

# **GENOME EDITING IN INDUCED PLURIPOTENT AND EMBRYONIC STEM CELLS**

## **MASTER THESIS**

submitted at the  
**IMC Fachhochschule Krems**  
**(University of Applied Sciences)**



**Master Program**  
***Medical and Pharmaceutical Biotechnology***

by

**Tanja ROTHGANGL, BSc**

for the award of the academic degree

**Master of Science in Engineering (MSc)**

Specialization: Advanced Therapeutics Development

**External Supervisor:** Laurence Daheron, PhD

**Internal Supervisor:** Prof.(FH) Dr. Barbara Entler

Date of Submission: July 2<sup>nd</sup> 2018

## I. Statutory Declaration

“I declare on my word of honor that I have written this paper on my own and that I have not used any sources or resources other than stated and that I have marked those passages and/or ideas that were either verbally or textually extracted from sources. This also applies to drawings, sketches, graphic representations as well as to sources from the internet.

The paper has not been submitted in this or similar form for assessment at any other domestic or foreign post-secondary educational institution and has not been published elsewhere. The present paper complies with the version submitted electronically.”

Date: July 2<sup>nd</sup> 2018

Tanja Rothgangl

Tanja Rothgangl

## II. Clause of confidentiality

### **SPERRVERMERK/CLAUSE OF CONFIDENTIALITY**

Auf Wunsch der Firma/Institution/*upon request of* Name der Firma/Institution, ist die vorliegende BA/MA Arbeit für die Dauer von maximal 5 Jahren für die öffentliche Nutzung zu sperren/will the present BA/MA thesis be retained from public access for the period of max.5 years

Dauer der Sperre / period: 3 Jahre / years

Veröffentlichung, Vervielfältigung und Einsichtnahme sind ohne ausdrückliche Genehmigung der o.a. Firma und des Verfassers/ der Verfasserin bis zum genannten Datum nicht gestattet/*Unauthorized reading, publication and duplication will not be allowed without explicit consent given by the above-mentioned company and the author before:*

zu veröffentlichen am/*publication allowed:* 02.07.2021

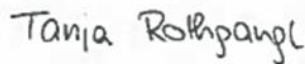
Unterschrift/*Signature:*



Name/Funktion. *Name/position:* Laurence Daheron, PhD / Head of HSCI iPS Core Facility

VerfasserIn Abschlussarbeit/*Author thesis:* Tanja Rothgangel, BSc

Unterschrift/*Signature:*



StudiengangsleiterIn /*Programme director:* Dr. Harald Hundsberger

Unterschrift/*Signature:*



Stempel FH/ *Seal:*



Bestätigt am: /*Notified as of:* 02.07.2018

### **III. Acknowledgements**

First and foremost, I would like Laurence Daheron, PhD for giving me the opportunity to conduct the experiments mentioned in this paper, her skilled supervision and her continual support throughout my time at the iPS Core. I am grateful for her enthusiasm and passion while mentoring me and I admire her broad organizational, technical and scientific skillset.

I would also like to thank Prof.(FH) Dr. Barbara Entler for helping me with any questions that came up during my time in Cambridge, supporting me when applying for my scholarships, and helping me with the visa process.

My sincerest gratitude to the biotech team at the IMC Krems. Every lecture of every course contributed to my preparation to conduct research at one of the most renowned universities in the world. I would like to thank Dr. Harald Hundsberger, who believed in me from the moment we met. Also, I won't forget the study service team who significantly helped organize my study period both in Krems and during my entire research semester.

I dearly want to thank all of my lab colleagues at the iPS Core facility: Caroline Becker, Laëtitia Pint, Mizuki Nagayama, Nikolaos Poulis, George Savvidis, Fatimah Alsalah, Ying Shao, Li Li, Kamar Reda and Nailia Mukhamedshina. They all made me feel very welcome and comfortable during my stay. The entire team was very supportive and helped me whenever I had questions. I especially want to thank Li Li for his supervision and guidance. Most of my work was done with his support and feedback.

I am also grateful for the financial support for my research provided by the Austrian Marshall Plan Foundation and the Julius Raab Stiftung.

Last but not least, I would like to thank my parents. They supported me all my life and always believed in me.

## IV. Abstract

Research in regenerative medicine and especially the use of induced pluripotent stem (iPS) cells has advanced greatly over the last couple of years. Alongside this, recent progress in genome editing and the ultimate combination of both fields provides a broad spectrum of medical applications. The ability to make precise changes in the genome of a eukaryotic cell broadened our understanding of human genetics and the study of certain genes and their contribution to disease phenotypes. Successful and reproducible genome editing in iPS and embryonic stem (ES) cells is a helpful tool to further study how genes are involved in certain pathologies. Using genome editing, specific genetic loci can be targeted resulting in a dysfunctional gene. Such knock-out (KO) experiments are already common practice when studying gene function. On the other hand, mutated cells, directly derived from a patient, can be reprogrammed into iPS cells and repaired using genome editing. The repaired cells can be differentiated and serve as isogenic controls for *in vitro* experiments. In theory, mutated cells that were derived from a patient could be repaired and eventually used for transplantation procedures. Such patient-derived isogenic cell transplants are a big goal within the regenerative medicine community. After reprogramming and genome editing steps, repaired cells can be differentiated into any cell type or tissue to replace the malfunctioning cells in the patient. Even though the differentiation procedure and the generation of clinical grade iPS is still a challenge, intensive research is currently ongoing to enable this goal.

In this paper, iPS cells and their derivation will be described as well as the current advances in genome editing. We studied the effects of diverse delivery methods to introduce genomic modifications in iPS and ES cells using the Cas9 nuclease. We established an editing protocol with which we successfully increased the knock-in (KI) efficiency compared to current protocols. Moreover, a detailed evaluation of editing events will be discussed in this paper.

## V. List of Abbreviations

5mCs	5-methylated cytosines
ASCs	Adult stem cells
BME	Basement Membrane Extract
BFP	Blue fluorescent protein
BF	Brightfield
CICs	Cancer initiating cells
CSCs	Cancer stem cells
Cas 9	CRISPR associated protein 9 nuclease
CPP	Cell Penetrating Peptide
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	corresponding RNA
CRISPRi	CRISPR Interference
dCas9	deactivated Cas9
DSMO	Dimethyl sulfoxide
DSBs	Double-strand breaks
DOX	Doxycycline
DMEM	Dulbecco's Modified Eagle's Medium
EGFP	enhanced GFP
ESCs	Embryonic stem cells
Epi	Episomal
EBNA-1	Epstein-Barr Nuclear Antigen-1
FBS	Fetal Bovine Serum

FACS	Fluorescence-activated cell sorting
F	Fusion protein
GFP	green fluorescent protein
HN	Hemagglutinin-Neuraminidase
HDR	Homology directed recombination
hiPSCs	human induced pluripotent stem cells
iPSCs	Induced pluripotent stem cells
ICE	Inference of CRISPR Editing
ICM	Inner cell mass
ITRs	Inverted terminal repeats
KI	Knock-in
KO	Knock-out
L	Large protein
LTRs	Long terminal repeats
M	Matrix protein
MN	Meganucleases
miRs	microRNAs
MUT	Mutant
NEB	New England Biolabs
NHEJ	Non-homologous end joining, classical (C-NHEJ), alternative (A-NHEJ)
NP	Nucleocapsid protein
PFA	Paraformaldehyde
PB	PiggyBac

PSCs	Pluripotent stem cells
PCR	Polymerase Chain Reaction
PAM	protospacer-adjacent motif
RVD	Repeat variable di-residue
RNAi	RNA interference
SeV	Sendai-viral
shRNA	short-hairpin RNA
sgRNA	single guide RNA
SNP	Single-nucleotide polymorphism
ssODNs	single-stranded oligodeoxynucleotides
siRNA	small-interfering RNA
SCNT	Somatic Cell Nuclear Transfer
SCs	Stem cells
TIDE	Tracking of Indels by Decomposition
tracrRNA	trans-activating crRNA
TALEN	Transcription Activator-Like Effector-based Nucleases
TFs	Transcription factors
WT	Wild type
ZFN	Zinc-Finger Nucleases



## VI. Table of Contents

I. Statutory Declaration .....	I
II. Clause of confidentiality .....	II
III. Acknowledgements.....	III
IV. Abstract .....	IV
V. List of Abbreviations .....	V
VI. Table of Contents .....	VIII
1. Introduction.....	1
1.1. Induced pluripotent stem cells (iPSCs) .....	1
1.2. Reprogramming.....	4
1.2.1. Integrative reprogramming approaches .....	5
1.2.2. Non-Integrative reprogramming approaches.....	6
1.2.2.1. Non-Integrative, Viral .....	6
1.2.2.2. Non-Integrative, Non-Viral .....	7
1.2.2.3. Sendai Virus (SeV) .....	8
1.2.2.4. Episomal vectors (Epi).....	10
1.2.2.5. mRNA transfection .....	12
1.3. Genome editing .....	14
1.3.1. Meganuclease (MN).....	17
1.3.2. Zinc-finger nuclease (ZFN) .....	17
1.3.3. TALEN .....	19
1.3.4. CRISPR/Cas9 .....	21
1.3.5. Comparison of programmable nuclease platforms.....	24
1.4. Genome editing in induced pluripotent and embryonic stem cells .....	27
2. Materials and Methods .....	28
2.1. Cell culture .....	28
2.1.1. Pluripotent Stem Cells (PSCs).....	28
2.1.2. HEK 293T cells.....	28

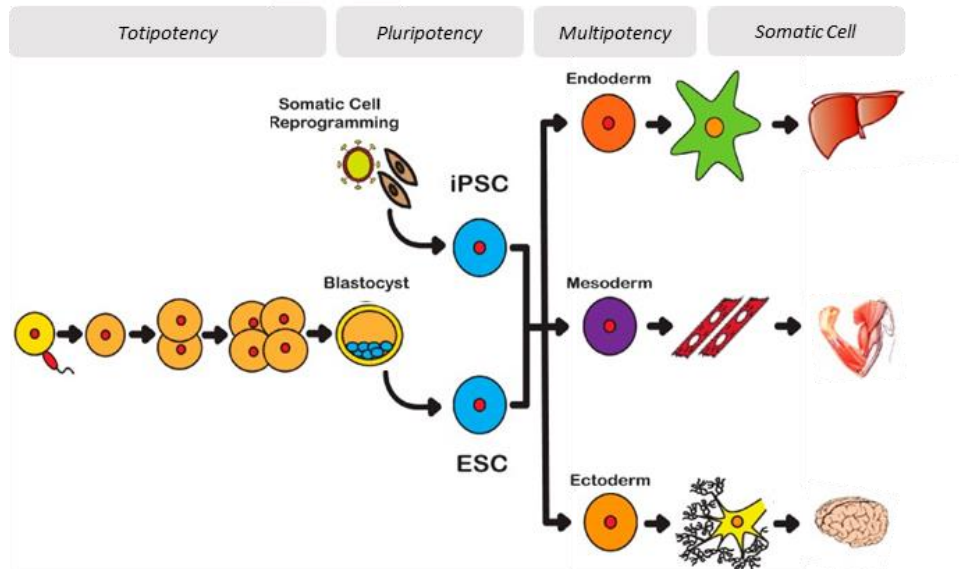
2.1.3.	Freezing.....	29
2.2.	Transfection .....	29
2.2.1.	CRISPRMAX Lipofectamine .....	29
2.2.2.	NEON Electroporation system .....	30
2.3.	Low density seeding, clone picking and passaging .....	31
2.4.	PCR, Sanger Sequencing and Analysis .....	32
2.5.	TIDE Analysis.....	33
2.6.	ICE Analysis.....	34
2.7.	GFP to BFP conversion .....	35
2.7.1.	Oligo Design .....	35
2.8.	Microscopy and Image processing .....	36
2.9.	FACS Analysis .....	37
2.10.	gRNA design .....	38
2.10.1.	gRNA design tool from MIT .....	38
2.10.2.	sgRNA efficiency testing.....	38
2.11.	Clone Analysis .....	39
3.	Results.....	40
3.1.	Optimization of RNP delivery .....	40
3.1.1.	Optimization of CRISPRMAX lipofection system for RNP delivery .....	40
3.1.2.	Optimization of NEON electroporation system for RNP delivery .....	42
3.1.3.	Transfection using the NEON Electroporation system results in higher KI yield than using CRISPRMAX Lipofection.....	44
3.1.4.	Delivery of two ssODNs increases the yield for WT/KI .....	46
3.2.	Repair of a heterozygous mutation in a patient-derived iPS line.....	48
3.2.1.	sgRNA design.....	48
3.2.2.	sgRNA testing.....	49
3.2.3.	ssODN design .....	49
3.2.4.	Pool Analysis .....	50
3.2.5.	Clone Analysis.....	52
3.2.6.	DNA repair Analysis .....	53
3.3.	RNP vs. Plasmid .....	55

3.4.	Evaluation of experiment variability .....	56
3.4.1.	Characterization of editing events after GFP to BFP conversion.....	56
3.4.2.	Comparison of editing events between experiments .....	57
4.	Discussion .....	59
4.1.	Optimization of RNP delivery .....	59
4.2.	Delivery of two ssODN increases the yield for WT/KI .....	60
4.3.	Repair of a heterozygous mutation .....	61
4.4.	RNP vs. Plasmid .....	62
4.5.	Evaluation of experiment variability .....	63
5.	Conclusion.....	64
5.1.	Further perspectives .....	65
VII.	Table of Figures.....	XI
VIII.	References .....	XIII

# 1. Introduction

## 1.1. Induced pluripotent stem cells (iPSCs)

Stem cells (SCs) are self-renewing cells that can generate progeny cells that are capable of following more than one single differentiation pathway [1]. SCs can be divided into four main classes: embryonic stem cells (ESCs), adult stem cells (ASCs), cancer stem cells (CSCs) and induced pluripotent stem cells (iPSCs) [1].



**Figure 1 iPSCs are capable to self-renew and differentiate into tissues of all three germ layers.** Depending on the specialization stage of the stem cells, the SCs can be described according to their potency to differentiate into various progenitor cells. Totipotency: During the first few cell divisions after fertilization, the cells maintain the ability to form all three germ layers as well as extra-embryonic tissues or placental cells. Pluripotency: The cells maintain the ability to self-renew and differentiate into the three germ layers and down many lineages but do not form extra-embryonic tissues or placental cells (ESCs and iPSCs). Multipotency: Adult stem cells (ASCs) are undifferentiated cells found in postnatal tissues. These specialized cells have very limited ability to self-renew and are committed to lineage specific differentiation. Adapted from [2].

ESCs are self-renewing cells that are widely known to be pluripotent, a term referring to a cell's ability to differentiate into progenitor cells of all three germ layers, namely endoderm, mesoderm and ectoderm. As a result, ESCs are capable of generating all cell lineages of the developing and adult organism. ESCs can be derived from the inner cell mass (ICM) of mammalian blastocysts. In ESCs, the pluripotent state is largely governed

by so-called core transcription factors (TFs) maintaining cells in their undifferentiated stage. Three of these core TFs are Oct4<sup>1</sup>, Sox2<sup>2</sup> and Nanog, which are often referred to as pluripotency genes [3][4][5]. ESCs are the best-studied SCs and knowledge derived from ESC research has guided the investigations of other types of SCs. However, the derivation of ESCs from human embryos evokes strong ethical concerns.

ASCs are stem cells with more restricted differentiation potential that are located throughout the body. They retain co-expression of at least three of the core TFs [3]. They are not considered pluripotent because their potency is restricted as they can only form progenitors of a certain branch. For such lineage-restricted SCs, the differentiation capability is limited. For example, hematopoietic SCs give rise to blood cells, whereas neural SCs generate various neuronal and glial populations [1]. CSCs, also referred to as cancer initiating cells (CICs), are defined as cells within a tumor that can self-renew, produce differentiated progeny, and drive tumorigenesis. Unlike ESCs, CSCs are highly heterogeneous with great variation among the markers for each tumor type [3].

Researchers have to deal with two major drawbacks when using ESCs. First, embryonic stem cells are derived from human embryos, an ethically controversial source of cells. Secondly, for therapeutic application of derivatives from ESCs, transplantation into a heterologous patient causes immune rejection and inflammation. To overcome the rejection issue, autologous pluripotent stem cells (PSCs) served as a preliminary solution. Two initial methods for the derivation of autologous PSCs evolved: Fusion with embryonic stem cells [6] and Somatic Cell Nuclear Transfer (SCNT). The latter refers to a method where somatic cells can be reprogrammed to pluripotent cells by transferring their nuclear content into oocytes [7][8]. However, in both methods embryo-derived ESCs were still required, which did not solve the ethical controversy. Not surprisingly, a new reprogramming method avoiding the use of embryos, ESC or oocytes was reported in 2006, resulting in the generation of induced Pluripotent Stem Cells (iPSCs).

iPSCs are cells derived from non-pluripotent cells. Yamanaka et al. showed that introducing four genes, encoding the transcription factors Oct4, Sox2, c-Myc and Klf4, can convert adult skin cells into PSCs [9]. In this first publication, Yamanaka et al. used

---

<sup>1</sup> octamer-binding transcription factor 4

<sup>2</sup> SRY (sex determining region Y)-box 2

the reactivation of endogenous Fbx15, an ESC-specific reporter gene, to select for the few reprogrammed cells [10].

In this first publication, Yamanaka et al. used the reactivation of endogenous Fbx15, an ESC-specific reporter gene, to select for the few reprogrammed cells. Similar to ESCs, these iPSCs had unlimited self-renewing potential and could differentiate into all three germ layers. Nevertheless, they still had some differences in their gene and epigenetic regulations. Although they were able to form embryoid bodies, teratomas and fetal chimeras, these first generation iPSCs could not produce a fully functional, viable chimera when they were injected into developing embryos [9].

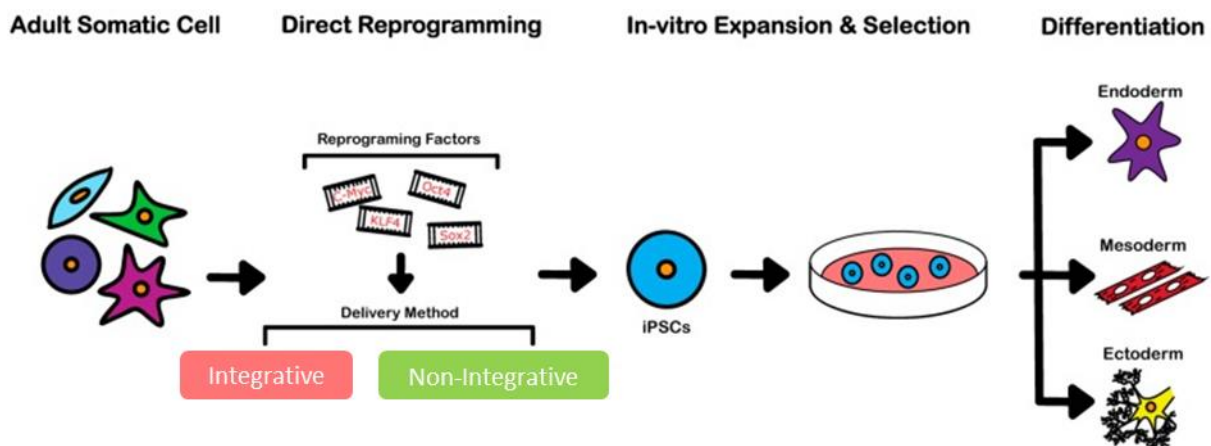
For “second generation iPSCs,” mouse fibroblasts were transfected with the same transcription factors via retroviral transduction. However, instead of using Fbx15 as a reporter gene, Nanog and Oct4 were used and the derived iPSCs that were generated were considered to be functionally identical to ESCs [11][12].

This method proved to be applicable for human cells as well. Yamanaka et al. successfully derived iPSCs from human fibroblasts by introducing Oct4, Sox2, Klf4, and c-Myc via a retroviral system [13]. Another successful reprogramming approach was performed by Thomson et al. using a lentiviral system to introduce the factors Oct4, Nanog, Sox2 and Lin28 [14].

Before iPSCs can be used in the clinic, additional work is required to avoid probable uncontrolled proliferation behavior resulting in tumor formation and to determine to what extent human iPSCs differ from ESCs. Nonetheless, human iPSCs are a useful platform for studying the development and function of human tissues. Patient-derived iPSCs eliminate the concern of immune rejection upon transplantation therapies to a large extent, with the exception of autoimmune diseases. For drug development, human iPSCs, in combination with genome editing, can be exploited to generate patient-specific, *ex-vivo* testing strategies.

## 1.2. Reprogramming

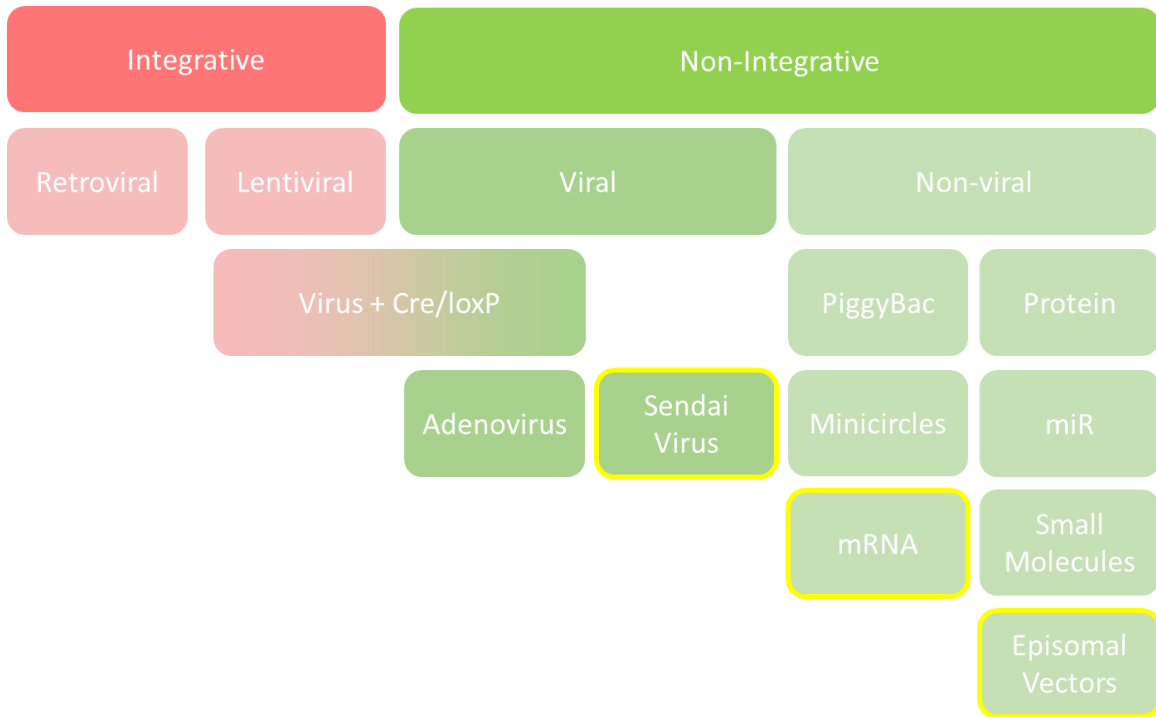
When deciding on a reprogramming method, several factors have to be taken into consideration, including the cell type, the capacity of the method to adequately reprogram the cell of interest and ensuring the integrated sequences will not hinder downstream application in iPSCs [15]. Reprogramming approaches can be divided into two main categories: integrative and non-integrative. The integrative techniques involve retroviral and lentiviral vectors which integrate their genome into the host chromosome. The non-integrative methods include: non-integrating vectors (episomal vector), non-integrative virus (Sendai virus) and mRNA or protein transfection. Apart from these, small molecules can be used to alter the regulatory pathways or the epigenetic state of cells to increase reprogramming efficiency [16].



**Figure 2** Induced pluripotent stem cells (iPSCs) can be derived from adult somatic cells and upon expansion and selection be differentiated into cells of all three germ layers. The reprogramming of adult somatic cells to iPSCs can be achieved through the ectopic expression of the transcription factors Oct3/4, c-Myc, Sox2 and Klf4. Once reprogrammed to a pluripotent state, iPSCs can be expanded *in vitro* and subsequently differentiated to ectoderm, mesoderm and endoderm lineages for use in cell therapy, disease modelling and drug discovery. Adapted from [2].

In most cases, the reprogramming methods deliver the four Yamanaka factors or OSKM factors, namely Oct3/4, Sox2, Klf4 and c-Myc [2]. Preferentially, non-integrative methods are used and in the case of human induced pluripotent stem cells (hiPSCs) the most common reprogramming methods are Sendai-viral (SeV), episomal (Epi) and mRNA

transfection [17]. This chapter will shortly describe the most popular reprogramming methods and will focus on the latter three methods (SeV, Epi and mRNA transfection), which are currently most relevant for the derivation of hiPSCs.



**Figure 3 Overview of most commonly used reprogramming approaches to derive iPS cells.** To create an overview, the reprogramming tools are clustered into integrative and non-integrative, where the latter is again separated into viral and non-viral approaches. The most reliable methods for reprogramming human cells are Sendai-viral (SeV), episomal (Epi) and mRNA transfection, indicated with a yellow frame.

### 1.2.1. Integrative reprogramming approaches

As the first successful reprogramming of cells was done using retroviral-mediated transduction [9], integrative reprogramming approaches were considered powerful to introduce foreign genes into host genomes. Primarily used integrative vectors are retroviral and lentiviral, where the latter can also infect cells that are not actively dividing [18]. Nonetheless, even though the efficiency of such integrative tools is higher than the non-integrative methods, there are many challenges when it comes to their application



for therapeutic purposes. The major drawback of integrating vectors is their random integration into the host genome, which can lead to insertional mutations and tumorigenesis [19]. Although the safety of integrative methods cannot be guaranteed, intensive research is ongoing in order to control the integration behavior of such vectors [16].

