VILNIUS UNIVERSITY
LIFE SCIENCES CENTRE

UGNĖ ŠIMULIŪNAITĖ

Establishment of Methods for HDR-Based Polyengineering of Human Cells

Master Thesis

Genetics Study Programme

Academic Supervisor
PhD, Jonathan Lee Arias Fuenzalida

Vilnius 2023

TABLE OF CONTENTS

VILNIAUS UNIVERSITETAS

GYVYBĖS MOKSLŲ CENTRAS

Ugnė Šimuliūnaitė

Homologine rekombinacija paremtų metodų kūrimas žmogaus ląstelių poliinžinerijai

Magistro baigiamasis darbas

SANTRAUKA

Ląstelių linijų genomų redagavimas yra kritiškai svarbus metodas vystymosi biologijos, ligų modeliavimo ir klinikiniuose tyrimuose. Tačiau paprastai vienoje ląstelių linijoje modifikuojamas vienas ar keli genai, o tai riboja genų sąveikos tyrimų galimybes. Daugelio lokusų redagavimo strategija toje pačioje ląstelių linijoje galėtų būti naudinga tiriant genų sąveikos įtaką išsamiam sveiko ir sutrikusio vystymosi tyrimui.

Šiame darbe pagrindinis dėmesys skiriamas genų išveiklinimo poliinžinerijos strategijos kūrimui homologine rekombnacija paremto genomo redagavimo būdu. Tyrimo tikslas buvo sukurti plazmidžių konstrukcijas genų *SOX1, PAX6, EOMES, GSC, FOXG1, OTX2, SOX17, EN1, HOXA2* ir *HOXB4* išveiklinimui bei sukurti žmogaus ląstelių liniją su įterptu Cas9, naudojant homologine rekombnacija paremtą genomo redagavimą. *In-Fusion* klonavimo būdu buvo sukonstruotos plazmidės su teigiamu atrankos modeliu genų *SOX1, PAX6, EOMES* ir *SOX17* išveiklinimui, kurie visi yra svarbūs diferenciacijai. Buvo sukurta Jurkat ląstelių linija su įterpta Cas9 seka, skirta stabiliai Cas9 raiškai pasiekti, kuri būtų naudinga redaguojant kelis tos pačios ląstelių linijos lokusus. Cas9 įterpimas į genomą buvo patvirtintas genotipuojant PGR metodu.

VILNIUS UNIVERSITY

LIFE SCIENCES CENTER

Ugnė Šimuliūnaitė

Establishment of Methods for HDR-Based Polyengineering of Human Cells

Master thesis

SUMMARY

Genomic editing of cell lines is essential in developmental biology, disease modelling and clinical research. However, usually, only one or a few genes are edited in the same cell line, limiting the capability of research into gene interactions. A strategy for editing many loci in the same cell line could be beneficial for research on interactions of genes, providing a method for the detailed study of both healthy and disordered development.

This project focuses on the design of a strategy for polyengineering knockouts by HDR-based genome editing. The aim of the research was to create plasmid constructs for the knockouts of the genes SOX1, PAX6, EOMES, GSC, FOXG1, OTX2, SOX17, EN1, HOXA2, HOXB4 and establish a human Cas9 knock-in cell line, using HDR-based genome editing. Plasmids with a positive selection model were created by In-Fusion cloning for the knockout of genes SOX1, PAX6, EOMES and SOX17, all vital for differentiation. A Jurkat cell line with Cas9 knock-in was created for stable expression of Cas9, useful in editing multiple loci of the same cell line. The success of the knock-in was verified by genotyping PCR.

INTRODUCTION

Cell cultures have become a widely used tool for disease modelling and comparing the biological and biochemical processes between healthy and diseased cells (Li et al., 2020; Segeritz & Vallier, 2017). Genome editing of cells is often employed to create cell lines with desired characteristics. In many studies, knockout cell lines are designed to study the function of the affected gene, its role in healthy development and disease (Paul et al., 2021). However, one or a couple of genes are usually investigated simultaneously, which limits the understanding of intricate interactions between these genes or factors encoded by them.

Using a strategy that involves genome editing in multiple loci would allow research on such interactions, providing a method for the detailed study of the development, differentiation, and complex disease modelling. iPSCs would be a good candidate for the cell line used for this type of research, as they are able to differentiate into all cell types and self-renewal capabilities (Medvedev et al., 2010). The knock-in of Cas9 into the genome of a cell line would provide a possibility for a more efficient way to introduce several different genetic modifications into one cell line (Liao et al., 2022).

Aim: To create plasmid constructs for the knockouts of the genes *SOX1, PAX6, EOMES, GSC, FOXG1, OTX2, SOX17, EN1, HOXA2, HOXB4* and establish a human Cas9 expressing cell line, using HDR-based genome editing.

Objectives:

1. Construct plasmids with EGFP and dTomato positive selection models for the knockout of genes *SOX1, PAX6, EOMES, GSC, FOXG1, OTX2, SOX17, EN1, HOXA2* and *HOXB4.*
2. Construct a plasmid for the knock-in of Cas9 into the human genome.
3. Create a human Cas9 expressing cell line and verify the success of the knock-in by genotyping PCR.

1. LITERATURE REVIEW

1.1. The Variety of Human Cell Cultures

Laboratory-cultivated human cells can be separated into three main types: primary, transformed, and self-renewing (Segeritz & Vallier, 2017). Primary cells are obtained directly from human tissues through biopsies or explants. While they are commonly used in biomedical research and are representative of the tissue of origin, they have significant disadvantages. The lifespan of primary cells is finite, and their capacity for replication is limited, which means that the culture is viable for a relatively short time (Gillooly et al., 2012).

Transformed cells are immortalized either naturally or by genetic manipulation. Immortalized cell lines maintain fast growth rates and stable conditions. However, the genome editing may cause abnormalities in their karyotype and phenotype (Segeritz & Vallier, 2017).

Self-renewing cells comprise embryonic stem cells, induced pluripotent stem cells (iPSCs), and adult stem cells. Stem cells have the ability to differentiate into other types of cells and can be maintained in vitro long-term because of their self-renewal (Segeritz & Vallier, 2017). These unique capabilities make self-renewing cells especially practical for many different applications.

1.2. iPSC reprogramming factors

Human iPSCs are a type of pluripotent cells which are obtained by reprogramming differentiated human cells with the use of reprogramming factors (Medvedev et al., 2010). There are two different sets of reprogramming factors used for iPSC reprogramming: Yamanaka's cocktail, which includes OCT4, SOX2, KLF4 and MYC, and Thomson's reprogramming factors, consisting of OCT4, SOX2, NANOG and LIN28 (Takahashi & Yamanaka, 2006; X.-B. Zhang, 2013).

The most important factor for inducing pluripotency appears to be OCT4, as it can directly reprogram CD34+ adult blood cells into mesenchymal stem cells without the use of any additional factors (Meng et al., 2013).

Another significant reprogramming factor is SOX2. Similarly to the ability of OCT4 to directly reprogram cells, it has been found that SOX2 can reprogram fibroblasts into neural stem cells without additional reprogramming factors. As well as that, OCT4 and SOX2 together have been discovered to be sufficient to reprogram fibroblasts into iPSCs (Huangfu et al., 2008).

The third factor in Yamanaka's reprogramming cocktail – KLF4 – has a different function in pluripotency induction. Even though KLF4 is not necessary for reprogramming induction, it is

essential for developing high-quality iPSCs (X.-B. Zhang, 2013). KLF4 has been shown to have a role in the mediation of higher-order chromatin structure, important in inducing and maintaining pluripotency (Wei et al., 2013).

Somatic cell reprogramming happens because of transcriptome and chromatin structure changes. Reprogramming transcription factors can bind to a pluripotency-associated sequence. This ability is affected by changes in chromatin structure, which is influenced by the methylation of DNA, ATP-dependent chromatin modelling, and histone modifications. Reprogramming factors form an autoregulatory circuit, which triggers their promoter genes and works together with other genes involved in pluripotency maintenance (Al Abbar et al., 2020; Takahashi & Yamanaka, 2016).

As mentioned before, the last Yamanaka factor is MYC. It has been characterized as a non-linear amplifier of gene expression (Nie et al., 2012). A drawback of MYC is that it acts as an oncogene. Activation of MYC can induce tumor formation (Carey et al., 2011). This can be considered troubling and the use of MYC should be avoided, if possible, especially in cases of clinical application. And that is possible through the improvement of episomal vectors, as it permits retaining the efficiency of reprogramming without the use of MYC (Su et al., 2013; X.-B. Zhang, 2013), or the replacement of MYC with alterantive variants devoid of oncogenic properties as MYCL (Akifuji et al., 2021).

Other factors, like LIN28 and NANOG are also significant for achieving high efficiency of reprogramming (X.-B. Zhang, 2013). LIN28 protein is important for early embryo development, differentiation of embryonic stem cells and somatic cell reprogramming (Sun et al., 2022). Also, in human cell reprogramming NANOG and LIN28 stimulate FOXH1, which is a downstream effector, important for improved reprogramming (Wang et al., 2019).

Overall, it may be said that there are several main reprogramming factors, which work by different mechanisms and by affecting many secondary downstream factors.

1.3. The methods of human iPSC reprogramming

Several different strategies to produce iPSCs by reprogramming have been developed. The strategies can be separated into two main groups: reprogramming using integrative or nonintegrative transfer systems. Each of these strategies can be done by using viral or nonviral methods (Al Abbar et al., 2020).

Integrating viral vectors were the first reprogramming strategy (Al Abbar et al., 2020). Widely used retroviral vectors provide temporary expression of the exogenous genes, because of epigenetic modification caused silencing of proviral transgene expression in the later stages of the

reprogramming process (Al Abbar et al., 2020; Matsui et al., 2010; Stadtfeld et al., 2008). A significant disadvantage of the retroviral reprogramming system is that for transduction to occur, cells must be actively dividing. Consequently, cells that do not usually divide or divide slowly are less susceptible to transduction. Additionally, if retroviral DNA is integrated into the genome, it can cause continuous expression of transgenes, which can cause mutagenesis and cancer (L. Shao & Wu, 2010).

Another commonly used type of integrating viral vector is lentiviruses. They are considered a subclass of retroviruses. Unlike other retroviruses, lentiviruses infect dividing and non-dividing cells. The site of lentiviral integration is unpredictable, which can cause mutations and promote the development of tumors (L. Shao & Wu, 2010). Lentiviral vectors have broad tropism and are more efficient than other retroviral vectors.

Nonviral transfer systems are considered safer for clinical applications. A gene transfer vector, which is dependent on integration, was designed by including the factors into the loxP sites of the reprogramming construct(Al Abbar et al., 2020). But this method involves risks related to integration and poor reprogramming efficiency (Yamanaka, 2009). These problems were resolved by using mobile genetic elements, like piggyBac (PB) transposons. Using PB transposons to deliver pluripotency factor genes has high efficiency, and allows for the element to be removed from the cell by transient transposase expression. (Al Abbar et al., 2020). However, the human genome contains mobile genetic elements similar to PB transposons (Medvedev et al., 2010). They can cause nonspecific genomic alterations upon transgene excision. They can induce nonspecific genome changes after transgene excision (Al Abbar et al., 2020).

Human iPSCs have also been created using adenoviruses, which are nonintegrating viral vectors (Zhou & Freed, 2009). Human iPSCs have also been created using adenoviruses, which are nonintegrating viral vectors. Even though the absence of integration eliminates the risk of mutation and tumorigenesis, the reprogramming efficiency of nonintegrating viral vectors is around 0.001 %. Yet, the usage of adenoviral methods is promising. Alternatively, Sandai-virus can be used instead of adenovirus. It introduces the foreign genes with higher efficiency, but it’s reprogramming efficiency is poor (Al Abbar et al., 2020).

To avoid vector integration into the genome, transcription factor genes can be delivered into cells by cytoplasmic RNA, episomal (self-replicating and selectable vectors) or polycistronic minicircle DNA nonviral vector systems (Al Abbar et al., 2020). Reprogramming efficiency of minicircle DNA was found to be ~0.005 %, which is higher than the efficiency of nonintegrating viral vectors and previous reprogramming methods based on plasmids (Jia et al., 2010).

1.4.Human cell engineering methods

The genome editing of human cells can be performed by a few different methods. The main three techniques used for for targeted genome editing are Zinc Finger Nuclease (ZFN), Transcription-Activator Like Effector Nucleases (TALEN) and Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR-Associated 9 (CRISPR-Cas9) (D. Gupta et al., 2019).

ZFNs consist of a fused bacterial protein FokI DNA cleavage domain and sequence-specific eukaryotic transcription factor zinc fingers (D. Gupta et al., 2019; Kim et al., 1996). The specificity of ZFNs is dependent on the zinc-finger region with three to six Cys2-His2 fingers, which recognize a separate codon each (Urnov et al., 2010). TALENs are created similarly, by fusing a catalytic domain of FokI endonucleases and a DNA-binding domain from transcription activator-like effectors (TALEs) (H.-X. Zhang et al., 2019).

The third genome engineering method is CRISPR-Cas9. The Cas 9 protein has six domains: REC I, REC II, Bridge Helix, PAM-interacting domain, HNH and RuvC. Rec I is involved in binding guide RNA. The function of REC II is not yet well understood. The bridge helix domain is critical for initiating cleavage after binding to target DNA. The PAM-interacting domain is responsible for PAM specificity and initiates binding to target DNA. HNH and RuvC are nuclease domains, which cut single-stranded DNA. Guide RNA is composed of a 20 bases long single stranded RNA sequence, which creates a T. The guide RNA is designed in a way that its 5′ end is complimentary to the target sequence, which makes it specific to the target DNA (D. Gupta et al., 2019).

The CRISPR/Cas9-based genome editing mechanism consists of three steps: recognition, cleavage, and repair (M. Shao et al., 2016). The sgRNA recognizes the target sequence of the gene of interest. Cas-9 nuclease creates double-stranded breaks (DSBs) 3 base pairs upstream to PAM (Asmamaw & Zawdie, 2021). DNA is then melted, RNA-DNA hybrid is formed. The complementary strand is cleaved by the HNH domain, the non-complementary strand is cleaved by RuvC. The DSB created by the nuclease is repaired by the natural repair pathways of the cell, which include Non-homologous end joining (NHEJ), and homology-directed repair (HDR)(Asmamaw & Zawdie, 2021; M. Shao et al., 2016).

Unlike ZFNs and TALENs, which require using large DNA segments (500–1500 bp), CRISPR-Cas9 can be easily adapted to target any sequence by editing the 20-bp protospacer of the guide RNA , while the Cas9 protein can remain the same (R. M. Gupta & Musunuru, 2014). Also,

recombinant Cas9 protein is easy to produce and complex with sgRNA and deliver by electroporation as RNP complexes, decreases unwanted off-target DSBs. which makes it very convenient for genome editing, especially when targeting multiple loci (Bak et al., 2018). CRISPR-Cas9-based genome editing permits in situ correction of a mutant gene, restoring gene function (de Carvalho et al., 2018).

1.5. HDR-based repair in genome editing

DSBs aren't only generated by engineering techniques but are naturally a common occurrence in eukaryotic cells. Unrepaired DSBs can cause cell apoptosis or senescence and incorrect repair of DSBs can result in genomic instability (Davis & Chen, 2013). There are two main pathways for repairing them: homologous recombination (HR) and NHEJ (Lieber, 2010).

NHEJ is a repair mechanism that doesn't involve homologous sequences (Davis & Chen, 2013). NHEJ can be involved in the repair of DSBs caused by genome editing. Because no homologous sequences are used in the process, NHEJ based repair of DSBs created by Cas9 can result in insertions or deletions (Mali et al., 2013). This can be both an asset and a problem. In cases where precise editing is needed and insertions or deletions would be problematic, NHEJ might cause unwanted mutations, however, these same deletions can be useful for gene knockout or disruption (T. Guo et al., 2018; Mali et al., 2013).

HR or homology-directed repair (HDR) can repair DSBs if the DSB occurs during the S or G2 cell cycle phase and a homologous sequence is present nearby. The ends of the DSB are processed by various nucleases, which results in DNA strand resection, creating single stranded DNA regions at the ends of the break. The overhangs invade the homologous DNA template with the help of the recombinase RAD51, creating the displacement loop (d-loop). Then, DNA polymerases begin DNA synthesis, using the template as a guide. The template strand is displaced by the newly synthesized DNA and nicks are sealed by DNA ligases. Because of the use of a homologous template, this type of DSB repair is extremely precise and can restore the sequence without mutations or induce specific mutations in case of genome editing (M. Liu et al., 2019). The scheme of the HDR mechanism can be seen in fig.1.1.

