.

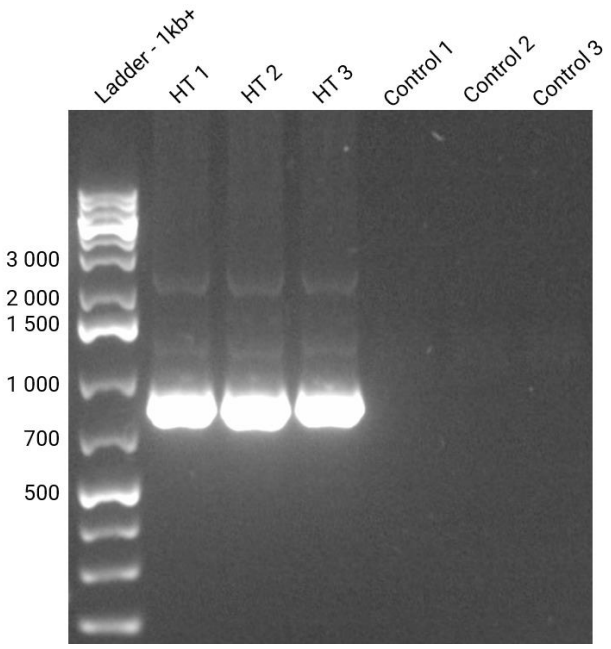


Figure 4.3: Validation of PCR product size and purity. PCR product of the area surrounding the rs1799993 SNP in the HT1080 cell line run on a 1.5% agarose gel. The first well shows the 1kb+ ladder, while HT1, HT2, and HT3 corresponds to the three parallel samples of the experiment. Control 1, no template; control 2, no primers; control 3, neither template nor primers. The expected size of the product is 890 bp, which corresponds to the location of the major band on the gel.

The PCR product was cleaned using the Illustra^{TM} ExoProStar 1-step kit and sequenced by Big Dye Sanger sequencing. The sequencing revealed that the HT1080 cell line is heterozygous for the SNP, containing both the protective allele (C) and the risk allele (A) (Figure 4.4). The sequencing was only successful on the reverse primer, but seeing as the number of parallels performed was high and the results on the reverse primers were deemed to be of high quality, the results were assessed to be reliable. A MAFFT multiple-sequence alignment in Benchling of the reference sequence and the sequencing results from the reverse primer gave a pairwise identity of 99.48%, and a coverage of 61.8% in the area between the forward and reverse primers. Heterozygosity for the SNP meant that the edit needed to be performed in both directions to create homozygous risk and protective cell lines.

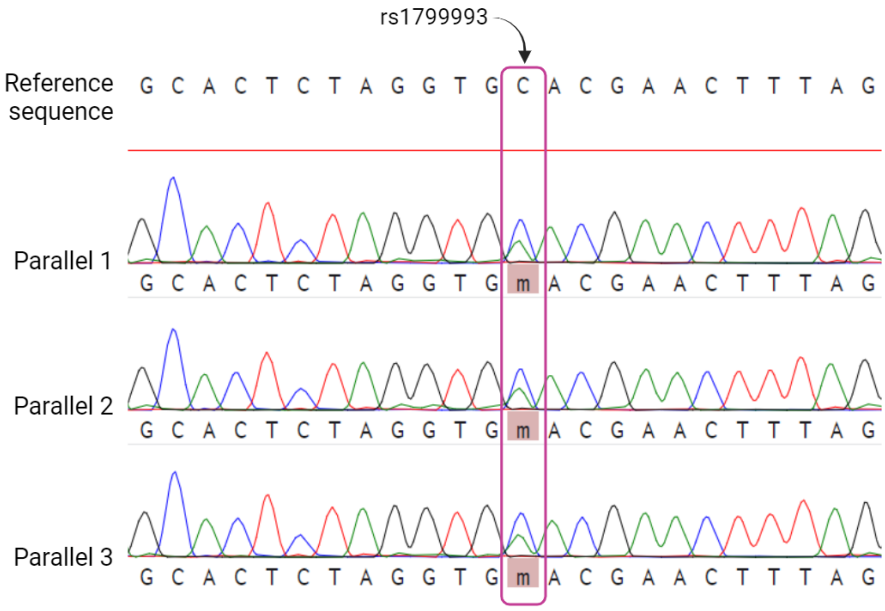


Figure 4.4: Genotyping the SNP rs1799993 in HT1080. Sanger sequencing using BigDye on genomic DNA revealed that the HT1080 cell line is heterozygous for the SNP rs1799993, containing both the protective allele (C, blue peak), and the risk allele (A, green peak), denoted as "m" in the sequences. The reference sequence shows only the protective allele, C. The arrow points to the SNP in all three parallels as well as in the reference sequence. Alignment and results were analyzed in the online software Bencling.com using the MAFFT auto multiple-alignment algorithm. Figure created with BioRender.com.

4.2.2. Optimizing transfection of the PE plasmid in HT1080 cells

In order to deliver the PE plasmids successfully into the HT1080 cell line, a pilot transfection was conducted using TransIT-LT1, pCMV-PEmax-P2A-GFP, and pMax_GFP to analyze the ratios of reagent and DNA that were optimal for the transfection. 3 200 cells/well in a 96-well plate were transfected with the following DNA:reagent ratios of 1:1, 1:2, 1:3, and 1:4 in combination with 0, 25, 50 and 100 ng of plasmid DNA. The cells were monitored for GFP expression in the IncuCyte at 0, 24, 48, 72, and 96 hours post transfection.

Based on the images taken by the IncuCyte, an analysis was conducted to determine the number of cells emitting green fluorescence at the different time points. The analysis showed the highest number of green objects detected at 72 hours post transfection for pMax_GFP (Figure 4.5 A), and at 92 hours post transfection for pCMV-PEmax-P2A-GFP (Figure 4.5 B). The highest amount of both DNA and transfection reagent (100 ng and a 1:4 ratio, respectively) resulted in the strongest expression of both plasmids. However, the larger PEmax plasmid resulted in fewer GFP positive cells than the small pMax_GFP plasmid (Figure 4.5 and Figure 4.6). The analysis did not consider the change in morphology that can be observed at 72 and 96 hours, which may indicate reduced cell viability beyond 48 hours (Figure 4.6). The change was consistent among all experimental parallels, including the untreated control, indicating that it may not be a direct effect of the transfection.

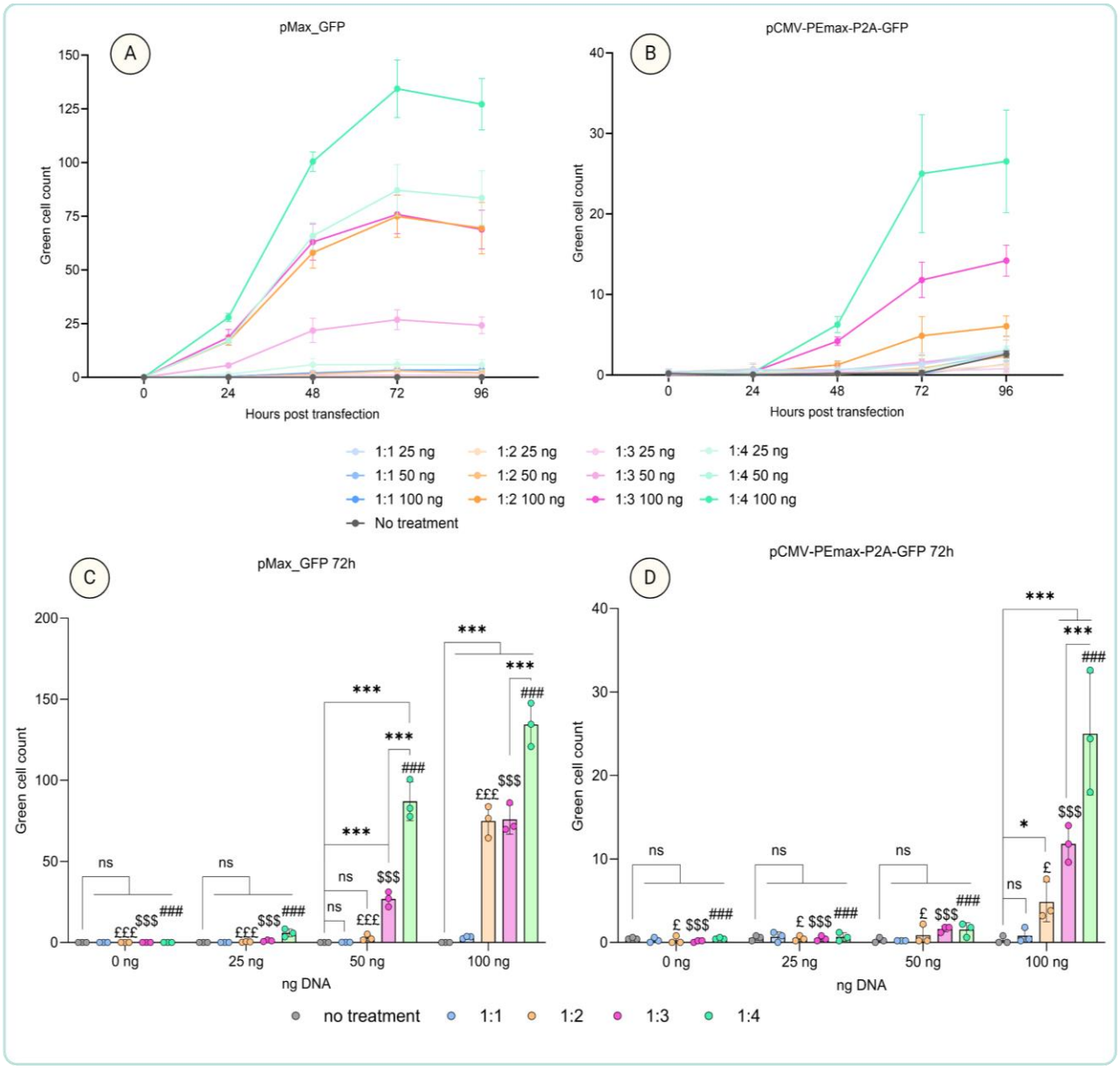


Figure 4.5: Green fluorescent cell counts in response to different transfection conditions in HT1080 cells. 3 200 HT1080 cells in a 96-well plate were transfected with different ratios of TransIT-LT1 with **(A)** pMax_GFP and **(B)** pCMV-PEmax-P2A-GFP and monitored in the IncuCyte S3 at 0, 24, 48, 72, and 96 hours post transfection. 25, 50, and 100 ng of plasmid DNA was transfected with 1:1, 1:2, 1:3, and 1:4 ratios between DNA and reagent, with three parallels conducted for each experimental combination. A green fluorescent cell count analysis was conducted across all time points and conditions. The y-axis denotes the total number of green objects (cells) counted per well, and the x-axis shows hours post transfection with n = 3 biological replicates. Panels **(C)** and **(D)** show the green fluorescent cell count at 72 hours post transfection as bar graphs. The y-axis denotes the total number of green objects (cells) counted per well, and the x-axis denotes different combinations of DNA:transfection reagent ratios and ng DNA with n = 3 biological replicates. The data are shown as mean ± SD. *** p_{adj}, <0.001, ### p_{adj} <0.001 for 1:4 ratios vs 0 ng, $$$ p_{adj} <0.001 for 1:3 ratios vs 0 ng, and £ p_{adj} <0.033, £££ p_{adj} <0.001 for 1:2 ratios vs 0 ng, based on 2-way ANOVA with Holm-Šídák correction for multiple testing. Figure and analysis created with GraphPad Prism and BioRender.com.

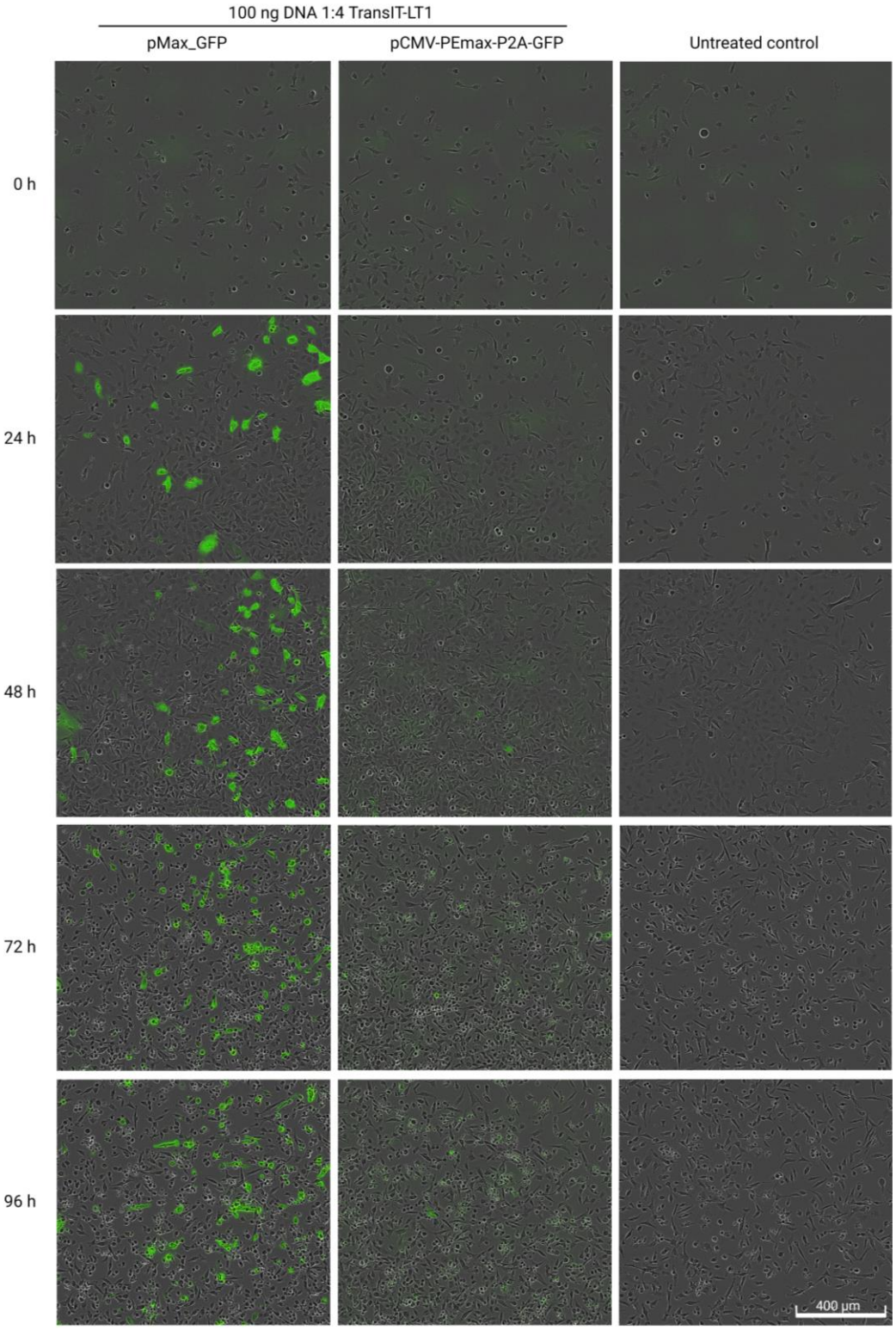


Figure 4.6: Images of changing GFP-expression in HT1080 at different time points after transfection. Merged brightfield and green fluorescent channel images of cells transfected with 100 ng pMax_GFP or pCMV-PEmax-P2A-GFP at 0, 24, 48, 72 and 96 hours post transfection along with an untreated control. Each image is a representative of five images taken pr well of n = 3 wells from a single experiment. The images were taken with phase contrast and GFP channel at 10x magnification, scale bar = 400 µm.

4.2.3. Transient transfection of the complete PE system in HT1080 cells

Based on the observations made during the transfection pilot, the PE experiment was conducted with the same best-performing ratio between TransIT-LT1 and plasmid DNA (1:4 and 100 ng) but scaled up to 6-well plates (96 000 cells/well and 2.5 µg DNA). Because changed morphology and indications of reduced viability was observed past 48 hours, the cells were sorted by FACS at the 48-hour time point. The transfection was performed with and without the MMR inhibitor MLH1 (PE4 and PE2, respectively). The three pegRNA-plasmids (Table 2.5) were pooled for each direction of the edit, one pool for peg A and one for peg C. The cells were imaged by the Nikon TE2000 Eclipse fluorescent microscope 24 hours post transfection. All samples transfected with the PE plasmid pCMV-PEmax-P2A-GFP displayed strong GFP expression in a relatively high proportion of the cells (Figure 4.7, PE2 not shown). Therefore, the experiment proceeded to FACS sorting for positively transfected cells the next day, at 48 hours post transfection.

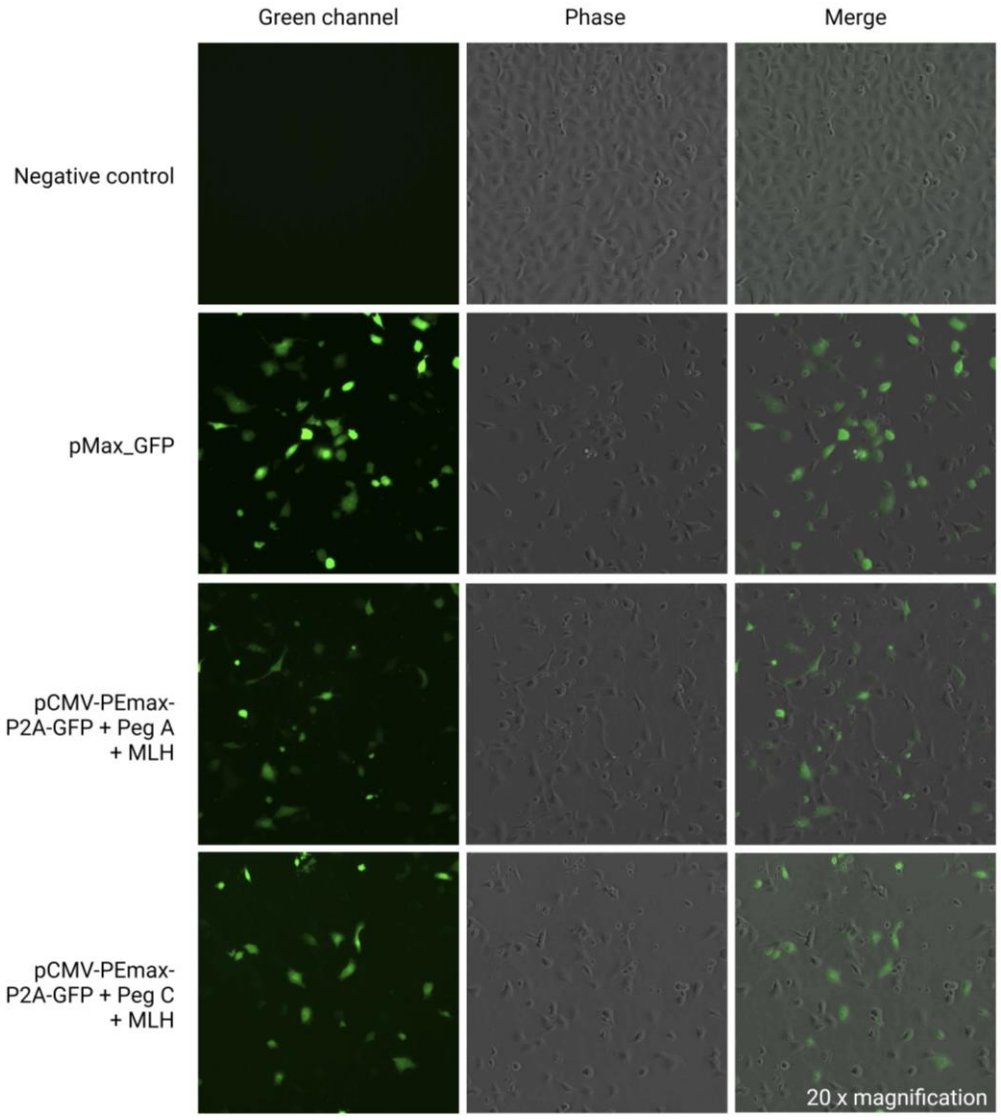


Figure 4.7: **Images of GFP-expression 24 hours after transient transfection of the PE-system in HT1080 cells.** HT1080 cells were transfected with 2.5 µg plasmid DNA with a 1:4 ratio between DNA and reagent, TransIT-LT1. 48 hours post transfection the GFP positive cells were sorted by FACS and pooled. 24 hours after transfection the cells were imaged at 20x magnification (scale bar = 400 µm) at phase contrast and green channel by the Nikon Eclipse TI2. The images were edited with Fiji ImageJ, where a merge between the channels was created, as can be seen on the right column.

