DISSERTATION

Utilising S/MAR DNA vectors for the genetic modification of human induced pluripotent stem cells for cell and gene therapy

Manuela Urban, 2021

Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of Doctor of Natural Sciences

Presented by

M. Sc. Manuela Vanessa Urban

Born in Biberach an der Riss, Germany

Oral examination: 25.11.2021

Utilising S/MAR DNA vectors for

the genetic modification of

human induced pluripotent stem cells for cell

and gene therapy

Referees:

apl. Prof. Dr. Stefan Wiemann

Dr. Richard P. Harbottle

1   Summary

In the last years, our laboratory has developed and refined a novel platform of episomal self-sustaining S/MAR DNA vectors for gene and cell therapy. Their non-viral and non-integrative nature avoids integrational genotoxicity problems, a known issue with currently used gene therapy vectors. Previous work demonstrated their ability to provide persistent expression in cell lines and primary T-Cells. Further work successfully applied them to genetically modify mouse stem cells. Without altering their pluripotency capabilities, sustained maintenance and expression during reprogramming, differentiation and chimaera formation were shown. In this project, we aimed to extend our knowledge and understanding of these episomal vectors and developed and applied the technology towards human induced pluripotent stem cells (hiPSCs).

As a first step, we demonstrate the potential of our generated S/MAR DNA vectors in cancer cell lines and implemented vector establishment protocols with antibiotic selection or by purification of expressing cells via FACS sorting. We then provided proof of principle evidence that we can genetically modify hiPSCs using our S/MAR DNA vector system and demonstrated their isogeneity and unaltered capabilities to act as genetically modified cell source for gene and cell therapy. Besides implementing xeno-free hiPSC culturing which can easily be transferred to a GMP conforming protocol, we used cells isolated from urine as a novel, non-invasive cellular source for the generation of hiPSCs.

For the first time, we then moved from using the stable expression of GFP as an easily trackable reporter gene towards the restoration of a functional transgene in these cells. The potential of our S/MAR DNA vectors for use in advanced cell models and as prophylactic gene therapy for Birt-Hogg-Dubé syndrome (BHD) was then investigated. BHD patients harbour germline mutations in the gene for Folliculin (FLCN). After second hit mutations, functional FLCN is lost which leads to the development of kidney cancer. Currently, the only treatment available is the surgical resection of these tumours. However, this treatment does not restrict the development of further second hit

Summary

mutations and tumours. Previous work by our group and others indicates that the pathways in which FLCN is involved and the tumorigenesis can be avoided by introducing a functional copy of FLCN. Thus, one potential treatment for this disease would be the introduction of mutation-proof copies of Folliculin into cells before the tumorigenic event takes place.

Utilising CRISPR/Cas9 we generated FLCN-knock-out urinary derived hiPSC cell lines. By establishing GFP or FLCN encoding S/MAR DNA vectors in both WT and KO cells, we generated important cell models for the study of FLCN. Characterisation of the cell lines as well as single-cell RNA sequencing suggested little impact of the S/MAR DNA vector on the cells. Also, FLCN expression levels were shown to be not required for the exit of pluripotency of these hiPSCs, rendering them a developmentally earliest cell model to study FLCN expression in a variety of cell models in the future. We finally demonstrated their unaltered nature and persistent transgene expression by implementing a kidney organoid differentiation protocol as an advanced cell model to investigate BHD.

Together, our data demonstrate the versatile application possibilities of the combination of our S/MAR DNA vector platform in combination with urinary and fibroblast derived hiPSCs for functional pathway analysis, disease modelling, patient-specific drug screenings or future cell therapies with optimised, non-viral gene therapy vectors.

2   Zusammenfassung

In den letzten Jahren hat unsere Arbeitsgruppe eine neuartige, sich selbst und episomal erhaltende S/MAR DNS Vektor Plattform für die Gen- und Zelltherapie weiterentwickelt. Häufig in der Gentherapie verwendete virale Vektoren bringen das Risiko von Genotoxizität durch Integration der Erbinformation mit sich. Durch den Ausschluss von viralen Elementen und das nicht-integrierende Verhalten kann man dies mit unsereren S/MAR DNS Vektoren vermeiden. Wir konnten zuvor bereits zeigen, dass unsere Vektoren in der Lage sind anhaltende Transgenexpression in Zelllinien und primäre T-Zellen zu liefern. Des Weiteren wurden sie erfolgreich für die genetische Veränderung von Maus Stammzellen verwendet. Ohne dabei das Pluripotenzpotential zu verändern konnte dauerhafter S/MAR DNS Vektor Erhalt und Expression sowohl während der Reprogrammierung, der Differenzierung und der Generierung von chimaeren Mäusen gezeigt werden. Die hier vorgelegte Arbeit baut auf diesem Wissen und der Erfahrung mit dem episomalen Vektorensystem auf und erweitert das Anwendungsgebiet zum ersten Mal auf humane induzierte pluripotente Stammzellen (hiPSCs). Die Zugabe von nicht mutierenden Kopien von FLCN in die Zellen, bevor die Tumorentwicklung stattfindet, stellt daher eine mögliche Behandlung dieser Krankheit dar.

Zunächst wurden die generierten S/MAR DNS Vektoren in Krebszelllinien getestet, verifiziert und verschiedene Etablierungsprotokolle mit und ohne aktiver Antibiotikaselektion implementiert. Als nächster Schritt wurde gezeigt, dass unsere S/MAR DNS Vektoren grundsätzlich in der Lage sind, hiPSCs zu verändern und dabei die für eine Gen- und Zelltherapie wichtigen Eigenschaften wie Pluripotenz und Differenzierungspotential erhalten bleiben. Neben dem Etablieren von Xeno-freien Kultivierungsmethoden die leicht in GMP konforme Protokolle überführt werden können konnten wir außerdem das Potential von Zellen die aus dem Urin isoliert wurden für die nicht-invasive, einfache Quelle für hiPSCs darlegen.

Neben GFP als leicht zu detektierendes Reportertransgen haben wir zum ersten Mal auch die Expression eines funktionalen Transgens angewendet. Dazu haben wir uns auf

Zusammenfassung

eine mögliche Anwendung unserer S/MAR DNS Vektoren Plattform für fortschrittliche Zellmodelle oder eine prophylaktische Gentherapie für das Birt-Hogg-Dubé Syndrom (BHD) konzentriert. Patienten, die an BHD leiden weisen eine Keimbahnmutation im Gen für Folliculin (FLCN) auf. Durch eine second hit Mutation auf Grund dessen funktionelles FLCN verloren geht wird dann im Laufe des Lebens die Entwicklung von Nierenkrebs ausgelöst. Momentan besteht die einzige Therapiemöglichkeit darin, den Tumor chirurgisch zu entfernen wenn er zu groß wird. Dies verhindert allerdings keine weitere second hit Mutation in einer anderen Nierenzelle und die Entstehung eines neuen Tumors. Bisher veröffentlichte Arbeiten und Forschung von unserer Arbeitsgruppe deutet darauf hin, dass diese Krebsentwicklung durch eine funktionale Kopie von FLCN verhindert werden kann.

Mittels CRISPR/Cas9 wurden in dieser Arbeit FLCN-knock-out hiPSC Linien von Urin als Ausgangsquelle hergestellt. Sowohl WT als auch KO Zelllinien wurden dann stabil genetisch mittels GFP oder FLCN-S/MAR DNS Vektor verändert um so eine gesamte Palette an wichtigen Zellmodellen für die Untersuchung von FLCN zu erhalten. Die Charakterisierung dieser Zellen zeigte erneut den minimalen Einfluss unserer Vektoren auf die Wirtszelle. Zusätzlich konnte gezeigt werden, dass die fehlende oder starke Expression von FLCN keinen Einfluss auf den Austritt aus der Pluripotenz der hiPSCs hat. Damit können die hier generierten Zelllinien in der Zukunft als entwicklungsbiologisch frühes Modelsystem für BHD oder für FLCN relevante Signalwege für die verschiedensten Zellmodele verwendet werden. In dieser Arbeit wurde zum Schluss diese Fähigkeit der Zellen durch das Etablieren eines komplexen Differenzierungsprotokols zu Nierenorganoiden gezeigt.

Schlussendlich veranschaulicht diese Arbeit die vielseitigen Anwendungsmöglichkeiten unserer S/MAR DNS Vektor Plattforms in Kombination mit hiPSCs die aus Fibroblasten oder Zellen aus dem Urin generiert wurden: von funktionellen Signalwegestudien, über die Modellierung von Krankheiten, Patientenspezifischen Medikamentenscreenings bis hin zu zukünftigen Zelltherapien mit optimierten, nicht-viralen Gentherapievektoren.

X

3   Abbreviations

°C	degree Celsius
µg	Microgram
µl	Microliter
~	approximately
ActA	Activin A
AP	alkaline phosphatase
BAP	bacterial alkaline phosphatase
BCA	bicinchoninic acid
BHD	Birt-Hogg-Dubé syndrome
bp	base pairs
BSA	bovine serum albumin
CAG	CMV early enhancer and chicken β-actin promoter
cDNA	complementary DNA
CHIR	CHIR-99021 – GSK-3 inhibitor
CHM	Choroideremia
CMAR	core S/MAR sequence, identified and isolated by Dr Matthias Bozza
c-Myc	V-Myc Avian Myelocytomatosis Viral Oncogene Homolog – reprogramming factor
coGFP	variant of the reporter GFP (green fluorescent protein) (Ex=488nm, Em=508nm) – GFP variant isolated from the pmaxGFP plasmid from Lonza, which is a variant of CopGFP from copepoda Pontellina plumata
CRISPR	clustered regularly interspaced short palindromic repeats
Da	Dalton
DMEM	Dulbecco's modified eagle's medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
do	donkey
dpt	days post transfection
ds	double stranded

Abbreviations

dTom	dTomato, red reporter gene (Ex=554nm, Em=581nm)
EBs	embryoid bodies
EBNA1	Epstein-Barr Nuclear Antigen 1
ECAD	E-cadherin - distal tubuli marker
EF1α	Elongation factor 1α promoter
eff.	efficiency
e.g.	for example
Ele40	Insulating Element 40
EtOH	ethanol
FACS	fluorescence activated cell sorting
FBS/FCS	fetal bovine/calf serum
FiPSCs	Fibroblast derived iPSCs
FLCN	Folliculin
FLCN-flag	Folliculin with an artificial flag-tag at the N-terminus
FM	freezing media
FSC -A/ -W	forward scatter - area/ - width
g	gram
GAPDH	Glyceraldehyde 3-Phosphate Dehydrogenase – housekeeping gene
gDNA	genomic DNA
GFP	green fluorescence protein – here we predominantly refer to the variant coGFP derived from the Lonza pmaxGFP plasmid
GFP+	GFP positive/GFP expressing
GFP-	GFP negative/not GFP expressing
gt	goat
Gy	Gray
h	unit: hour(s)
h	species: human
H_2O	water
hiPSCs	human induced pluripotent stem cells
IF	Immunofluorescence
IM	intermediate mesoderm
iMEF	immortalized mouse embryonic fibroblast

iPSCs	induced pluripotent stem cells
IRES	Internal Ribosome Binding Site
kb	kilo base pairs
kDa	Kilodalton
kg	Kilogram
Klf4	Krueppel-like factor 4 – EBNA reprogramming factor
KO	knockout
KSR	knockout serum replacement
L	Litre
LB	lysogeny broth
Lin28	RNA-binding protein Lin28 – EBNA reprogramming factor and human pluripotency marker
L-Myc	Viral Oncogene Lung Carcinoma Derived Homolog RNA – EBNA reprogramming factor
M	molar (moles per litre)
max.	maximal/maximum
MFI	mean fluorescence intensity
mg	Milligramm
min	minute(s)
min.	minimal
ml	Milliliter
mM	Millimolar
MQ water	Milli-Q purified water
mRNA	messenger RNA
ms	mouse
NaCl	sodium chloride
Nanog	Gene belonging to the Nanog homeobox Family – human pluripotency marker
NEAA	non-essential amino acids
ng	Nanogram
NHDF	normal human dermal fibroblasts
NPC	nephron progenitor cell

Abbreviations

nS/MAR	S/MAR Nanovectors™
NTC	Nature Technology Corporation
Oct3/4	Octamer-binding transcription factor 4 (also known as POU5F1) – EBNA reprogramming factor and human pluripotency marker
Oct4A	Isoform of Oct3/4 that is linked to human pluripotency – used for qPCR primer design for pluripotency marker expression
p2A	2A self-cleavage peptide sequence
p/s	penicillin/streptomycin
Pax2	Paired box 2 - IM marker
PBS	phosphate buffered saline
PBST	PBS with tween20
PCR	polymerase chain reaction
pDNA	plasmid DNA
PEI	polyethylenimine
PFA	paraformaldehyde
PI	propidium iodide
PODXL	podocalyxin
PPS	posterior primitive streak
pS/MAR	plasmid S/MAR DNA vector
puro	puromycin
puroR	puromycin resistance gene
qPCR	quantitative polymerase chain reaction
rb	rabbit
RCA	rolling circle amplification
RNA	ribonucleic acid
RNAseq	RNA sequencing
ROCK	Rho-associated, coiled-coil containing protein kinase
ROCKi	ROCK inhibitor
rpm	rounds per minute
RT	room temperature
RV	renal vesicles
SCs	stem cells

sec	second(s)
SFFV	Spleen Focus-Forming Virus promoter
S/MAR	Scaffold/Matrix Attachment Region
SOC media	super optimal broth
Sox2	SRY (Sex-determining region Y)-box 2 – EBNA reprogramming factor
ss	single stranded
SSC -A -W	sideward scatter - area/ - width
TBXT	Brachyury / T-Box transctiption factor T - PPS marker
Tm	melting temperature
TRA-1-60	Podocalyxin – human pluripotency marker
USCs	urinary stem cells – cells isolated from the urine
UiPSCs	urinary cell derived iPSCs, USCs as cell source
ULA	ultra-low attachment
UV	ultraviolet
V	volt
v	Volume
Vil1	Villin – proximal tubuli marker
vs	versus
w	weight
WB	Western Blot
wp	well plate
WT	wildtype
WT1	Wilms tumor 1 - glomeruli marker
w/v	weight/volume
xg	times G force
noRT	no reverse transcriptase
°	degree

4   Contents

5 INTRODUCTION

5.1 Overview

Gene and cell therapy are among the most promising therapeutic approaches of the last decades and the first approved gene and cell therapy products are reaching patients (Dunbar et al., 2018). To achieve this, a huge amount of research on suitable gene vector technologies was performed. While viral vectors have been intensively used due to their high efficiency, novel genetic technologies have evolved providing clear advantages such as avoiding integrational genotoxicity and providing easier purification and production to a GMP level. Our group has worked on improving such a non-integrating S/MAR DNA vector platform for gene therapy. We have recently shown that we can apply them for effectively engineering cells for CAR T-Cell immunotherapy (Bozza et al., 2021), and persistently modify mouse stem cells with a GFP expressing S/MAR DNA vector while maintaining transgene expression in stem cells as well as in their differentiated progeny (Roig Merino & Urban et al., in press).

The discovery of induced pluripotent stem cells (iPSCs) opened a new promising approach for patient-specific cell therapy. However, by nature, they are refractory to genetic modification. Typically, they are reprogrammed from dermal fibroblasts or blood cells. It has recently been shown that cells derived from the urine can also be effectively used as a non-invasive cell source for iPSCs (Drozd et al., 2015). In combination with novel gene editing tools like CRISPR/Cas9 (Song & Ramakrishna, 2018) as well as advanced differentiation technologies such as kidney organoids (O'Neill & Ricardo, 2013) further enable new sophisticated and personalised disease modelling systems.

In this work, we build on the experience in mouse stem cells and genetically modify human iPSCs with our optimised S/MAR DNA vector platform thereby providing evidence for their joint application for gene and cell therapy.
Next to proof of principle experiments with genetic modification with GFP we further apply our system to a disease using a functional transgene. Here, we focus on the

INTRODUCTION

monogenic Birt-Hogg-Dubé Syndrome (BHD), where the loss of folliculin (FLCN) leads to the development of kidney cancer. We aim to provide evidence that our DNA vector system can be applied as preventive gene therapy in this setting, as well as generate state-of-the-art disease model systems to facilitate a better understanding of the disease.

5.2 Gene and cell therapy

When scientists started to decipher and understand the genetic code, they also soon made the connection that many genetically inherited diseases might not only be treated but rather cured when providing functional copies of the mutated gene. In particular monogenic disorders, where researchers have pinpointed the disorder to a single gene, are interesting targets for gene therapy. Genetic diseases can be either caused by gain of function (GOF) or loss of function (LOF) mutations. In GOF mutations, the mutation leads to a change where the gene product gains a new function in the body which ultimately causes the disease. In contrast, a LOF mutation leads to a non-functional protein, which thus cannot fulfil its supposed role in the body and this missing function causes the disease (B. Chen & Altman, 2017). Gene therapy is most commonly targeted to the latter, where it aims to deliver a functional copy of a disease-causing gene, thereby restoring its function and reversing the disease phenotype. Notably, conventional drug discovery strategies don't work well for these LOF diseases as it is difficult to restore an absence of protein by pharmacological means (Segalat 2007). While bioinformatic attempts are often made to screen promising candidates (Chen 2017), gene therapy might be the most promising therapy option for patients with monogenic diseases.

Cell therapy research is closely related to gene therapy. Here, cells are either used directly as therapeutic agent or to replenish missing cells, or, in combination with gene therapy, they are genetically modified ex vivo and subsequently re-implanted. A prominent example outside traditional gene therapy for genetic diseases is cellular cancer immunotherapy. For chimeric antigen receptor (CAR) T-cell therapies, for example, T-cells from patients are isolated, genetically modified with a CAR that recognises the cancer cells and then reintroduced into the patient (Dunbar et al., 2018). When in 2017 the first CAR-T therapy Kymriah by Novartis gained FDA approval it was considered a historic moment following decades of research and finally obtaining clinical approval and drug designation and many hopes are now placed on the future benefits of these therapies ("First-Ever CAR T-cell Therapy Approved in U.S," 2017).

INTRODUCTION

5.3 Genetic modification of cells

To achieve the correction of genetic information in cells for gene therapy purposes, different methods and tools were developed which are briefly summarised below.

5.3.1 Gene editing

In cases where mutated genetic information results in proteins which cause the disease, rather than the absence of the functional gene, the simple supplementation of corrected genes is not sufficient. Here, the genetic information in the patient cells needs to be modified and corrected. In the last decades, new and highly promising technologies have emerged which make advances over TALEN or zinc finger nucleases for gene editing. The principal technology **CRISPR/Cas9** could enable treatment options not only for LOF mutations. Briefly, this editing technology enables precise changes in DNA by the combination of a guide-RNA that binds and labels the target site and the Cas9 nuclease that cleaves the DNA at the respective site (Cong et al., 2013). The importance and wide expectations for the future of this method was emphasised by the Nobel prize in 2020 for the researchers Emmanuelle Charpentier and Jennifer Doudna, only 8 years after first reports of its possible application to modify targeted DNA (Gasiunas & Siksnys, 2013). However, there is still lots of essential research to be done to avoid off-target modification and optimise the delivery and production. Besides gene editing, this method also enables the simple removal of a whole gene. Reintroduction of a target sequence into a specific site however is by default less efficient since it relies on further repair systems in addition to the nuclease activity of Cas9 (Kebriaei et al., 2017).

Today this tool already plays a key role in *in vitro* and *in vivo* disease modelling and studies of cellular mechanisms. Following the claimed first application of CRISPR/Cas9 in human embryos (which was performed without ethical approval in 2018), the responsible researchers in China were sentenced to serve time in prison. The way this first CRISPR/Cas9 therapy on human embryos took place and the harsh criticism by many scientists and media all over the world might have left a remaining dark shadow over future promising therapeutic applications with this tool and illustrates the big

barriers that still need to be overcome beyond the pure scientific feasibility (Cyranoski, 2020).

5.3.2   Gene supplementation technologies

Besides the alteration of the genetic code, suitable gene transfer vectors targeting LOF mutations and aiming to supplement the missing functional protein have been researched for decades. For this work, we predominantly focus on gene supplementation technologies and a novel vector system in this context.

5.3.2.1   Viral vectors

Initially, due to their high efficiency and innate ability to deliver genetic information into mammalian cells, researchers focused on developing viral delivery vectors for gene therapy. The first vectors used in gene transfer trials were based on **retroviruses** (Dunbar et al., 2018; Miller, 1992). They can act simultaneously as delivery system and vector of the genetic cargo. By utilising only genes encoding packaging information without other viral proteins, these modified vectors delivered genetic information with high efficiency and integration into dividing cells (Miller, 1992). However, they demonstrate drawbacks such as limitations of target cells to dividing cells and problems with their purity after production in cells (Miller, 1992).

Most importantly, the virus itself can cause immunity and the genetic information can be damaging or silenced when integrated or toxic to host cells in other ways (Dunbar et al., 2018). The risk of damaging important parts of the genome by integration into the wrong location is a major drawback of any integrating gene delivery system. The observation of cancer development after successful gene therapy with retroviral modified HSCs for SCID-X1 in clinical trials revealed the genuine threat of this risk (Hacein-Bey-Abina et al., 2003). Current gene therapy vectors applications are still mainly used for life-threatening diseases due to risks that result from the utilised vectors (Seow 2009).

INTRODUCTION

By moving from γ-Retroviruses to Lenti-viral vectors researchers made it possible to target non-dividing cells (Dunbar et al., 2018). Lentiviruses are known to integrate into coding regions, which renders their risk by integration at least in hematopoietic stem cells (HSCs) lower than γ-Retroviruses which integrate into to 5' untranslated regions (Larochelle & Dunbar, 2013).

Adenoviruses are non-integrating vectors, that can very efficiently transduce also non-replicating cells and be produced in high titres, have a high packaging capacity and were thought to become a key gene therapy vector (Miller, 1992). Unfortunately, they induce a strong immune response and partial immunity due to common infection with the Adenovirus AdV5 is a significant drawback (Seow & Wood, 2009). While this rendered Adenoviruses unsuitable for gene therapy, they were subsequently investigated for their oncolytic ability for cancer treatment and their use as vaccinations (Wold & Toth, 2013).

It is noteworthy, that following more than 30 years of research in gene therapy and particularly the work on Adenoviruses and their triggered inflammatory responses provided the foundation for the speedy development of some SARS-CoV-2 vaccinations such as the vaccine by Johnson & Johnson or Vaxzevria by Astrazeneca to counter the 2020 coronavirus pandemic (Kremer, 2020). The application of these gene therapy vectors in the broad population also raised concerned voices of future immunity against developed gene therapies that are generated with the same vectors (Aledo-Serrano et al., 2021).

Adeno-associated viruses (AAVs) represent the next step taken towards a non-integrative viral vector. While limiting the packageable DNA size to 5kb, these vectors remain predominantly episomal (Dunbar et al., 2018; Kebriaei et al., 2017). AAVs require other viruses such as Adenoviruses to replicate and are less immunogenic than adenoviruses which provides a route towards a more tolerable therapy (Chan et al., 2021; Dunbar et al., 2018)

In general, production and limitations in packaging size present still difficulties to overcome before these vectors can be used in large scale and for broad applications (Seow & Wood, 2009). Also, remaining occurrences of integration-mediated tumorigenicity is closely monitored during trials and remains a major concern also for the use of AAV vectors. It further has been reported that not only the nucleus has to be monitored, but also integration hotspots in mitochondria have been observed (Kaeppel et al., 2013).

5.3.2.2 Non-viral vectors

Despite the continuous improvement of viral vectors, further possibilities of viral-free gene delivery platforms are increasingly being investigated to avoid virus-induced immunogenicity. Besides bacteriophages, biological or synthetic liposomes, or exosomes (Seow & Wood, 2009), **transposons** present a commonly used and intensively researched tool (Hudecek & Ivics, 2018). The importance of transposons was also emphasised by the award of a Nobel prize for their discovery. These elements are DNA sequences that can jump to different locations in the genome and lead to the development of the sleeping beauty vector system. Briefly, the target sequence gets flanked with a transposon specific element in a donor plasmid which is then cut and integrated into the genome (Kebriaei et al., 2017). Although being integrative and harbouring the risk of integrational genotoxicity, the system has been intensively researched and improved in regard to integration site and is currently used in several clinical trials (Kebriaei et al., 2017). Despite the high reliability of the integration site, the integration and consequential risk for genotoxicity remains a major drawback of this method.

Researchers have therefore continued to investigate a simple, non-viral and non-integrative vector utilising the raw DNA in **plasmid** form. Plasmids are easy to purify and can be produced in large amounts with high purity, but by default, they usually only provide transient expression. For gene therapy, however, stable expression of the functional gene is usually required. Therefore, features of viruses that allow episomal replication within the host such as SV40 or the EBV Nuclear Antigen 1 (EBNA1) from

INTRODUCTION

the Epstein-Barr virus are utilised in plasmid DNA sequences (Glover et al., 2005). For EBNA vectors it was reported that the episomes were lost over time and continuous antibiotic selection led to integration opposing the aimed advantages of these vectors over viral ones (Glover et al., 2005).

To further improve plasmid vectors, **minicircles** were developed (Bigger et al., 2001; Darquet et al., 1997). Here, the bacterial backbone containing for example a selection marker for antibiotic resistance and a bacterial origin of replication is removed usually by recombination sites (Stenler et al., 2014). This does not only benefit the vector performance, but also clinical regulatory agencies recommend to avoid such sequences in vectors to avoid immune responses against the vector, unmethylated CpG regions or inadvertently expressed bacterial proteins (Mayrhofer et al., 2009; Stenler et al., 2014; Viecelli et al., 2014). These optimisations therefore led to safer vectors that were also reported to provide more sustainable and better transgene expression levels (Darquet et al., 1997; Mayrhofer et al., 2009).

Minicircle production relies on recombinational removal of the bacterial sequences, and these sequences subsequently need to be purified from the plasmids. A novel method to generate minimally sized bacterial backbone plasmid DNA was generated and patented by Nature Technology Corporation (NTC). These **Nanovectors™** are produced antibiotic-free, with a sucrose dependent selection system that has been reported to be less immunogenic (J. Luke et al., 2009; J. M. Luke et al., 2014). Briefly, these vectors encode a small-sized mRNA that complements the mRNA for the levansucrase protein that is encoded by the bacterial genome. Bacteria without the vector produces levansucrase that then converts sucrose from the media into a toxic molecule that leads to the death of the bacteria. Upon maintenance of the plasmid, the binding of the RNA-OUT sequence inhibits levansucrase production and thus the formation of this toxic molecule and the bacteria does not die (J. M. Luke et al., 2014).