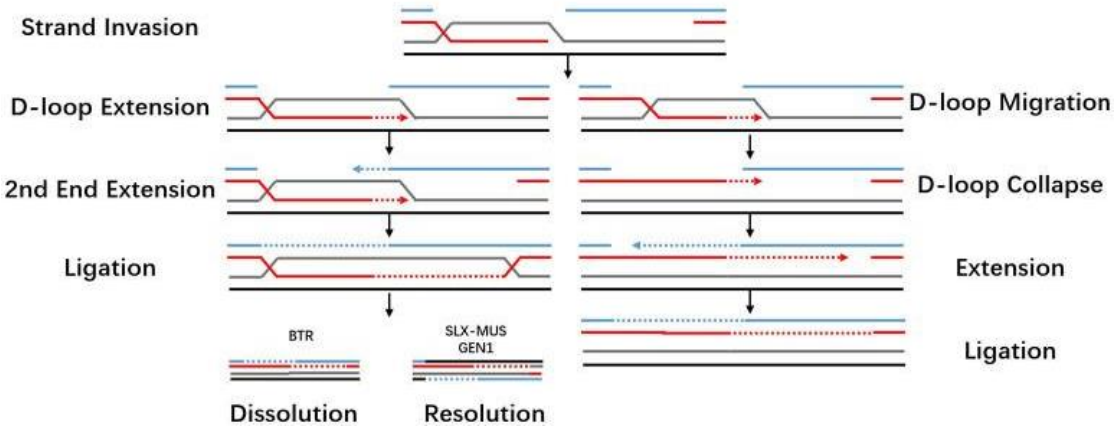

Fig. 1.1. The mechanism of HDR. (M. Liu et al., 2019).

In contrast to NHEJ, HDR can be used to introduce precise modifications, like single nucleotide replacement, codon replacement or insertion of an exogenous gene or reporter. This is done by HR with vectors, containing homologous sequences or homology arms, which work as repair templates (Chu et al., 2015). Some studies suggest that inhibiting NHEJ might improve the precision of CRISPR-Cas9-based genome editing by promoting HDR (Maruyama et al., 2015). However, this approach is not appropriate for clinical applications, because of the negative effect on NHEJ. A safer alternative to promote HDR could be the use of HDR agonists (M. Liu et al., 2019).

Overall, HDR is a precise method of introducing specific mutations by employing the natural cell repair mechanism.

1.6.Target Genes for Knockout Strategy

This project is focused on the knockouts of genes associated with differentiation into different lineages. The genes *SOX1*, *PAX6*, *EOMES*, *SOX17*, *GSC*, *FOXG1*, *OTX2*, *EN1*, *HOXA2* and *HOXB4* were chosen.

SOX1 (SRY-Box Transcription Factor 1) is closely related to *SOX2*, but instead of maintaining pluripotency, it is crucial for neural differentiation (Ahmad et al., 2017). In *SOX1* knockout studies, it was found that homozygous mutant mice, deficient in *Sox1,* have abnormal ventral forebrain development, which causes them to experience epilepsy (Malas et al., 2003). *Sox1* in mice also regulates genes essential for lens development. The deletion of *Sox1* results in abnormally small eye size and cataracts (Nishiguchi et al., 1998). Also, a reporter mouse line was created with *egfp* inserted into the *Sox1* locus. This results in a fluorescent signal in the areas where *Sox1* expression is abundant – the nervous system (Aubert et al., 2003). The fluorescence can be seen in fig. 1.2. *SOX1* has also been found to be important for rostral hindbrain regionalization of

neural precursor cells. Knockout of SOX1 leads to the upregulation of midbrain genes in human embryonic stem cells (X. Liu et al., 2020).

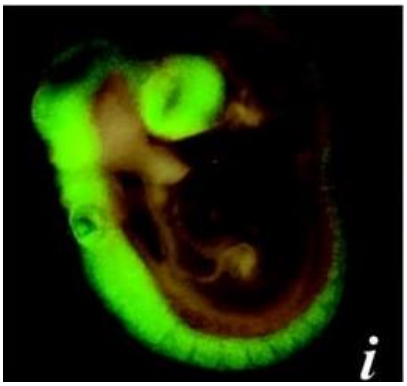

Fig. 1.2. *Sox1*^{GFP} expression in a mouse embryo. *Sox1*^{GFP} embryo at E9.5 showing expression throughout the length of the neural tube (Aubert et al., 2003).

The protein encoded by the *PAX6* (Paired Box 6) gene is also involved in neural development and differentiation. *Pax6* homozygous knockout mice are lethal (Krakowski et al., 2000). *PAX6* is essential for the development of nervous system at olfactory neuroepithelium, retinal pigmented epithelium (Fujimura et al., 2015). It is also important for thalamic neuron development, forebrain neuronal connections, diencephalic neurons (Clegg et al., 2015; Jones et al., 2002; Mastick & Andrews, 2001). *PAX6* is involved in the differentiation of glial cells in the cortex (Götz et al., 1998). It has also been discovered that retinal stem cells require *PAX6* (Xu et al., 2007).

EOMES (Eomesodermin) is another gene important for the development of the nervous system. *EOMES* silencing has been shown to lead to abnormal brain development and microcephaly (Baala et al., 2007). It has been discovered that *EOMES* controls neurogenesis in the embryonic cortical subventricular zone, which is a critical site for generating cortical projection neurons (Arnold et al., 2008). However, *EOMES* is also involved in the induction of mesodermal differentiation and cardiogenesis (X. Guo et al., 2018). Also, it is important for the function of immune cell lineages, such as T cells and natural killer cells (Llaó-Cid et al., 2021; Wagner et al., 2020). Because of this, the knockout of *EOMES* would likely affect not only neural differentiation, but also other differentiation of mesodermal cell lineages.

SOX17 (SRY-Box Transcription Factor 17) is also involved in neural development. *SOX17* has been identified regulate oligodendrocyte development (Sohn et al., 2006). However, it is mostly known as a factor for endodermal differentiation. *Sox17* has been shown to play a role in the differentiation of the extraembryonic endoderm (Shimoda et al., 2007). *SOX17* overexpression in embryonic stem cells (ESCs) causes differentiation of extraembryonic or definitive endoderm cells

(Takayama et al., 2011). In vitro studies have shown that *Sox17* causes the differentiation of mouse ESCs into endodermal lineages, including primitive and definitive endoderm (Qu et al., 2008).

GSC (Goosecoid Homeobox) is involved in the mesodermal differentiation and is involved in the formation of the notochord (Yasuo & Lemaire, 2001). *GSC* has also been shown to be important in the development of neural tissues. It has a double role in ectodermal differentiation. *Xenopus* embryo experiments have suggested that *gsc* downregulates neuroectoderm formation when expressed in the same cell and upregulates when expressed in neighboring cells (Umair et al., 2021). *GSC* is also important for endoderm formation and is considered to be one of the markers of definitive endoderm (Yasunaga et al., 2005).

The FOXG1 (Forkhead Box G1) protein is a factor mainly involved in neuronal differenciation. The down-regulation of *FOXG1* is a prerequisite for the onset of neuronal differentiation during cerebellar development (Adesina et al., 2015). Studies on murine stem cells have shown homozygous knockout of *Foxg1* is associated with reduced differentiation of cortical progenitors (Mall et al., 2017). The FOXG1 protein is important in various neurodevelopmental disorders. It has been discovered that *Foxg1* knockout mice have severe deformities in the inner ear (He et al., 2019).

OTX2 (Orthodenticle Homeobox 2) encodes a neural differenciation factor. It is crucial in early eye development. OTX2 regulates retinal cell differentiation (Diacou et al., 2022). The transcription factor Otx2 is also important for the development of brain, cerebellum and pineal gland. OTX2 deletions cause delays in pituitary development (Mortensen et al., 2015). Heterozygous mutations of OTX2 include brain malformations, ocular or pituitary abnormalities (Beby & Lamonerie, 2013).

The *EN1* (Engrailed 1) gene codes a transcription factor essiantial in embryonic development of cerebellum, midbrain, skeleton, and limbs (Györfi et al., 2021). It is also involved in the development of dopaminergic neurons and might be associated with certain conditions, like schizophrenia and Parkinson's disease (Alves dos Santos & Smidt, 2011). Also, it is impotant for the induction of myofibroblast differentiation. Fibroblast-specific *En1* knockout mice have reduced fibroblast-to-myofibroblast transition (Györfi et al., 2021).

HOXA2 (Homeobox A2) is another gene associated with neural development. Hoxa2 is important tocranial neural crest cells (Kitazawa et al., 2015). *Hoxa2* is expressed in the neural tube and the neural crest-derived mesenchyme (Ohnemus et al., 2001). Also, *Hoxa2* is involved in murine palate development (Iyyanar & Nazarali, 2017; Smith et al., 2009).

The gene *HOXB4* (Homeobox B4) is important for hematopoietic stem cell regulation. It has been demonstrated that HOXB4 increases the efficiency of iPSC differentiation into hematopoietic stem cells (Forrester & Jackson, 2012). The overexpression of *HOXB4* encourages hematopoietic development from human embryonic stem cells (Bowles et al., 2006).

1.7. Cas9 Knock-In Cell Lines

Usually for genome editing of cell lines, exogenous CRISPR-Cas9 is introduced during transfection in the form of a plasmid. However, there have been several studies, which created a Cas9 knock-in cell line. In the study of Platt et al., a Cas9 knock-in mouse model was established. The scientists inserted the gene sequence of Cas9 into the mouse genome using CRISPR-Cas9. It was inserted into the *Rosa26* locus, which is considered a safe harbor locus in mice. This resulted in ubiquitous expression of Cas9 in different cell types, which can be useful for genome editing (Platt et al., 2014).

Similarly, Dow et al. also created a Cas9 knock-in mouse model. In this case, the researchers developed an inducible Cas9 model, by utilizing the Cre-LoxP into the Rosa26 locus in a way that it could be activated in response to Cre recombinase expression. This makes it possible to control Cas9 expression and allows genome editing in specific cell types. Inducible CRISPR-Cas9–mediated genome editing is a simple strategy to create conditional knockout models in under 6 months, allowing the simple study of gene function in vivo (Dow et al., 2015).

The knock-in of Cas9 into the genome of a cell line can facilitate easier and more efficient genome editing. The integration of the Cas9 gene into the genome, lets the cell line have a stable and continuous supply of Cas9, which means that transient transfection of Cas9 expression vectors for each experiment is not necessary, which can save time and improve efficiency (Foley et al., 2022; Platt et al., 2014).

2. MATERIALS AND METHODS

2.1.Equipment And Software

The equipment that was used for the experiments: mini centrifuge Corning Mini Microcentrifuge (Corning), vortex mixer ZX3 Vortex Mixer (Fisherbrand), gas burner Labogaz 206 (Camping Gaz), microcentrifuge AccuSpin Micro 17 Centrifuge (Fisher Scientific), ThermoMixer C (Eppendorf), Vacuubrand BVC Fluid Aspiration Systems (BrandTech Scientific), Navigator NV3202 (Ohaus), power supply PowerPro 300 (Fisherbrand), electrophoresis chamber Owl EasyCast B2 Mini Gel Electrophoresis Systems (Thermo Scientific), spectrophotometer NanoDrop OneC Microvolume UV-Vis Spectrophotometer (Thermo Scientific), thermocycler ProFlex PCR System (Applied Biosystems), cell counter Countess 3 Automated Cell Counter (Invitrogen), Pipetboy acu 2 Pipette Controller (Integra Biosciences), electroporation chamber CU540 Cuvette Chamber (Nepagene), microwave Sharp R971STW COMBI Magnetron 40L RVS (Sharp), gel imager GelDoc Go Imaging System (Bio-Rad), centrifuge GT4 Expert Centrifuge (Fisherbrand), cell culture incubator CB 150 (Binder), laminar flow hood Biowizard GoldenLine GL-200 (Kojair), microscope AE2000 (Motic), shaking incubator Innova 4000 Incubator Orbital Shaker (New Brunswick Scientific), incubator Incucell (MMM Group), electroporator NEPA21 Electroporator (Nepagene). The software used for sequence analysis was ApE (version v3.1.3) and SnapGene (version 6.2). The software used for the creation and editing of images was AffinityPublisher (version 2.0.4).

2.2.Materials

2.2.1. Bacterial Strains and Cell Lines

For plasmid construction and amplification, the bacterial strains NEB 10-beta (New England Biolabs Inc) and Stellar (Takara) were used. The cell lines used for experiments were KISCO-i001.A and Jurkat.

2.2.2. Plasmids

Plasmids, used for the experiments are listed in table 2.1.

Table 2.1. Plasmids, used in experiments.

Plasmid name	Application
p161A	Knockout donor plasmid backbone with EGFP
p162A	Knockout donor plasmid backbone with dTomato
p1005CHA	Cas9 nock-in donor plasmid backbone
pB39555	Cas9 and safe-harbour locus sgRNA containing plasmid
pCR-BluntII-TOPO	TOPO cloning

2.3. Methods

2.3.1. Preparation of Competent *E. coli* Cells

A volume of 10 µl of bacterial stock was thawed and inoculated into 10 ml of liquid LB medium and grown overnight (for 16-18 hours). The next day, 5 ml of overnight culture was transferred into 100 ml of liquid LB medium and grown until the optical density reached 0.5. Then the culture was incubated for 15 min in an ice bath and centrifuged at 3220 × g for 10 min at 4 °C. The supernatant was removed, the pellet was resuspended 25 ml of prechilled 0.1 M $MgCl_2$ and centrifuged again using the same settings. The supernatant was removed. The pellet resuspended in 25 ml of prechilled 0.1 M $CaCl_2$ and incubated on ice for 20 min. After the incubation, the culture was centrifuged again using the same settings. The supernatant was removed, the pellet was resuspended 5 ml of 0.1 M CaCl2 and 15 % glycerol solution. The prepared competent bacteria were aliquoted into 70 µl stocks, flash frozen in liquid nitrogen and stored in -80 °C.

2.3.2. Transformation

Cloning mixture (2 µl of TOPO cloning, 5 µl of In-Fusion cloning or 10 µl of sgRNA ligation mixture) was added to 70 µl of barely thawed competent *E. coli* and incubated on ice for 30 min. Then the mixture was transferred to 42 °C for 45 seconds and quickly moved back to the ice. 1 mL of LB medium without antibiotics was added. The bacteria in then were incubated at 37 °C for 1 hour, shaking at 300 rpm. 500 µl of the mixture is plated on one plate of LB agar with appropriate antibiotic (Carbenicillin disodium salt (Thermo Scientific) or Kanamycin sulfate (Tocris)). The bacteria were grown at 37 °C overnight (for 16-18 hours).

2.3.3. Genomic DNA extraction

Genomic DNA was purified from the cell lines using the DNeasy Blood & Tissue Kit (Qiagen) according to a modified version of the protocol from the manufacturer. 2×10^6 cells were centrifuged at 300 × g for 5 min. The pellet was then resuspended in 200 µl of DPBS (Gibco). 20 µl of Proteinase K and 200 µl of lysis buffer AL was added, then the sample was mixed by vortexing and incubated for 10 min at 56 °C. After the incubation, 200 µl of 100 % ethanol (Honeywell) was added, the sample was mixed by vortexing. The mixture was transferred to a DNeasy Mini spin column and centrifuged at 6000 × g for 1 min. After discarding the flow-through, 500 µl of wash buffer AW1 was added, and the sample was centrifuged at the same conditions. The process was repeated 500 µl of wash buffer AW2, the centrifugation was performed at 20000 × g for 3 min. The spin column was transferred to a new 1.5 ml tube, DNA was eluted in 100 µl nuclease free water, incubated for 1 min at room temperature and centrifuged at 6000× g for 1 min.

2.3.4. Purification of PCR and restriction reaction products

PCR products and restriction reaction products were purified using the the GeneJET Gel Extraction Kit (Thermo Fisher) according to a modified version of the manufacturer's protocol. The PCR product was mixed with Binding Buffer at 1:3 volume ratio. 4^{o}C isopropanol (Honeywell) was added at a 1:1 volume ratio. The solution was mixed well by pipetting and loaded to a spin column and centrifuged at 9600 × g for 1 min. The flowthrough was discarded, 650 µl of Wash Buffer was added and the column was centrifuged using the same settings. The flowthrough was discarded, the empty column was then centrifuged at 13800 × g for 1 min. The column was transferred to a new 1.5 ml tube. The sample was diluted in 100 µl (1 µl for every 1 µl of the initial PCR product) of nuclease free water. The sample was centrifuged at 13800 × g for 1 min.

2.3.5. PCR for Target Gene Sequence Amplification from The Genome and for the creation of Homology Arms

In order to get the gene fragments for *SOX1, PAX6, EOMES, GSC, FOXG1, OTX2, SOX17, EN1*, *HOXA2* and *HOXB4* knockouts, Genomic DNA was purified from the KISCO-i001.A cell line as described in chapter 2.3.3. PCR reactions were performed to amplify the homology arms of the genes of interest. The PCR mixture composition for one sample can be seen in table 2.2.

Table 2.2. PCR mixture composition for target gene sequence amplification from the genome.

Component	Concentration/amount
5X PrimeSTAR® GXL Buffer (Mg2+ plus) (Takara)	1X
dNTP mixture (Takara)	200 µM each
Forward primer	0.2 µM
Reverse primer	0.2 µM
Genomic DNA template	100 ng
PrimeSTAR GXL polymerase (Takara)	2.5 U
Water	to 100 µl

The PCR reaction was performed, the thermocycler settings are provided in table 2.3.

Table 2.3. Thermocycler settings for target gene sequence amplification from the genome.