FACS of the PE-transfected HT1080 cells showed a transfection efficiency for the live single cell population of 98% for pMax_GFP and on average 34% for pCMV-PEmax-P2A-GFP (Table 4.2), indicating that the pCMV-PEmax-P2A-GFP plasmid was relatively successfully transfected into the cells. The gates were initially set based on a pMax_GFP positive control and validated with cells transfected with pCMV-PEmax-P2A-GFP only (no pegRNAs; Figure 4.8). While the cells transfected with the pMax_GFP plasmid produced a defined cell population with uniform GFP intensity, the PE-transfected cells showed a large spread in intensity within the GFP positive gate (Figure 4.8). Approximately 24 000 GFP positive cells were sorted from each sample (except gating controls). A relatively small proportion of each sample was sorted as live cells, indicating a large proportion of cells that did not survive the transfection or the FACS treatment (Table 4.2).

Table 4.2 | Percentage of live, single, and GFP-positive cells sorted by FACS

Sample:	Freq. of Parent (%)	Live cells Freq. of Parent (%)	Live cells/Single Cells Freq. of Parent (%)	Live cells/Single Cells/GFP neg Freq. of Parent (%)	Live cells/Single Cells/GFP pos Freq. of Parent (%)
GFP negative control	100.0	73.3	97.6	99.7	0.0
pMax_GFP positive control	100.0	33.8	96.9	0.8	98.0
PEmax-GFP	100.0	61.3	96.7	61.9	32.7
PEmax-GFP + MLH	100.0	49.4	97.4	60.9	33.7
PEmax-GFP + Peg A	100.0	52.5	96.9	55.7	38.4
PEmax-GFP + Peg A + MLH	100.0	48.9	96.8	52.0	41.3
PEmax-GFP + Peg C	100.0	43.3	96.9	65.4	29.4
PEmax-GFP + Peg C + MLH	100.0	35.6	96.6	61.2	32.8

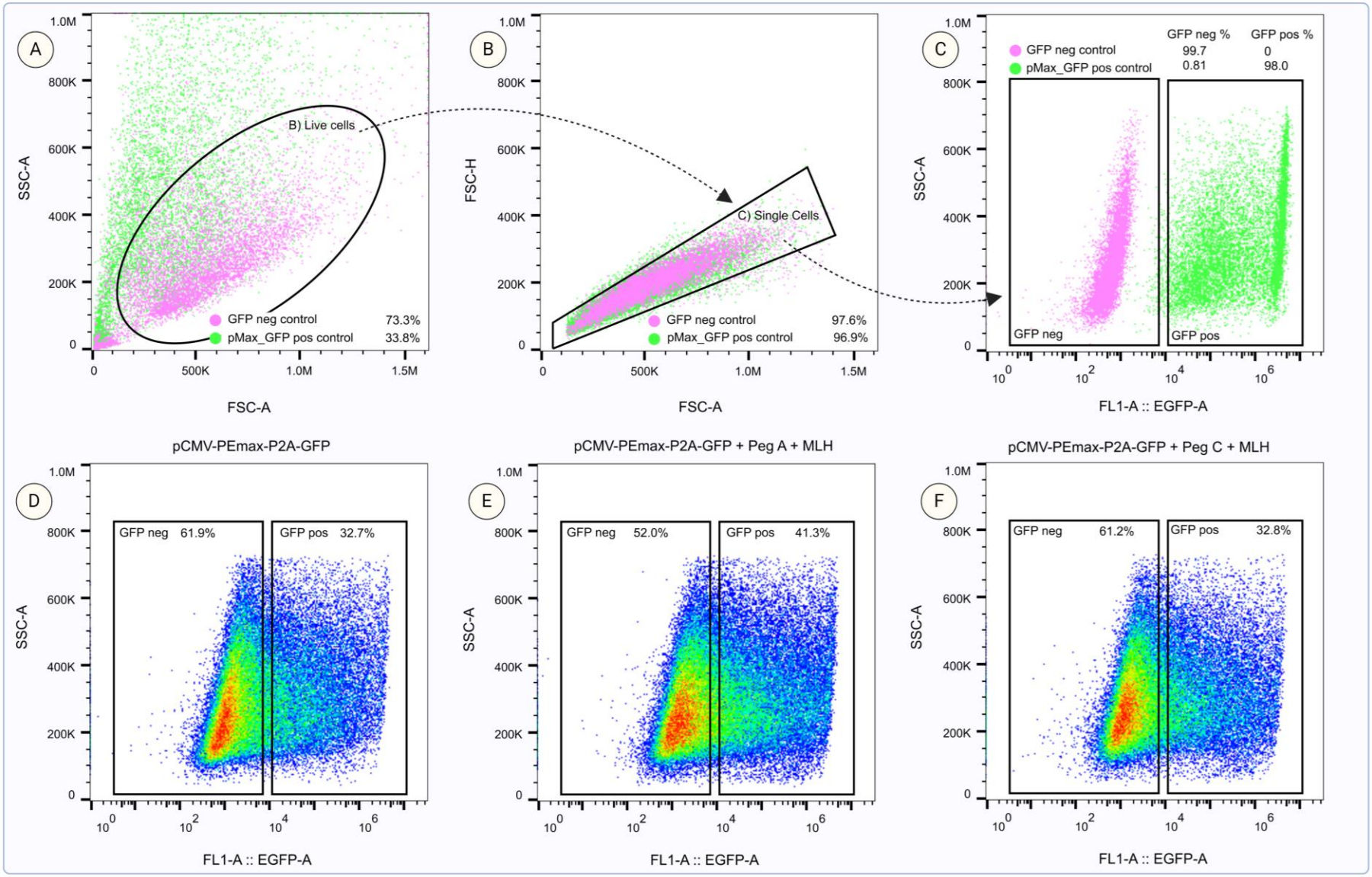

A
B) Live cells
GFP neg control 73.3%
pMax_GFP pos control 33.8%
SSC-A
FSC-A
B
C) Single Cells
GFP neg control 97.6%
pMax_GFP pos control 96.9%
FSC-H
FSC-A
C
GFP neg % GFP pos %
GFP neg control 99.7 0
pMax_GFP pos control 0.81 98.0
GFP neg
GFP pos
SSC-A
FL1-A :: EGFP-A
pCMV-PEmax-P2A-GFP
D
GFP neg 61.9%
GFP pos 32.7%
SSC-A
FL1-A :: EGFP-A
pCMV-PEmax-P2A-GFP + Peg A + MLH
E
GFP neg 52.0%
GFP pos 41.3%
SSC-A
FL1-A :: EGFP-A
pCMV-PEmax-P2A-GFP + Peg C + MLH
F
GFP neg 61.2%
GFP pos 32.8%
SSC-A
FL1-A :: EGFP-A

Figure 4.8: Density plot of GFP-positive HT1080 cells sorted by FACS. The cells were transfected by plasmids according to the PE4 system with a total of 2.5 µg DNA per well in 6-well plates. The cells were sorted with the SONY SH800S Cell Sorter. Figures (A), (B), and (C) show the gating strategy using positive and negative controls for GFP, where the positive control was transfected with only the pMax_GFP plasmid. **(A)** The figure shows the whole cell population with the gating set on forward scatter area (FSC-A) against side scatter area (SSC-A) to get an estimation of the main population of live cells, assess granularity and to filter out debris. **(B)** The live cell population was shown in a new plot as FSC-A against forward scatter height (FSC-H) and the second gate (single cells) was set diagonally to exclude doublets. **(C)** The single cells were shown in a new plot as FL1-A (GFP) against SSC-A. The negative and positive controls were used to set the GFP gates. **(D)** Cells transfected with only pCMV-PEmax-P2A-GFP acted as a GFP positive control (as well as a negative PE control) to sort the cells according to the gates set in figure (C). **(E)** Cells transfected with pCMV-PEmax-P2A-GFP, pegRNA encoding the A-allele of rs1799993, and mismatch repair inhibitor (MLH in the figure) corresponding to the PE4-system. **(F)** Cells transfected with pCMV-PEmax-P2A-GFP, pegRNA encoding the C-allele of rs1799993, and mismatch repair inhibitor (MLH in the figure) corresponding to the PE4-system. Figure created with FlowJo and BioRender.com.

4.2.4. Single-cell colony expansion of PE-transfected HT1080 cells

To achieve homozygous cell lines containing the risk or protective alleles, colonies starting from a single cell needed to be achieved. The cells were therefore sorted by FACS as single cells (one cell per well) in 96-well plates, or pooled and kept as a carry-on culture which was then utilized to seed cells thinly on larger plates for colony formation. Thus, 95 cells were sorted as single cells in a 96-well plate, with one well containing 100 cells (per sample) and the carry-on cultures were utilized as needed to seed 1000 cells on 10 cm or 15 cm plates for single-cell colony formation.

The cells that were sorted as single cells in 96-well plates had a varying survival rate from sample to sample. Only 6 colonies across all four samples were successfully expanded this way. However, for the cells that were seeded out as 1000 cells on 10 cm and 15 cm plates formed many defined colonies with clean margins and little evidence of cell migration on the plates (Figure 4.9). It took on average between 7 and 14 days for the colonies to form. When they had become sufficiently confluent (as can be seen on day 9 in Figure 4.9), each colony was independently removed from the plate by gently scraping with a pipette tip or cell scraper, and transferred to the well of a 96-well plate to expand further. In total, 162 colonies were successfully expanded: 40 colonies transfected with peg A, 34 transfected with peg A and MLH, 40 transfected with peg C, 40 transfected with peg C and MLH, and 8 negative PE control colonies (WT).

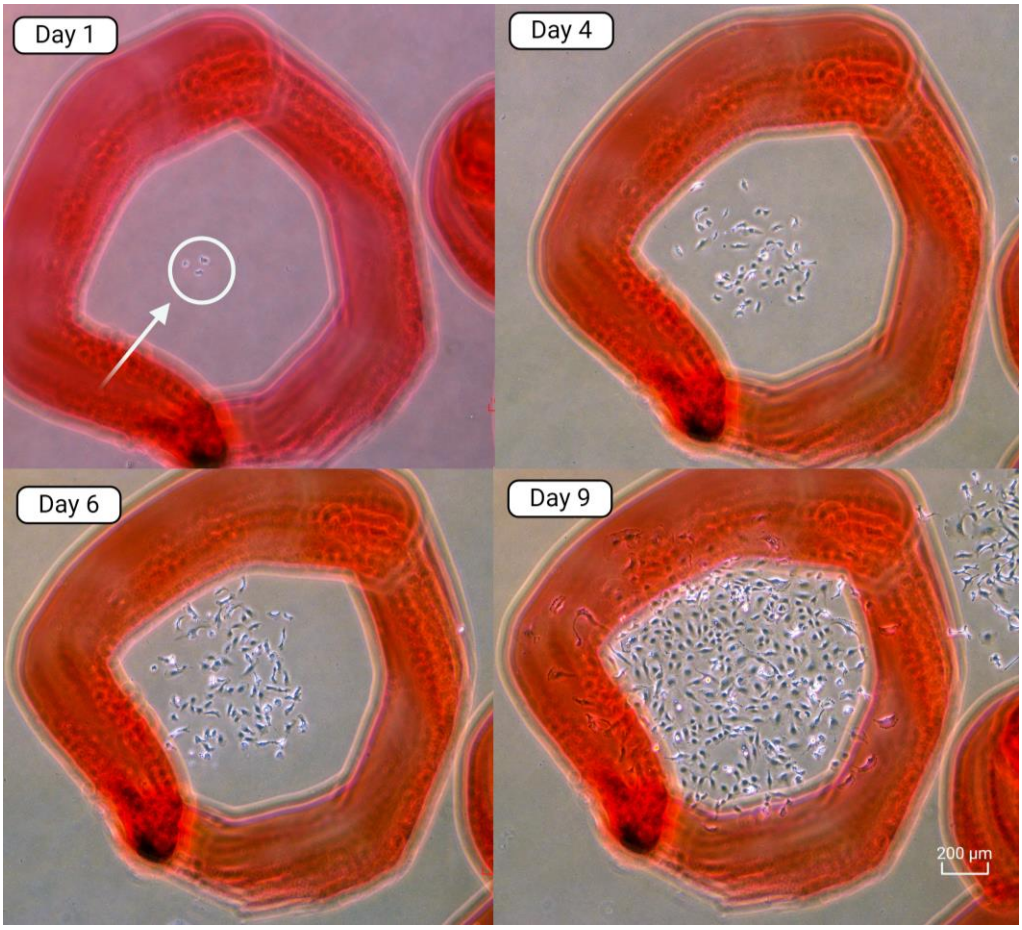


Figure 4.9: Colony expansion of transfected HT1080 cells. The method of colony expansion that gave the highest number of successfully expanded colonies consisted of seeding a few cells over a large area, then tracking single-cell expansion by drawing a red ring around single cells and observing them over the following days with the Nikon TS100 Eclipse. The colonies that formed were subsequently transferred to single wells in a 96-well plate for further expansion. The figure shows the development of a colony grown on a 15-cm culture plate for 9 days after seeding 1000 FACS-sorted HT1080 cells. On day 1 (24 hours after seeding) the single cell had already split into three cells.

4.2.5. Screening for edited HT1080 cells using qPCR with SNP-specific primers

To screen the 162 colonies that had been expanded for a potential edit, SNP-specific qPCR primers were designed that would only give hits on their respective alleles. To validate that the primers worked as intended, they were tested on gDNA from patient samples where the alleles for rs1799993 was known to be homozygous for A (V586_A) or for C (V602_C). As expected, the primers only amplified gDNA with the matching genotype (Figure 4.10). An estimation of gene copy number (GCN) for the alleles were made based on Eq. 2 (section 3.9.6). The GCN was found to be about 1.0 for the A-alleles in V586_A and 0.3 for the C-alleles in V602_C, relative to the control genomic locus. Although the expected GCN for a homozygous allele lies

around 2, suggesting these primers were not fully efficient, there was nevertheless a significant difference between the primers in each sample, indicating that the primers worked as intended.

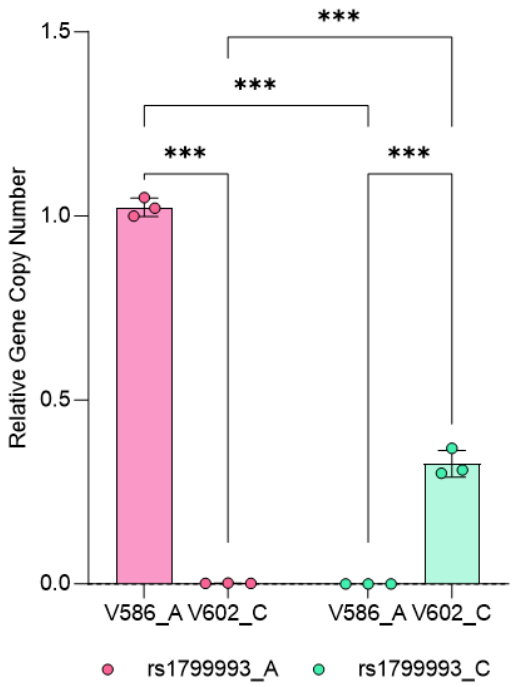


Figure 4.10: SNP-specific primers tested on patient gDNA samples. To screen edited cells for successful editing of rs1799993, SNP specific primers were made. The primers were tested on patient samples where the genotype was known to be homozygous for the two alleles (A or C). The approximate gene copy number (y-axis) was calculated based on a genomic locus on chromosome 16 as reference. The x-axis denotes the biological samples with n = 3 technical replicates. The data are shown as mean ± SD. *** p_{adj} <0.001 based on 2-way ANOVA with Holm-Šídák correction for multiple testing. Figure and analysis created in GraphPad Prism.

Once the primers had been validated, they could be applied to the expanded colonies. Due to the large number of samples, Chelex 100 was applied to create a crude lysate containing gDNA from the samples. The gDNA was then analyzed by qPCR using the SNP-specific primers, and the resulting Ct values were analyzed by comparing the difference between the A-primer and C-primer scaled to represent positive values for A-edited colonies and negative values for C-edited colonies, as described in section 3.8.5. There were several attempts at using Chelex for the gDNA harvest, and several attempts made with different amounts of gDNA for the reaction, which all gave varying success (data not shown). The resulting graph is based on 4 µl Chelex 100 gDNA diluted 1:5 from colonies grown in 24-well plates until 70-90% confluent (Figure 4.11). A total of 10 colonies gave an A/C-ratio swaying towards A or C (ratio of at least +/- 10) and were chosen for further analysis, which included colonies C21, C30, C48, C55, C59, C66, C107, C134, C142, and C148.

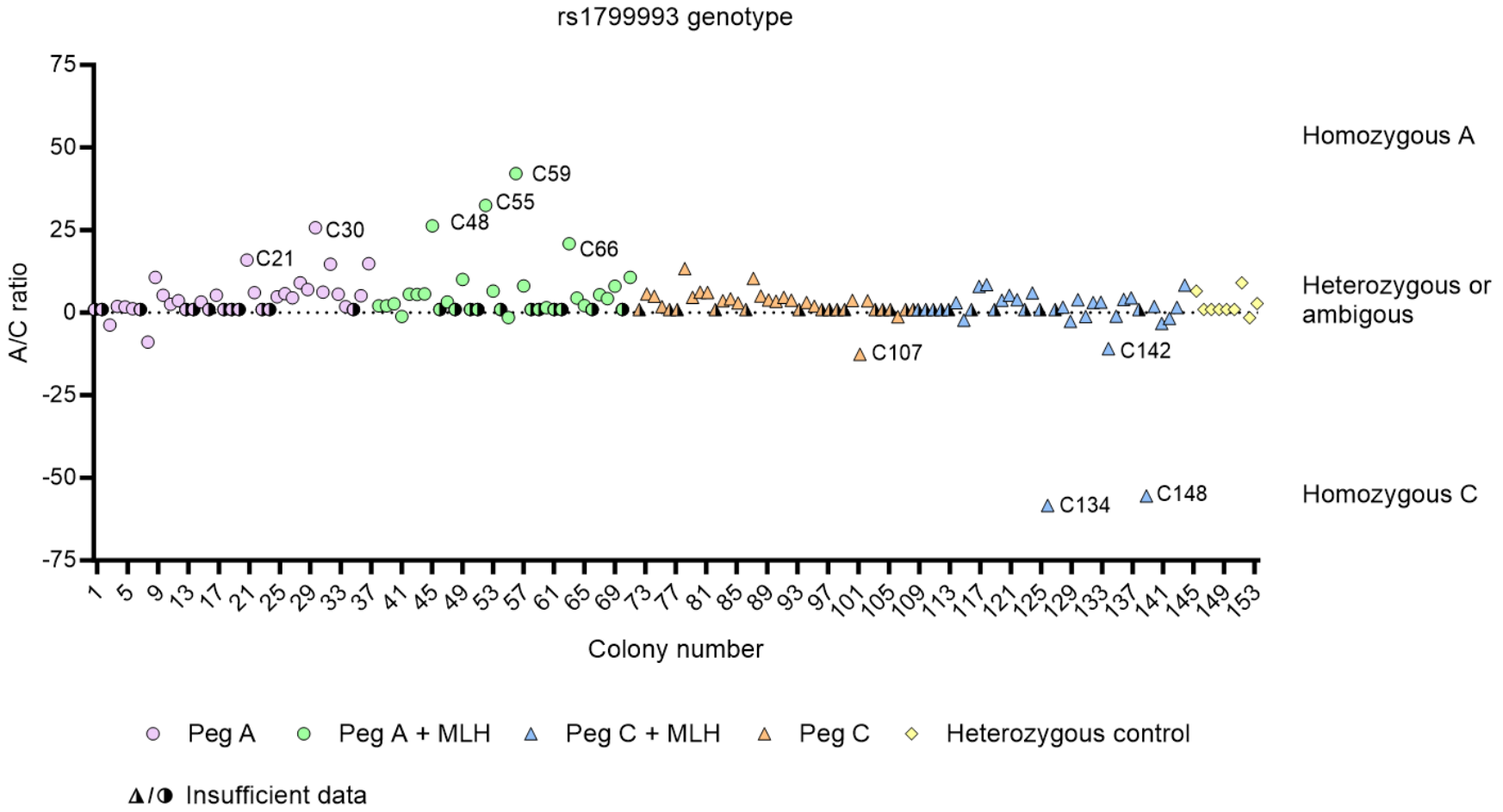


Figure 4.11: Analysis of Ct values between SNP-specific primers on successfully expanded PE-edited HT1080 colonies. Chelex 100 was applied to create a crude lysate containing genomic DNA (gDNA) from 70-90% confluent colonies grown in 24-well plates (diluted 1:5). The gDNA was then analyzed by qPCR and the resulting Ct values were analyzed by comparing the difference between the A- and C-primer. Antilog values of the difference between the Ct values ($2^{-\Delta Ct}$) represented A and scaled to represent C by negative one over the antilog value ($-1/2^{-\Delta Ct}$). The colonies are represented as dots (peg A), triangles (peg C), or diamonds (WT, heterozygous control), sorted by colony number on the x-axis (n = 1 colony) and the A/C-ratio on the y-axis. Colonies that did not result in detectable Ct values are shown as insufficient data with half-filled symbols. Colonies of interest are labeled with name. Figure created in GraphPad Prism.