5.4 S/MAR DNA vectors – an optimised non-integrating vector platform for gene therapy

To further improve episomal replication and persistence while completely avoiding viral sequences, S/MAR DNA vectors were developed. S/MAR stands for scaffold/matrix attachment region and comprises AT rich regions that occur in the human genome and play an important role in compartmenting the genome into loop forming domains such as actively transcribed and untranscribed parts as well as functional domains (Baiker et al., 2000; Mirkovitch et al., 1984). While the exact mechanism is not yet known, S/MAR regions are thought to localize with transcription factories, at certain chromatin areas and have a non-dynamic behaviour during mitosis (Hagedorn, Gogol-Doring, et al., 2017; Jenke et al., 2004).

DNA vectors containing S/MAR elements were shown to remain episomal and could therefore circumvent the risk of insertional mutagenesis. The S/MAR motif can act as origin of replication enabling persistent maintenance throughout unlimited cell divisions (Hagedorn et al., 2011; Hashizume & Shimizu, 2007). Also, a regulatory effect on expression and the prevention of silencing similar to insulator elements were reported for some S/MAR motifs (Arope et al., 2013; Sun et al., 2016).

The first published DNA plasmid vector that contained a human β-interferon derived S/MAR sequence and was thus devoid of viral sequences was pEPI (Piechaczek et al., 1999). It was shown that this vector can persist episomally without selection pressure for several hundred divisions in cell lines (Glover et al., 2005). Even the successful generation of transgenic pig embryos via sperm-mediated gene transfer was reported with this vector (Manzini et al., 2006).

To enable the delivery of these non-viral S/MAR DNA vectors, a variety of chemical and mechanical methods such as lipofection, nucleofection or electroporation is typically used. Some interesting research was also done by combining S/MAR DNA vectors with AAVs: The promise of efficient delivery of AAVs and the stable expression and episomal maintenance of S/MARs combined while circumventing delivery issues of

INTRODUCTION

S/MAR DNA vectors and unstable, integrational maintenance of AAVs is appealing but the application was limited and more inefficient than anticipated (Hagedorn, Schnodt-Fuchs, et al., 2017).

In the last years, our research group has worked extensively on optimising these S/MAR containing plasmid vectors. Based on the original pEPI vector, the S/MAR sequence itself, but also each element of the vector was scrutinised and challenged for improvement (Figure 1 A, B) (Bozza, 2017; Roig Merino, 2018). An overview of these optimisation steps performed by Dr Matthias Bozza is shown in Figure 1.

In pEPI, the antibiotic resistance gene resided in the bacterial part of the vector. We observed breakage of the vector upon antibiotic selection in southern blots. By using the puromycin resistance gene in the mammalian expression cassette, we aimed to avoid vector breakage, while simultaneously forcing the expression of the transgene and preventing the over expression of the bacterial genes which could cause toxicity in mammalian cells (Figure 1 C, referred to as pCAG in this thesis). In general, all unnecessary DNA elements which could induce immune reactions or silencing events were removed. In collaboration with NTC, instead of minicircles (Figure 1 B) their minimally sized bacterial backbone plasmid DNA was used for some vectors (Nanovectors™, Figure 1 D, F; vector D is referred to as nCAG in this thesis). To further stabilize expression and inhibit the spread of silencing events, the insulator Element40 (Ele40) was added between the bacterial backbone and the promoter region (Figure 1 E). Additionally, suitable promoters were tested within the different arrangements for different applications (Figure 1).

During the course of this work, further optimisations were performed such as testing smaller S/MAR motifs or adding splicing sites to align the plasmid mRNA with endogenous mRNA splicing events (Figure 1F, G).
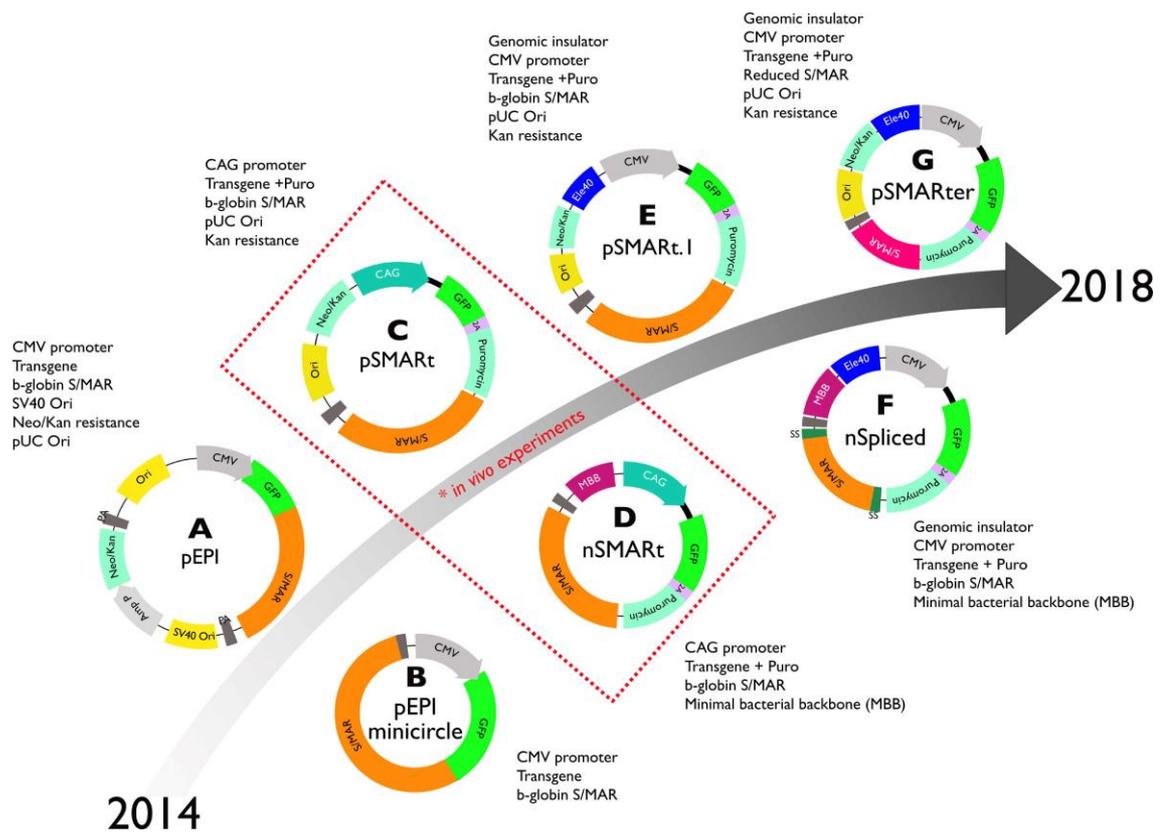
Figure 1: Overview of S/MAR DNA vector development. Reprinted with permission from the dissertation of Dr Alicia Roig Merino (Roig Merino, 2018).
The development of vector feature optimisation is shown schematically. The important changes are further described in the text in chapter 5.4. The vectors used in this study are based on the depicted vectors C and D which are subsequently referred to as pCAG and nCAG, respectively.

With this S/MAR DNA vector platform we now have an optimised gene therapy vector on hands that overcomes the aforementioned limitations of the viral and non-viral gene therapy vectors. The main advantages of this system can be summarized as follows:

S/MAR DNA vectors…

- Are non-integrative: no risk of integration induced genotoxicity

- Provide persistent expression due to episomal maintained with S/MAR as an origin of replication

- Are fast and easy to produce and purify

- Present no size limitation towards the target gene

- have low impact on host cell

- trigger minimised immune reaction due to optimised sequences and absence of viral sequences (and minimalised bacterial backbone in Nanovectors™)

- can be easily adjusted and optimised to achieve specialised vectors for the respective target cell

5.5   Previous applications of our S/MAR DNA vectors

Our optimised S/MAR DNA platform has been extensively refined and evaluated over the last years. We have shown repeatedly that our system outperforms the previously described pEPI S/MAR DNA vector in a variety of primary cells and established lines (Bozza et al., 2020).

Initially, their functionality was proven for a variety of cancer cell lines. Cells were successfully stably modified while persisting in low copy numbers and remaining episomal. In this context, the vectors enable the generation of isogenic cell lines where the vector itself causes little impact on the cells (Bozza et al., 2020). Therefore, investigation of signalling pathways or cellular behaviour is predominantly caused by the added transgene, and not changes due to the vector or extensive drug selection. Predominantly, Lentivirus is used to generate stable cell lines for cell models. Our S/MAR DNA vector system can be generated with competent bacteria and routine plasmid preparation methods, therefore avoiding the need for an S2 laboratory setting and laborious virus production and quantification.

Recently, the vectors were utilised to implement a fast and safe generation of persistently modified T-Cells for CAR T-Cell therapy (Bozza et al., 2021). Here, single-cell RNAseq analysis of the engineered T-Cells also shows that the S/MAR DNA vectors trigger less changes in the host cell than the usually used, state of the art Lentiviral vectors.

Another challenging use of our system has been its application in stem cells. These experiments are further elaborated in chapter 5.10, after introducing the principles of stem cells in the following chapters.

5.6   Stem cells

Stem cells represent a special class of primary cells, whose discovery has provided the potential for a novel, personalised or off-the-shelf, cell-based therapy platform for many diseases (Hombach-Klonisch et al., 2008; Zakrzewski et al., 2019). All stem cells have in common their ability to self-renew and to further differentiate into more specialised cells. Based on their capability to differentiate into a broad amount or only very few specialised cells, they are categorised into omni-, pluri- or multipotent cells (Fus-Kujawa et al., 2021; Zakrzewski et al., 2019).

Embryonic stem cells derived from blastocysts represent the very first stem cells, that are capable of renewing themselves and further differentiating until they form a whole body with all its specialised cells in it (Thomson et al., 1998). They therefore represent the single kind of omnipotent stem cells we know as of today.

Pluripotent stem cells need further cells to form a whole body, but they are capable to differentiate into cells from all three germ layers, Endoderm, Ectoderm and Mesoderm. While these omni and pluripotent cells seem to exist only in early development, our adult body still harbours a variety of further specialised, adult stem cells (Wagers & Weissman, 2004; Zakrzewski et al., 2019).

These multipotent cells reside in a specific tissue and are capable of generating and replenishing cells only from that origin (A. I. Caplan, 2010; Arnold I. Caplan, 2015; He et al., 2007). Prominent examples are the intestinal stem cells that reside in the crypts of the intestine, and which continuously provide replacement for dying cells. The recent discovery of kidney originated progenitor cells that can be isolated from the urine are further elaborated in the following chapter 5.9. Another therapeutically important example for adult stem cells are hematopoietic stem cells (HSCs), which are multipotent cells in the blood system that enable for example the regeneration of our blood cells upon blood loss. The isolation of HSCs has spread hope that they could be used for cell therapy. In a difficult process, they can be activated with specific drugs,

causing their relocation from the bone marrow into the bloodstream, and subsequently harvested from the blood.

In the early 1990s, not just HSCs, but also "mesenchymal stem cells" were successfully isolated. Mesenchymal stem or stromal cells (MSCs) are multipotent cells that can differentiate into cells like osteoblasts, chondrocytes, or adipocytes (Arnold I. Caplan, 2015). These cells are heavily researched for their potential to regenerate whole tissues. They can be purified from adipose tissue or the umbilical cord after birth which enables personalised as well as general cell banking of these cells.

All these adult stem cell populations present very heterologous, not distinctively characterised populations. Attempts to better standardise the progenitor population with specific markers such as expression of CD73, CD90, CD105 and absence of CD11b, CD14, CD19, CD34, CD45, CD79a for HSCs are continuously improved for example by the international society for cellular therapy (ISCT) (Mindaye et al., 2013; Weiss et al., 2011). Current work also focuses on the possible synergistic effect of the combinational use of MSCs and HSCs (Battiwalla & Hematti, 2009).

5.7　Induced pluripotent stem cells

It was always believed that cellular differentiation is a one-way street, and there was no turn back or sideways. However, in 1962 John Gurdon showed that the nucleus of an adult cell can provide the whole genetic information to form a full organism, after being transferred into a fertilized, enucleated egg cell (Gurdon, 1962).

Building on this evidence, that the genetic information of also earlier stages can be made accessible in adult cells, Yamanaka and others showed in 2006 how they successfully generated induced pluripotent stem cells (iPSCs) from mouse fibroblasts (Takahashi & Yamanaka, 2006). By delivering only a combination of four transcription factors, namely Oct3/4, Sox2, Klf4 and c/Myc, they were able to reprogram the cell fate. They turned highly specialised, differentiated fibroblasts towards pluripotent cells that were able to subsequently again differentiate into any cell of the three germ layers, not limited to fibroblasts.

The experiments were extended further with human fibroblasts and confirmed the universal ability of these transcription factors to reprogram cells from different species (Takahashi et al., 2007). These four initial reprogramming factors are commonly known as Yamanaka factors. The impact of this discovery and its potential for autologous cell therapy was further acknowledged by awarding the Nobel prize in medicine in 2012 to John Gurdon and Shinya Yamanaka.

The technology of reprogramming enabled a whole new area of possible applications. The relatively straightforward, generation of pluripotent cells, which then can be turned into any needed cell type has huge potential for autologous patient-derived cell therapy, tissue repair or even whole organ transplantation in the future. Here, patient-specific cells can circumvent rejection reactions like graft-versus-host disease (GVHD), a common problem in cell transplantation (Zakrzewski et al., 2019). The unlimited growth capacity offers the opportunity for off-the-shelf cell therapy generation. While there are numerous protocols to differentiate iPSCs into progenitor cells, there is still a long way of research ahead to obtain reliable tissue or organ differentiation protocols.

Also, thorough purification protocols need to be established since a single remaining iPSC can lead to teratoma formation. In the meantime, besides the usage of the cells as direct therapy, they are used for patient-specific drug screenings as well as the generation of disease models to study the pathomechanisms behind the disease. (Fus-Kujawa et al., 2021; Papp & Plath, 2011; Peng et al., 2019).

INTRODUCTION

5.8 reprogramming delivery systems

5.8.1 Lentivirus 4-in-1

The reported Lentivirus 4-in-1 system combines the efficient infection and gene delivery with an optimised cassette of reprogramming factors in one expression cassette (Warlich et al., 2011). The expression cassette encodes the Yamanaka factors Oct4, Klf4, Sox2 and c-Myc in a human, codon optimised sequence (Figure 2). They are connected via 2a sites and are followed by IRES and dTomato as reporter gene. Another notable feature is the SFFV-promoter, which is known to be not active in pluripotent cells (Herbst et al., 2012). Therefore, the dTomato expression enables visualisation of successful transduction as well as easy identification of successfully fully reprogrammed stem cell colonies.
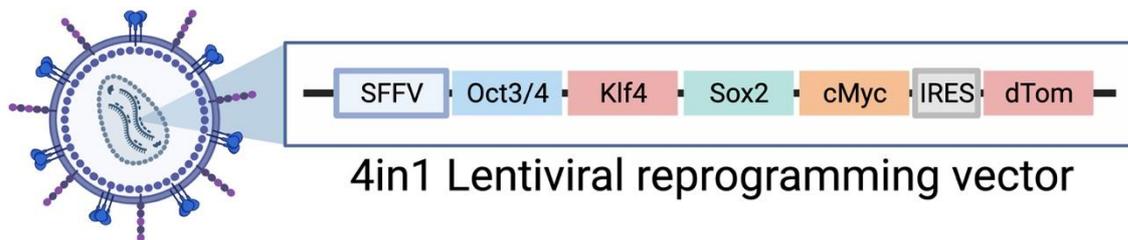
Figure 2: Schematic of Lentiviral 4-in-1 reprogramming vector.
The expression cassette of the Lentiviral 4-in-1 reprogramming vector as described in (Warlich et al., 2011) is shown and further described in the main text of chapter 5.8.1.

While presenting a useful and efficient reprogramming tool, as introduced before, Lentiviral vectors integrate the delivered genetic information into the genome (chapter 5.3.2.1). By doing so, the obtained iPSC clones are heterogeneous due to different integration sites, and the integration could impair the function or differentiation capacity of the cells and integration mediated genotoxicity, which renders them not suitable for the therapeutic generation of iPSCs (Shi & Cheung, 2021).

5.8.2 Sendai virus

The non-integrating Sendai virus is an RNA virus that delivers the genetic information only to the cytoplasm. While the only transiently achieved transgene expression renders Sendai viruses not applicable for gene therapy, these features fit the needs for reprogramming (Nishimura et al., 2011). Sendai reprogramming virus encoding the Yamanaka factors can be obtained commercially and with xeno-free, GMP compliant production (e.g. thermo fisher scientific). While several optimisations have been

performed, traditionally still the Yamanaka factors are utilised which contain the oncogene c-Myc.

5.8.3   EBNA reprogramming

As for gene therapy vectors, in light of future cell therapy applications of hiPSCs researchers also thrive to use safe, non-altering methods for reprogramming. As a step towards this, the episomal EBNA reprogramming system got established and is commonly used and commercially available (Drozd et al., 2015; Mack et al., 2011; Okita et al., 2011). Here, four different vectors with the EBNA backbone as depicted in Figure 3 must be delivered simultaneously to the cells. Three of these vectors contain the reprogramming factors Sox2, Klf4, Lin28, L-myc and Oct4. All vectors encode additionally for the EBNA-1 protein, while the fourth vector encodes solely for it. EBNA-1 is needed to drive replication and maintenance of all the EBNA vectors. All genes are regulated by a CAG promoter. Additional shRNA for p53 is added to enhance reprogramming efficiency (Zhao et al., 2008). It was reported, that over time EBNA vectors get lost at a rate of 5% per cell cycle (Nanbo et al., 2007). However, it is good practice to screen the obtained stem cell clones for being negative for the reprogramming vectors (Zakrzewski et al., 2019). While the system replaced the oncogene c-Myc with the isoform L-Myc, the presence of the oncoprotein EBNA-1 that is based on a viral feature demonstrate the needed research for further viral-free, non-integrative reprogramming tools.
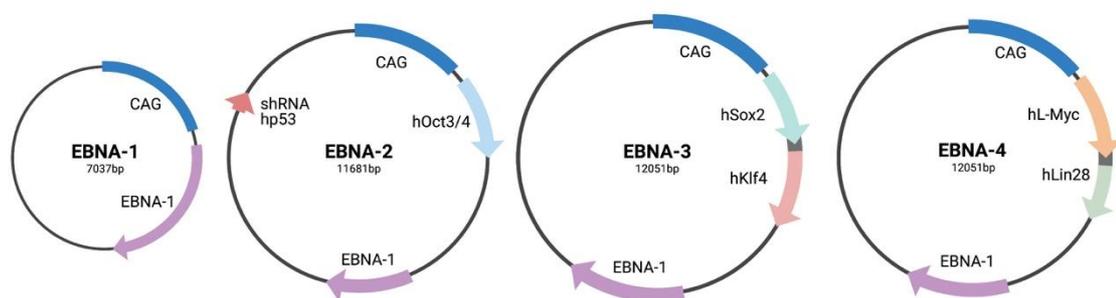
Figure 3: Schematic of EBNA reprogramming vectors.
The expression cassettes of the four EBNA reprogramming vectors are depicted and further described in the main text in chapter 5.8.3.

5.8.4   SMAR reprogramming tested

In previous work by Dr Alicia Roig Merino (Roig Merino, 2018), our S/MAR DNA vector system was tested as an alternative episomal reprogramming tool. The Lenti-4-in-1 expression cassette as well as the three reprogramming factors containing EBNA reprogramming vectors expression cassettes were cloned into S/MAR DNA vector backbones and transgene expression were verified. The Lenti-4-in-1 based S/MAR reprogramming vector (nPOP) was shown to be able to reprogram patient derived fibroblasts (Roig Merino, 2018). However, further work needs to be done to properly establish efficient and reproducible reprogramming. Especially the total expression levels of the reprogramming factors and the expression levels in relation to each other seem to be very crucial and can be easily altered when changing vector size, vector count and delivery method.

5.8.5   mRNA/small molecule reprogramming

Latest work on reprogramming further focuses on avoiding the use of the oncogenes or viral features by replacing the Yamanaka factors with small molecules (Ichida et al., 2009; Velychko et al., 2019) or mRNA (Poleganov et al., 2015). Also, evaluations on novel reprogramming factors as well as factor combinations or replacements are performed and show impressively that we are still at the very beginning of understanding the depths of the reprogramming process (Jung et al., 2021).

5.9   Cells isolated from the urine as cell source for hiPSCs

Recently, researchers are working on isolating and analyzing Urinary Stem Cells (USCs), which can expand *in vitro* after isolation from the urine (L. Chen et al., 2018; Choi et al., 2017; Falzarano & Ferlini, 2019; Oliveira Arcolino et al., 2015; T. Zhou et al., 2011). These cells can be obtained using non-invasive techniques, decreasing ethical concerns, and can be expanded to high numbers in a short period of time. The easy isolation and culturing procedure which can be performed with common cell culture laboratory equipment further presents a lower threshold for their use for e.g. personalized drug screenings (Mulder et al., 2020; T. Zhou et al., 2012). Also, first studies are working on the possible storage and preservation of USCs for delayed cell isolation (Lang et al., 2013).

It has been shown that these isolated and proliferating cells seem to be adult stem cells shed from the kidney (Bharadwaj et al., 2013; Oliveira Arcolino et al., 2015)*.* Routinely, two distinct morphologically distinguishable cell types are isolated: cobblestone-like cells that seem to are originated from the nephron tubule and spindle-like cells that are originated from the renal mesenchyme (Shi & Cheung, 2021).

Interestingly, these USCs share similar differentiation capacities like mesenchymal stem cells (MSCs) as they can successfully differentiate into bone, cartilage, and neurons (Manaph et al., 2018; Oliveira Arcolino et al., 2015). They are reported to express some markers in common with MSCs, HSCs and iPSCs (Bharadwaj et al., 2013; Pavathuparambil Abdul Manaph et al., 2018; Rahman et al., 2020), which makes them an interesting candidate for new cell therapy approaches.

Although their differentiation capacity is not yet fully elucidated, and their potential applications are only starting to be evaluated their utilisation for iPSC generation provides great potential for a variety of applications in research and therapy. It has been reported that USCs can be used as an alternative cell source for the generation of hiPSCs, where USCs reprogram even faster than other primary cells like fibroblasts (T. Zhou et al., 2012; T. Zhou et al., 2011). Especially for underaged patients this painless

INTRODUCTION

obtention of cells can lead to easier availability of patient-derived cells and improve patient well-being.

5.10 Genetic modification of stem cells using S/MAR DNA vectors

As elaborated in the previous chapters, stem cells in general provide the potential for many future therapies. For some of these therapies, not only the amplification and generation of specific cells and tissues is needed, but also the genetic alteration of these cells. Therefore, having a suitable technology to genetically modify these cells is key for future applications (Szablowska-Gadomska et al., 2013).

In the body, stem cells represent an important source for future cells and long-term maintenance of body functions. Therefore, they comprise extensive protection mechanisms for genetic alterations such as DNA damage and repair systems. Also, the delivery into the cells, as well as vector dilution due to extensive proliferation and vector silencing, hinder the genetic modification of these cells (Thyagarajan et al., 2009). The report of the repetitive need of antibiotic selection upon each passaging when utilising EBNA vectors to genetically modify stem cells remarkably illustrates that (Nassor et al., 2020).

CRISPR/Cas9 has been repeatedly reported to successfully enable the genetic modification of stem cells (E. J. Kim et al., 2017; Musunuru, 2013; Song & Ramakrishna, 2018). However, as previously discussed (chapter 5.3.1) there remain open questions on possible off-target effects and especially for their use beyond in vitro or animal model research towards cell therapies. Traditional gene therapy vectors such as retroviral vectors, AAVs, transposons or EBNA vectors have been tested and applied in stem cells, some more successful than others, and always less efficient than in differentiated or cancer cells (Byrne et al., 2014; Nassor et al., 2020).

When considering the genetic modification of these cells the same principles and wish-list as described for gene therapy before applies to the vectors used in cell therapy. As introduced, our S/MAR DNA vector platform presents an optimised gene therapy vector that provides sustained expression levels while being maintained episomal and affecting the host cells as little as possible. We therefore thrived to extend our application portfolio on these special types of cells.

INTRODUCTION

Before the work presented here, previous efforts were taken to evaluate the performance of S/MAR DNA vectors in these difficult to modify cell types. In these experiments, predominantly mouse ESCs and iPSCs were used as cell models. This work has been published in the dissertation of Dr Alicia Roig Merino as well as in an accepted shared-first authored paper in stem cell reports together with most of the work from the here presented thesis (Roig Merino, 2018; Roig Merino et al., in press).
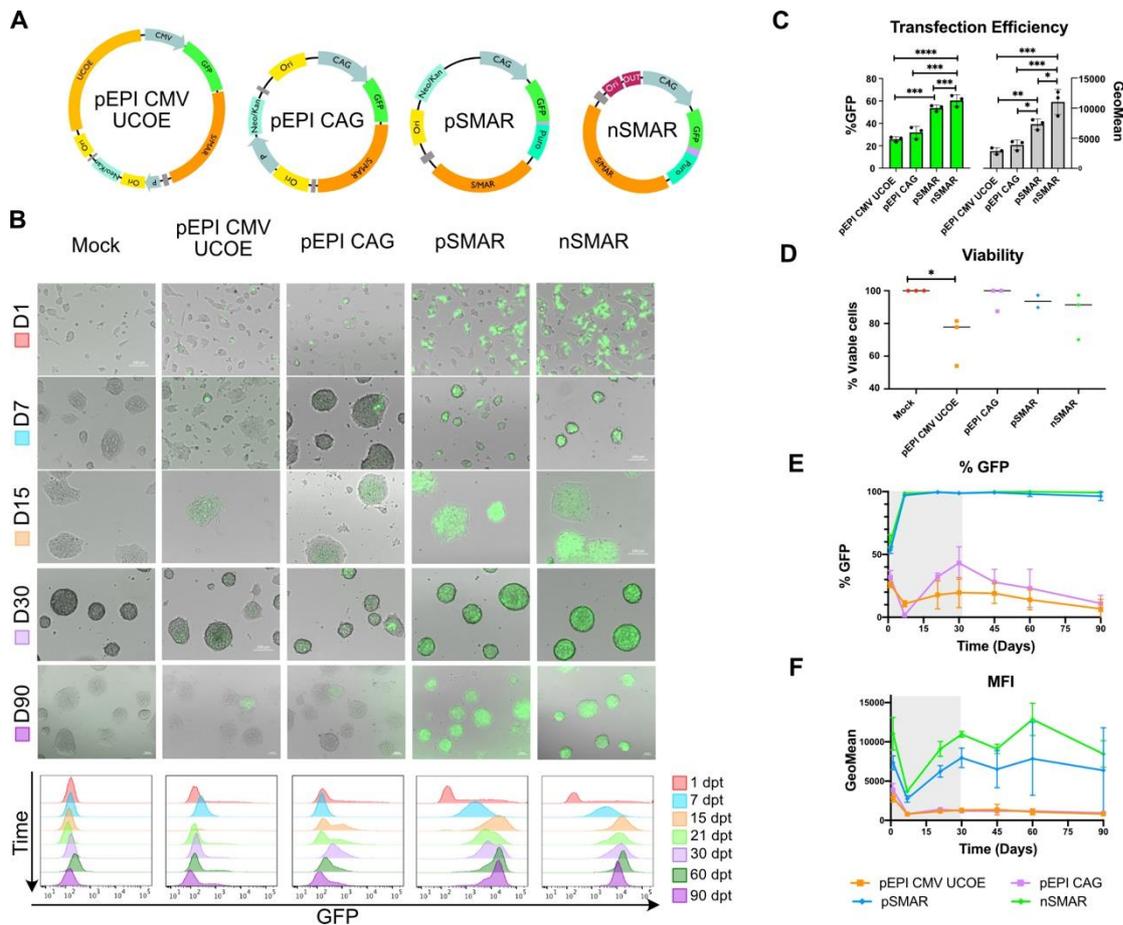
Figure 4: Generation of stable modified mESCs using different S/MAR DNA vectors. Reprinted with permission from the paper (Roig Merino et al., in press)
A) Schematics of utilised vectors are shown. B) Overview of GFP expression in fluorescence microscopy and FACS analysis on single, alive cells over 90 days. C) Comparison of Transfection efficiency of the different vectors in regard to the percentage of expressing cells and the geometric mean of the expressing cell population. D) Cell viability performance of the different vectors 24hpt. E) Time course FACS data of GFP expressing single, alive cells over three months. F) Time course FACS data of geometric mean of GFP expressing, single, alive cells over three months. Data is originally published and explained in detail in (Roig Merino et al., in press).