Temperature	Time	Number of cycles
98 °C	10 s	30
60 °C	15 s	
68 °C	10 s/1 kb	
4 °C	∞	1

The results of the PCR were visualized by electrophoresis, which was performed at 180 V for 40 min. 1 % agarose gel with SYBR Safe I DNA Gel Stain (Invitrogen), and TAE buffer (Thermo Scientific) were used. The ladder used during the electrophoresis was GeneRuler 1 kb DNA Ladder (Thermo Scientific).

For *SOX1* amplification, the PrimeSTAR HS with GC buffer (Takara) was used. The PCR mixture composition for one sample can be seen in table 2.4.

Table 2.4. PCR mixture composition for target gene sequence amplification from the genome with PrimeSTAR HS DNA Polymerase.

Component	Concentration/amount
2X PrimeSTAR GC Buffer (Mg2+ plus) (Takara)	1X
dNTP mixture (Takara)	200 μM each
Forward primer	0.2 μM
Reverse primer	0.2 μM
Genomic DNA template	100 ng
PrimeSTAR HS DNA Polymerase (Takara)	2.5 U
Water	to 100 μl

The thermocycler settings are provided in table 2.5.

Table 2.5. Thermocycler settings for target gene sequence amplification from the genome with PrimeSTAR HS DNA Polymerase.

Temperature	Time	Number of cycles
98°C	20 s	1
98°C	10 s	30
60°C	15 s	
68°C	10 s/1 kb	
4°C	∞	1

The thermocycler settings for the gradient PCR are provided in table 2.6.

Table 2.6. Thermocycler settings for gradient PCR with PrimeSTAR HS DNA Polymerase.

Temperature	Time	Number of cycles
98°C	20 s	1
98°C	10 s	30
60°C/62°C/64°C/66°C/68°C	15 s	
68°C	10 s/1 kb	
4°C	∞	1

PCR products were then purified using the the GeneJET Gel Extraction Kit according to a modified version of the manufacturer's protocol, as described in chapter 2.3.4. The samples were then sent for sequencing at Microsynth Seqlab. PCR products with acceptable sequencing results were used directly for the 2-step PCR reactions, performed as described in chapter 2.3.7. TOPO cloning was used to separate the sequences of different alleles of the genes with worse quality sequencing results. This was done to screen for SNPs and achieve a more homogenous sequence for the knockouts. In this case, the plasmid constructs from TOPO cloning were used for the 2-step PCRs.

2.3.6. Donor plasmid backbone preparation for cloning

The linearized donor plasmid backbones for the cloning of *SOX1, PAX6, EOMES, GSC, FOXG1, OTX2, SOX17, EN1, HOXA2* and *HOXB4* homology arms were prepared by digesting the p161A and p162A plasmids with KspAI (HpaI) (Fisher Scientific) restriction enzyme. The restriction reaction mixture consisted of 4 µl of KspAI enzyme, 8 µl of 10X B buffer (Fisher Scientific), 4 µg of plasmid and up to 80 µl of nuclease free water. The samples were incubated at 37 °C for 2 hours. The results of the PCR were visualized by electrophoresis, which was performed at 180 V for 40 min. 1 % agarose gel with SYBR Safe I dye, and TAE buffer were used. The plasmids were then purified using the the GeneJET Gel Extraction Kit according to a modified version of the manufacturer's protocol.

The linearized donor plasmid backbone p1005CHA for cloning of Cas9 was prepared by Monika Roliūtė.

2.3.7. 2-step PCR for plasmid construction

For the creation of homology arms of the genes *SOX1, PAX6, EOMES, GSC, OTX2, SOX17, HOXA2* knockouts, 2-step PCR reactions were performed. The PCR products of gene fragment extraction from the genome or plasmids from TOPO cloning were used.

The insert for cloning of Cas9 into the plasmid p1005CHA was was created by performing a 2-step PCR reaction. The plasmid pB39555 was used as the template. The reaction mixture is described in table 2.7.

Table 2.7. PCR mixture composition 2-step PCR.

Component	Concentration/amount
5X PrimeSTAR® GXL Buffer (Mg2+ plus)	1X
dNTP mixture	200 µM each
Forward primer	0.2 µM

Extension of table 2.7. PCR mixture composition 2-step PCR.

Reverse primer	0.2 μM
DNA template	100 ng
PrimeSTAR GXL polymerase	1.25 U/μl
Water	to 100 μl

The PCR reaction was performed, the thermocycler settings are provided in table 2.8

Table 2.8. Thermocycler settings for 2-step PCR.

Temperature	Time	Number of cycles
98°C	10 s	3
60°C	15 s	
68°C	10 s/1 kb	
98°C	10 s	27
68°C	10 s/1 kb	
4°C	∞	1

The results of the PCR were visualized by electrophoresis, which was performed at 180 V for 40 min. 1 % agarose gel with SYBR Safe I DNA Gel Stain (Invitrogen), and TAE buffer (Thermo Scientific) were used. The ladder used during the electrophoresis was GeneRuler 1 kb DNA Ladder (Thermo Scientific). The samples were then purified with the GeneJET Gel Extraction Kit.

For the 2-step PCR products of the Cas9 fragment and the homology arms created from TOPO plasmids, an additional step of DpnI (Fisher Scientific) treatment was performed. The restriction reaction mixture consisted of 1 μl of KspAI enzyme, 8 μl of 10X Tango buffer (Fisher Scientific), up to 1 μg of PCR product and up to 20 μl of nuclease free water. The samples were incubated at 37 °C for 2 hours. The plasmids were then purified using the the GeneJET Gel Extraction Kit according to a modified version of the manufacturer's protocol.

2.3.8. TOPO cloning

TOPO Cloning reaction mixture was created, the consisting of 50 ng of PCR product, 1 μl of Salt solution, 1 μl of pCR-BluntII-TOPO plasmid and nuclease free water up to 6 μl. The cloning mixture was mixed gently and incubated for 5 min at room temperature, then used for transformation.

2.3.9. In-Fusion cloning

For In-Fusion cloning of Cas9, the cloning mixture was created, which consisted of 30 pmol of backbone plasmid p1005CHA, 30 pmol of insert, 2 µl 5X In-Fusion HD Enzyme Premix (Takara) µl and nuclease free water up to 10 µl.

For the cloning of knockout homology arms, the cloning mixtures were created, which consisted of 30 pmol of backbone plasmid p161A or p162A, 30 pmol of left homology arm, 30 pmol of right homology arm, 2 µl 5X In-Fusion HD Enzyme Premix (Takara) µl and nuclease free water up to 10 µl. The mixture was incubated at 50 °C for 20 min. then used for transformation.

The mixtures were incubated at 50 °C for 20 min. then used for transformation.

2.3.10. Glycerol stock preparation and plasmid purification

Overnight cultures of transformed *E. coli* were prepared by inoculating a single colony in 2 ml of LB broth with the appropriate antibiotic and grown overnight at 37 °C, shaking at 220 rpm (for 16-18 hours). 500 µl of the culture was mixed with 500 µl of 100% sterile glycerol (Sigma-Aldrich), frozen and placed at -80 °C to create glycerol stocks.

The plasmids were purified by using the the QIAprep Spin Miniprep Kit (Qiagen) according to a modified version of the manufacturer's protocol. 1 ml of the bacterial overnight culture was transferred to a 1,5 ml tube and centrifuged at 17000 × g for 3 min. The supernatant was discarded. Pelleted bacterial cells were resuspended in 250 µl of Buffer P1 by pipetting. 250 µl of Buffer P2 was added and mixed by inverting the tube 4-6 times. 350 µl of Buffer N3 was added and mixed by inverting the tube 4-6 times. The samples were centrifuged for 10 min at 17000 × g. The supernatant was transferred to a spin column and centrifuged for 1 min at 7 000 × g. The flow-through was discarded, 650 µl of Buffer PE was added to the column and it was centrifuged for 1 min at 7 000 × g. The flow-through was discarded, the empty column was centrifuged for 1 min at 7 000 × g. The column was placed into a new 1,5 ml tube. To elute the DNA, 50 µl of nuclease free water was added to the center of the column, and left to stand for 1 min. Then the samples were centrifuged for 2 min at 17 000 × g.

2.3.11. Plasmid-typing

For plasmid screening before sequencing, plasmid-typing was performed using DreamTaq Green PCR Master Mix (Thermo Scientific). The PCR reaction mix for one sample consisted of 12.5 µL DreamTaq Green PCR Master Mix (2X), 0.2 µM forward primer, 0.2 µM reverse primer, 6

ng of plasmid and nuclease free up to 15 µL. For gene knockout plasmids, plasmid-typing was performed before the purification of plasmids directly from the colonies.

Thermocycler settings are described in table 2.9.

Table 2.9. Thermocycler settings for plasmid-typing.

Temperature	Time	Number of cycles
95°C	2 min	1
95°C	30 s	30
60°C	30 s	
72°C	1 min/2kb	
72°C	5 min	1
4°C	∞	1

2.3.12. Plasmid amplification and purification from glycerol stocks

To achieve a higher yield of the plasmids for transfection, 400 ml bacterial cultures were prepared from the glycerol stocks. 40 µl of the glycerol stock was inoculated in 3 ml of LB broth with appropriated antibiotic and grown at 37 °C, shaking at 220 rpm for 8 h. Afterward, 1 ml of the culture is transferred to 400 ml and grown for 16 hours at 37 °C, shaking at 220 rpm.

QIAGEN Plasmid Maxi Kit (Qiagen) was then used to purify the plasmids according to a modified version of the manufacturer's protocol. The bacterial culture was centrifuged at 4000 × g for 20 min at 4°C. The pellet was resuspended in 10 ml Buffer P1, 10 ml of Buffer P2 was added. The solution was mixed by inverting and incubated at room temperature (15–25°C) for 5 min. 10 ml of prechilled Buffer P3 was added, the solution was mixed by inverting. The mixture was centrifuged at 4000 × g for 10 min at 4°C. 6. A QIAGEN-tip was equilibrated by applying 10 ml of Buffer QBT and the column was allowed to empty by gravity flow. The supernatant was transferred to the QIAGEN-tip and allowed to enter the resin by gravity flow. The QIAGEN-tip was washed with 60 ml Buffer QC. DNA was eluted with 15 ml Buffer QF into a clean 50 ml vessel. The plasmids were precipitated by adding 10.5 ml of isopropanol (Honeywell) to the eluted DNA, mixing and incubating for 5 min. The QIAprecipitator was attached to the nozzle of a 30 ml syringe, the mixture was added to the syringe and filtered. The QIAprecipitator was washed with 2 ml of 70% ethanol. The QIAprecipitator was then dried by passing air through it with an empty syringe and leaving it to dry for 2 min. QIAprecipitator was transferred onto a 5 ml syringe and the DNA was eluted with 500 of nuclease free water. QIAprecipitator was dried by passing air through it with an empty syringe into the 1.5 ml tube that was used for elution.

2.3.13. Cell culture cultivation

For the cultivation of the Jurkat cell line, RPMI 1640 Medium (Gibco) with 10 % FBS (Sigma) and 1 % Penicillin Streptomycin (Gibco) was used. The cells were plated at the density of 5.0×10^5 cells/ml and were maintained at the density of 5.0×10^5 – 2.0×10^6 cells/ml. At the density of 2.0×10^6 cells/ml the cells were passaged by centrifuging at 300 xg for 3 min, resuspending in 1 ml of medium, counting the number of live cells and adding enough medium to bring the density down to 5.0×10^5 cells/ml.

For the cultivation of KISCO-i001.A, Essential 8 Medium (Gibco) with Gibco Essential 8 Supplement (Gibco) and 1 % penicillin-streptomycin was used. The 9.6 cm2 plates were coated with 1.5 ml of 10 μg/ml Vitronectin XF (StemCell Technologies) diluted in essential 8 medium with supplement, 1% penicillin-streptomycin and 10μM ROCK inhibitor Y-27632 dihydrochloride (Tocris) and incubated at 37 °C and 5% CO2 for 2 hours before use. The cells were plated at a density of 1.0 ×105 cells per 9.6 cm2 on a plate coated with and the medium was increased to 2.5 ml.

2.3.14. Transfection and selection of Jurkat-Cas9

A mixture of 1 μg of Cas9 and sgRNA containing plasmid pB39555 and 2 μg donor plasmid with Cas9 p1005CHA-Cas9 was created. Cultured Jurkat cells were centrifuged at 300 × g for 3 min, 2×10^6 cells were taken and resuspended in 100 μl Opti-MEM I Reduced Serum Medium (Gibco). The plasmid mixture was added to the cells, then they were transferred to an electroporation chamber. Transfection was performed by electroporation using the settings provided in table 2.10.

Table 2.10. Electroporation settings.

	Voltage, V	Pulse length, ms	Interval length, ms	Pulse number	Decay rate, %
Poring pulse	125	10	50	3	10
Transfer pulse	20	50	50	5	40

After the electroporation, 1 ml of RPMI medium (Gibco) with ROCK inhibitor, 10 % FBS, 1 % penicillin-streptomycin was added, the cells were then transferred to 1 well in a 6 well plate with 4 ml of RPMI medium with 10μM ROCK inhibitor and incubated at 37 °C.

After the cell number exceeded 12×10^6 cells, 2×10^6 were set aside for genotyping. 10×10^6 cells were used for puromycin-based selection. The cells were grown in RPMI medium with 10 %

FBS, 1 % penicillin-streptomycin and 1 µg/ml puromycin for 48 hours. After the first selection, the cells were grown in usual conditions for 72 hours until recovered and then the second round of selection was performed in the same way as the first.

2.3.15. Genotyping

Genotyping was performed before and after puromycin selection. Genomic DNA for the genotyping of the wild type Jurkat and the newly created Jurkat-Cas9 line was extracted using the DNeasy Blood & Tissue Kit (Qiagen) according to the protocol described in chapter 2.3.3.

Genotyping was done using the GXL polymerase. The PCR mixture composition for one sample can be seen in table 2.11.

Table 2.11. PCR mixture composition for genotyping.

Component	Concentration/amount
5X PrimeSTAR® GXL Buffer (Mg2+ plus)	1X
dNTP mixture	200 µM each
Forward primer	0.2 µM
Reverse primer	0.2 µM
Genomic DNA template	100 ng
GXL polymerase	1.25 U/µl
Water	to 25 µl

The PCR reaction was performed, the thermocycler settings are provided in table 2.12.

Table 2.12. Thermocycler settings for genotyping PCR.

Temperature	Time	Number of cycles
98 °C	10 s	30
60 °C	15 s	
68 °C	10 s/1 kb	
4 °C	∞	1

3. RESULTS

In this project, I attempted to develop a strategy for polyengineering of human cells. The aim was to create plasmid constructs for the knockouts of the genes *SOX1, PAX6, EOMES, GSC, FOXG1, OTX2, SOX17, EN1, HOXA2, HOXB4* and to establish a human Cas9 knock-in cell line, using HDR-based genome editing. The main objectives included the construction plasmids for the knockout of genes *SOX1, PAX6, EOMES, GSC, FOXG1, OTX2, SOX17, EN1, HOXA2, HOXB4;* creation of a plasmid for the knock-in of Cas9 into the human genome; and the creation of a human Cas9 knock-in cell line and verification of the knock-in success by genotyping PCR.

3.1. Knockout Target Gene Sequence Amplification from The Genome

Firstly, the target gene sequences were amplified from the genomic DNA of the KISCO-i001.A cell line. The results of the PCR amplification of genes *PAX6, EOMES* and *GSC* can be seen in fig. 3.1.

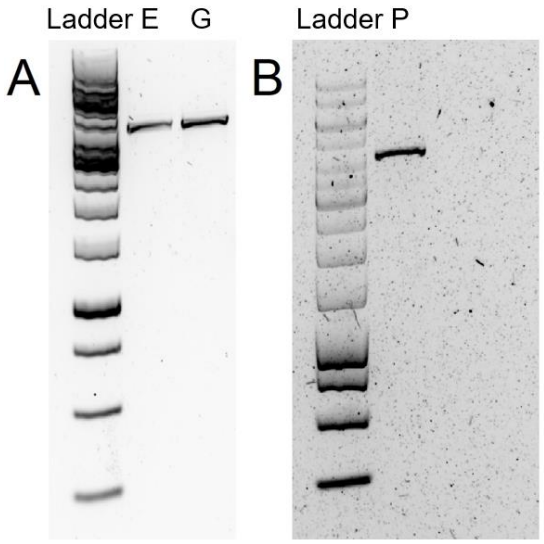

Fig 3.1. Amplification of genes *GSC, EOMES* and *PAX6.* A – the results of *GSC* and *EOMES* amplification. Ladder – GeneRuler 1 kb DNA Ladder, E – *EOMES,* G – *GSC.* B – the results of *PAX6* amplification. Ladder – GeneRuler 1 kb DNA Ladder, P – *PAX6.*

As can be observed in the electrophoresis gel image, in the samples of *EOMES* and *GSC* 4 kb DNA fragments were obtained, which was the expected length. No extra fragments can be observed in either of these samples. Also, in fig. 3.1.A, the ladder appears to be not completely separated. In the case of PAX6, also only a 4 kb length fragment is seen, with no other fragments of any size apparent.

The amplification of *SOX1* with GXL polymerase did not result in any visible bands. A a gradient PCR using the HS polymerase with GC buffer was performed to optimize the conditions. The results are seen in fig. 3.2.