In order to control the safety of viral integration, lentiviral delivery in combination with Cre recombinase can be used. One example would be lentiviral vectors containing LoxP sites that are introduced into the 5' and 3' long terminal repeats (LTRs) together with Doxycycline (DOX) inducible Cre recombinase. The virus reliably delivers the OSKM factors and upon Cre-mediated recombination, the transgenes are removed again. However, a potential risk of insertional mutations from residual vector sequences may not be ruled out [20]. Even minimal residual vector sequences would be problematic when used for clinical applications [2].

## **1.2.2. Non-Integrative reprogramming approaches**

To further enhance the safety of cell reprogramming, completely integration- or even transgene-free approaches were pursued. Non-integrative methods can be further divided into viral and non-viral methods. The most commonly used non-integrative viral vectors are adenovirus and Sendai virus [16].

### **1.2.2.1. Non-Integrative, Viral**

Adenovirus undergoes receptor-mediated endocytosis to infect the host cell. It is released in the cytosol and reaches the nucleus, where the cargo DNA is not integrated but remains extra-chromosomally [21]. The virus infects dividing and non-dividing cells. However, since the DNA cargo of the virus is not integrated into the host genome, the construct will ultimately be diluted out after several cell divisions [22][23].

Another integration-free, virus-mediated reprogramming method utilizes Sendai virus (SeV), which is efficient and highly reliable. The workload is relatively low and the

procedure is considered to be safe since viral sequences vanish after several passages [17]. More detailed information on SeV reprogramming will be described below (1.2.2.3. Sendai Virus).

#### **1.2.2.2. Non-Integrative, Non-Viral**

A very robust non-viral, non-integrative reprogramming tool is the PiggyBac (PB) transposon system. For this method, a mobile genetic element is transposed between the delivered polycistronic expression vector and the chromosome. The PB transposase recognizes specific inverted terminal repeats (ITRs) that are located at the ends of transposons and temporarily integrates them at TTAA sequences in the chromosome. The process can be imagined as a “cut and paste” mechanism which provides high efficiency but can also leave behind residual sequences that are likely to stay in the host’s genome due to the difficulty in removing all transposon copies [24].

The OSKM factors can also be delivered in their purified proteinaceous form [25]. However, since delivery of protein into the intracellular compartment is very difficult, Cell Penetrating Peptide (CPP) is used. CPP is a short segment of basic amino acids [26]. The use of proteins for reprogramming purposes is possible but is rather inefficient and the proteins are difficult to synthesize in large quantities [2].

Minicircle DNA is a very straightforward and easy method to use for cell reprogramming. Minicircles are compact vectors that are free of bacterial DNA and are capable of persistent high-level expression in cells. There is no need for subsequent drug selection or vector-excision. Minicircle DNA is small (~4kb) and because they contain no bacterial DNA sequences, they are less likely to be perceived as foreign and to be destroyed. The smaller the size of the minicircles the longer their persistence, but efficiency is still relatively poor and they do not work in all cell types [27].

Lately, microRNAs (miRs) were found to be applicable for reprogramming purposes as well. MiRs are expressed as longer transcripts, which are processed into smaller 18–23-nucleotide mature miRs that bind to the target transcripts, inducing cleavage or inhibiting translation. For example, using the mir-302/367 has resulted in the establishment of iPSC

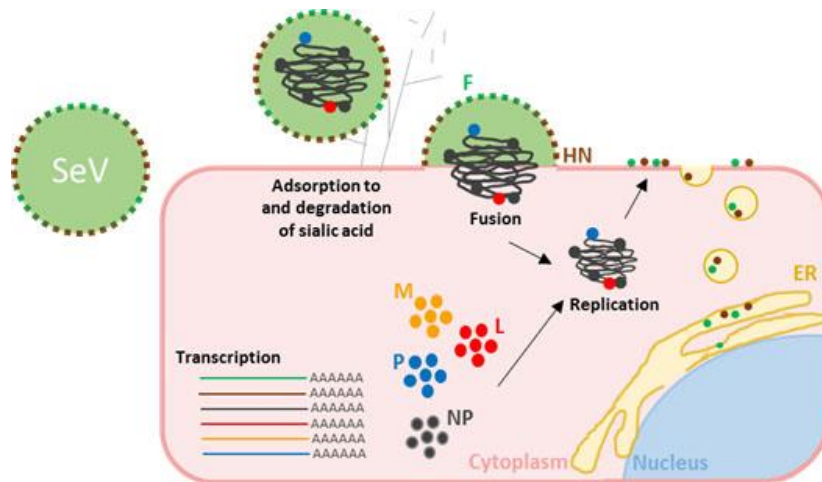
with high efficiencies [28]. Reprogramming using miRs does not require vectors nor is it integrative [2].

Small molecules can be used to generate iPSCs as they can interfere with the cell's physiology in manners not involving transcription factors. These include epigenetic modifiers, WNT signal modulators, cell senescence attenuators, metabolism modulators, and regulators of cell apoptosis/senescence pathways. There are several modes of action, but overall, small molecules can either improve reprogramming efficiency or can even fully replace reprogramming factors. However, the reprogramming of somatic cells with small molecules only was reported for murine cells but not for human cells so far. This represents an attractive alternative to TF-dependent reprogramming methods, as they are not only cost-effective but can also be combined to function in a temporary, controllable and/or reversible manner [2]. For example, the reprogramming efficiency of episomal vectors was successfully enhanced by the use of small molecules [29].

In order to derive hiPSCs, SeV, episomal and mRNA transfection are the most frequently used and studied reprogramming tools. The advantages and disadvantages of these methods are explained below.

### **1.2.2.3. Sendai Virus (SeV)**

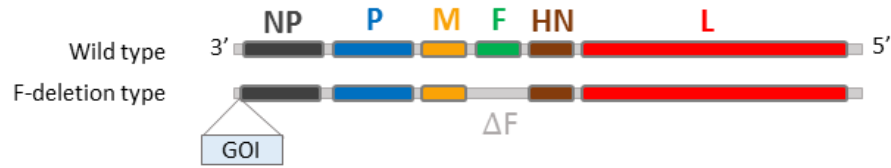
Sendai virus is a negative-strand RNA virus and a member of the paramyxovirus family. It replicates in the cytoplasm of infected cells and does not integrate into the host genome, facilitating a safe expression of transgenes without the risk of modifying the genome. Infected cells stay genetically intact and carry the same genomic DNA as the original cells – therefore SeV is an ideal tool for hiPSC reprogramming [30]. While Sendai virus is not harmful to humans, it can infect a wide range of cell types of various hosts since SeV infects cells by attaching to the sialic acid receptor present on the surface of many different cells [31][32].



**Figure 4 The life-cycle of the Sendai Virus.** Hemagglutinin-Neuraminidase (HN) recognizes the cell surface receptor sialic acid and Fusion protein (F) fuses the viral envelope with the cell membrane when the virus enters the cell. After infection, the virus goes through genome replication, transcription and protein synthesis, followed by the assembly and release of daughter virus particles. For reprogramming purposes, the gene encoding the F protein is deleted in order to make the virus incapable of producing infectious particles within infected cells. Large protein (L) is the large subunit of RNA polymerase and phosphoprotein (P) is the small subunit of the RNA polymerase. Nucleocapsid protein (NP) forms the core nucleocapsid complex with the genomic RNA. Matrix protein (M) supports the envelope structure from the inside. Adapted from [33].

As depicted in Figure 4, the high level of safety stems from the fact that the entire life cycle of Sendai virus is carried out exclusively in the cytoplasm. RNA (and not DNA) serves as a template for replication and, therefore, carries no risk of disturbing the genetic information in host chromosomes [34][35].

As is typical for a non-integrative method, the introduced vector will dilute out after several cell cycles. For reprogramming purposes, the viral RNA is modified in order to inhibit the infection of surrounding cells. For this purpose, the fusion (F) protein, required for the functional fusion of a virus particle, is missing in the modified SeV and will prohibit the virus from producing infectious particles. After Sendai virus transfection, approximately 10 passages are required before the generated iPSCs are footprint-free [36].



**Figure 5 Sendai Virus RNA vector.** Activation of F protein by a protease is required for the virus-cell fusion process to take place. The gene encoding the F protein is deleted from commercially available reprogramming vectors such as the CytoTune® reprogramming vectors. Without the F protein, the virus cannot produce infectious particles and hence the particles generated in the infected cells are incapable of infecting surrounding cells. Adopted from [37].

Commercial vendors provide SeV reprogramming kits (CytoTune-iPS 2.0 Sendai Reprogramming Kit, Thermo Fisher Scientific), which are very convenient for their application for reprogramming skin fibroblasts and blood cells. However, there is limited flexibility for the modification of these SeV viruses, such as introducing other reprogramming factors. Another major drawback of SeV reprogramming is the lack of clinical-grade SeV and the length of time it takes SeV RNA to clear out of the infected cell [17].

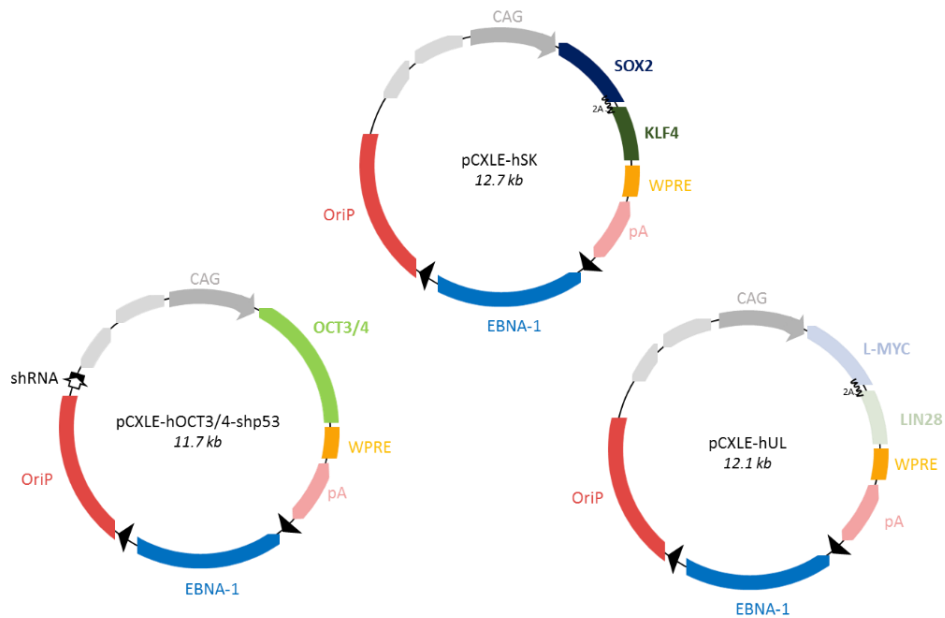
#### 1.2.2.4. Episomal vectors (Epi)

Episomal vectors (or episomal plasmids) are another integration-free reprogramming method [38]. A combination of such episomal vectors provides relatively efficient and safe reprogramming results [39]. Episomal vectors are based on the Epstein-Barr Nuclear Antigen-1 (oriP-EBNA-1) system. For sufficient retention and replication of the vectors, each vector contains a viral origin of replication (OriP) and a sequence that encodes the DNA binding protein EBNA-1 [40]. Relatively high expression of the reprogramming factors is sustained since the EBNA-1 protein recognizes OriP and induces replication of the episomes in parallel with DNA amplification in the host cell [41]. Generally, the use of episomal plasmids has proven to show good efficiency and works for many cell types.

This procedure is also applicable for derivation of iPSCs from cord blood and peripheral blood cells [40]. Initially, the efficiency for reprogramming PBMCs by episomal vectors was generally low [42] but could be enhanced by generating a mixture of vectors that encode the typical factors *Oct3/4*, *Sox2*, *Klf4*, c-Myc with Lin28 and Nanog. It has been

shown that both suppression of the p53 pathway (by expressing a p53shRNA or a dominant negative form of p53) and replacing c-Myc and Nanog with L-Myc enhances iPSC generation [43][44].

To further enhance the reprogramming efficiency, small molecules can be added and an extra plasmid that lacks OriP can be introduced. Without OriP, the vector is unable to replicate but still expresses additional EBNA-1 transiently [40][29]. Accordingly, different combinations and designs of episomal vectors are available, each respectively yielding a different level of efficiency.

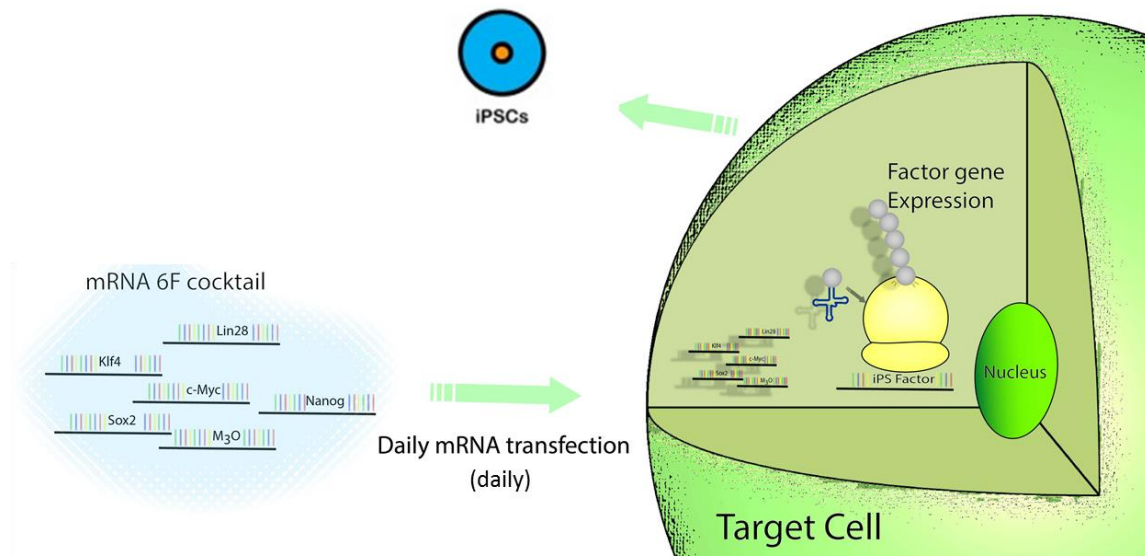


**Figure 6 Episomal vectors.** The depicted set of “Epi” delivers the transcription factors OCT3/4, SOX2, KLF4, L-MYC and LIN28 in three episomal plasmids. An additional p53 shRNA in one of the three plasmids further enhances the reprogramming efficiency. 2A, 2A self-cleavage sequence; CAG, CAG promoter; WPRE, woodchuck hepatitis post-transcriptional regulatory element; and pA, polyadenylation signal. Adapted from [39].

Overall, the key advantages of Epi reprogramming are the high reliability of hiPSC generation from fibroblast and blood samples as well as the relatively quick loss of the reprogramming agent. However, retention of the vectors can be found in up to 30% of the lines. Moreover, due to the use of a shRNA for p53, this reprogramming method does raise concerns regarding the genetic integrity of the resulting hiPSC lines. Nonetheless, episomal reprogramming is an cost-effective method for creating hiPSCs in a non-viral and non-integrating manner [17].

### 1.2.2.5. mRNA transfection

Another non-integrative way of reprogramming somatic cells into iPSCs is mRNA transfection [45]. Reprogramming with mRNA is considered absolutely footprint free since mRNA are degraded within 24-36h [46]. Modified mRNA is used for this procedure to limit the innate immune response that is usually triggered after mRNA transfection.



**Figure 7 mRNA transfection.** A “reprogramming cocktail” of several mRNAs is added daily to the somatic cells in order to derive iPSCs. The mRNAs will be quickly translated into the cell. Adapted from [47] and Allele Biotech, 2018.

Although this procedure can reprogram cells at a relatively high efficiency, it is time-consuming and labor intensive [2]. When it comes to implementing this technique, it is relatively difficult and the success rate varies between cell types. At this time, there have been no reports of successful generation of hiPSCs via mRNA transfection in blood cells.

Apart from the increased workload, the low success rate of implementation, and the need for a tissue culture incubator with 5% O<sub>2</sub> control, there are several advantages of using mRNA for reprogramming, such as the speed of colony emergence, high efficiency, a complete absence of integration, a very low aneuploidy rate and a low donor cell requirement (see Table 1). Nonetheless, the method needs to be optimized and be made applicable for a wider range of cell types.

Promising approaches for improving reprogramming efficiencies are, for example, the combination of mRNA transfection with pluripotency-inducing miRNAs or fusion of Oct4 to a heterologous transactivation domain, which has been shown to increase efficiency and accelerate colony emergence [17][47].

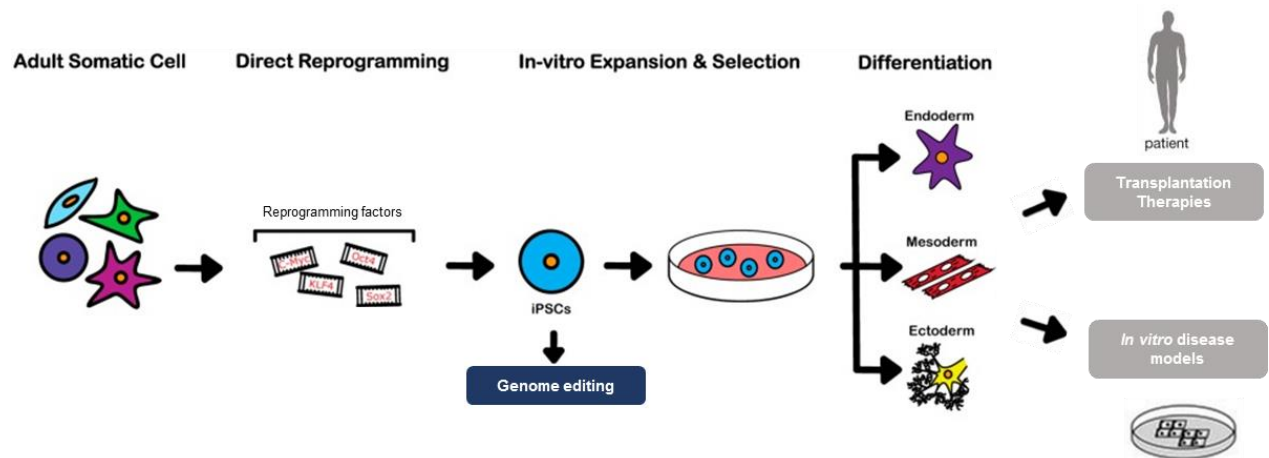
**Table 1 A comparison of the most commonly used non-integrating reprogramming methods to derive human induced pluripotent stem cells (hiPSCs).** The most prominent non-integrative reprogramming tools are listed and compared to a common integrative method. Here, the RNA (mRNA) from Stemgent, the Sendai Virus (SeV) Kit “CytoTune” from Life Technologies, the Episomal plasmids (Epi) from Yamanaka/Addgene and the integrating Lentivirus from Mostoslavsky/BU are compared. F, Fibroblasts; B, Blood. Adapted from [17].

	mRNA	SeV	Epi	Lenti
<b>Efficiency (F)</b>	~1%	~0.1%	~0.01%	~0.5%
<b>Efficiency (B)</b>	-	High	High	High
<b>Reliability (F)</b>	<50% (+miRNA: >70%)	>90%	>90%	>90%
<b>Reliability (B)</b>	-	High	High	High
<b>Reprogramming workload</b>	High	Low	Low	Very high (excision)
<b>Aneuploidy rate</b>	2.3.3%	4.6%	11.5%	4.5%
<b>Time to colony emergence</b>	Fast	Moderate	Moderate	Moderate
<b>Lines free of reprogramming agents by passage 5</b>	100%	0%	60.9%	0%
<b>Lines free of reprogramming agents by passage 9-11</b>	100%	78.8%	66.7%	0%
<b>Major drawback</b>	Not working for blood cells	High passage number to reach complete absence of virus	Concerns about genetic integrity (especially with shRNA for p53)	Integrative



### 1.3. Genome editing

The term “genome editing” generally describes the tasks of deleting, inserting or replacing DNA in the genome of a living organism [48].

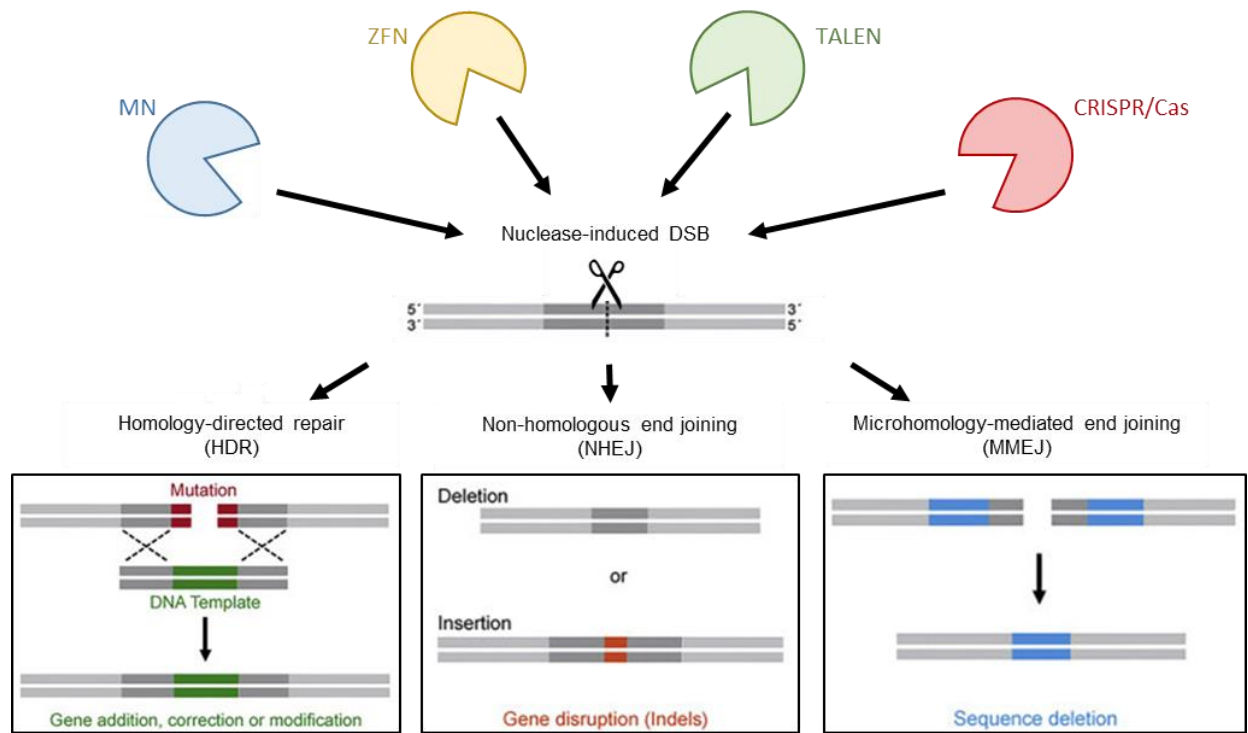


**Figure 8** Genome-edited iPSCs can be expanded, selected and differentiated into somatic cells that could be used either in transplantation therapies or alternatively to model human diseases. Adapted from [2], [48].

Before the era of genome editing, temporary interference with mRNA was widely used to study gene targets and protein expression *in vivo* and *in vitro*. This was facilitated by the use of RNA interference (RNAi). RNAi is a well-studied method and several improvements have already been established, such as small-interfering RNA (siRNA) and short-hairpin RNA (shRNA). However, these tools lead to temporary downregulation of gene expression and the efficacy of downregulation is mostly limited [49].

In contrast, genome editing allows permanent modifications on the DNA level and embodies a very reliable method to study genes and their functions. For genome editing, the tools of choice are nucleases. Depending on the genetic edit that is desired, diverse types of nucleases are available. The most commonly used methods for genome editing are Meganucleases (MN), Zinc-Finger Nucleases (ZFN), Transcription Activator-Like Effector-based Nucleases (TALEN), and the clustered regularly interspaced short palindromic repeats (CRISPR)/Cas 9 (CRISPR associated protein 9 nuclease) system [50]. All these nucleases cut the DNA at a specific site – usually creating double-strand

breaks (DSBs) or single-strand breaks (nicks) – whose subsequent repair results in the genomic modification. As shown below (Figure 9), the resulting DSBs or nicks can be repaired by non-homologous end joining (NHEJ), Microhomology-mediated end joining (MMEJ) or homology directed recombination (HDR).



**Figure 9** Meganucleases (MN), Zinc-Finger Nucleases (ZFN), Transcription Activator-Like Effector-based Nucleases (TALEN) and the Cas enzyme of the CRISPR/Cas system are all nucleases that can be used for specific DNA cleavage. Nuclease-induced DNA double-strand breaks (DSBs) can either be repaired via homology directed recombination (HDR), non-homologous end joining (NHEJ) or microhomology-mediated end joining (MMEJ), resulting in either random or precise insertions, deletions or substitutions. Adapted from [51].

HDR results in precise repair and takes place in the S- and G2-phase of the cell cycle, as it uses the sister chromatid as a template. Accordingly, for genome editing purposes, HDR is essential for precise editing of point mutations or SNPs, and requires an exogenous DNA template [52].

The two most commonly used DNA templates for HDR-mediated gene editing are single-stranded oligodeoxynucleotides (ssODNs) and plasmid DNA templates. The use of

ssODNs or plasmids mostly depends on the insertion size. For the repair of point mutations, ssODNs are most often used. For larger gene insertions, plasmid DNA templates are primarily used. A plasmid DNA template should include homology arms on the 5' and 3' ends typically ranging from 500 to 1,000bp. Compared to plasmids, ssODNs are much smaller, ranging from 100 to 200 nucleotides in length. Richardson et al. found that asymmetric homology arms in ssODNs support HDR efficiencies up to 60% [53]. Plasmids require substantial cloning to be derived, whereas ssODNs can readily be obtained from commercial vendors [54] [55].