While we were not able to introduce GFP encoding pEPI vectors stably in these cells, the choice of a CAG promoter and optimisation of the vector features renders our pCAG and nCAG (pSMAR and nSMAR) S/MAR DNA vectors as a suitable expression system in stem cells. We were able to demonstrate improved transfection efficiency, viability as well as performance for stable modification (Figure 4 A, B). Antibiotic

selection was only applied to purify the expressing cells in the first 30 days. The generated stable pCAG and nCAG cell lines presented stable expression levels over a period of three months even without selection pressure (Figure 4 E, F). Analysis of pluripotency markers suggested no alteration of mouse stem cell features by the S/MAR DNA vector expression and its maintenance.

Since our vector system remains episomal, we further tested the effect on our vectors through extensive expression restructuring processes such as reprogramming or differentiation. Importantly, persistent GFP S/MAR DNA vector expression was demonstrated throughout reprogramming and random differentiation processes into cells of all three germ layers as well as minimal loss of pGAC expression upon 6 days of HSC differentiation. Episomal maintenance was also confirmed by southern blot and plasmid rescue experiments.

Further data derived from microarray analysis of this work also demonstrates the minimal changes in expression levels upon transient delivery of our vectors in hESCs. To provide more evidence, that mouse stem cell abilities are not altered through GFP S/MAR DNA vector expression, chimeric mice were generated via microinjection of stably modified mESCs. Fur colour confirmed successful contribution of the modified ESCs to the generated chimaeras and GFP expression was detectable in adult organs like muscle, skin, liver, heart, kidney as well as bone marrow, blood and spleen.

Subsequent backcrossing of the transgenic animals showed no GFP S/MAR DNA vector expression or detectable genetic material in the F1 generation. Investigation of spermatogenesis suggested meiosis as limiting factor for S/MAR DNA vector persistence. Importantly, this data highly supports the non-integrative nature of our S/MAR DNA vector system.

The presented data contains predominantly proof of principle studies with GFP as transgene and performance in mouse pluripotent stem cells and transgenic animal generation. While this is time-consuming and necessary to determine the mode of action of our system and the suitable configuration of the suitable S/MAR DNA vectors

for this cell type, we aimed to extend this work and move towards more translational experiments using functional transgenes and a disease application as well as human stem cells in the work presented here.

5.11 Birt-Hogg-Dubé syndrome: a model disease for the application of S/MAR DNA vectors

When moving towards a therapeutic application of our S/MAR DNA vector system monogenic diseases are prime targets. Here, the absence of a functional protein is the cause for the pathological phenotype. Also, the absence of a dominant-negative effect of the damaged endogenous protein is important for the ability of the supplementation of the functional protein with the S/MAR DNA vector system to restore the healthy phenotype. In this work, we focus on such a monogenic, hereditary disease, which is further introduced in the following.

5.11.1 Symptoms of Birt-Hogg-Dubé Syndrome

Birt-Hogg-Dubé syndrome is an autosomal dominantly inherited disease named after three Canadian physicians. In 1977, they were the first to describe the accumulation of skin lesions as classic features of this disease. Soon it became understood that the symptoms like lung cysts, spontaneous pneumothoraxes as well as the development of renal tumours are all indications of this disease (Schmidt, 2013; Toro et al., 1999).

The most common manifestation is skin fibrofolliculomas which occur predominantly on the upper torso, the face and the neck and occur in more than 85% of BHD patients (Leter et al., 2008; Toro et al., 2008). While this mainly leads to aesthetic problems, 75% of these patients develop several lung cysts, and in 25% of BHD cases, patients suffer from lung pneumothoraxes before their 40s (Gupta et al., 2013). Most severely, by their 50s kidney cysts and tumours start to arise in 1/3rd of the patients (Pavlovich et al., 2005). Interestingly, there is no predominant form of renal tumours, they can have a clear cell, chromophobe, papillary or oncocytic phenotype (Khoo et al., 2003). A recent study supports previous indications of a connection of BHD to the development of colon cancer (Sattler et al., 2021). Definite evidence for this connection however is difficult to obtain since BHD is a very rare disease with for example currently only ~75 diagnosed families in Germany and ~700 families worldwide. Due to the interdisciplinary symptoms and unfamiliarity of the diverse phenotypes, the disease is

thought to be highly underdiagnosed and incident numbers might change in the next years with raising awareness and genotyping opportunities.

5.11.2 Folliculin as genetic cause for Birt-Hogg-Dubé syndrome and its role in the cell

In 2002 Nickerson et al. identified a gene on chromosome 17 (17p11.2) to be associated with BHD (Nickerson et al., 2002). It was found that patients with this syndrome have a germline mutation in the BHD gene encoding for the protein folliculin. So far, more than 150 different mutations have been identified (Schmidt & Linehan, 2018). By being predisposed by this germline mutation, BHD patients develop second hit mutations predominantly in some hotspots in the FLCN gene. These lead to loss of functional FLCN and the subsequent development of renal cancer (Vocke et al., 2005). This and further research characterised the tumour suppressor characteristic of FLCN (Hudon et al., 2010). It is not yet clear which combination of mutations are linked to which phenotype and severity of the syndrome, which can vary significantly from patient to patient. However, it seems that LOH or somatic second hit mutations are not involved in the formation of skin lesions, rather more, haploinsufficiency of FLCN already leads to the development of fibrofolliculomas (van Steensel et al., 2007).

Researchers have attempted to understand how the loss of FLCN leads to tumour formation by the investigation of its interaction partners and its involvement in different pathways. It was found that the FLCN interacting proteins 1 and 2 (FNIP1/2) directly interact with FLCN and mediate AMPK interaction for regulation of the mTOR pathway (Baba et al., 2006). Besides this, FLCN seems to be involved in many fundamental pathways like Wnt and TGF-beta signalling (Cash et al., 2011; Luijten et al., 2013).

Interestingly, FLCN has also an impact on stem cell maintenance: By inhibiting nuclear TFE3 localisation, FLCN enables the exit from pluripotency of embryonic stem cells (Betschinger et al., 2013). This is further supported by the finding that homozygous ubiquitous FLCN knock-out mice are embryonic lethal, suggesting that the cells are not

able to differentiate upon loss of functional FLCN (Baba et al., 2008; Hasumi et al., 2009). The same phenotype of homozygous lethality occurs in the Nihon renal cell carcinoma (RCC) rat model, which harbours an insertion in Exon three of FLCN resulting in a truncated form of FLCN. Heterozygous Nihon rats start to develop renal carcinomas from three weeks on after their birth (Togashi et al., 2006).

5.11.3 Available *in vitro* models, suitability for gene therapy and current treatment

Previous work on FLCN was typically performed using the BHD-patient derived renal carcinoma cell line (RCC line) UOK257 which lacks functional FLCN (Y. Yang et al., 2008). Our group and others found that upon stable expression of wt FLCN the functionality of pathways disrupted by the loss of FLCN can be restored (Baba et al., 2006; Wong & Harbottle, 2013). Furthermore, it was shown that by introducing a wt FLCN transgene into the Nihon rat renal cancer (RC) development could be inhibited, emphasising FLCN as the driver mutation of RC (Togashi et al., 2006). Based on this, we think that supplementing a patients' kidney cells with a functional, stable and mutation-proof copy of FLCN before a second hit mutation occurs should maintain the molecular integrity of the cell if the genomic FLCN is lost. This treatment should accordingly suppress the development of RCs leading to a novel preventive gene therapy for BHD.

Currently, the only treatment available for this eventually life-threatening kidney cancer caused by BHD is surgical resection of affected tissue when a tumour exceeds a certain size. However, taking into account the genetic predisposition of BHD patients, this does not impede the development of further tumours due to novel second hit mutations, and a lasting or preventive solution such as gene therapy with a functional copy of FLCN is desirable.

5.11.4 Previous work with older versions of S/MAR DNA vectors on Folliculin and focus of this study

When working on FLCN researchers commonly compare UOK257 cells with a UOK257-2 cell line which was restored with a wt FLCN copy via lentiviral transduction (Baba et al., 2006). The advantages and disadvantages of lentiviral as well as other gene delivery vectors have been discussed extensively before (chapter 5.3). To enable a model system with less impact of the vector on the host cell, a stable FLCN restored UOK257 cell line (UOK257-FS) generated with a previous version of our S/MAR DNA vector platform was generated before (Wong & Harbottle, 2013). In this project, we aimed to apply our novel, improved S/MAR DNA vector platform to establish an updated FLCN restored UOK257 cell line model to study FLCN.

5.12 Stem cell derived disease modelling through kidney organoids

An important application of patient derived hiPSCs or hiPSCs in general is disease modelling. While patient derived cells can often help to investigate pathomechanisms or drug responses in a personalized manner, gene editing technologies such as CRISPR/Cas9 (chapter 5.3.1) enable the generation of disease models from the same genetic and episomal background. With these complementary cell lines, direct comparison without artefacts is possible.

Based on these cell lines, the ability of pluripotent stem cells to differentiate into all cells of the body can be taken advantage of. They can be used to engineer advanced 3D cell models, named organoids, that are composed of a mixture of different cell types and contain structures which resemble the full organ. Thus, organoids help to recapitulate the *in vivo* situation and provide a useful *in vitro* model as close to the *in vivo* situation as currently possible (Garreta et al., 2021). These systems have been successfully used to generate disease and cancer models for tissues such as e.g. the brain or kidneys and used for the study of diseases like Parkinson disease, polycystic kidney disease or (Forbes et al., 2018; Hofer & Lutolf, 2021; Li et al., 2014). In regard to the investigated BHD syndrome and the development of kidney tumours, the generation of kidney organoids as advanced multicellular model is especially of interest for this disease.

Due to organ shortages for transplantation medicine, in the last years, lots of research has been done on the development of organs like the kidney, and how to achieve directed differentiation of iPSCs to kidney progenitor cells and functional 3D organoids (Hasegawa et al., 2019; Hoogduijn et al., 2020). Several different protocols, all aiming at mimicking embryonic development with small molecules were generated (Morizane & Bonventre, 2017; Takasato & Little, 2015). All generate slightly different developmental stage kidney organoids and there is still much research to be done to generate kidney organoids as close to a functional kidney as possible (Garreta et al., 2019; Little & Combes, 2019). In the meantime, the current protocols are used to generate patient derived models for renal diseases (Little & Quinlan, 2020) and

protocol based on the publication of Freedman et al. was recently commercialized with stem cell technologies, enabling comparable differentiation conditions (Freedman, 2015).

The application of stable S/MAR DNA vector modified hiPSCs in kidney organoid differentiation protocols will provide further insights into the possible utilisation of our vectors in such sophisticated and highly differentiated systems. Furthermore, kidney organoids from BHD patient-derived iPSCs and KO iPSCs, and FLCN-S/MAR DNA vector restored cells will be a novel and useful tool for drug screens and further research on FLCN. While it has been reported that cyst formation in polycystic kidney disease can be monitored in kidney organoids, it has to be seen if FLCN-dependent tumorigenesis can be observed in this cell model (Freedman, 2015). Finally, the application of our FLCN-S/MAR DNA vector system can provide proof-of-principle data, if overexpression of FLCN with this system harms cell behaviour or differentiation and lay the groundwork for a potential preventive gene therapy for this disease.

5.13 Aim of the study

In this project, we aimed to bring together our expertise on the improved S/MAR DNA vector platform from the previous application in mouse stem cells and move towards their application in more clinically relevant human induced pluripotent stem cells. Additionally, we aimed to move from proof of principle studies with GFP towards a functional protein and to continue the former work on the genetic correction or rescue of Birt-Hogg-Dubé syndrome in an optimised, more advanced vector platform and cellular model systems.

Specifically, we aimed to:

- generate a novel set of S/MAR DNA vectors based on our knowledge of vector features and cell-type specificity

- verify and utilise our S/MAR DNA vectors that encode GFP or the functional transgene FLCN in cancer cell lines

- restore FLCN in FLCN deficient cell lines to generate stable cell models with reduced impact of the vector that can aid to study important roles of FLCN in the cell

- combine the knowledge of our improved S/MAR DNA vector platform with the background in mouse pluripotent stem cells and explore its application on human iPSCs and their potential for combined gene and cell therapy

- investigate alternative sources for the generation of hiPSCs

- improve hiPSC culturing to enable the delivery of our S/MAR DNA vectors

- be attentive to establish protocols that can be easily transferred towards clinical requirements and with little impact on the host cells

INTRODUCTION

- establish our S/MAR DNA vectors for the first time in reprogrammed hiPSCs

- characterise the generated cell line for expression, pluripotency and differentiation potential maintenance

- move on from GFP-reporter gene towards the stable expression of a functional transgene in pluripotent cells

- provide evidence that our novel episomal S/MAR DNA vector platform is capable of restoring gene function in knockout cells and a potential protective gene therapy for BHD by establishing and characterising respective cell lines

- generate kidney organoids as an advanced cell model for BHD and explore if FLCN-dependent tumorigenesis can be investigated

6 MATERIALS

6.1 Mammalian cells:

Table 1 - Mammalian cells used in this study

cell type	media	source
Feeder dependent hiPSCs	Feeder dependent hiPSC media	Generated in this study from NHDF and UCs
Feeder free hiPSCs	Stemfit basic02	Generated in this study from NHDF and UCs
FTC113	RPMI + 10% FCS (+ 1% p/s)	DNA vector lab, DKFZ
H1703	RPMI + 10% FCS (+ 1% p/s)	DNA vector lab, DKFZ
HEK293T	DMEM + 10% FCS (+ 1% p/s)	DNA vector lab, DKFZ
HeLa	DMEM + 10% FCS (+ 1% p/s)	DNA vector lab, DKFZ
iMEF	DMEM + 10% FCS + 1% p/s + 1% NEAA	Kind gift from Anne Rademacher (Dr. Karsten Rippe, DKFZ)
NHDF	DMEM + 10% FCS + 1% p/s + 1% NEAA	Promcell (C-12300); thermo fisher (C0045C)
U2OS	DMEM + 10% FCS (+ 1% p/s)	DNA vector lab, DKFZ
UOK257	DMEM + 10% FCS (+ 1% p/s)	NIH
USCs	USC medias	Isolated in this study

6.1.1 Mammalian cell media compositions

Table 2 – media compositions for mammalian cells

media	components	composition	volume
Advanced RPMI 1640 basal medium	Advanced RPMI	98%%	500ml
	L-GlutaMAX	1%	5ml
	p/s	1%	5ml
DMEM	DMEM	88%	500ml
	FCS	10%	50ml
	(p/s)	1%	5ml
	(NEAA)	1%	5ml
EB media	DMEM	80%	40ml
	FCS	20%	10ml
Feeder dependent hiPSC media (supplemented aliquot stable for 2 weeks)	Stock:		
	DMEM/F12		400ml
	KSR	20%	100ml
	p/s	1%	5ml
	NEAA	1%	5ml
	β-mercaptoethanol	0.1%	0.5ml

	<u>50ml aliquot</u>, supplemented with FGF2 (10µg/ml)	10ng/ml	50µl
Feeder free hiPSC media, Stemfit basic02 (supplemented aliquot stable for 2 weeks)	Bottle A Bottle B thawed at 4°C, mixed, 50ml aliquots stored at -20°C	80% 20%	400ml 100ml
	<u>50ml aliquot</u>, supplemented with FGF2 (10µg/ml or 50µg/ml)	10ng/ml or 30ng/ml	50µl or 30µl
RPMI	RPMI FCS p/s NEAA	88% 10% 1% 1%	500ml 50ml 5ml 5ml
USC primary media (supplemented aliquot stable for 2 weeks)	<u>Stock</u>: DMEM/F12 FCS p/s	89% 10% 1%	500ml 50ml 5ml
	<u>50ml aliquot</u>, supplemented with REGM SingleQuots: hEGF insulin hydrocortisone GA-1000 Transferrin Triiodothyronine Epinephrine FBS	0.1% 0.1% 0.1% 0.1% 0.1% 0.1% 0.1% 0.5%	50µl 50µl 50µl 50µl 50µl 50µl 50µl 250µl
USC secondary media REBM (supplemented aliquot stable for 2 weeks)	REBM p/s REGM SingleQuots: hEGF insulin hydrocortisone GA-1000 Transferrin Triiodothyronine Epinephrine FBS	97.8% 1% 0.1% 0.1% 0.1% 0.1% 0.1% 0.1% 0.1% 0.5%	50ml 0.5ml 50µl 50µl 50µl 50µl 50µl 50µl 50µl 250µl

6.2 Bacterial cells:

Table 3 – bacterial cells used in this study

cell type	company	catalogue number
ElectroMAX DH10β cells	Life Technologies	18290-015
one shot stbl3 chemically competent e.coli	Life Technologies	C7373-03
stellar competent cells	Takara Bio Europe	636766
subcloning efficiency DH5α competent cells	Invitrogen	18265017

6.3 Chemicals and reagents:

6.3.1 Cell culture components

Table 4 – cell culture components

reagent	company	catalogue number
β-mercaptoethanol	Gibco	31350010
Advanced RPMI 1640	Life technologies	12633-012
Distilled water, cell culture grade	Life technologies	15230-089
DMEM	Sigma aldrich	D5796-6X500ML
DMEM/F12 + GlutaMAX	Gibco	31331028
DMSO	Carl Roth GmbH	4720.1
FCS	Sigma aldrich	F7524-500ml
iMatrix Laminin-511	amsbio	AMS.892 012
KSR	Gibco	10828010
L-GlutaMAX (200mM)	Life technologies	35050-038
NEAA (MEM non-essential amino acids)	Gibco	11140-035
optimem	Thermo fisher scientific	31985070
p/s (penicillin-streptomycin)	Thermo fisher scientific	10378016
PBS	Thermo fisher scientific	10010023
PBS + Ca^{2+} +Mg^{2+},	Sigma aldrich	D8662-100ML
REBM Basal Medium	Lonza	CC-3191
REGM SingleQuots supplements CC-4127	Lonza	CC-4127
ReLesR	Stemcell technologies	05872
RPMI 1640	Sigma aldrich	R8758-6X500ML
StemFit Basic02 media (Ajinomoto)	Amsbio	SFB-500

StemPro accutase	Thermo fisher scientific	A1110501
TrypLE Express	Thermo fisher scientific	12604013
Trypsin	Sigma aldrich	T4049-100ML

6.3.2   Transfection reagents

Table 5 – transfection reagents

transfection reagent	company	catalogue number
JetPEI	VWR	101-10N
lipofectamine stem	Life technologies	STEM00003
XtremeGene9	Sigma aldrich	06365787001
XtremeGeneHP	Sigma aldrich	6366244001

6.3.3   Small molecules and supplements

Table 6 – small molecules and supplements

reagent	stock	usage	company	catalogue number
Activin A	100µg/ml in sterile filtered 4mM HCl; 10µg in 100µl 4mM HCl; 10µl aliquots at -20°C	1:10.000 for 10ng/ml	R&D Systems	338-AC-010
Ampicillin	100mg/ml in water	1:1.000	Agilent Technologies	300021-61
Carbenicillin	50mg/ml in EtOH	1:1.000	Applichem	A1491,0001
CHIR99021	10mM in DMSO; 5mg in 1.08ml DMSO; 10µl aliquots at -20°C	1:1.250 for 8µM; 1:3.333 for 3µM	Sigma	SML1046-5MG
DAPI	2µg/ml	1:1.000	Sigma aldrich	000000010236276001
Geneticin (G418)	50mg/ml	1mg/ml	Roth	2039.2
heparin	1mg/ml in water; u/mg lot specific: Lot SLCC4711 203u/mg	1:1.000 for 1µg/ml	Sigma	H3149-10KU

	potency = 49.3mg; at 4°C			
Kanamycin	30mg/ml in water	1:1.000	Roche	10106801001
puromycin	1mg/ml in water	0.5-1μg/ml	Applichem	A2856,0010
recombinant human FGF2	10μg/ml or 50ug/ml in 5mM Tris-HCl pH 7.6 + 0.1% BSA; 50μl or 30μl aliquots at -80°C	10ng/ml; 50μl in 50ml or 30ng/ml; 30μl in 50ml	Peprotech	100-18B
recombinant human FGF9	100μg/ml in water; 20μg in 200μl water; 20μl aliquots at -80°C	1:500 for 200ng/ml	Peprotech	100-23B
ROCK inhibitor / Y-27632 dihydrochloride	10mM; 1mg in 312μl PBS (MW 320.3mg/mmol); 10μl aliquots at -20°C	10μM; 1:1.000	Peprotech	1293823-1mg

6.3.4   Chemicals and reagents

Table 7 – chemicals and reagents

reagent	company	catalogue number
10% Formalin	Sigma Aldrich	HT501128-4L
10x TGS buffer	Biorad	161-0732
1kb DNA ladder	NEB	N3232S
4% PFA in PBS (for IF, no methanol)	Neolab	TCL119100ML
4% PFA in PBS (for IF, no methanol)	VWR	J61899.AK
acetone	Fisher chemicals	A/0600/17
Agar	Carl Roth	S210-3
Agarose	Sigma aldrich	A9539-500G
BSA	Serva	11930.03
chloroform	Sigma aldrich	3221
Endotoxin free water	MP Biomedicals	2430304
Ethanol absolute	Fisher chemicals	E10650DF/C17
Gel loading dye purple 6x	NEB	B7024S
glycerol	Sigma aldrich	15523
HALT proteasae and phosphatase inhibitor cocktail	Thermo fisher scientific	78440
HCl 37%	Roth	4625.2

Histogel specimen processing gel	Thermo Scientific	HG-4000-012
isopropanol	Sigma aldrich	33539-2
Laemmli sample buffer 4x	Biorad	161-0747
LB		
Maxcyte electroporation buffer	Maxcyte	
methanol	Sigma aldrich	32213-2
Milk powder	Carl Roth	T145.2
PageRuler Plus pre-stained Protein Ladder	Thermo fisher scientific	26619
Pellet paint	Merck millipore	69049-3
Peqgreen	Peqlab	37-5010
Phenol:chloroform:isoamyl-alcohol (25:24:1)	Sigma aldrich	77617-100ML
Restore PLUS Western Blot stripping buffer	Life Technologies	46430
RIPA	Boster Bio	AR0105
RNaseOUT	G-Biosciences	786-70
Shandon tissue-marking dye, green	Thermo Fisher Scientific	3120128
SOC media	invitogen	15544-34
sodium acetate	Invitrogen	AM9740
SuperSignal West Pico Plus chemiluminescent substrate	Thermo Fisher Scientific	34580
triton X-100	Applichem	A4975,0500
Trizol reagent	Cellutron Life Technologies	15596-026
Trypan blue	invitrogen	T10282
tween20	Applichem	A1389,0500
Type A gelatin from porcine skin powder	Sigma aldrich	G1890-100G
Water nuclease-free, autoclaved, DEPC-treated	Carl Roth	T143.5

6.4   Buffer compositions:

Table 8 – Buffer compositions

buffer	components
1x EB buffer	50ml 50x EB buffer 4.5L water
1x TBS	950ml water 50ml 20x TBS
1x TBST	1L 1x TBS 0.1% tween20 (= 1ml)
1x TGS	900ml water 100ml 10x TGS
20x TBS	1M Tris ph7.4 3M NaCl 54mM KCl
5% milk in TBST	5g milk powder 100ml 1x TBST
50x EB buffer	484,6g Tris (2M) 41g sodium acetate (0.25M) 37.2g EDTA (0.05M) adjust to pH 7.8 with acetic acid fill up to 2L with water
triton lysis buffer	99ml 1x TBS 1% triton X-100 (= 1ml) freshly add 1:100 HALT to aliquot

6.5   Enzymes and respective buffers

Table 9 – enzymes and respective buffers

reagent	company	catalogue number
bacterial alkaline phosphatase (BAP)	Invitrogen	18011-015
BamHI-FD	Life Technologies	FD0054
BcuI-FD	Life Technologies	FD1253
BglI-FD	Life Technologies	FD0074
BglII-FD	Life Technologies	FD0084
BpiI-FD (BbsI)	Life Technologies	FD1014
BshTi-FD	Life Technologies	FD1464
Bsp1407I-FD	Life Technologies	FD0934
CloneAmp Hifi PCR premix	Clontech	639298
DpnI-FD	Fisher Scientific	10819410
Eco32I-FD	Life Technologies	FD0303
EcoRI-FD	Life Technologies	FD0275
HindIII-FD	Life Technologies	FD0504

KpnI-FD	Life Technologies	FD0534
MluI-FD	Life Technologies	FD0564
NdeI-FD	Thermo fisher scientific	10349709
Phusion high-fidelity DNA polymerase	NEB	M0530S
PstI-FD	Life Technologies	FD0615
SalI-FD	Life Technologies	FD0644
SmaI-FD	Life Technologies	FD0663
T4 ligase	NEB	M0202S
T4 ligase reaction buffer	NEB	B0202S
T4 PNK	NEB	B0201S
XbaI-FD	Thermo fisher scientific	10151320
XhoI-FD	Life Technologies	FD0695