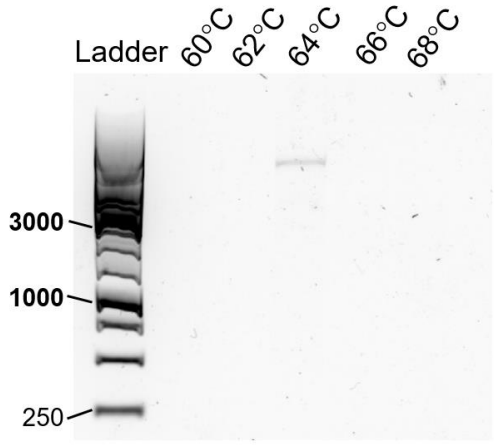

Fig 3.2. Gradient PCR of *SOX1*. Ladder – GeneRuler 1 kb DNA Ladder, 60 °C - 68 °C – the annealing temperatures used for each sample.

As evident from the gel image, no bands were visible at any annealing temperature, except 64 °C. The length of this DNA fragment is difficult to estimate because of the inadequate separation of the ladder. This annealing temperature was chosen and a PCR reaction of a larger volume was performed for the purification of the PCR product. The results are available in fig. 3.3.

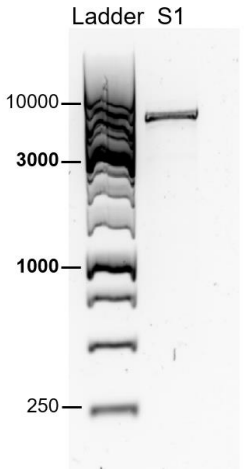

Fig 3.3. Amplification of the gene *SOX1*. Ladder – GeneRuler 1 kb DNA Ladder, S1 – *SOX1*.

A single fragment can be seen in the *SOX1* sample. The specific length was again difficult to estimate because of ladder separation issues, but it was assumed that the fragment was a specific PCR product, and this was examined by sequencing. The gene fragment of the 100 µl PCR appears to be brighter than the fragment from the gradient PCR.

The genes *FOXG1, OTX2, SOX17, HOXB4, EN1* and *HOXA2* were also amplified from the genome. The results are available in fig. 3.4.

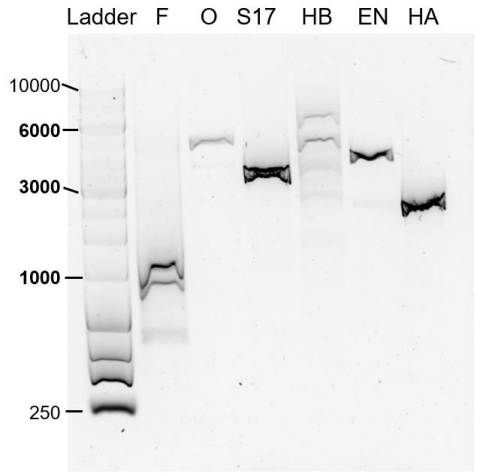

Fig 3.4. Amplification of the genes *FOXG1, OTX2, SOX17, HOXB4, EN1* and *HOXA2.* Ladder – GeneRuler 1 kb DNA Ladder, F – *FOXG1*, O – *OTX2,* S17 – *SOX17*, HB – *HOXB4*, EN – *EN1* and, HB – *HOXA2.*

As can be observed in the image of the gel, in the sample of *FOXG1* 3 bands are visible – around 1 kb long, around 900 bp and 75 bp. For *OTX2*, a 5 kb fragment is visible, for *SOX17* – a 3.5 kb band is apparent. The *HOXB4* sample contained 5 fragments of various sizes (7.5 kb, 6 kb, 4 kb, 3 kb and 1.5 kb). Amplification of *EN1* resulted in a 5 kb fragment and a 3 bp dimmer fragment. In the sample of *HOXA2* a 3 kb band is visible.

Only the bands of *SOX17* and *HOXA2* were of the expected length. For the rest of the genes, amplification was repeated with new primers. The only gene whose amplification resulted in visible fragments was *OTX2*, the results of the PCR of *OTX2* can be observed in fig 3.5.

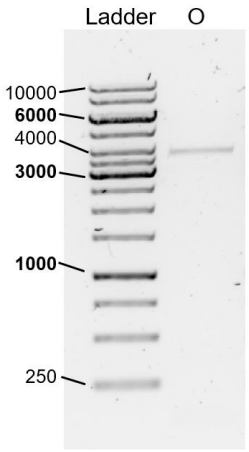

Fig 3.5. Amplification of the gene *OTX2.* Ladder – GeneRuler 1 kb DNA Ladder, O – *OTX2.*

The PCR of *OTX2* resulted in a 4 bp length DNA fragment, which was the expected size.

3.2.Sequencing of amplified gene fragments

After being amplified from genomic DNA, the PCR products were purified and sent to Microsynth Seqlab for sequencing. The sequencing results of *SOX1* can be seen in fig. 3.6.

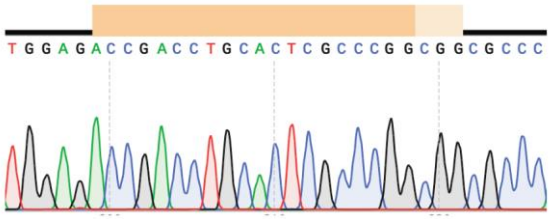

Fig. 3.6. Sequencing results of *SOX1*. The orange bar represents the sgRNA target sequence, the lighter orange bar represents the PAM sequence.

The *SOX1* fragment, amplified from the genome, did not contain any mutations in the sgRNA target sequence, the PAM sequence, or surrounding areas. The signal was clear, and no SNPs were detected.

As seen in fig. 3.7., no clear mutations were detected in the sgRNA target sequence or the PAM sequence of *SOX17* fragment for either of the sgRNA target sequences.

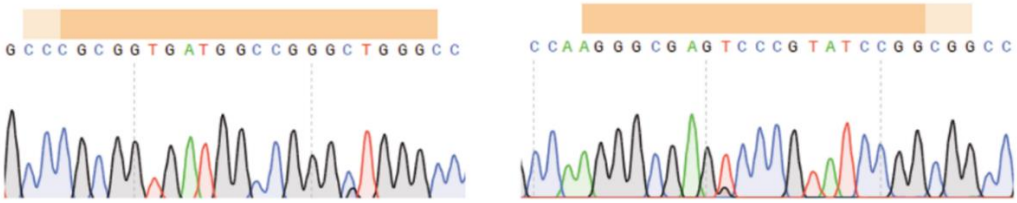

Fig. 3.7. Sequencing results of *SOX17*. The orange bar represents the sgRNA target sequence, the lighter orange bar represents the PAM sequence.

HOXA2 sequencing results are provided in fig. 3.8**.**

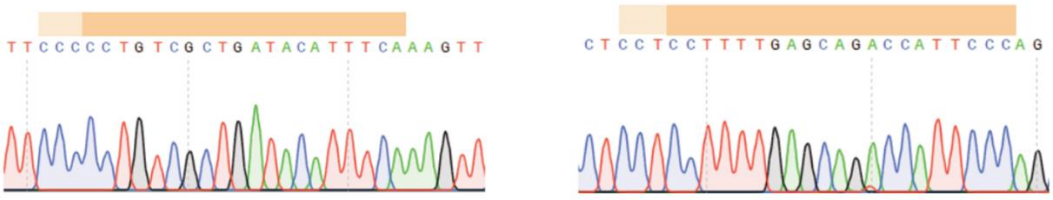

Fig. 3.8. Sequencing results of *HOXA2*. The orange bar represents the sgRNA target sequence, the lighter orange bar represents the PAM sequence.

In the case of *HOXA2*, no mutations were observer in either of the sgRNA target sequences or the PAM sequences. The signal was clear, no double signal or signs of possible SNPs were visible.

The sequencing results of *OTX2* were similarly positive, as seen in fig. 9.

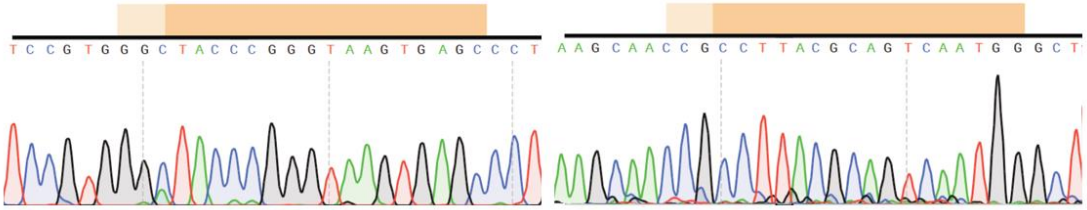

Fig. 9. Sequencing results of *OTX2*. The orange bar represents the sgRNA target sequence, the lighter orange bar represents the PAM sequence.

Both *OTX2* sgRNA target sequences or the PAM sequences did not seem to contain any mutations or SNPs, the signal was clear for the first sgRNA target sequence. However, the sequencing results of the second sgRNA target sequence seemed to have a low level of background signal.

The results of *PAX6* sequencing were less promising, as apparent from fig. 3.10.

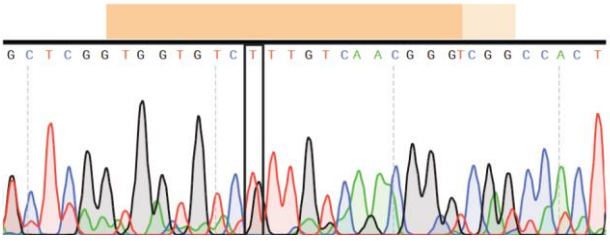

Fig. 3.10. Sequencing results of *PAX6*. The orange bar represents the sgRNA target sequence, the lighter orange bar represents the PAM sequence. The black rectangle marks a possible SNP.

In the case of *PAX6*, there was a strong background signal throughout the sequence and a possible SNP was detected.

The sequencing results of *EOMES*, seen in fig. 11, looked similar.

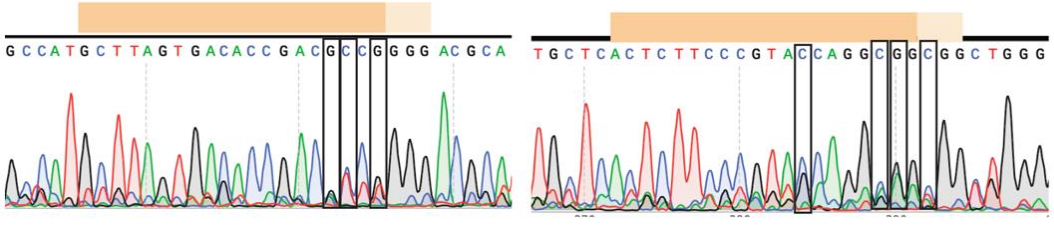

Fig. 3.11. Sequencing results of *EOMES*. The orange bar represents the sgRNA target sequence, the lighter orange bar represents the PAM sequence. The black rectangles mark possible SNPs.

In both sgRNA target sequences and PAM sequences of *EOMES* there was a strong background signal. Multiple possible SNPs were detected in both sequences – 3 in the first on and 4 in the second. One of the SNPs of the second sgRNA was detected in the PAM sequence.

The results of *GSC* sequencing can be seen in fig. 3.12.

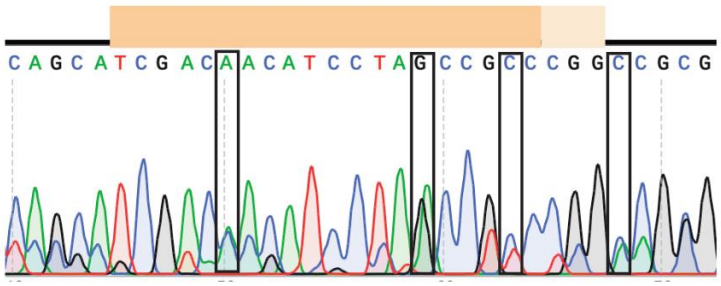

Fig. 3.12. Sequencing results of *GSC.* The orange bar represents the sgRNA target sequence, the lighter orange bar represents the PAM sequence. The black rectangles mark possible SNPs.

In the case of *GSC*, a strong background signal was also observed and there were 4 possible SNPs detected in the sgRNA target sequence and surrounding sequences.

Overall, *SOX1, SOX17, HOXA2* and *OTX2* homology arms could be amplified directly from the sequences, however *PAX6, GSC* and *EOMES* need a TOPO cloning step.

3.3.TOPO cloning results

Because of poor quality sequencing results, the PCR products of genes *GSC, EOMES* and *PAX6* were cloned into the pCR-BluntII-TOPO plasmid. TOPO cloning was used to differentiate between the sequences of different alleles of the genes. The purpose of this was screening for SNPs and achieving a more homogenous sequence for the knockouts.

TOPO cloning sequencing results of *PAX6* are provided in fig. 3.13.

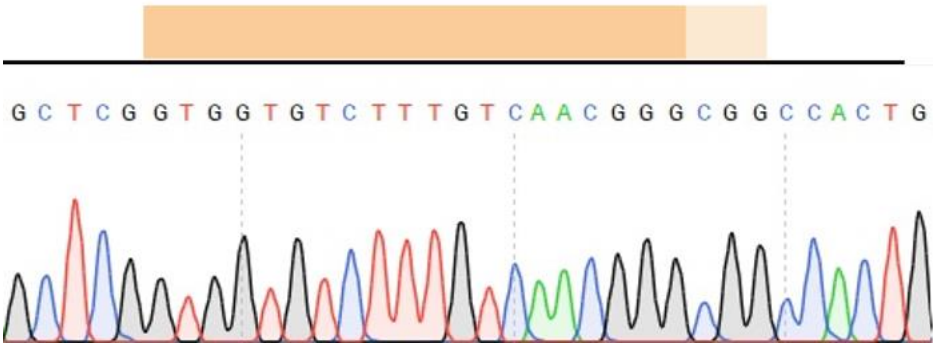

Fig. 3.13. Sequencing results of *PAX6.* The orange bar represents the sgRNA target sequence, the lighter orange bar represents the PAM sequence.

In the TOPO plasmid sequencing results of *PAX6* no background signal was observed. Also, there was no evidence of SNPs. The sample is homogeneous and can be used for homology arm amplification.

In fig. 3.14. the sequencing results of *GSC* TOPO cloning can be seen.

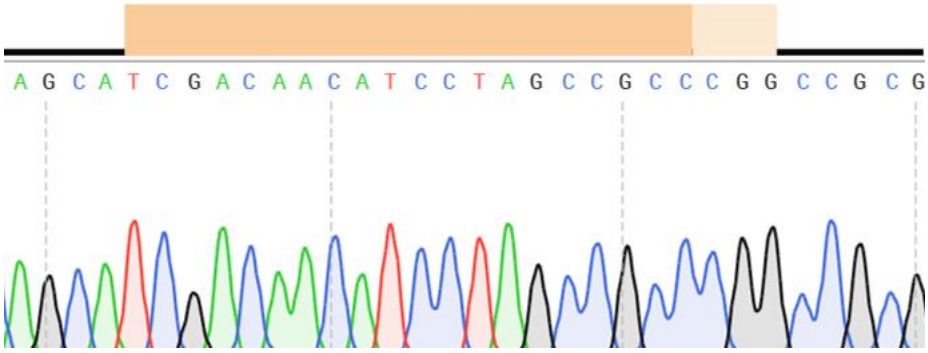

Fig. 3.14. Sequencing results of *GSC.* The orange bar represents the sgRNA target sequence, the lighter orange bar represents the PAM sequence.

In the TOPO cloning sequencing results of *GSC* no background signal and no evidence of SNPs was detected. The sample is homogeneous and of sufficient quality and can be used for homology arm amplification.

The sequencing results of *EOMES* TOPO cloning are available in fig. 3.15.

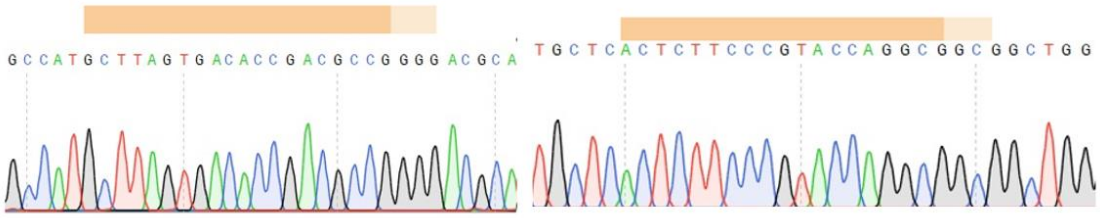

Fig. 3.15. Sequencing results of *EOMES.* The orange bar represents the sgRNA target sequence, the lighter orange bar represents the PAM sequence.

In the TOPO cloning sequencing results of *EOMES* no background signal and no evidence of SNPs was detected for either of the sgRNA target sequences. The sample is homogeneous enough to be used for homology arm amplification.

3.4. 2-step PCR of homology arms

The next step was the construction of homology arms by 2-step PCR. The results of *SOX17* and *HOXA2* homology arm construction by 2-step PCR can be seen in fig. 3.16.