4.2.6. Genotype validation of edited HT1080 colonies

The colonies that the qPCR results indicated might be homozygous for A or C were genotyped, as well as a selection of the ones that did not give any data. The colonies were expanded and gDNA was isolated by silica-based spin columns. The area surrounding the SNP was amplified by PCR using flanking primers (primer for *HMBS*, Table 2.11Table 2.1) and sequenced by Sanger sequencing. A total of 87 colonies were sequenced; the ones listed above in section 4.2.5 as well as a random selection of ambiguous colonies. A single colony was found to be homozygous for the C-allele, suggesting a successful edit (Figure 4.12). Interestingly, none of the colonies that the qPCR-results indicated had been edited turned out to be homozygous for the SNP. The only colony that had been successfully edited was C120 (Figure 4.12). The colony had been transfected with a pool of peg C and included MLH1 (PE4), as described in section 3.8.3. This gives an estimated editing efficiency of approximately 1%.

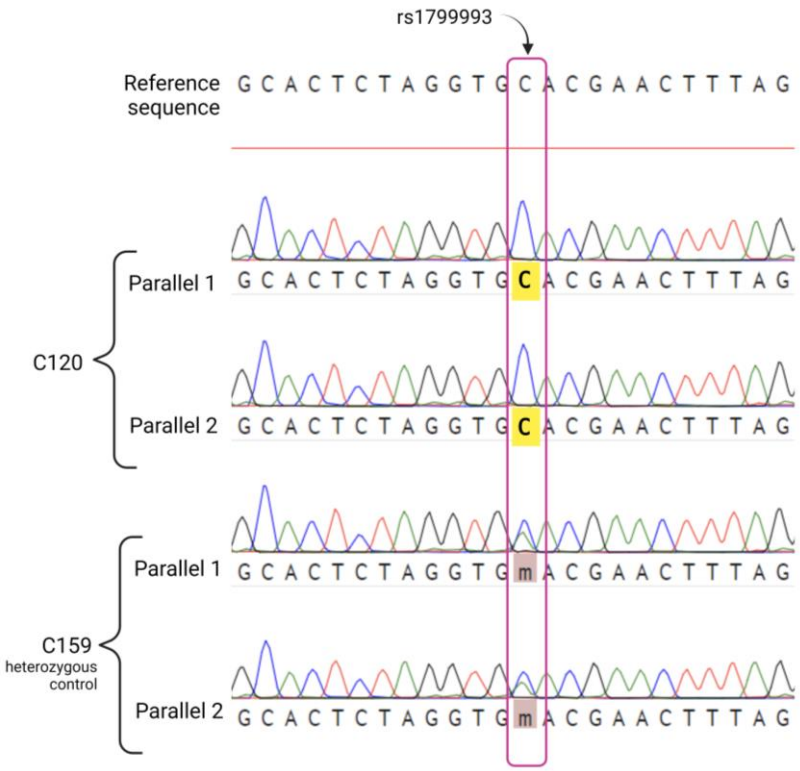


Figure 4.12: HT1080 colony edited by prime editing (PE) at the rs1799993 SNP from heterozygous (A and C) to homozygous for C. The figure shows chromatograms from BigDye sequencing of colonies 120 (C120) and 159 (C159) aligned to the reference sequence for the area surrounding the SNP. The arrow points to the SNP rs1799993, which on the reference sequence displays the C-allele, is homozygous for C in C120 (blue peak, C highlighted in yellow), and heterozygous for both alleles (A and C highlighted as 'm', green and blue peaks, respectively) for C159. The colonies were grown from single cells post transfection of PE-plasmids. C159 was not transfected with PE-plasmids and kept as a WT control. Alignment and results were analyzed in the online software Bencling.com using the MAFFT auto multiple-alignment algorithm. Figure created with BioRender.com.

4.3. Prime editing in ASC52telo cells using lentiviral delivery

4.3.1. Virus production in HEK293T

Having proved the principle of PE in HT1080 cells, we next turned our attention to the more biologically relevant ASC52telo cells. Due to their resistance to transfection of large plasmids (Krill, 2023), we reasoned

that lentiviral delivery could potentially be an alternative strategy of PE delivery. This strategy would entail first generating cells that stably overexpress the PE enzyme, then transiently overexpressing the smaller plasmids harboring the pegRNAs to initiate the genomic editing.

HEK293T were used to produce viral particles carrying PEmax, as described in section 3.9.1. Three viral harvests were conducted from a single transfection (24, 36, and 44 hours post transfection), with the Lenti-X^{TM} GoStix being used to estimate the number of viral particles from each harvest before and after concentrating the harvests with the Lenti-X^{TM} Concentrator, and after a freeze-thaw cycle (Table 4.3). The viral particles were concentrated to 1/100th of their original volume and resuspended in Opti-MEM. While all three harvests produced viruses, the third harvest was found to have the highest concentration of the lentiviral capsid protein p24. After freezing and thawing of the sample, a total of 986 ng p24 was identified, which corresponds to 9.81 x 10^9 physical viral particles (PP).

Table 4.3 | Viral capsid protein (p24) detected before and after concentration, after freeze-thaw cycle, and estimated number of physical viral particles (PP)

Harvest (hours post transfection)	**p24 (ng/ml)**	**p24 post-concentrator (ng/ml)**	**Recovery**	**p24 post freeze-thaw (ng/ml)**	**Recovery**	**Total p24 in 0.9 ml sample (ng)**	**PP post freeze-thaw**
24	9						
36	10	2288	90%	1095	48%	986	9.81 x 10^9
44	57						
$\bar{x}$	25.3						

4.3.2. Determination of blasticidin dose on ASC52telo

Because the Lentiviral plasmid (pLenti-PE2max-BSD) contains resistance to blasticidin (BSD), which would be used as a selectable marker for positively transduced cells, a BSD titration assay was conducted on WT ASC52telo cells to identify the lowest concentration that would kill all non-transduced cells after 7 days (as described in section 3.9.2). The results were analyzed by IncuCyte S3 to calculate the confluency of the cells in each well. The results showed a significant difference in cell confluency between 10 and 20 µg/ml of BSD compared to the remaining concentrations (Figure 4.13). While 20 µg/ml completely killed all cells, some cells were still attached at 10 µg/mL (Figure 4.14). However, the difference in confluency between these two doses was not statistically significant. Based on these results, the decision was made to subject virus-transduced cells to a 20 µg/ml concentration of BSD to make sure all non-transduced cells were killed.

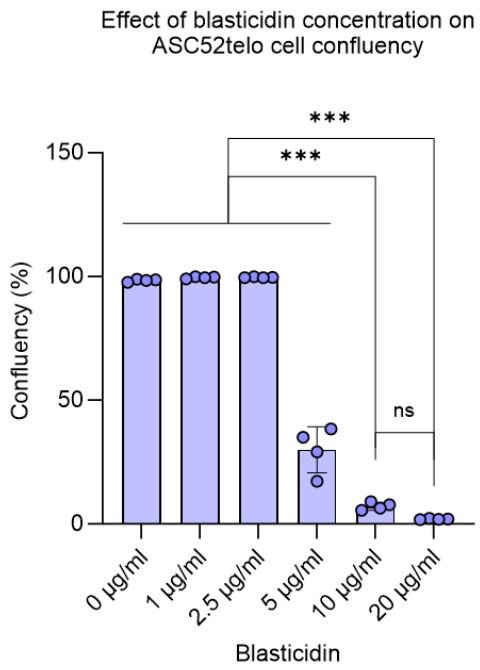


Figure 4.13: IncuCyte S3 analysis on cell confluency in response to blasticidin (BSD) concentration after 7 days. The pLenti-PE2max-BSD lentiviral packaging plasmid encodes a gene for BSD resistance. Successfully transduced cells would therefore be resistant to BSD. WT ASC52telo cells were subjected to six different concentrations of BSD (0, 1, 2.5, 5, 10, and 20 µg/ml over the span of 7 days), and the cell confluency was subsequently analyzed by the IncuCyte S3. The x-axis denotes BSD concentration with n = 4 biological replicates for each experiment, and the y-axis denotes confluency. The data are shown as mean ± SD. p_{adj} <0.001 (***); ns, non-significant, based on an ordinary one-way ANOVA with Holm-Šídák correction for multiple testing.

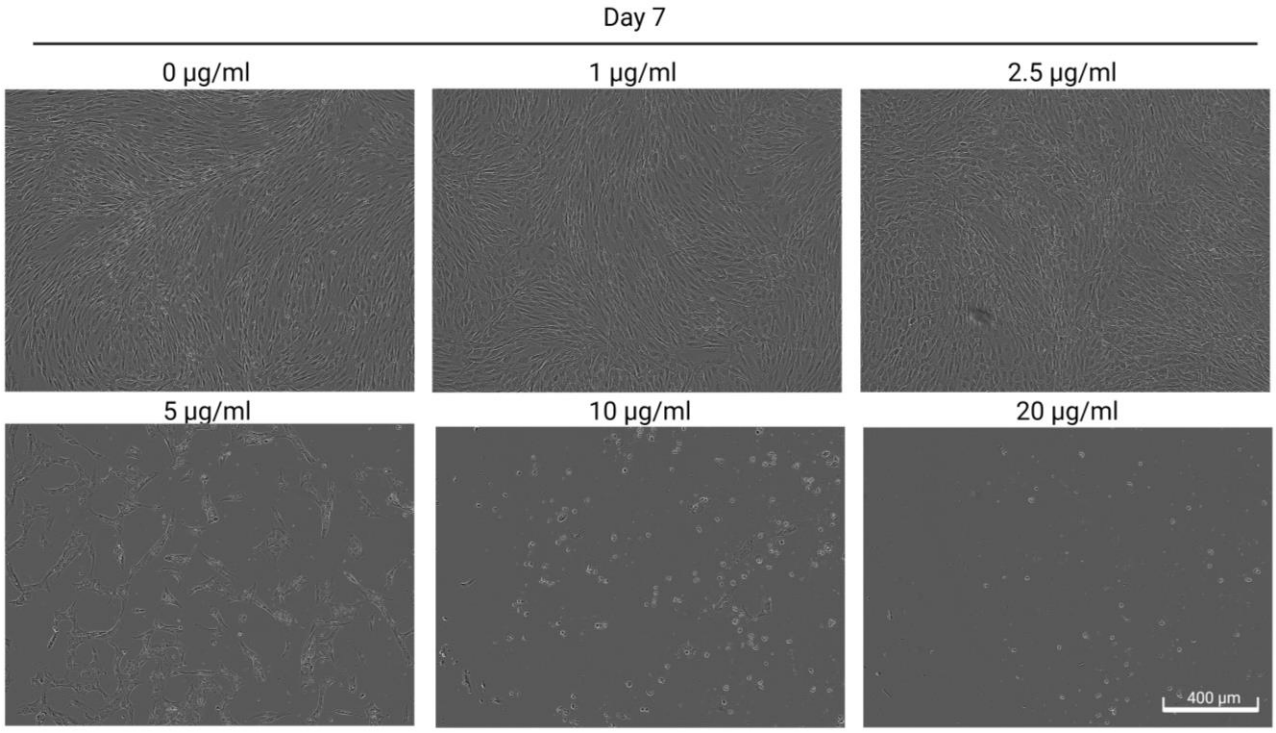


Figure 4.14: IncuCyte S3 images of ASC52telo and HT1080 cell confluency in response to blasticidin (BSD) concentration after 7 days. 5000 cells/cm^2 ASC52telo cells were exposed to six different concentrations of BSD (0, 1, 2.5, 5, 10, and 20 µg/ml) over the span of 7 days. On day 7, the cells were imaged by the IncuCyte S3 with phase contrast at 10x magnification, scale bar = 400 µm. Each image is a representative of four images taken pr well of n = 4 wells from a single experiment. A cell confluency analysis was run based on these images.

4.3.3. Determination of functional viral titer

After production of virus containing the PEmax construct, a viral titer test was performed on WT ASC52telo cells in order to see how many viral particles were capable of infection (IFU). RetroNectin is a recombinant human fibronectin fragment reported to enhance transduction efficiency (Kurachi et al., 2017), and thus two experiments were performed, one with the wells coated with RetroNectin, and one without, to see if the transduction was benefitted by the addition of RetroNectin. The transduction was performed in a 5-fold serial dilution with 5 points, where 20 µg/ml of BSD was added to the medium to start the selection 2 days later.

13-days post transduction (11 days after introduction of BSD) the number of colonizing-forming units (CFU) was counted, and the corresponding IFU/ml was calculated based on the dilution factor and ng p24 (as measured by the GoStix™) of the viral sample. Pictures were taken with the Nikon Eclipse TS100. ASC52telo did not seem to readily have formed colonies, and the number of surviving cells varied greatly between the dilutions, so a rough estimation of the CFU was therefore made based on the number of surviving cells in the wells where the cells were easiest to count (Figure 4.15 dilution 1:25 and 1:125, and Table 4.4). There was no observable difference between the wells coated with and without RetroNectin (Figure 4.15). Based on the rough estimation of CFU, the IFU in the viral stock was found to be 1.93×10^5 IFU/ml (Table 4.4). This makes the ratio of IFU to PP 1.8×10^{-5} IFU/PP (Table 4.4). This provided the foundation for testing different multiplicities of infection (MOIs), representing the IFU-to-cell ratio, for the transduction.

Table 4.4 | IFU estimated based on surviving cells 13 days post-transduction

Viral dilution	CFU (surviving cells)/well	IFU/ml	IFU/PP
1:25	129	120 000	-
1:125	85	266 875	-
$\bar{x}$	107	193 438	1.8×10^{-5}

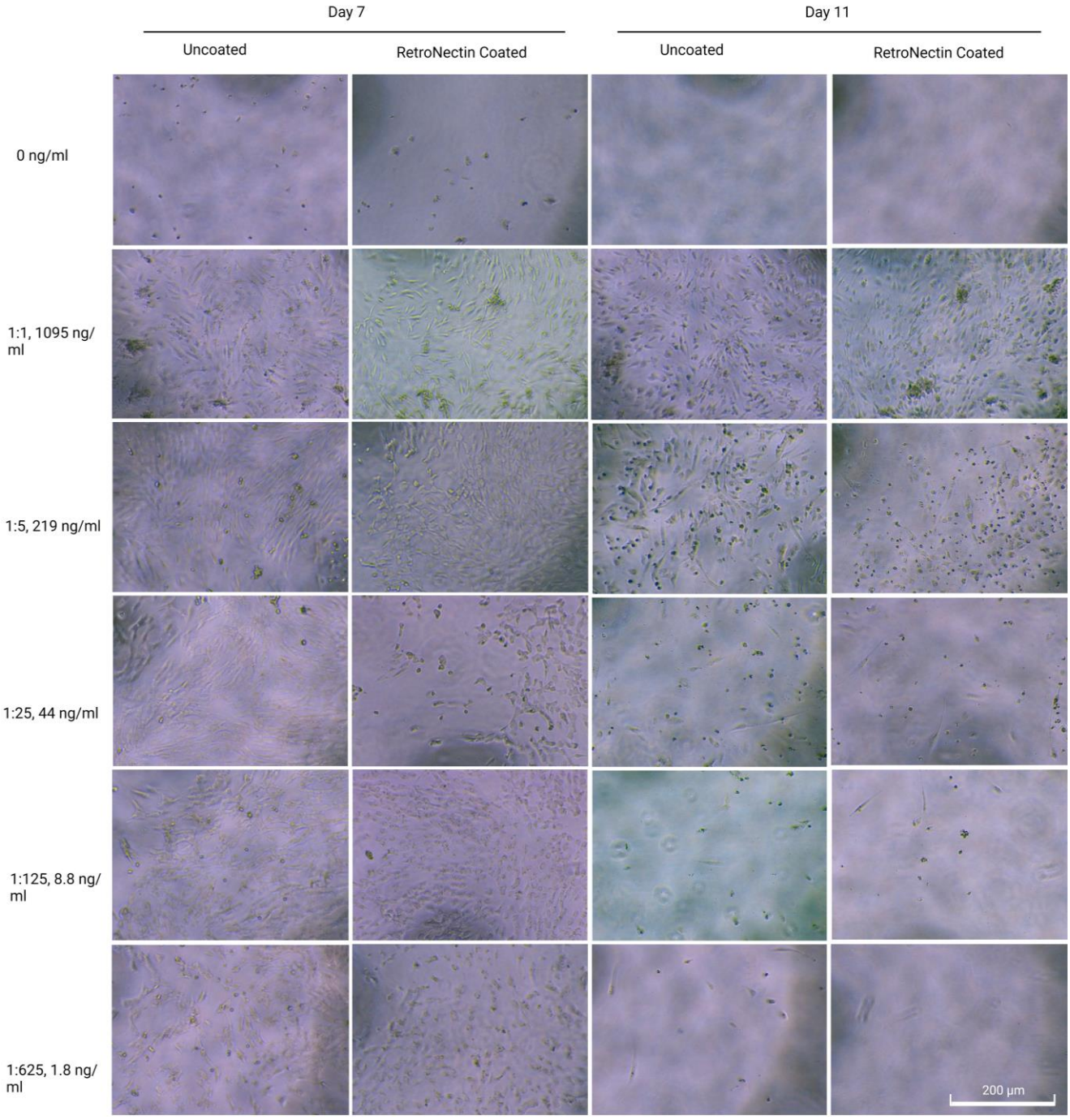


Figure 4.15: Estimation of viral titer on ASC52telo. 4000 cells (96-well plate) of WT ASC52telo were subjected to different concentrations of virus and selected by 20 µg/ml blasticidin (BSD). Images were taken with the Nikon Eclipse TS100 at 10 x magnification, Scale bar = 200 µm, and each image represents one biological replicate. The images were taken 7 and 11 days after introduction of BSD (9 and 13 days post transduction). The viral titer was estimated based on the images taken on day 11 in the wells coated with 5 µg/cm^2 RetroNectin.