6.6 Kits and ready-to-use materials

Table 10 – kits and ready-to-use materials

kit	company	catalogue number
10x Chromium Next GEM Single Cell V(D)J Reagent Kit v1.1	10x Genomics	
Alkaline Phosphatase Staining Kit II	stemgent	00-0055
Chromium i7 Multiplex Kit	10x Genomics	120262
Dead Cell Removal Kit	Miltenyi Biotec	130-115-660
DNA-free DNA removal kit	Invitrogen	AM1906
DNeasy blood and tissue kit	Qiagen	69504
EndoFree plasmid maxi kit	Qiagen	12362
GenElute	sigma aldrich	NA1111-1KT
human dermal fibroblast nucleofector kit	Lonza	VAPD-1001
iBlot2 NC regular stacks	Invitrogen	IB23001
iBlot2 PVDF regular stacks	Invitrogen	IB24001
Illustra TempliPhi amplification kit	GE Healthcare	25-6400-10
InFusion HD cloning kit	Takara Bio	639649
Lentiviral reprogramming supernatant	Kindly generated and provided by lab members from the group of Dr Marco Binder, DKFZ	
LS Columns	Miltenyi Biotec	130-042-401
Mini-PROTEAN TGX gels 4-15%, 10-well	biorad	456-1083
Mini-PROTEAN TGX gels 4-15%, 12-well	biorad	456-1085

Mini-PROTEAN TGX gels 4-20%, 10-well	biorad	456-1093
Mini-PROTEAN TGX gels 4-20%, 15-well	biorad	456-1096
Neon 10µl and 100µl tip kit	Thermo fisher scientific	MPK1025
Neon transfection 100µl kit	Thermo fisher scientific	MPK10025
Phire tissue direct PCR master mix	Thermo fisher scientific	F170S
Pierce BCA Protein Assay Kit	Thermo fisher scientific	23225
primaQUANT 2x qPCR-CYBR-Green-BLUE-MasterMix ohne ROX	Steinbrenner	SL-9902B-5ml
QIAprep spin miniprep kit	Qiagen	27106
RevertAid H minus first strand cDNA synthesis kit	Thermo fisher scientific	K1622
RNase-free DNase set	Qiagen	14325
RNeasy kit	Qiagen	12577
Single index Kit N Set A	10x Genomics	10000212
StemMACS trilineage differentiation kit	Miltenyi Biotec	130-090-101
Wizard genomic DNA extraction kit	Promega	A1120

6.7   Plasticware and single-use materials

Table 11 – plasticware and single-use materials

material	company	catalogue number
10µl filter tips XL	Nerbe	07-612-8300
1000µl filter tips	Nerbe	07-693-8300
10cm bacterial dishes	Sarstedt	82.1473
10cm cell culture dishes	Sarstedt	83.3902
10ml serological plastic pipettes	Corning	CLS4488
12-well plates	Corning	3512
15ml falcon tubes	Falcon	2096
20µl filter tips	Nerbe	07-622-8300
200µl filter tips	Nerbe	07-662-8300
200µl wide orifice pipette tips	Starlab	E1011-8400
24-well plates	Thermo Scientific Nunc	10604903
25ml serological plastic pipettes	Corning	CLS4489
50ml falcon tubes	Falcon	2070
5ml serological plastic pipettes	Corning	CLS4487

6-well plates	Greiner	657160
96-well flat bottom for BCA assay	Greiner	655180
96-well plate, V-bottom, Nunc	Thermo Scientific	249935
96-well u-bottom plate	Greiner Bio-One	650180
Corning Transwell polyester membrane cell culture inserts, 6-well plate	Sigma	CLS3450
Cryotubes	Kisker Biotech	366656
Filter FACS tube	BD	
Glass Pasteur pipettes	Karl Hecht	HECH40567001
LightCycler 480 Multiwell Plate 96	Roche	04729692001
LightCycler 480 sealing foil	Roche	04729757001
Luna cell counting slides	biozym	872011
Multidispense tips 125ul	Integra	4425
OC1002x PA	Maxcyte	
PCR tubes with individual attached flat caps	Kisker Biotech	G003-F
Reagent reservoir 50ml	Corning	4870
Safe lock 2ml tubes	Eppendorf	300120.094
Scalpel		
T175 cell culture flask	Greiner	660160
T75 cell culture flask	Greiner	658175
Tissue-Tek Cryomold, Sakura	Fisher Scientific	10690461
Towerpack D1000ST tips	Gilson	F167204
Towerpack D10ST tips	Gilson	F167202
Towerpack D200ST tips	Gilson	F167203
u-plate 96-well glass bottom microscopy plate ibiTreat	ibidi	89626
Ultra-low attachment 24-well plates	Corning	3473
0.1cm gap Micro Pulser electroporation cuvette	Biorad	1652089
0.2ml flat cap tubes, assorted colors	Thermo Scientific	AB-0622
Safe lock 1.5ml tubes	Eppendorf	30120.086
Multidispense tips 12,5ul	Integra	4415

6.8   Antibodies

Antibodies used in this study are listed below. All antibodies were diluted in the respective blocking buffer, implying 5% Milk in TBST for Western Blot (WB) (7.1.11) and PBS supplemented with 1% FCS, 0.5% BSA and 0.1% triton X-100 or 3% BSA and 0.1% tween20 in PBS for the respective immunofluorescence (IF) method (7.2.13).

6.8.1 Primary antibodies

Table 12 – primary antibodies used in this study

antibody	origin	dilution	method	company	catalogue number
aSMA	ms	1:100	IF methanol	santa cruz	sc53142
aSMA	ms	1:200	IF PFA	invitrogen	MA5-15805
b3-Tubulin	ms	1:100	IF	santa cruz	sc80005
Flag M2	ms	1:1000	WB	sigma aldrich	F1804
FLCN D14G9	rb	1:500	WB	CST	3697S
FoxA2	ms	1:100	IF	santa cruz	sc374376
GAPDH G-9	ms	1:1000	WB	santa cruz	sc365062, Lot# J0516
Lin28	ms	1:50, 1:500	IF, WB	santa cruz	sc374460
Nanog, 1E6C4	ms	1:50	IF	santa cruz	sc293121
Oct3/4, n-19	gt	1:100, 1:500	IF, WB	santa cruz	sc8628
TRA-1-60	ms	1:50	IF	santa cruz	sc21705

6.8.2 Secondary antibodies

Table 13 – secondary antibodies used in this study

antibody	origin	dilution	method	company	catalogue number
Goat-Alexa Fluor 647	do	1:1.000	IF	Abcam	ab150131
Mouse-Alexa Fluor 488	do	1:1.000	IF	Abcam	ab150105
Mouse-Alexa Fluor 594	do	1:1.000	IF	Abcam	ab150108
Mouse-HRP	gt	1:10.000	WB	Life technologies	A31430
Rabbit-Alexa Fluor 647	do	1:1.000	IF	Abcam	ab150075
Rabbit-HRP	do	1:10.000	WB	Life technologies	A16023
Rat-Alexa Fluor 594	gt	1:1.000	IF	Abcam	ab150160

6.9   Plasmids

Table 14 – plasmids used in this study

plasmid name	details	source
EBNA-1 = pCXLE-EBNA1	EBNA-1 encoding reprogramming vector	Kindly provided by Prof Tristan Mckay (MMU), Addgene ID 37624
EBNA-2 = pCXLE-hOCTshp53	hOct3/4 encoding reprogramming vector with shRNA for hp53	Kindly provided by Prof Tristan Mckay (MMU), Addgene ID 27077
EBNA-3 = pCXLE-hSK	hSox2 and hKlf4 encoding reprogramming vector	Kindly provided by Prof Tristan Mckay (MMU), Addgene ID 27078
EBNA-4 = pCXLE-hUL	hL-Myc and hLin28 encoding reprogramming vector	Kindly provided by Prof Tristan Mckay (MMU), Addgene ID 27080
lentiviral vector pRRL.PPT.SF.hOKSM-IRES-dTom	Lentiviral reprogramming plasmid encoding Oct3/4, Klf4, Sox2, cMyc and dTom under an SFFV promoter	Kindly provided by Dr Michael Milsom (HiStem, DKFZ) and Prof Dr Axel Schambach (Warlich et al., 2011)
N2	Nanovector™ encoding Ele40-CAG-coGFP-p2a-FlCN-flag-SMAR	Cloned in this study in collaboration with NTC
N4	Nanovector™ encoding Ele40-CAG-coGFP-p2a-FlCN-flag-splicedSMAR	Cloned in this study in collaboration with NTC, SMAR designed by DNA vector lab member Dr Matthias Bozza
N6	Nanovector™ encoding Ele40-CAG-coGFP-p2a-FlCN-flag-splicedCMAR	Cloned in this study in collaboration with NTC, SMAR designed by DNA vector lab member Dr Matthias Bozza
nCAG	Nanovector™ based on pCAG, encoding CAG-coGFP-p2a-PuroR-SMAR	Cloned by DNA vector lab member Dr Alicia Roig Merino in collaboration with NTC
NV6	Nanovector™ encoding Ele40-CAG-coGFP-p2a-Rep1-splicedSMAR	Cloned by DNA vector lab member Dr Patrick V. Almeida in collaboration with NTC
P13	encoding CAG-coGFP-p2a-FlCN-flag-SMAR	Cloned in this study
P14	encoding CAG-coGFP-p2a-FlCN-SMAR	Cloned in this study
P15	encoding Ele40-CAG-coGFP-p2a-FlCN-flag-SMAR	Cloned in this study

P17	based on pCAG, encoding Ele40-CAG-coGFP-p2a-PuroR-SMAR	Cloned in this study
P3	encoding Ele40-CAG-PuroR-p2a-FlCN-flag-SMAR	Cloned in this study
P9	encoding Ele40-CAG-PuroR-p2a-FlCN-flag-IRES-coGFP-SMAR	Cloned in this study
pCAG	encoding CAG-coGFP-p2a-PuroR-SMAR	Cloned by DNA vector lab member Dr Alicia Roig Merino
px458-RFP-FLCN	CRISPR/Cas9 vector encoding gRNA for FLCN, Cas9 and a RFP reporter gene	Cloned in this study; px458-RFP kindly provided by Magdalena Büscher, HiStem, DKFZ

6.10 Primers

All primers were ordered and synthesised by sigma aldrich. The lyophilised primer stock was resuspended to 100µM using nuclease free water, and then diluted for working stocks as described for the respective application in the methods section.

6.10.1 qPCR primers

Table 15 – qPCR primers used in this study

name	sequence	Eff. Tm	amplicon length
qPCR13a_GAPDH_fwd	TCCTCTGACTTCAACAGCGACA	1.99	
qPCR13b_GAPDH_rev	TGTAGCCAAATTCGTTGTCATACCAG	>60	115bp
qPCR14a_FLCN_fwd	GGAGTGGATGAACAAAGTGAAGGTG	2.09	
qPCR14b_FLCN_rev	TCAGGCCAGTCATCCAGAACT	>60	146bp
qPCR15a_TGFb2_fwd	ACAGCAGGGTCCTGAGCTTAT	2.06	
qPCR15b_TGFb2_rev	AGAAAGCTGTTCAATCTTGGGTGTT	65	134bp
qPCR16a_SMAD7_fwd	CCTTACTCCAGATACCCGATGGAT	2.05	
qPCR16b_SMAD7_rev	CTCCAGAAGAAGTTGGGAATCTGAAAG	65	135bp
qPCR17a_Axin2_fwd	CTGGCTCCAGAAGATCACAAAGAG	1.90	
qPCR17b_Axin2_rev	CAGCATCCTCCGGTATGGAATTT	65	120bp
qPCR20a_SMAD3_fwd	GAGCGGAGTACAGGAGACAGAC	1.96	
qPCR20b_SMAD3_rev	ACTGGAACAGCGGATGCTTG	65	128bp
qPCR21a_RAB27B_fwd	GCATCTTCAGCTTTGGGACACT	2.01	
qPCR21b_RAB27B_rev	GCAGTTGGCTCATCCAGTTTCT	65	149bp
qPCR24a_coGFP_fwd	ACCAACAAGATGAAGAGCACCAAA	2.04	
qPCR24b_coGFP_rev	TTGTTGATGGCGTGCAGGAA	65	137bp
qPCR27a_Pax2_fwd	atcaacagaatcatccggaccaaag	1.99	

qPCR27b_Pax2_rev	tcattggaggcgctggaaac	64	137bp
qPCR29a_Oct4A_fwd	ccgtgaagctggagaaggagaa	1.95	
qPCR29b_Oct4A_rev	ctgcttcaggagcttggcaaat	66	104bp
qPCR30a_TBXT_fwd	gcttcaaggagctcaccaatga	1.96	
qPCR30b_TBXT_rev	agtccagcaggaaggagtacat	65	120bp
qPCR32a_FLCN-FLAG_fwd	CAAGCTGCTGAAGTTCTGGATGA	2.06	
qPCR32b_FLCN-FLAG_rev	cgtcatcgtctttgtagtcGTTCC	65	116bp
qPCR33a_WT1_fwd	GCGGAGCCCAATACAGAATACAC	2.07	
qPCR33b_WT1_rev	TCTCAGATGCCGACCGTACAA	65	105bp
qPCR34a_nanog_fwd	aatacctcagcctccagcagat	1.95	
qPCR34b_nanog_rev	cctgcgtcacaccattgctatt	65	150bp
qPCR36a_WT1_fwd	GTGCTTACCCAGGCTGCAATAA	2.06	
qPCR36b_WT1_rev	CGTTCACAGTCCTTGAAGTCACAC	65	112bp
qPCR39a_WT1_fwd	GTGACTTCAAGGACTGTGAACG	2.01	
qPCR39b_WT1_rev (Przepiorski et al., 2018)	CGGGAGAACTTTCGCTGACAA		118bp
qPCR42a_VIL1_fwd	AGACCCAGTCTTGCTGCTATCT	2.21	
qPCR42b_VIL1_rev	CTTCCACCACCACTTGCTTCTC	65	117bp
qPCR43a_ECAD_fwd	tgcagaccttcctcccaataca	2.05	
qPCR43b_ECAD_rev	acccacctctaaggccatcttt	65	136bp
qPCR46a_PODXL_fwd	GATAAGTGCGGCATACGGCT	1.99	
qPCR46b_PODXL_rev (Garreta et al., 2019)	GCTCGTACACATCCTTGGCA	65	103bp

Eff. = efficiency; Tm = melting temperature

6.10.2 Sequencing primer

Table 16 – sequencing primers used in this study

name	sequence
seq1_FLuc_fwd	TGCAACACCCCAACATCTTC
seq2_hPGK_fwd	ggttgcgccttttccaag
seq3_ori_fwd	ACGACCTACACCGAACTGAG
seq4_FLCNstart_fwd	CTCTCTGCCACTTCTGCGA
seq5_FLCNmiddle1_fwd	GCCCCATCTTCTTCGGAGA
seq6_FLCNmiddle2_fwd	CTGCCAGTCTTCAAGTCCCT
seq7_FLCNend_fwd	AAGATTGAAGCGGCTCTGAC
seq8_UbCend_fwd	GCTGTGAGGTCGTTGAAACA
seq9_SMARmiddle1_fwd	GTCTAAATGGAGAAGGCAAAACT
seq10_SMARend_fwd	AGGTAATGCTGGCCTAGTTGA

seq11_oristart_rev	gctcttgatccggcaaacaa
seq12_CMVstart_rev	GGCTATGAACTAATGACCCCG
seq13_IRESend_fwd	acgttgtgagttggatagttgt
seq14_FLCNendV2_fwd	GCCCCACCATCCTGAATAAG
seq15_pCS-SV40start_rev	AAAGGGAACAAAAGCTGGGT
seq16_pCS-CMVend_fwd	GAAGGACGCCAGGGTACATT
seq17_FLCNstart_rev	TCATGGCTGATATATCCCGGG
seq18_U6_fwd	tttgcatatacgatacaaggctg
seq19_Cas9end_fwd	TAAGCCCATCAGAGAGCAGG
seq20_bGHpolyA_rev	AGGAAAGGACAGTGGGAGTG
seq21_chimericintron_fwd	GGTTACAAGACAGGTTTAAGGAGGCC
seq22_SMAR_rev	TCCCCACACATCTCATGCTT
seq23_coGFPmiddle_fwd	ACAAGATCATCCGCAGCAAC
seq24_Epo_rev	AAATTAACTTTGGTGTCTGGGAC
seq25_CMVaddgene_fwd	CGCAAATGGGCGGTAGGCGTG
seq26_PCR19fwd	CGCATGACCAACAAGATG
seq27_PCR19rev	TTGCTGTGCAGCTCCTC

6.10.3 PCR primer

Table 17 – PCR primers used in this study

name	sequence
FLCN_gDNA_amplicon_fwd	GGAGGTTTCATGGAGTCAATAGG
FLCN_gDNA_amplicon_rev	CACTGCTCTCAGGTCCTCC
MU10a_cogfp_P2A_fwd	CCATCGCCTTCGCCAGATCTcgagctcgaggcagcggcgccaccaacttc
MU10b_FLAG_smar_rev	atcgagctCGAGATCTCActtgtcgtcatcg
MU11a_flag_SD_fwd	aagTGAAGATCTCGagatgcatgcagaagttggtcgtgagg
MU11a_hPGK_fwd	CCGCCATGTTACTAGTccacggggttgg
MU11b_bb_SA_rev	tagatccggtggatccgagtggacacct
MU11b_hPGK_rev	CAATAAAGCTACTAGTcctggggagagag
MU12a_flcn_FLCN_fwd	TCCGGGGAATCATCGATGAGCTCCAGG
MU12a_SpeI_UBC_fwd	CCGCCATGTTACTAGTGGCCTCCGCGCCG
MU12b_sd_FLAG_rev	gcatgcatctCGAGATCTCActtgtcgtcatcg
MU12b_UBC_SpeI_rev	CAATAAAGCTACTAGTTCTAACAAAAAAGCCAAAAACG
MU13a_cag_FLCN_fwd	TACCGAAGCCGCTAGCATGAATGCCATCGTGGCTCT
MU13a_flag_SD_fwd	TCTCGagatgcatgcagaagttggtcg
MU13b_FLAG_smar_rev	TAtcatcgagctCGAGcttgtcgtcatcgtctttg
MU14a15a_2_coGFP_fwd	CCGCATGACCAACAAGATGA
MU14a15a_chimeric_fwd	AGAAACTGGGCTTGTCGAGA
MU14b_FLCN_rev	CTCAGGCCAGTCATCCAGAA

MU15b_Puro_rev	tCGAGATCTtcaggcaccg
MU16a_P2A_hEpo_fwd	aagcaggccggcgacgtggaggaaaaccctgggcccATGGGGGTGCACGAATGTCCTGC
MU16b_18b_SMAR_hEpo_rev	TAtcatcgagctCGAGGATATCTCATCTGTCCCCTGTCCTGC
MU17a_intron_coGFP_fwd	TACCGAAGCCGCTAGCGCTACCGGTCGCCACCATGCCCGCCATGAAG
MU17b_hEpo_P2A_rev	ACCCCCATgggcccagggtttt
MU18a_P2A_hEpo_fwd	tgggcccATGGGGGTGCACGAATG
MU19a_GFPgeno_ARM_fwd	GCCGCATGACCAACAAGATG
MU19b_GFPgeno_ARM_rev	GTTGCTGTGCAGCTCCTCCA
MU1a_ar71_PURO_fwd	CGCTAGCGCTACCGGTatgaccgagtacaagcccacg
MU1b_PURO_p2a_rev	cagcaggctgaagttggtggcgccgctgccggcaccgggcttgcg
MU20a_FLCNgeno_Tm60_fwd	CAGATGAACAGTCGGATGCG
MU20b_FLCNgeno_Tm60_rev = MU14b	CTCAGGCCAGTCATCCAGAA
MU21a_FLCNgeno_Tm65_fwd	CTGCACGCCCCACTTCCTCA
MU21b_FLCNgeno_Tm65_rev	GCTGCGGACCGTGGACATGA
MU22a_coGFP-FLCNgeno_Tm60_fwd	AACACCCGCATCGAGAAGTA
MU22b_coGFP-FLCNgeno_Tm60_rev	CACTGGTCACCACAAACTCG
MU23a_coGFP-FLCNgeno_Tm65_fwd	GCCCCTACCTGCTGAGCCAC
MU23b_coGFP-FLCNgeno_Tm65_rev	GCCCCAGGAAGTTGCACCGA
MU24a_coGFP-PUROgeno_Tm60_fwd	AACACCCGCATCGAGAAGTA
MU24b_coGFP-PUROgeno_Tm60_rev	tcgtagaaggggaggttgc
MU24c_coGFP-KANgeno_Tm60_rev	TTGGGTGGAGAGGCTATTCG
MU25a_coGFP-PUROgeno2_Tm60_fwd	CGCCATGAAGATCGAGTGC
MU25b_coGFP-PUROgeno2_Tm60_rev	ggccttccatctgttgctg
MU26a_OriP_fwd	TCG GGG GTG TTA GAG ACA AC
MU26b_OriP_rev	TTC CAC GAG GGT AGT GAA CC
MU27a_EBNA-1_fwd	CCC AGG AGT CCC AGT AGT CA
MU27b_EBNA-1_rev	ATC GTC AAA GCT GCA CAC AG
MU2a_p2a_FLCN_fwd	aacttcagcctgctgaagcaggccggcgacgtggaggaaaaccctgggcccATGAATGCCATCGTGGCTCT
MU2b_1_FLCN_flag_ar71_rev	atcgagctCGAGATCTTCActtgtcgtcatcgtctttgtagtcGTTCCGAGACTCCGAGGCT
MU2b_2_FLCN_ar71_rev	atcgagctCGAGATCTTCAGTTCCGAGACTCCGAGGCT

MU3a_luc_IRES_fwd	AAGATCGCCGTGTAAcgccccccccccccctaacg
MU3b_IRES_bsd_rev	caaaggcttggccatttatcatcgtgttttcaaaggaaaaccacg
MU4a_BSD_fwd	atggccaagcctttgtctcaag
MU4b_LUC_stop_rev	ttaCACGGCGATCTTGCCG
MU5a_puro_p2a_FLCN_fwd	cgcaagcccggtgccggcagcggcgccaccaacttcagcctgctgaagcaggccggcgacgtggaggaaaaccctgggcccATGAATGCCATCGTGGCTCTCTGC
MU5b_PURO_p2a_rev	ggtggcgccgctgccggcaccgggcttgcg
MU6a_2_flcn_IRES_fwd	AACTGAAGATCTCGacgcccccccccccct
MU6a_flag_IRES_fwd	aagTGAAGATCTCGacgcccccccccccct
MU6b_IRES_cogfp_rev	GCGGGCATgcttatcatcgtgttttcaaaggaaaac
MU7a_ires_coGFP_fwd	gataagcATGCCCGCCATGAAGATCG
MU7b_coGFP_smar_rev	TAtcatcgagctCGAtcatcgagctCgAGATCTGGC
MU8a_bamhI_flag_hCDK4_fwd	cccGGATCCgccaccatggactacaaggacgacgatgacaagggaggacatatggctacctctcga
MU8b_hCDK4_ecorI_rev	CCCGAATTCTTActccggattaccttc
MU9a_bamhi_flag_hTSSK6_fwd	cccGGATCCgccaccatggactacaaggacgacgatgacaagggaggacatatgtcgggagacaaa
MU9b_hTSSK6_ecorI_rev	CCCGAATTCTTAgccggagtccccggc
Oligo_Crisprcloning_fwd	CACCGAGGCACCATGAATGCCATCG
Oligo_Crisprcloning_rev	AAACCGATGGCATTCATGGTGCCTC

6.11 Equipment

Table 18 – Equipment used in this study

device	details
-20°C freezer	
-80°C freezer U725 innova	New Brunswick (Eppendorf)
+4°C fridge	Liebherr
Amaxa II Nucleofector	Lonza
Bacterial hood	
Bacterial Shaker Centromat SII	B. Braun Biotech International
BD FACSAria Fusion	BD Bioscience
BD LSR Fortessa	BD Bioscience
Cell culture hood Safe2020	Thermo fisher scientific
Cell culture Incubator	
Centrifuge Ependorff 5430R	Eppendorf
Centrifuge Eppendorf 5424R	Eppendorf
Centrifuge Eppendorf 5810	Eppendorf
Centrifuge rotor F10-6x-500y FiberLite rotor	Thermo fisher scientific
Centrifuge Sorvall RC6+	Thermo fisher scientific
CoolCell	biozym 210001

MATERIALS

Eppendorf Thermomixer R	Eppendorf
Evos XL core cell imaging system	Thermo fisher scientific
Fusion SL chemiluminescence imagine system	Analis
Gamma cell 1000 irradiator	Gamma cell
Gel basic power supply PowerPacTM	biorad
Gel chambers PerfectBlue Wide Gel System ExM	Peqlab
Gel printer Mitsubishi P93D	Mitsubishi
iBlot 2	invitrogen
Integra Vacusafe cell culture pump	Integra
Integra VIAFLO electronic multidispense pipettes	Integra
Liquid Nitrogen cryostorage system K Series	Tec-Lab
LUNA automated cell counter	logosbio
MicroPulser	biorad
Microscope Keyence BZ-X700 Fluorescent	Keyence
Microscope phase contrast Olympus CK40	Olympus
Microwave	Siemens
Millipore machine	
Mini spin centrifuge	
Mini-PROTEAN Tetra Cell vertical electrophoresis cell	biorad
MS2 minishaker (vortex)	IKA
Multiskan Go plate reader	Thermo fisher scientific
Nanodrop 2000C	Thermo fisher scientific
Neon transfection system	life technologies
Neubauer chamber	
Nikon Eclipse Ti/X-Cite120Led microscope	Nikon
PCR hood	Labcaire
PCR thermocycler peqSTAR 2X/96X Universal Gradient	VWR peqlab
Pipettes 10μl, 1000μl	Eppendorf
Pipettes 20μl, 100μl, 200μl, 1000μl	Gilson
Pipette Boy	Integra
Plate shaker	
QIAcube robot	Qiagen
QuadroMACS Separator	Miltenyi biotec
Real-Time PCR System LightCycler® 96	Roche
Tube roller	NeoLab
UV transluminator N90 LW366	Konrad Benda
Water bath	

6.12 Software and websites

Table 19 – software used in this study

software	origin
affinity designer	Pantone
Endnote	Clarivate Analytics
Excel	Microsoft
Fiji	(Schindelin et al., 2012)
FlowJo	FlowJo, LLC
FUSION-CAP software	Analis
Graphpad Prism	GraphPad Software Inc.
LightCycle96 Real-Time PCR system	Roche
SnapGene	GSL Biotech LLC
Word	Microsoft

Table 20 – websites used in this study

name	website
BioRender	most figures created with BioRender.com
Ensembl	https://www.ensembl.org/index.html
GenBank	https://www.ncbi.nlm.nih.gov/genbank/
IDT PrimerQuest	https://eu.idtdna.com/Primerquest/Home/Index
NEBioCalculator	https://nebiocalculator.neb.com/#!/ligation
primer blast	https://www.ncbi.nlm.nih.gov/tools/primer-blast/
Primer3	https://primer3plus.com/primer3web/primer3web_input.htm
Tm calculator	https://www.thermofisher.com/de/de/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html

7 METHODS

7.1 Molecular biology methods

7.1.1 Cloning of S/MAR plasmids

For the cloning of S/MAR plasmid constructs, SnapGene was used to design the cloning strategy. After the determination of appropriate restriction sites or the design of primers, the following methods were used to obtain the desired constructs.