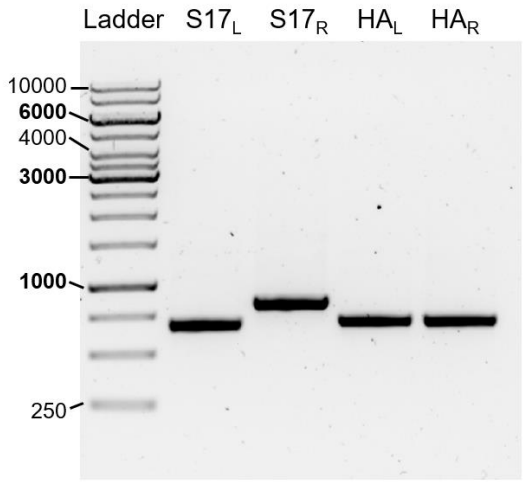

Fig 3.16. 2-step PCR of *SOX17* and *HOXA2*. Ladder – GeneRuler 1 kb DNA Ladder, $S17_L$ – *SOX17* left homology arm, $S17_R$ – *SOX17* right homology arm, HA_L – *HOXA2* left homology arm, HA_R – *HOXA2* right homology arm.

In the *SOX17* left homology arm sample, a 700 bp DNA fragment can be seen, in *SOX17* right homology arm sample – 800 bp DNA fragment. Both left and right *HOXA2* homology arm bands were 750 bp long. All these fragments were of the expected size.

The results of *SOX1* homology arm construction can be seen in fig. 3.17.

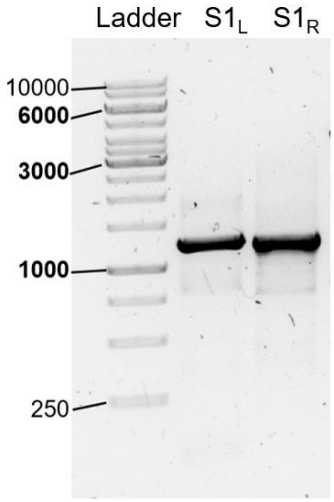

Fig 3.17. 2-step PCR of *SOX1*. Ladder – GeneRuler 1 kb DNA Ladder, $S1_L$ – *SOX1* left homology arm, $S1_R$ – *SOX1* right homology arm.

In the case of *SOX1* left homology arm sample, a 1.3 kb DNA fragment can be seen, as well as in the right homology arm sample, the expected size for both fragments. No additional fragments could be detected.

The results of *GSC* 2-step PCR can be seen in fig. 3.18.

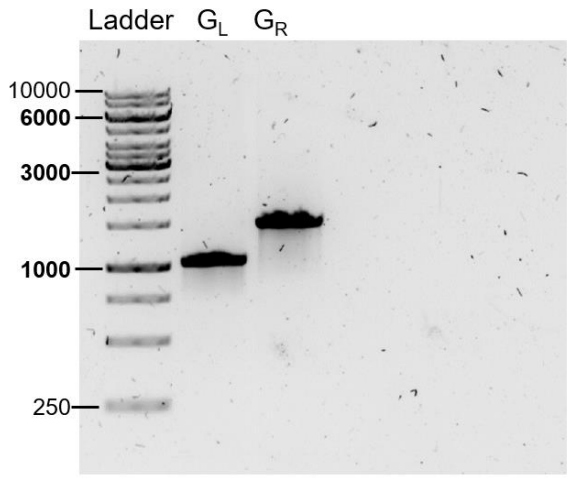

Fig 3.18. 2-step PCR of *GSC*. Ladder – GeneRuler 1 kb DNA Ladder, G_L – *GSC* left homology arm, G_R – *GSC* right homology arm.

In the case of *GSC* left homology arm sample, a 1 kb DNA fragment was obtained, the right homology arm sample band was 1.5 kb long. Both fragments were of the expected size, no additional fragments could be detected.

The results of *OTX2, EOMES* and *PAX6* homology arm construction can be seen in fig. 3.19.

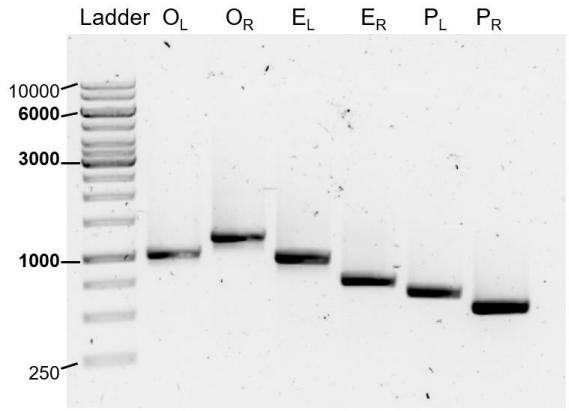

Fig 3.19. 2-step PCR of *OTX2, EOMES* and *PAX6*. Ladder – GeneRuler 1 kb DNA Ladder, O_L – *OTX2* left homology arm, O_R – *OTX2* right homology arm, E_L – *EOMES* left homology arm, E_R – *EOMES* right homology arm, P_L – *PAX6* left homology arm, P_R – *PAX6* right homology arm.

In the case of *OTX2* left homology arm sample, a 1 kb DNA fragment was obtained, the right homology arm sample band was 1.2 kb long. In the *EOMES* left homology arm sample, a 1 kb DNA fragment can be seen, in *EOMES* right homology arm sample – 750 bp DNA fragment. The *PAX6* left homology arm sample contained a 700 bp DNA fragment, the right homology arm sample band was 600 bp long. All these fragments were of the expected size, no unspecific fragments were obtained.

The Cas9 insert for the In-Fusion cloning was also prepared by 2-step PCR, the results are provided in fig. 3.20.

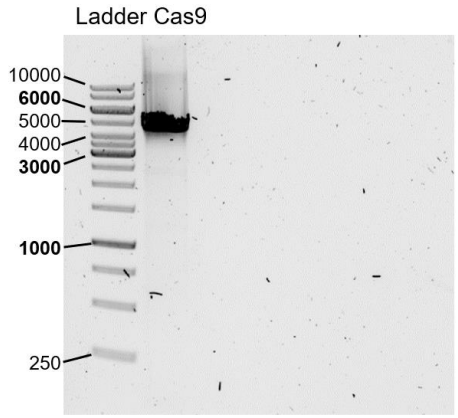

Fig. 3.20. 2-step PCR of Cas9. Ladder – GeneRuler 1 kb DNA Ladder, Cas9 – Cas9 insert.

As evident from the image, the 2-step PCR of the Cas9 insert resulted in a 4.5 kb long DNA fragment, which was the expected length of the insert. No unspecific DNA fragments were obtained.

3.5. The digestion of plasmids p161A and p162A

The backbone linearized donor plasmid backbone for the cloning of homology arms was prepared by digesting the p161A and p162A plasmids with KspAI restriction enzyme. The digestion results are depicted in fig. 3.21.

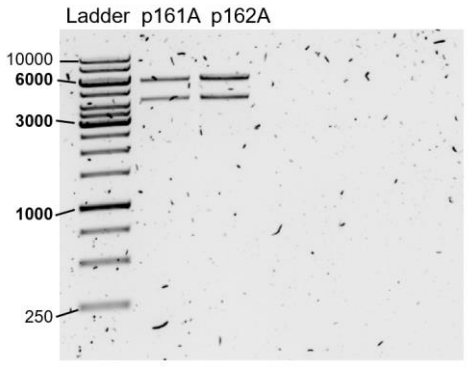

Fig. 3.21. The KspAI digestion of plasmids p161A and p162A. Ladder – GeneRuler 1 kb DNA Ladder, p161A – plasmid p161A, p162A – plasmid p162A.

As evident from the image, in the case of both plasmids, KspAI digestion resulted in 2 fragments – 6kb and 4.5 kb long, which were the fragment lengths expected. The digested plasmid samples were then purified and prepared for cloning.

3.6. Plasmid-typing

The plasmids for the gene knockouts and Cas9 knock-in were created by using In-fusion cloning. Plasmid-typing of p1005CHA-Cas9 plasmids was performed to screen them before sequencing. The results can be seen in fig. 3.22.

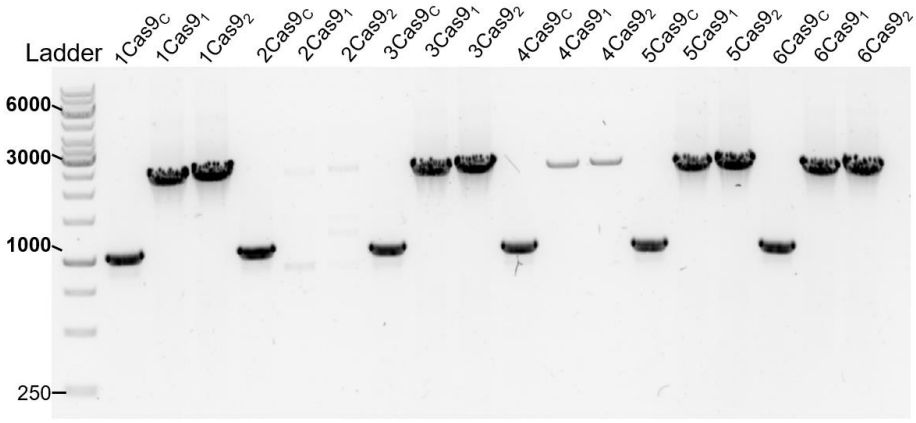

Fig. 3.22. Plasmid-typing of p1005CHA-Cas9. Ladder – GeneRuler 1 kb DNA Ladder. Sample names: $Cas9_C$ – p1005CHA-Cas9 with control primers, $Cas9_1$– p1005CHA-Cas9 with specific primer pair 1, $Cas9_2$ – p1005CHA-Cas9 with specific primer pair 2; the first number in the name is the number of the bacterial clone.

The expected length of the fragments is 1 kb for control primers and 2.5 kb for both specific primer pairs. As can be observed in the image, plasmid samples from all 6 bacterial clones had the fragments of expected length. However, the plasmid from clone 2 and clone 4 had much dimmer bands, also clone 2 had extra bands. So only clones 1, 3, 5 and 6 were used for further experiments.

Plasmid-typing of knockout plasmids was performed to screen the colonies before purifying the plasmids. This step was not performed for plasmid p161A-SOX1 and p161A-SOX1. The results of p161A-EOMES can be seen in fig. 3.23.

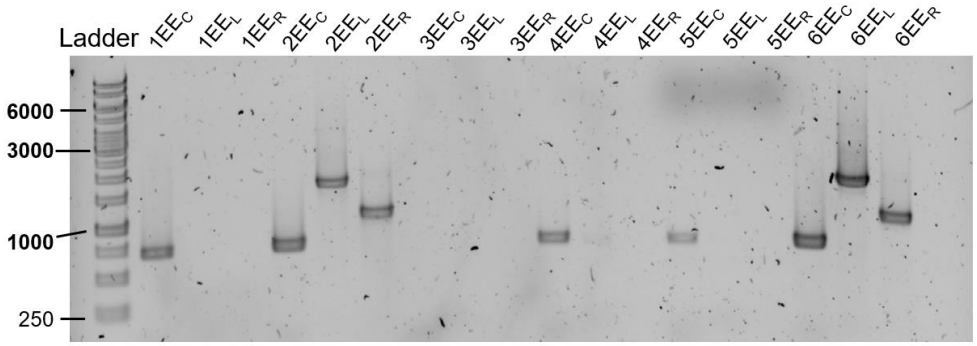

Fig. 3.23. Plasmid-typing of p161A-EOMES. Ladder – GeneRuler 1 kb DNA Ladder. Sample names: EE_C – p161A-EOMES with control primers, EE_L– p161A-EOMES left homology arm, EE_R – p161A-EOMES right homology arm; the first number in the name is the number of the bacterial clone.

The expected length of the fragments of p161A-EOMES plasmid-typing was 750 bp for control primers, 1.5 kb for the left homology arm and 1 kb for the right homology arm. As apparent from the image, only plasmid samples from bacterial clones 2 and 6 had the fragments of expected length. Clones 1, 4 and 5 had only the control fragment and clone 3 did not have any fragments.

The results of the other *EOMES* plasmid p162A-EOMES are provided in fig. 3.24.

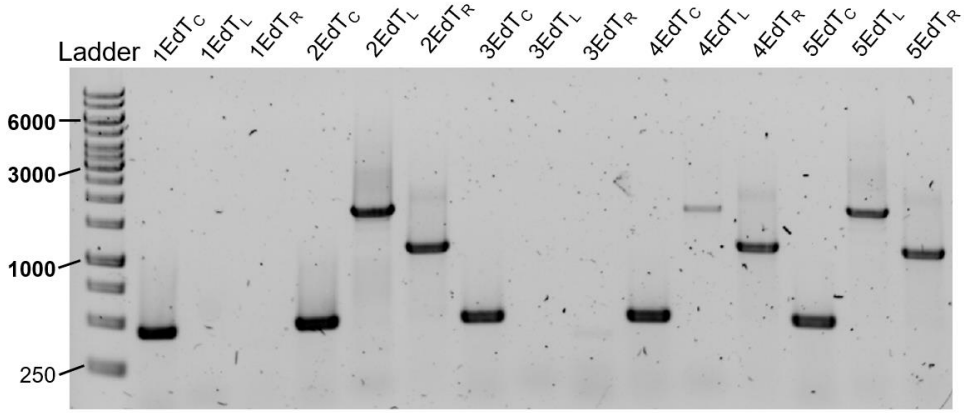

Fig. 3.24. Plasmid-typing of p162A-EOMES. Ladder – GeneRuler 1 kb DNA Ladder. Sample names: EdT_C – p162A-EOMES with control primers, EdT_L– p162A-EOMES left homology arm, EdT_R – p162A-EOMES right homology arm; the first number in the name is the number of the bacterial clone.

The expected length of the fragments was 500 bp for control primers, 1.5 kb for the left homology arm and 1 kb for the right homology arm. The only plasmid samples that had the fragments of expected length were from bacterial clones 2, 4 and 5. Clones 1 and 3 only had the control fragments.

The results *SOX1* plasmid p162A- SOX1 can be observed in fig. 3.25.

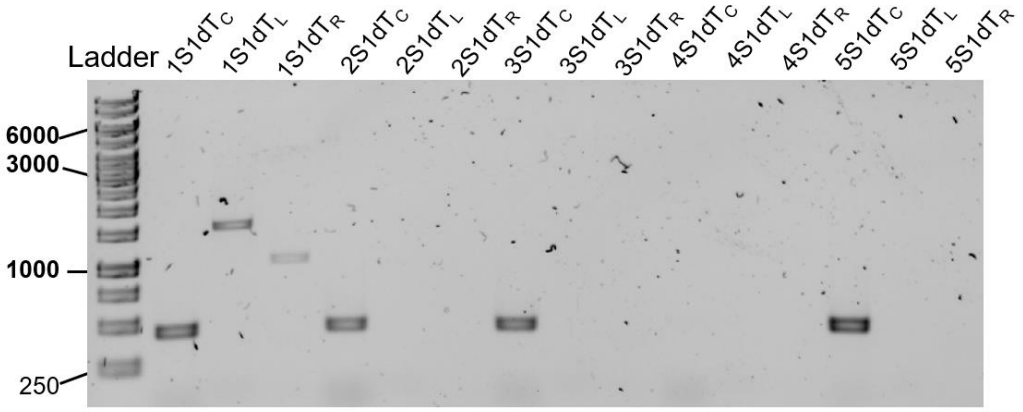

Fig. 3.25. Plasmid-typing of p162A- SOX1. Ladder – GeneRuler 1 kb DNA Ladder. Sample names: $S1dT_C$ – p162A- SOX1 with control primers, $S1dT_L$– p162A- SOX1 left homology arm, $S1dT_R$ – p162A- SOX1 right homology arm; the first number in the name is the number of the bacterial clone.

The expected length of the fragments after p162A- SOX1 plasmid-typing PCR was 500 bp for control primers, 1.5 kb for the left homology arm and 1 kb for the right homology arm. As can be seen in the image, the only plasmid sample that had all the fragments of expected length were from the bacterial clone 1. Clones 2, 3, 4 and 5 only had the control fragments.

The results *PAX6* plasmid p161A- PAX6 are available in fig. 3.26.

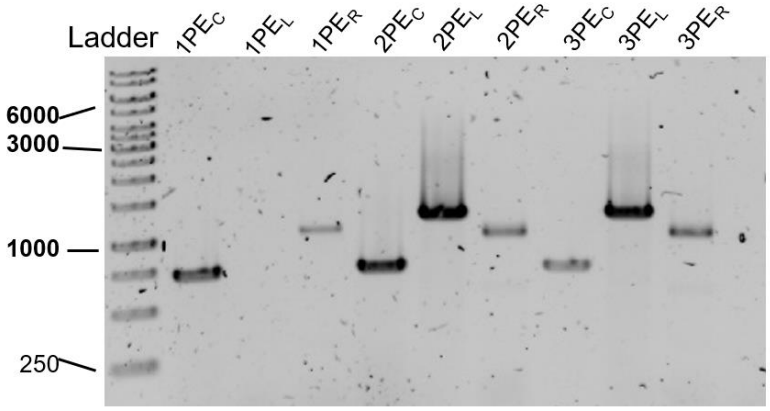

Fig. 3.26. Plasmid-typing of p161A- PAX6. Ladder – GeneRuler 1 kb DNA Ladder. Sample names: PE_C – p161A- PAX6 with control primers, PE_L– p161A- PAX6 left homology arm, PE_R – p161A- PAX6 right homology arm; the first number in the name is the number of the bacterial clone.