4.3.4. Lentiviral transduction of ASC52telo with pLenti-PE2max-BSD, and single-cell colony expansion

To generate ASC52telo cells stably overexpressing PEmax, WT cells were transduced with the lentiviruses encoding PEmax at MOIs of 0.05, 0.1, 1, 2 and 5. Two days after transduction the cells were subjected to 20 µg/ml BSD. This concentration of BSD was kept until one week after the kill-well was confirmed dead. After that, the cells were kept in 10 µg/ml as maintenance concentration.

The transduced cells were seeded out at varying densities to find the optimal density for colonies to form. The lower MOIs of 0.05, 0.1, and 1 were moved from the 96-well plate of the transduction to 6-well plates and then later 10 cm plates for the selection. They died approximately one month after the kill-well had died, suggesting these viral doses were too low to robustly transduce the target cells. There were very few surviving cells on the plate after the selection with BSD, and the ones that were left did not proliferate. Before they died, the cells were attempted revived by addition of conditioned medium harvested from WT cultures. Of note, many of the cells that were monitored in the microscope did not proliferate, but instead moved around on the plate for a while before they died (Figure 4.16).

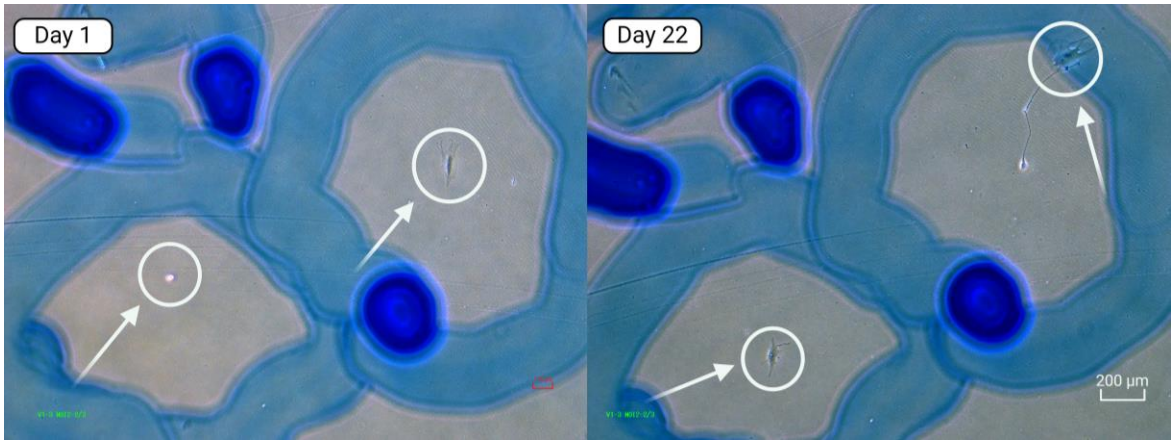


Figure 4.16: ASC52telo single cells transduced with prime editing (PE) -carrying lentivirus showing no signs of colony formation. The cells were transduced with lentivirus carrying PEmax and resistance to blasticidin (BSD) at an MOI of 2. The cells were seeded at approximately 1000 cells in a 10 cm plate 2 days post transduction where BSD was introduced at 20 µg/ml. Single cells were tracked for colony formation by marking the underside of the plate with marker (blue lines in the images) and monitored microscopically every day for signs of proliferation and colony growth. The single cells are indicated in the images with white circles and arrows. On day 22 the cells still showed no sign of division. Images were taken by the Nikon TS100 Eclipse at 10x magnification Scale bar = 200 µm. Figure created with BioRender.com.

Sorting single cells by FACS and growing them in medium with added FGF (1 ng/ml, v/v, f.c.) and EGF (10 ng/ml, v/v, f.c.) was attempted as a method of expanding single-cell colonies. None of the cells survived.

For MOIs of 2 and 5, the cells were moved directly from the 96-well plate of the transduction to the 10 cm plate for selection (2000 cells/plate). Cells were observed to have varying phenotypes, with some cells being very small and growing dense (Figure 4.17 A and B), while others were very elongated and grew far apart (Figure 4.17 C and D). The monitored cells did not seem to readily form colonies, and had more of an elongated look, while migrating a lot on the plate. It was discovered, however, that some cells had formed what looked like colonies outside of the monitored areas (Figure 4.17 A and B). It is unknown whether these colonies were true single-cell colonies, or if they were a collection of cells that had gathered in the same spot.

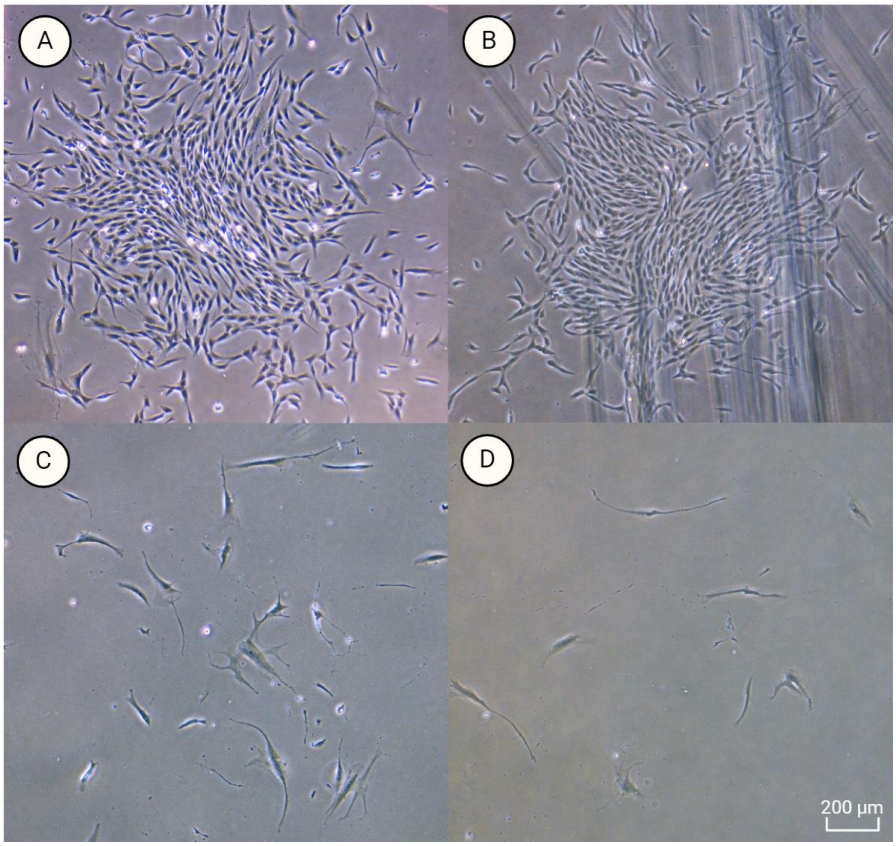


Figure 4.17: Varying phenotypes and colony formation in ASC52telo transduced with prime editing (PE) -carrying lentivirus. The cells were transduced with lentivirus carrying PEmax and resistance to blasticidin (BSD) at an MOI of 2 (A, B, and C) and an MOI of 5 (D). The cells were seeded at approximately 2000 cells in 10 cm plates 2 days post-transduction where BSD was introduced at 20 µg/ml. Figure **(A)** and **(B)** shows images of transduced ASC52telo (MOI 2) colonies that formed outside of the monitored areas of the plate. These colonies were picked and expanded. Figure **(C)** and **(D)** shows images of varying phenotypes between cells in contrast to the phenotypes that can be observed in images (A) and (B). The majority of cells were observed to grow like the cells depicted in (C) and (D). Images were taken by the Nikon TS100 Eclipse at 10x magnification Scale bar = 200 µm.

The cells were observed to migrate on the plate, often in the direction of other cells. This was the case both on plates where cells were very far apart (< 35 cells/cm^2), and for plates where cells were denser. To try to get cells to proliferate without being too far apart, a method was attempted of seeding the cells densely on the plate, and subsequently isolate cells by scraping away surrounding cells. This method appeared to work at first, but then surrounding cells migrated towards the isolated cells, and no colonies with clean margins that were distinguishable from surrounding cells formed (Figure 4.18).

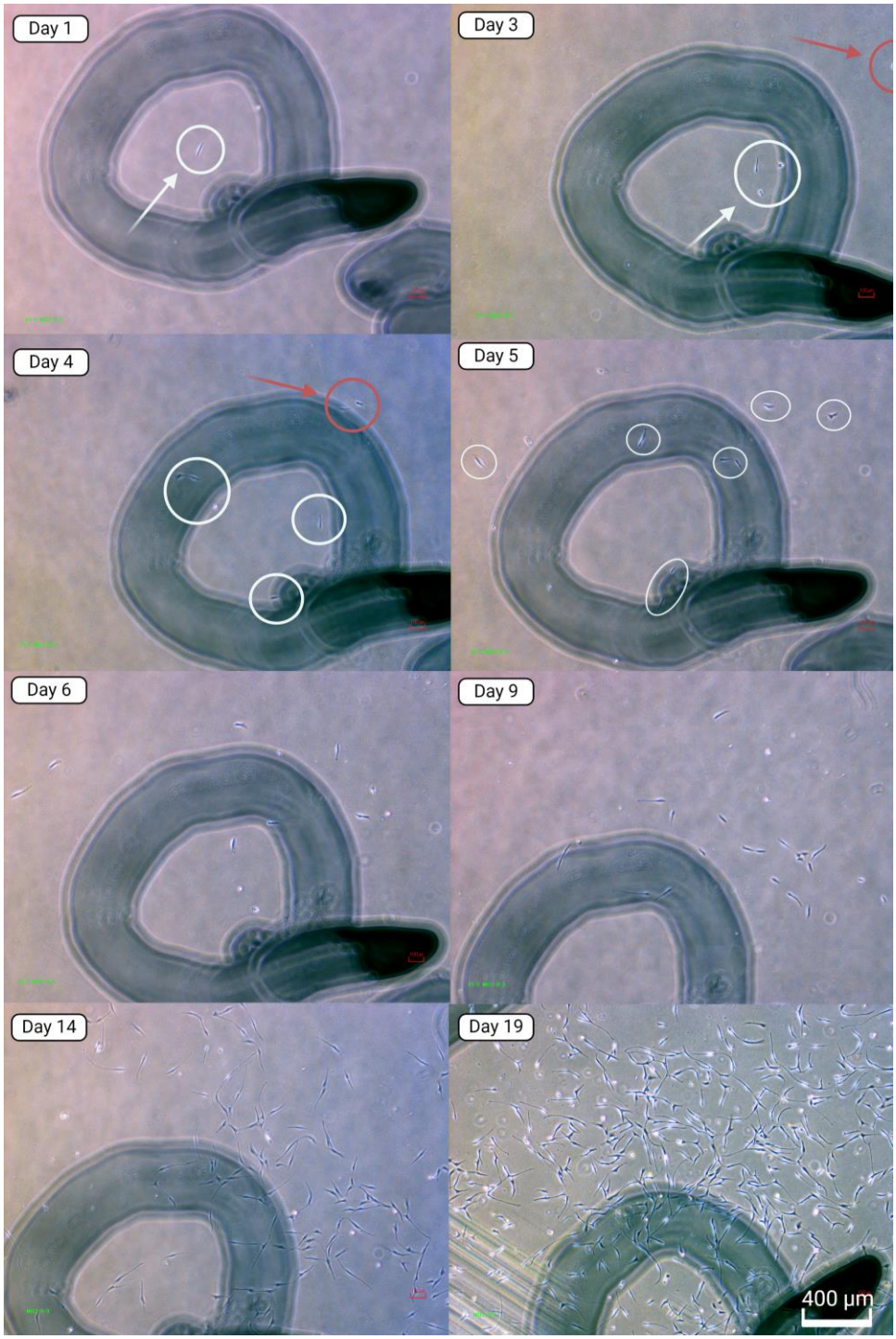


Figure 4.18: ASC52telo single cells transduced with prime editing (PE) -carrying lentivirus. The cells were transduced with lentivirus carrying PEmax and resistance to blasticidin (BSD) at an MOI of 2. The cells were seeded at approximately 5000 cells in a 10 cm plate from a subset of proliferating cells from MOI 2 approximately three weeks post transduction and selection of BSD at 20 µg/ml for 14 days, before the BSD concentration was reduced to 10 µg/ml for maintenance. Single cells were tracked for colony formation by marking the underside of the plate with marker (black lines in the images) and monitored microscopically every day for signs of proliferation and colony growth. Single cells that were too close to each other were isolated by scraping away surrounding cells. On day 1 (one day after seeding) there

was a single cell inside the lines, which had split to three cells by day 3 (white circles). In the corner of the image a cell can be observed that is migrating either towards the colony or away from it (red circle). On day 3 the cell had moved even closer to the colony and by day 5 it was impossible to tell which cells had originated inside the marked area and which had come from the outside. This trend continued on until day 19, when it was clear that these cells did not form colonies, but rather migrated to and from other cells to form a heterogeneous mixture of cells from different locations on the plate. Images were taken by the Nikon TS100 Eclipse at 10x magnification, scale bar = 400 µm. Figure created with BioRender.com.

As the expansion of single-cell colonies was unsuccessful, a series of heterogeneous cell cultures were expanded from varying cells that had either formed colonies without being monitored, or from cells that were proliferating, but migrated too much to see clean colonies form. These cultures were given names based on the MOIs they were harvested from, MOI 2.1, 2.2, 2.3, 2.4, and 5.1 (Figure 4.19). The cultures had great variance in phenotypes and growth patterns, with some of them looking very different from WT cells. The cultures were cryopreserved for future experiments and attempts at creating single-cell colonies.

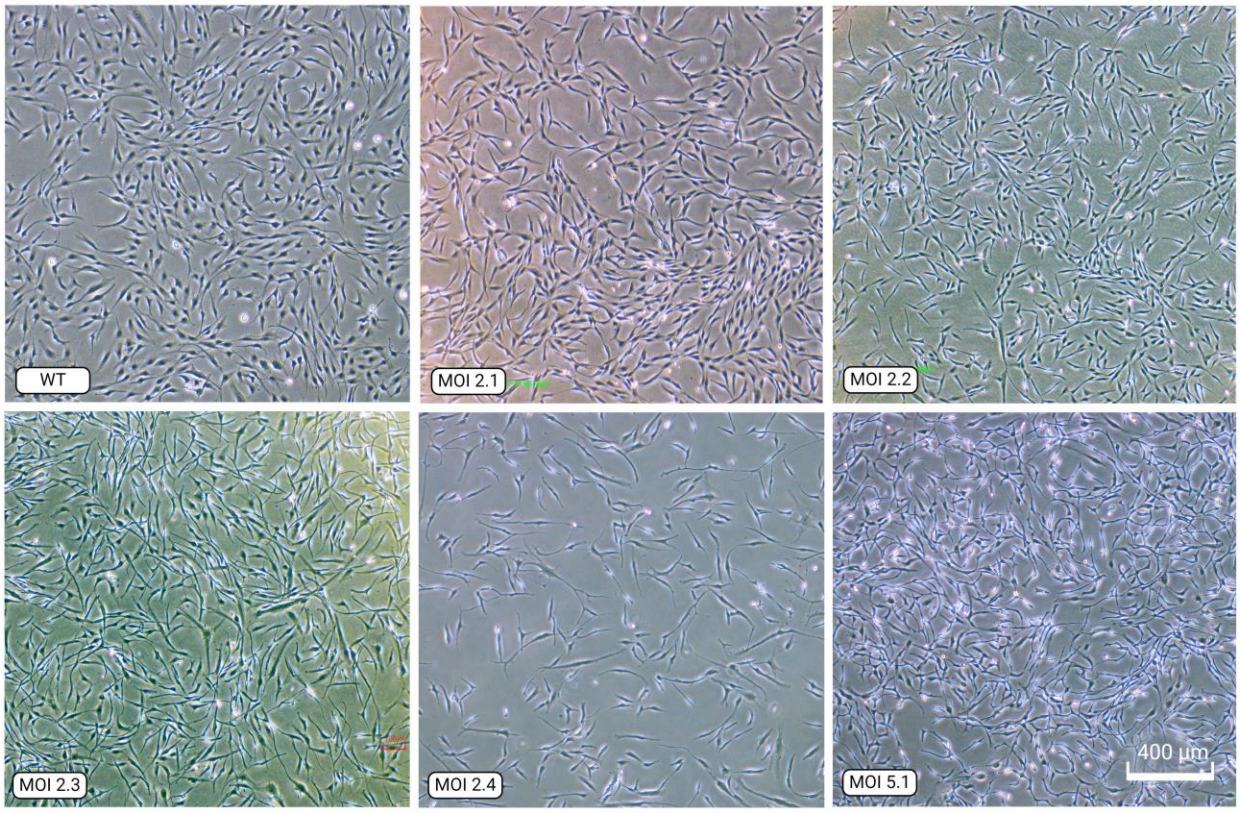


Figure 4.19: ASC52telo transduced with prime editing (PE) -carrying lentivirus, expanded cultures compared to wild type (WT) cells. The cells were transduced with lentivirus carrying PEmax and resistance to blasticidin (BSD) at an MOI of 2 and 5. 20 µg/ml of BSD was introduced to the medium 2 days post transduction for selection of transduced cells. After 14 days the concentration was reduced to 10 µg/ml for maintenance. While trying to produce clean single-cell colonies, cultures of heterogeneous mixes of cells were expanded from different subsets of proliferating cells to test downstream applications of the cells, as well as screen the cells for PEmax expression. In the upper left corner untreated

WT ASC52telo can be seen as elongated cells with long fibers extending from the points and dense-looking centers. These cells appear to grow separated by their fibers. MOI 2.1 can be seen as similar, but more diamond shaped and more compact, growing more densely, and with shorter extending fibers. In MOI 2.2 the cells appear smaller, but similar to 2.1. Both MOI 2.1 and 2.2 appear homogeneous in their phenotypes. MOI 2.3 and 2.4 appear more heterogeneous in the shapes and sizes of the cells, with longer fibers extending further, inhibiting the dense growth that can be seen for MOI 2.1 and 2.2. MOI 5.1 was the only surviving subset of cells from MOI 5, and were slow growing compared to the other cultures. Images were taken by the Nikon TS100 Eclipse at 10x magnification, scale bar = 400 µm. Figure created with BioRender.com.

4.3.5. Validation of PEmax expression in expanded cultures

To validate the presence of PEmax in the transduced cells, gDNA, mRNA and protein were isolated from the expanded cultures (MOI 2.1, 2.2, 2.3, 2.4, and 5.1) using the Norgen RNA/DNA/Protein kit. qPCR on gDNA and mRNA show that the cultures all expressed PEmax, and that the average GCN of PEmax was close to 2 (Figure 4.20 A and B, respectively). However, the PEmax enzyme could not be detected on WB, using 10 µg protein lysate and anti-Cas9 (as described in section 3.6.33.6.3) (Figure 4.20 C). Thus, although the virus was shown to be integrated into the genome of the ASC52telo cells, leading to expression of PEmax mRNA, it was not clear whether a functional protein was produced.