7.1.1.1 Agarose gel electrophoresis

For the separation of nucleic acids e.g., after restriction digestion reaction or PCR amplification, agarose gel electrophoresis was performed. To do so, first, an agarose gel with the appropriate number of wells and size was prepared. The concentration of agarose depended on the desired nucleic acid size and separation. Most times, 1% agarose in EB buffer was carefully boiled in a microwave until all agarose was dissolved. Depending on the total volume, 1-4µl of peqgreen was added to allow nucleic acid visualization, and the solution was poured in a prepared chamber with an appropriate comb. When the agarose gel was solidified, the nucleic acid samples were mixed with a loading buffer and loaded on the gel together with an appropriate DNA ladder. The samples were separated by size after applying a voltage ranging from 70-120V if not stated otherwise. Progress of separation and analysis of the gel was performed using a UV lamp or the Fusion SL chemiluminescence imagine system.

7.1.1.2 Phenol-chloroform extraction of DNA

To extract DNA from solutions a phenol-chloroform extraction was performed. First, the DNA solution was adjusted to a total volume of 200µl using nuclease-free water. Next, an equal volume (200µl) of phenol-chloroform-isoamyl-alcohol was added. The mixture was thoroughly vortexed and spun for 5min at RT at max. speed. The resulting two phases consisted of a lower, yellowish, organic phase, and an upper, clear aqueous phase, which contained the DNA. Therefore, the upper phase was carefully transferred to a new 1.5ml tube using a 200µl pipette. If needed, a residual aqueous phase was left to avoid carryover of the organic phase. The DNA was reextracted using

a 1:29 solution of the phenol-chloroform-isoamyl-alcohol and chloroform. Another round of vortexing and centrifugation as described before was performed and again the resulting upper aqueous phase was transferred in the same way to a new 1.5ml tube. After this DNA extraction, often precipitation of DNA was performed as described in the following (chapter 7.1.1.3).

7.1.1.3   Precipitation of nucleic acids

Often, DNA had to be precipitated to concentrate the solution or to change the buffer. Sometimes, phenol-chloroform extraction as described before was performed beforehand to clean up the DNA solution (chapter 7.1.1.2). For precipitation of DNA in a solution, 1/10th of its volume of 3M sodium acetate pH 5.2 was added. For RNA precipitation, 5M ammonium acetate was used instead. To aid the visualisation of the nucleic acid pellet, 1µl of pellet paint was added to the solution. After further addition of 3x volume of 100% EtOH, the mixture was incubated at -80°C for at least 30min up to overnight. After the incubation, a spinning at max. speed for 10min pelleted the precipitated nucleic acid and a small white (without pellet paint) or pink (with pellet paint) pellet was visible. The supernatant was carefully aspired, and the pellet washed with 75% EtOH. After a brief incubation at RT, the tube was spun again for 2min max. speed, and all supernatant was carefully aspired. If needed, another brief spin was performed to aspire residual liquid using a small volume pipette. Immediately upon removal of all EtOH, the pellet was resuspended in the appropriate volume of water or desired other buffers. Usually, 10µl of nuclease-free water was used for this step.

7.1.1.4   Digestion of Plasmid DNA

For control digestions, where restriction patterns were only analysed via UV light after agarose gel electrophoresis extraction, the input ranged from 200-1000ng of plasmid DNA. For DNA digestions, where the agarose bands were subsequently isolated and used for cloning and thus big amounts of resulting cut DNA was needed, around 4µg of plasmid DNA was used as input DNA. Plasmid digestions were usually performed in a total volume of 20µl, consisting of 2µl 10x digestion buffer, 1.5µl per enzyme, the sample DNA and filled up with nuclease-free water. The reaction was incubated for

20min at 37°C for digestion pattern analysis with small DNA input, or 1-4h at 37°C for digestion of large amounts of DNA.

7.1.1.5   PCR of Plasmid DNA

7.1.1.5.1   CloneAmp HiFi PCR premix

For cloning of plasmid DNA, mainly the CloneAmp HiFi PCR premix was used to amplify specific DNA regions. Primers were designed using the In-Fusion cloning tool in SnapGene with the pre-set parameters of 15bp overlap and target Tm of 60°C. After ordering the primers from sigma aldrich, the lyophilised primers were resuspended to a 100µM stock and further diluted to 10µM working stocks. For the PCR reaction, ~5-10ng of template DNA was used, mixed with 12.5µl of the CloneAmp HiFi PCR premix, 0.75µl of forward and 0.75µl of reverse primer of the 5µM working concentration. The mixture was filled to a total of 25µl using nuclease-free water. After brief mixing via pipetting and quick spin, the PCR was performed using the following conditions: denaturation at 95°C for 2min, 30 cycles of amplification, including brief denaturation for 10sec at 98°C, annealing for 10sec at 55°C, elongation at 72°C for 5sec/kb with 2-3 seconds addition and a final elongation for 10min at 72°C. The PCR product was analysed by running the whole PCR mixture on a 1% agarose gel (chapter 7.1.1.1). The desired band was cut, and the PCR product extracted as described elsewhere (chapter 7.1.1.6). The annealing temperature of 55°C usually was robust, however, sometimes it had to be adjusted when many unspecific amplification bands were observed in the agarose gel.

7.1.1.5.2   Phusion PCR reaction

For PCRs using the Phusion polymerase, the reaction mix consisted of 10ng DNA template, 5µl of 5x HF buffer, 5µl of 2mM dNTPs, 0.5µl of Phusion polymerase and 1µl of a 12.5µM primer mix containing both forward and reverse primers. The reaction was filled up to 25µl using nuclease-free water. The primer annealing temperature was determined using the Tm calculator website from thermo fisher and adjusted if necessary. The cycling conditions were 98°C for 1min denaturation, then 35 cycles of denaturation at 98°C for 20sec, annealing for 20sec at the calculated annealing

temperature, elongation at 72°C for 1min per kb and final elongation at 72°C
for 10min.

7.1.1.6   Gel extraction

Upon separation of DNA via an agarose gel, sometimes specific bands needed to be
isolated for downstream processing. Therefore, the agarose gel was handled under UV
light as little as possible to avoid UV-mediated mutation. The desired band was cut
under UV light using a scalpel and transferred to a 1.5ml tube. The slice was stored at -
20°C or directly used for the next step. The DNA was extracted from the gel using the
GenElute kit (sigma aldrich) as of manufacturers' recommendation. All centrifugation
steps were performed at 16.000xg for 1min. Briefly, the gel slice was weighted, and
mixed with 3x the volume of gel solubilisation solution and dissolved at 55°C for up to
10min with vortexing in between. In the meantime, the binding column was prepared
by the addition of 500µl column preparation solution and 1min centrifugation. Upon
successful dissolving, 1x gel volume of 100% isopropanol was added to the tube and
the whole solution was loaded on the prepared column. After another round of
centrifugation, the DNA was loaded on the column and the flow-through was
discarded. The column was washed with 700µl wash solution and centrifugation and
subsequently dried of residual ethanol using an empty centrifugation round. Finally,
the DNA was eluted using either nuclease-free water or the provided elution buffer. To
increase elution efficiency, the elution solution was incubated after addition to the
column for up to 2min before elution via centrifugation into a fresh 1.5ml tube. The
volume of the elution solution was adjusted to the desired concentration and ranged
from 20-100µl.

7.1.1.7   Infusion cloning

For the generation of plasmids, mainly the InFusion Kit (Takara) was used. The
sequences of the designed vector were designed using the In-Fusion reaction tool in
SnapGene and the DNA fragments of the vector and insert appropriately prepared
using restriction digestion and PCR amplification (chapter 7.1.1.4 and 7.1.1.5). 50ng of
the insert was mixed with 100ng of the vector, together with 2µl of the 5x infusion mix

and everything was filled up with nuclease-free water to a total of 10µl. In special cases with unusual sizes of insert and vector, the molar ratio had to be adjusted as recommended by the manufacturer. Since the volume of the infusion reaction was very little, the DNA fragments had to be very concentrated by either eluting in a little amount of water or concentration via precipitation as described before (chapter 7.1.1.3). The mixed infusion reaction was incubated at 50°C for 15min in a water bath, during which time the recombination of the 15bp overlapping region of insert and vector was happening. After the incubation time, the reaction was put on ice and 2.5µl was used for subsequent transformation of stellar competent bacteria as described in chapter 7.1.1.9. Alternatively, the whole reaction or the leftover was frozen and stored at -20°C for transformation at a later time point.

7.1.1.8   T4 ligation

For a simple cut and paste cloning strategy using T4 ligase, both vector and insert were cut with the same restriction enzymes to achieve fitting sticky ends. When cut with only one enzyme, to avoid relegation of the vector, it was dephosphorylated by the addition of 1µl of bacterial alkaline phosphatase (BAP) and incubation at 65°C for 1h. The sample was run on a gel and extracted as described before (chapter 7.1.1.1 and 7.1.1.6), in order to purify the cut and dephosphorylated band from uncut vector, enzymes and buffer components.

If needed, the insert was first produced using PCR amplification with primers constructed in a way that they contain the appropriate restriction enzyme cutting sites. The insert remained phosphorylated, to enable subsequent T4 ligation, however, it was handled similarly as the vector after digestion with the appropriate restriction enzymes.

For the T4 ligation reaction, a vector : insert ratio of 1:20 was calculated using the NEBioCalculator website. Then, the appropriate amount of vector and insert, to achieve a maximum of 200ng total DNA input was mixed, together with 2µl 10x T4 ligase buffer and 1µl T4 DNA ligase and everything was filled up with nuclease-free

water to a total volume of 20μl. Everything was handled on ice while pipetting, and subsequently, the reaction was incubated at 16°C overnight in a thermocycler. The next day, the ligase was heat-inactivated at 65°C for 10min. The ligated plasmid was extracted from the mixture using phenol-chloroform extraction (chapter 7.1.1.2) and precipitated and resuspended in 5μl nuclease-free water as described before (chapter 7.1.1.3). 2.5μl of the obtained plasmid solution was used for transformation of competent bacteria (chapter 7.1.1.9) or stored at -20°C.

7.1.1.9　Transformation of competent bacteria

For cloning purposes, different competent bacteria were used and transformed using heat shock. 50μl Stellar competent cells were used for the direct transformation of 2.5μl of infusion reaction without clean up or precipitation. For longer storage of plasmid DNA in bacterial glycerol stocks, plasmid DNA was retransformed into DH5α competent cells. For more stable maintenance of plasmid DNA, avoiding recombination events, stbl3 cells were used.

All competent bacteria were thawed on ice. If aliquoting had to be performed, bacteria was gently handled on ice, aliquoted in autoclaved, pre-cooled 1.5ml tubes and rapidly frozen by incubation 2-3min on a mixture of dry ice with EtOH and subsequently stored at -80°C. When used for transformation, the appropriate amount of plasmid DNA (2.5μl of infusion reaction, ~10ng of plasmid DNA for retransformation) was added to the thawed bacteria on ice. The bacteria with the plasmid DNA was incubated on ice for 30min to allow equilibration of bacteria and mixing with the plasmid DNA. Then the tube was held in a water bath at 42°C to perform a heat shock for 20sec (DH5α) or 45sec (stbl3 and stellar), respectively. The tube was immediately put back on the ice and the bacteria were mixed with 500μl prewarmed SOC media and recovered at 37°C with 180rpm shaking for 1h to allow expression of the antibiotic resistance gene provided by the transformed plasmid DNA. During the recovery time, LB plates with the appropriate antibiotic were pre-warmed at RT. Then, the transformed and recovered bacteria suspension was pelleted at 4.000rpm for 5min, resuspended in 100-200μl SOC media and plated on the prewarmed LB plate. The

bacteria were grown by incubating the plate at 37°C overnight. If needed, the number of plated bacteria had to be adjusted to enable single colony picking the next day. The whole or remaining bacteria solution could be stored at 4°C for a maximum of one week until plating.

For rescue experiments (chapter 7.1.7) the electrocompetent bacteria strain DH10β was used. The bacteria were handled in a similar way as the other competent bacteria and thawed on ice. However, after the addition of a maximum of 1/10th of the total volume of DNA, bacteria were transferred to a pre-cooled 0.1cm gap electroporation cuvette and immediately electroporated using a biorad MicroPulser and the bacterial program. Different amounts of DNA were tried ranging from 200ng to 2μg. Subsequently, bacteria were resuspended in 500μl pre-warmed SOC media and transferred to a fresh 1.5ml tube. Again, recovery at 37°C with gentle shaking was performed and the pelleted and resuspended bacteria were plated on an appropriate LB plate and incubated overnight at 37°C.

7.1.1.10 Preparation of LB agar plates

LB agar plates were prepared by addition of 15mg agar to 1L of LB solution. After autoclaving the solution was handled in a water bath at 50°C until it cooled down so it could be touched for more than 10sec. Then, the temperature was assumed to be stable and low enough that antibiotics were not degraded. The appropriate antibiotic was added, mixed, and approximately 25ml was poured per 10cm bacterial plate in a sterile environment under a bacterial hood. After solidification, the lids were closed, and the plates stored upside down in a plastic wrap at 4°C until usage.

7.1.1.11 Glycerol stocks of bacteria

For long-term storage of bacteria transformed with a specific plasmid, a storage solution consisting of 50% glycerol diluted in water was prepared and autoclaved. Approximately 500μl of the bacteria-LB solution was mixed 1:1 with 500μl of the 50% glycerol solution in a 1.5ml tube and stored at -80°C. For the regrowth of frozen

bacteria, some of the frozen glycerol stock was scraped off using a sterile pipette tip and used to inoculate a fresh LB solution.

7.1.2   Plasmid DNA preparation

7.1.2.1   Small scale plasmid isolation

For small scale production of plasmid DNA, e.g., to analyse and verify a cloning, single bacterial colonies were picked using a sterile pipette tip and grown overnight in 5ml LB solution at 37°C with 180rpm shaking. The next day, 4ml of this small-scale bacterial culture was used to isolate its plasmid DNA with the QIAprep spin miniprep kit (Qiagen). Briefly, 2x 2ml of bacterial culture was pellet on top of each other at 6.800xg for 5min and the pellet was resuspended in 250μl buffer P1 supplemented with RNaseA. At this step, plasmid isolation was either continued using the QIAcube robot and the rapid plasmid isolation protocol or manually as follows: the resuspended bacteria was lysed by the addition of 250μl of buffer P2. The tube was inverted four to six times and the lysis reaction was incubated at RT for max. 5min. The addition of 350μl buffer N3 neutralized the lysis reaction and caused precipitation of genomic bacterial DNA as well as cellular protein and membrane debris. After spinning at 13.000rpm for 10min the plasmid DNA containing supernatant was transferred to a provided QIAprep spin column and bound by centrifugation at 13.000rpm for one minute. A washing step using 500μl buffer PB and subsequently 750μl buffer PE was performed using the same centrifugation steps. After an empty spin to remove residual buffer the column was transferred to an empty 1.5ml tube and the plasmid DNA was eluted using 50μl EB buffer or nuclease-free water. The obtained plasmid DNA was analysed as described in the quality control section (chapter 7.1.3) and stored at - 20°C.

7.1.2.2   Large scale endotoxin free plasmid isolation

For large scale production and purification of plasmid DNA, 250ml of bacterial culture of LB media with the appropriate antibiotic was inoculated and grown overnight at 37°C with 180rpm shaking. The next day, the EndoFree plasmid maxi kit (Qiagen) was

METHODS

used for plasmid isolation. Briefly, the culture was transferred to a centrifugation tube and spun at 6.000xg for 15min at 4°C. The obtained pellet was resuspended in 10ml buffer P1 and transferred to a 50ml tube. Lysis of bacterial cells was performed by addition of 10ml buffer P2, inversion for 4-6 times and incubation at RT for 5min. In the meantime, the QIAfilter cartridge was closed using the provided cap. After the incubation, the lysis was stopped and genomic DNA, as well as cell debris, precipitated by the addition of 10ml of cold buffer P3 and inversion of the tube. The solution was poured on the cartridge and incubated for 10min at RT to allow sedimentation of the precipitate. Then, the plunger was inserted, and the lysate filtered through the cartridge into a new 50ml tube. 2.5ml of ER buffer was added to the filtered lysate containing plasmid DNA and incubated on ice for 30min to subsequently prevent present bacterial endotoxins from binding to the resin of the QIAGEN-tip. In the meantime, the QIAGEN-tip was equilibrated by addition of 10ml buffer QBT and elution by gravity flow. After the incubation time, the filtered and treated lysate was added on the prepared tip. Plasmid DNA bound to the tip when the solution went through the tip by gravity flow. Subsequently, the tip was washed 2x using 30ml of buffer QC and the plasmid DNA was eluted in a fresh 50ml tube using 15ml of elution buffer QN. The eluted DNA was precipitated by addition of 10.5ml of isopropanol and pelleted by centrifugation for 30min at 4°C at 15.000xg. The supernatant was carefully decanted, the pellet gently washed by addition of 70% EtOH and another 10min centrifugation at 4°C at 15.000xg. All alcohol solution was carefully removed, and the plasmid DNA pellet dried until the pellet started to turn from white to clear. Before it became completely dry, the DNA was resuspended in 300-400µl endotoxin-free water. The concentration and quality were measured as described (chapter 7.1.3) and the plasmid DNA was used for transfection of mammalian cells or further cloning and stored at - 20°C.

7.1.3 Quality control of plasmid DNA

For many purposes, a nanodrop machine was used to determine nucleic acid concentration and purification. For blanking, the respective buffer in which the nucleic acid was contained was used.

Furthermore, control digestions with enzymes that were non-cutters, single cutters were performed and run on an agarose gel to analyse the quality of a DNA plasmid purification and check, if the plasmid was predominantly in a nicked, linear or supercoiled state. Restriction pattern analysis with multiple cutters was predominantly performed to verify cloning success or plasmid identity. Digestion reactions were performed as described before (chapter 7.1.1.4).

Finally, DNA samples were sent for sequencing with eurofins genomics to verify base pair sequence of a specific DNA segment. For this, sequencing primers were designed using pre-set values of the primer3 website tool, inserting the beginning or end of the desired DNA segment and searching using the pick sequencing primers task. Chosen primers were ordered and synthesised by sigma aldrich. Upon arrival, lyophilised primers were resuspended in nuclease-free water to a stock concentration of 100μM. For sequencing, a working dilution to 5μM was performed, and approximately 500μl of sample DNA in 5μl was mixed with 5μl of the 5μM sequencing primer stock and sent to eurofins. Analysis of the obtained sequence was performed using SnapGene.

7.1.4 Genomic DNA extraction

7.1.4.1 DNeasy blood and tissue kit

The DNeasy blood and tissue kit was used to isolate genomic DNA (gDNA) from cells for rescue experiments and southern blot. Briefly, around 1.5×10^5 cultured cells were spun for 5min at 200xg, washed with PBS and resuspended in 200μl PBS. Next, 20μl provided proteinase K and 200μl buffer AL were added. The solution was mixed by thorough vortexing. Then, 200μl of 100% EtOH was added and again mixed by vortexing. The mixture was then placed on a DNeasy mini spin column, and the

genomic DNA was loaded on the column by centrifugation at 6.000xg for 1min. The column was washed using 500µl of the AW1 buffer, spun again at 6.000xg for 1min and washed again this time using the buffer AW2 and centrifugation at 20.000xg for 3min. The column was transferred to a new 1.5ml tube and the gDNA eluted using 2x 50µl of elution buffer AE, centrifugation for 1min at 6.000xg and 1min incubation of elution buffer before spinning.

7.1.4.2   Phire tissue direct PCR master mix kit

For rapid gDNA isolation without any column or precipitation where plasmid vectors might get lost, the phire tissue direct PCR master mix kit was used. Here, the desired cell count was harvested, washed with PBS and the cells were pelleted. Based on the pellet size, the cells were resuspended in a multiple of 20µl dilution buffer mixed with 0.5µl of DNArelease additive. The cell suspension was pipetted up and down until no clumps remained. If necessary, more solution was added to the cell suspension. After 2min incubation at RT, the cells were lysed by heating at 98°C for 2min. After a quick spin, the supernatant was directly used as input for a PCR reaction using the provided PCR master mix or stored at -20°C until further usage.

7.1.4.3   Wizard genomic DNA extraction kit

For some experiments like copy number or rolling circle amplification assays, gDNA was isolated using the wizard genomic DNA extraction kit. Here, as with the phire tissue direct PCR master mix kit, no column was used to purify the DNA, avoiding the loss of plasmid DNA in the gDNA extraction. However, in this kit, precipitation of the gDNA was performed to obtain a purer gDNA. Briefly, $0.5\text{-}1\times10^6$ cells were harvested and washed with PBS. The cell pellet was then resuspended and lysed in 600µl nuclei lysis solution by pipetting until no visible cell clumps remained. Then, 3µl of RNase solution was added, mixed by inversion of the tube, and incubated for 15-30min at 37°C. After reaching RT again, 200µl of protein precipitation solution was added and vortexed vigorously for 20sec. Subsequently, the sample was incubated on ice for 5min and then spun for 4min at 16.000xg. The precipitated protein was found in a tight white pellet, and the DNA containing supernatant was transferred to a 1.5ml tube

containing 600µl RT isopropanol while avoiding the carryover of protein precipitant. The isopropanol DNA mix was gently inverted until the DNA formed a visible white mass. Centrifugation at 16.000xg for 1min at RT was performed to pellet the DNA and the supernatant was carefully removed. After washing with 600µl of 70% EtOH and gentle inversion, the pellet was spun again and this time all liquid was carefully removed, and the pellet air-dried for 10-15min at RT. The dry pellet was resuspended in 50-100µl of DNA rehydration solution by incubation overnight at 4°C and subsequent mixing by pipetting before further usage.

7.1.5   PCR on genomic DNA / RCA

7.1.5.1   Phire tissue direct PCR master mix kit

PCRs using the phire tissue direct PCR master mix kit were performed in a reaction containing 1µl of 25-100ng/µl gDNA isolated or 10ng/µl plasmid DNA, 0.5µl of each primer in a 20µM dilution and 10µl of the 2x phire mix. The reaction was filled with nuclease-free water to 20µl and run using the following conditions: initial denaturation at 98°C for 5min, then 40 cycles of denaturation for 5sec at 98°C, annealing at the primer specific annealing temperature for 5sec, elongation at 72°C for 20sec/kb with 5sec and final elongation at 72°C for 1min. For the provided internal control primers, both annealing and elongation were performed for 20sec at 72°C as of manufacturers' recommendation. The reaction was then directly loaded on an agarose gel and the PCR bands were analysed (chapter 7.1.1.1).

7.1.5.2   Rolling circle amplification

To specifically amplify and prove the presence of circular, not integrated plasmid DNA, the rolling circle amplification using the TempliPhi amplification kit was used. This amplification works with an initial denaturation. Subsequently, 18h of annealing and amplification is happening, where circular templates are continuously amplified by the strand displacement ability of the polymerase, and newly amplified circular DNA serves as further template. Also, the polymerase is thought to have an excellent processivity and proof-reading ability, which makes it possible to even amplify hard to

PCR features like the CAG promoter. The kit was stored at -80°C and the components were always thawed and handled on ice. Usually, 20-100ng gDNA and 10ng plasmid DNA was used as input. For the reaction mix, first 5µl of sample buffer was aliquoted in PCR tubes per reaction. After the addition of 0.5-1µl sample DNA, the mixture was denatured for 3min at 95°C. First annealing occurred during the subsequent cooling to 4°C. Then, 5µl of the in the meantime prepared master mix consisting of 5µl reaction buffer and 0.2µl enzyme mix per reaction was added. The amplification reaction was performed at 30°C for 18h overnight, followed by a heat inactivation step for 10min at 65°C. The next day, 3-5µl of the RCA reaction was digested with a single or multiple cutters of the target vector for 1-2h using the previously described digestion protocol (7.1.1.4). Concurrent control digestion of plasmid DNA was performed, and all reactions were run on an agarose gel (chapter 7.1.1.1) to check the digestion patterns.

Usually, target plasmid DNA was used as input for positive control of the RCA reaction, water and gDNA of not modified cells were used as a negative control input. Initially, gDNA for this assay was obtained using the wizard genomic DNA extraction kit (chapter 7.1.4.3), however, gDNA isolated with the phire tissue direct PCR master mix kit also proved to work in this reaction (chapter 7.1.4.2).

7.1.6   Southern Blot

To provide evidence, that the DNA vectors were retained as episomal units and not integrated, southern blot experiments were performed. Up to 30µg of gDNA from the cell line of interest was isolated with the previously described techniques and digested with a single cutter of the utilised S/MAR DNA vector at 37°C for at least 4h. The digested gDNA was then separated on a 0.8% agarose gel at 15V overnight. 5ng of digested the utilised plasmid was run in another lane as control. The next day, digestion of gDNA was confirmed by visualisation of a smear. While gDNA was digested randomly and integrated fraction would appear at any hight, the unique cutter of the S/MAR DNA vector should result in a single band specific for the size of the S/MAR DNA vector.

The gel was incubated for 10min in depurination buffer consisting of 250mM HCl and then washed two times with MQ water. After two incubation steps for 15min in denaturation buffer containing 1.5M NaCl, 0.5M Trisodium citrate pH 7.0 and two equilibration steps in 20x SSC buffer (3M NaCl, 0.5M Trisodium citrate pH 7.0) for 10min, the DNA was transferred to a nylon membrane. For this step, the gel was placed on Whatmann paper which was soaked and inversed in a reservoir of 10x SSC buffer. The nylon membrane was also soaked in 10x SSC buffer and put on top of the gel while avoiding any air bubbles. The construction was continued by addition of more Whatmann paper as well as paper towels and a weight and incubated overnight was performed to ensure transfer of the DNA by capillary effect. The next day, the membrane was exposed to UV-B radiation for cross-linking of the DNA with the membrane. Subsequent hybridisation and was performed with radioactively labelled DNA probes generated based on the GFP gene by Dr Matthias Bozza and membrane was exposed for 3-4h up to several days.

7.1.7 Plasmid rescue

To provide evidence, that the DNA vectors were retained as episomal, circular units, rescue experiments were performed, and different protocols were tested. In general, gDNA was isolated with the different gDNA isolation methods with precipitation or purification via a mini prep column and with or without prior digestion with proteinase K. Then, the obtained gDNA was used to transform DH10β cells as described before (chapter 7.1.1.9). The bacteria was plated on antibiotic-free agarose plates, incubated overnight at 37°C and colonies were picked and further expanded in 5ml LB cultures containing the S/MAR DNA vector specific antibiotic. The DNA of these bacterial cultures was isolated, digested, and run on an agarose gel to confirm the digestion pattern of the respective S/MAR DNA vector.