NHEJ, in contrast, is responsible for fast DSB repair and can occur during the whole cell cycle. NHEJ leads to insertions or deletions (indels) of variable length [56]. NHEJ can be divided into classical NHEJ (C-NHEJ) and alternative NHEJ (A-NHEJ). In the classical pathway, the DSBs are directly ligated, which leads to deletions, insertions or substitutions. The C-NHEJ is especially prone to errors.

During A-NHEJ, which is also known as microhomology-mediated end joining (MMEJ), the DSBs are linked via micro-homologous domains. These domains are about 5-25 nucleotides long and lead exclusively to deletions at this position [57].

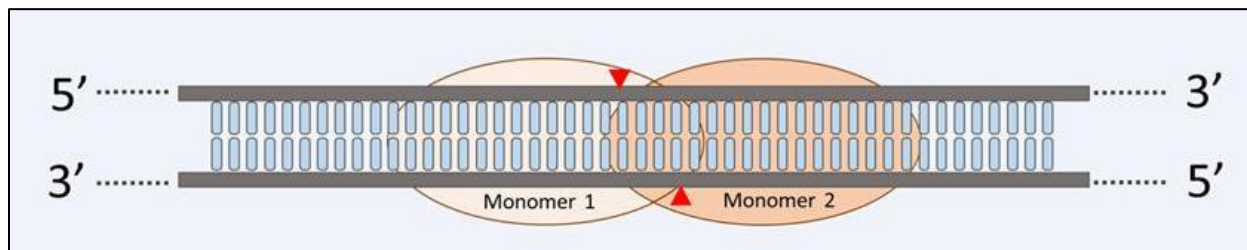
All repair pathways differ in their machineries. The current belief is that the repair proteins compete for the free DNA ends. While the exact mechanisms are unknown, there are important determining factors influencing the editing outcome, such as the stage in the cell cycle and the presence of a template DNA [51].

Overall, for successful DNA cutting (Knock-Out event), a specific nuclease is sufficient. If a DNA repair or modification (Knock-In event) is pursued upon cutting, a repair-template in the form of ssODN or plasmid DNA has to be provided. The mechanism of function for the respective nucleases, as well as their benefits and drawbacks, will be discussed in the following chapters.

### 1.3.1. Meganuclease (MN)

Meganucleases (or homing endonucleases) are endodeoxyribonucleases which recognize DNA sequences of 12–42 bp in length and are very specific for their target sequence [58]. However, to apply MNs for gene targeting, their DNA recognition properties have to be redesigned in order to cleave any given sequence.

There are several approaches for creating custom MNs. Two commonly used methods are mutagenesis and combinatorial assembly [51].



**Figure 10** Simplified illustration of a homodimer structure of a Meganuclease (MN) system. Meganucleases or homing endonucleases recognize DNA sequences between 12 and 42 bp in length. Adapted from [59].

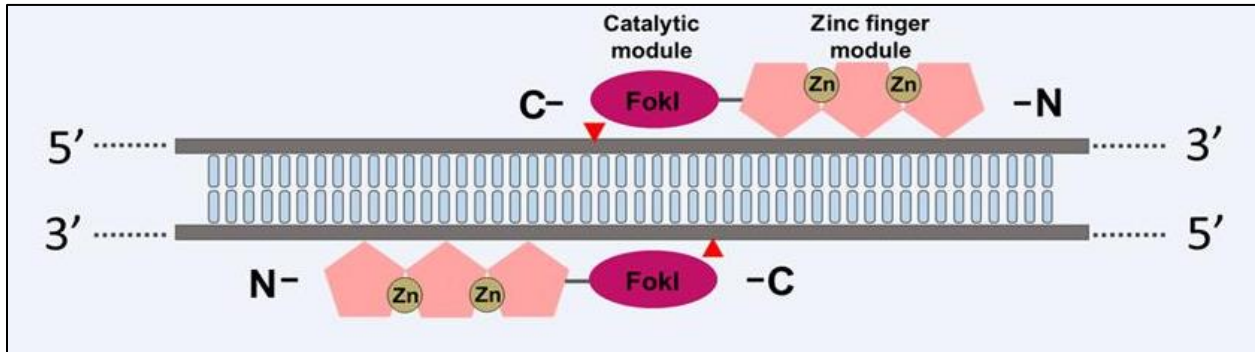
Meganucleases are considered to exhibit less toxicity in cells than other methods. However, engineering of those enzymes is very difficult, costly and time-consuming [60].

### 1.3.2. Zinc-finger nuclease (ZFN)

Zinc-finger nucleases consist of a site-specific DNA binding domain fused to the nonspecific cleavage domain of the FokI endonuclease. The DNA binding domain consists of zinc-fingers where each zinc-finger-molecule is specific for a base triplet.

At least two ZFN-molecules are required for genetic modification, as the two ZFN-molecules bind to the targeted DNA in a tail-to-tail orientation separated by a 5–7 bp-long spacer sequence. Each ZFN module is linked to a FokI cleavage domain that becomes active when dimerized. The double-stranded DNA cleavage occurs in the spacer region

[61]. An effective ZFN should even contain more than three zinc-finger domains in each DNA-binding module to increase specific DNA recognition [62].



**Figure 11 Zinc-Finger Nuclease (ZFN).** Two ZFN monomers are depicted that are bound to DNA. Each ZFN consists of a zinc finger DNA-binding domain (pentagons in rose) linked to a catalytic FokI domain (ellipse in pink). The ZFN molecules bind in a tail-to-tail orientation separated by a spacer, in which the cleavage occurs. Adapted from [59].

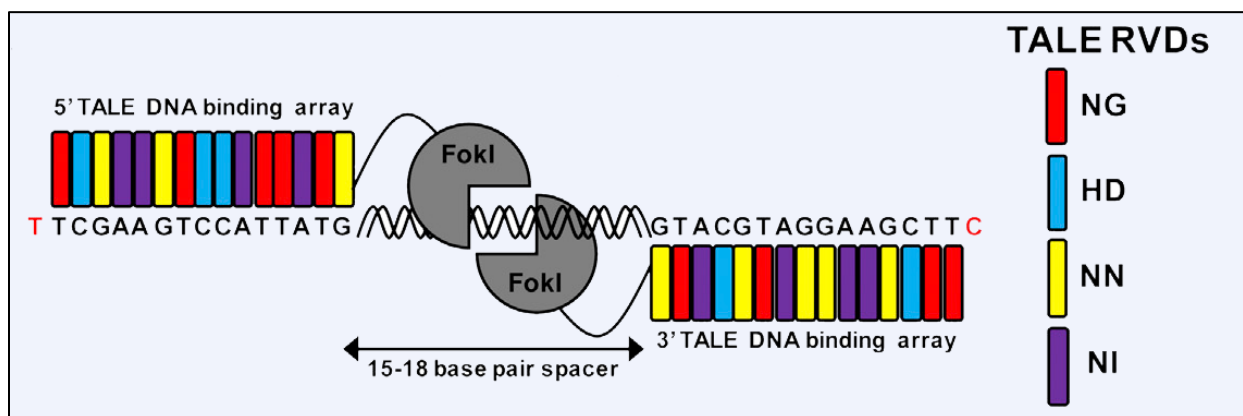
A big superfamily of ZFN is C2H2 zinc-fingers. These zinc-fingers are TFs that contain paired cysteines and histidines that bind a central zinc ion. Linkers combine several of the zinc-fingers, forming a DNA specific DNA binding unit [63].

Today, it is possible to generate zinc-finger arrays where different zinc-fingers are combined to target nearly any desired DNA sequence [51]. In order to reduce off-target cleavage, only engineered FokI cleavage domains can be used that are exclusively active as a heterodimer [64].

ZFN is a relative easy tool to use for *in vivo* delivery. Despite several advances and improvements, engineering of ZFN is still quite complex and difficult [60].

### 1.3.3. TALEN

Transcription activator-like effector nucleases (TALENs) consist of site-specific transcription activator-like effectors (TALEs) fused to the nonspecific cleavage domain of the FokI endonuclease. TALEs are proteins expressed and secreted by *Xanthomonas* bacteria, a plant pathogen. They are important virulence factors that act as transcriptional activators in the plant cell nucleus. Each TALE repeat, which serves as the DNA binding domain, consists of a 33-34 amino acid sequence which is highly conserved with the exception of the 12<sup>th</sup> and 13<sup>th</sup> amino acids. The amino acids at positions 12 and 13 (repeat variable di-residue, RVD) recognize a single base and determine binding specificity [65].



**Figure 12 Transcription activator-like effector nuclease (TALEN).** Depicted are two monomers bound to DNA. TALEN consist of a DNA binding domain containing the repeat variable di-residue (RVD) arrays of amino acids to recognize DNA specific sequences linked to a catalytic FokI domain (grey open cycle). Different RVDs encode for certain DNA bases. Adapted from [66][67][59].

The most common RVDs are NG, HD, NN and NI, which respectively bind thymine (T), cytosine (C), adenine/guanine (A/G), and adenine (A). For example, an Asparagine-Isoleucine (N-I) repeat in the TALE sequence will specifically recognize an A-base in the DNA sequence, whereas a Histidine-Aspartic Acid (H-D) repeat binds to a C-base (Figure 12) [68].

TALENs are primarily delivered in two plasmids: one plasmid containing the 5' TALEN DNA-binding array fused to a FokI-nuclease monomer and the other plasmid containing the 3' TALEN DNA-binding array which is also fused to a FokI-nuclease monomer [69].

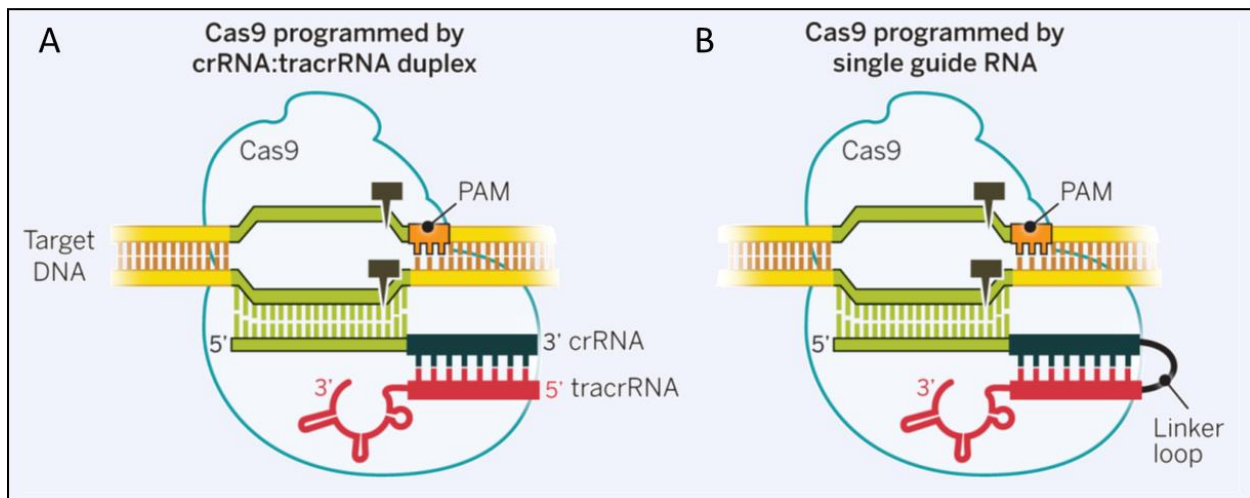
Regarding target site accessibility, one requirement for TALEN binding was a thymine (T) immediately upstream of the DNA-binding domain [68]. Recent adaptations led to the creation of synthetic RVDs which can overcome that 5' T-Rule [70]. Binding to the target site can also be influenced by methylation. Conventional TALE DNA-binding domains cannot bind and cleave methylated DNA sequences that contain one or more 5-methylated cytosines (5mCs), which are often found in CpG islands in promoter regions and proximal (5') exons [71][72]. To overcome this 5mC sensitivity, a de-methylating agent can be used during cell culture [73]. Such de-methylating agents are potentially harmful and can be avoided by using an alternative cytosine-binding RVD, referred to as "N\*", which can be used as a universal 5mC binding module and overcomes the 5mC sensitivity of TALENs [68][74].

Apart from 5mC sensitivity, TALENs are considered to be quite specific. However, the specificity can decrease if the binding affinity is too strong, whereas too-low affinity would lead to reduced TALEN activity. The decrease of specificity upon excess binding affinity may explain why NN RVDs bind either A or G [75]. In order to prevent re-cutting of the edited sequence after successful knock-in, it is recommended to introduce a silent mutation in the TALEN-binding site [66].

The difficulty of engineering of TALENs is moderate. Nonetheless, the *in vivo* delivery can be problematic due to the relatively large size of each TALEN [60].

### 1.3.4. CRISPR/Cas9

The CRISPR/Cas9 system depends on the Cas9 (CRISPR-associated) endonuclease in combination with Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR). In contrast to the previously described genome editing tools, the Cas9 endonuclease is guided via RNA molecules to its target site. This phenomenon was discovered in bacteria as part of their adaptive immune system. The CRISPR/Cas9 system cuts small fragments from invading virus or plasmid DNA and integrates them as protospacers between short palindromic repeats in the CRISPR array. These fragments are about 20 bp in length, and when the whole CRISPR array is transcribed, the corresponding RNA fragments (crRNA) will be excised and integrated in the Cas9 endonuclease with the help of a bacterial trans-activating crRNA (tracrRNA).

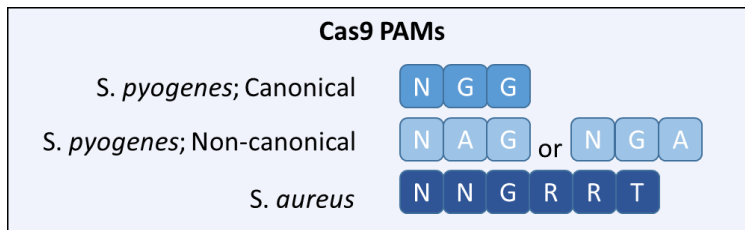


**Figure 13 The Cas9 endonuclease.** The enzyme requires a protospacer-adjacent motif (PAM) and a crRNA-tracrRNA duplex to be fully functional. The Cas9 nuclease is guided by a two-RNA structure formed by an activating tracrRNA base-paired to a crRNA (A). The two RNA structures can also be linked: A chimeric RNA is generated by fusing the 3' end of crRNA to the 5' end of tracrRNA (B). Adapted from [76] [77].

The crRNA is base-paired to tracrRNA to form a two-RNA structure that directs the Cas9 enzyme to the DNA locus where a DSB should be introduced [78][77]. The crRNA and tracrRNA can be linked by fusing the 3' end of crRNA to the 5' end of tracrRNA, generating a chimeric RNA referred to as single guide RNA (sgRNA). In order to exploit this system



for genome editing purposes, endonuclease activity can be directed to almost any sequence by delivering the complementary RNA together with the Cas9 enzyme. This activated Cas9 endonuclease specifically cleaves DNA sequences complementary to the sgRNA only if an additional 3-5 bp-long sequence, the PAM (protospacer adjacent motif), is present at the 3'-end of the target [79]. In the original CRISPR/Cas9 system, derived from *S. pyogenes*, the target DNA must immediately precede a 5-NGG PAM [80]. The 5-NGG PAM motif is found at about every 8–12 base pairs in the human genome, which already confers flexibility in target sequence selection. Nonetheless, this still excludes access to some very precise sequences. For instance, repair of point mutations via HDR-mediated repair requires very specific binding of the Cas9 endonuclease. To increase the variety of Cas9 target sequences, additional variability is given by using different orthologues that have different PAM requirements as depicted in Figure 14 [81].

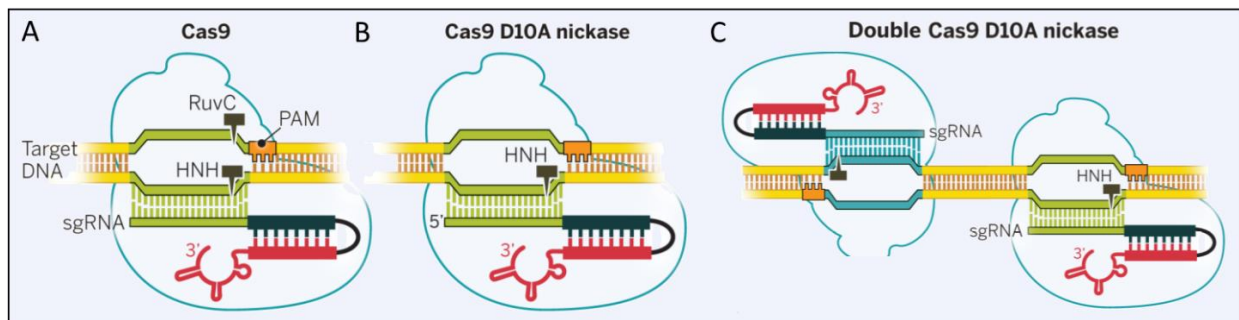


**Figure 14 Different Cas9 orthologues have different PAM requirements.** Each type of Cas9 nuclease uses its own protospacer adjacent motif (PAM), which can be exploited for genome editing purposes [66].

To improve the specificity of Cas9-mediated genome editing, novel methods and modifications have been developed.

For instance, Cong et al. modified the Cas9 enzyme via an aspartate-to-alanine substitution (D10A) to introduce an inactivating mutation in the HNH domain in order to derive a specific DNA nickase: the D10A mutant nickase (Cas9n). Nicked genomic DNA is typically repaired rather flawlessly. A combination of a pair of such Cas9n enzymes with a pair of offset sgRNAs complementary to opposite strands of the target site leads to site-specific DSBs. Since this method requires binding of two complexes, the off-target mutagenesis by each individual Cas9n-sgRNA complex is minimized while maintaining on-target modification rates comparable to those of wild type Cas9. Likewise, similar to the Cas9 D10A nickase, the H840A nickase mutant, which exhibits an inactivating mutation in the RucC domain of the Cas9 active site, can be utilized for a double nickase

set-up [76]. Paired nicking can be applied to minimize the risk of unspecific cutting as it significantly reduces off-target activity within the genome [82].



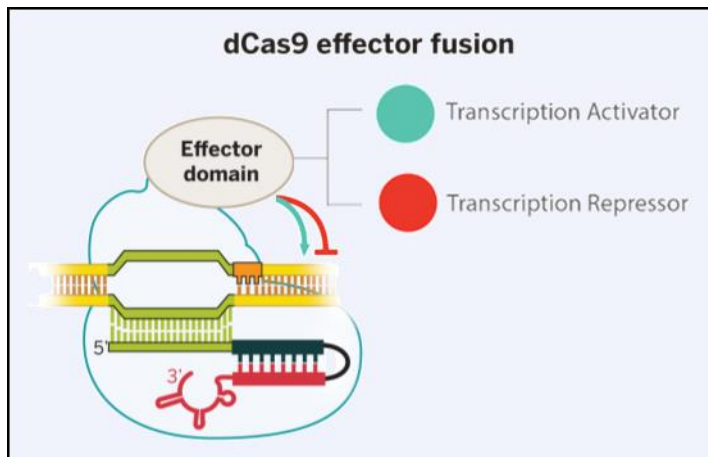
**Figure 15 Cas9 nickase and double nicking.** The Cas9 nuclease can function as a nickase (Cas9n) when engineered to contain an inactivating mutation in the active site. (A) A fully functional Cas9 enzyme complexed with the sgRNA and targeting the DNA at the PAM site. (B) The Cas9 D10A nickase: A mutation in the HNH domain in the active region of the Cas9 enzyme leads to a mutant that cuts the DNA on one strand only. (C) When Cas9n is used with two sgRNAs that recognize different sequences close to the target, a staggered DSB is created. Adapted from [76].

Another form of modification is coupling deactivated Cas9 (dCas9) to a certain effector domain. Such an effector domain can consist of almost any tag. Most commonly used linked domains are transcription activators or transcription repressors. Representatives of activator domains are VP64 transcriptional activator in combination with the transcription factors p65 and Rta – a system referred to as the dCas9-VPR system [83]. Another widely used system for transcriptional activation is the SunTag activator system, which allows multiple antibodies fused to VP64 to bind to dCas9-SunTag and recruit RNA polymerase, resulting in increased gene expression [84].

When it comes to transcriptional repression, it needs to be considered that dCas9 alone is sterically interfering with the binding and function of the transcription machinery, which is often referred to as CRISPRi (or CRISPR Interference). Nonetheless, CRISPRi seems to perform rather unreliably in eukaryotic cells, which is why coupling of dCas9 to transcriptional repressors was designed. For example fusion of KRAB<sup>3</sup> to dCas9 decreased the transcriptional activity further [85].

<sup>3</sup>Krüppel-associated box, the transcriptional repressor domain of Kox1: a member of the KRAB C2H2 zinc finger family, confers strong transcriptional repressor activities even to remote promoter positions [101].

The dCas9 variant has already found a wide range of enhancements and new applications. Coupling dCas9 to fluorescent domains like green fluorescent protein (GFP) can be a useful tool for live-cell imaging of chromosomal loci. The use of dCas9 coupling can also be extended to a linkage to DNA acetyltransferases, DNA demethylases or other epigenetic regulators.



**Figure 16 dCas9 effector fusion.** The catalytically deactivated Cas9 (dCas9) fused to an effector domain. When engineered to contain inactivating mutations in both of its active sites, the Cas9 enzyme can be used as an RNA-guided DNA binding protein. The dCas9 can mediate transcriptional down-regulation or activation, depending on the attached effector domain. Adapted from [85] [76].

The CRISPR/Cas9 system is a very versatile tool with a wide range of applications.

Compared to other genome editing tools (such as TALEN), CRISPR/Cas9 is more efficient for hyper-methylated DNA sequences [86]. When it comes to DSB repair, the efficiency of HDR is similar between TALEN and CRISPR/Cas9. NHEJ-mediated knockout, in contrast, seems to be more efficient with CRISPR/Cas9 [87]. However, CRISPR/Cas9 target sites are limited to loci that harbor a PAM, whereas TALEN or ZFN DNA-binding domains can be designed to target almost any sequence [88].

### 1.3.5. Comparison of programmable nuclease platforms

Even though genome editing turned out to be a very valuable tool with a big range of applications, it still carries some risks. Currently, the biggest concern is DNA off-target cleavage. Since genome editing is a permanent modification of DNA sequences, off-target effects could have tremendous side effects. Meganucleases and TALENs are considered to be the most specific of the above-mentioned genome editing toolset [89][90][75]. ZFNs, in contrast, can have some off-target effects, dependent on the target

site [91]. The same applies for the CRISPR/Cas9 system. Just recently, it was shown that the CRISPR/Cas9 system is not as specific as initially assumed [92]. There is a lot of research ongoing to overcome and/or control this problem.

**Table 2 Comparison of different programmable nuclease platforms.** Adapted from [93]

	<b>MN</b>	<b>ZFN</b>	<b>TALEN</b>	<b>Cas9</b>
Nuclease Size	1 kb	1 kb	3 kb	4.2 kb (Cas9) + 0.1 kb (gRNA)
Recognition site	14 - 40 bp	18 - 36 bp <i>per ZFN pair</i>	28 - 40bp <i>per TALEN pair</i>	22bp (20bp guide sequence + 2bp PAM sequence)
Off-target Effects	Low	Variable	Low	Variable
Targeting constraints	Targeting novel sequences often results in low efficiency	Difficult to target non-G-rich sequences	5' targeted base must be a T for each TALEN monomer	Targeted sequence must precede a PAM
Ease of multiplexing	Difficult	Difficult	Difficult	Easy
Ease of engineering	Difficult	Difficult	Moderate	Relatively easy
Immunogenicity	Unknown	Likely low	Unknown	Unknown
Ease of ex vivo delivery	Relatively easy	Relatively easy	Relatively easy	Relatively easy
Ease of in vivo delivery	Relatively easy	Relatively easy	Difficult	Moderate

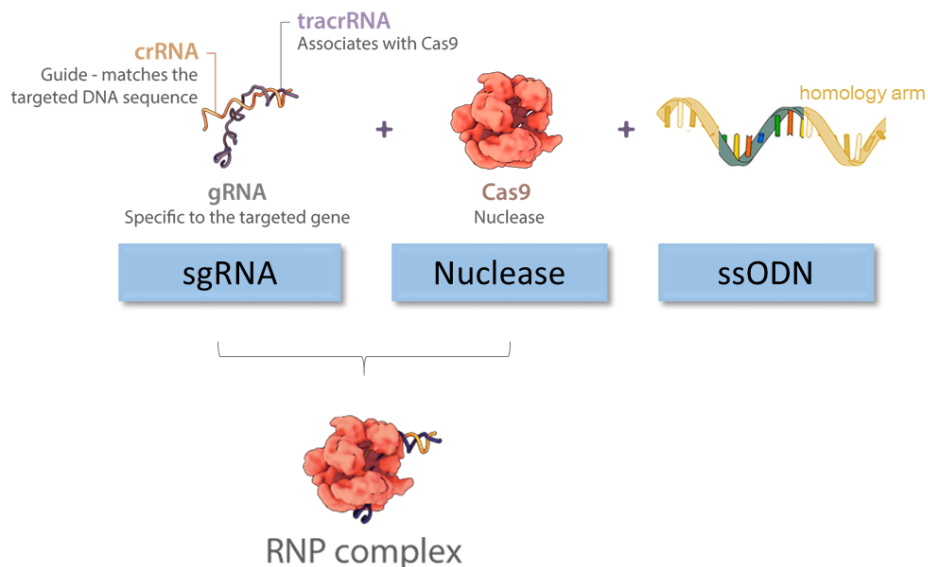
Apart from off-target effects, some other factors play essential roles when selecting a genome editing tool. *Ex vivo* delivery is considered to already be relatively easy using methods such as electroporation and viral transduction. *In vivo* delivery, however, is more complex. The size of the nuclease might be relevant for vector design and affects *in vivo* delivery. For instance, due to the large size of each TALEN and the repetitive nature of DNA encoding TALENs, *in vivo* delivery can be very difficult as it can lead to unwanted recombination events when packaged into vectors [60].