The expected length of the fragments was 750 bp for control primers, 1 kb for the left homology arm and 1.5 kb for the right homology arm. The plasmid samples from the bacterial clones 2 and 3 had all the fragments of expected length. Clone 1 had only the control and the right homology arm fragment.

The results *PAX6* plasmid p162A- PAX6 can be seen in fig. 3.27.

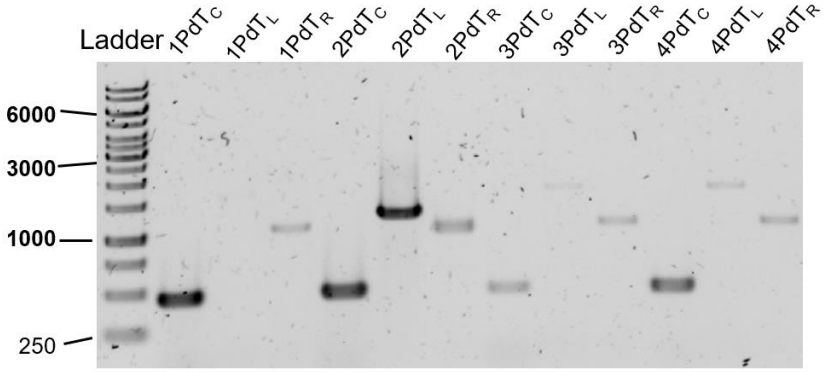

Fig. 3.27. Plasmid-typing of p162A- PAX6. Ladder – GeneRuler 1 kb DNA Ladder. Sample names: PdT_C – p162A - PAX6 with control primers, PdT_L– p162A - PAX6 left homology arm, PdT_R – p162A - PAX6 right homology arm; the first number in the name is the number of the bacterial clone.

The expected length of the fragments after p162A- PAX6 plasmid-typing PCR was 500 bp for control primers, 1.2 kb for the left homology arm and 1 kb for the right homology arm. As can be observed in the image, plasmid samples from clones 2, 3, 4 had all the fragments of expected length. Clones 1 had the control fragment and the right homology arm fragment.

No clones, positive for all 3 fragments were obtained for p162A- SOX17, p161A- GSC, p162A- GSC, p161A- OTX2, p162A- OTX2, p161A- HOXA2 and p162A- HOXA2 plasmids.

Plasmids were purified from the positive clones and sent to be sequenced.

3.7. Plasmid sequencing

The p1005CHA-Cas9 clone 1 plasmid with positive plasmid-typing results was sequenced to verify if it can be used for the knock-in. The results can be seen in fig. 3.28.

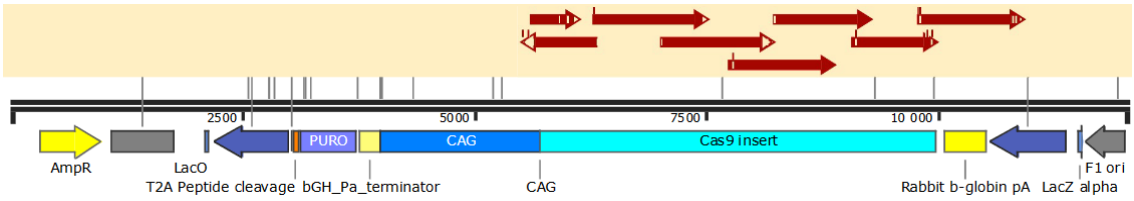

Fig. 3.28. Sequencing results of plasmid p1005CHA-Cas9. The cas insert is highlighted in cyan, the red arrows mark the sequencing reads.

The whole Cas9 insert was covered by 8 sequencing reads. No mutations were detected in the Cas9 insert sequence. The plasmid was selected for electroporation.

In the case of knockout plasmids, p161A-SOX17 did not undergo previous screening by plasmid-typing. The sequencing results are available in fig. 3.29.

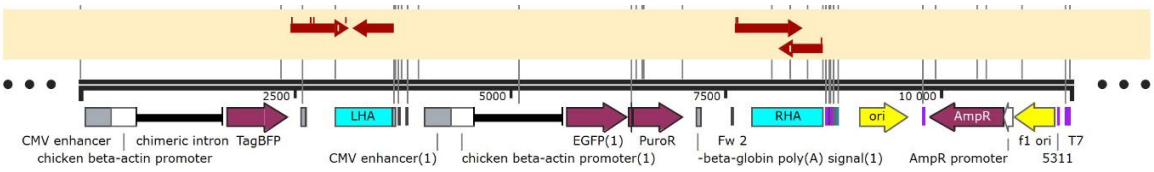

Fig. 3.29. Sequencing results of plasmid p161A-SOX17. The homology arms (LHA – left homolgy arm, RHA – right homology arm) are highlighted in cyan, the red arrows mark the sequencing reads.

As seen in the picture, the plasmid contains both homology arms. There are several mismatches unaccounted for by other reads, however, they would not interfere with our knockout strategy.

Another plasmid previously not screened by plasmid-typing was p161A-SOX1, its sequencing results can be observed in fig 3.30.

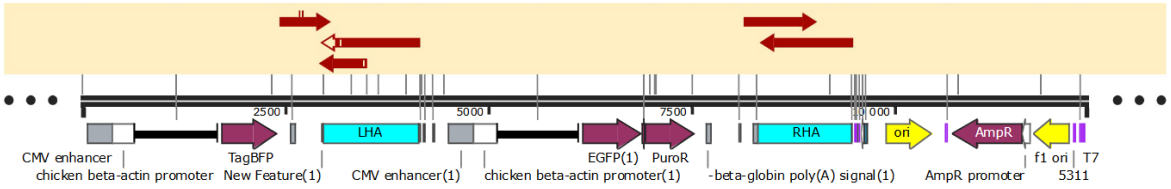

Fig. 3.30. Sequencing results of plasmid p161A-SOX1. The homology arms (LHA – left homolgy arm, RHA – right homology arm) are highlighted in cyan, the red arrows mark the sequencing reads.

The plasmid p161A-SOX1 contained both the left and the right homology arms. The entire sequence of both homology arms is covered by the reads. No mutations were present in either homology arm.

The sequencing results of the plasmid p162A-SOX1 are available in fig. 3.31.

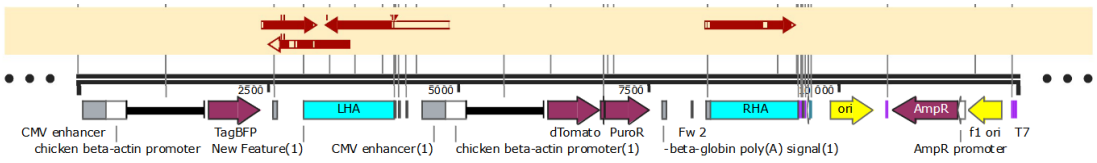

Fig. 3.31. Sequencing results of plasmid p162A-SOX1. The homology arms (LHA – left homolgy arm, RHA – right homology arm) are highlighted in cyan, the red arrows mark the sequencing reads.

As expected because of plasmid-typing results, the plasmid p162A-SOX1 contained both the left and the right homology arms. No mutations significant for the knockout strategy were discovered.

The plasmid p161A-PAX6 was also sequenced, the results can be seen in fig. 3.32.

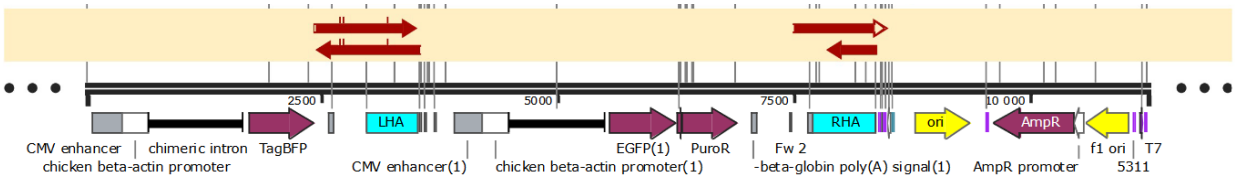

Fig. 3.32. Sequencing results of plasmid p161A-PAX6. The homology arms (LHA – left homolgy arm, RHA – right homology arm) are highlighted in cyan, the red arrows mark the sequencing reads.

As evident from the picture, the plasmid also contains both homology arms. No mutations significant to the knockout approach were present. The reads covered the homology arms well, read quality was sufficient.

In fig. 3.33 the sequencing results of the second plasmid for the knockout of *PAX6* gene – p162A-PAX6 – can be observed.

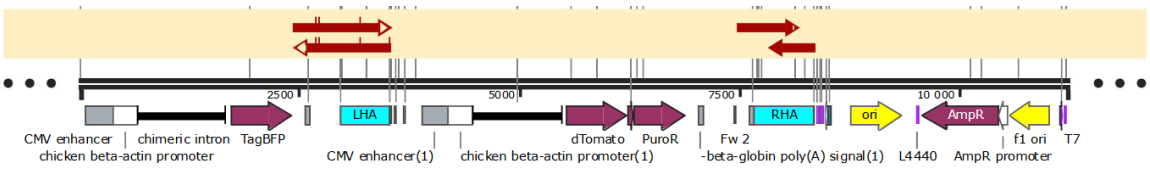

Fig. 3.33. Sequencing results of plasmid p162A-PAX6. The homology arms (LHA – left homolgy arm, RHA – right homology arm) are highlighted in cyan, the red arrows mark the sequencing reads.

The p162A-PAX6 plasmid also contains both homology arms. Any mismatches detected were incidental and insignificant to the knockout strategy.

For the *EOMES* both p161A and p162A plasmids were also obtained and sequenced. The results of the p161A-EOMES sequencing are available in fig. 3.34.

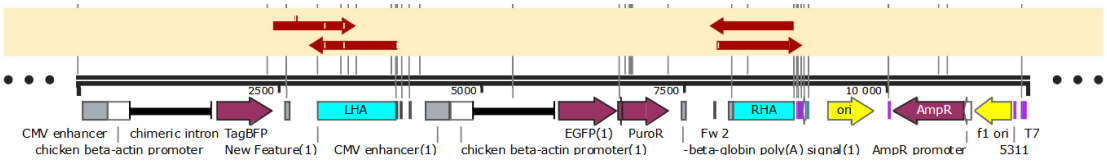

Fig. 3.34. Sequencing results of plasmid p161A-EOMES. The homology arms (LHA – left homolgy arm, RHA – right homology arm) are highlighted in cyan, the red arrows mark the sequencing reads.

The p161A-EOMES plasmid sequencing results also revieled that it contains both homology arms. No mutations, that could cause problems for the knockout strategy were detetected, read quality was sufficient.

The sequencing results of the last plasmid – p162A-EOMES – are provided in fig. 3.35.

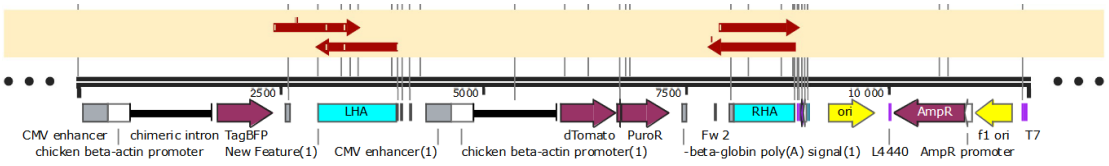

Fig. 3.35. Sequencing results of plasmid p16A-EOMES. The homology arms (LHA – left homolgy arm, RHA – right homology arm) are highlighted in cyan, the red arrows mark the sequencing reads.

The p162A-EOMES plasmid sequencing results also displayed the fact that it contains both *EOMES* homology arms. No mutations, problematic for the knockout strategy were identified.

Overall, the plasmid for Cas9 knock-in, both plasmids for the knockout of *SOX1*, *PAX6*, *EOMES* and the 161A plasmid for the knockout of *SOX17* were constructed.

3.8.The cultivation of KISCO-i001.A

As previously mentioned, the initial goal was to perform the genome editing steps with the iPSC cell line KISCO-i001.A. Because of this, iPSCs were thawed and plated, however, they died 2 days after plating, this can be seen in fig. 3.36.

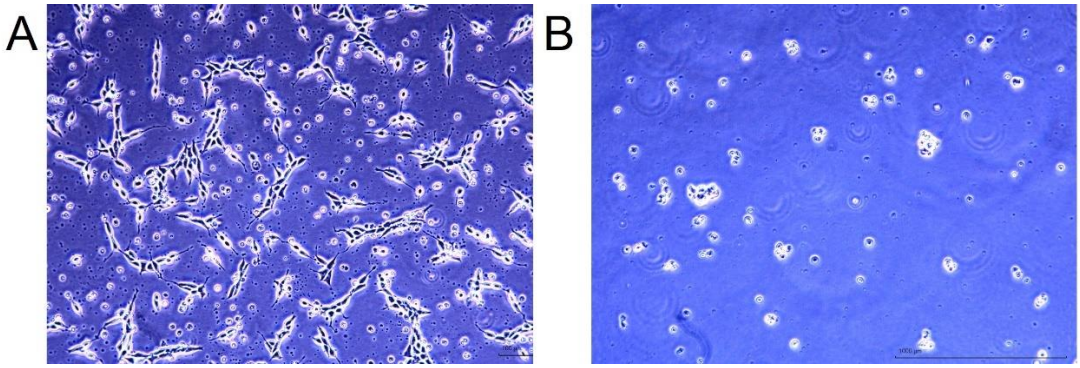

Fig. 3.36. iPSC culture. A – iPSC culture 1 day after plating, 100 scale µm. B - iPSC culture 2 days after plating 1000 µm scale.

As seen in the photos, there were many dead cells even 1 day after plating. However, 2 days after plating no living cells remained. This problem pecisted and troubleshooting attempts were unsuccessful, and Jurkat cells were transfected instead as proof of concept.

3.9. Jurkat-Cas9 selection and genotyping

Jurkat cells were electroporated to create the Cas9 knock-in Jurkat-Cas9 cell line. To verify the success of the Cas9 knock-in into the genome of the Jurkat cell line, a genotyping PCR reaction was performed on the newly created Jurkat-Cas9 cell line before puromycin selection. The results can be seen in fig. 3.37.

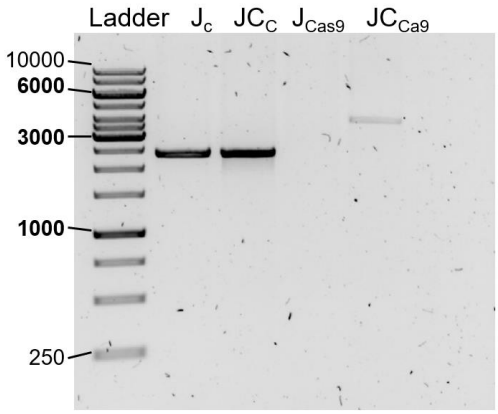

Fig. 3.37. Jurkat-Cas9 genotyping results before selection. Ladder - GeneRuler 1 kb DNA Ladder, J_c – WT Jurkat cell genomic DNA with control primers, JC_C – Jurkat-Cas9 cell genomic DNA with control primers **,** J_{Cas9} –WT Jurkat cell genomic DNA with Cas9 specific primers JC_{Ca9} – Jurkat-Cas9 cell genomic DNA with Cas9 specific primers.

As seen in the elctrophoresis gel picture, both the WT Jurkat cell and Jurkat-Cas9 cell genomic DNA samples, amplified with control primers, could produce around 2329 bp long PCR product. The longer 9865 bp PCR product is not seen in the Jurkat-Cas9 cell sample. However, using the Cas9 specific primers, a 3824 bp long PCR product. A band of such size cannot be seen in the sample of WT Jurkat cells.

Jurkat-Cas9 cells were then selected by puromycin selection. 2 rounds of selection were performed. During the first round of selection, in 2 days the number of live cells decreased by 53.60 % from the starting number. During the second round, the decrase of live cells in 2 days reached 63.70 % from the beginning of the second selection.

After the selection, the genotyping PCR reaction was repeated. The results can be seen in fig. 3.38.