To test whether the missing PEmax band on the membrane could be due to failure of the Cas9 antibody to detect the nCas9+RT fusion protein, an overexpression of the enzyme was performed in HT1080 cells as described in section 3.5.1. Protein was subsequently isolated using both the Norgen RNA/DNA/Protein kit and RIPA WB lysis buffer. Lysing the cells in Norgen resulted in higher concentration than RIPA, with a total yield of 80.1 µg and 34.3 µg, respectively. The SDS-PAGE gel was then loaded to full capacity, which equated to 10 µg protein in the RIPA samples and 20 µg for the Norgen samples. 200 ng recombinant Cas9 was also included as a positive control, which has an expected size of 160 kDa compared to 250 kDa for PEmax. Ponceau staining was applied to verify that all the protein was transferred, which showed a difference in band intensity between the RIPA and Norgen samples for mid-size and large proteins (Figure 4.20 D). Staining with anti-Cas9 revealed weak bands for the Norgen samples of expected size around 250 kDa. Intriguingly, RIPA resulted in stronger bands of expected size 250 kDa, despite less total protein being loaded. However, the beta-actin loading control suggest a similar amount of total protein was loaded for both types of lysates. Finally, a very strong band of expected size (160 kDa) was observed for the recombinant Cas9 (Figure 4.20 E). Taken together, these data demonstrate that the Cas9 antibody is able to detect the PEmax fusion protein. Reasons why it was not picked up in the transduced cells are discussed in section 5.1.9.

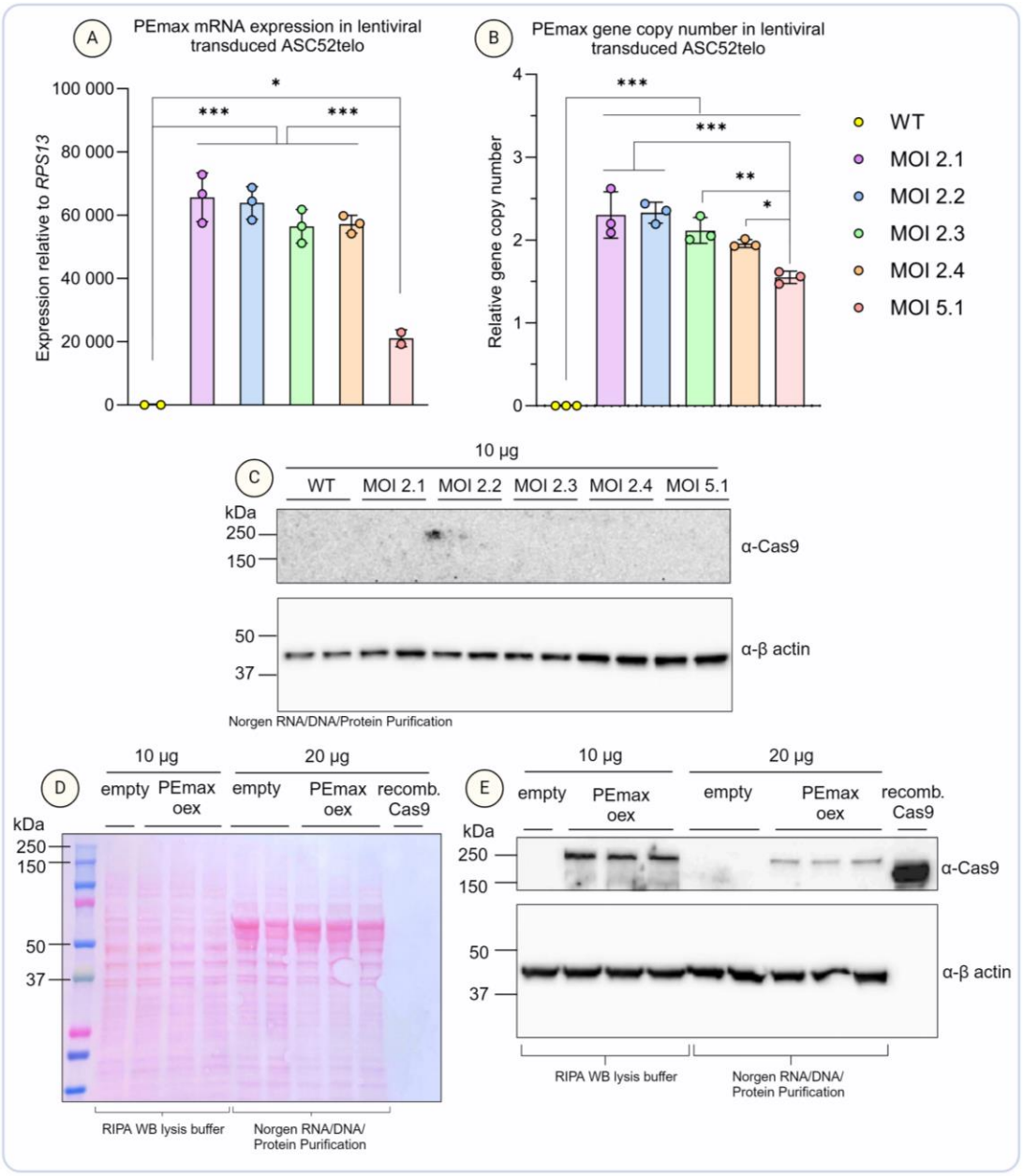


Figure 4.20: Validation of PEmax expression in prime editing (PE) -transduced ASC52telo. (A) qPCR on mRNA isolated from expanded heterogeneous mixes of PE-transduced ASC52telo cells. PEmax expression was calculated using the ΔΔCt method relative to the reference gene *RPS13*. The data are shown as ± SD, with n = 3 biological replicates. **(B)** qPCR on gDNA isolated from expanded heterogeneous mixes of PE-transduced ASC52telo cells. The approximate gene copy number (GCN) was calculated relative to a genomic locus on chromosome 16 as reference gene. The x-axis denotes the cell culture, while the y-axis denotes the relative GCN. The data are shown as ± SD, with n = 3 biological replicates. **(C)** Western Blot on 10 µg protein isolated from ASC52telo transduced with lentivirus carrying PEmax using the Norgen DNA/RNA/Protein kit. The membrane was stained with antibodies anti-Cas9 and anti-beta-actin (diluted 1:1000 and 1:5000 (respectively) in 3% BSA (w/v, f.c.) in PBST) over night. **(D-E)** Transient overexpression of PEmax in HT1080 cells. The cells were lysed in either RIPA WB lysis buffer or buffer SKP from the Norgen RNA/DNA/Protein kit with

10 µg and 20 µg protein loaded, respectively. In addition, 200 ng recombinant Cas9 protein was included as a positive control. **(D)** Ponceau staining of the membrane. **(E)** Western blot of the same membrane, stained with anti-Cas9 and anti-beta-actin (diluted 1:1000 and 1:5000 (respectively) in 3% BSA (w/v, f.c.) in PBST) over night. * p_{adj} <0.033, ** p_{adj} <0.002, *** p_{adj}, <0.001, based on 2-way ANOVA with Holm-Šídák correction for multiple testing. Figure created with GraphPad Prism and BioRender.com.

4.3.6. Nucleofection of ASC52telo

With a plausible presence of PEmax in the expanded cultures, genome editing was attempted by introducing the remaining PE components (corresponding to PE4) by nucleofection. The cells were nucleofected according to the method described in sections 3.5.2 and 3.9.7, with pMax_GFP acting as a proxy marker of positive transfection. As there was no significant statistical difference between PEmax mRNA expression of the MOIs, MOI 2.4 was chosen for the genome editing as it was growing and proliferating well.

Following nucleofection, the cells were incubated for 48 hours before sorting as single cells by FACS to achieve single-cell colonies. During FACS, transfection efficiencies of the live single cell population ranged from 40-60% (Table 4.5). The gates were set based on cells transfected with pMax_GFP, and WT cells as a negative control (Figure 4.21). Samples transfected with pMax_GFP, peg A or C, and MLH1, were sorted as single cells in 96-well plates containing medium with added FGF (1 ng/ml, v/v, f.c.) and EGF (10 ng/ml, v/v, f.c.). Prior to sorting, the nucleofected cells were observed in the microscope to have very high levels of cell death (data not shown), and as such only between 473-796 cells were successfully sorted for each sample. Based on previous experience of sorting ASC52telo as single cells by FACS (section 4.2.4), a very high survival rate was not expected. None of the sorted nucleofected cells survived.

Table 4.5 | Percentage of live, single, and GFP-positive cells sorted by FACS

Sample:	Freq. of Parent (%)	Live cells Freq. of Parent (%)	Live cells/Single Cells Freq. of Parent (%)	Live cells/Single Cells/GFP neg Freq. of Parent (%)	Live cells/Single Cells/GFP pos Freq. of Parent (%)
GFP negative control	100	91.2	90.8	99.1	0
pMax_GFP + MLH	100	59.6	87.4	51.2	38.4
pMax_GFP + Peg A + MLH	100	80.5	88.6	38.2	51.4
pMax_GFP + Peg C + MLH	100	47.1	88	42.1	46.2

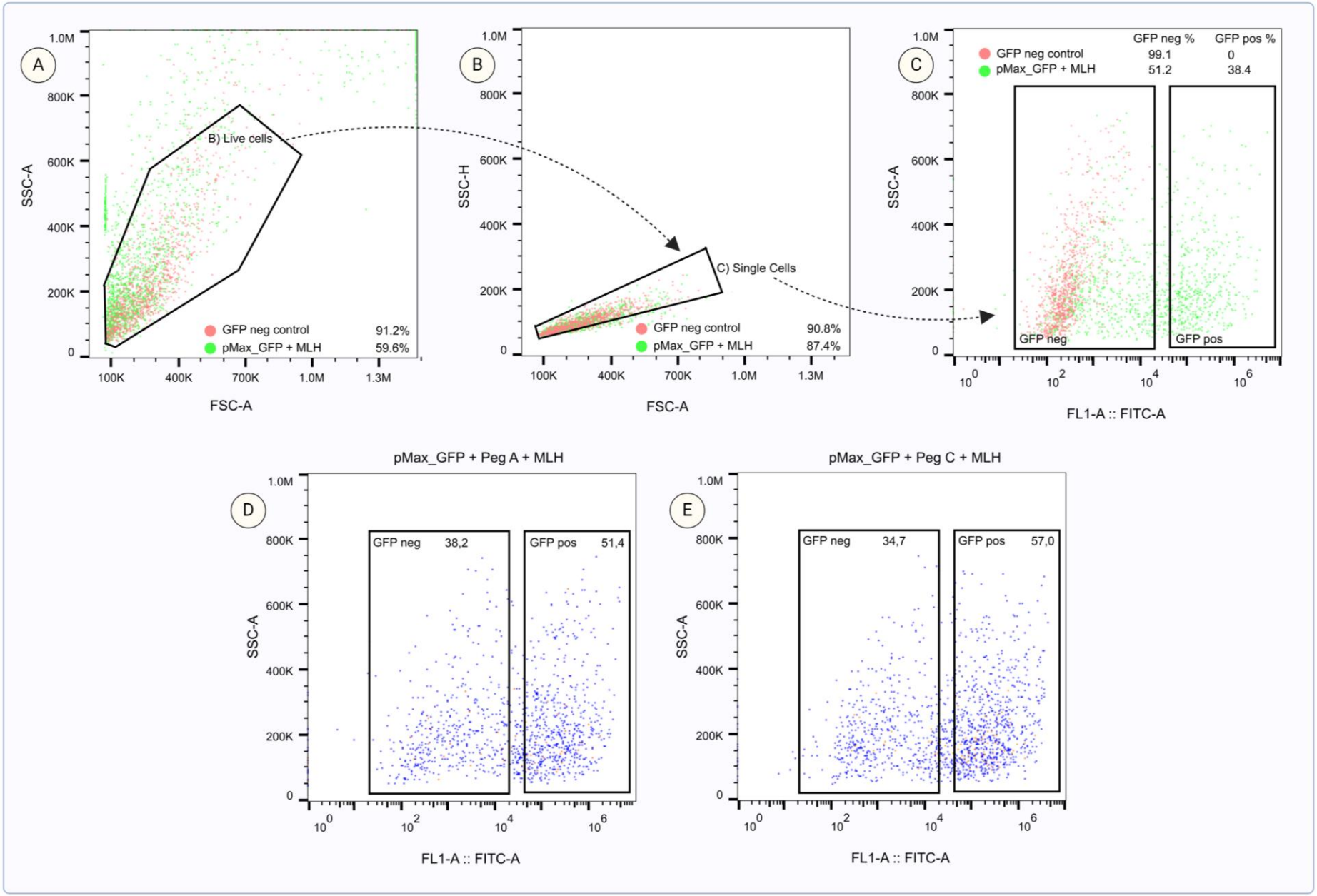

A
B) Live cells
GFP neg control 91.2%
pMax_GFP + MLH 59.6%
SSC-A
FSC-A
B
C) Single Cells
GFP neg control 90.8%
pMax_GFP + MLH 87.4%
SSC-H
FSC-A
C
GFP neg % GFP pos %
GFP neg control 99.1 0
pMax_GFP + MLH 51.2 38.4
GFP neg
GFP pos
SSC-A
FL1-A :: FITC-A
pMax_GFP + Peg A + MLH
D
GFP neg 38,2
GFP pos 51,4
SSC-A
FL1-A :: FITC-A
pMax_GFP + Peg C + MLH
E
GFP neg 34,7
GFP pos 57,0
SSC-A
FL1-A :: FITC-A

Figure 4.21: FACS on prime editing (PE) -transduced ASC52telo post nucleofection. The cells were transfected by nucleofection with plasmids according to the PE4 system. The cells were sorted with the SONY SH800S Cell Sorter. Figures (A), (B), and (C) show positive and negative controls for GFP, where the positive control was transfected with pMax_GFP and mismatch repair inhibitor (MLH). **(A)** The figure shows the whole cell population of both positive and negative controls with the gating set on forward scatter area (FSC-A) against side scatter area (SSC-A) to get an estimation of the main population of live cells, assess granularity and to filter out debris. **(B)** The live cell population was shown in a new plot as FSC-A against forward scatter height (FSC-H) and the second gate (single cells) was set diagonally to exclude doublets. **(C)** The single cells were shown in a new plot as FL1-A (GFP) against SSC-A and. The negative and positive controls were used to set gates to positive cells accordingly. **(D)** The figure shows cells transfected with pMax_GFP, pegRNA encoding the A-allele for rs1799993, and MLH. **(E)** The figure shows cells transfected with pMax_GFP, pegRNA encoding the C-allele for rs1799993, and MLH. Figure created with FlowJo and BioRender.com.

5. Discussion

5.1. General and methodological discussion

This thesis investigated prime editing (PE) as a step in translating epidemiological data of complex diseases into underlying biological processes. Specifically, PE of predicted causal variants can both evaluate the causality of the edited SNP as well as identify the downstream target gene(s), hence addressing two of the major challenges within functional genomics (Gallagher and Chen-Plotkin, 2018, Lappalainen and MacArthur, 2021). By gaining biological insights into new gene regulatory mechanisms, new therapies and even precision medicine may be developed. In this thesis, obesity, which is one of the most prevalent complex diseases globally (Abarca-Gómez et al., 2017), has been used as a model disease. However, the PE strategies used in this thesis are applicable to almost any complex disease for which GWAS data exists.

Two different strategies for PE were tested in this thesis. One involved transient delivery of the editing machinery and the other lentiviral delivery of PEmax (stably integrated into the genome of the target cell) and transient delivery of pegRNAs and supporting components. Using transient delivery, PE of rs1799993 from heterozygous for A and C to homozygous C (protective allele), but not to homozygous A (risk allele) was achieved in HT1080 cells. However, editing efficiency was low, at about 1%, and not all colonies were analyzed, and thus analyzing more colonies could potentially lead to identification of a colony homozygous for the risk allele as well. Regardless, the PE principle was successfully demonstrated. A clear advantage of the transient delivery strategy is that it could be performed relatively fast due to the ease of delivering the PE machinery into the target cells.

The ASC52telo cells, on the other hand, required a comprehensive and time-consuming lentiviral transduction strategy to deliver the PE enzyme. This approach is necessary for cells that are difficult to transfect, including AD-MSCs. An additional challenge with these cells was the inability to grow them as single colonies despite considerable efforts to modify the protocol. While this might eventually be achieved given enough time for optimization, it illustrates an important drawback of the clonal PE strategy. While a heterogenous population of transduced ASC52telo cells expressing the PEmax enzyme was achieved, the cells did not survive the subsequent nucleoporation of the pegRNA and FACS-sorting required for genome editing to take place. Due to time restrictions, this nucleoporation method was not further optimized, but is likely solvable. However, due to the expected low editing efficiency, single-cell colonies will still be needed to clearly demonstrate an effect of the variant editing on cell function. Nonetheless, the variation between cell types in editing efficiency and single-cell expansion implies that success using PE as a strategy depends largely on the cell type to be investigated.

5.1.1. Prime editing as a tool to experimentally validate likely causal SNPs

While computational methods can give valuable clues and predictions regarding which SNP(s) in a disease-associated locus is causal (fine mapping, epigenetic screens, etc.), as well as well-educated guesses of downstream target genes (e.g., eQTL analyses), experimental testing is required to verify or falsify these predictions (Gallagher and Chen-Plotkin, 2018, Claussnitzer et al., 2020, Lappalainen and MacArthur, 2021). Luciferase reporter assays are an important experimental starting point but fall short in defining how the predicted causal variant works in the native biological network and genomic context and do not identify the specific target gene(s) but rather only an effect on transcriptional activity reflected by expression of the luciferase gene. In addition, genome editing, when successfully hitting causal variants, presents as a more physiologically relevant method that not only identifies direct target genes but also downstream affected pathways (such as by transcriptome profiling of edited vs. non-edited cells). Genome editing, PE included, has had a big part in addressing this issue, but this revolutionizing technology comes with many challenges, as demonstrated by the present project.

One of the issues experienced in this thesis was the low editing efficiency achieved by PE in HT1080. With only 1% of cells showing the intended edit, the editing efficiency was lower than expected compared to other studies reporting efficiencies laying generally between 10-30% but even reported to be as high as 65% (Anzalone et al., 2019, Chemello et al., 2021, Chen et al., 2021). Only 87 out of 162 colonies were sequenced, however, and if the remaining are sequenced, more successfully edited colonies might be discovered. There are several possible explanations for the low efficiency, including design of the pegRNAs, inclusion/exclusion of nicking guide RNA (PE3 and PE5 systems), inclusion/exclusion of the MMR inhibitor (PE4), other intrinsic repair mechanisms in the cell disfavoring the edited strand or the delivery method of the components. None of these explanations can be ruled out before the remainder of the colonies have been sequenced, and the experiment replicated with more colonies.

As mentioned in the introduction, there have been many improvements to the original PE system: PE2, which contains a mutated PE enzyme leading to increased editing efficiency; PE3, which includes a nicking guide RNA designed to nick the non-edited strand to favor incorporation of the edited strand; PE4, which is PE2 with co-expression of an MMR inhibitor; and PE5, which is PE3 with co-expression of an MMR inhibitor. Co-expression of the MMR inhibitor utilized in this thesis (MLH1) has been reported to increase prime editing efficiencies by sevenfold over PE2 alone and twofold over PE3 (Chen and Liu, 2023). The one colony that was successfully edited, C120, was transfected with the complete PE4 system, suggesting that repeating the experiment with more colonies all transfected with MLH1 could aid in improving the editing efficiency and understanding why it was low the first time. Designing nicking guide RNAs for the PE3 and PE5 systems might also improve the efficiencies.

Different delivery systems also affect the editing efficiency of PE. In cultures such as HeLa and HEK293T cells, transient lipid-mediated transfection and electroporation of plasmids encoding PE components can result in high efficiencies (Anzalone et al., 2019, Chen et al., 2021). Electroporation has also led to successful editing in iPSCs using the PE3 system (Chemello et al., 2021). In the case of HT1080, the FACS sorting of cells post-transfection of PE components show that an average of 34% of cells were sorted as GFP-positive, while the smaller pMax_GFP gave a live single cell transfection efficiency of 98%. This suggests a relatively high transfection efficiency for the PEmax-encoding plasmid when one considers how large it is in comparison to pMax_GFP. It also suggests that the delivery of the plasmid is not the main problem for the low editing efficiency. However, if stable transduction into the genome of PEmax were conducted (such as with ASC52telo) and the pegRNAs were later transfected transiently and at a higher concentration, the chance of a successful edit might increase, as more pegRNA plasmids have the chance to enter the cell.