The accessibility to engineering plays a crucial role for experimental set-ups and indicates the feasibility of precise sequence modifications. For MNs and ZFN, engineering requires substantial protein engineering which is time-consuming and difficult. TALENs, in contrast, can be engineered via complex molecular cloning methods that, once

established, can be used with moderate effort. The CRISPR/Cas9 system can easily be re-targeted using standard cloning procedures and oligo synthesis.

Immunogenicity in humans is relevant in terms of using a programmable nuclease platform for gene therapy approaches. ZFN are based on human protein scaffold and are likely to have low immunogenic effects. However, FokI is derived from bacteria and may be immunogenic. All other nuclease platforms are derived from different species, such as Meganucleases, which originate from many organisms including eukaryotes. The TALEN protein is derived from *Xanthomonas* sp. and the CRISPR/Cas9 system was isolated from various bacterial species. Therefore, the immunogenicity in humans is largely unknown and a matter of concern [60].

In summary, for DNA cutting, a nuclease is required which has to be programmed in a way to bind specifically to the site of editing. For successful repair, a DNA template is required that can be designed for the modification of interest. As an example, the major editing factors for a successful Knock-In experiment using the Cas9 nuclease are depicted below in Figure 17.



**Figure 17 Major editing factors for a Knock-In experiment using CRISPR/Cas9.** The gRNA is responsible for the specificity of the Cas9-binding, the Cas9 nuclease cuts the DNA upon forming a complex with the gRNA (RNP complex). The repair template for HDR is here illustrated as a single-stranded oligonucleotide (ssODN) with homology arms. For bigger insertions or modifications, a DNA plasmid with homology arms can be used as well. Adapted from [94].

#### **1.4. Genome editing in induced pluripotent and embryonic stem cells**

After successful reprogramming, iPS cells can be edited genetically. For disease modelling purposes, genome editing is used to introduce a mutation of interest or to repair a certain genetic locus. Gene functions and the impact on disease phenotypes can be studied thoroughly using genome editing approaches. Edited iPS cells or ES cells can be differentiated into diverse tissue or 3D culture technologies, such as organoids. Such 3D organoids have led to the development of novel and more physiological human healthy tissue and cancer models [95]. The efficiency, safety and reproducibility of genome editing in iPS and ES cells is therefore essential and needs to be optimized to further improve related research. In this paper, improvements and applications of genome editing in iPS and ES cells will be described and discussed.

## **2. Materials and Methods**

### **2.1. Cell culture**

#### **2.1.1. Pluripotent Stem Cells (PSCs)**

HUES8 cells are male pluripotent cells isolated from the inner cell mass of a blastocyst and are property of Harvard University. Clones with either two copies (homozygous) or one copy (heterozygous) of EGFP (enhanced GFP) in the AAVS1 locus were derived by genome editing using TALEN. Both clones were derived by Susanne Waizenauer [96]. The clone HUES8GFP002 has two copies of EGFP inserted in the AAVS1 locus and is homozygous, whereas the clone HUES8GFPF is heterozygous for EGFP. Ery ED3 SiPS-A AAVS1 GFP cells were derived from blood cells using SeV reprogramming. Upon successful reprogramming, EGFP was inserted in the AAVS1 locus in both alleles. This homozygous cell line was derived by Nikolaos Poulis.

iPSCs were cultured in a feeder-free system on tissue culture dishes or plates coated with Cultrex (Cultrex® Basement Membrane Extract (BME), Trevigen) diluted in cold DMEM/F12 (Invitrogen). Undifferentiated cells were maintained with mTeSR1 medium (Stemcell Technologies) supplemented with 1% (vol/vol) penicillin/streptomycin (P/S, Invitrogen). Cells were not allowed to reach confluency greater than 90% and were passaged every 3–4 days using Gentle Cell Dissociation Reagent (Stemcell Technologies) treatment and, upon scraping, the cells were passaged as aggregates. After thawing, transfection or any other harsh treatment, the cells were re-plated in mTeSR1 supplemented with 10 µM Rock Inhibitor (Insolution Y-27632, EMD Chemicals Inc, RI) for 24 h. Medium was replenished daily with fresh mTeSR1 full medium. Cells were maintained at 37°C and 5% CO<sub>2</sub>.

#### **2.1.2. HEK 293T cells**

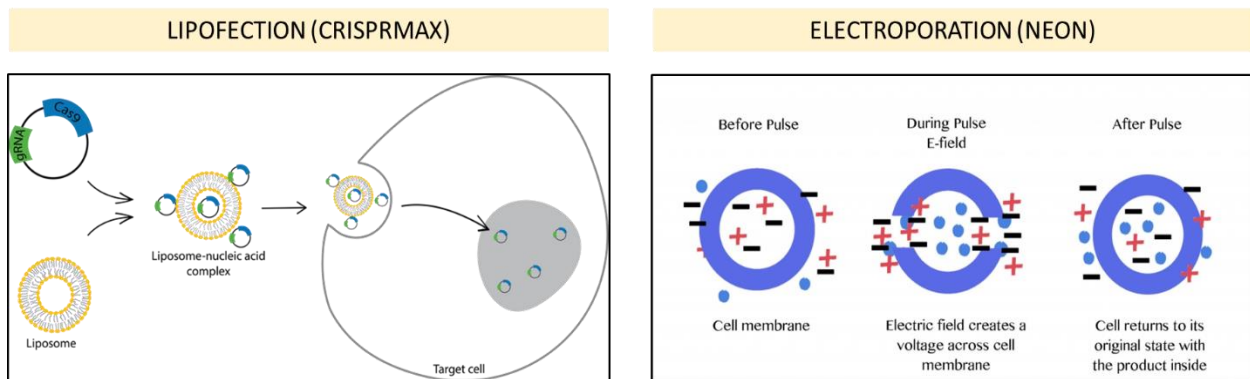
HEK 293T cells were cultured in T75 flasks and maintained using basal medium: Dulbecco's Modified Eagle's Medium (DMEM, Invitrogen) supplemented with 10% heat inactivated Fetal Bovine Serum (FBS, Gibco, Thermo Fisher Scientific) and 1% P/S (Invitrogen). Medium was changed every 2-3 days and cells were maintained at 37°C and 5% CO<sub>2</sub>.

### 2.1.3. Freezing

For freezing, the cells were detached, centrifuged at 1200 rpm for 3 min. and resuspended in freezing medium (FBS with 10% dimethyl sulfoxide (DMSO, Sigma Aldrich)). The vials were frozen down stepwise in an isopropanol-containing Mr. Frosty box at -80 °C and transferred to a -150°C freezer the next day.

## 2.2. Transfection

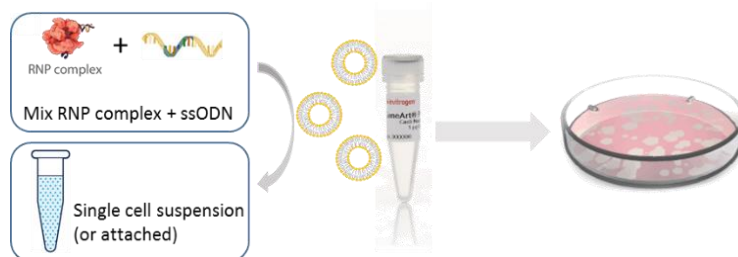
For transfection, either Lipofection or Electroporation was used. During Lipofection, the genome editing factors were delivered via a Liposome. In contrast, the electroporation system takes advantage of an electric pulse that briefly creates pores in the cell membrane through which the factors can enter the cell (Figure 18).



**Figure 18 Transfection methods.** Lipofection delivers editing factors via formation of a liposome-complex. During electroporation, the membrane is shortly depolarized which forms small pores in the membrane through which the editing factors can be delivered into the cell.

### 2.2.1. CRISPRMAX Lipofectamine

For lipofection, cells were either dissociated into single-cell suspension using Accutase (Invitrogen, 1:3 dilution in PBS) or kept attached.



**Figure 19 Transfection using CRISPRMAX Reagent.** First, the RNP complex is formed, before formation of the liposome complex, the ssODN is added. The final Liposome-RNP-ssODN-complex is added to either suspension cells or attached cells. Adapted from [94].



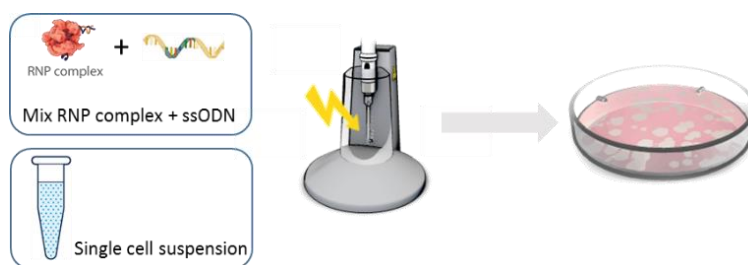
For the transfection of 500,000 cells, Lipofectamine™ CRISPRMAX™ Cas9 Transfection Reagent (Thermo Fisher Scientific) was added according to the following set-up:

**Table 3 Set-up for Lipofection of iPS cells** using Lipofectamine™ CRISPRMAX™ Cas9 Transfection Reagent (Thermo Fisher Scientific). This set-up was applied for transfecting 500,000 cells. The volumes were adjusted to other cell numbers.

Tube A	Tube B
125 µl OptiMEM (Invitrogen)	125 µl OptiMEM (Invitrogen)
0.8 µl Cas9 (5 µg / µl, TrueCut™ Cas9 Protein v2, Thermo Fisher Scientific)	0.5 µl ssODN (1 µg / µl, IDT)
1 µl sgRNA (150 pmol / µl, Synthego)	7.5 µl Lipofectamine CRISPRMAX Reagent
5 µl Lipofectamine Cas9 Plus Reagent	
5 min. separate incubation	5 min. separate incubation
Mix Tube B to Tube A and incubate for at least 5 min	

### 2.2.2. NEON Electroporation system

For electroporation purposes, the NEON® system (Thermo Fisher Scientific) was used. Cells were dissociated into single-cell suspension using Accutase (Invitrogen, 1:3 dilution in PBS). The RNP complex was formed in the provided R buffer and incubated for at least 10 min. prior to electroporation.

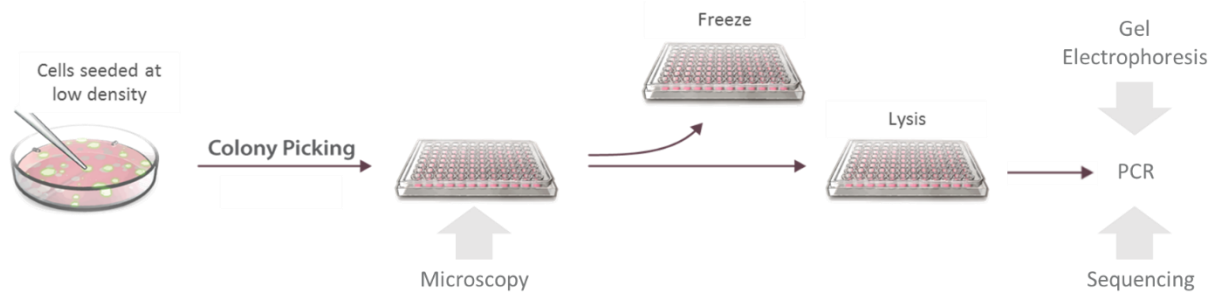


**Figure 20 Transfection using NEON electroporation System.** First, the RNP complex is formed in R Buffer. After adding the ssODN, the editing factors are mixed with the cell suspension. Upon mixing, the cells are electroporated immediately and plated. Adapted from [94].

The electroporation was adjusted according to the experimental set-up. The ratio between Cas9 (TrueCut™ Cas9 Protein v2, Thermo Fisher Scientific) and gRNA (Synthego) was kept at 1:6. The ssODN (IDT) was added 5 min. after the Cas9 and the gRNA to promote the RNP complex formation.

### 2.3. Low density seeding, clone picking and passaging

For purposes of single clone analysis, cells were fully dissociated using Accutase (Invitrogen, 1:3 dilution in PBS), counted and re-plated at low density. Densities between 7,500 to 15,000 cells per 10 cm dish were used. After at least 6 days, solid colonies formed from each single cell (clone) that were picked into a 96 well plate. Fluorescent clones could be analyzed at this point. After another period of expansion, when all of the colonies reached above 60% confluency, the clones were passaged. For passaging, the cells were accutased and passaged into a freezing plate and a lysis plate.



**Figure 21 Passaging.** After low density seeding the formed cell colonies are picked into a 96 well plate. Fluorescent cells can be analyzed at that point. After reaching about 60% confluency, the clones are passaged into a freezing plate and a lysis plate. The clones in the freezing plate can be thawed any time. The DNA extracted from the lysis plate is amplified using PCR and further analyzed using gel electrophoresis or Sanger sequencing. Adapted from [94].

For cell lysis, 10  $\mu$ l of lysis buffer (4x Lysis Buffer: 10 mM Tris-HCl at pH8, 2% Triton X, 1mM EDTA and 1% freshly added Proteinase K) was added to 30  $\mu$ l of the cell suspension. The lysis was performed in the thermocycler using the following program: 60°C, 60 min; 95°C, 10 min; 4°C, hold. The lysate was diluted to a final volume of 100  $\mu$ l. 12.5  $\mu$ l was used for subsequent PCR.

## 2.4. PCR, Sanger Sequencing and Analysis

The region of interest was amplified using Polymerase Chain Reaction (PCR). The template for PCR was either purified DNA, purified by QIAamp DNA mini kit (Qiagen) or taken directly from diluted lysates. The PCR conditions were optimized according to the region of interest. The following polymerase enzymes were used: Taq with 10x Standard Taq Buffer (New England Biolabs, NEB), Phusion (Phusion High-Fidelity DNA Polymerase, Thermo Fisher Scientific or NEB) with 5x HF Buffer or 5x High GC Buffer and dNTPs (NEB). If required, DMSO (NEB) or MgCl<sub>2</sub> (NEB) was added to enhance the amplification. The primers were ordered from Integrated DNA Technologies (IDT). The quality of the PCR product was checked using gel electrophoresis. For gel electrophoresis, the PCR product was loaded onto a 1.5 – 2% agarose gel together with a 1kb ladder for analysis of product size.

In case of nonspecific amplifications, the amplicon of interest was cut out of the gel and purified using QIAquick Gel Extraction Kit (Qiagen). In most cases the PCR product was purified using QIAquick PCR Purification Kit (Qiagen). Sequencing was outsourced and performed by Genewiz. Sequencing files were examined using Geneious.

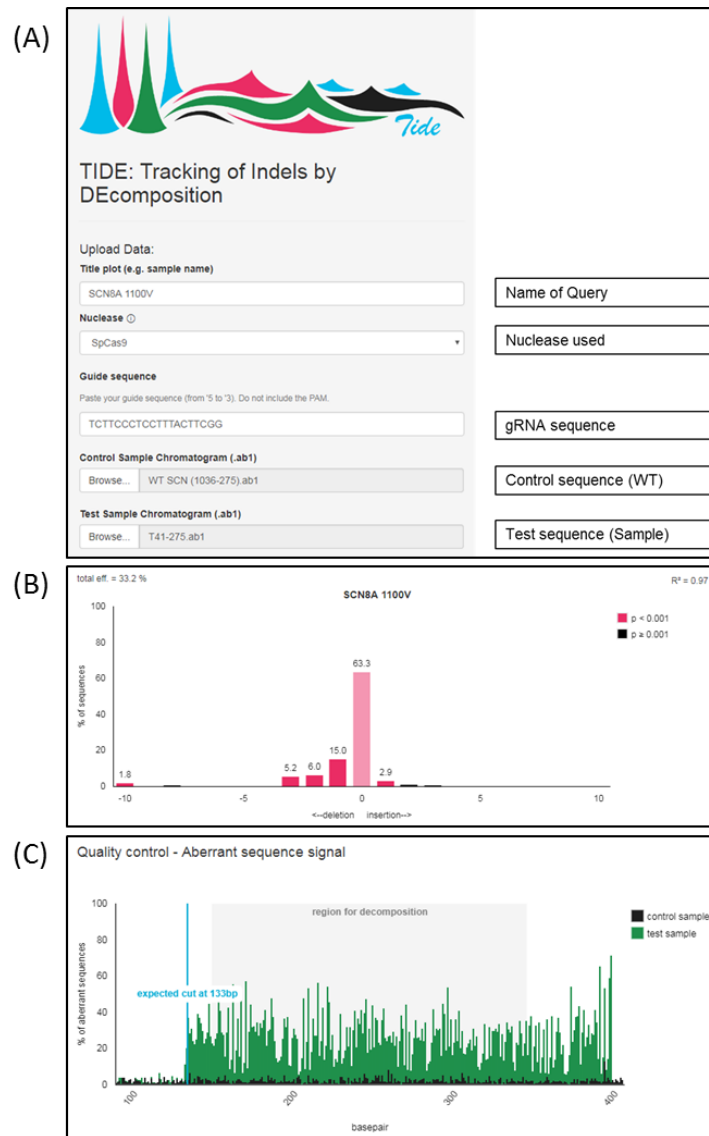
**Table 4 Primer sequences and respective purpose.** A pair of forward (F) and reverse (R) primer is required to amplify a region of interest. For sequencing, an in-sequence-primer is used.

Primer Sequence	Purpose
TTCACCGGGGTGGTGCCCAT	PCR amplification, GFP region (F)
TTGCCGTCCTCCTTGAAGTC	PCR amplification, GFP region (R)
GACGTAAACGGCCACAAGTT	Sequencing GFP region (F)
ATCCCAGCATACTCCAAGAC	PCR amplification, SCN8A region (F)
AATCCCTGTTCTTCTGCTGT	PCR amplification, SCN8A region (R)
GAGCTGAATGATTATCGCGG	Sequencing SCN8A region (F)

Cutting efficiency was evaluated using TIDE or ICE analysis. The sequences were analyzed in .ab1 format. The sequences were compared to a control sample that was sequenced with the same sequencing primer.

## 2.5. TIDE Analysis

Tracking of Indels by Decomposition (TIDE) is an online-based web tool which rapidly assesses cutting efficiency. Based on the quantitative sequence trace data from two standard capillary (Sanger) sequencing reactions (.ab1 File, DNA Electropherogram File), the TIDE software quantifies cutting efficacy and identifies the predominant types of indels of a cell pool [97].



**Figure 22 Tracking of Indels by Decomposition (TIDE) Analysis.** Example input and output. (A) Input of the required sequences and information in the online web tool. (B) Indel spectrum of the query. The total efficiency is indicated at the top left. (C) Quality control – Aberrant sequence signal, showing the test sample signals vs. the signals of the control sample and indication of the expected cut site [97].

## 2.6. ICE Analysis

Inference of CRISPR Editing (ICE) Analysis is an open-source, free to use web-based tool (<https://ice.synthego.com>) provided by the company Synthego. Similar to TIDE Analysis, it identifies the indel composition in a targeted cell pool in order to estimate the cutting efficiency. Moreover, batch analysis is possible with this software.



**Figure 23 Inference of CRISPR Editing (ICE) Analysis.** Example input and output. (A) Input of the required sequences and information in the online web tool. (B) Summary of the query. The total efficiency is indicated at top (field called “ICE”). (C) Discord & Indel chart –sample signals vs. the signals of the control sample and indication of the expected cut site. The Indel spectrum is depicted on the right.

## 2.7. GFP to BFP conversion

The HSCI iPS core has generated eGFP-expressing pluripotent stem cells to assess gene editing efficiency [98] by measuring the GFP to BFP conversion. This quick and elegant assay can be used to optimize the CRIPSR/Cas9 protocols used in the lab [98]. As illustrated in Figure 24, one bp (196 T>C) change is enough to convert GFP into BFP (Y66H). Two additional bp changes were introduced into the ssODN. One of the additional mutations (194 C > G) increases the fluorescence intensity. It also mutates the PAM sequence to avoid re-cutting after editing. The third bp change (201 C>G) introduces a mismatch to prevent re-binding of the sgRNA to the edited sequence.



**Figure 24 GFP to BFP conversion.** A cell expressing GFP can be targeted via genome editing to turn into a BFP-expressing cell. Only one bp change is enough to convert GFP into BFP. Except for this essential 1 bp change, the introduction of two further silent mutations prevents re-cutting by the Cas9 enzyme and increases the fluorescence intensity (red arrows indicate the introduced modifications after HDR using the designed ssODN).

### 2.7.1. Oligo Design

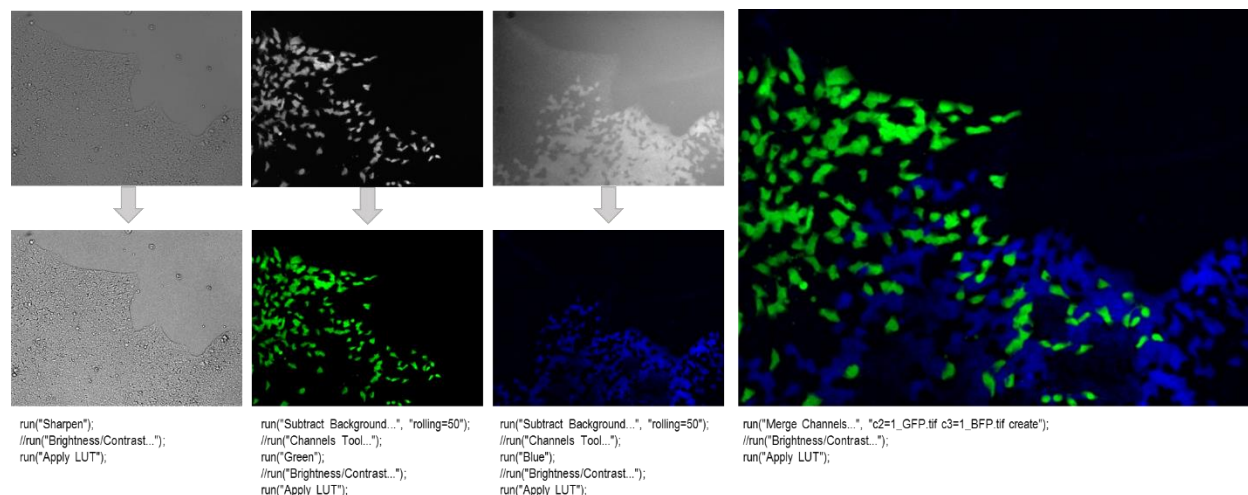
For the BFP conversion, the ssODN depicted in Figure 24 was designed to introduce the one amino acid change to convert GFP into BFP. For the experiment where two oligos were introduced, another ssODN (with silent mutations) corresponding to the GFP sequence was designed. The sequences used are shown in Table 5.

**Table 5 ssODNs used for the GFP to BFP conversion experiments.** The bold letters indicate the silent mutations whereas the blue letters indicate the required sequence for BFP and the green letters the respective sequence for GFP expression.

	Sequence
<b>WT GFP at region of interest</b>	CCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACC CTG <b>ACC/TAC/GGG</b> GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTT CTTCAAGTCCGCCAT
<b>ssODN encoding BFP</b>	CCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACC CTG <b>AGC/CAC/GGG</b> GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTT TCTTCAAGTCCGCCAT
<b>ssODN encoding GFP</b>	CCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACC CTG <b>ACG/TAC/GGG</b> GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTT TCTTCAAGTCCGCCAT

## 2.8. Microscopy and Image processing

Live-cell images were acquired using an Olympus IX71 microscope. Pictures were taken in Brightfield (BF), green and blue channels. The respective images were taken separately using the proper filters. The pictures were processed using ImageJ. Background subtraction, Brightness/Contrast adjustment and Merge were accomplished using the following Macro codes (Figure 25):



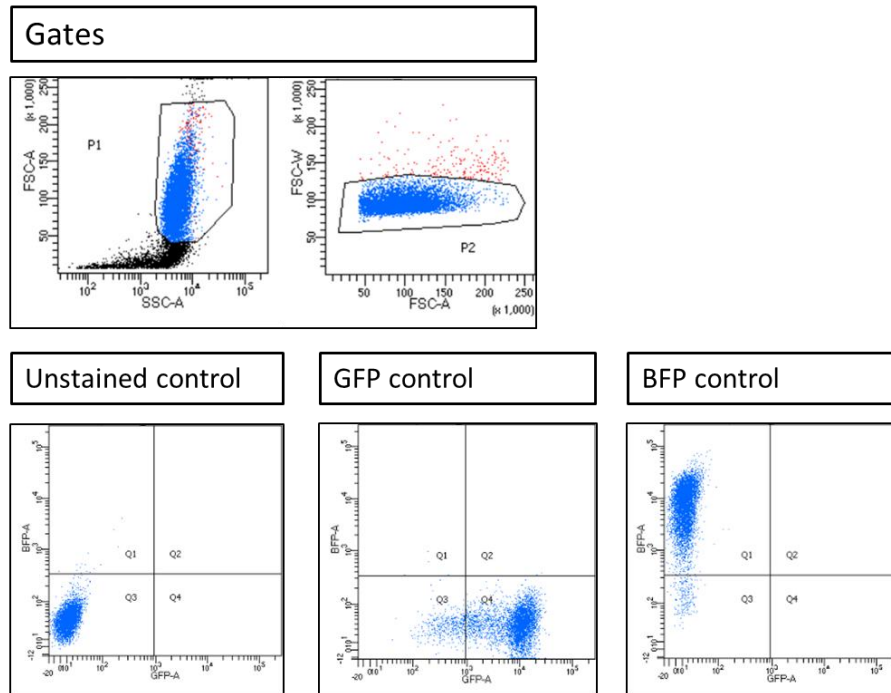
**Figure 25 Image processing with ImageJ software.** The upper panel shows the acquired pictures before processing. The lower panel shows the images after the indicated processing step. The rightmost image shows the merge of the separately acquired pictures in the respective channels. The codes below describe the Macro code used in ImageJ for the respective adjustments.