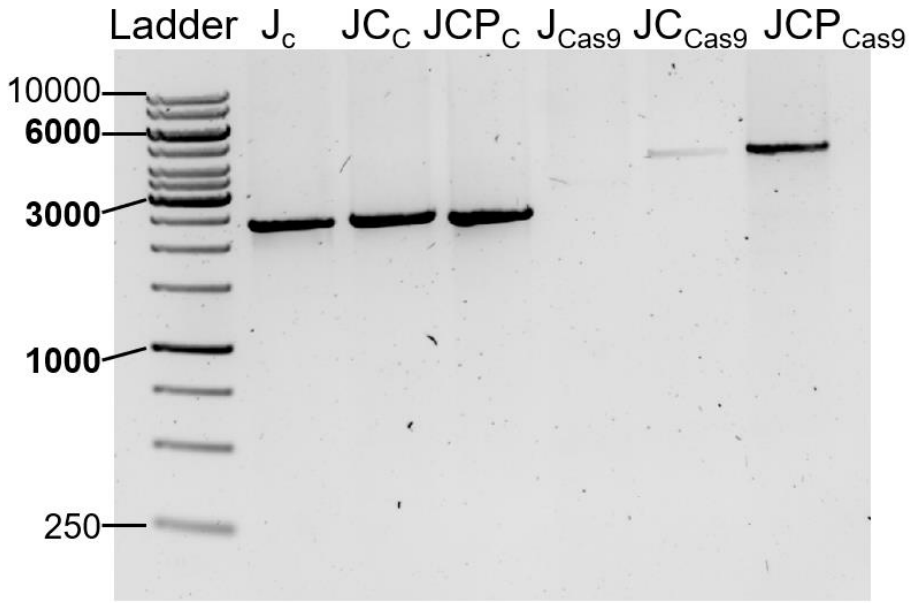

Fig. 3.38. Jurkat-Cas9 genotyping results after puromycin selection. Ladder - GeneRuler 1 kb DNA Ladder, J_c – WT Jurkat cell genomic DNA with control primers, JC_C – genomic DNA of Jurkat-Cas9 cells before selection with control primers, JCP_C – genomic DNA of Jurkat-Cas9 cells after selection with control primers**,** J_{Cas9} –WT Jurkat cell genomic DNA with Cas9 specific primers, JC_{Ca9} – genomic DNA of Jurkat-Cas9 cells before selection with Cas9 specific primers, JCP_{Cas9} – genomic DNA of Jurkat-Cas9 cells after selection with Cas9 specific primers.

As can be observed in the photo of the elctrophoresis gel, the genomic DNA samples of WT Jurkat cells, Jurkat-Cas9 cells before selection and Jurkat-Cas9 cells after selection, amplified with control primers, produced the 2.5 kb PCR product, indicating no integration in the genome. The longer PCR product, suggesting the integration of Cas9, was not observed even after selection as polymerases cannot pass through the high-GC content CAG element of the cargo inserted.

Nevertheless, the Cas9 specific primers (located in the 3’ section of the cargo gene). Nevertheless, the Cas9 specific primers amplified the 4 kb PCR product in Jurkat-Cas9 cell gnomic DNA before and after selection, indicating the presence of Cas9 in the safe harbor locus. The band of Jurkat-Cas9 cell after puromycin selection appears brighter than before selection, which could be the result of increased concentration of genomic DNA with the iserted sequence. A band of such size is not seen in the sample of WT Jurkat cells, indicating that the product observed in the knock-in cells is specific.

4. DISCUSSION

This project was focused on developing a strategy for human cell polyengineering. The aim was to construct plasmids for knocking out the genes *SOX1, PAX6, EOMES, GSC, FOXG1, SOX17, OTX2, EN1, HOXA2* and *HOXB4,* and to create a human Cas9 knock-in cell line, using HDR-based genome editing.

When attempting to amplify knockout target gene sequences the genome, *SOX1, PAX6, EOMES, GSC, SOX17, OTX2* and *HOXA2* sequences were successfully amplified. The sequence of *SOX1* could only be extracted from the genome using PrimeSTAR HS polymerase with GC buffer. This likely happened because of the high GC content of the *SOX1* sequence. The efficiency of PCR is dependent on the nucleotide composition of template DNA and GC-rich regions are considered problematic (Green & Sambrook, 2019). PrimeSTAR HS polymerase with GC buffer is designed to amplify GC-rich regions more efficiently, which is why it helped with amplification of *SOX1*. *FOXG1, EN1* and *HOXB4* amplification failed likely due to different reasons. These genes contained repetitive elements near their target regions. Repetitive elements often cause problems with PCR as they can cause the primers to bind in several loci and tend to generate undesired artifact products as a result (Hommelsheim et al., 2014).

The amplified regions were then sequenced. The sequencing results showed no mutations for the genes *SOX1, SOX17, OTX2* and *HOXA2.* For the genes *PAX6, EOMES* and *GSC* there were high levels of background signal and possible SNPs, were discovered. After TOPO cloning, the sequencing results had no significant background signal, and no mutations or SNPs were seen. The low quality of previous sequencing results could have been due to additional unspecific PCR products that were not seen in the gel due to low concentrations or SNPs affecting only one allele.

The Cas9 insert fragment and the homology arms for the knockouts of *SOX1, PAX6, EOMES, GSC, SOX17, OTX2* and *HOXA2* were successfully amplified, the plasmids p161A and p162A were digested. In-Fusion cloning was performed. Plasmid-typing and sequencing results showed that plasmids were successfully created for the knock-in of Cas9 and the knockouts of *SOX1, PAX6, EOMES* and *SOX17.* For *SOX1, PAX6, EOMES* both versions of the plasmid – p161A and p162A – were constructed, however, for *SOX17* only the plasmid p161A was created. This lowered success rate could have been caused by lower cloning efficiency for 4 fragment cloning.

The iPSC line KISCO-i001.A cultivation was attempted, but the cells didn't seem to be viable and would die on the second day after thawing. Some possible reasons for this could have been

compromised freezing of the cells during transport (Uhrig et al., 2022) or cell culture atmosphere maintenance as iPS cells are fragile to sub-optimal incubation conditions..

Jurkat-Cas9 cell line was then created by electroporation. After genotyping PCR, the fragments indicating no integration into the safe-harbour locus could be seen in the WT Jurkat cells and Jurkat-Cas9 before and after puromycin selection. That is because selection was incomplete and the Jurkat-Cas9 cell line still consists of a mix of WT Jurkat cells and Cas9 containing cells. The large fragment indicating the whole module insertion of Cas9 into the safe-harbour locus was absent Jurkat-Cas9 cells before and after selection. This likely happened because of the high CG content of the CAG element. This problem might be mitigated by using PrimeSTAR HS polymerase with GC buffer for genotyping PCR. However, the fragment, amplified with forward primer that binds on Cas9 sequence downstream from the CAG element and a reverse primer on the safe-harbour locus gene, was obtained in Jurkat-Cas9 cells both before and after puromycin selection, with the band after selection being brighter. This is evidence that the knock-in was successful and may indicate that selection increased the number of Cas9 containing cells sufficiently to make a difference in the efficiency PCR.

CONCLUSIONS

1. Plasmids with EGFP and dTomato positive selection models were constructed for the knockout strategy of genes *SOX1*, *PAX6* and *EOMES*, plasmid with EGFP was created for the knockout strategy of *SOX17*.
2. The plasmid for Cas9 knock-in into the human genome was successfully constructed.
3. Cas9 knock-in Jurkat cell line was established and the knock-in verified with the use of genotyping PCR.

DESCRIPTION OF PERSONAL CONTRIBUTION

My personal contribution to this project was performing all the experiments described, creating the primers and other oligonucleotides, performing sequence and literature analysis, and writing the thesis.

ACKNOWLEDGMENTS

I would like to express gratitude to the people and organizations, who have had significant contributions to the completion of my thesis.

First and foremost, I would like to thank my supervisor, Dr. Jonathan Lee Arias Fuenzalida, for his support, guidance, and encouragement. His expertise and feedback have been invaluable to my growth.

I am grateful to my colleagues, who have helped me and supported me throughout the project. The close-knit and friendly environment of our group was very comforting and encouraging.

I would also like extend my gratitude to Vilnius university and VU LSC-EMBL PI for providing me with the environment and the resources for my research.

REFERENCES

Adesina, A. M., Veo, B. L., Courteau, G., Mehta, V., Wu, X., Pang, K., Liu, Z., Li, X.-N., & Peters, L. (2015). FOXG1 expression shows correlation with neuronal differentiation in cerebellar development, aggressive phenotype in medulloblastomas, and survival in a xenograft model of medulloblastoma. *Human Pathology*, *46*(12), 1859–1871. https://doi.org/10.1016/j.humpath.2015.08.003

Ahmad, A., Strohbuecker, S., Tufarelli, C., & Sottile, V. (2017). Expression of a SOX1 overlapping transcript in neural differentiation and cancer models. *Cellular and Molecular Life Sciences*, *74*(22), 4245–4258. https://doi.org/10.1007/s00018-017-2580-3

Akifuji, C., Iwasaki, M., Kawahara, Y., Sakurai, C., Cheng, Y.-S., Imai, T., & Nakagawa, M. (2021). MYCL promotes iPSC-like colony formation via MYC Box 0 and 2 domains. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-03260-5

Al Abbar, A., Ngai, S. C., Nograles, N., Alhaji, S. Y., & Abdullah, S. (2020). Induced Pluripotent Stem Cells: Reprogramming Platforms and Applications in Cell Replacement Therapy. *BioResearch Open Access*, *9*(1), 121–136. https://doi.org/10.1089/biores.2019.0046

Alves dos Santos, M. T., & Smidt, M. P. (2011). En1 and Wnt signaling in midbrain dopaminergic neuronal development. *Neural Development*, *6*, 23. https://doi.org/10.1186/1749-8104-6-23

Arnold, S. J., Huang, G.-J., Cheung, A. F. P., Era, T., Nishikawa, S.-I., Bikoff, E. K., Molnár, Z., Robertson, E. J., & Groszer, M. (2008). The T-box transcription factor Eomes/Tbr2 regulates neurogenesis in the cortical subventricular zone. *Genes & Development*, *22*(18), 2479–2484. https://doi.org/10.1101/gad.475408

Asmamaw, M., & Zawdie, B. (2021). Mechanism and Applications of CRISPR/Cas-9-Mediated Genome Editing. *Biologics : Targets & Therapy*, *15*, 353–361. https://doi.org/10.2147/BTT.S326422

Aubert, J., Stavridis, M. P., Tweedie, S., O'Reilly, M., Vierlinger, K., Li, M., Ghazal, P., Pratt, T., Mason, J. O., Roy, D., & Smith, A. (2003). Screening for mammalian neural genes via fluorescence-activated cell sorter purification of neural precursors from Sox1-gfp knock-in mice. *Proceedings of the National Academy of Sciences*, *100*(suppl_1), 11836–11841. https://doi.org/10.1073/pnas.1734197100

Baala, L., Briault, S., Etchevers, H. C., Laumonnier, F., Natiq, A., Amiel, J., Boddaert, N., Picard, C., Sbiti, A., Asermouh, A., Attié-Bitach, T., Encha-Razavi, F., Munnich, A., Sefiani, A., & Lyonnet,

S. (2007). Homozygous silencing of T-box transcription factor EOMES leads to microcephaly with polymicrogyria and corpus callosum agenesis. *Nature Genetics*, *39*(4), Article 4. https://doi.org/10.1038/ng1993

Bak, R. O., Dever, D. P., & Porteus, M. H. (2018). CRISPR/Cas9 genome editing in human hematopoietic stem cells. *Nature Protocols*, *13*(2), 358–376. https://doi.org/10.1038/nprot.2017.143

Beby, F., & Lamonerie, T. (2013). The homeobox gene Otx2 in development and disease. *Experimental Eye Research*, *111*, 9–16. https://doi.org/10.1016/j.exer.2013.03.007

Bowles, K. M., Vallier, L., Smith, J. R., Alexander, M. R. J., & Pedersen, R. A. (2006). HOXB4 overexpression promotes hematopoietic development by human embryonic stem cells. *Stem Cells (Dayton, Ohio)*, *24*(5), 1359–1369. https://doi.org/10.1634/stemcells.2005-0210

Carey, B. W., Markoulaki, S., Hanna, J. H., Faddah, D. A., Buganim, Y., Kim, J., Ganz, K., Steine, E. J., Cassady, J. P., Creyghton, M. P., Welstead, G. G., Gao, Q., & Jaenisch, R. (2011). Reprogramming Factor Stoichiometry Influences the Epigenetic State and Biological Properties of Induced Pluripotent Stem Cells. *Cell Stem Cell*, *9*(6), 588–598. https://doi.org/10.1016/j.stem.2011.11.003

Chu, V. T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., & Kühn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nature Biotechnology*, *33*(5), 543–548. https://doi.org/10.1038/nbt.3198

Clegg, J. M., Li, Z., Molinek, M., Caballero, I. M., Manuel, M. N., & Price, D. J. (2015). Pax6 is required intrinsically by thalamic progenitors for the normal molecular patterning of thalamic neurons but not the growth and guidance of their axons. *Neural Development*, *10*(1), 26. https://doi.org/10.1186/s13064-015-0053-7

Davis, A. J., & Chen, D. J. (2013). DNA double strand break repair via non-homologous end-joining. *Translational Cancer Research*, *2*(3), 130–143. https://doi.org/10.3978/j.issn.2218-676X.2013.04.02

de Carvalho, T. G., Schuh, R., Pasqualim, G., Pellenz, F. M., Filippi-Chiela, E. C., Giugliani, R., Baldo, G., & Matte, U. (2018). CRISPR-Cas9-mediated gene editing in human MPS I fibroblasts. *Gene*, *678*, 33–37. https://doi.org/10.1016/j.gene.2018.08.004

Diacou, R., Nandigrami, P., Fiser, A., Liu, W., Ashery-Padan, R., & Cvekl, A. (2022). Cell fate decisions, transcription factors and signaling during early retinal development. *Progress in Retinal and Eye Research*, *91*, 101093. https://doi.org/10.1016/j.preteyeres.2022.101093

Dow, L. E., Fisher, J., O'Rourke, K. P., Muley, A., Kastenhuber, E. R., Livshits, G., Tschaharganeh, D. F., Socci, N. D., & Lowe, S. W. (2015). Inducible in vivo genome editing with CRISPR-Cas9. *Nature Biotechnology*, *33*(4), Article 4. https://doi.org/10.1038/nbt.3155

Foley, R. A., Sims, R. A., Duggan, E. C., Olmedo, J. K., Ma, R., & Jonas, S. J. (2022). Delivering the CRISPR/Cas9 system for engineering gene therapies: Recent cargo and delivery approaches for clinical translation. *Frontiers in Bioengineering and Biotechnology*, *10*, 973326. https://doi.org/10.3389/fbioe.2022.973326

Forrester, L. M., & Jackson, M. (2012). Mechanism of Action of HOXB4 on the Hematopoietic Differentiation of Embryonic Stem Cells. *Stem Cells*, *30*(3), 379–385. https://doi.org/10.1002/stem.1036

Fujimura, N., Klimova, L., Antosova, B., Smolikova, J., Machon, O., & Kozmik, Z. (2015). Genetic interaction between Pax6 and β-catenin in the developing retinal pigment epithelium. *Development Genes and Evolution*, *225*(2), 121–128. https://doi.org/10.1007/s00427-015-0493-4

Gillooly, J. F., Hayward, A., Hou, C., & Burleigh, J. G. (2012). Explaining differences in the lifespan and replicative capacity of cells: A general model and comparative analysis of vertebrates. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1744), 3976–3980. https://doi.org/10.1098/rspb.2012.1129

Götz, M., Stoykova, A., & Gruss, P. (1998). Pax6 Controls Radial Glia Differentiation in the Cerebral Cortex. *Neuron*, *21*(5), 1031–1044. https://doi.org/10.1016/S0896-6273(00)80621-2

Green, M. R., & Sambrook, J. (2019). Polymerase Chain Reaction (PCR) Amplification of GC-Rich Templates. *Cold Spring Harbor Protocols*, *2019*(2), pdb.prot095141. https://doi.org/10.1101/pdb.prot095141

Guo, T., Feng, Y.-L., Xiao, J.-J., Liu, Q., Sun, X.-N., Xiang, J.-F., Kong, N., Liu, S.-C., Chen, G.-Q., Wang, Y., Dong, M.-M., Cai, Z., Lin, H., Cai, X.-J., & Xie, A.-Y. (2018). Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/Cas9-mediated genome editing. *Genome Biology*, *19*(1), 170. https://doi.org/10.1186/s13059-018-1518-x

Guo, X., Xu, Y., Wang, Z., Wu, Y., Chen, J., Wang, G., Lu, C., Jia, W., Xi, J., Zhu, S., Jiapaer, Z., Wan, X., Liu, Z., Gao, S., & Kang, J. (2018). A Linc1405/Eomes Complex Promotes Cardiac Mesoderm Specification and Cardiogenesis. *Cell Stem Cell*, *22*(6), 893-908.e6. https://doi.org/10.1016/j.stem.2018.04.013

Gupta, D., Bhattacharjee, O., Mandal, D., Sen, M. K., Dey, D., Dasgupta, A., Kazi, T. A., Gupta, R., Sinharoy, S., Acharya, K., Chattopadhyay, D., Ravichandiran, V., Roy, S., & Ghosh, D. (2019). CRISPR-Cas9 system: A new-fangled dawn in gene editing. *Life Sciences*, *232*, 116636. https://doi.org/10.1016/j.lfs.2019.116636

Gupta, R. M., & Musunuru, K. (2014). Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. *The Journal of Clinical Investigation*, *124*(10), 4154–4161. https://doi.org/10.1172/JCI72992

Györfi, A.-H., Matei, A.-E., Fuchs, M., Liang, C., Rigau, A. R., Hong, X., Zhu, H., Luber, M., Bergmann, C., Dees, C., Ludolph, I., Horch, R. E., Distler, O., Wang, J., Bengsch, B., Schett, G., Kunz, M., & Distler, J. H. W. (2021). Engrailed 1 coordinates cytoskeletal reorganization to induce myofibroblast differentiation. *The Journal of Experimental Medicine*, *218*(9), e20201916. https://doi.org/10.1084/jem.20201916

He, Z., Fang, Q., Li, H., Shao, B., Zhang, Y., Zhang, Y., Han, X., Guo, R., Cheng, C., Guo, L., Shi, L., Li, A., Yu, C., Kong, W., Zhao, C., Gao, X., & Chai, R. (2019). The role of FOXG1 in the postnatal development and survival of mouse cochlear hair cells. *Neuropharmacology*, *144*, 43–57. https://doi.org/10.1016/j.neuropharm.2018.10.021

Hommelsheim, C. M., Frantzeskakis, L., Huang, M., & Ülker, B. (2014). PCR amplification of repetitive DNA: A limitation to genome editing technologies and many other applications. *Scientific Reports*, *4*(1), Article 1. https://doi.org/10.1038/srep05052

Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., & Melton, D. A. (2008). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nature Biotechnology*, *26*(11), 1269–1275. https://doi.org/10.1038/nbt.1502

Iyyanar, P. P. R., & Nazarali, A. J. (2017). Hoxa2 Inhibits Bone Morphogenetic Protein Signaling during Osteogenic Differentiation of the Palatal Mesenchyme. *Frontiers in Physiology*, *8*, 929. https://doi.org/10.3389/fphys.2017.00929

Jia, F., Wilson, K. D., Sun, N., Gupta, D. M., Huang, M., Li, Z., Robbins, R. C., Kay, M. A., Longaker, M. T., & Wu, J. C. (2010). A Nonviral Minicircle Vector for Deriving Human iPS Cells. *Nature Methods*, *7*(3), 197–199. https://doi.org/10.1038/nmeth.1426

Jones, L., López-Bendito, G., Gruss, P., Stoykova, A., & Molnár, Z. (2002). Pax6 is required for the normal development of the forebrain axonal connections. *Development*, *129*(21), 5041–5052. https://doi.org/10.1242/dev.129.21.5041

Kim, Y. G., Cha, J., & Chandrasegaran, S. (1996). Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(3), 1156–1160.