METHODS

7.1.8 RNA isolation

7.1.8.1 RNeasy kit

RNA isolation was sometimes performed using the RNeasy kit from Qiagen. RNA isolation was performed as of manufacturers' recommendation. Briefly, cells were harvested, washed with PBS and pelleted. The cell pellet was resuspended and lysed in 350μl RLT buffer and briefly stored at -20°C overnight. Then, the working area was sprayed with RNaseOUT and the tube thawed at RT. All following centrifugations were performed at 8.000xg for 15sec. The thawed lysate was mixed with 350μl 70% EtOH, mixed by pipetting and transferred on a RNeasy mini spin column. RNA was bound to the column by centrifugation. Then, a on column DNA digestion step was performed using the RNase-free DNase set. First, the bound RNA was washed using the buffer RW1 and spinning. Then, a master mix of 70μl buffer RDD and 10μl DNase I per sample was prepared, 80μl of this solution added on top of the column and incubated at RT for 15min. The DNA digestion was stopped and washed away by addition of 350μl buffer RW1 and another centrifugation step. Then, two rounds of washing with each 500μl buffer RPE were performed, and the column spun empty for 2min to remove any remaining buffer. For elution of the RNA, the column was placed in a new collection tube, 30-50μl nuclease-free water was added, incubated for 1min at RT and then spun for 1min. The obtained RNA concentration and quality was measured with a nanodrop and used for subsequent cDNA synthesis.

7.1.8.2 Trizol isolation

Since some cells like UOK257 cells have little amount of RNA, RNA isolation was switched to a protocol using trizol. Here, cells were also harvested, washed and pelleted, but then lysed and resuspended in 500μl trizol, incubated for 3-5min at RT and stored -80°C until further usage. Upon RNA isolation, the lysate was thawed on a RNase free workspace and mixed vigorously with 100μl chloroform. The mixture was incubated for 2-3min at RT and then spun at 4°C for 15min at 12.000xg, resulting in a lower red organic phase, an interphase and an upper, colourless aqueous phase. This RNA containing upper aqueous phase was carefully transferred to a new tube. Then, the RNA was precipitated by the addition of 250μl isopropanol and incubation at -80°C

for 1-2h. The precipitate was pelleted at 4°C for 10min at 12.000xg, the supernatant discarded, and the pellet washed with 500µl 75% EtOH made with nuclease-free water. Another centrifugation step for 5min at 4°C at 7.500xg allowed subsequent careful aspiration of all supernatant. The pellet was dried on ice for 5min with an open lid and then resuspended in 20µl nuclease-free water by 5min incubation on ice and pipetting.

For digestion of remaining DNA in the sample, the DNA-free DNA removal kit (Invitrogen) was used. Since a maximum concentration of 200ng/µl was required for efficient DNA digestion, sample concentration was measure with a nanodrop and diluted if required. 2µl of 10x DNase I buffer and 1µl rDNase I were added to 20µl sample and gently mixed. After incubation at 37°C for 20-30min the digestion was stopped by the addition of 2.3µl of vortexed DNase inactivation reagent and vortexing. After another 2min incubation at RT with mixing in between, the inactivation was pelleted using 1.5min centrifugation at 10.000xg at RT. The final supernatant containing the DNase treated RNA was transferred to a new tube, the concentration and quality was determined using a nanodrop and the RNA was subsequently used for cDNA synthesis as described in the following section (chapter 7.1.9).

7.1.9   cDNA synthesis

For the synthesis of cDNA from isolated RNA (chapter 7.1.8), the RevertAid H minus first strand cDNA synthesis kit was used. Usually, 1µg of input RNA was used per sample and mixed with 1µl of an equal mixture of the provided oligo dTs and random hexamers. The reaction was filled with nuclease-free water to 12µl and incubated at 65°C for 5min. After this initial denaturation and annealing, the samples were cooled down and handled on ice. In the meantime, a master mix containing 4µl of 5x reaction buffer, 1µl RiboLock RNase inhibitor, 2µl of the 10mM dNTP mix was prepared. For the negative control without reverse transfection (noRT control), to check for remaining DNA in samples, 7µl of this master mix and 1µl of water was mixed to this sample. For the remaining samples, 1µl of RevertAid H Minus M-MuLV Reverse Transcriptase per

sample was added to the master mix and 8µl of it was then pipetted to each sample, resulting in a final volume of 20µl. The reverse transcription reaction was then performed using the following parameters: 5min 25°C, 60min 42°C, 10min 72°C and cooling to 4°C. Since 1µg of input RNA is supposed to yield 1µg of cDNA, the sample concentration of the 20µl reaction was assumed to be 50ng/µl cDNA. The cDNA was stored at -20°C until usage for qPCR.

7.1.10 qPCR

7.1.10.1 Primer design

To design suitable primers for qPCR, first cDNA and gDNA from the respective gene of interest was collected from Ensembl or GenBank and saved in SnapGene files. Next, exons and introns were labelled in both files. Then, the IDT PrimerQuest tool was used to design appropriate primers. After inserting the cDNA sequence, in the "custom design parameters" the primer Tm at the "primer criteria" was changed to min. 62°C, opt. 65°C and max. 67°C. To get optimal primer lengths for assays with sybr green, the amplicon size in the "amplicon criteria" was adjusted to min. 100bp, opt. 120bp and max. 150bp. As the last step, in the "custom target region", under "design across junctions" an overlap junction was chosen by insertion of the bp number of the cDNA, where an exon junction in one of the last exons of the protein occurred. Using these parameters, several primer pairs were suggested. These were then blasted using the direct primer blast link and controlled *in silico* for unspecific binding in *homo sapiens* by checking the E values. The most suitable primer pair was then inserted in the cDNA and gDNA SnapGene file to control again correct binding, before ordering the primers from sigma aldrich. With these design settings, all designed primers exhibit similar annealing temperature, with the optimal length for sybr green assays and similar elongation times, allowing the usage of all primers on one qPCR assay. Furthermore, the exon-spanning primer design excluded possible gDNA amplification and thus signal from residual gDNA in the qPCR data. Upon usage, lyophilised primers were resuspended in nuclease-free water to 100µM stock. From these, 5µM working concentrations were done, aliquoted in 50µl-150µl aliquots in PCR tubes and stored at

-20°C for single usage to avoid cross-contamination and influence by repeated freezing and thawing.

7.1.10.2 Sample preparation and plating

All samples were handled on ice throughout the pipetting process and under a PCR hood. Primer efficiencies for each primer pair were determined by making six sequential 1:10 dilutions of a 50ng/µl cDNA sample (chapter 7.1.9) which was thought to express the protein to which the primer binds. Samples to test, including the noRT control, were diluted to 5ng/µl by the addition of 180µl of nuclease-free water. The plate layout was designed in a 96-well format in a way, that technical triplicates were obtained. Usually, water, noRT control or the cDNA of a sample was distributed along one row, with different rows containing different cDNA or control inputs. Primers were then distributed column-wise, with three columns containing the same primer to enable technical triplicates. After successful plate design, the required wells per primer pair plus three additional wells were calculated and a master mix was performed. In this master mix, per well 1µl of each forward and reverse 5µM primer was mixed with 12.5µl primaQUANT 2x qPCR-CYBR-green-BLUE MasterMix and 8.5µl nuclease-free water, resulting in a total of 23µl per well. Then, 2µl of each prepared control or cDNA input sample was plated according to the plating scheme using a multistep pipette, always pipetting in the same corner of the plate. Then, the plate was turned 90° and 23µl of the primer master mix was added to each well according to the plating scheme also using a multistep pipette and pipetting in the other corner of the well. The tip was changed whenever new master mix was aspired, to avoid cross-contamination. Finally, the plate was sealed, spun for 1min at 1.000rpm and inserted into the LightCycler96 instrument. The 25µl reaction with 0.2µM primer and 10ng sample cDNA was then performed using the following program: 1) preincubation: 95°C for 10min, acquisition mode none; 2) 40 cycles of 3 step amplification: 95°C for 10sec, acquisition mode none; 60°C for 20sec, acquisition mode none; 72°C for 30sec, acquisition mode single; 3) melting: 98°C for 60sec, acquisition mode none; 60°C for 1sec, acquisition mode none; 98°C for 1sec, acquisition mode continuous. Data was analysed using the LightCycler96 software, Excel and Graphpad Prism. To further analyse unspecific

products, sometimes the technical triplicates were pooled and run on a 2% agarose gel (chapter 7.1.1.1).

7.1.11 Protein isolation and Western Blot

7.1.11.1 Protein isolation

For protein isolation, cells were harvested using the appropriate method. Then, cells were spun and washed with PBS. After pelleting, the cells were resuspended in 40-100µl lysis buffer, depending on the pellet size. As lysis buffer either triton lysis buffer consisting of 1x TBS with 1% triton X-100 or RIPA, both supplemented with 1:100 HALT protease and phosphatase inhibitor. The lysate was kept on ice and incubated there for 30min with mixing in between. Then, cellular debris were pelleted by centrifugation at 4°C for 10min max. speed. The supernatant was transferred to a new 1.5 tube and this protein lysate was stored at -20°C until further use.

7.1.11.2 Protein concentration determination using BCA assay

The concentration of protein lysates was determined using the Pierce BCA protein assay kit. Standard dilutions of provided 2mg/ml BSA solution was prepared in water or lysis buffer without HALT and stored at -20°C. A master mix of 196µl reagent A and 4µl reagent B per well was prepared and 200µl was aliquoted in a 96-well flat bottom plate. Then, either 1µl of a standard or sample was added per well and mixed, or 25µl of standard and 1µl of sample mixed with 24µl lysis buffer was mixed per well, dependent on how concentrated the samples were. Then, the plate was incubated at 37°C for 30min. During this time, a formation of a purple-coloured product correlation to the amount of protein in the sample was observed and its absorbance at 562nm was measured using a multiskan go plate reader (Thermo fisher scientific). The absorption data was plotted in an Excel sheet and using the standards and a linear regression, the concentration of the lysates was calculated. If appropriate, the 1:25 dilution factor of the lysate sample was included.

7.1.11.3 SDS-PAGE

After determining the concentration (chapter 7.1.11.2) of the protein samples (chapter 7.1.11.1), the amount of each sample required to load 20-40µg of protein was calculated. Then, this amount of lysate was mixed with the appropriate amount of 4x loading buffer and boiled for 5min at 95°C. After a quick spinning the samples were ready to load on a SDS-PAGE. Therefore, precast Mini-PROTEAN gels 4-20% were assembled in an electrophoresis chamber and 1x TGS was used as running buffer. The samples and the pre-stained protein ladder were loaded on the gel and run at 80-130V for 1-2h until the required protein separation was observed.

7.1.11.4 Transfer and protein detection

After appropriate separation, the SDS-PAGE was stopped and the gel taken out of the plastic plates, the stacking gel including the wells was removed and handled in a box with water. To transfer total protein from the gel onto a PVDF or nitrocellulose membrane, a commercial transfer stack was prepared assembled with the gel as of manufacturers' recommendation in an iBlot2 device. The transfer took place with the program P3, using 20V for 7min. Then, the membrane was cut with a scalpel and blocked for 1h at RT in 5% milk in TBST. Sometimes, ponceau staining was performed in between: the membrane was incubated for 2min in ponceau solution, then gently washed in water until the protein bands were nicely visible. Before blocking the membrane was completely destained using water or TBST. After blocking, the membrane was incubated overnight at 4°C with 4ml of the appropriate antibody dilution in 5% milk in TBST in a 50ml tube with rotation. The next day, the membrane was washed three times with TBST for a total of 1h and subsequently incubated for 1h shaking at RT in the respective dilution of secondary antibody conjugated to HRP in 5% milk in TBST. After another three washing steps in TBST for at least 1h, the membrane was developed by addition of ECL reagent and recorded using a FusionSL system. If needed, the membrane was stripped subsequently, blocked for 1h, and the process was repeated with another antibody combination. Antibodies used in this study and their dilution are listed in chapter 6.8.

7.2   Cell culture methods

7.2.1   General cell culture methods

7.2.1.1   Cell counting using Luna cell counter

For general cell counting, 10µl of cell suspension was mixed with 10µl of trypan blue reagent and approximately 10µl was transferred to a single-use Luna cell counting slide. Cell concentration was determined with brightfield counting using the LUNA automated cell counter. The concentration of live cell counts was used for further calculations.

7.2.1.2   Cell counting using Neubauer chamber

For hiPSCs, automatic counting resulted in unreliable concentrations, possibly due to different cell size and often observed small cell clumps. Thus, hiPSCs were counted manually using a Neubauer chamber. As for automatic counting, 10µl of cell suspension was mixed with 10µl of trypan blue reagent. 10µl of this mixture was pipetted on a glass Neubauer chamber. Cells in all four grid squares of one chamber were counted and cell concentration was determined as follows: The cell number of the four grid squares were summed and divided by 4. Next, multiplication by 2 was performed to include the dilution factor of 1:1 cell suspension with trypan blue reagent. The resulting number was multiplied by $1x10^4$ to obtain the cell concentration measured in cells/ml.

7.2.1.3   Gelatine coating

Cell culture ware sometimes was coated with gelatine to support cell adherence. For this, a 0.1% (w/v) gelatine solution was made by mixing 0.1g gelatine powder per 100ml MQ water and was dissolved and sterilised by autoclaving. Cell culture ware was covered in this solution and incubated for at least 20min at 37°C before removal of excess gelatine by aspiration and plating of the respective cells.

7.2.1.4   Feeder preparation

Immortal mouse embryonic fibroblasts (iMEFs) were used as feeder cells for stem cell culturing and prepared as follows: iMEFs were cultured and expanded on gelatine coated plates with DMEM supplemented with 10% FCS, 1% p/s and 1% NEAA. Cells from 1-4 confluent T125 were trypsinised, pooled and irradiated with 60Gy of γ-irradiation (Gamma cell 1000). Subsequently, irradiated iMEFs were counted and either 1.5×10^6 cells were distributed equally per gelatine coated multi-well plate. If not needed immediately and for subsequent experiments, 2×10^6 irradiated iMEFs were frozen per tube, and after thawing one vial was used per multi-well plate.

7.2.1.5   Mycoplasma contamination test

All cultured cells were tested for mycoplasma contamination using the applied genomics mycoplasma test service of eurofins genomics. For this, 500µl of media of 80% confluent cells which were cultured for at least 2 weeks and without media change for at least 3 days was collected in a 1.5ml tube. After heating for 10min at 95°C, the sample was spun at 13.000rpm for 5sec to remove debris. 200µl of the supernatant was transferred to a new 1.5ml tube, labelled with a barcode and sent to eurofins for analysis.

All tested cancer cell lines, as well as primary cells used in these studies were tested negative for mycoplasma contamination. However, the iMEFs used for feeder production were tested positive. Since the first stem cells were all reprogrammed on these feeder cells, it has to be assumed, that the resulting iPSCs were also contaminated. We did not observe any impact of the contamination on the handling or production of stable cell lines. With later experiments, the system was switched to feeder-free reprogramming, avoiding the use of these iMEFs and routine mycoplasma contamination checks were performed.

7.2.1.6   Flow cytometry analysis and cell sorting

To analyse GFP expression in a quantitative way as a single readout or over time, flow cytometry analysis was performed using a BD LSR Fortessa. To purify the GFP positive

cell population e.g., during vector establishment, fluorescence activated cell sorting (FACS) was performed using a BD FACSAria Fusion. In both cases, cells were harvested with the appropriate method, washed with PBS and pelleted. Cells were resuspended in PBS mixed with 1:100 propidium iodide (PI) to allow dead-live distinguishing and filtered through a filter FACS tube. Samples were handled on ice to avoid clumping of cells. Analysis of the data was performed using FlowJo. Usually, first, the cell population was gated for in a FSC-A vs SSC-A plot. Then, single cells of the chosen population were gated in a SSC-A vs SSC-W plot. Dead cells were excluded by choosing the PI negative population in a FSC-A vs PI plot, and GFP expression of this final population was used as final data and depicted as appropriate. As gating controls, untreated cells resuspended in PBS alone or with PI were used.

7.2.2   Routine maintenance of cancer cell lines

If not stated otherwise, all cells were grown in a humidified cell incubator at 37°C with 5% CO_2. An overview of the used cancer cell lines and their media are provided in chapter 6.1. In general, cancer cell lines were passaged upon 80-90% confluency, which was usually two to three times per week. To do so, media was aspired, and cells were washed with PBS. Cell detachment was performed using trypsin and incubation at 37°C for 5min. Upon successful detachment of cells, which was checked under the microscope and by gently tapping against the cell culture plasticware, trypsin was inactivated with FCS containing cell type specific media. Then, cells were resuspended by pipetting up and down until a single cell suspension was obtained. Cell line specific cell numbers or percentages of the cell suspension were plated in the appropriate cell culture plasticware. For fast growing cells like HEK293T, HeLa or U2OS cells, routine maintenance was done without cell counting and ~1/10th of the obtained single-cell suspension was plated into the same size of cell culture plasticware. For UOK257 and FTC113 cells the initial plating density was higher, and 3-5/10th of the cell suspension was plated for routine maintenance.

7.2.3 Chemical transfection of cancer cell lines

7.2.3.1 JetPEI

Easy transfectable cancer cell lines like HEK293T and U2OS were transfected using chemical transfection with JetPEI. On day 0, cells were plated in 6-well wells in 2ml media in a fashion, that they reached 60-70% confluence on day 1. The plating density had to be adjusted for each cell line, for example for HEK293T cells this required plating of approximately $4x10^5$ cells per well, while only $1x10^5$ U2OS cells but $7x10^5$ H1703 cells were plated per 6-well well. The next day, the transfection mix was prepared: 1µg of DNA was diluted in 50µl NaCl provided from the JetPEI kit. In another tube, 2µl JetPEI reagent was diluted in 50µl NaCl. Both dilutions were briefly vortexed and quick spun. Then, the diluted JetPEI reagent was added to the diluted DNA. The mix was immediately briefly vortexed and quick spun and incubated at RT for 15min. When more than one well was transfected, a master mix of diluted JetPEI reagent and diluted DNA was made if possible. After the incubation time, the formed DNA-JetPEI complexes were added to the plated cells by dropwise addition of the mixture to the cells. By gentle shaking of the plate, equal distribution in the cell media and the cells was achieved before putting the cells back into the incubator. For this reagent, antibiotics in the media did not alter the transfection efficiency.

7.2.3.2 XtremeGene9

For transfection using XtremeGene9, cells were plated in a similar manner as described for JetPEI transfection before (chapter 7.2.3.1). Notably, using XtremeGene9, the transfection was hampered using antibiotic containing media. Thus, cells were either plated in media lacking antibiotics, or media was changed to antibiotic free media on day 1 before addition of the transfection mix.

For transfections in a 6-wp (well plate) format, 6µl of XtremeGene9 was diluted in 94µl optimem, briefly vortexed and quick spun and incubated at RT for 5min. In the meantime, 2µg plasmid DNA was diluted in 100µl optimem, briefly vortexed and quick spun. After the incubation time, the diluted XtremeGene9 mixture was added to the diluted DNA and the mix was also vortexed and quick spun. In general, a

DNA : XtremeGene9 ratio of 1:3 was used. Whenever possible, a master mix of dilutions was performed when transfecting several wells at the same time. After another incubation at RT for 40min, the complete 200µl mixture was added dropwise to the prepared cells.

7.2.3.3 XtremeGeneHP

XtremeGeneHP is a more potent transfection reagent than XtremeGene9, specifically designed for hard to transfect cells and was used for the transfection of UOK257 cells. XtremeGeneHP was stored at -20°C and had to be adjusted to RT before usage. The transfection with this reagent was performed in a similar manner to transfections with XtemeGene9 (chapter 7.2.3.2). Again, antibiotics in the media hampered transfection efficiency, thus cells were cultured in antibiotic deficient media at least during the time of transfection. Also, a DNA : XtremeGeneHP ratio of 1:3 was used, resulting in dilution of 2µg DNA and 6µl XtremeGeneHP in each time 100µl optimem. Dilutions were performed as described for XtremeGene9, however, the mixture of the XtremeGeneHP and DNA dilutions was only incubated for 15min at RT before addition to the cells.

7.2.4 Generation of stable S/MAR DNA vector modified cancer cell lines
7.2.4.1 Via antibiotic selection

For the generation of stable S/MAR modified cancer cell lines with a S/MAR DNA vector containing a puromycin resistance gene in the mammalian expression cassette, stable cell lines were often obtained using antibiotic selection. Therefore, cells were transfected with the S/MAR DNA vector using the appropriate transfection reagent as described before (chapter 7.2.3). Successful transfection was verified 24h later using fluorescence microscopy. If the S/MAR DNA vector did not contain a fluorescent reporter gene, an additional sample using a GFP encoding S/MAR construct was transfected with the same method to verify successful transfection. At least 24h after the plasmid delivery, the cell culture media was changed to fresh media containing 0.5-1µg/ml puromycin. Cells were cultured in antibiotic containing media for one month, then S/MAR DNA vectors were assumed to be established. If the cells were

very sensitive to the antibiotic selection, cells were plated in antibiotic free medium upon expansion and puromycin was added after the cells had attached. Also, if needed, a break of antibiotic selection pressure was done until cell count had recovered.

7.2.4.2   Via FACS sorting

To generate S/MAR modified stable cell lines in a gentler way without using antibiotic selection, mammalian expression cassettes containing GFP as reporter gene were used. After delivery of the respective S/MAR DNA vector, successful transfection was verified using fluorescence microscopy. To get rid of cells that were not transfected, or which lost expression of the S/MAR DNA vectors, cells were sorted using FACS (chapter 7.2.1.6). Usually, sorting in the first 48h after transfection resulted in very low viability. A first significant loss of expression was reproducibly observed throughout different cell lines around 7-11 days post transfection (dpt). Thus, to increase the percentage of establishing cells, a first FACS was performed around 7-11 dpt. Then, cells were expanded and kept in culture until another sorting of GFP positive cells between 20-30dpt. In between, cells were monitored using fluorescence microscopy or flow cytometry analysis. To verify the successful establishment and maintenance of S/MAR DNA vector expression, cells were continued to be monitored after the establishment period.

7.2.5   Handling of primary cells

7.2.5.1   NHDF expansion

Normal human dermal fibroblasts (NHDFs) were purchased by either promocell (C-12300) or thermo fisher (C0045C). Upon receiving, cells were thawed and plated on gelatine coated cell culture flasks and cultured using DMEM supplemented with 10% FCS, 1% p/s and 1% NEAA at 37°C and 5% CO_2. Since cells are required to be at a low passage number for successful reprogramming, they were expanded for a maximum of 4 passages using trypsin and aliquots of $5x10^5$ cells were frozen and separately used for experiments.

7.2.5.2 Urinary cell isolation

Cells derived from the urine were isolated and used for reprogramming as described in (Mulder et al., 2020; T. Zhou et al., 2012; T. Zhou et al., 2011). Ethical approval to work with these cells was granted by the ethics committee of the medical faculty of the University of Heidelberg and registered under the study number S-550/2019.

When possible, urinary samples were processed immediately after obtention or stored at 4°C for a maximum of 4h if not stated otherwise. To isolate cells derived from the urine (USCs), urine samples were distributed into 50ml tubes and spun at 400xg for 10min. After careful removal of the supernatant, pellets were resuspended in PBS and pooled in one 50ml tube with 40-50ml total volume. Cells were pelleted again at 400xg for 10min and all PBS except for ~200µl was carefully removed. Cells were resuspended in 1ml of urinary cell media (USC media), plated on a gelatinised 12-well well and incubated at 37°C and 5% CO_2.

In the first attempts of USC isolation, different media were tested based on (T. Zhou et al., 2011), as described further in 8.4.2 and 8.6.2. However, best isolation efficiency was achieved with the following protocol and media based on (Mulder et al., 2020).

In order to aid attachment, the plates were not moved during the first 24h of culturing. Only 24h post isolation (day 1), 1ml of fresh primary USC media was gently added to the well. This step was repeated on day 2 and 3. Stock primary USC media was composed of DMEM/F12, supplemented with 10% FBS and 1% p/s. Before usage, 50ml of the stock media was aliquoted and further supplemented with 50µl of each supplement of the REGM SingleQuots (hEGF, insulin, hydrocortisone, GA-1000, transferrin, triiodothyronine and epinephrine), except for the provided FBS, of which 250µl was added. This complete media was used as primary USC media and stable at 4°C for 2 weeks.

On day 4 the media was carefully removed and 1ml of secondary USC media was added. Secondary USC media consisted of 50ml aliquots of REBM media,

supplemented only with 1% p/s, and again 50µl of each supplement of the REGM SingleQuots (hEGF, insulin, hydrocortisone, GA-1000, transferrin, triiodothyronine and epinephrine), except for the provided FBS, of which 250µl was added. This complete media was also stable at 4°C for two weeks. The next days, half of the culture media was gently removed every second day and supplemented with the same amount of secondary USC media. First colonies were usually observed within 5-9 days post isolation. As soon as the 12-well well was confluent, cells were redistributed on a fresh gelatinised 12-well well. Passaging of USCs was performed by removal of the media, washing with PBS and detachment using TrypLE for 5min at 37°C. TrypLE was inactivated by dilution with secondary USC media. Upon 80-90% confluence, cells were expanded to a gelatinised 6-well well and subsequently a gelatinised 10cm dish. From passaging on, half of the media was replaced with fresh media as described only three times a week instead of every second day. After two weeks of culturing, cells were tested for mycoplasma contamination.

7.2.6   Reprogramming of primary cells

7.2.6.1   Delivery of EBNA reprogramming vectors to NHDFs

For reprogramming, passage 4 commercial NHDFs were thawed and cultured as described before. After one passage, cells were harvested with trypsin and trypsinisation was stopped using DMEM supplemented with 10% FCS and 1% NEAA. After pelleting the cells at 220xg for 5min, they were resuspended in PBS and counted. 5×10^5 cells per transfection were aliquoted in a 1.5ml tube. Cells were again spun at 220xg for 5min and all liquid was carefully removed. In the meantime, 90µl of nucleofector solution and 20µl supplement from the Lonza human dermal fibroblast nucleofector kit was mixed per reaction. 2µg of each of the four episomal EBNA reprogramming plasmids (pCXLE-hSK, pCXLE-hUL, pCXLE-hOCTshp53 and pCXLE-EBNA1, in the following referred to as EBNA-1-4) was added per reaction mix. The cell pellet was resuspended in the plasmid containing reaction mix, transferred to a cuvette and nucleofected with the Amaxa II nucleofector using program P-022. Subsequently, the cell solution was gently transferred from the cuvette to a prepared

gelatine coated 6-well well containing prewarmed DMEM containing 10% FCS and 1% NEAA but no antibiotics for the first 24h. Cells were cultured in NHDF conditions until day 8 post reprogramming factor delivery. When reaching confluency, cells were passaged like NHDFs using trypsin and plated on gelatine coated dishes with the NHDF media. If not passaged, media was replaced with fresh media every third day.