## 2.9. FACS Analysis

The cells were dissociated into single-cell suspension using Accutase (Invitrogen, 1:3 dilution in PBS). Before fixing, the cells were washed with PBS. For fixing, the cells were incubated for at least 15 min. in 4% PFA (Paraformaldehyde) at RT. After another washing step, the cells were resuspended in PBS with 1% FBS and transferred into FACS blue tubes (VWR) through the filter lids.

For adjustment of the FACS machine, GFP-positive cells were used as GFP-positive samples. Cells that did not contain any fluorophore were stained with DAPI and used as BFP-positive controls.

Viable cells were gated using forward and side scatter. The GFP expression was measured using a 488 nm laser for excitation. For BFP detection, a 383 nm laser was used.



**Figure 26 FACS Analysis.** The population of living cells was selected by setting the appropriate gates in the side-scatter to forward-scatter analysis. The laser settings were adjusted using unstained cells, green-fluorescent cells (GFP control) and blue-fluorescent cells (BFP control) as depicted in the lower panel of graphs.



## 2.10. gRNA design

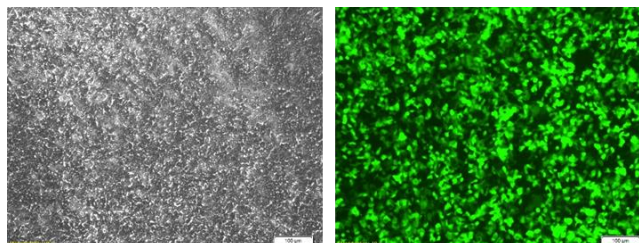
### 2.10.1. gRNA design tool from MIT

Guide RNA design was achieved using a CRISPR design web tool created by the Zhang Lab at the Massachusetts Institute of Technology. It is open-source and can be found under this link: [crispr.mit.edu](http://crispr.mit.edu). The web tool uses a specific algorithm to scan a sequence of choices for possible guides (20 nucleotides followed by a PAM sequence) and then scan for possible off-target matches throughout the selected genome (human genome for hiPSCs). Then the web tool assigns scores to the guides based on the possible off-target effects genome-wide and by target specificity.

Because the guides designed by this web tool differ in their properties, several guides were chosen and tested for their ability to target the right sequence. The guides were ordered as modified synthetic single guide-RNA (sgRNA) from Synthego.

### 2.10.2. sgRNA efficiency testing

Cutting efficiency testing was performed in HEK 293T cells, which are far easier to transfect than stem cells. RNP delivery via Lipofectamine™ CRISPRMAX™ Cas9 Transfection Reagent (Thermo Fisher Scientific) was performed as stated above. The performance was compared to delivery of Cas9 via a plasmid using Lipofectamine 3000 according to the suppliers' recommendations. The plasmid contains GFP, which was used to check the transfection efficiency the day after transfection.

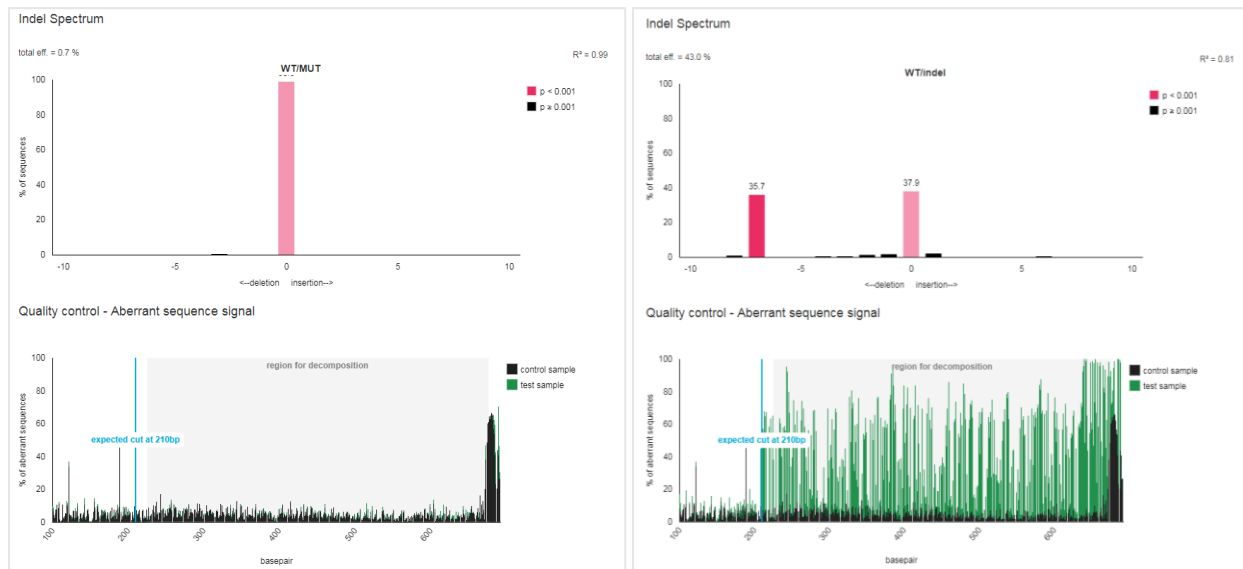


**Figure 27 Transfection efficiency testing** after plasmid delivery in HEK 293T cells. GFP-positive cells were successfully transfected with the plasmid that contains the sequence encoding the Cas9 protein.

After 72 hours, the cells were harvested, centrifuged and the cell pellet was lysed in order to isolate the DNA. Upon PCR amplification of the respective region, the PCR product was sequenced and the cutting efficiency was determined using TIDE and ICE Analysis as described above.

## 2.11.Clone Analysis

Upon lysis, the DNA of each individual clone was amplified using PCR and sequenced. The sequences were analyzed with Geneious software via screening for the KI-sequence (sequence of the ssODN). For the analysis of the indel contribution, the sequence of the respective clone was compared to the control sequence using TIDE or ICE analysis as depicted in Figure 28.



**Figure 28 Clone Analysis. Determination of indel contribution.** Depicted are the TIDE Analysis results of an unmodified clone (left panel) vs. a clone that was cut and repaired via NHEJ, resulting in indels (right panel). Analysis of the indels showed that one allele remained intact whereas the other allele lost 7 base pairs (-7).

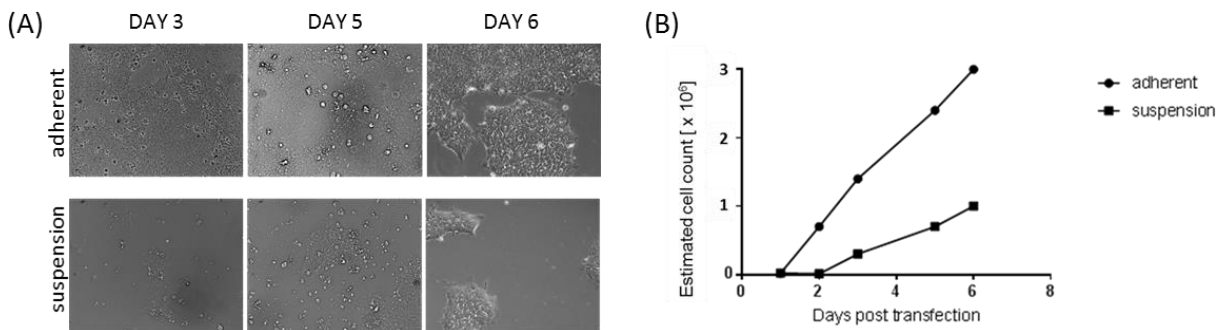
### 3. Results

#### 3.1. Optimization of RNP delivery

Literature states that transfection via electroporation is more efficient than lipofection-based transfection methods [99]. We decided to test these two delivery methods in iPSCs. CRISPRMAX lipofection and electroporation with NEON were optimized and compared for their efficiencies. GFP to BFP conversion was used to evaluate the editing efficiency.

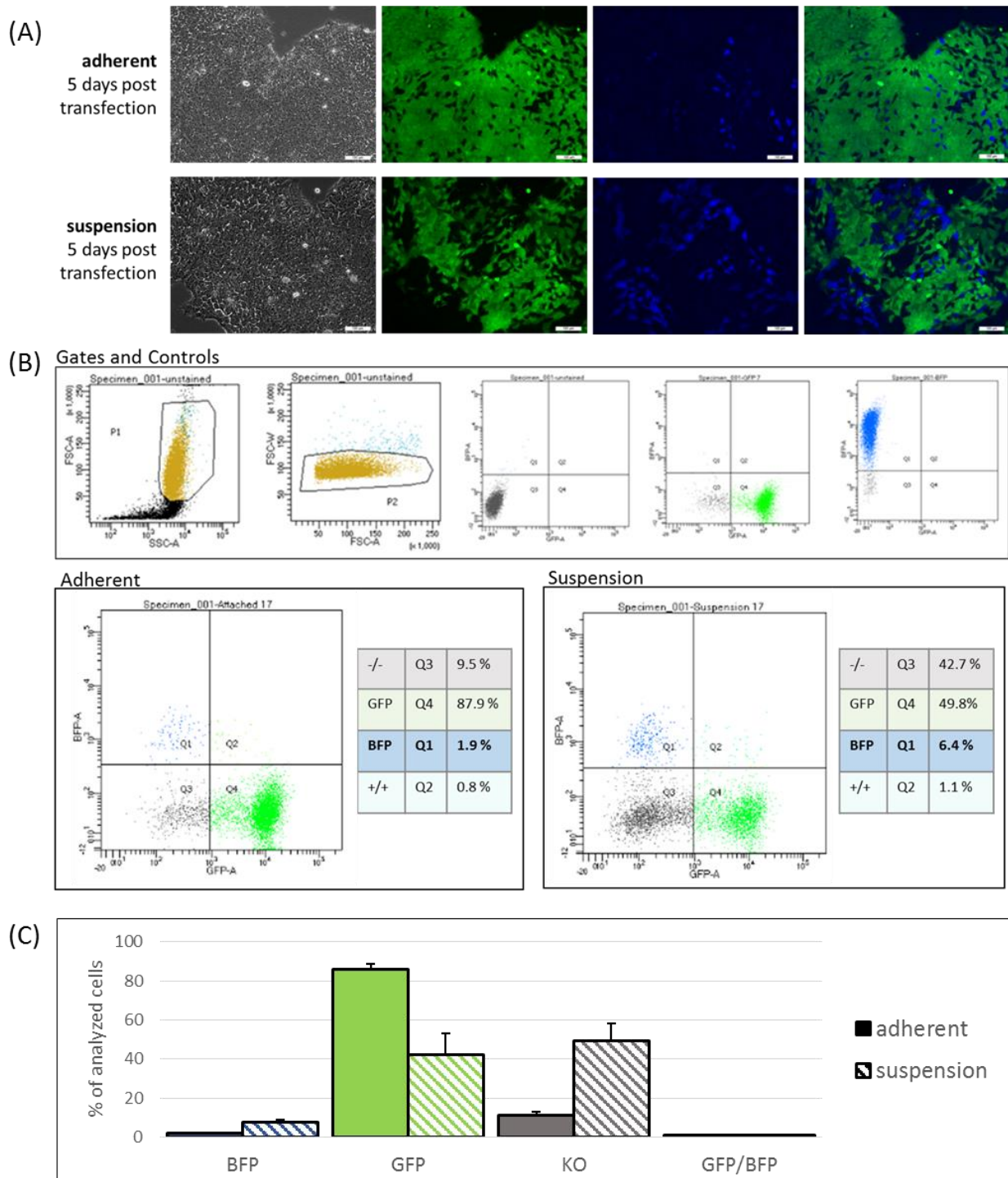
##### 3.1.1. Optimization of CRISPRMAX lipofection system for RNP delivery

For the optimization of RNP delivery with CRISPRMAX Lipofection, homozygous GFP (GFP/GFP) cells were used (Ery ED3 siPS-A AAVS1 GFP). Cells in suspension yield a larger surface area and, consequently, higher accessibility for cationic lipid-mediated endocytosis. To test this hypothesis, CRISPRMAX lipofection in adherent cells versus cells in suspension was examined. First, the effects of the transfection method on the cells' survival was examined. As shown in Figure 29, the cell survival was lower when the cells were transfected in suspension.



**Figure 29 Survival post transfection of cells that were adherent vs. in suspension** during the transfection procedure. (A) Microscopy pictures of transfected cells (B) Graphical illustration of the estimated cell count showing the delay of recovery of cells transfected in suspension.

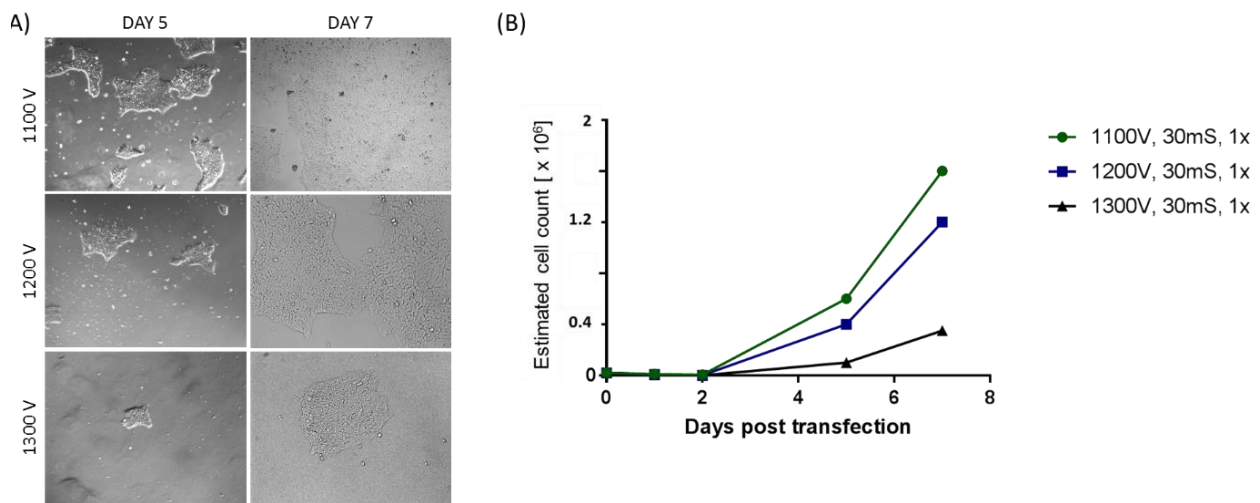
Second, the delivery efficiency was measured using the GFP to BFP conversion. Successful cutting of the gene results in the loss of functional GFP and green signal (KO event). Introduction of the ssODN converts GFP to BFP which gives a blue signal (KI event). Succeeding visual evaluation via fluorescence microscopy, the percentage of KO and KI events was determined by FACS analysis. Overall, the gene editing efficiency was higher when the cells were transfected in suspension (Figure 30).



**Figure 30 Lipofection with CRISPRMAX in homozygous GFP cells** transfected in an adherent state vs. in suspension. (A) Microscopy 5 days post transfection comparing the amount of GFP-positive, GFP-negative and BFP-positive cells. (B) FACS Analysis 6 days post transfection. The upper panel displays the gate set-up and the respective controls: unstained control, GFP control and BFP control. Below, adherent cells were compared to cells in suspension. The BFP-positive cells represent a successful KI event, indicated in bold. n=2.

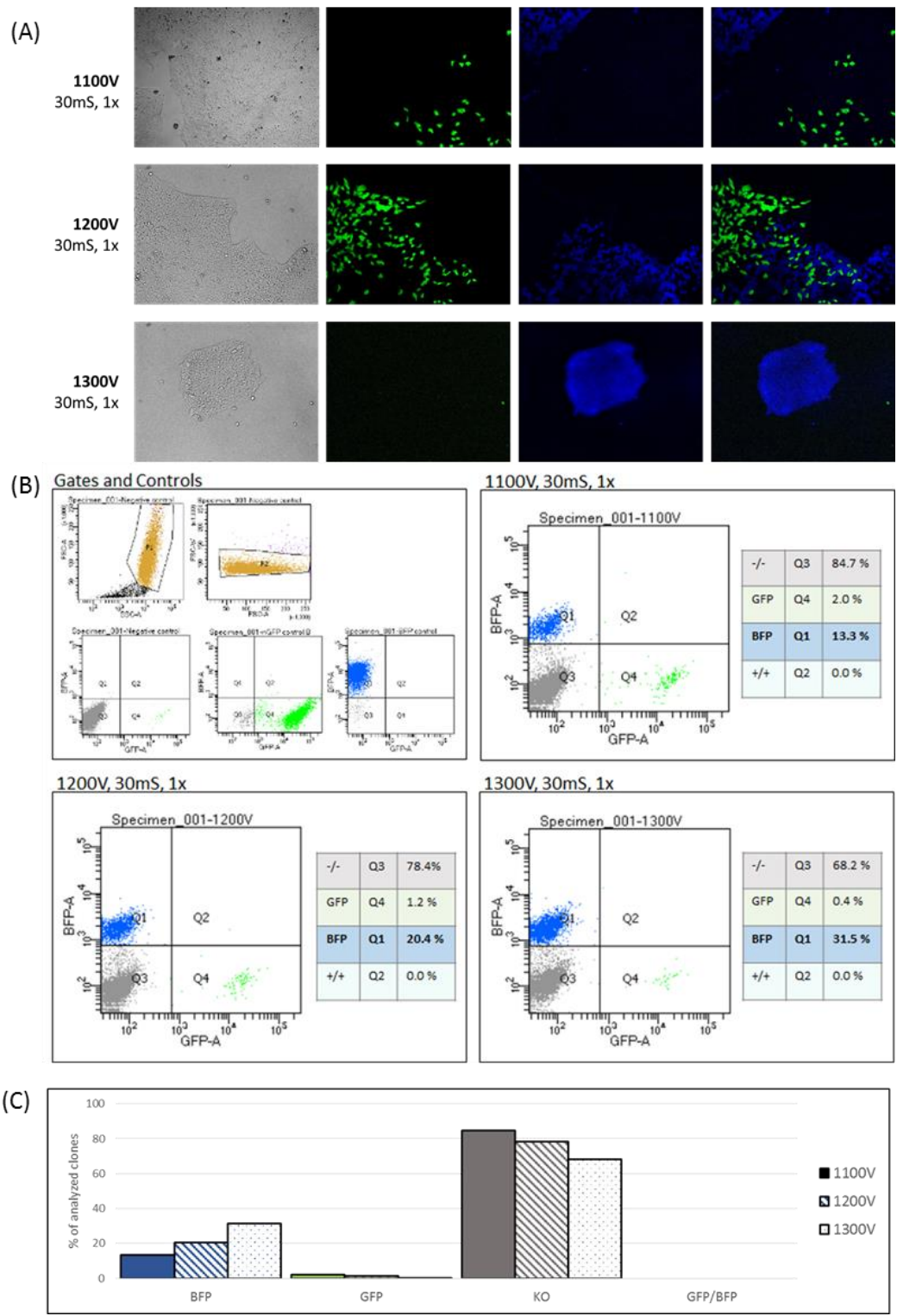
### 3.1.2. Optimization of NEON electroporation system for RNP delivery

As described in the Materials and Methods section above, the general NEON set-up was done according to the supplier's instructions and a range of the most commonly used voltages for electroporation in iPSCs was tested and compared. A heterozygous GFP-positive cell line (HUES8 GFP F with one copy of GFP in the AAVS1 locus) was used to optimize the conditions to yield the highest transfection efficiency [96]. As depicted in Figure 31, the survival after applying each condition was examined first. At higher voltage (1300V) the survival was significantly decreased.



**Figure 31 Survival post transfection of cells that were transfected using NEON electroporation system** applying different voltages. (A) Microscopy pictures taken from the respective voltage applied at day 5 and day 7 (B) Graphical illustration of the estimated cell count showing the delay of recovery proportional to the voltage applied.

Then the delivery efficiency was evaluated using fluorescence microscopy and confirmed by FACS Analysis. The yield of KO events as well as KI events increased with the increase of applied voltage. Microscopy pictures display the decrease of survival and the proportional increase of BFP-positive cells, representing a successful KI event (Figure 32, A). FACS Analysis results agreed with these findings. The higher the voltage, the higher the yield of edited cells. Overall, the highest percentage of BFP-positive cells were found after transfection with the highest voltage but the survival was very low as well. The transfection with the settings 1200V, 30mS, 1x pulse delivered the best outcome in terms of survival to efficiency ratio.



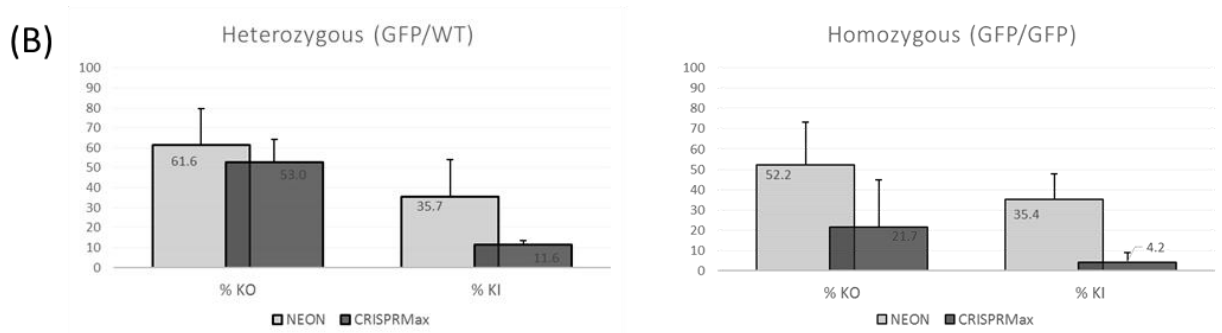
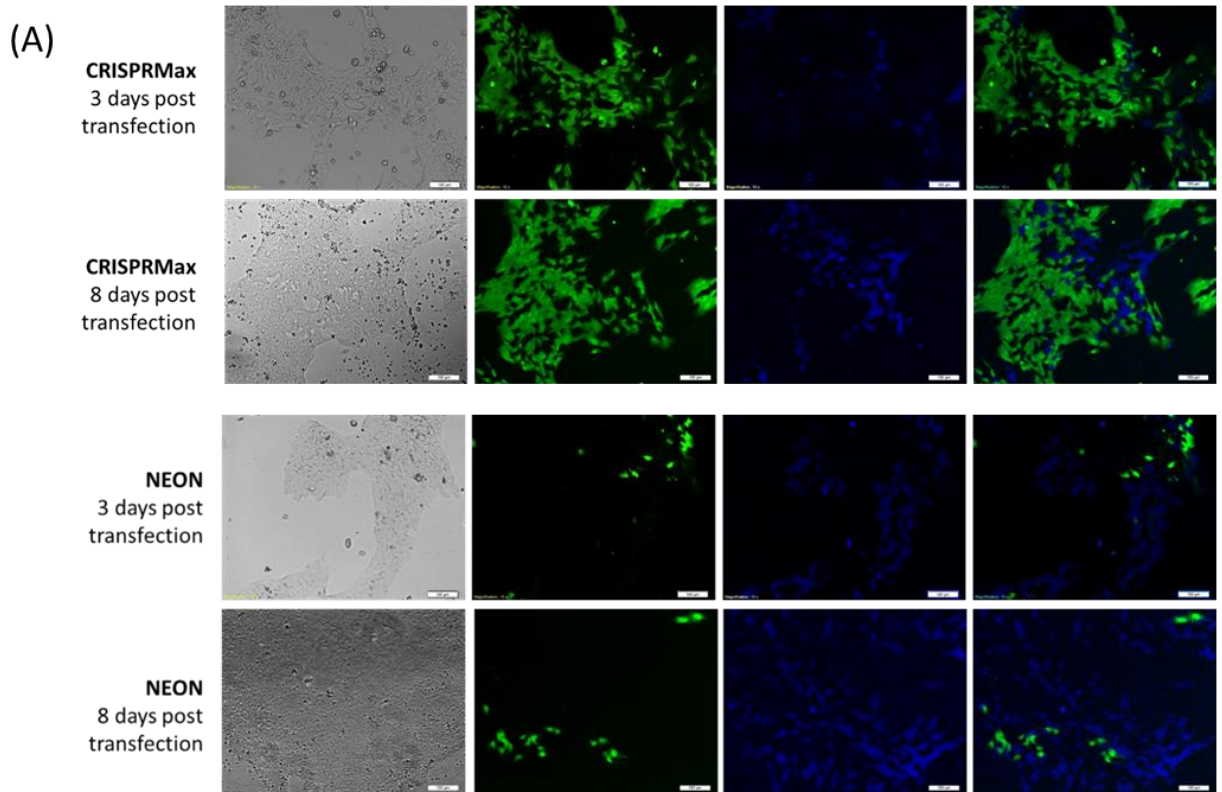
**Figure 32 NEON Electroporation in heterozygous GFP cells.** (A) Microscopy 7 days post transfection comparing the amount of GFP-positive, GFP-negative and BFP-positive cells. (B) FACS Analysis 9 days post transfection. The upper left panel displays the gate set-up and the respective controls: unstained control, GFP control and BFP control. Different voltage settings were compared. (C) Graphical illustration of the FACS data. The BFP-positive cells represent successful KI event, indicated in bold. n=1.