Kitazawa, T., Fujisawa, K., Narboux-Nême, N., Arima, Y., Kawamura, Y., Inoue, T., Wada, Y., Kohro, T., Aburatani, H., Kodama, T., Kim, K.-S., Sato, T., Uchijima, Y., Maeda, K., Miyagawa-Tomita, S., Minoux, M., Rijli, F. M., Levi, G., Kurihara, Y., & Kurihara, H. (2015). Distinct effects of Hoxa2 overexpression in cranial neural crest populations reveal that the mammalian hyomandibular-ceratohyal boundary maps within the styloid process. *Developmental Biology*, *402*(2), 162–174. https://doi.org/10.1016/j.ydbio.2015.04.007

Krakowski, M., Yeung, B., Abdelmalik, R., Good, A., Mocnik, L., Sosa-Pineda, B., St-Onge, L., Gruss, P., & Sarvetnick, N. (2000). IFN-γ Overexpression Within the Pancreas Is Not Sufficient to Rescue Pax4, Pax6, and Pdx-1 Mutant Mice from Death. *Pancreas*, *21*(4), 399.

Li, H., Yang, Y., Hong, W., Huang, M., Wu, M., & Zhao, X. (2020). Applications of genome editing technology in the targeted therapy of human diseases: Mechanisms, advances and prospects. *Signal Transduction and Targeted Therapy*, *5*(1), Article 1. https://doi.org/10.1038/s41392-019-0089-y

Liao, J. Q., Zhou, G., & Zhou, Y. (2022). Generation of Monoclonal iPSC Lines with Stable Cas9 Expression and High Cas9 Activity. In A. Nagy & K. Turksen (Eds.), *Induced Pluripotent Stem (iPS) Cells: Methods and Protocols* (pp. 575–588). Springer US. https://doi.org/10.1007/7651_2020_304

Lieber, M. R. (2010). The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End Joining Pathway. *Annual Review of Biochemistry*, *79*, 181–211. https://doi.org/10.1146/annurev.biochem.052308.093131

Liu, M., Rehman, S., Tang, X., Gu, K., Fan, Q., Chen, D., & Ma, W. (2019). Methodologies for Improving HDR Efficiency. *Frontiers in Genetics*, *9*, 691. https://doi.org/10.3389/fgene.2018.00691

Liu, X., Fang, Z., Wen, J., Tang, F., Liao, B., Jing, N., Lai, D., & Jin, Y. (2020). SOX1 Is Required for the Specification of Rostral Hindbrain Neural Progenitor Cells from Human Embryonic Stem Cells. *IScience*, *23*(9), 101475. https://doi.org/10.1016/j.isci.2020.101475

Llaó-Cid, L., Roessner, P. M., Chapaprieta, V., Öztürk, S., Roider, T., Bordas, M., Izcue, A., Colomer, D., Dietrich, S., Stilgenbauer, S., Hanna, B., Martín-Subero, J. I., & Seiffert, M. (2021). EOMES is

essential for antitumor activity of CD8+ T cells in chronic lymphocytic leukemia. *Leukemia*, *35*(11), 3152–3162. https://doi.org/10.1038/s41375-021-01198-1

Malas, S., Postlethwaite, M., Ekonomou, A., Whalley, B., Nishiguchi, S., Wood, H., Meldrum, B., Constanti, A., & Episkopou, V. (2003). Sox1-deficient mice suffer from epilepsy associated with abnormal ventral forebrain development and olfactory cortex hyperexcitability. *Neuroscience*, *119*(2), 421–432. https://doi.org/10.1016/S0306-4522(03)00158-1

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., Yang, L., & Church, G. M. (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology*, *31*(9), Article 9. https://doi.org/10.1038/nbt.2675

Mall, E. M., Herrmann, D., & Niemann, H. (2017). Murine pluripotent stem cells with a homozygous knockout of Foxg1 show reduced differentiation towards cortical progenitors in vitro. *Stem Cell Research*, *25*, 50–60. https://doi.org/10.1016/j.scr.2017.10.012

Maruyama, T., Dougan, S. K., Truttmann, M. C., Bilate, A. M., Ingram, J. R., & Ploegh, H. L. (2015). Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nature Biotechnology*, *33*(5), Article 5. https://doi.org/10.1038/nbt.3190

Mastick, G. S., & Andrews, G. L. (2001). Pax6 Regulates the Identity of Embryonic Diencephalic Neurons. *Molecular and Cellular Neuroscience*, *17*(1), 190–207. https://doi.org/10.1006/mcne.2000.0924

Matsui, T., Leung, D., Miyashita, H., Maksakova, I. A., Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M. C., & Shinkai, Y. (2010). Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature*, *464*(7290), 927–931. https://doi.org/10.1038/nature08858

Medvedev, S. P., Shevchenko, A. I., & Zakian, S. M. (2010). Induced Pluripotent Stem Cells: Problems and Advantages when Applying them in Regenerative Medicine. *Acta Naturae*, *2*(2), 18–28.

Meng, X., Su, R.-J., Baylink, D. J., Neises, A., Kiroyan, J. B., Lee, W. Y.-W., Payne, K. J., Gridley, D. S., Wang, J., Lau, K.-H. W., Li, G., & Zhang, X.-B. (2013). Rapid and efficient reprogramming of human fetal and adult blood CD34+ cells into mesenchymal stem cells with a single factor. *Cell Research*, *23*(5), 658–672. https://doi.org/10.1038/cr.2013.40

Mortensen, A. H., Schade, V., Lamonerie, T., & Camper, S. A. (2015). Deletion of OTX2 in neural ectoderm delays anterior pituitary development. *Human Molecular Genetics*, *24*(4), 939–953. https://doi.org/10.1093/hmg/ddu506

Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D. R., Tessarollo, L., Casellas, R., Zhao, K., & Levens, D. (2012). C-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell*, *151*(1), 68–79. https://doi.org/10.1016/j.cell.2012.08.033

Nishiguchi, S., Wood, H., Kondoh, H., Lovell-Badge, R., & Episkopou, V. (1998). Sox1 directly regulates the γ-crystallin genes and is essential for lens development in mice. *Genes & Development*, *12*(6), 776–781.

Ohnemus, S., Bobola, N., Kanzler, B., & Mallo, M. (2001). Different levels of Hoxa2 are required for particular developmental processes. *Mechanisms of Development*, *108*(1), 135–147. https://doi.org/10.1016/S0925-4773(01)00502-0

Paul, A., Anand, R., Karmakar, S. P., Rawat, S., Bairagi, N., & Chatterjee, S. (2021). Exploring gene knockout strategies to identify potential drug targets using genome-scale metabolic models. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-020-80561-1

Platt, R. J., Chen, S., Zhou, Y., Yim, M. J., Swiech, L., Kempton, H. R., Dahlman, J. E., Parnas, O., Eisenhaure, T. M., Jovanovic, M., Graham, D. B., Jhunjhunwala, S., Xavier, R. J., Langer, R., Anderson, D. G., Hacohen, N., Regev, A., Feng, G., Sharp, P. A., & Zhang, F. (2014). CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. *Cell*, *159*(2), 440–455. https://doi.org/10.1016/j.cell.2014.09.014

Qu, X.-B., Pan, J., Zhang, C., & Huang, S.-Y. (2008). Sox17 facilitates the differentiation of mouse embryonic stem cells into primitive and definitive endoderm in vitro. *Development, Growth & Differentiation*, *50*(7), 585–593. https://doi.org/10.1111/j.1440-169X.2008.01056.x

Segeritz, C.-P., & Vallier, L. (2017). Cell Culture. In *Basic Science Methods for Clinical Researchers* (pp. 151–172). Elsevier. https://doi.org/10.1016/B978-0-12-803077-6.00009-6

Shao, L., & Wu, W.-S. (2010). Gene-delivery systems for iPS cell generation. *Expert Opinion on Biological Therapy*, *10*(2), 231–242. https://doi.org/10.1517/14712590903455989

Shao, M., Xu, T.-R., & Chen, C.-S. (2016). The big bang of genome editing technology: Development and application of the CRISPR/Cas9 system in disease animal models. *Zoological Research*, *37*(4), 191–204. https://doi.org/10.13918/j.issn.2095-8137.2016.4.191

Shimoda, M., Kanai-Azuma, M., Hara, K., Miyazaki, S., Kanai, Y., Monden, M., & Miyazaki, J. (2007). Sox17 plays a substantial role in late-stage differentiation of the extraembryonic endoderm in vitro. *Journal of Cell Science*, *120*(21), 3859–3869. https://doi.org/10.1242/jcs.007856

Smith, T. M., Wang, X., Zhang, W., Kulyk, W., & Nazarali, A. J. (2009). Hoxa2 plays a direct role in murine palate development. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *238*(9), 2364–2373. https://doi.org/10.1002/dvdy.22040

Sohn, J., Natale, J., Chew, L.-J., Belachew, S., Cheng, Y., Aguirre, A., Lytle, J., Nait-Oumesmar, B., Kerninon, C., Kanai-Azuma, M., Kanai, Y., & Gallo, V. (2006). Identification of Sox17 as a Transcription Factor That Regulates Oligodendrocyte Development. *The Journal of Neuroscience*, *26*(38), 9722–9735. https://doi.org/10.1523/JNEUROSCI.1716-06.2006

Stadtfeld, M., Maherali, N., & Hochedlinger, K. (2008). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell*, *2*(3), 230–240. https://doi.org/10.1016/j.stem.2008.02.001

Su, R.-J., Baylink, D. J., Neises, A., Kiroyan, J. B., Meng, X., Payne, K. J., Tschudy-Seney, B., Duan, Y., Appleby, N., Kearns-Jonker, M., Gridley, D. S., Wang, J., Lau, K.-H. W., & Zhang, X.-B. (2013). Efficient Generation of Integration-Free iPS Cells from Human Adult Peripheral Blood Using BCL-XL Together with Yamanaka Factors. *PLoS ONE*, *8*(5), e64496. https://doi.org/10.1371/journal.pone.0064496

Sun, Z., Yu, H., Zhao, J., Tan, T., Pan, H., Zhu, Y., Chen, L., Zhang, C., Zhang, L., Lei, A., Xu, Y., Bi, X., Huang, X., Gao, B., Wang, L., Correia, C., Chen, M., Sun, Q., Feng, Y., … Zhang, J. (2022). LIN28 coordinately promotes nucleolar/ribosomal functions and represses the 2C-like transcriptional program in pluripotent stem cells. *Protein & Cell*, *13*(7), 490–512. https://doi.org/10.1007/s13238-021-00864-5

Takahashi, K., & Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, *126*(4), 663–676. https://doi.org/10.1016/j.cell.2006.07.024

Takahashi, K., & Yamanaka, S. (2016). A decade of transcription factor-mediated reprogramming to pluripotency. *Nature Reviews Molecular Cell Biology*, *17*(3), 183–193. https://doi.org/10.1038/nrm.2016.8

Takayama, K., Inamura, M., Kawabata, K., Tashiro, K., Katayama, K., Sakurai, F., Hayakawa, T., Furue, M. K., & Mizuguchi, H. (2011). Efficient and Directive Generation of Two Distinct

Endoderm Lineages from Human ESCs and iPSCs by Differentiation Stage-Specific SOX17 Transduction. *PLOS ONE*, *6*(7), e21780. https://doi.org/10.1371/journal.pone.0021780

Uhrig, M., Ezquer, F., & Ezquer, M. (2022). Improving Cell Recovery: Freezing and Thawing Optimization of Induced Pluripotent Stem Cells. *Cells*, *11*(5), 799. https://doi.org/10.3390/cells11050799

Umair, Z., Kumar, V., Goutam, R. S., Kumar, S., Lee, U., & Kim, J. (2021). Goosecoid Controls Neuroectoderm Specification via Dual Circuits of Direct Repression and Indirect Stimulation in Xenopus Embryos. *Molecules and Cells*, *44*(10), 723–735. https://doi.org/10.14348/molcells.2021.0055

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., & Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics*, *11*(9), 636–646. https://doi.org/10.1038/nrg2842

Wagner, J. A., Wong, P., Schappe, T., Berrien-Elliott, M. M., Cubitt, C., Jaeger, N., Lee, M., Keppel, C. R., Marin, N. D., Foltz, J. A., Marsala, L., Neal, C. C., Sullivan, R. P., Schneider, S. E., Keppel, M. P., Saucier, N., Cooper, M. A., & Fehniger, T. A. (2020). Stage-Specific Requirement for Eomes in Mature NK Cell Homeostasis and Cytotoxicity. *Cell Reports*, *31*(9), 107720. https://doi.org/10.1016/j.celrep.2020.107720

Wang, L., Su, Y., Huang, C., Yin, Y., Zhu, J., Knupp, A., Chu, A., & Tang, Y. (2019). FOXH1 Is Regulated by NANOG and LIN28 for Early-stage Reprogramming. *Scientific Reports*, *9*, 16443. https://doi.org/10.1038/s41598-019-52861-8

Wei, Z., Gao, F., Kim, S., Yang, H., Lyu, J., An, W., Wang, K., & Lu, W. (2013). Klf4 Organizes Long-Range Chromosomal Interactions with the Oct4 Locus in Reprogramming and Pluripotency. *Cell Stem Cell*, *13*(1), 36–47. https://doi.org/10.1016/j.stem.2013.05.010

Xu, S., Sunderland, M. E., Coles, B. L. K., Kam, A., Holowacz, T., Ashery-Padan, R., Marquardt, T., McInnes, R. R., & van der Kooy, D. (2007). The Proliferation and Expansion of Retinal Stem Cells Require Functional Pax6. *Developmental Biology*, *304*(2), 713–721. https://doi.org/10.1016/j.ydbio.2007.01.021

Yamanaka, S. (2009). A Fresh Look at iPS Cells. *Cell*, *137*(1), 13–17. https://doi.org/10.1016/j.cell.2009.03.034

Yasunaga, M., Tada, S., Torikai-Nishikawa, S., Nakano, Y., Okada, M., Jakt, L. M., Nishikawa, S., Chiba, T., Era, T., & Nishikawa, S.-I. (2005). Induction and monitoring of definitive and visceral

endoderm differentiation of mouse ES cells. *Nature Biotechnology*, *23*(12), Article 12. https://doi.org/10.1038/nbt1167

Yasuo, H., & Lemaire, P. (2001). Role of Goosecoid, Xnot and Wnt antagonists in the maintenance of the notochord genetic programme in Xenopus gastrulae. *Development*, *128*(19), 3783–3793. https://doi.org/10.1242/dev.128.19.3783

Zhang, H.-X., Zhang, Y., & Yin, H. (2019). Genome Editing with mRNA Encoding ZFN, TALEN, and Cas9. *Molecular Therapy*, *27*(4), 735–746. https://doi.org/10.1016/j.ymthe.2019.01.014

Zhang, X.-B. (2013). Cellular Reprogramming of Human Peripheral Blood Cells. *Genomics, Proteomics & Bioinformatics*, *11*(5), 264–274. https://doi.org/10.1016/j.gpb.2013.09.001

Zhou, W., & Freed, C. R. (2009). Adenoviral Gene Delivery Can Reprogram Human Fibroblasts to Induced Pluripotent Stem Cells. *Stem Cells*, *27*(11), 2667–2674. https://doi.org/10.1002/stem.201