Optimization of the pegRNA design is critical for maximizing prime editing efficiency, as different lengths of the PBS and RT template causes a wide variety of editing efficiencies (Anzalone et al., 2019). Several online tools and libraries have been developed where the performance of the pegRNAs can be accurately predicted (Chen and Liu, 2023). Screening the pegRNAs utilized in this thesis with more tools could aid in understanding their performance. For example, DeepPE by Kim et al. (2021) is a deep learning model trained on pooled pegRNA target site library screens in human cells. While transfection of the large PEmax-plasmid was seemingly not the issue, redesigning the pegRNAs and testing different options may be a step to improve the editing efficiency.

Another important consideration is the possibility of off-target effects. Traditional CRISPR/Cas relies on double-stranded breaks (DSBs) in the DNA, which are potentially related to severe off-target effects like large deletions, complex rearrangements, and apoptotic responses in the cell (Trerotola et al., 2015). PE reduces the risk of these effects as the nick only occurs on one strand. It therefore offers a much lower risk of off-target activity and fewer byproducts than traditional CRISPR/Cas. It even offers a similar efficiency to Cas9-initiated HDR (Anzalone et al., 2019), which is quite efficient at installing precise DNA changes by relying on exogenous donor DNA templates. HDR does, however, generate a lot of excess indels from end-joining repair of the DSBs (Anzalone et al., 2019). The risk of off-target effects might also vary between cell lines (Petrova and Smirnikhina, 2023). The introduction of the MMR inhibitor in PE4 and PE5 has also shown a reduced number of off-target effects, with PE5 reducing indel byproducts by twofold compared to PE3 (Chen and Liu, 2023). In the current thesis, off-target effects were not investigated. This will be a natural next step to explore once the editing efficiency of both HT1080 and ASC52telo is sufficient.

Taken together, numerous areas of improvement of PE and other genome editing tools exist, but there is immense potential for these to become more effective and applied to the development of diagnostic strategies

and precise medicine directly targeting the root of the cause, or to simply uncover the fundamentals of biology and thus further scientific knowledge.

5.1.2. HT1080 for proof-of-concept genome editing

The HT1080 cell line was selected due to its convenient handling, transfection capabilities, and diploid nature, making it a suitable candidate for genome editing. It was, however, later noticed that this cell line often exhibits pseudodiploidy, meaning that while the chromosome count may be accurate, chromosomal rearrangements result in an abnormal karyotype. According to the manufacturer, approximately 40% of the cells demonstrate rearranged karyotypes, including the presence of an additional E-group chromosome (chromosomes 16-18) and the absence of a C-group chromosome (chromosomes 6-12 and X). It is suspected that the missing chromosome from the C-group is chromosome 11 (ATCC, 2023), which poses a challenge for the project as the target SNP, rs1799993, is located on chromosome 11.

Furthermore, selecting reference genes and proteins for experiments such as qPCR and western blotting presents a challenge. To assess gene overexpression relative to a reference gene, a reference gene that accurately represents the entire cell population without being influenced by treatment is required. *RPS13*, for instance, resides on chromosome 11 (Martin et al., 2022), and was utilized as a reference gene for the overexpression of pLenti-PE2max-BSD in HT1080 cells (data not shown). *RPLP0*, located on chromosome 12, was included as a secondary reference gene, and was selected for the final ΔΔCt analysis. The difference in Ct values between *RPS13* and *RPLP0* was ~2 on average; however, it is uncertain whether this discrepancy is due to gene expression level, primer efficiency, or the potential absence of chromosome 11 in some cells. To address this uncertainty, standard curves for both primers and single-cell sequencing of the cell batch used in the experiment are necessary. Sequencing of the region surrounding rs1799993 in the cell line revealed positive results for the reverse primer, indicating the presence of at least a significant number of cells still carrying chromosome 11. Additionally, genotyping the single-cell colonies exhibited positive results for the region, suggesting that the cell batch in this experiment still contains chromosome 11.

Due to the reported chromosomal instability in HT1080 cells, there is also a possibility that the single-cell colony that was assumed successfully edited from heterozygous for A and C to homozygous for C, originated from a cell that had lost one chromosome 11 during cell division, and that in reality, the editing efficiency was 0%. It would therefore be beneficial to repeat the proof-of-concept experiment in a separate cell line.

5.1.3. gDNA isolation using Chelex 100

When expanding the HT1080 colonies, one question that arose was how to analyze them while saving time and resources. To estimate editing efficiency on a large number of colonies, it was necessary to develop a

protocol that negated the need for silica-based spin columns, as this would take too much time. Therefore, to isolate gDNA from a small number of cells and multiple samples simultaneously, Chelex 100 resin was used. The resin has been widely used in forensic DNA extraction protocols for its simplicity and speed in extracting DNA from minimal starting material (Walsh et al., 1991).

The gDNA extracted from the colonies was then analyzed by qPCR using SNP-specific primers to determine which colonies had been successfully edited before further sequencing. Initially, gDNA was extracted from colonies in 96-well plates, with 2 µl used as a template for each reaction. However, this yielded minimal to no results. Several more attempts were made before expanding the colonies to 24-well plates, performing a second extraction, and diluting the samples 1:5 before using 4 µl as template. While this approach provided varied results across the samples, due to time constraints, no further attempts were made.

The Ct-values were used to determine whether the sample was likely homozygous for A or C, or ambiguous. The analysis suggested that colonies C21, C30, C48, C55, C59, C66, C107, C134, C142, and C148 might be homozygous for A or C, indicating a successful edit of the SNP. However, subsequent sequencing of these colonies revealed that they were all heterozygous, suggesting that the template used in the qPCR inhibited the reaction in some way which caused misleading results.

Chelex 100 works by chelating transition metal ions thus binds to the magnesium ions in the sample. Mg^+ are cofactors for several nucleases, and their binding to the resin bead renders them unavailable to the nucleases, which hinders DNA degradation (Singh et al., 2018). The addition of proteinase K aids in digesting proteins by hydrolyzing peptide bonds. However, if not inactivated prior to PCR, it may digest the polymerase and inhibit the reaction. In the protocol used in this thesis, proteinase K was allowed to digest the protein samples at 56°C for 4 hours, but no step was taken to inactivate it before PCR. Furthermore, carrying Chelex 100 beads into the reaction mix might cause them to bind to the Mg^+ needed by the polymerase. Therefore, optimizing this method would involve adding a step to inhibit proteinase K and ensuring no beads are carried over in the reaction mix. Developing such a protocol would aid in downstream applications of PE to streamline the process of screening numerous single-cell colonies without the need for time-consuming silica-based spin columns.

5.1.4. Choosing the optimal cell model

ASC52telo was specifically chosen as the model cell line for this project because ENCODE epigenetic data showed that rs1799993 is situated in a region with enhancer signal specifically in AD-MSC, and not in mature adipocytes (Samuelsen, 2021, Mirza, 2022). ASC52telo is a human AD-MSC line with the ability to differentiate into mature adipocytes, allowing for the study of adipose-related genes prior to, during, and after differentiation. This is especially important considering that the risk-allele has been linked to a potential increase in VAT (Karlsson et al., 2019). While a single cell line can offer important mechanistic insight,

validation in other cell lines is important to assess the universality of the initial finding. Thus, the experiments conducted in this thesis should be repeated in other AD-MSC cell lines. Moreover, screening for cell lines that are capable of clonal expansion from single cells could be essential to ensure a complete editing procedure can be carried out. While non-immortalized patient-derived AD-MSC would normally be closer to the *in vivo* setting than immortalized cell lines, such primary cells will quickly undergo senescence and functional decline with extended passaging (Lee et al., 2004), preventing their use for PE which involves extensive sub-passaging that non-immortalized cells would not survive. However, the ASC52telo cells were immortalized using the hTERT enzyme that preserves telomer lengths. This leads to immortalized cells that, unlike cells immortalized with the SV40 large T antigen, are physiologically very close to primary cells, while avoiding senescence and preserving the karyotype (Bodnar et al., 1998, Rambhatla et al., 2002, Lee et al., 2004, Yin et al., n.d.), making the ASC52telo cell line a good option as a cell model in the context of the experiments conducted in the present thesis.

Despite being immortalized using hTERT, two studies reported that ASC52telo cells have decreased differentiation potential and decreased insulin sensitivity compared to primary cells (Masnikov et al., 2021, Kulebyakin et al., 2021). While the differentiation media used in these studies differed slightly, the patient-derived non-immortalized cells differentiated better than the ASC52telo cells in both studies by these protocols. However, experiments conducted in our lab showed that the ASC52telo cells differentiate well in optimized medium and with a longer differentiation protocol (Jan-Inge Bjune, pers. com., unpublished). Thus, this cell model, once edited, should provide insight into the effect of rs1799993 in different stages of differentiation.

A potential limitation with ASC52telo as a cell model is the fact that it was isolated from SAT (Wolbank et al., 2009). However, changes in SAT can likely affect the expansion of VAT. For example, according to the "spillover theory", excess nutrients are first stored in SAT, but when the capacity of these cells is reached, lipids will be secreted from SAT and could be stored in other tissues such as VAT (Nelson et al., 2007, Koutsari et al., 2011, Koutsari et al., 2012, Abildgaard et al., 2021). Of note, eQTL data for rs1799993 gave signals in both VAT and SAT (Samuelsen, 2021). Still, performing the experiment in a VAT cell line would clarify whether the SNP primarily affects VAT tissues directly or indirectly via SAT. Unfortunately, VAT cell lines are hard to come by, and experience inefficient differentiation in vitro (Tallapragada, 2021).

5.1.5. Lentiviral transduction of ASC52telo

Previous comprehensive testing of transfection methods in the ASC52telo cell line confirms that this cell line is hard to transfect, with nucleofection being the method showing the most promise for transient delivery of plasmids (Krill, 2023). A pilot of lentiviral transduction on ASC52telo also suggests that, with further optimization of the lentiviral packaging protocol, the method shows promise for introducing the PE-plasmids

(Krill, 2023). Studies on transfection of MSCs generally reveal that transfection of MSCs is, in general, ineffective (McMahon et al., 2006, Oggu et al., 2017). Thus, lentiviral transduction is suggested as an alternative, with successful experiments reported (McMahon et al., 2006, Zielske et al., 2009, Lin et al., 2011, Lin et al., 2012, Quiroz-Reyes et al., 2023) Establishing a robust lentiviral protocol for ASC52telo gives future opportunities for the study of downstream processes these cells are involved in. Furthermore, establishing a cell line that stably expresses PEmax would allow for the editing of virtually any part of the genome, given the right pegRNAs.

Lentiviruses are a subtype of retroviruses whose genome consists of a single-stranded RNA that is reverse-transcribed into DNA and then integrated into the genome of the host cell. Unlike other retroviruses, lentiviruses can infect non-dividing cells, broadening their potential for use in research and gene therapy. Most lentiviral vectors are HIV-1-based. In the present thesis, the second-generation lentiviral packaging plasmid and third-generation (Dull et al., 1998) envelope and cargo plasmids were utilized. The packaging plasmid encodes structural and enzymatic proteins (GAG, POL, TAT, and REV), the envelope plasmid encodes the envelope protein, and the cargo-plasmid carries the transfer-gene(s). A major breakthrough was achieved when the parental HIV-1 envelope was substituted for the vesicular stomatitis virus G protein (VSV-G). This gave the vectors the ability to transduce a greater number of cell types and tissues, stabilizes the vector particles during centrifugation, and directing vector entry through an endocytic pathway (Aiken, 1997, Cockrell and Kafri, 2007). The cargo-plasmid encodes the 5’ and 3’ LTRs, which act as weak promoters(Alba et al., 2005, Nayerossadat et al., 2012). The U3 region of the 3’ LTR is deleted. During reverse transcription, the same deletion would be copied onto the 5’-LTR promoter region, which hinders replication of the virus after it has incorporated into the host cell. The HIV-1 ψ packaging signal is also fused to the 5’ LTR and is involved in regulating the process of packaging the RNA into the viral capsid during replication. The deletion in the 3’ LTR offers a self-inactivation of the vector, increasing the safety level (Dull et al., 1998). The cargo-plasmid also carry the rev response element (RRE), which is required for the transport and processing of the viral RNA. The development of further generations of lentiviral vector systems ensures a level of biosafety that makes the use of these vectors in research and medicine easier, safer, and more efficient.

There are several viral vectors available along with lentiviruses. These include, but are not limited to, retroviruses, adenoviruses and adeno-associated virus (AAVs). Standard retroviruses, called γ-retroviruses, infect cells by passing through the nuclear pores of mitotic cells, where they integrate their cargo linearly into the host cell genome (Nayerossadat et al., 2012). Adenoviral vectors can transfer large DNA particles into both dividing and non-dividing cells. However, as opposed to retroviruses and lentiviruses, the DNA they deposit into the host cells are not integrated, and as such the gene expression is short term (Nayerossadat et al., 2012). Producing AAV vectors is a complicated process, but several generations of AAVs have been developed in to

both make this process less complicated and make their capacity to carry cargo larger (Alba et al., 2005, Nayerossadat et al., 2012).

Wang et al. (2021) has developed a highly efficient PE delivery system using adenoviral vectors, demonstrating up to 90% efficiencies in HeLa cells. One of the major challenges with PE is the delivery of the large plasmids. Wang et al. (2021) therefore utilize fully viral gene-deleted adenoviral vectors (also called high-capacity adenoviral vectors or gutless adenoviral vectors) due to their large packaging capacity, high genetic stability, and efficient transduction of both dividing and quiescent cells which produces high titers (Ricobaraza et al., 2020). Grünewald et al. (2023) and Davis et al. (2024) showed, however, that the limited packaging capacity of traditional non-gutless AAVs can be overcome by dividing up the PE-enzyme and transducing the two nCas9 and RT sequences separately in their own AAV construct, while still achieving relatively efficient editing.

One of the most desirable traits of lentiviruses as gene vectors is their capability to permanently integrate their cargo into the host genome. However, an important question is how this integration affects the health of the infected cell. After the viral RNA genome goes through reverse transcription to make a cDNA copy, the cDNA interact with the viral-encoded integrase and other complexes to form pre-integration complexes. The integrase protein involved in this process is likely the main determinant of integration-site selection through tethering to cellular proteins bound near their preferred genomic region (Lewinski et al., 2006). Research suggests that lentiviruses tend to integrate into actively transcribed genes, with a potential preference for housekeeping genes (Schröder et al., 2002, Wang et al., 2007). When the viral DNA inserts into a gene it might alter its expression, either by reducing or increasing it. This could potentially lead to severe mutagenesis, rendering the cell incomparable to the control cell line, or unable to proliferate and thrive in culture.

In order to improve transduction efficiencies, an additional solution to aid transduction is often included. The most used is Polybrene, which is a cationic polymer that neutralizes the negative charges of the cell membrane, facilitating the attachment of retroviruses (lentivirus included) to the cell surface. It is noted that Polybrene may have adverse effects on proliferation and senescence of stem cells, MSCs in particular (Lin et al., 2011). RetroNectin is a recombinant human fibronectin fragment that promotes retroviral attachment to target cells. It has been reported that RetroNectin results in slightly higher transduction efficiency compared to Polybrene in T-cells (Kurachi et al., 2017). RetroNectin is, however, more expensive than Polybrene. During the viral titer test, there was no considerable observable difference in the viability of the transduced cells between the parallels coated with and without RetroNectin. There are therefore inconclusive results on whether or not RetroNectin truly aided in the transduction. However, based on previous studies citing that transduction efficiencies are greatly increased in the presence of a polycation or other aid such as RetroNectin (Lin et al.,

2011, Denning et al., 2013, Kurachi et al., 2017), the decision was made to include it in downstream transductions.

In the experimental setup for measuring viral titer, ASC52telo cells were subjected to varying concentrations of the virus, both with and without RetroNectin coating. 13 days after infection (11 days after introduction of BSD) the number of CFUs were roughly estimated based on the number of surviving cells. Based on these calculations, five different MOIs were estimated and utilized for the transduction. An MOI of 1 assumes one infectious unit per cell. However, a critical consideration with MOIs is their sensitivity to experimental setup, such as changes in volume and surface area. If the number of viral particles remain consistent, an increased volume in the dish leads to the particles being more dispersed, altering the likelihood of infection. Only viral particles in close proximity to the cell have the potential to infect it (Shabram and Aguilar-Cordova, 2000). Adding more viral particle suspension to increase the number of particles in the medium alters the MOI, but not necessarily the number of infectious events, as it also increases the depth, leaving many particles suspended too far away from the cells.

Furthermore, the IFU calculation assumes that most viral particles lack infectivity. When HEK293T cells produce viral particles, the resulting mix comprises complete, functional viral particles, nonfunctional (i.e. empty) particles, and free-floating capsid protein (p24) (Quinn and Rowe, 2020). Measurement of viral particles using Lenti-X™ GoStix involves quantifying both functional and nonfunctional particles by assessing the amount of p24 in the medium. However, during a functional titer test, the objective is to quantify infective particles. Yet, due to the aforementioned variables, estimating MOIs becomes a little like a guessing game. MOIs also become virtually unreproducible for other researchers without the knowledge of all experimental details. For instance, an MOI of 5 in one experiment may equate to an MOI of 10 in another due to difference in particle concentration, duration of infection, surface area of the target cells, volume of the medium added, constituent added to aid infection (Polybrene, RetroNectin, etc.), addition of a centrifugation step, and so on. Therefore, comparing the effects of different MOIs without considering these variables could yield inaccurate and misleading outcomes.

The approximated CFUs that were counted in the well represents the number of surviving cells after the transduction. CFU stands for colony forming units, and due to ASC52telo not forming colonies, it was estimated based on the number of surviving cells after the titer test. The survival rate or growth rate between the dilutions also varied greatly, making it even more difficult to get an accurate estimation. The CFU does also not take into account the effect of BSD on the cells. It is unknown whether the BSD-enzyme possesses a cumulative effect, meaning that cells that have been infected more or express more BSD-enzyme will have an advantage over the cells that have a lower expression or have only been infected once. In the titer test, the number of surviving cells was higher in the wells with higher concentrations of viral particles. There is

therefore reason to suggest that there is some sort of cumulative effect we are unaware of, or that the cells proliferate quicker when surrounded more by other cells. We therefore have no real estimate of how efficient the transduction itself was, but there was a significant number of surviving cells that we were able to expand into larger cultures. However, there are a number of steps to take to increase transduction efficiency, such as closing the gap between the virus and cells by pelleting the cells in U-shaped wells prior to transduction (Pirona et al., 2020). There are also a number of additives that are supposed to enhance transduction efficiencies by forming different complexes between the cells, between the cell and virus, or promoting interactions, as discussed above (Gouvarchin Ghaleh et al., 2020)

In short, there are many considerations to make when using lentiviral transduction on cell lines, and due to the many variables, these experiments are rarely accurately reproducible. However, the transduction achieved in the present thesis can be considered successful, as there was a great number of cells that survived the selection process which were expanded and cryopreserved for future experiments.

5.1.6. Single cell colony expansion in ASC52telo

Our experience in working with ASC52telo as a MSC model cell line is that they do not readily form single-cell colonies. The cells migrate on the plate, mixing with other cells from other colonies, and if seeded with a density that is too low, they fail to proliferate. The colonies that did form were not monitored or tested for heterogeneity, so it is unknown whether they originate from one cell or many. Therefore, the heterogeneous colonies that did form where cultured and expanded further, and tested for PEmax expression, GCN of PEmax, and protein was isolated to assess the presence of PEmax. They were subsequently named after their respective MOIs, as MOI 2.1, 2.2, 2.3, 2.4, and MOI 5.1.