### **3.1.3. Transfection using the NEON Electroporation system results in higher KI yield than using CRISPRMAX Lipofection**

For comparison of NEON electroporation versus CRISPRMAX lipofection as delivery methods, the efficiencies were tested in both, a heterozygous and a homozygous GFP line (HUES8 GFP 002, HUES8 GFP F). The optimized protocols for each delivery method were compared. Each set-up was repeated 4 times and an average was taken to compare the editing efficiencies in both cell lines. The overall editing distribution was similar as described in the sections 3.1.1. Optimization of CRISPRMAX lipofection system for RNP delivery and 3.1.2. Optimization of NEON electroporation system for RNP delivery. For comparison, we focused on the percentage of KO and KI events.

When analyzing the KO and KI events, the NEON electroporation system was found to yield a higher KO and higher KI percentage compared to CRISPRMAX lipofection (Figure 33). For both cell lines, using the NEON electroporation system a KI yield of above 35% was achieved. The CRISPRMAX system delivered poorer editing success when compared to the NEON system. The difference was even more striking in the homozygous cell line.

The NEON system delivered consistent editing results in the heterozygous as well as in the homozygous cell line. Interestingly, while conducting these experiments, we noticed a very small percentage of cells were GFP+/BFP+ for the homozygous line, indicating that it is very infrequent to target only one allele.



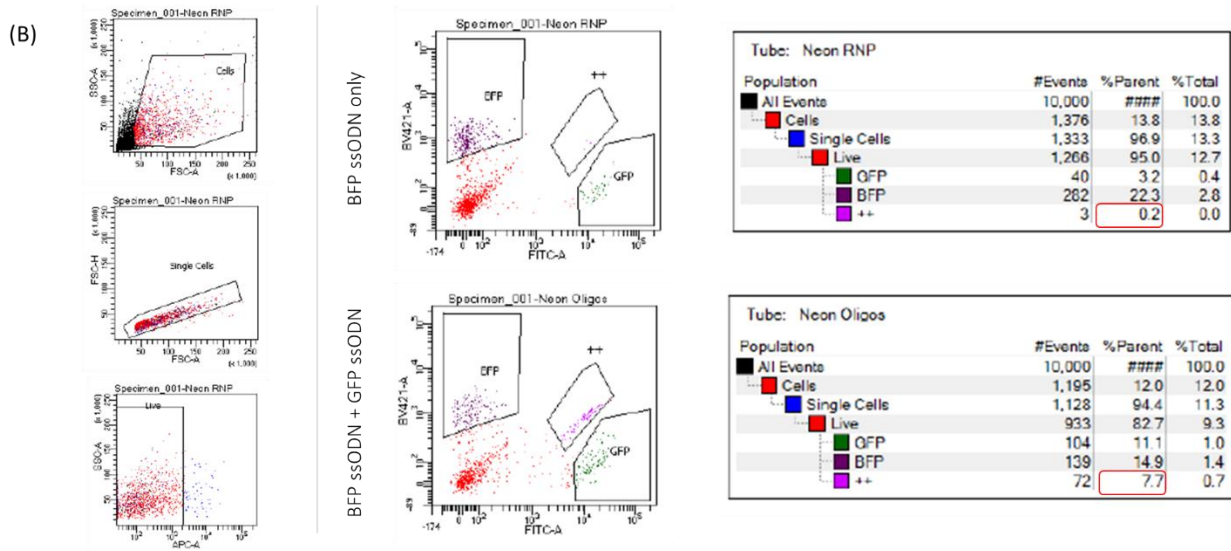
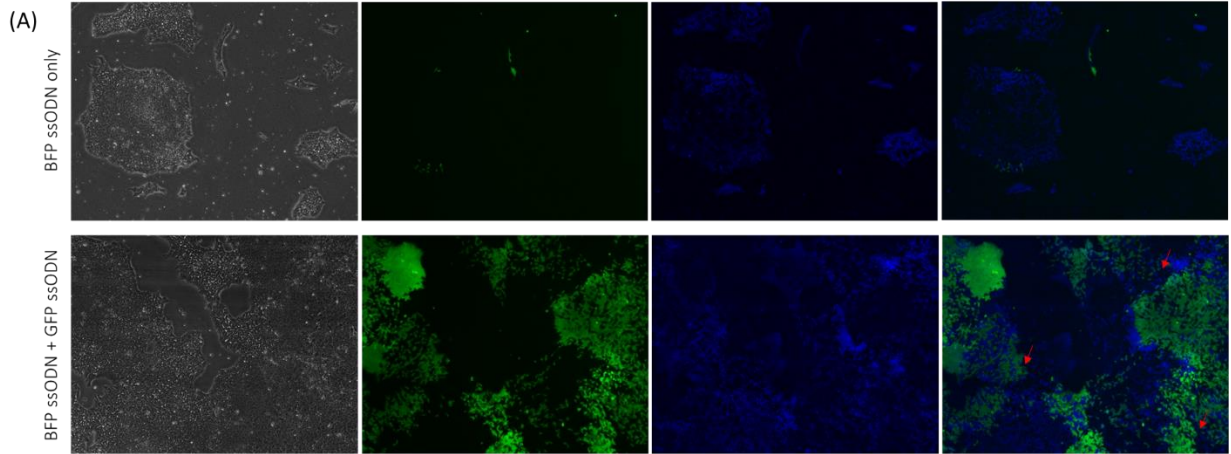
**Figure 33 Comparison of CRISPRMAX Lipofection vs. NEON Electroporation.** The same batch of cells was dissociated and transfected using the respective method and analyzed using microscopy and FACS. (A) Imaging was performed 3 days and 8 days post transfection. The image shown is an example from an experiment in a homozygous GFP line. (B) The graphs show the percentage of KO and KI events for the heterozygous line (left graph) vs. the homozygous line (right graph). The experiment was repeated 4 times for each set-up, n=4.



#### **3.1.4. Delivery of two ssODNs increases the yield for WT/KI**

As already described in Figure 33, the event of single-allelic editing was found to be very rare. However, a number of gene editing projects conducted at the HSCI iPS core require the repair of heterozygous iPS lines or the introduction of a mutation exclusively to one allele. If the editing is too efficient, both alleles are targeted, and WT/KI cells are almost absent. Based on the data presented in Section 3.1.3, when one allele is targeted, it is very unlikely that the second allele is kept uncut. Therefore, we proposed to use two ssODNs (one with BFP sequence and one with GFP sequence) to increase our chance to obtain BFP+/GFP+ cells (relying on two KI events). Homozygous GFP-positive cells (HUES8 GFP 002) were transfected using the NEON electroporation system. The cells were treated exactly the same way as in the previous experiment except one batch received only one type of ssODN encoding the BFP sequence whereas the other batch received two ssODNs where one encodes the BFP sequence and the other a modified version of GFP (see Materials and Methods section above).

When comparing both set-ups, the number of GFP-positive cells was found to be higher using the two ssODNs, suggesting that ssODNs encoding GFP were used as repair templates. But more importantly, upon examination of the GFP and BFP signals, the overlay of the respective channels displayed a detectable amount of double-positive cells (Figure 34, A). In this situation, both types of ssODN were present and used by chance for DSB repair. FACS sorting confirmed these findings (Figure 34, B). The percentage went up from <1% with the BFP ssODN to 7.7% with both ssODNs.



**Figure 34 Comparison delivering proteinaceous RNP with either one ssODN or two ssODNs.** One ssODN for BFP (BFP ssODN only, upper panel) or two ssODNs for BFP and GFP (BFO ssODN + GFP ssODN, lower panel) were used. Both ssODNs were designed with silent mutations to prevent re-cutting by the Cas9 enzyme. The overlay of fluorescence microscopy images taken from GFP- or BFP-positive cells displays that cyan-colored cells represent double-positive cells (GFP+/BFP+, red arrows). (B) FACS Analysis of the double oligo approach. The left panel depicts the gates and selection of surviving cells after Sytox Red staining. The gate labelled “++” indicates double-positive cells (red frames).

### 3.2. Repair of a heterozygous mutation in a patient-derived iPS line.

The gene SCN8A (sodium voltage-gated channel alpha subunit 8) encodes a protein essential for sodium passage into the cell. If mutated, it was found to be strongly linked to Infantile Epileptic Encephalopathy [100]. Cells were derived from a patient with a heterozygous mutation in the SCN8A gene.

The starting cell (WT/MUT) inherits a heterozygous mutation where one base is mutated: thymine is mutated to a cytosine (T → C), causing a switch from Threonine to Methionine (Thr → Met).

#### 3.2.1. sgRNA design

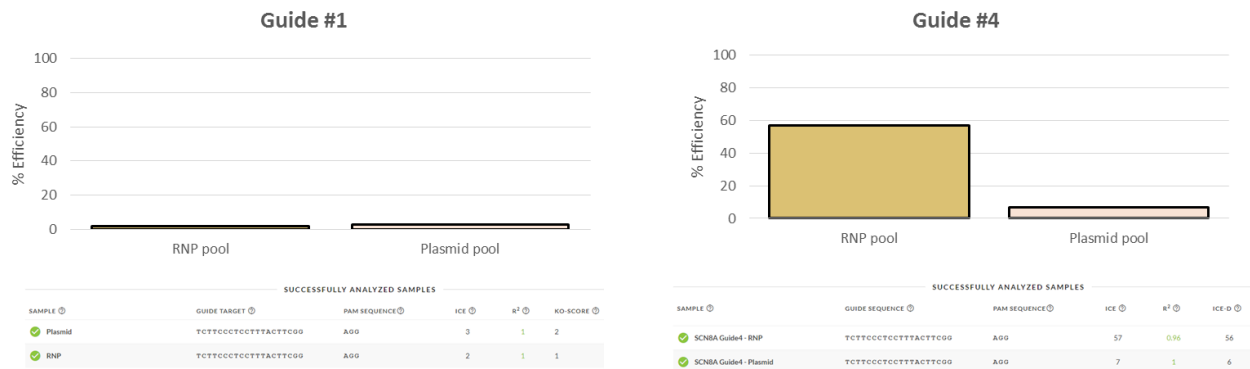
For the design of a fitting single guide RNA, the region of interest was screened using /crispr.mit.edu. It was shown that HDR efficiency increases when the double-strand break occurs close to the mutation. For this locus, there was no sgRNA screened by the web tool that spans the mutant allele exclusively. Five sgRNA with slightly different scores were screened. They are all located upstream of the mutation (Figure 35). Therefore we had to choose gRNAs that bind both alleles.



**Figure 35** sgRNA design for the SCN8A gene. First, the region of interest was defined. Second, the region was screened for binding gRNAs using /crispr.mit.edu. The binding region was examined and confirmed using Geneious software.

### 3.2.2. sgRNA testing

Two of these five sgRNAs were chosen and tested for successful cutting. As explained in the Materials and Method section above, HEK293T cells were transfected, DNA was isolated, the region of interest was amplified using PCR and sequenced before ICE analysis of the pool. The cells were transfected using Lipofection (CRISPRMAX for proteinaceous RNP delivery, Lipofectamine 3000 for plasmid delivery). Figure 36 depicts the ICE analysis results after HEK 293T cell transfection with the respective gRNA. Guide 1 was shown to cut poorly while guide 4 introduced a DSB in more than 50% of the cells.

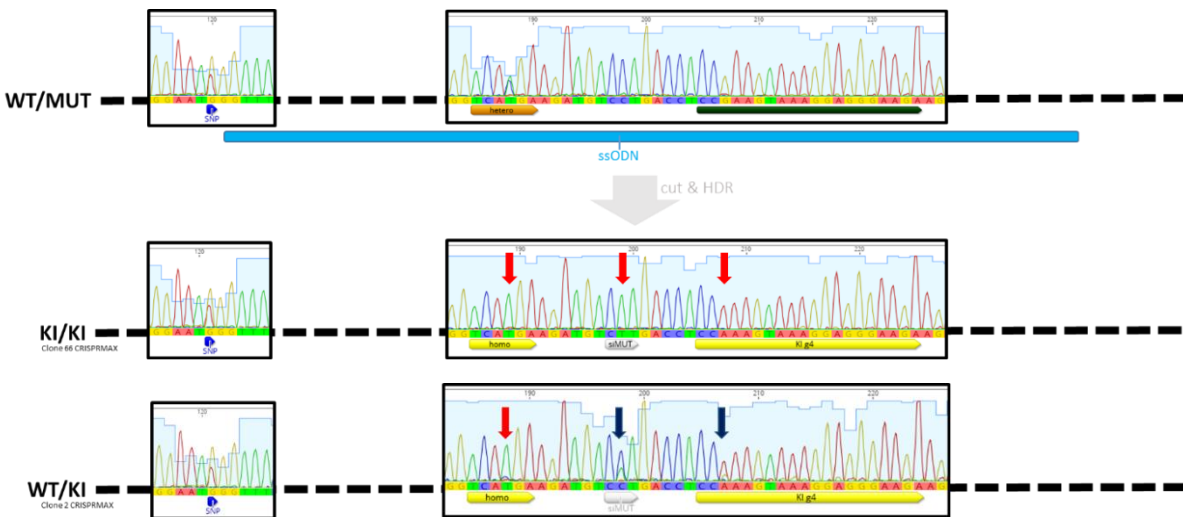


**Figure 36 sgRNA testing using ICE Analysis (Synthego).** Here, the efficiencies of the guide RNAs were tested and both RNP and Plasmid deliveries were compared. The left panel shows the cutting efficiency after transfection with guide #1. The right panel depicts the ICE Analysis results after transfection with guide #4. The tables below each graph show the original ICE Analysis output.

### 3.2.3. ssODN design

As a repair template, a 120 bp-long ssODN was designed. We detected a SNP in this patient iPS line around 68 nucleotides upstream of the mutation (indicated in Figure 37). The ssODN does not span over the SNP sequence. The ssODN contains the repaired sequence at the site of mutation and two additional silent mutations to limit rebinding of the sgRNA to the sequence post editing. With these additional mutations, double KI (KI/KI) or single KI (WT/KI) can be differentiated. Figure 37 depicts the original genetic situation (WT/MUT) with the heterozygous mutation and the SNP. The heterozygous mutation will be targeted upon successful editing using the ssODN, whereas the SNP should remain unmodified. After the binding to the RNP complex and the introduction of a DSB, the cut site will be repaired by the cell. In case of HDR, not only the mutation site

will be changed (C/T → T) but additional two silent mutations will be introduced (C → T and G → A, indicated with red arrows). If a double KI occurs, the silent mutations will be present on both alleles (KI/KI), whereas a single KI event will be visible as a heterozygous modification (WT/KI).



**Figure 37 ssODN design. Illustration of the genetic locus of interest within the SCN8A gene.** Upon successful cutting and HDR, the KI event can be tracked using the introduced silent mutations. Double KI leads to a clean homozygous expression of the introduced base whereas a single KI scenario results in a heterozygous base change. The ssODN does not span the present SNP. The pictures are derived from Sanger sequencing files using an anti-sense primer and are displayed using Geneious software.

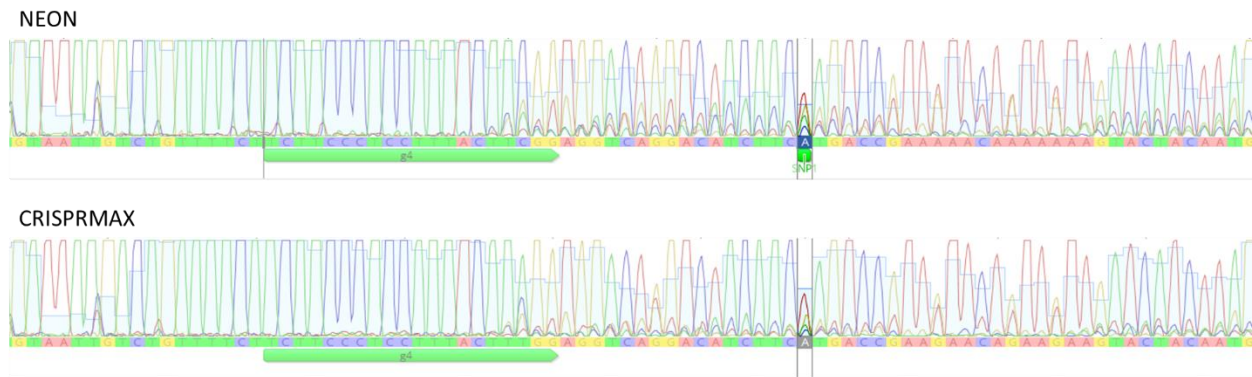
### 3.2.4. Pool Analysis

The mutated iPS cells (WT/MUT) were transfected using NEON electroporation system or CRISPRMAX Lipofection Reagent. 9 days post transfection, an aliquot of the cells was taken, DNA isolated and the region of interest amplified using PCR. The amplified PCR products of the cell pool were sequenced via Sanger sequencing. The editing efficiency was estimated via pool analysis using ICE, which compared the amplified 434 bp-long PCR product of the pool with the sequence of the same amplified region of the starting heterozygous mutant (WT/MUT). Pool analysis results indicated that 33% of the present DNA in the pool was successfully cut compared to 23% in the CRISPRMAX pool (Figure 38).



**Figure 38 ICE Analysis of transfected cell pools.** The cutting efficiency after transfection with NEON (in grey) or CRISPRMAX (in magenta) was evaluated. The table below depicts the original ICE Analysis output.

When working with the ICE software, we experienced a prediction bias of the algorithm. A KI event is not different enough to the WT sequence since they just differ from each other by 3 base pairs. The algorithm will therefore predict a KI event as unedited. To exclude the prediction bias of ICE Analysis the sequences of the pool were examined visually using Geneious software. Visual analysis of the transfected pools displayed the relative amount of the repaired sequence over the heterozygous mutation. The relative amount of repaired sequences was perceived to be higher in the CRISPRMAX pool compared to the NEON pool (Figure 39).

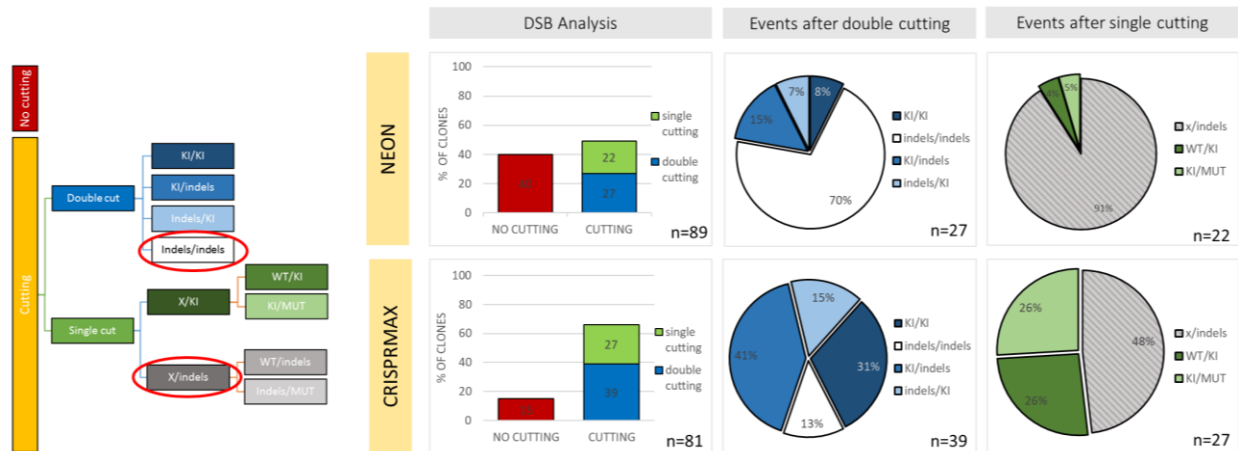


**Figure 39 Visual analysis of transfected cell pools.** Displayed are the Sanger sequencing files of the respective DNA locus. The upper panel displays the sequencing profile of the NEON pool and the lower panel displays the CRISPRMAX pool. The colored peaks indicate the relative amount of each individual base that was detected at the respective locus. The sense primer was used for sequencing – a higher relative amount of an A-base at the indicated locus indicates a higher proportional amount of KI events.

### 3.2.5. Clone Analysis

After transfection, cells were seeded in low density and individual clones were picked into one well of a 96 well plate each. For each condition, 96 clones were picked, sequenced and analyzed. The targeted sequence for each clone was analyzed manually using Geneious software and both ICE and TIDE Analysis for indel-identification. During low density seeding or clone picking the event of two cells merging can occur by chance. Some clones were identified as ‘mixed,’ referring to two clones that merged. Mixed cells were excluded from analysis.

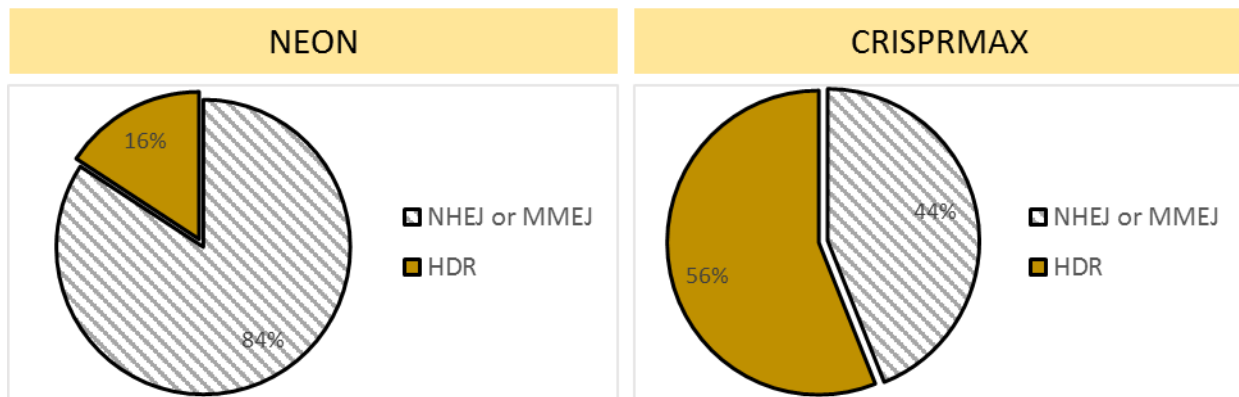
Upon clone analysis, the cutting efficiency was found to be higher using CRISPRMAX transfection compared to NEON electroporation. In both cases, targeting of both alleles (double cutting) was more likely than one allele only (single cutting). The indel ratio was higher in the cells that were transfected using NEON. When analyzing the events of a one-allelic modification, the likelihood of targeting was the same for both alleles (either WT or MUT). Overall, in the iPS cell line mutated in the SCN8A gene, the number of edited cells was higher after CRISPRMAX transfection (Figure 40).



**Figure 40 Clone Analysis of the modified genetic locus within the SCN8A gene.** The illustration on the left indicates the possible genetic situations present in the pool. If no editing occurs, the cells remain unmodified (no cutting, red). If editing occurs, the events can be categorized via the amount and locus of the cutting as well as the amount and locus of the KI event. The respective events were summarized and illustrated graphically. Overall cutting was higher in the CRISPRMAX pool, as well as the overall KI events. The percentage of indels was higher in the NEON pool.

### 3.2.6. DNA repair Analysis

As shown in Figure 40, more HDR events were detected after CRISPRMAX transfection compared to NEON electroporation. In Figure 41, the editing events were compiled according to their repair mechanism. The events were analyzed per allele (two events per clone). A successful editing event is dependent on HDR. In iPS cells that were treated with NEON transfection, we found that HDR was less likely to occur than in cells treated with CRISPRMAX. After NEON transfection, most alleles were repaired via NHEJ or MMEJ.

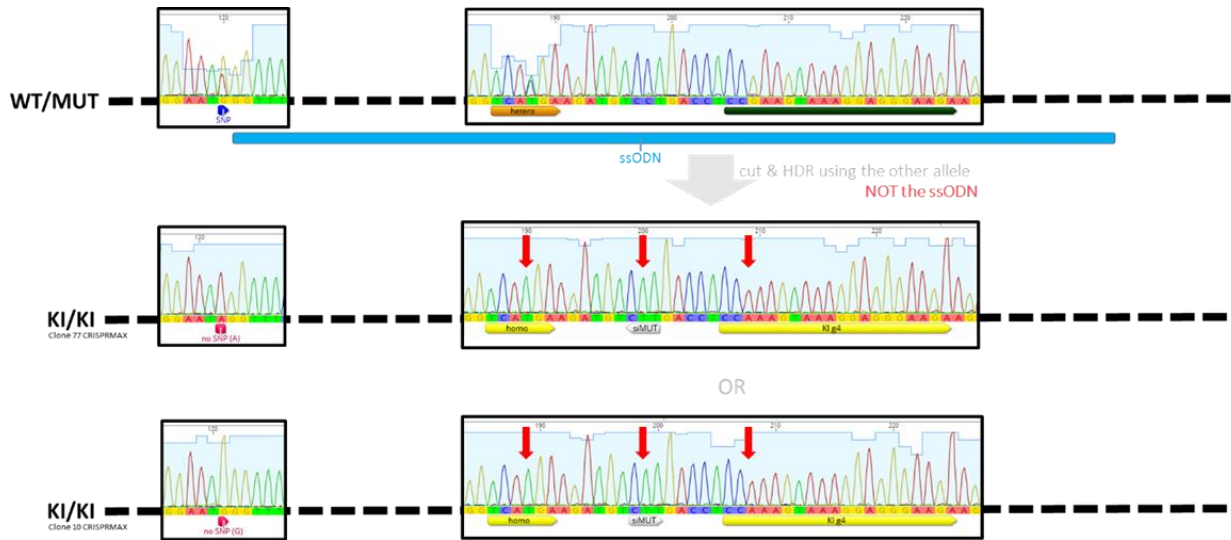


**Figure 41 Repair Analysis after NEON or CRISPRMAX transfection.** The likelihood of a HDR repair upon cutting was 16% in the NEON pool whereas CRISPRMAX transfected yielded a 56% chance of an HDR repair.