7.2.6.2   Delivery of reprogramming factors to USCs

7.2.6.2.1   Reprogramming of USCs using Lentivirus

To deliver reprogramming factors to USCs, different methods were tried. First obtained USC derived iPSCs were reprogrammed via transduction of USCs with a lentiviral 4-in-1 reprogramming vector (introduced in chapter 5.8.1, (Warlich et al., 2011). The vector was kindly produced by the members of the virus-associated carcinogenesis department (F170, DKFZ). Briefly, HEK293T cells were transfected with the three individual plasmids encoding for the lentiviral particles and the lentiviral vector pRRL.PPT.SF.hOKSM-IRES-dTom. The supernatant containing the produced virus was harvested, filtered, aliquoted and stored at -80°C until further use. Viral load was not measured, instead, a three-step transduction protocol was performed, and if needed, dilution of viral stock with media was performed to obtain the best concentration of the viral stock solution.

Transduction was performed in a 24-well plate format. Cells were plated on a gelatinised 24-well plate to achieve a 70% confluency on day 1, which required ~20.000 USCs per well. The next morning, media was aspirated and replaced with viral media. This was repeated in the evening and the next morning. The following evening, viral media was removed, and USC media was added. Successful transduction was verified by the observation of the dTom reporter gene expression under a fluorescence microscope. Cells were cultured as USCs (chapter 7.2.5.2) for the next days until day 8 post transduction with half media change three times a week, and expansion using TrypLE and plating on gelatine coated cell culture ware when required.

7.2.6.2.2 Reprogramming of USCs using EBNA vectors

Upon optimisation of USC isolation and handling, media components changed and allowed efficient transfection using lipofectamine stem or neon transfection system. Different conditions of Lipofectamine stem and plasmid concentration, as well as media and coatings, were tried with USCs. Although Lipofectamine stem seemed to be the only efficient delivery to USCs isolated with the first media, the conditions proved to be inconsistent across USCs from different isolations and media compositions.

While electroporation using the neon transfection system resulted in cell death with the first USC isolations, with the optimised isolation protocol and media as described (chapter 7.2.5.2), electroporation with this device resulted in easy, consistent, and efficient delivery and was used for a gentler reprogramming using EBNA vectors. For this, first, the electroporation system was prepared under a cell culture laminar flow hood by pipetting 3ml of electrolytic buffer E2 into the neon tube and inserting it into the pipette station. Also, a 100µl neon tip was properly attached to the neon pipette and 1.5ml tubes with sterile water, PBS and 70% EtOH for washing of the tip were prepared. Then, USCs were detached as described before using TrypLE, washed with PBS and counted. 1×10^6 cells were aliquoted and spun at 200xg for 5min. In the meantime, 110µl buffer R was mixed with 1.5µg of each EBNA plasmid (EBNA-1-4) per sample. After spinning, the supernatant was carefully removed, and cells were resuspended in the vortexed and briefly spun mixture of buffer R and DNA. The obtained cell suspension-DNA mix was aspired with the 100µl neon tip and the neon pipette while making sure that no air bubble was in the tip. The pipette was then inserted into the neon tube in the pipette station, and the cells were electroporated at 1.400V with 2 pulses with 20ms pulse width. Immediately after the electroporation, cells were gently pipetted into a prepared gelatinized 6-well well with USC media deficient of antibiotics and cultured in a cell culture incubator. For further electroporations with the same cells and DNA, the same neon tip was used. Therefore, the tip was washed by multiple pipetting of the prepared reagents: first, PBS was used to remove possible remaining cells in the tip. Next, reminiscent DNA was removed using the 70% EtOH solution, and finally, the tip was washed in water to remove traces

of EtOH. Then, the tip was either used for aspiration of the next cell-DNA mixture or was stored in a sterile tube at RT until further use. The electrolytic buffer in the neon tube was reused for the same cells and DNA electroporation batch. Upon finished electroporation of all samples, the buffer was aspired, and the tube washed with PBS. The neon tube was then stored together with the neon tip in a sterile 50ml tube at RT until the next usage. After delivery of the reprogramming factors, cells were maintained and handled as USCs with routine USC culturing conditions (chapter 7.2.5.2) and expanded, if necessary, until day 8 past delivery.

7.2.6.3   Reprogramming in a feeder dependent or independent manner

In general, during the first 7 days after delivery of reprogramming factors, cells were handled as their original primary cell type. The successful start of reprogramming was indicated by a change in morphology of the cells during these days: the wide and elongated morphology of the primary cells was changed in dense, colony-like growing cells with a narrow, compact, and round cell shape. On day 8, cells were detached using trypsin (NHDFs) or TrypLE (USCs), counted and 30.000 cells were plated per well of a 6-wp or 15.000 cells per well of a 12-wp in stem cell coated plates (gelatine coating and feeder cells or iMatrix Laminin-511, respectively), but with the respective media of the primary cell. Remaining cells were either frozen (chapter 7.2.8) and could be thawed at a later point, assuming day 8 of reprogramming and directly plated on stem cell conditions or lysed and used for Western Blot (WB) control for reprogramming factor expression (chapter 7.1.11). Usually, the 6-wp format was used to obtain stem cells for further expansion, and 12-well plates were used for AP staining to get an estimation of the efficiency and success of reprogramming.

Thus, for feeder dependent reprogramming, appropriate feeder cell destination plates were prepared on day 7 post reprogramming factor delivery. A detailed description of feeder cell preparation was described before (chapter 7.2.1.4). Briefly, the appropriate plate was coated with gelatine, and fresh or frozen feeder cells were plated and cultured in feeder cell conditions until the next day.

For feeder independent reprogramming, destination plates were coated with the stem cell matrix iMatrix Laminin-511 if not otherwise stated before plating of primary cells on day 8. Detailed coating description is described elsewhere (chapter 7.2.7.2.1).

From day 9 on, after the reprogramming cells attached to the stem cell culture plates, cells were completely treated as stem cells: primary cell media was removed and the respective media for feeder dependent or independent hiPSC culturing, respectively, was added (chapter 7.2.7.1.1 and 7.2.7.2.1). The appropriate media was changed at least three times a week throughout the whole reprogramming process, with increased feeding over two days of no media change. First colonies usually appeared around day 22-28 of reprogramming.

7.2.6.4   Expansion of reprogrammed iPSCs

Most successfully reprogrammed iPSC colonies were usually obtained around day 30 of reprogramming, and each colony is thought to have originated from one single primary cell. Thus, colonies were either picked and expanded in a clonal or polyclonal manner, by either expanding one colony per well in a small well format, or several colonies in a bigger well format. Either way, destination plates were previously prepared by either plating of feeder cells (chapter 7.2.1.4) or coating with stem cell matrix (chapter 7.2.7.2.1) as described before. Also, the appropriate stem cell media was added to the destination plates (chapter 7.2.7.1.1 and 7.2.7.2.1). Next, distinct iPSC colonies were manually scraped off the reprogramming plate using a 200μl pipette tip and gently transferred to a new destination plate. This was at the beginning performed using a disinfected, stereoscopic microscope under the laminar flow hood or with closed door alone in the open air under a compact cell culture microscope. Later, the Evos XL core imaging system was used under the laminar flow hood, which allowed detailed and exact iPSC colony picking as with the normal cell culture microscope but within a sterile environment. For feeder-free culture conditions ROCK inhibitor was added to the stem cell media of the destination plate, even though cells were obtained as a mixture of single cells but also cell clumps.

7.2.7   Handling of human iPSCs

7.2.7.1   Culturing on feeder cells

7.2.7.1.1   Culture conditions of iPSCs on feeder cells

Initially, human iPSCs were cultured on a feeder layer consisting of gelatine coated plates and irradiated iMEFs as described before (chapter 7.2.1.4). With these conditions, the stem cell media consisted of DMEM/F12 supplemented with 20% knockout serum replacement (KSR), 1% p/s, 1% NEAA and 0.1% β-mercaptoethanol. 50ml of this media stock was aliquoted, freshly supplemented with 50µl of 10µg/ml FGF2 to achieve 10ng/ml final concentration and used within two weeks. Media was changed every other day.

7.2.7.1.2   Passaging of iPSCs on feeder cells

For routine passaging, one to three days before passaging a fresh feeder plate was prepared in the appropriate cell culture plate (chapter 7.2.1.4). Upon passaging, the media was changed to fresh feeder dependent stem cell media in both the feeder coated destination plate as well as the plate with the stem cells. Then, distinct, undifferentiated and compact stem cell colonies were manually scraped off from the plate using a 200µl pipette tip. This was at the beginning performed using a disinfected, stereoscopic microscope under the laminar flow hood or with closed door alone in the open air under a compact cell culture microscope. Later on, the Evos XL core cell imaging system was used under the laminar flow hood, which allowed detailed and exact iPSC colony picking as with the normal cell culture microscope but within a sterile environment. The media with the detached iPSC clumps was gently transferred to the prepared new feeder plate and incubated for 24h without further moving the plate to allow attachment of iPSC clumps. Usually, cells were passaged once a week.

7.2.7.2   Culturing in a feeder free system

7.2.7.2.1   Culture conditions of iPSCs in feeder free condition

Several feeder-free stem cell media and matrixes were tried when switching the hiPSC culture to a feeder-free system. Routinely, StemFit Basic02 media (Ajinomoto) in combination with iMatrix Laminin-511 (Amsbio) was used since it demonstrated the best culture maintenance for our stem cells.

StemFit Basic02 media was delivered in two frozen bottles, which were thawed at 4°C in the fridge overnight. Then, the 100ml supplement bottle B was added to the 400ml of bottle A. After thorough mixing, 50ml aliquots were prepared and stored at -20°C. Upon usage, the media aliquot was thawed at 4°C or RT and supplemented with 50μl of 10μg/ml or 30μl of 50μg/ml FGF2 and used within two weeks. The 50ml tube was wrapped in aluminium foil to protect the media from light, and the media was never warmed above RT.

Precoating of plates was performed by mixing PBS with the appropriate amount of iMatrix Laminin-511 and incubation of the culture vessel at 37°C for 1h. As of manufacturers' recommendation, the iMatrix Laminin-511 solution (0.5mg/ml) was to be used at 0.5μg/cm^2. Thus, the 9.6cm^2 well of a 6-well plate was coated with 2ml PBS and 9.6μl iMatrix Laminin-511. Respectively, 12-well wells were coated using 1ml PBS and 4.8μl iMatrix, 24-well well using 0.5ml PBS and 2.4μl iMatrix. However, 96-well microscopy glass-bottom plates were coated using 300μl PBS and 1.175μl iMatrix per well. Sometimes, precoating of culture plates was not necessary and premixing of stem cell suspension with iMatrix just before plating to uncoated cell culture plates was possible. In these cases, half the amount of iMatrix used for precoating was added to the cell suspension directly before plating.

7.2.7.2.2   Passaging of iPSCs in feeder-free system

In a feeder-free system, iPSCs were passaged either in clumps or as single cells. For clump passaging, the media was aspired, and cells were washed twice with PBS. Then, 1ml of ReLesR was added per 6-well well (or scaled appropriately) and incubated for

1-3min at RT. During this time, cells were monitored under a cell culture microscope, and the colonies started to get less tightly packed and soaked from the outside to the inside. As soon as the colonies had a uniform look at the border and the inner part of the colony, ReLesR was aspired and the cells were continued to incubate at RT until the colonies were broken into small clumps for further 5-9min. Then, the cells were quickly flushed with 1ml media and immediately shaken at 12.000xrpm for 2min on a plate shaker. Differentiated cells remained attached to the plate, while regular stem cell clumps were obtained. 100-400µl of the obtained cell suspension was transferred to a new plate. The plate was not moved for the first 24h of culturing to aid cell detachment and avoid clumping of cells in one area of the plate.

For single-cell passaging, cells were also washed twice with PBS and then incubated with 500µl StemPro accutase for 5min at 37°C. Cell detachment was checked under a cell culture microscope and tested by strong tapping against the culture plate. Then, cells were resuspended in the accutase solution, and accutase was inactivated by mixing with StemFit Basic02 media supplemented with 10µM ROCK inhibitor. After counting of the cells as described before (chapter 7.2.1.2), 50.000cells/well were plated in a 6-well plate and cultured for more than 24h in ROCK inhibitor-containing media. Also, after assuring equal distribution of the plated cells, the plate was not moved for the first 24h of incubation to enable evenly distributed attachment.

7.2.8   Freezing and thawing of mammalian cells

For long-term storage, mammalian cells were frozen. Therefore, freezing media for the respective cell lines was made. For cells like primary cells and cancer cell lines, which were cultured in FCS containing media, a freezing media (FM) consisting of FCS supplemented with 10% DMSO. For hiPSCs and UCs, which were cultured in media with knockout serum replacement (KSR) (chapter 7.2.7.1.1), feeder-free media (chapter 7.2.7.2.1), or more specific media with little FCS (chapter 7.2.5.2), respectively, the corresponding media was supplemented with 10% DMSO and used as FM. Approximately 1.5x the number of cells that were used for plating for routine

maintenance was aliquoted from a cell suspension, pelleted and resuspended in the respective FM. To keep the cells from long exposure to DMSO in a thawed situation, cells were immediately transferred into a CoolCell or wrapped in paper towels in a styrofoam box to aid gentle freezing and stored at -80°C at least over one night. Upon complete freezing, cells were transferred to a liquid nitrogen tank for long-term storage.

To get stored cells back into culture, the cryotube was removed from the liquid nitrogen tank and cells were thawed rapidly by holding and gently moving the cryotube in a water bath at 37°C. Upon almost complete thawing, the cryotube was thoroughly disinfected using 70% EtOH and taken under a cell culture laminar flow hood. The cell suspension was gently transferred to a 15ml tube and mixed with at least 5ml of the respective cell culture media to dilute the DMSO in the FM. The cryotube was rinsed once with cell culture media to avoid a loss of cells. The cells were pelleted using 220xg for 5min. Subsequently, the supernatant was aspired, the cells resuspended in the respective cell culture media and plated on the suitable cell culture ware.

Since stem cells are more sensible than cancer cell lines, they required a gentler thawing process. Therefore, cells were always pipetted very slowly and using 5ml plastic serological pipettes since their wide bore end causes less shearing stress to the cells. Furthermore, after transfer of the cells to a 15ml tube, the cell suspension was diluted with fresh ROCK inhibitor-containing stem cell media in a dropwise manner with gentle shaking of the tube to decrease osmotic shock of the cells. After pelleting with only 100xg for 4min and aspiration of DMSO containing supernatant, the cell pellet was gently loosened by tapping against the tube. Then, cells were only resuspended by slowly adding fresh ROCK inhibitor-containing stem cell media and gentle aspiration of the pellet. The mixture was immediately transferred to a prepared, coated cell culture plasticware and not moved for 24h upon immediate placement in a cell culture incubator.

METHODS

7.2.9   Transfection of human iPSCs

7.2.9.1   Lipofectamine stem

For transfection of feeder-free maintained hiPSCs, lipofectamine stem was used. Cells were prepared by plating clumps from ReLesR detachment (chapter 7.2.7.2.2) in a way, that a confluency of 50-70% was obtained one day after plating in 24-well plates. The next day, 2µl lipofectamine stem was diluted in 25µl optimem and 500ng plasmid DNA was diluted in another tube with 25µl optimem. If possible, master mixes of the dilutions were performed, and the lipofectamine dilution was aliquoted per reaction. After vortexing and quick spinning, the diluted DNA was added to the diluted lipofectamine stem, vortexed and quick spun and the mixture was incubated for 10min at RT. Then, the solution was added to the cells in a dropwise manner and put back into the incubator.

Alternatively, $5x10^4$ single cells were plated in precoated 24-well plates in ROCK inhibitor-containing media. The next day, more than 24h after plating, a transfection mix was prepared, this time by dilution of 1µl of lipofectamine stem and 500ng of plasmid DNA per 24-well well. Again, master mixes were performed when possible, and the DNA dilution was added to the lipofectamine dilution, vortexed and quick spun and incubated for 10min at RT. In the meantime, the media of the cells to transfect was changed to stem cell media without ROCK inhibitor. Then, the incubated transfection mix was added dropwise to the cells and the cells were put back into the incubator.

7.2.9.2   Maxcyte electroporator

An initial electroporation experiment with feeder-free cultured hiPSCs was performed using the Maxcyte electroporator. Cells were detached as single cells with accutase, washed with the provided Maxcyte buffer and resuspended in Maxcyte buffer to a concentration of $10x10^6$ cells/ml. 10µg or 15µg of DNA (5µg/µl concentrated) was aliquoted into 1.5ml tubes. The DNA was mixed with 50µl of the prepared cell suspension and the solution was transferred to an OC1002x PA. After electroporation with the Opt6 or Opt9 program, the cells were incubated in the incubator for 20min to

recover. Subsequently, cells were gently transferred to prepared iMatrix coated 6-well plates containing StemFit Basic02 media supplemented with ROCK inhibitor. Viability and transfection efficiency was measured 24h later by pooling attached and unattached cells and analysed via FACS with propidium iodide (PI) for dead/live staining.

7.2.10 Generation of stable S/MAR DNA vector modified iPSC lines

7.2.10.1 With antibiotic selection

To generate stably S/MAR DNA vector modified iPSCs, as with cancer stem cells, plasmid DNA with a puromycin resistance gene was used and delivered as described before (chapter 7.2.9). 24-48h post transfection, cells were selected using 0.5μg/ml puromycin. With these cells, cell growth had to be closely monitored during selection. Also, depending on the density of the cells, the selection had to be removed from time to time to avoid the death of all cells. Especially during the expansion of cells, antibiotics had to be taken away and added again at a later timepoint. Throughout the outgrowth of the cells for one month, cells were handled like this, until stable transgene expression was observed.

7.2.10.2 With FACS

To enable gentler, drug-free establishment of stably modified iPSCs, an establishment protocol using FACS was generated. Here, S/MAR DNA vectors encoding for GFP in the mammalian expression cassette were generated and used. After delivery of the plasmid DNA as described (chapter 7.2.9), cells were cultured for 7-10 days and expanded using ReLesR if necessary. Between day 7-10 post transfection, all cells were harvested as single cells and GFP positive population was directly sorted (chapter 7.2.1.6) into the appropriate precoated well containing media supplemented with ROCK inhibitor. Cells were continued to culture and expanded, and the sorting was repeated between day 22-30 post transfection. Then the S/MAR DNA vectors were presumed to be established. To verify this, as during the establishment period, cells were continued to be monitored using fluorescence microscopy and flow cytometry.

METHODS

To avoid possible contaminations during the sorting process, often stem cell media was supplemented with 1% p/s during the time of establishment.

7.2.11 AP staining

Staining of alkaline phosphatase (AP) was used to verify and quantify reprogramming efficiency. The expression of AP happens early during the reprogramming process to iPSCs and thus serves as an indicator of reprogramming success. The staining was an endpoint analysis and performed using the alkaline phosphatase staining kit II as of manufacturers' recommendation protected from light. First, PBST was made by the addition of 5µl tween20 to 10ml PBS and thorough mixing. Then, the total needed volume of AP substrate solution was calculated and divided by three. Equal volumes of the calculated amount of substrate A and B were mixed, incubated for 2min and substrate C was added. Then, the media of the cells to be tested was aspired, the cells were washed with PBST and the provided fixing solution was added to cover the well surface. After 5min incubation at RT, the fixing solution was aspired, and the well was washed with PBST. The PBST was exchanged with the prepared AP substrate solution and incubated at RT for 5 to 15min while monitoring the colour change of the staining. Upon successful staining or latest after 15min, the substrate was aspired, the wells washed twice with PBS and fresh PBS was added. Pictures of the staining were taken using a Keyence microscope. The following volumes to cover the surface were used: 90µl for 96-wp wells, 150µl for 24-wp wells, 300µl for 12-wp wells and 500µl for 6-wp wells.

7.2.12 Differentiation into cells of the three germ layers

7.2.12.1 Trilineage differentiation via embryoid bodies

To observe the capability of hiPSCs to differentiate randomly into cells of all three germ layers, a random differentiation assay via embryoid bodies (EBs) was performed. First, EBs were obtained by clumping hiPSCs together in suspension. Therefore, cells were plated in clumps in an ultra-low attachment (ULA) plate using stem cell media

supplemented with 10μM ROCK inhibitor (ROCKi). EB formation using hanging drops, where 600 single cells in 20μl drops were plated on the lid of a bacterial petri dish and cultured upside down, was also performed, however, it resulted in EB formation. On day 2 of the ULA EB formation, the media was carefully half replaced with fresh stem cell media containing ROCKi by tilting the plate and waiting for 5min until the building EBs had sedimented. On day 4, media was changed to EB media consisting of DMEM supplemented with 20% FCS to allow random differentiation through withdrawal of stem cell factors. On day 6, a 96-well glass-bottom plate was gelatinised and two to three EBs were plated in fresh EB media per well. The EBs were carefully moved to the middle of the well if possible, and then cultured for 24h without movement. On day 7, 50-100μl fresh EB media was carefully added per well without disturbing the attaching EBs. In the following days, the media was replaced or just carefully supplemented with some fresh EB media every other day. Cell differentiation was monitored, and usually, distinct differentiated structures were obtained around day 14-17. Then, differentiation into cells of the three germ layers was assayed by immunofluorescence staining as described elsewhere (chapter 7.2.13).

7.2.12.2 Trilineage differentiation kit

To achieve a more reliable and faster differentiation of hiPSCs into representatives of the three germ layers, the trilineage differentiation kit was used. First, 96-well glass-bottom plates were coated as described (chapter 7.2.7.2.1). In the meantime, hiPSCs were detached as single cells using accutase and counted as described (chapter 7.2.7.2.2 and 7.2.1.2). Then, cells were plated in stem cell media supplemented with ROCKi in the appropriate density as of manufacturers' recommendation adjusted to 96-well format: mesoderm: 40.000cells/cm^2 (2.24x10^4 cells per 96-well well); endoderm: 66.000cells/cm^2 (3.7x10^4 cells per 96-well well); ectoderm: 53.000cells/cm^2 (3.0x10^4 cells per 96-well well). For ectoderm, media was very carefully changed to EctoDiff media supplemented with ROCK after successful cell attachment a couple of hours after plating. The following six days, the media was changed in the respective wells according to manufacturers' instructions as follows: Mesoderm: media was changed to 300μl MesoDiff I on day 1, then media was changed on day 4, 5 and 6 with

each 200μl MesoDiff II. Endoderm: on day 2, 3, 4, 5 and 6, media was aspired, cells were washed with PBS containing Ca^{2+} and Mg^{2+} and 150μo of fresh EndoDiff media was added. Ectoderm: media was changed every day using 150μl EctoDiff. On day 7, cells were fixed, and germ layer specific protein expression was analysed using immunofluorescence staining as described in chapter 7.2.13.

For quantitative analysis regarding maintenance of GFP expression upon differentiation, the experiment was upscaled in a 48-well cell culture plate and control cells were handled simultaneously to compare GFP expression on day 7 via FACS analysis.

7.2.13 Immunofluorescence staining

Immunofluorescence staining was performed to either verify stem cell pluripotency marker expression, or markers of cell types of the three germ layers after differentiation. For the staining, cells were cultured in a 96-well glass bottom microscopy plate. When 80% confluency of cells or differentiation was achieved, cells were washed twice with cold PBS and then fixed using 100% ice cold methanol and incubation at -20°C for 7min. Then, the methanol was aspired, and cells were briefly rinsed in ice cold acetone for 20sec. Subsequent permeabilization of cells was performed with 0.1% tween20 in PBS for 5min at RT and three further washing steps with the same buffer. Then, cells were blocked for 30min at RT with PBS supplemented with 1% FCS, 0.5% BSA and 0.1% triton X-100. In the meantime, primary antibody solutions were prepared using the same blocking buffer as dilution buffer. Primary antibody solutions were incubated at 4°C overnight. The next day, cells were washed three times with blocking solution and then incubated with blocking buffer containing the respective secondary antibody and 2μg/ml DAPI. Cells were incubated for 1h at RT in the dark, then washed with PBS three times and imaged in PBS using a Nikon Eclipse Ti/X-Cite120Led microscope.

To enable the preservation of endogenous fluorophores like GFP, the staining protocol was switched to paraformaldehyde (PFA) fixation, and everything was performed

protected from direct light. In this protocol, again cells were washed twice with PBS and then incubated in 4% PFA (Himedia) for 10min at RT. After aspiration of the PFA and three washed for 5min at RT with PBS, the cells were permeabilised in PBS supplemented with 0.1% triton X-100 for 5min at RT. Subsequently, three 5min washes at RT with PBS containing 0.1% tween20 were performed before cells were blocked for 1h at RT in blocking buffer consisting of 3% BSA and 0.1% tween20 in PBS. During the incubation, primary antibody dilutions in blocking buffer were prepared. After the addition of the primary antibody solutions to the cells, they were incubated overnight at 4°C. The next day, cells were washed three times 5min with PBS at RT, and then incubated for 1h at RT in a secondary antibody dilution with 2μg/ml DAPI in blocking buffer. Finally, cells were washed three times with PBS and then imaged in PBS using the Nikon fluorescence microscope. Pictures were processed using Fiji and affinity designer.

The primary and secondary antibodies used in this study with their respective dilution are listed in the materials section 6.8.

7.2.14 Kidney organoid differentiation

Directed differentiation of hiPSCs into kidney organoids, a 3D structure containing a variety of cells from the kidney, was performed with kind support of the transplantation lab of Martin Hoogduijn at MC Rotterdam and the support and protocol of Nuria Montserrat at IBEC Barcelona (Garreta et al.).