Several studies have been published in which clonal expansion of MSC, even AC52telo, has been performed. Katz et al. (2020) used lentiviral transduction and subsequent clonal expansion in which the transduced cells were plated at low densities on 15 cm plates until colonies formed. Assumed single-cell colonies were transferred from the plate into a well of a 24-well plate. The number of cells originally seeded on the 15 cm plate was not reported. Whether they had trouble expanding the colonies or getting them to proliferate was also not reported. Pitrone et al. (2019) performed a colony-forming assay to determine ASC52telo's (and other cell lines) ability to form colonies before and after treatment as a way of assessing proliferation. After 14 days, the cells were fixed with formaldehyde and stained with crystal violet, and groups containing more than 50 cells were considered colonies. Matveeva et al. (2024) also performed a colony-forming unit assay with ASC52telo, seeding 100 cells/25 cm^2, before and after treatment to assess proliferation. None of the studies mentioned whether the cells were continuously monitored to make sure they originated from single cells. None of them used any growth-enhancing substances in the cell culture medium.

The colonies attempted grown in the present thesis were seeded at varying densities in 6-well, 10 cm, and 15 cm plates. Some cells were also seeded as single cells by FACS in 96-well plates. A large number of cells (upwards of 40 cells per plate) were continuously monitored in the microscope for signs of colony formation, but none of them formed colonies that did not mix with surrounding cells. Either the cells stopped proliferating after a certain time period, or the cells moved around so much that it was not possible to assess where the colony originated, or which cells belonged to it. When the cell density became too low, such as for MOIs 0.05, 0.1, and 1 once the untransduced cells died, the surviving cells did not proliferate. Trying to seed cells with a very low density after selection was done was also not successful, as the cells died seemingly due to not being in the presence of other cells. Conditioned medium was briefly added to MOIs 0.05, 0.1, and 1 to attempt to revive the remaining cells, but no proliferation was observed. Conditioned medium should be tested under controlled conditions on WT cells before a conclusion can be made on its effectiveness. The cells should also be attempted seeded out using limited dilution instead of FACS, as the FACS treatment might be harsh on the cells and cause lower viability. More testing using FACS using a larger number of WT untreated cells should be conducted. If very few cells in a population of ASC52telo have the ability to form colonies, there might not have been enough time and patience exhibited during clonal expansion of the transduced cells to observe them form colonies.

If it is so that the cells will not proliferate due to not being in the presence of other cells, then transwell cell culture dishes can also be attempted, where single cells can be seeded on top and multiple cells on the bottom, allowing the cells to communicate and exchange growth factors through a membrane that still keeps them separate.

One can argue that it is not necessary to expand single-cell colonies before the introduction of the pegRNAs - as long as the cells express PEmax, there is no need for them to be exactly the same. However, to improve editing efficiencies later, picking a colony with the highest PEmax expression would be beneficial. To ensure that the transduced cells were as similar to the WT cells as possible, a low GCN would also be beneficial. It would also ensure fewer differences between the colonies after the edit has been performed, as all the cells within the colony would have PEmax integrated in the same spot. Picking several colonies would serve as controls, to see whether the placement of PEmax in the genome affects the outcome. Even if single-cell colony expansion was not performed after transduction, it would still remain an issue after introduction of the pegRNAs. To screen the cells for the potential edit, single-cell colonies are necessary.

5.1.7. Gene expression and gene copy number

As both plasmids, pCMV-PEmax-P2A-GFP and pLenti-PE2max-BSD have the PEmax sequence separated by a P2A sequence from their selection sequences (GFP and BSD, respectively), the PEmax and selection sequences are transcribed together. 2A sequences (P2A included) cause ribosomal skipping, causing the

polypeptides to separate (Doleschall et al., 2022). This ensures that the GFP expression and BSD resistance stays constant along with the PEmax expression. The presence of a 2A-sequence is associated with the chance of ribosome drop-off, meaning that there is a chance that the whole sequence will not get translated. Ribosome drop-off is, however, positively associated with longer constructs, and as such some inefficient translation is expected (Karlsson et al., 2019). The PEmax enzyme is very large, and this might increase the chance of inefficient translation. This might explain the relatively low PEmax expression in the transduced cultures, as well as the inability to detect the protein on the WB. GFP intensity in the transfected HT1080 cells was also much lower in the PEmax-transfected cells than the pMax_GFP-transfected cells. This could also be explained by the polymerase having to transcribe a much longer sequence before reaching the GFP sequence on the plasmid (same with the ribosome during translation), causing the production of GFP to be much more inefficient in these cells compared to the cells transfected with pMax_GFP. The low GFP intensity was observed on both the Incucyte analysis and the FACS-results.

To estimate the GCN of incorporated PEmax sequences, the equation described by Doleschall et al. (2022) was utilized. By utilizing a reference genomic region that occurs twice in a diploid genome, the GCN can be calculated for a gene with an unknown number of integrations, for example, after lentiviral transduction. However, other methods exist to determine the copy number of integrated viruses in cell culture. The Lenti-X Provirus Quantification Kit by Takara Bio utilizes qPCR on gDNA in a similar manner. The primers included in this are designed to amplify a conserved region of the HIV-1 genome adjacent to the packaging signal which are compatible with most commonly used lentiviral vectors. Re-doing the experiment with numerous methods will help to understand the accuracy of the one utilized in the current thesis.

5.1.8. qPCR, reference genes, primer design, and sequencing

Apart from the importance of template quality for successful qPCR, as discussed above, there are several other crucial factors to consider to ensure a high-quality and reproducible experiment. qPCR experiments are challenging to replicate due to various variables, including differences in nucleic acid extraction methods (both mRNA and gDNA), reverse transcription synthesis (for mRNA), primer efficiencies (for both target and reference), pipetting errors, and the choice of reference genes (Kozera and Rapacz, 2013). When quantifying mRNA expression levels, selecting a reference gene with minimal expression variability under experimental conditions and across different tissues and physiological states is crucial. While housekeeping genes involved in basic cell functions are often chosen as reference genes, no single gene is perfect for all experiments. Using multiple reference genes can therefore help to reduce errors and improve resolution.

In this thesis, only one reference gene was used to calculate mRNA expression in both HT1080 and ASC52telo. To test whether the pLenti-PE2max-BSD plasmid expressed PEmax, it was transfected into HT1080 cells and compared to *RPLP0* as a reference gene. The primers used to target PEmax (AK63F and AK64R) after the

overexpression in HT1080 have one mismatch towards the 5' end compared to the PE plasmids utilized in the present thesis, potentially affecting primer efficiency, and leading to inaccurate results. Therefore, new primers were designed specifically for analyzing gDNA and mRNA isolated from PEmax-transduced ASC52telo cells.

When genotyping the region surrounding rs1799993 in HT1080, the forward primer (HMBS fwd) did not provide a readable sequence beyond a certain point. The primer lies upstream of a poly-A region, leading to slippage in the sequencing results. The slippage occurs approximately 300 bp before the location of the SNP. The sequencing up until this region is of high quality, and the reverse primer (HMBS rev) provided readable results with no slippage. Designing a forward primer that extends across the poly-A region and anchors a few nucleotides on the 3' end might address this issue and provide a readable sequence (Grayburn and Sims, 1998).

5.1.9. Protein harvest methods and WB

To check if the transduced ASC52telo cells produced PEmax, three methods were used: qPCR to estimate the GCN from gDNA, qPCR to measure mRNA expression, and WB to detect the PEmax protein. The Norgen RNA/DNA/Protein purification kit was used to collect gDNA, mRNA, and protein samples. In each well of the WB, 10 µg of protein was added. Initially, we tried staining the membrane with anti-Cas9 antibody diluted at 1:1000, but no bands could be seen. So, a second attempt was made with a 1:500, but this resulted in a lot of background noise and still no bands (data not shown).

To troubleshoot, an overexpression of PEmax using the pLenti-PE2max-BSD plasmid was made in HT1080 cells and the protein was harvested using both the Norgen kit and RIPA WB lysis buffer. The maximum volume of each method was loaded, resulting in 10 µg for the RIPA samples and 20 µg for the Norgen samples. A positive control of 200 ng recombinant Cas9 was also included. When the membrane was subsequently stained with the same antibody using a 1:1000 dilution, weak bands could be observed on the Norgen samples and slightly stronger bands on the RIPA samples. The positive control showed a very strong band, confirming that the antibody works. However, when we stained the membrane with Ponceau S staining solution, the membrane showed strong bands for mid-size proteins and weaker bands towards the top of the membrane for the Norgen samples. In contrast, the RIPA samples showed more evenly distributed band strength. This suggests that even though twice as much protein was loaded for the Norgen samples, the purification method might filter out the larger proteins, including PEmax. To address this, different lysing methods should be tried on the transduced ASC52telo cell lines, such as RIPA WB lysis buffer. Wang et al. (2021) showed success using the same antibody to detect PE2 by lysing the cells in Laemmli buffer and performing all blocking and staining steps with TBST instead of PBST. Testing out different Cas9 antibodies is also an option. By changing one factor at a time, a reliable protocol for detecting the PEmax protein in the transduced cells could be developed.

5.2. Conclusions

The main goals of the current thesis were to validate the PE method in HT1080 cells and develop a PE protocol that can be utilized in AD-MSCs to decode the underlying biology of the SNP rs1799993 in relation to VAT. During the course of this thesis, the PE principle was demonstrated in the HT1080 cell line using the PE4 system. The SNP rs1799993 was successfully edited from heterozygous for A and C to homozygous for C. Furthermore, a lentiviral protocol for transduction of the AD-MSC cell line ASC52telo was successfully developed, and heterogeneous cell lines expressing the PEmax enzyme were expanded. Finally, while pegRNAs were transfected into the cell lines stably expressing PEmax, these cells did not survive the subsequent FACS sorting. Thus, rs1799993 was not edited in these cells. Specific achievements include:

- The HT1080 cell line was genotyped for rs1799993 and found it to be heterozygous for A and C.
- A transfection protocol for PE4 in HT1080 was established.
- FACS sorting of GFP-positive cells and single-cell colony expansion of HT1080 was optimized.
- 162 HT1080 single-cell colonies were successfully expanded.
- One HT1080 colony was successfully edited from heterozygous for rs1799993 (A and C) to homozygous for C using the PE technique.
- A lentiviral approach to deliver the PEmax construct into the genome of ASC52telo was adapted.
- Cell lines of ASC52telo that stably expresses PEmax mRNA were expanded.

Overall, this thesis has demonstrated that rs1799993 is targetable by PE and has progressed the development of the PE technique in ASC52telo cells.

6. Future perspectives

Several aspects of the research conducted in the present thesis were limited by time constraints, leaving important areas for future exploration. One key area is continuation of the development of a high-throughput screening method for the edited colonies following single-cell colony expansion. Due to the repeated attempts at obtaining results from the qPCR using SNP-specific primers for rs1799993 to screen the HT1080 colonies for homozygous for A or C, no time was left to sequence all the expanded colonies. Currently, 87 colonies have been sequenced, but the remaining 75 colonies still need to be analyzed to better understand the editing efficiency. Therefore, continue to develop a high-throughput screening method for a large number of colonies and sequencing the remaining 75 colonies are natural steps to take. Developing such a protocol would also aid to streamline downstream applications of PE to explore the underlying biology of other SNPs.

Due to time restrictions, no further steps were conducted past nucleofection of transduced ASC52telo. The challenges that were encountered during single-cell colony expansion of the transduced cells led to the

expectation that colonies would not form after nucleofection. Indeed, all the cells died after FACS sorting, but the exact cause – whether FACS treatment, nucleofection, or the inability for the cells to proliferate alone – remains unclear. Future work should involve controlled experiments to isolate each factor. Numerous approaches can be taken. An alternative includes pooling the cells instead of seeding them directly as single cells during FACS. Testing both WT cells and transduced ASC52telo will also rule out whether the transduced cells have a lower capability of colony formation. Utilizing transwell cell culture dishes to allow single cells to communicate with other cells might also improve outcomes. As there have been reports of ASC52telo successfully forming colonies, contacting other researchers that have performed these experiments to share experiences will also be a natural next step.

Optimizing nucleofection of pegRNAs in transduced ASC52telo cells before perfecting single-cell colony expansion is another potential step that can be taken. By pooling the GFP-positive cells after FACS sorting instead of sorting them directly as single cells in 96-well plates could allow them to recover, and subsequent single-cell sequencing can also be conducted on the pooled culture to assess editing efficiency.

Given that ASC52telo is an AD-MSC isolated from SAT, extending PE to both SAT and VAT samples from the same patient would provide a clearer understanding of whether the SNP affects one or both cell types. Our collaborators on this project are currently working on this approach, and the samples are expected to arrive in the near future. Additionally, continuing the work of Tallapragada (2021) in developing an AD-MSC line from VAT is of interest, although it should not be prioritized until the aforementioned problems are resolved.

As previously discussed, exploring AAV-mediated delivery for co-transduction of nCas9, RT, and pegRNAs simultaneously is another potential step. However, this is a more comprehensive experiment, and requires a lot of preparation. Therefore, the focus should remain on the already established lentiviral protocol that has been demonstrated successfully in the present thesis.

Taken together, exploring further development of robust PE protocols will aid our lab in the present project, as well as other researchers and downstream projects to translate epidemiological data of complex diseases into underlying biological processes. As each step becomes more efficient, the underlying biological and gene regulatory mechanisms diseases can be uncovered at a quicker rate than before, and by sharing these findings with other researchers, new therapies and even precision medicine may be developed. In the present thesis, the focus was on a single nucleotide variant related to obesity. However, the PE strategies used in this thesis are applicable to almost any complex disease for which GWAS data exists.

7. References

Abarca-Gómez, L., Abdeen, Z. A., Hamid, Z. A., Abu-Rmeileh, N. M., Acosta-Cazares, B., Acuin, C., Adams, R. J., Aekplakorn, W., Afsana, K., Aguilar-Salinas, C. A., 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. *The Lancet,* 390**,** 2627-2642.

Abildgaard, J., Ploug, T., Al-Saoudi, E., Wagner, T., Thomsen, C., Ewertsen, C., Bzorek, M., Pedersen, B. K., Pedersen, A. T. & Lindegaard, B. 2021. Changes in abdominal subcutaneous adipose tissue phenotype following menopause is associated with increased visceral fat mass. *Sci Rep,* 11**,** 14750.

Aguet, F., Brown, A. A., Castel, S. E., Davis, J. R., He, Y., Jo, B., Mohammadi, P., Park, Y., Parsana, P., Segrè, A. V., et al. 2017. Genetic effects on gene expression across human tissues. *Nature,* 550**,** 204-213.

Aiken, C. 1997. Pseudotyping human immunodeficiency virus type 1 (HIV-1) by the glycoprotein of vesicular stomatitis virus targets HIV-1 entry to an endocytic pathway and suppresses both the requirement for Nef and the sensitivity to cyclosporin A. *J Virol,* 71**,** 5871-7.

Alba, R., Bosch, A. & Chillon, M. 2005. Gutless adenovirus: Last-generation adenovirus for gene therapy. *Gene therapy,* 12 Suppl 1**,** S18-27.

Altshuler, D., Donnelly, P. & The International HapMap, C. 2005. A haplotype map of the human genome. *Nature,* 437**,** 1299-1320.

Anzalone, A. V., Koblan, L. W. & Liu, D. R. 2020. Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nature Biotechnology,* 38**,** 824-844.

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.

ATCC. 2023. *HT-1080 [HT1080] CCL-121* ™ [Online]. ATCC. Available: https://www.atcc.org/products/ccl-121#required-products [Accessed 09.04 2024].

Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., et al. 2015. A global reference for human genetic variation. *Nature,* 526**,** 68-74.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A. & Horvath, P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science,* 315**,** 1709-12.

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

Bick, D., Bick, S. L., Dimmock, D. P., Fowler, T. A., Caulfield, M. J. & Scott, R. H. 2021. An online compendium of treatable genetic disorders. *Am J Med Genet C Semin Med Genet,* 187**,** 48-54.

Bjune, J.-I., Laber, S., Lawrence-Archer, L., Zhao, X., Yamada, S., Al-Sharabi, N., Mustafa, K., Njølstad, P. R., Claussnitzer, M., Cox, R. D., et al. 2023. Mechanisms of the FTO locus association with obesity: Irx3 controls a sumoylation-dependent switch between adipogenesis and osteogenesis. *bioRxiv*.

Bodnar, A. G., Ouellette, M., Frolkis, M., Holt, S. E., Chiu, C.-P., Morin, G. B., Harley, C. B., Shay, J. W., Lichtsteiner, S. & Wright, W. E. 1998. Extension of Life-Span by Introduction of Telomerase into Normal Human Cells. *Science,* 279**,** 349-352.

Chemello, F., Chai, A. C., Li, H., Rodriguez-Caycedo, C., Sanchez-Ortiz, E., Atmanli, A., Mireault, A. A., Liu, N., Bassel-Duby, R. & Olson, E. N. 2021. Precise correction of Duchenne muscular dystrophy exon deletion mutations by base and prime editing. *Sci Adv,* 7.

Chen, P. J., Hussmann, J. A., Yan, J., Knipping, F., Ravisankar, P., Chen, P. F., Chen, C., Nelson, J. W., Newby, G. A., Sahin, M., et al. 2021. Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell,* 184**,** 5635-5652.

Chen, P. J. & Liu, D. R. 2023. Prime editing for precise and highly versatile genome manipulation. *Nature Reviews Genetics,* 24**,** 161-177.

Chong, Z. X., Yeap, S. K. & Ho, W. Y. 2021. Transfection types, methods and strategies: a technical review. *PeerJ,* 9**,** e11165.

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,* 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.

Cockrell, A. S. & Kafri, T. 2007. Gene delivery by lentivirus vectors. *Molecular Biotechnology,* 36**,** 184-204.

Collins, F. S., Doudna, J. A., Lander, E. S. & Rotimi, C. N. 2021. Human Molecular Genetics and Genomics - Important Advances and Exciting Possibilities. *N Engl J Med,* 384**,** 1-4.

Daly, M. J., Rioux, J. D., Schaffner, S. F., Hudson, T. J. & Lander, E. S. 2001. High-resolution haplotype structure in the human genome. *Nature Genetics,* 29**,** 229-232.

Davis, J. R., Banskota, S., Levy, J. M., Newby, G. A., Wang, X., Anzalone, A. V., Nelson, A. T., Chen, P. J., Hennes, A. D., An, M., et al. 2024. Efficient prime editing in mouse brain, liver and heart with dual AAVs. *Nature Biotechnology,* 42**,** 253-264.

Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. 2002. Capturing chromosome conformation. *Science,* 295**,** 1306-11.

Denning, W., Das, S., Guo, S., Xu, J., Kappes, J. C. & Hel, Z. 2013. Optimization of the transductional efficiency of lentiviral vectors: effect of sera and polycations. *Mol Biotechnol,* 53**,** 308-14.

Doane, A. S. & Elemento, O. 2017. Regulatory elements in molecular networks. *Wiley Interdiscip Rev Syst Biol Med,* 9.

Doleschall, M., Darvasi, O., Herold, Z., Doleschall, Z., Nyirő, G., Somogyi, A., Igaz, P. & Patócs, A. 2022. Quantitative PCR from human genomic DNA: The determination of gene copy numbers for congenital adrenal hyperplasia and RCCX copy number variation. *PLoS One,* 17.

Dull, T., Zufferey, R., Kelly, M., Mandel, R. J., Nguyen, M., Trono, D. & Naldini, L. 1998. A third-generation lentivirus vector with a conditional packaging system. *J Virol,* 72**,** 8463-71.

Edraki, A., Mir, A., Ibraheim, R., Gainetdinov, I., Yoon, Y., Song, C. Q., Cao, Y., Gallant, J., Xue, W., Rivera-Pérez, J. A., et al. 2019. A Compact, High-Accuracy Cas9 with a Dinucleotide PAM for In Vivo Genome Editing. *Mol Cell,* 73**,** 714-726.e4.

ENCODE 2012. The ENCODE Project Consortium: An integrated encyclopedia of DNA elements in the human genome. *Nature,* 489**,** 57-74.