As mentioned earlier, the designed ssODN did not cover the SNP detected upstream of the mutation site. For this reason, changes in the SNP sequence were not expected. Surprisingly, we noticed some clones where the SNP was gone. If the SNP is absent, it can be assumed that the other allele was used as repair template instead of the provided ssODN. In such an scenario, one allele was successfully repaired using the ssODN and the DSB in the other allele was fixed using the repaired other allele.

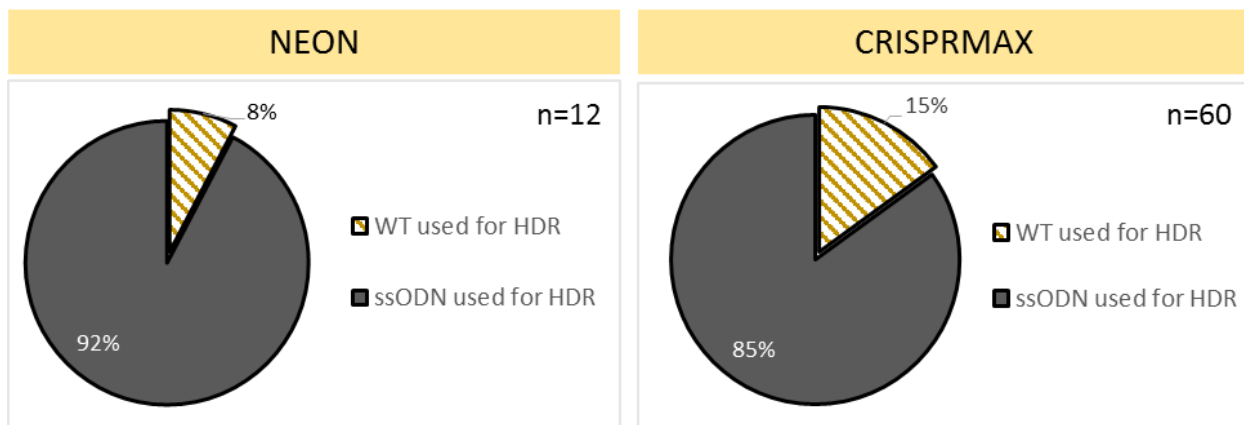
Upon sequencing, on the mutated allele, an A base is detectable whereas there is a G base on the WT allele. Figure 42 illustrates the two possible outcomes upon HDR using the other present allele. In both cases a double KI (KI/KI) is shown where either the MUT allele was used as repair template (Clone 77 CRISPRMAX) or the WT allele (Clone 10 CRISPRMAX).





**Figure 42 Non-ssODN-mediated HDR.** Illustrated is the starting cell (WT/MUT) with the ssODN that does not span over the SNP. If the silent mutations are present but the SNP is lost, it can be assumed that one allele was repaired using the ssODN whereas the second allele was repaired using the already repaired other allele.

We took advantage of this finding and evaluated according to the presence of this SNP how frequently the ssODN was used as repair template. Therefore, if HDR occurred, we were able to differentiate whether the ssODN was used for repair or the respective other allele. We found that in 85% (CRISPRMAX) and 92% (NEON) of the HDR events, the ssODN is used as template (Figure 43).



**Figure 43 HDR Analysis after NEON or CRISPRMAX transfection.** The likelihood of repair using the other allele (meaning either WT or MUT) was less than HDR using the provided ssODN. n=12 (NEON), n=60 (CRISPRMAX)

### 3.3. RNP vs. Plasmid

Upon initial editing using plasmid delivery, poor editing efficiency lead us to the adaptation of direct RNP delivery, where the Cas9 is delivered in its proteinaceous form and the gRNA as single modified version together with the ssODN. Recent studies could show that direct delivery of the Cas9 protein and gRNA reduce the risk of off-target effects and leads to immediate activity of the RNP complex, compared to plasmid delivery where the RNP complex becomes active after transcription and translation steps within the cell. The delay in RNP complex activity with plasmid delivery leads lower success of KI due to the decay of the ssODN in the meantime (Tang et al, in preparation). The editing results upon plasmid delivery were compared to RNP delivery and the success of the two approaches was analyzed within this project.

The cells were transfected with the same gRNA and ssODN using either plasmid delivery or RNP delivery. For plasmid delivery, the SCN8A-mutated iPS cells were transfected using Lipofectamine 3000 by Li Li. The poor editing results upon plasmid delivery lead us to the adaptation of the RNP delivery approach.

As can be seen in Table 6, when using plasmid delivery 270 clones were analyzed and only one was correctly modified. Using the optimized RNP delivery method, altogether 170 clones were analyzed and 14 correctly modified clones were selected.

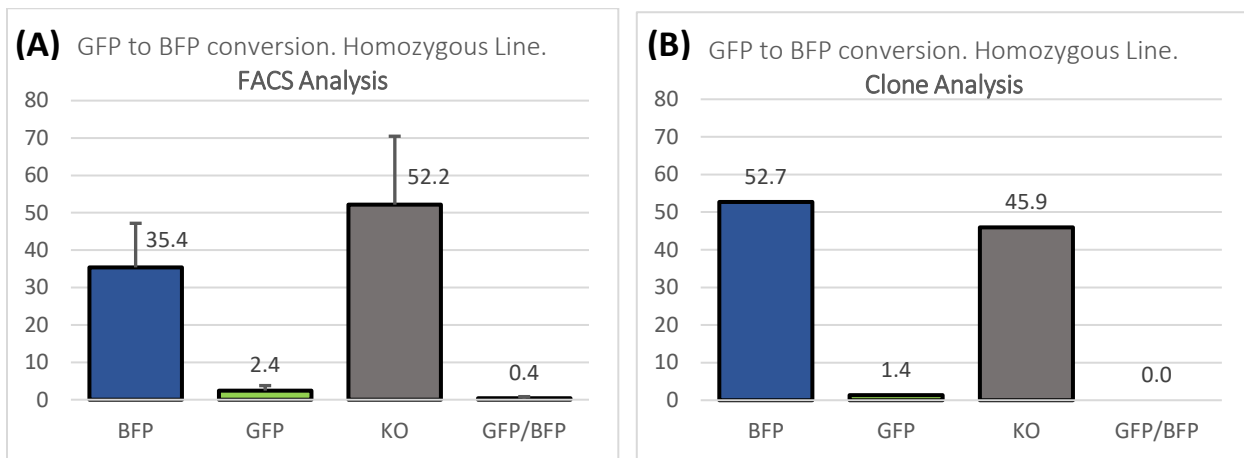
**Table 6 Plasmid delivery vs. RNP delivery.** Cas9 enzyme and gRNA were either delivered via a plasmid using Lipofectamine 3000 or in proteinaceous / modified sgRNA using either CRISPRMAX or NEON electroporation. The same ssODN was used for all methods.

Plasmid Delivery		RNP Delivery	
Analyzed clones	270	Analyzed clones	170
Correctly modified clones	1	Correctly modified clones	14

### 3.4. Evaluation of experiment variability

#### 3.4.1. Characterization of editing events after GFP to BFP conversion

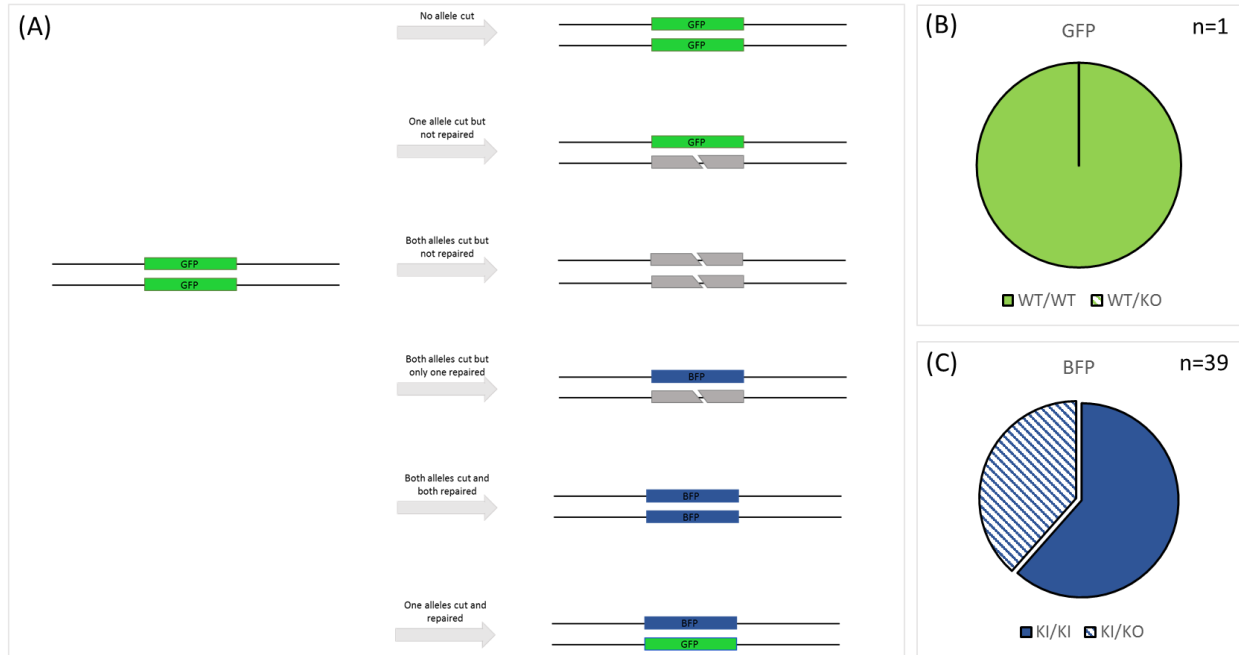
In order to further define the editing events, single clones were analyzed after NEON transfection in homozygous (GFP/GFP) cells (HUES8 GFP 002). The clones were first characterized using fluorescence microscopy and we counted the number of clones that were BFP-positive, GFP-positive, non-fluorescent and double-positive. Table 44 shows a comparison between FACS data and clone analysis.



**Figure 44 Analysis of editing events via FACS Analysis vs. Clone Analysis.** Analysis of a homozygous GFP/GFP line after transfection with NEON Electroporation using 1200V. (A) FACS Analysis. 10,000 cells were analyzed per experiment. 4 individual experiments were performed. n=4 (B) Clone Analysis via fluorescent microscopy and Sanger sequencing. Out of 96 clones, 74 were used for analysis. Mixed clones were excluded.

After characterizing the clones for BFP or GFP signals, some clones were sequenced in order to differentiate and confirm if one or two alleles were targeted. As shown in Figure 45, after transfection, several editing outcomes are possible: No editing, single cutting, double cutting, double cutting with single repair, double cutting with double repair and single cutting with repair.

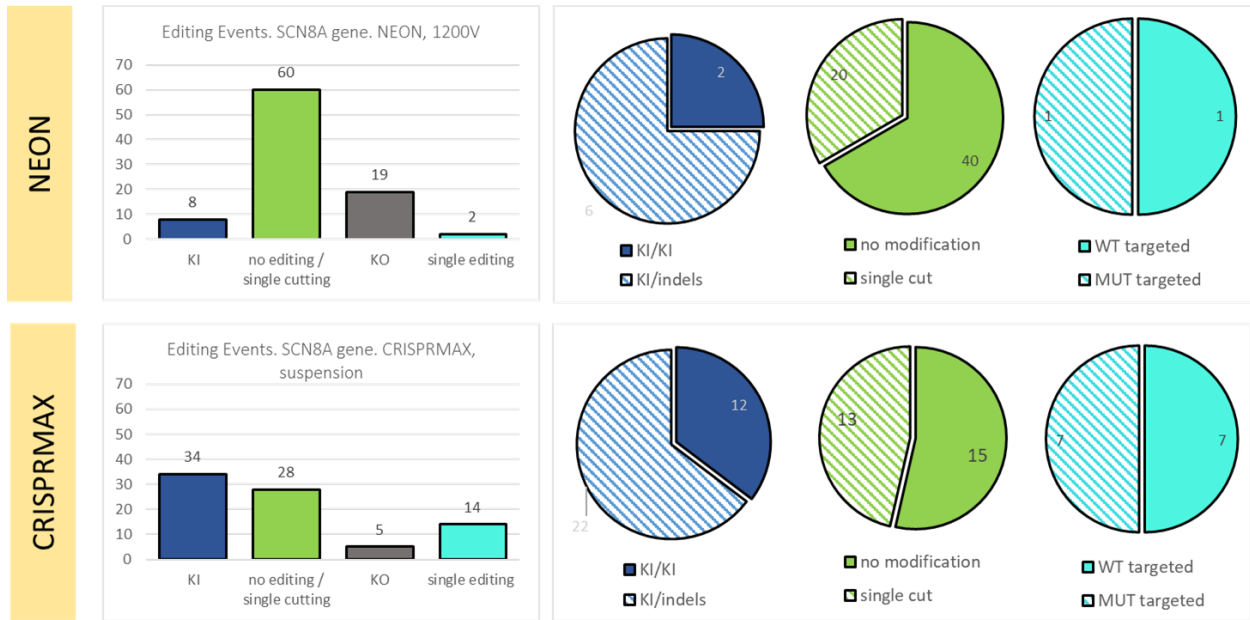
Ultimately, the BFP-positive clones were sub-grouped into BFP/BFP and BFP/KO, as well as the GFP-positive clones into GFP/GFP and GFP/KO, in order to differentiate the respective clones within the green or blue pool.



**Figure 45 Characterization of editing events after GFP to BFP conversion.** (A) Illustration of possible editing events starting from a homozygous GFP/GFP cell line. (B) Sub-characterization of GFP-positive clones. Untargeted clones remained homozygous GFP (WT/WT) whereas clones that had only one copy of GFP left were targeted but not repaired. (C) Sub-characterization of BFP-positive clones. All clones were targeted twice but only the double-positive ones were repaired twice (KI/KI). A single KI event results in a heterozygous BFP clone (KI/KO).

### 3.4.2. Comparison of editing events between experiments

Several factors influence the editing outcome. The variations between different set-ups were evaluated. The GFP to BFP conversion in ES cells (Set-up A) was compared to targeting iPS cells in the SCN8A locus (Set-up B). As shown in Figure 46, the editing events were compiled accordingly: KI/KI with KI/indel were pooled and compared to the BFP-positive pool, the cells that received no editing or single cutting were summed up to be equivalent to the GFP-positive ones, double cutting was equivalent to non-fluorescent and single edited cells equivalent to GFP/BFP double-positive (Figure 46). In Set-up B, a clear reduction in the overall editing was observed compared to Set-up A, as can be seen in the unedited/single cutting pool (green bar). A double KI event (KI/KI) was more likely to occur in the GFP to BFP conversion set-up. Single editing (cyan) occurred more often in set-up B than set-up A. If a single editing event occurred, there is a 50:50 chance that either the mutant or the wildtype allele was targeted.



**Figure 46 Comparison of Clone Analysis to GFP to BFP conversion.** The clones were pooled to correspond to the outcomes of GFP to BFP conversion. The green bar indicated unedited or single cut clones, which were found to be far higher than in the GFP to BFP conversion experiment. When comparing NEON vs. CRISPRMAX, the pool was bigger in the NEON pool, where unedited cells contributed more than single cutting (green pie chart). KI/KI or KI/indel events (blue) were far more likely to occur in the CRISPRMAX pool where a KO event (indel/indel, grey) occurred more often in the NEON pool. Double KI (KI/KI) was less likely to occur than single KI (KI/indel) in both pools (blue pie chart). If a single allele was targeted and repaired (cyan pie chart), both alleles were targeted with the same likelihood.

## 4. Discussion

### 4.1. Optimization of RNP delivery

When comparing KO and KI efficiencies upon CRISPRMAX Lipofection in adherent cells vs. cells in suspension, an increase in efficiency was detected when the cells were in suspension. Fully dissociated cells provide a larger surface area that can be in contact with the liposome-complexes which leads to an increase of endocytotic events. However, pluripotent stem cells do prefer to grow in clumps as full dissociation leads to increased cell death. Overall, even though the survival was lower after full dissociation, after several days of recovery, the increase in efficiency was significant enough to continue CRISPRMAX transfection in suspended cells.

NEON electroporation depends on several variables such as the voltage of the electric pulse, the period of the pulse (mS) and the number of pulses. We searched the literature for appropriate conditions for pluripotent stem cells and found the applied conditions to be the most promising: 1100V, 30mS, 1x; 1200V 30 mS, 1x and 1300V 30 mS, 1x. In order to evaluate the impact of each set-up, the cells were all derived and dissociated from the same plate and passaged and treated exactly the same way with the exception of the applied voltage. A clear reduction of survival proportional to the increase of voltage was observed. However, the KO and KI efficiency increases proportionally to the voltage. Therefore, since the 1300V yielded a very low number of surviving cells, we decided to continue with a compromise between survival and efficiency: 1200V 30ms and 1 pulse.

When comparing CRISPRMAX Lipofection with NEON electroporation, KO and KI yields were evaluated after transfecting heterozygous and homozygous GFP-positive cells. The experiments were repeated 4 times for each condition. We already observed rather high variations between each experiment. Several factors were maintained consistent between experiments to sustain reproducibility. Cell count, reagent concentrations as well as the steps during the dissociation procedure were kept consistent. Nonetheless, other factors can have an impact on the editing efficiency and were not kept rigorously

consistent. We noticed deviations in the confluency before the experiment, the passage number or even the time-point of medium change after lipofection. All these factors could have led to the observed deviations between experiments. Nonetheless, an overall trend of increased KO and KI yield after NEON transfection compared to CRISPRMAX was striking.

The lower efficiency of CRISPRMAX transfection compared to NEON electroporation was observed in both, the heterozygous and homozygous cell lines. Especially in the homozygous cell line, where double cutting is required to observe a full KO, NEON performed better. NEON transfection could not only target both alleles with a higher chance but could also yield a higher number of BFP-positive cells, representing a KI event.

When comparing the KO events after NEON electroporation, the efficiency was very similar suggesting that if the Cas9 enzyme is present at a certain concentration, the DSB occurs very efficiently. This was shown in the heterozygous line once and in the homozygous line twice. Since such a pattern was not noticed after CRISPRMAX transfection, we assume that less editing factors are delivered using CRISPRMAX and therefore the editing events are less likely to occur.

Overall, under the given conditions, we found the NEON electroporation system to lead to a higher number of editing events and the yield of KI events was far higher compared to Lipofection with CRISPRMAX.

#### **4.2. Delivery of two ssODN increases the yield for WT/KI**

Delivery of two ssODN, instead of one, did in fact increase the amount of double-positive cells. This approach is only relevant if the KI efficiency is already high. A high number of HDR events is required for such an approach. Delivery of two oligos requires the targeting of both alleles and their repair using the delivered ssODNs. The two oligos delivered, one

encoding BFP and the other GFP, were used as repair template by chance. Correspondingly, there is a 50:50 chance that one of the two oligos is used as repair template. Overall, delivery of two ssODNs is only feasible when the editing efficiency is above a certain threshold.

If the editing efficiency is low, the ratio of single-allele edited cells is increased, whereas the ratio of two-alleles edited cells is decreased. With lower editing efficiency, it is more likely that just one allele will be targeted instead of two. Less efficient access for Cas9 binding (location on the chromosome) or the binding affinity of the sgRNA are factors that can diminish the cutting efficiency. Moreover, HDR events are required for a KI event. As soon as the probability of HDR is reduced, the likelihood for a KI is decreased proportionally. For example, the speed of cell recovery after transfection could affect the HDR likelihood, since cells need to be in a proliferative state for HDR event to occur.

### **4.3. Repair of a heterozygous mutation**

For the repair of a heterozygous mutated cell, a sgRNA that only targets the mutated (MUT) allele would be ideal. However, since no gRNA was found in close proximity to the site of mutation, all sgRNAs that were found targeted both the WT and MUT allele. When checking the efficiency of the selected sgRNAs, the efficiency in HEK 293T cells was found to be 57% using proteinaceous RNP delivery. It is known that HEK 293T cells are easy to transfect, especially when compared to iPS or ES cells. Therefore, it can be assumed that the efficiency even decreases when applied to the iPS cell line. Analysis of the transfected iPS cells showed that the overall efficiency upon proteinaceous RNP transfection was 33% when transfected with NEON electroporation and 23% using CRISPRMAX.

However, visual analysis of the respective pools indicated that the actual KI proportion seemed to be higher in the CRISPRMAX pool.

When examining each clone individually, this finding was confirmed: the CRISPRMAX pool indeed led to a higher amount of correctly repaired clones. The prediction bias of



ICE analysis can be explained by the software's underlying mechanism of prediction. ICE analysis compares the sequence of interest with the sequence of the original cell. It aligns the whole sequence of the PCR amplicons of the sample sequence with the control sequence and interprets only very different sequences as modified. If just a few bases, such as the 3 bases in this case, are modified, the sequence is still interpreted as similar to the original and not counted as edited. Therefore, the tool predicts the editing efficiency to be lower as it does not interpret KI events as edited.

When looking at each clone individually, the percentage of indels was far higher within the NEON transfected cells than with CRISPRMAX, leading to the assumption that HDR was more likely to occur after CRISPRMAX transfection.

When compiling all editing events, the probability of HDR in the NEON pool was only 16% compared to 56% in the CRISPRMAX pool. As mentioned, for HDR to happen, a proliferative state of the cell is required. Since each cell line responds differently to certain treatments and modifications, it can be assumed that editing using NEON was more harmful to the cells leading to a decrease of proliferation and ultimately a decrease of HDR events. CRISPRMAX was less toxic to the cells and resulted in a higher yield of KI events.

Moreover, when analyzing the HDR events in detail, according to the missing SNP, it can be deduced that either the ssODN (SNP still present) or the respective other allele (SNP gone) was used as a repair template. Regardless of the transfection procedure, it was more likely that the ssODN was used as repair template over the other allele.

#### **4.4. RNP vs. Plasmid**

As already discussed, direct RNP delivery performed better in HEK 293T cells when testing for guide efficiency. Plasmid delivery was performed by Li Li using the same guide sequence, but instead of delivering it as a modified single guide RNA, the RNA sequence was delivered via a plasmid. The Cas9 protein is immediately present in the RNP delivery

set-up, whereas plasmid delivery requires transcription and translation before the protein is expressed and active. The same ssODN was used for both set-ups. When the RNP complex is delivered in its active conformation, it is immediately functional and the ssODN is present in a high enough concentration to be used for HDR. If delivered via plasmid, the editing factors need to be expressed first and it takes about 48 hours until an active RNP complex is formed. In the meantime, some ssODNs are already degraded, which decreases the KI yield. Apart from that, the editing factors are maintained in the cells for a longer period of time, which can have a negative impact in regard to off-target effects. In contrast, direct RNP delivery assures that the editing factors are present immediately and are degraded after 48h which significantly reduces the chance of off-target events. In summary, it can be concluded that direct RNP delivery is more reliable and efficient than plasmid delivery.

#### **4.5. Evaluation of experiment variability**

Clone analysis of the GFP to BFP conversion experiment correlated with the FACS results. However, a higher yield of BFP-positive cells was observed after clone analysis. The BFP fluorophore is relatively dim. Therefore, the settings for the gates during FACS Analysis could have been set too high resulting in an underestimation of BFP-positive cells.

Nevertheless, clone analysis could confirm the overall editing distribution and a high yield of BFP-positive cells representing a successful KI event. Using microscopy and sequencing, the green cells and blue cells were analyzed. Within the analyzed clones, only one clone (n=1) remained green and was totally unedited (GFP/GFP). The overall editing was very efficient where more than half of the clones were repaired after a DSB. If repaired, the likelihood of a repair in both alleles was slightly more than that of a single repair (KI/KI > KI/KO).

The editing events after CRISPRMAX and NEON transfection in the SCN8A cells, were compiled into the same groups as in the GFP to BFP conversion experiment. After

compilation, a clear reduction of editing events was noticed. The events of single cutting were more likely and especially CRISPRMAX transfection increased the yield of single edited cells.

In conclusion, NEON transfection is a very efficient method to deliver editing factors for genome editing using CRISPR/Cas9. However, the method is relatively harsh and leads to variations between different cell lines. In this case, the HUES8 cell line seems to be more robust than the SCN8A-mutated iPS cell line. It cannot be excluded that the passage number had an impact on this finding. HUES8 cells were transfected at passage 52 whereas the iPS cell line was edited at passage 16. Even though ES and iPS cells are very similar, it cannot be ruled out that the cell type has an impact. HUES8 cells are ES cells whereas the patient-derived cells are iPS.

Regardless, the delivery of editing factors has to be adapted for individual cell lines. In the case of the SCN8A-mutated iPS cell line, CRISPRMAX transfection was less toxic to the cells and resulted in a higher survival. Since a proliferative state is required for a successful HDR event, the higher yield of KI events may be a result of their faster recovery.

## **5. Conclusion**

Genome editing is a very powerful tool and has proved itself especially useful in the field of regenerative medicine. It requires precision and detailed planning. Several editing tools are available, where the CRISPR/Cas9 system is the most convenient in terms of experiment design. Before editing a certain genome locus, the DNA sequence of that locus needs to be examined and the guide RNA has to be designed properly. The predicted off-targets should be kept low and testing the cutting efficiency of the selected gRNA is recommended. The same holds true of the design of ssODN. It has to be assured that the genetic modification upon a KI event will not result in missense mutations, and introduction of silent mutations should prevent re-cutting.

Upon successful design, the delivery of the editing factors is essential. The GFP to BFP conversion is an elegant platform to study the effects of different delivery methods. Evaluation of the editing events can be performed using microscopy and FACS Analysis, and does not require the protracted procedure of clone analysis and sequencing. Overall, electroporation was more efficient in delivering the editing factors. Regardless of the delivery method, a clear increase in efficiency was observed upon direct RNP delivery over plasmid delivery.

Correspondingly, when examining the variability between experiments, direct RNP delivery certainly performed better than plasmid delivery. Regarding delivery, the method of choice has to be optimized for the respective cell type. NEON electroporation proved itself as a powerful tool, but survival and recovery after transfection require further optimization.

### **5.1. Further perspectives**

Continuous improvement of genome editing is essential to increase its efficiency and reduce workload. Based on our findings, the iPS Core will switch from plasmid delivery to RNP delivery and will introduce NEON electroporation as one of the delivery methods. The settings for electroporation might vary between cell lines and need to be optimized for each individual line. The experimental set-up and the individual steps are described and a protocol was established. Direct RNP delivery will speed up the genome editing service in the iPS Core facility. Cell lines can efficiently be repaired, or a mutation of interest can be introduced. These genetic modifications are valuable to study gene function and derive isogenic controls. Such isogenic controls are essential in drug screening experiments and for disease modeling.

## VII. Table of Figures

Figure 1 iPSCs. Adapted from [2].	1
Figure 2 Induced pluripotent stem cells (iPSCs). Adapted from [2].	4
Figure 3 Overview of most commonly used reprogramming approaches.	5
Figure 4 The life-cycle of the Sendai Virus. Adapted from [33].	9
Figure 5 Sendai Virus RNA vector. Adopted from [37].	10
Figure 6 Episomal vectors. Adapted from [39].	11
Figure 7 mRNA transfection. Adapted from [47] and Allele Biotech, 2018.	12
Figure 8 Genome edited iPSCs. Adapted from [2], [48]	14
Figure 9 Nucleases. Adapted from [51].	15
Figure 10 Meganuclease (MN). Adapted from [59]	17
Figure 11 Zinc-Finger Nuclease (ZFN). Adapted from [59]	18
Figure 12 TALEN. Adapted from [66] [67][59]	19
Figure 13 The Cas9 endonuclease. Adapted from [76] [77]	21
Figure 14 Different Cas9 orthologues have different PAM requirements. [66].	22
Figure 15 Cas9 nickase and double nicking. Adapted from [76].	23
Figure 16 dCas9 effector fusion. Adapted from [85] [76].	24
Figure 17 Major editing factors for a Knock-In experiment. Adapted from [94].	26
Figure 18 Transfection methods. <i><a href="https://blog.benchling.com/2016/03/24/how-to-express-crispr-in-your-target-cells/">https://blog.benchling.com/2016/03/24/how-to-express-crispr-in-your-target-cells/</a></i> .	29
Figure 19 Transfection using CRISPRMAX Reagent.	29
Figure 20 Transfection using NEON electroporation System.	30
Figure 21 Passaging.	31
Figure 22 Tracking of Indels by Decomposition (TIDE) Analysis.	33
Figure 23 Inference of CRISPR Editing (ICE) Analysis.	34
Figure 24 GFP to BFP conversion.	35
Figure 25 Image processing with ImageJ software.	36
Figure 26 FACS Analysis.	37
Figure 27 Transfection efficiency testing	38
Figure 28 Clone Analysis. Determination of indel contribution.	39
Figure 29 Survival post transfection of cells that were adherent vs. in suspension.	40

Figure 30 Lipofection with CRISPRMAX in homozygous GFP cells.....	41
Figure 31 Survival post transfection of cells that were transfected using NEON electroporation system .....	42
Figure 32 NEON Electroporation in heterozygous GFP cells. ....	43
Figure 33 Comparison of CRISPRMAX Lipofection vs. NEON Electroporation..	45
Figure 34 Comparison delivering proteinaceous RNP with either one ssODN or two ssODNs.....	47
Figure 35 sgRNA design for the SCN8A gene. ....	48
Figure 36 sgRNA testing using ICE Analysis (Synthego). ....	49
Figure 37 ssODN design. ....	50
Figure 38 ICE Analysis of transfected cell pools. ....	51
Figure 39 Visual analysis of transfected cell pools.....	51
Figure 40 Clone Analysis of the modified genetic locus within the SCN8A gene.	52
Figure 41 Repair Analysis after NEON or CRISPRMAX transfection.....	53
Figure 42 Non-ssODN-mediated HDR. ....	54
Figure 43 HDR Analysis after NEON or CRISPRMAX transfection. ....	54
Figure 44 Analysis of editing events via FACS Analysis vs. Clone Analysis. ....	56
Figure 45 Characterization of editing events after GFP to BFP conversion. ....	57
Figure 46 Comparison of Clone Analysis to GFP to BFP conversion.....	58

## VIII. References

- [1] M. Spivakov and A. G. Fisher, "Epigenetic signatures of stem-cell identity," *Nature Reviews Genetics*, vol. 8, no. 4. pp. 263–271, 2007.
- [2] S. Menon, S. Shailendra, A. Renda, M. Longaker, and N. Quarto, "An overview of direct somatic reprogramming: The ins and outs of iPSCs," *International Journal of Molecular Sciences*, vol. 17, no. 1. 2016.
- [3] M. Herreros-Villanueva, L. Bujanda, D. D. Billadeau, and J.-S. Zhang, "Embryonic stem cell factors and pancreatic cancer.," *World J. Gastroenterol.*, vol. 20, no. 9, pp. 2247–54, 2014.
- [4] G. R. Martin, "Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.," *Proc. Natl. Acad. Sci.*, vol. 78, no. 12, pp. 7634–7638, 1981.
- [5] R. A. Young, "Control of the embryonic stem cell state," *Cell*, vol. 144, no. 6. pp. 940–954, 2011.
- [6] C. A. Cowan, J. Atienza, D. A. Melton, and K. Eggan, "Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells," *Science (80-. )*, vol. 309, no. 5739, pp. 1369–1373, 2005.
- [7] I. Wilmut, A. E. Schnieke, J. McWhir, A. J. Kind, and K. H. S. Campbell, "Viable offspring derived from fetal and adult mammalian cells," *Nature*, vol. 385, no. 6619, pp. 810–813, 1997.
- [8] K. Hochedlinger and R. Jaenisch, "Nuclear reprogramming and pluripotency," *Nature*, vol. 441, no. 7097. pp. 1061–1067, 2006.
- [9] K. Takahashi and S. Yamanaka, "Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors," *Cell*, vol. 126, no. 4, pp. 663–676, 2006.
- [10] X.-L. Guo and J.-S. Chen, "Research on induced pluripotent stem cells and the application in ocular tissues," *Int J Ophthalmol*, vol. 8, no. 4, pp. 818–825, 2015.
- [11] K. Okita, T. Ichisaka, and S. Yamanaka, "Generation of germline-competent induced pluripotent stem cells," *Nature*, vol. 448, no. 7151, pp. 313–317, 2007.
- [12] M. Wernig *et al.*, "In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state," *Nature*, vol. 448, no. 7151, pp. 318–324, 2007.
- [13] K. Takahashi *et al.*, "Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors," *Cell*, vol. 107, no. 5, pp. 861–872, 2007.

- [14] J. Yu *et al.*, “Induced pluripotent stem cell lines derived from human somatic cells,” *Science*, vol. 318, no. 5858, pp. 1917–20, 2007.
- [15] N. Malik and M. S. Rao, “A review of the methods for human iPSC derivation,” in *Methods in Molecular Biology*, 2013, vol. 997, pp. 23–33.
- [16] N. Rawat and M. K. Singh, “Induced pluripotent stem cell: A headway in reprogramming with promising approach in regenerative biology,” *Veterinary World*, vol. 10, no. 6, pp. 640–649, 2017.
- [17] T. M. Schlaeger *et al.*, “A comparison of non-integrating reprogramming methods,” *Nat. Biotechnol.*, vol. 33, no. 1, pp. 58–63, 2015.
- [18] N. Takenaka-Ninagawa, Y. Kawabata, S. Watanabe, K. Nagata, and S. Torihashi, “Generation of rat-induced pluripotent stem cells from a new model of metabolic syndrome,” *PLoS One*, vol. 9, no. 8, 2014.
- [19] M. Imamura *et al.*, “Derivation of induced pluripotent stem cells by retroviral gene transduction in mammalian species,” *Methods Mol Biol*, vol. 925, pp. 21–48, 2012.
- [20] C. W. Chang *et al.*, “Polycistronic lentiviral vector for ‘hit and run’ reprogramming of adult skin fibroblasts to induced pluripotent stem cells,” *Stem Cells*, vol. 27, no. 5, pp. 1042–1049, 2009.
- [21] O. Meier and U. F. Greber, “Adenovirus endocytosis,” *Journal of Gene Medicine*, vol. 5, no. 6, pp. 451–462, 2003.
- [22] J. Roy-Chowdhury and M. S. Horwitz, “Evolution of adenoviruses as gene therapy vectors,” *Mol. Ther.*, vol. 5, no. 4, pp. 340–344, 2002.
- [23] L. Zhong, A. Granelli-Piperno, Y. Choi, and R. M. Steinman, “Recombinant adenovirus is an efficient and non-perturbing genetic vector for human dendritic cells,” *Eur. J. Immunol.*, vol. 29, no. 3, pp. 964–972, 1999.
- [24] L. E. Woodard and M. H. Wilson, “PiggyBac-ing models and new therapeutic strategies,” *Trends in Biotechnology*, vol. 33, no. 9, 2015.
- [25] D. Kim *et al.*, “Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins,” *Cell Stem Cell*, vol. 4, no. 6, pp. 472–476, 2009.
- [26] G. Guidotti, L. Brambilla, and D. Rossi, “Cell-Penetrating Peptides: From Basic Research to Clinics,” *Trends in Pharmacological Sciences*, vol. 38, no. 4, pp. 406–424, 2017.
- [27] F. Jia *et al.*, “A nonviral minicircle vector for deriving human iPS cells,” *Nat.*



*Methods*, vol. 7, no. 3, pp. 197–199, 2010.

- [28] T. B. P. V. L. Sandmaier S.E.S., *MicroRNA-Mediated Reprogramming of Somatic Cells into Induced Pluripotent Stem Cells.*, Vol 1330. Verma P., Sumer H. (eds) Cell Reprogramming. Methods in Molecular Biology, Humana Press, New York, NY, 2015.
- [29] J. Yu, K. F. Chau, M. A. Vodyanik, J. Jiang, and Y. Jiang, “Efficient feeder-free episomal reprogramming with small molecules,” *PLoS One*, vol. 6, no. 3, 2011.
- [30] Y. Fujie. *et al.*, “New type of sendai virus vector provides transgene-free iPS cells derived from chimpanzee blood,” *PLoS One*, vol. 9, no. 12, 2014.
- [31] Y. Kishino, T. Seki, S. Yuasa, J. Fujita, and K. Fukuda, “Generation of Induced Pluripotent Stem Cells from Human Peripheral T Cells Using Sendai Virus in Feeder-free Conditions.,” *J. Vis. Exp.*, no. 105, pp. 1–5, 2015.
- [32] J. L. Hurwitz *et al.*, “Intranasal Sendai virus vaccine protects African green monkeys from infection with human parainfluenza virus-type one.,” *Vaccine*, vol. 15, no. 5, pp. 533–540, 1997.
- [33] Invitrogen and LifeTechnologies, “CytoTune™ -iPS Reprogramming Kit,” *User Guid.*, 2014.
- [34] U. Griesenbach, M. Inoue, M. Hasegawa, and E. W. F. W. Alton, “Sendai virus for gene therapy and vaccination.,” *Curr. Opin. Mol. Ther.*, vol. 7, no. 4, pp. 346–52, 2005.
- [35] M. Bitzer, S. Armeanu, U. M. Lauer, and W. J. Neubert, “Sendai virus vectors as an emerging negative-strand RNA viral vector system,” *Journal of Gene Medicine*, vol. 5, no. 7. pp. 543–553, 2003.
- [36] S. Diecke *et al.*, “Recent technological updates and clinical applications of induced pluripotent stem cells,” *Korean Journal of Internal Medicine*, vol. 29, no. 5. pp. 547–557, 2014.
- [37] H.-O. Li *et al.*, “A Cytoplasmic RNA Vector Derived from Nontransmissible Sendai Virus with Efficient Gene Transfer and Expression,” *J. Virol.*, vol. 74, no. 14, pp. 6564–6569, 2000.
- [38] B. Thyagarajan *et al.*, “A single EBV-based vector for stable episomal maintenance and expression of GFP in human embryonic stem cells.,” *Regen. Med.*, vol. 4, no. 2, pp. 239–50, 2009.

- [39] K. Okita *et al.*, “A more efficient method to generate integration-free human iPS cells,” *Nat. Methods*, vol. 8, no. 5, pp. 409–12, May 2011.
- [40] K. Okita *et al.*, “An efficient nonviral method to generate integration-free human-induced pluripotent stem cells from cord blood and peripheral blood cells,” *Stem Cells*, vol. 31, no. 3, pp. 458–466, 2013.
- [41] A. Ehrhardt, R. Haase, A. Schepers, M. J. Deutsch, H. J. Lipps, and A. Baiker, “Episomal vectors for gene therapy,” *Curr. Gene Ther.*, vol. 8, no. 3, pp. 147–61, 2008.
- [42] A. A. Mack, S. Kroboth, D. Rajesh, and W. B. Wang, “Generation of induced pluripotent stem cells from CD34+ cells across blood drawn from multiple donors with non-integrating episomal vectors,” *PLoS One*, vol. 6, no. 11, 2011.
- [43] H. Hong *et al.*, “Suppression of induced pluripotent stem cell generation by the p53-p21 pathway,” *Nature*, vol. 460, no. 7259, pp. 1132–1135, 2009.
- [44] M. Nakagawa, N. Takizawa, M. Narita, T. Ichisaka, and S. Yamanaka, “Promotion of direct reprogramming by transformation-deficient Myc,” *Proc. Natl. Acad. Sci.*, vol. 107, no. 32, pp. 14152–14157, 2010.
- [45] J. Liu and P. J. Verma, “Synthetic mRNA reprogramming of human fibroblast cells,” in *Methods in Molecular Biology*, vol. 1330, 2015, pp. 17–28.
- [46] L. Warren *et al.*, “Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA,” *Cell Stem Cell*, vol. 7, no. 5, pp. 618–630, 2010.
- [47] L. Warren, Y. Ni, J. Wang, and X. Guo, “Feeder-free derivation of human induced pluripotent stem cells with messenger RNA,” *Sci Rep*, vol. 2, no. 657, pp. 1–7, 2012.
- [48] D. Hockemeyer and R. Jaenisch, “Induced pluripotent stem cells meet genome editing,” *Cell Stem Cell*, vol. 18, no. 5, pp. 573–586, 2016.
- [49] M. Fennell *et al.*, “Impact of RNA-guided technologies for target identification and deconvolution,” *Journal of Biomolecular Screening*, vol. 19, no. 10, pp. 1327–1337, 2014.
- [50] K. M. Esvelt and H. H. Wang, “Genome-scale engineering for systems and synthetic biology,” *Molecular Systems Biology*, vol. 9, 2013.

- [51] M. Yanik *et al.*, “In vivo genome editing as a potential treatment strategy for inherited retinal dystrophies,” *Progress in Retinal and Eye Research*, vol. 56, pp. 1–18, 2017.
- [52] A. Choulika, A. Perrin, B. Dujon, and J. Ois Nicolas, “Induction of Homologous Recombination in Mammalian Chromosomes by Using the I-SceI System of *Saccharomyces cerevisiae*,” *Mol. Cell. Biol.*, vol. 15, no. 4, pp. 1968–1973, 1995.
- [53] C. D. Richardson, G. J. Ray, M. A. DeWitt, G. L. Curie, and J. E. Corn, “Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA,” *Nat. Biotechnol.*, vol. 34, no. 3, pp. 339–344, 2016.
- [54] A. Hendel *et al.*, “Quantifying genome-editing outcomes at endogenous loci with SMRT sequencing,” *Cell Rep.*, vol. 7, no. 1, pp. 293–305, 2014.
- [55] F. A. Ran, P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, and F. Zhang, “Genome engineering using the CRISPR-Cas9 system,” *Nat. Protoc.*, vol. 8, no. 11, pp. 2281–2308, 2013.
- [56] M. R. Lieber, “The mechanism of human nonhomologous DNA end joining.,” *J. Biol. Chem.*, vol. 283, no. 1, pp. 1–5, 2008.
- [57] S. K. Deng, B. Gibb, M. J. De Almeida, E. C. Greene, and L. S. Symington, “RPA antagonizes microhomology-mediated repair of DNA double-strand breaks,” *Nat. Struct. Mol. Biol.*, vol. 21, no. 4, pp. 405–412, 2014.
- [58] B. L. Stoddard, “Homing endonuclease structure and function,” *Quarterly Reviews of Biophysics*, vol. 38, no. 1, pp. 49–95, 2005.
- [59] G. Romay and C. Bragard, “Antiviral defenses in plants through genome editing,” *Frontiers in Microbiology*, vol. 8, no. JAN, 2017.
- [60] D. B. T. Cox, R. J. Platt, and F. Zhang, “Therapeutic genome editing: Prospects and challenges,” *Nature Medicine*, vol. 21, no. 2, pp. 121–131, 2015.
- [61] B. Petersen and H. Niemann, “Molecular scissors and their application in genetically modified farm animals,” *Transgenic Research*, vol. 24, no. 3, pp. 381–396, 2015.
- [62] T. Gaj, A. C. Mercer, S. J. Sirk, H. L. Smith, and C. F. Barbas, “A comprehensive approach to zinc-finger recombinase customization enables genomic targeting in human cells,” *Nucleic Acids Res.*, vol. 41, no. 6, pp. 3937–3946, 2013.

- [63] R. D. Knight and S. M. Shimeld, "Identification of conserved C2H2 zinc-finger gene families in the Bilateria.," *Genome Biol.*, vol. 2, no. 5, p. RESEARCH0016, 2001.
- [64] J. C. Miller *et al.*, "An improved zinc-finger nuclease architecture for highly specific genome editing," *Nat. Biotechnol.*, vol. 25, no. 7, pp. 778–785, 2007.
- [65] H. Scholze and J. Boch, "TAL effectors are remote controls for gene activation," *Current Opinion in Microbiology*, vol. 14, no. 1. pp. 47–53, 2011.
- [66] W. T. Hendriks, C. R. Warren, and C. A. Cowan, "Genome Editing in Human Pluripotent Stem Cells: Approaches, Pitfalls, and Solutions," *Cell Stem Cell*, vol. 18, no. 1. pp. 53–65, 2016.
- [67] N. J. Baltes and D. F. Voytas, "Enabling plant synthetic biology through genome engineering," *Trends in Biotechnology*, vol. 33, no. 2. pp. 120–131, 2015.
- [68] J. Boch *et al.*, "Breaking the code of DNA binding specificity of TAL-type III effectors," *Science (80-. )*, vol. 326, no. 5959, pp. 1509–1512, 2009.
- [69] J. C. Miller *et al.*, "A TALE nuclease architecture for efficient genome editing," *Nat. Biotechnol.*, vol. 29, no. 2, pp. 143–150, 2011.
- [70] B. M. Lamb, A. C. Mercer, and C. F. Barbas, "Directed evolution of the TALE N-terminal domain for recognition of all 5' bases," *Nucleic Acids Res.*, vol. 41, no. 21, pp. 9779–9785, 2013.
- [71] S. Bultmann *et al.*, "Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers," *Nucleic Acids Res.*, vol. 40, no. 12, pp. 5368–5377, 2012.
- [72] Y. Kim *et al.*, "A library of TAL effector nucleases spanning the human genome," *Nat. Biotechnol.*, vol. 31, no. 3, pp. 251–258, 2013.
- [73] S. S. Palii, B. O. Van Emburgh, U. T. Sankpal, K. D. Brown, and K. D. Robertson, "DNA Methylation Inhibitor 5-Aza-2'-Deoxycytidine Induces Reversible Genome-Wide DNA Damage That Is Distinctly Influenced by DNA Methyltransferases 1 and 3B," *Mol. Cell. Biol.*, vol. 28, no. 2, pp. 752–771, 2008.
- [74] J. Valton *et al.*, "Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation," *J. Biol. Chem.*, vol. 287, no. 46, pp. 38427–38432, 2012.

- [75] J. P. Guilinger *et al.*, “Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity,” *Nat. Methods*, vol. 11, no. 4, pp. 429–435, 2014.
- [76] J. A. Doudna and E. Charpentier, “The new frontier of genome engineering with CRISPR-Cas9,” *Science*, vol. 346, no. 6213. 2014.
- [77] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier, “A Programmable Dual-RNA – Guided,” *Science (80-. )*, vol. 337, pp. 816–822, 2012.
- [78] H. Puchta and F. Fauser, “Synthetic nucleases for genome engineering in plants: Prospects for a bright future,” *Plant Journal*, vol. 78, no. 5. pp. 727–741, 2014.
- [79] P. Horvath and R. Barrangou, “CRISPR/Cas, the immune system of Bacteria and Archaea,” *Science*, vol. 327, no. 5962. pp. 167–170, 2010.
- [80] L. Cong *et al.*, “Multiplex Genome Engineering Using CRISPR/Cas Systems,” *Science (80-. )*, vol. 339, no. 6121, pp. 819–823, 2013.
- [81] Z. Hou *et al.*, “Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*,” *Proc. Natl. Acad. Sci. USA*, vol. 110, no. 39, pp. 15644–15649, 2013.
- [82] F. A. Ran *et al.*, “Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity,” *Cell*, vol. 154, no. 6, pp. 1380–1389, 2013.
- [83] A. Chavez *et al.*, “Comparison of Cas9 activators in multiple species,” *Nat. Methods*, vol. 13, no. 7, pp. 563–567, 2016.
- [84] M. E. Tanenbaum, L. A. Gilbert, L. S. Qi, J. S. Weissman, and R. D. Vale, “A protein-tagging system for signal amplification in gene expression and fluorescence imaging,” *Cell*, vol. 159, no. 3, pp. 635–646, 2014.
- [85] L. a Gilbert *et al.*, “CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes,” *Cell*, vol. 154, no. 2, pp. 442–451, 2013.
- [86] P. D. Hsu *et al.*, “DNA targeting specificity of RNA-guided Cas9 nucleases,” *Nat. Biotechnol.*, vol. 31, no. 9, pp. 827–832, 2013.
- [87] L. Yang *et al.*, “Optimization of scarless human stem cell genome editing,” *Nucleic Acids Res.*, vol. 41, no. 19, pp. 9049–9061, 2013.
- [88] S. Q. Tsai *et al.*, “Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing,” *Nat. Biotechnol.*, vol. 32, no. 6, pp. 569–576, 2014.

- [89] K. Jacoby, A. R. Lambert, and A. M. Scharenberg, "Characterization of homing endonuclease binding and cleavage specificities using yeast surface display SELEX (YSD-SELEX)," *Nucleic Acids Res.*, vol. 45, no. 3, p. e11, 2017.
- [90] C. N. Kanchiswamy, M. Maffei, M. Malnoy, R. Velasco, and J. S. Kim, "Fine-Tuning Next-Generation Genome Editing Tools," *Trends in Biotechnology*, vol. 34, no. 7, pp. 562–574, 2016.
- [91] V. Pattanayak, C. L. Ramirez, J. K. Joung, and D. R. Liu, "Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection," *Nat. Methods*, vol. 8, no. 9, pp. 765–772, 2011.
- [92] K. A. Schaefer, W. H. Wu, D. F. Colgan, S. H. Tsang, A. G. Bassuk, and V. B. Mahajan, "Unexpected mutations after CRISPR-Cas9 editing in vivo," *Nature Methods*, vol. 14, no. 6, pp. 547–548, 2017.
- [93] B. C. Cox, Z. Liu, M. M. M. Lagarde, and J. Zuo, "Conditional gene expression in the mouse inner ear using cre-loxP," *JARO - J. Assoc. Res. Otolaryngol.*, vol. 13, pp. 295–322, 2012.
- [94] B. Roberts *et al.*, "Systematic gene tagging using CRISPR/Cas9 in human stem cells to illuminate cell organization," *Mol. Biol. Cell*, vol. 28, no. 21, pp. 2854–2874, 2017.
- [95] J. Drost and H. Clevers, "Organoids in cancer research," *Nature Reviews Cancer*, pp. 1–12, 2018.
- [96] S. Waizenauer, "GENOME EDITING BY USING A CONVERSION OF GFP TO BFP IN HUMAN PLURIPOTENT STEM CELL LINES," 2018.
- [97] E. K. Brinkman, T. Chen, M. Amendola, and B. Van Steensel, "Easy quantitative assessment of genome editing by sequence trace decomposition," *Nucleic Acids Res.*, vol. 42, no. 22, 2014.
- [98] A. Glaser, B. McColl, and J. Vadolas, "Corrigendum to GFP to BFP Conversion: A Versatile Assay for the Quantification of CRISPR/Cas9-mediated Genome Editing (Molecular Therapy - Nucleic Acids (2016) 5 (e334) (S2162253117300616) (10.1038/mtna.2016.48))," *Molecular Therapy - Nucleic Acids*, vol. 5, pp. 1–2, 2016.
- [99] H. L. Li, P. Gee, K. Ishida, and A. Hotta, "Efficient genomic correction methods in human iPS cells using CRISPR-Cas9 system," *Methods*, vol. 101, pp. 27–35, 2016.

- [100] K. R. Veeramah *et al.*, “De novo pathogenic SCN8A mutation identified by whole-genome sequencing of a family quartet affected by infantile epileptic encephalopathy and SUDEP,” *Am. J. Hum. Genet.*, vol. 90, no. 3, pp. 502–510, 2012.
- [101] P. Lorenz, D. Koczan, and H. J. Thiesen, “Transcriptional repression mediated by the KRAB domain of the human C2H2 zinc finger protein Kox1/ZNF10 does not require histone deacetylation,” *Biol. Chem.*, vol. 382, no. 4, pp. 637–644, 2001.