Farh, K. K.-H., Marson, A., Zhu, J., Kleinewietfeld, M., Housley, W. J., Beik, S., Shoresh, N., Whitton, H., Ryan, R. J. H., Shishkin, A. A., et al. 2015. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature,* 518**,** 337-343.

Ferreira, C. B., Sumner, R. P., Rodriguez-Plata, M. T., Rasaiyaah, J., Milne, R. S., Thrasher, A. J., Qasim, W. & Towers, G. J. 2020. Lentiviral Vector Production Titer Is Not Limited in HEK293T by Induced Intracellular Innate Immunity. *Mol Ther Methods Clin Dev,* 17**,** 209-219.

Freedman, M. L., Monteiro, A. N., Gayther, S. A., Coetzee, G. A., Risch, A., Plass, C., Casey, G., De Biasi, M., Carlson, C., Duggan, D., et al. 2011. Principles for the post-GWAS functional characterization of cancer risk loci. *Nat Genet,* 43**,** 513-8.

Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M., et al. 2002. The structure of haplotype blocks in the human genome. *Science,* 296**,** 2225-9.

Gallagher, M. D. & Chen-Plotkin, A. S. 2018. The Post-GWAS Era: From Association to Function. *Am J Hum Genet,* 102**,** 717-730.

Gibbs, R. A., Belmont, J. W., Hardenbol, P., Willis, T. D., Yu, F., Yang, H., Ch'ang, L.-Y., Huang, W., Liu, B., Shen, Y., et al. 2003. The International HapMap Project. *Nature,* 426**,** 789-796.

Gouvarchin Ghaleh, H. E., Bolandian, M., Dorostkar, R., Jafari, A. & Pour, M. F. 2020. Concise review on optimized methods in production and transduction of lentiviral vectors in order to facilitate immunotherapy and gene therapy. *Biomedicine & Pharmacotherapy,* 128**,** 110276.

Grayburn, W. S. & Sims, T. L. 1998. Anchored Oligo(dT) Primers for Automated Dye Terminator DNA Sequencing. *BioTechniques,* 25**,** 340-346.

Grünewald, J., Miller, B. R., Szalay, R. N., Cabeceiras, P. K., Woodilla, C. J., Holtz, E. J. B., Petri, K. & Joung, J. K. 2023. Engineered CRISPR prime editors with compact, untethered reverse transcriptases. *Nature Biotechnology,* 41**,** 337-343.

Harvey, I., Boudreau, A. & Stephens, J. M. 2020. Adipose tissue in health and disease. *Open Biol,* 10.

Hille, F., Richter, H., Wong, S. P., Bratovič, M., Ressel, S. & Charpentier, E. 2018. The Biology of CRISPR-Cas: Backward and Forward. *Cell,* 172**,** 1239-1259.

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

Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. 2018. CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response. *Nature Medicine,* 24**,** 927-930.

Ihry, R. J., Worringer, K. A., Salick, M. R., Frias, E., Ho, D., Theriault, K., Kommineni, S., Chen, J., Sondey, M., Ye, C., et al. 2018. p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. *Nature Medicine,* 24**,** 939-946.

Inoue, F. & Ahituv, N. 2015. Decoding enhancers using massively parallel reporter assays. *Genomics,* 106**,** 159-164.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J Bacteriol,* 169**,** 5429-5433.

Jaenisch, R. & Bird, A. 2003. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genetics,* 33**,** 245-254.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science,* 337**,** 816-21.

Kaisanlahti, A. & Glumoff, T. 2019. Browning of white fat: agents and implications for beige adipose tissue to type 2 diabetes. *J Physiol Biochem,* 75**,** 1-10.

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.

Katz, D. B., Huynh, N. P. T., Savadipour, A., Palte, I. & Guilak, F. 2020. An immortalized human adipose-derived stem cell line with highly enhanced chondrogenic properties. *Biochem Biophys Res Commun,* 530**,** 252-258.

Keijzers, G., Bohr, V. A. & Rasmussen, L. J. 2015. Human exonuclease 1 (EXO1) activity characterization and its function on flap structures. *Biosci Rep,* 35.

Kim, H. K., Yu, G., Park, J., Min, S., Lee, S., Yoon, S. & Kim, H. H. 2021. Predicting the efficiency of prime editing guide RNAs in human cells. *Nature Biotechnology,* 39**,** 198-206.

Kosicki, M., Tomberg, K. & Bradley, A. 2018. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol,* 36**,** 765-771.

Koster, A., Murphy, R. A., Eiriksdottir, G., Aspelund, T., Sigurdsson, S., Lang, T. F., Gudnason, V., Launer, L. J. & Harris, T. B. 2015. Fat distribution and mortality: the AGES-Reykjavik Study. *Obesity (Silver Spring),* 23**,** 893-907.

Koutsari, C., Ali, A. H., Mundi, M. S. & Jensen, M. D. 2011. Storage of Circulating Free Fatty Acid in Adipose Tissue of Postabsorptive Humans: Quantitative Measures and Implications for Body Fat Distribution. *Diabetes,* 60**,** 2032-2040.

Koutsari, C., Mundi, M. S., Ali, A. H. & Jensen, M. D. 2012. Storage Rates of Circulating Free Fatty Acid Into Adipose Tissue During Eating or Walking in Humans. *Diabetes,* 61**,** 329-338.

Kozera, B. & Rapacz, M. 2013. Reference genes in real-time PCR. *Journal of Applied Genetics,* 54**,** 391-406.

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

Kulebyakin, K., Tyurin-Kuzmin, P., Efimenko, A., Voloshin, N., Kartoshkin, A., Karagyaur, M., Grigorieva, O., Novoseletskaya, E., Sysoeva, V., Makarevich, P., et al. 2021. Decreased Insulin Sensitivity in Telomerase-Immortalized Mesenchymal Stem Cells Affects Efficacy and Outcome of Adipogenic Differentiation in vitro. *Frontiers in Cell and Developmental Biology,* 9.

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.

Kurachi, M., Kurachi, J., Chen, Z., Johnson, J., Khan, O., Bengsch, B., Stelekati, E., Attanasio, J., McLane, L. M., Tomura, M., et al. 2017. Optimized retroviral transduction of mouse T cells for in vivo assessment of gene function. *Nat Protoc,* 12**,** 1980-1998.

Lange, L. A., Hu, Y., Zhang, H., Xue, C., Schmidt, E. M., Tang, Z. Z., Bizon, C., Lange, E. M., Smith, J. D., Turner, E. H., et al. 2014. Whole-exome sequencing identifies rare and low-frequency coding variants associated with LDL cholesterol. *Am J Hum Genet,* 94**,** 233-245.

Lappalainen, T. & MacArthur, D. G. 2021. From variant to function in human disease genetics. *Science,* 373**,** 1464-1468.

Lee, K. M., Choi, K. H. & Ouellette, M. M. 2004. Use of exogenous hTERT to immortalize primary human cells. *Cytotechnology,* 45**,** 33-38.

Lewinski, M. K., Yamashita, M., Emerman, M., Ciuffi, A., Marshall, H., Crawford, G., Collins, F., Shinn, P., Leipzig, J., Hannenhalli, S., et al. 2006. Retroviral DNA integration: viral and cellular determinants of target-site selection. *PLoS Pathog,* 2.

Li, X., Sun, B., Qian, H., Ma, J., Paolino, M. & Zhang, Z. 2022. A high-efficiency and versatile CRISPR/Cas9-mediated HDR-based biallelic editing system. *J Zhejiang Univ Sci B,* 23**,** 141-152.

Lin, P., Correa, D., Lin, Y. & Caplan, A. I. 2011. Polybrene inhibits human mesenchymal stem cell proliferation during lentiviral transduction. *PLoS One,* 6.

Lin, P., Lin, Y., Lennon, D. P., Correa, D., Schluchter, M. & Caplan, A. I. 2012. Efficient lentiviral transduction of human mesenchymal stem cells that preserves proliferation and differentiation capabilities. *Stem Cells Transl Med,* 1**,** 886-97.

Lin, Y. C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., et al. 2014. Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nat Commun,* 5.

Liu, Y., Kao, H. I. & Bambara, R. A. 2004. Flap endonuclease 1: a central component of DNA metabolism. *Annu Rev Biochem,* 73**,** 589-615.

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

Lvovs, D., Favorova, O. O. & Favorov, A. V. 2012. A Polygenic Approach to the Study of Polygenic Diseases. *Acta Naturae,* 4**,** 59-71.

Martin, F. J., Amode, M. R., Aneja, A., Austine-Orimoloye, O., Azov, Andrey G., Barnes, I., Becker, A., Bennett, R., Berry, A., Bhai, J., et al. 2022. Ensembl 2023. *Nucleic Acids Research,* 51**,** D933-D941.

Masnikov, D., Stafeev, I., Michurina, S., Zubkova, E., Mamontova, E., Ratner, E., Menshikov, M. & Parfyonova, Y. 2021. hTERT-immortalized adipose-derived stem cell line ASC52Telo demonstrates limited potential for adipose biology research. *Analytical Biochemistry,* 628.

Matveeva, D., Kashirina, D., Ezdakova, M., Larina, I., Buravkova, L. & Ratushnyy, A. 2024. Senescence-Associated Alterations in Matrisome of Mesenchymal Stem Cells. *International Journal of Molecular Sciences,* 25.

McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P. & Hirschhorn, J. N. 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet,* 9**,** 356-369.

McMahon, J. M., Conroy, S., Lyons, M., Greiser, U., O'shea, C., Strappe, P., Howard, L., Murphy, M., Barry, F. & O'brien, T. 2006. Gene Transfer into Rat Mesenchymal Stem Cells: A Comparative Study of Viral and Nonviral Vectors. *Stem Cells and Development,* 15**,** 87-96.

McVean, G. A., Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., et al. 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature,* 491**,** 56-65.

Merck. n.d. *Sigma Aldrich: Successful Transduction Using Lentivirus* [Online]. Sigma Aldrich. Available: https://www.sigmaaldrich.com/NO/en/technical-documents/technical-article/genomics/advanced-gene-editing/successful-transduction-lentivirus [Accessed 24.05 2024].

Merten, O. W., Hebben, M. & Bovolenta, C. 2016. Production of lentiviral vectors. *Mol Ther Methods Clin Dev,* 3.

Mika, A., Macaluso, F., Barone, R., Di Felice, V. & Sledzinski, T. 2019. Effect of Exercise on Fatty Acid Metabolism and Adipokine Secretion in Adipose Tissue. *Front Physiol,* 10.

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

Mota de Sá, P., Richard, A. J., Hang, H. & Stephens, J. M. 2017. Transcriptional Regulation of Adipogenesis. *Compr Physiol,* 7**,** 635-674.

Nayerossadat, N., Maedeh, T. & Ali, P. A. 2012. Viral and nonviral delivery systems for gene delivery. *Adv Biomed Res,* 1.

Nelson, R. H., Basu, R., Johnson, C. M., Rizza, R. A. & Miles, J. M. 2007. Splanchnic Spillover of Extracellular Lipase–Generated Fatty Acids in Overweight and Obese Humans. *Diabetes,* 56**,** 2878-2884.

NIH. 2023a. *Human Genomic Variation* [Online]. Available: https://www.genome.gov/about-genomics/educational-resources/fact-sheets/human-genomic-variation [Accessed 25.05 2024].

NIH. 2023b. *National Human Genome Research Institute: Complex Disease* [Online]. Available: https://www.genome.gov/genetics-glossary/Complex-Disease [Accessed 31.10 2023].

Oggu, G. S., Sasikumar, S., Reddy, N., Ella, K. K. R., Rao, C. M. & Bokara, K. K. 2017. Gene Delivery Approaches for Mesenchymal Stem Cell Therapy: Strategies to Increase Efficiency and Specificity. *Stem Cell Reviews and Reports,* 13**,** 725-740.

Pennacchio, L. A., Bickmore, W., Dean, A., Nobrega, M. A. & Bejerano, G. 2013. Enhancers: five essential questions. *Nat Rev Genet,* 14**,** 288-295.

Petrova, I. O. & Smirnikhina, S. A. 2023. The Development, Optimization and Future of Prime Editing. *Int J Mol Sci,* 24.

Pirona, A. C., Oktriani, R., Boettcher, M. & Hoheisel, J. D. 2020. Process for an efficient lentiviral cell transduction. *Biol Methods Protoc,* 5.

Pitrone, M., Pizzolanti, G., Coppola, A., Tomasello, L., Martorana, S., Pantuso, G. & Giordano, C. 2019. Knockdown of NANOG Reduces Cell Proliferation and Induces G0/G1 Cell Cycle Arrest in Human Adipose Stem Cells. *Int J Mol Sci,* 20.

Quinn, T. & Rowe, M. 2020. *Understanding viral titration—behind the science* [Online]. Takara Bio. Available: https://www.takarabio.com/about/bioview-blog/tips-and-troubleshooting/understanding-viral-titration-behind-the-science [Accessed 28.05 2024].

Quiroz-Reyes, A. G., Gonzalez-Villarreal, C. A., Limon-Flores, A. Y., Delgado-Gonzalez, P., Martinez-Rodriguez, H. G., Said-Fernandez, S. L., Soto-Dominguez, A., Rivas-Estilla, A. M., Islas, J. F., Molina-De la Garza, J. F., et al. 2023. Mesenchymal Stem Cells Genetically Modified by Lentivirus-Express Soluble TRAIL and Interleukin-12 Inhibit Growth and Reduced Metastasis-Relate Changes in Lymphoma Mice Model. *Biomedicines,* 11.

Rambhatla, L., Chiu, C. P., Glickman, R. D. & Rowe-Rendleman, C. 2002. In vitro differentiation capacity of telomerase immortalized human RPE cells. *Invest Ophthalmol Vis Sci,* 43**,** 1622-1630.

Rao, S., Yao, Y. & Bauer, D. E. 2021. Editing GWAS: experimental approaches to dissect and exploit disease-associated genetic variation. *Genome Medicine,* 13.

Rees, H. A. & Liu, D. R. 2018. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet,* 19**,** 770-788.

Richard, A. J., White, U., Elks, C. M. & Stephens, J. M. 2020. *Adipose Tissue: Physiology to Metabolic Dysfunction* [Online]. Endotext: South Dartmouth (MA): MDText.com, Inc. Available: https://www.ncbi.nlm.nih.gov/books/NBK555602/ [Accessed 20.03.2024].

Ricobaraza, A., Gonzalez-Aparicio, M., Mora-Jimenez, L., Lumbreras, S. & Hernandez-Alcoceba, R. 2020. High-Capacity Adenoviral Vectors: Expanding the Scope of Gene Therapy. *International Journal of Molecular Sciences,* 21.

Rosen, Evan D. & Spiegelman, Bruce M. 2014. What We Talk About When We Talk About Fat. *Cell,* 156**,** 20-44.

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

Schröder, A. R. W., Shinn, P., Chen, H., Berry, C., Ecker, J. R. & Bushman, F. 2002. HIV-1 Integration in the Human Genome Favors Active Genes and Local Hotspots. *Cell,* 110**,** 521-529.

Shabram, P. & Aguilar-Cordova, E. 2000. Multiplicity of Infection/Multiplicity of Confusion. *Molecular Therapy,* 2**,** 420-421.

Shlyueva, D., Stampfel, G. & Stark, A. 2014. Transcriptional enhancers: from properties to genome-wide predictions. *Nature Reviews Genetics,* 15**,** 272-286.

Singh, U. A., Kumari, M. & Iyengar, S. 2018. Method for improving the quality of genomic DNA obtained from minute quantities of tissue and blood samples using Chelex 100 resin. *Biol Proced Online,* 20.

Smemo, S., Tena, J. J., Kim, K.-H., Gamazon, E. R., Sakabe, N. J., Gómez-Marín, C., Aneas, I., Credidio, F. L., Sobreira, D. R., Wasserman, N. F., et al. 2014. Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature,* 507**,** 371-375.

Smith, G. I., Mittendorfer, B. & Klein, S. 2019. Metabolically healthy obesity: facts and fantasies. *J Clin Invest,* 129**,** 3978-3989.

Sollis, E., Mosaku, A., Abid, A., Buniello, A., Cerezo, M., Gil, L., Groza, T., Güneş, O., Hall, P., Hayhurst, J., et al. 2022. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Research,* 51**,** D977-D985.

Tallapragada, D. S. P. 2021. *Novel adipose genes in substrate- and depot-dependent lipid accumulation.* Degree of Philosophiae doctor (PhD), University of Bergen.

Trerotola, M., Relli, V., Simeone, P. & Alberti, S. 2015. Epigenetic inheritance and the missing heritability. *Hum Genomics,* 9.

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T. & Posthuma, D. 2021. Genome-wide association studies. *Nature Reviews Methods Primers,* 1**,** 59.

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

Wall, J. D. & Pritchard, J. K. 2003. Haplotype blocks and linkage disequilibrium in the human genome. *Nature Reviews Genetics,* 4**,** 587-597.

Walsh, P. S., Metzger, D. A. & Higuchi, R. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques,* 10**,** 506-513.

Walter, K., Min, J. L., Huang, J., Crooks, L., Memari, Y., McCarthy, S., Perry, J. R. B., Xu, C., Futema, M., Lawson, D., et al. 2015. The UK10K project identifies rare variants in health and disease. *Nature,* 526**,** 82-90.

Wang, G., Sarkar, A., Carbonetto, P. & Stephens, M. 2020. A Simple New Approach to Variable Selection in Regression, with Application to Genetic Fine Mapping. *Journal of the Royal Statistical Society Series B: Statistical Methodology,* 82**,** 1273-1300.

Wang, G. P., Ciuffi, A., Leipzig, J., Berry, C. C. & Bushman, F. D. 2007. HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Res,* 17**,** 1186-94.

Wang, Q., Liu, J., Janssen, J. M., Tasca, F., Mei, H. & Gonçalves, M. A. F. V. 2021. Broadening the reach and investigating the potential of prime editors through fully viral gene-deleted adenoviral vector delivery. *Nucleic Acids Research,* 49**,** 11986-12001.

Wang, S., Pan, M. H., Hung, W. L., Tung, Y. C. & Ho, C. T. 2019. From white to beige adipocytes: therapeutic potential of dietary molecules against obesity and their molecular mechanisms. *Food Funct,* 10**,** 1263-1279.

Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L., et al. 2014. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res,* 42**,** D1001-1006.

Wolbank, S., Stadler, G., Peterbauer, A., Gillich, A., Karbiener, M., Streubel, B., Wieser, M., Katinger, H., van Griensven, M., Redl, H., et al. 2009. Telomerase immortalized human amnion- and adipose-derived mesenchymal stem cells: maintenance of differentiation and immunomodulatory characteristics. *Tissue Eng Part A,* 15**,** 1843-1854.

Yan, J., Cirincione, A. & Adamson, B. 2020. Prime Editing: Precision Genome Editing by Reverse Transcription. *Mol Cell,* 77**,** 210-212.

Yin, D., Wells, J. A., Clinton, J. & Zou, C. n.d. Comparative analysis of cell proliferation, immunosuppressive action, and multi-lineage differentiation of immortalized MSC and MSC from bone marrow, adipose tissue, and umbilical cord blood. ATCC.

Zhang, M., Hu, T., Zhang, S. & Zhou, L. 2015. Associations of Different Adipose Tissue Depots with Insulin Resistance: A Systematic Review and Meta-analysis of Observational Studies. *Sci Rep,* 5.

Zhao, Z., Tavoosidana, G., Sjölinder, M., Göndör, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Singh Sandhu, K., Singh, U., et al. 2006. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nature Genetics,* 38**,** 1341-1347.

Zielske, S. P., Livant, D. L. & Lawrence, T. S. 2009. Radiation increases invasion of gene-modified mesenchymal stem cells into tumors. *Int J Radiat Oncol Biol Phys,* 75**,** 843-853.