Initially a range of plating densities were tried and optimised for our cell lines. Usually, 1.5×10^5 hiPSCs were plated as single cells per precoated 12-well on day -5 well as described before (chapter 7.2.7.2.1). Upon plating, cells were evenly distributed and not moved for the next 24h upon culturing in the incubator. The following day, the media was aspired, cells were rinsed once with PBS and differentiation to primitive streak was induced by the addition of 1ml advanced RPMI supplemented with 8μM CHIR per well. This media change was performed the next three days every 24h on day

METHODS

-4, -3 and -2. On day -1, media was again aspired, cells were rinsed gently with PBS and 1ml of a new differentiation media consisting of advanced RPMI, 200ng/ml FGF9, 1µg/ml heparin and 10ng/ml activin A was added per well. The next day, on day 0, cells appeared as tight monolayer and were harvested after a gentle rinse with PBS and detached using accutase. Cells were resuspended in little amount of advanced RPMI containing 3µM CHIR, 200ng/ml FGF9 and 1µg/ml heparin to obtain a concentrated cell suspension with at least $2.5x10^6$ cells/ml. Per 12-well well, approximately $3x10^6$ cells were obtained. Cells were counted and $5x10^5$ cells per well were aliquoted into a 96-well v-bottom plate and filled up to 150µl total volume with the resuspension media. Cells were pelleted by centrifugation at 300xg for 3min and immediately cultured in an incubator for two days without media change to allow self-aggregation and nephron progenitor cell (NPC) differentiation. On day 2, the aggregated spheroids were carefully transferred to a transwell 6-well plate by gentle aspiration using a wide orifice 200µl tip, plating up to seven pellets per 6-well well. 1.2ml of advanced RPMI supplemented with 3µM CHIR, 200ng/ml FGF9 and 1µg/ml heparin was added to the bottom of the transwell plate. On day 3, the media was carefully aspirated using a glass Pasteur pipette and switched to 1.2ml of advanced RPMI media containing only 200ng/ml FGF9 and 1ug/ml heparin. Cells were cultured in this condition for 48h without media change. Then, on day 5, media was replaced with fresh advanced RPMI media containing only 200ng/ml FGF9 and 1µg/ml heparin in a similar manner. On day 7, media was aspired again and now switched to advanced RPMI without any supplemented growth factors. Cells were maintained in this condition until final analysis between day 11-16, with media changes every second day.

The experiment was scaled up respectively (6-wp, $3x10^6$ cells/well, 2ml media) to obtain more cells on day 0. The advanced RPMI media stock was always supplemented with 1% L-GlutaMAX and 1% p/s. Detailed handling, dilution and preparation of reagents and media can be found in chapter 6.1.1 and 6.3.3.

Bright field and fluorescence microscopy pictures were taken throughout the differentiation process to follow the generation of tubular structures and renal vesicles and GFP expression in S/MAR modified hiPSCs. To confirm differentiation, qPCR

samples were taken at day 0 and between day 11-16. The following markers were investigated: loss of Oct4 and nanog to confirm loss of pluripotency, GFP, FLCN and FLCN-flag for S/MAR expression validation; Pax2 upregulation on day 0 for intermediate mesoderm differentiation; TBXT (brachyury) expression for primitive streak differentiation on day 0; PODXL, WT1 (glomeruli), villin (proximal tubuli) and ECAD (distal tubuli) for further nephron differentiation upon 3D culturing, WT1 as indicator on day 0 for successful differentiation as suggested by the group from Rotterdam. Furthermore, protocols for fixation and processing for H&E and IHC staining were tested (chapter 7.2.14.1).

7.2.14.1 Handling of kidney organoids for H&E or IHC staining

H&E staining was performed in collaboration with Vanessa Vogel, HiStem, DKFZ. IHC staining was performed in collaboration with the lab of Martin Hoogduijn, MC Rotterdam, and their pathology core facility.

Kidney organoids between day 11 to 16 with visible structures were used, however we observed increased accumulation of dead cells and thus disruption of the 3D structure on organoids the longer they were cultured. Careful incubation of the organoids in a 1:100 in PBS diluted tissue marking dye solution for 5min at RT proved to aid the later process in tracking the embedded organoid when cutting the block. After gentle aspiration of the staining solution and optional rinse with PBS, the transwell membrane surrounding the organoid was cut with a scalpel and placed in a 10x10x5mm cryomold. Histogel was liquefied at 65-70°C in a water bath, and approximately 200-300µl was carefully added on top of the organoid. After the Histogel was solidified, the block was carefully pushed out by cutting the bottom of the cryomold with a scalpel, transferred in a tissue cassette and submerged into 10% formalin until further processing by our collaborators with their established protocols.

7.2.15 Generation of CRISPR/Cas9 FLCN knock-out iPSCs

7.2.15.1 sgRNA design and cloning of a CRISPR/Cas9 plasmid

For the generation of FLCN KO cells, sgRNAs were designed against exon 4 out of 14, which represents the first protein coding exon of FLCN. The UCSC Genome Browser was used and suggested target sites at the location chr17:17,227,819-17,228,247 were checked. A guide sequence at the location chr17:17228122-17228144, overlapping the protein start codon with a MIT Guide Specificity Score of 84, and 2 predicted off-targets upon 2 mismatches was chosen. Through the available direct link, CRISPOR was then used to obtain the oligo sequence for cloning (Concordet & Haeussler, 2018). The obtained sequence was modified at the 5' ends as indicated in the cloning protocol of the Zhang lab for target sequence cloning into px330-based CRISPR plasmids available on addgene (Cong et al., 2013). The respective oligos were ordered at sigma aldrich and cloned into a px458-RFP CRISPR plasmid by following the instructions of the cloning protocols. Briefly, 5µg of the plasmid was digested with 1.5µl BpiI-FD in a 20µl reaction for 1h at 37°C and subsequently dephosphorylated by the addition of 1µl bacterial alkaline phosphatase and further incubation for 1.5h at 65°C. In the meantime, 1µl of each oligo (100µM) were annealed with 0.5µl T4 PNK in T4 ligase buffer in a 10µl reaction for 30min at 37°C followed by 95°C for 5min and a ramp down to 25°C at 5°C/min. The digested and dephosphorylated plasmid was run on an agarose gel and extracted as described before (chapter 7.1.1.6). Ligation of 50ng vector and 1:250 diluted phosphorylated and annealed oligos was performed with 1µl T4 ligase in a 20µl reaction over night at 16°C. The next day, DH5α were transformed with the reaction, bacterial clones were expanded, plasmid DNA isolated and the successful cloning was confirmed by restriction digestion and sequencing.

7.2.15.2 Generation of a knock-out iPSC line

The cloned FLCN-px458-RFP plasmid was transfected to hiPSCs using lipofectamine stem (chapter 7.2.9.1). Three days after transfection, RFP-expressing alive cells were sorted as single cells in iMatrix Laminin-511 precoated 96-well plates. Wells where single colonies grew out were further expanded. To confirm successful knock-out (KO), gDNA of the clones was isolated from 10×10^4 cells with 21µl reagent from the fire

tissue kit (chapter 7.1.4.2). A 352bp fragment surrounding the sgRNA binding site was amplified with 100µM or 10µM primers and the fire tissue kit. After separation on an agarose gel, the respective band was cut, the DNA was extracted and sent for sequencing using the PCR primers.

7.2.16 10x single cell RNA sequencing

Single cell sequencing of WT, KO, and S/MAR modified hiPSCs was performed in collaboration with the lab of Prof. Michael Platten (Immunology and Cancer, DKFZ) and support of Dr Edward Green (Immunogenomics, DKFZ). The experiment was performed with the help of Chin Leng Tan (Immunogenomics, DKFZ), Gordon Haltenhof (Immunogenomics, DKFZ) and Julia Peterson (DNA vector lab, DKFZ). Cells were detached with accutase, and remaining doublets were excluded by filtering the cell suspension three times through FACS filter tubes. The different samples were labelled with a process referred to as hashing and designed by the group of Dr Edward Green. Briefly, single cells were incubated with 9µl a specific barcode oligo (1µM) and 9µl of an anchor1 oligo (1µM) with 72µl PBS. After 5min incubation on ice, 10µl of anchor2 was added, facilitating the binding of the barcode to the cell surface. After further 5min on ice, cells were washed thoroughly for 3 times to get rid of unbound barcodes.

Then, dead cells were removed with the dead cell removal kit (Miltenyi). Cells were resuspended in 100µl beads and incubated at RT for 15min. A 1.5ml was precoated with sterile 5% BSA in PBS at RT to avoid sticking of cells to the plastic. LS autoMACS columns (Miltenyi) were equilibrated with 3ml 1x solution from the kit diluted in autoclaved MQ H_2O. After incubation, 300µl solution was added to the cell suspension to obtain a total of 400µl, what matches the void volume of the column. The mixture was loaded on the prepared column and the flowthrough was trashed. Alive cells were then eluted with 1.4ml 1x solution into the prepare 1.5ml tube. Cells were washed with PBS to remove EDTA and counted.

METHODS

Some samples were subsequently mixed and up to 40.000cells were loaded on one GEM of a 10x Chromium Next GEM Single Cell V(D)J Reagent Kit v1.1.

RNA was isolated, barcoded and further processed as of manufacturer's instruction following the protocol with the reference number CG000208 Rev F skipping step 4 and 5. Samples were sequenced by the core facility and data was aligned, annotated, and analysed by Chin Leng Tan and Julia Peterson.

8 RESULTS

8.1 A variety of S/MAR DNA vectors

This project builds on previous work in the lab optimising the features of S/MAR DNA vectors (Dr Matthias Bozza) and their first proof of functionality in murine pluripotent stem cells (Dr Alicia Roig Merino). Based on this knowledge, a range of S/MAR DNA vectors were generated for this study. They all build upon the pCAG and nCAG vectors, expressing GFP and puromycin resistance (PuroR) under a CAG promoter as depicted in Figure 5A, which were previously successfully utilised for pluripotent cells (Roig Merino, 2018; Roig Merino et al., in press).

While using these vectors as valuable GFP expressing controls and as the basis for further vectors, an insulator element (Element 40, Ele40) was introduced to pCAG, with the intention to prevent possible silencing events, leading to the S/MAR DNA vector P17 (Figure 5A).

After having the first proof-of-concept validation, that the S/MAR DNA vector platform can be applied to pluripotent stem cells, this project aimed to move towards a more clinically relevant proof-of-concept study by introducing functional transgenes into the S/MAR DNA vectors and applying them to human induced pluripotent stem cells (iPSCs) for the first time.

Besides establishing the workflow, culturing and GFP-proof-of-concept experiments, this work mainly focused on the hereditary Birt-Hogg-Dubé syndrome (BHD) and its functional gene folliculin (FLCN) as models. New DNA vectors were generated based on pCAG, the expression cassette was replaced with a PuroR-p2a-FLCN-flag cassette (P3 vector, Figure 5B). After transcription into mRNA, the self-cleaving p2a sequence is translated into an amino acid sequence, which induces a ribosome skip and results in a break between two amino acids and two equally produced proteins. The amino acids of the p2a sequence however are added to the C- and N-terminus of both proteins and could interfere with protein function. Since the second protein only has one additional

proline at the N-terminus, the gene of interest was placed as the second gene while previously no impairment of GFP fluorescence was observed with the C-terminal p2a sequence.

To preserve both, puromycin resistance to aid S/MAR DNA vector establishment and GFP for easy visualisation while including the transgene FLCN, vector P9 was cloned (Figure 5C). Here, and coGFP was added to the expression cassette of P3 linked via an IRES element which acts as an additional translation starting site.

The next level of FLCN-S/MAR DNA vectors in Figure 5C – vector P13, P14 and P15 – were designed for a drug-free approach and lack the puromycin resistance gene. Instead, the expression cassette simply contains coGFP-p2a-FLCN with or without flag-tag (P13, P14). While the first two vectors are based on pCAG and thus lack an insulator element, P15 was cloned by adding Ele40 into the GFP-p2a-FLCN-flag vector P13.

Finally, besides focusing on the pS/MAR DNA vectors, also some FLCN expressing Nanovectors™ were cloned in collaboration with Nature Technology Corporation (NTC) (Figure 5E). Unlike the Nanovector™ nCAG, these contain the insulator Ele40 and the coGFP-p2a-FLCN-flag expression cassette as in P13 and P15. While N2 otherwise has the same S/MAR backbone as the other vectors, N4 and N6 contain further modifications of the S/MAR introduced by Dr Matthias Bozza, namely splicing sites in N4 and splicing sites and a shorter S/MAR sequence focusing on the core element of the S/MAR sequence (CMAR).

Some concurrent work was also performed with iPSCs derived from Choroideremia (CHM) patient cells in a proof of concept study. For this, Dr Patrick V. Almeida (DNA Vector lab, DKFZ) designed and cloned the Nanovector™ NV6, which was synthesised by NTC. The vector harbours the same spliced S/MAR sequence as N4, but the FLCN sequence was replaced with the Rep1 cDNA, the responsible gene for the monogenic disease CHM (Figure 5E).
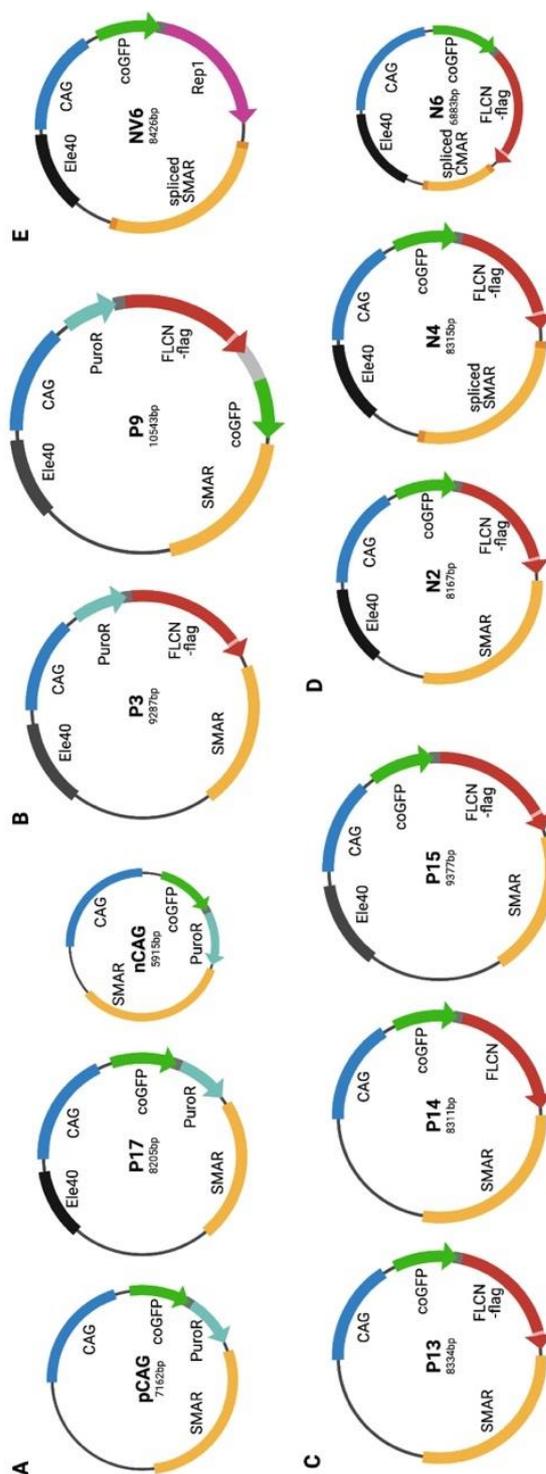
Figure 5: A variety of S/MAR DNA vectors.
Schematics of the vectors used in this study are depicted in categories. Details of the vectors are described in the main text in section 8.1. A) The GFP vectors pCAG and nCAG, cloned and applied in previous studies by Dr Matthias Bozza and Dr Alicia Roig Merino and the modified pCAG vector P17 are shown. B) pS/MAR DNA vectors containing FLCN and PuroR genes. C) pS/MAR DNA vectors encoding for GFP and FLCN. D) GFP-FLCN expressing Nanovectors™. E) The Nanovector™ NV6 is based on N4 but encodes the protein Rep1.

8.2 Genetic modification of cancer cell lines using Folliculin-S/MAR DNA vectors

8.2.1 Establishment via puromycin drug selection

8.2.1.1 S/MAR modification of Folliculin deficient cell lines UOK257 and FTC133

As a first step, the cloned FLCN-S/MAR DNA vectors were tested for expression and functionality. By applying them in the FLCN deficient cell lines UOK257 and FTC133, we aimed to generate FLCN-restored cell models which could be exploited for further studies on the role of FLCN. This data was part of a bachelor thesis and was generated with the support of Julia Emmenecker.
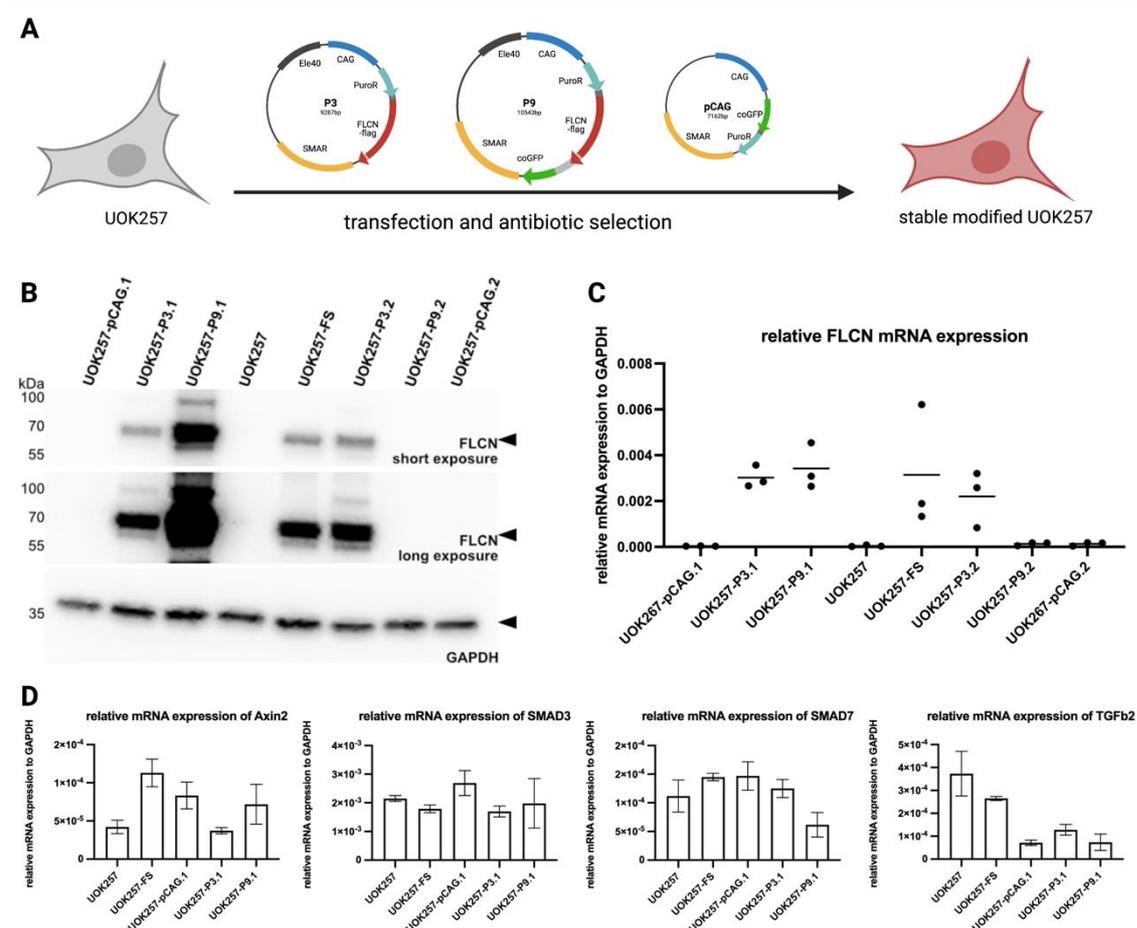
Figure 6: Stable S/MAR modified UOK257 cells.
A) Overview schematic of the used vectors (P3, P9, pCAG) in UOK257 cells. Vectors are depicted in detail and described in Figure 5 chapter 8.1. B) Western Blot of the different established cell lines showing FLCN protein levels and GAPDH as loading control. Samples taken >40dpt C) qPCR data of three biological replicates (samples taken >40dpt) showing FLCN expression relative to GAPDH. D) qPCR of technical replicates (samples taken >40dpt) for genes important in signalling pathways which are known to be affected by FLCN (Axin2, SMAD3, SMAD7, TGFβ2) relative to GAPDH.

Several optimisation steps of delivering the DNA vectors to the two cell types using the Neon electroporator or the lipofection reagents XtremeGene9 and XtremeGene HP were performed (data not shown) and efficient delivery of the vectors P3, P9 or pCAG was achieved (schematic Figure 6A/Figure 7A).
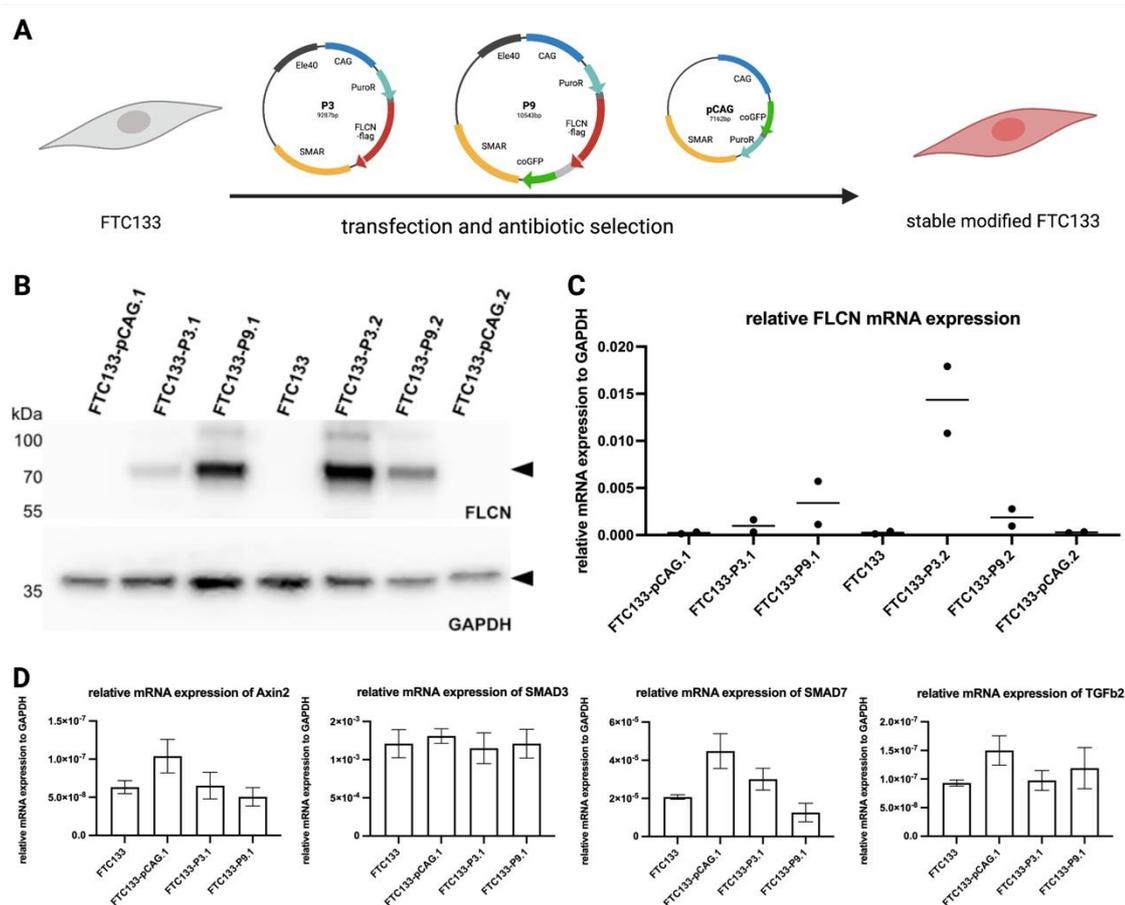

Figure 7: Stable S/MAR modified FTC133 cells.
A) Overview schematic of the used vectors (P3, P9, pCAG) in UOK257 cells. Vectors are depicted in detail and described in Figure 5 chapter 8.1. B) Western Blot of the different established cell lines showing FLCN protein levels and GAPDH as loading control. Samples taken >40dpt C) qPCR data of three biological replicates (samples taken >40dpt) showing FLCN expression relative to GAPDH. D) qPCR of technical replicates (samples taken >40dpt) for genes important in signalling pathways which are known to be affected by FLCN (Axin2, SMAD3, SMAD7, TGFβ2) relative to GAPDH.

Stable S/MAR modified UOK257 and FTC133 cells were generated by several rounds of puromycin selection during a total of approximately 2-3 weeks of culturing. For each of the cell lines, two stable cell lines per vector were established. Expression or absence of FLCN was confirmed by Western Blot and qPCR with samples taken after at least 40 days post transfection (dpt) (Figure 6B, C/Figure 7B, C). While the cell line UOK-P9.2 did not show FLCN expression in either method, all other established cell lines confirmed presence (P3, P9) or absence (parental line, pCAG) of FLCN in the respective samples. Notably, corresponding levels of mRNA and protein were observed in FTC133

stable cell lines, while the strong expression of UOK257-P9.1 was not equally observed in mRNA samples. For UOK257 cells, the FLCN restored UOK257-FS cell line, which has been established previously with an earlier version of S/MAR DNA vectors was used as a positive control (Wong & Harbottle, 2013).

To illustrate the utility of the genetically corrected cell lines the UOK257-P3.1 cell line was shared with collaborators and utilised to study the impact of FLCN on the survival of neisseria gonococci upon infection and showed upregulation of E-cadherin (T. Yang et al., 2020).

While our FLCN-restored cell lines proved to be a valuable tool for such studies, we were also interested in the tumorigenesis effect of FLCN causing kidney cancer in BHD patients. Thus, we had a look at target genes of known pathways affected by FLCN expression such as Wnt (Axin2) and the TGFβ pathway (SMAD3, SMAD7, TGFβ2) (Figure 6D, Figure 7D). However, in a preliminary experiment, we could not observe changes in target gene expression distinctively brought about by FLCN-S/MAR DNA vector expression, as e.g., TGFβ2 levels in UOK257 cells were most affected by pCAG expression. Together the data supports, that the isogenic restoration of FLCN can be a helpful tool to investigate the function, mechanism of action and involvement of this protein. However, the unstable nature of cancer cell lines renders them unsuitable for FLCN-induced tumorigenesis studies and alternative advanced models are needed.

8.2.1.2   Validation of insufficient GFP expression in stable HEK293T-P9 cells

While puromycin establishment with the P3-S/MAR DNA vector in UOK257 and FTC133 cells was performed without any visual expression control, some stable P9 cells heterogeneously expressed GFP. GFP expression in these cells was expected to be less reliable than PuroR and FLCN since it was the third transgene in the expression cassette and despite having the same mRNA, the IRES element requires a new translation initiation. To further investigate this observation in a commonly used cancer cell line, a stable HEK293T-P9 cell line was established using puromycin

selection and analysed 35 dpt (Figure 8A). While a lack of cell death upon drug selection suggested a pure stable cell population and FLCN expression was confirmed via Western Blot (Figure 8B), fluorescence microscopy suggested very inconsistent GFP expression with varying intensity and many cells with hardly observable or no GFP expression (Figure 8C).

Together with observed GFP- cells in stable UOK257-pCAG cell lines, this additionally hints at a possible emerging of puromycin resistance during establishment (data not shown). The impracticality of the PuroR-FLCN-GFP vector P9 and the risk of emerging drug resistance led to the development of an optimised approach for S/MAR DNA vector establishment.
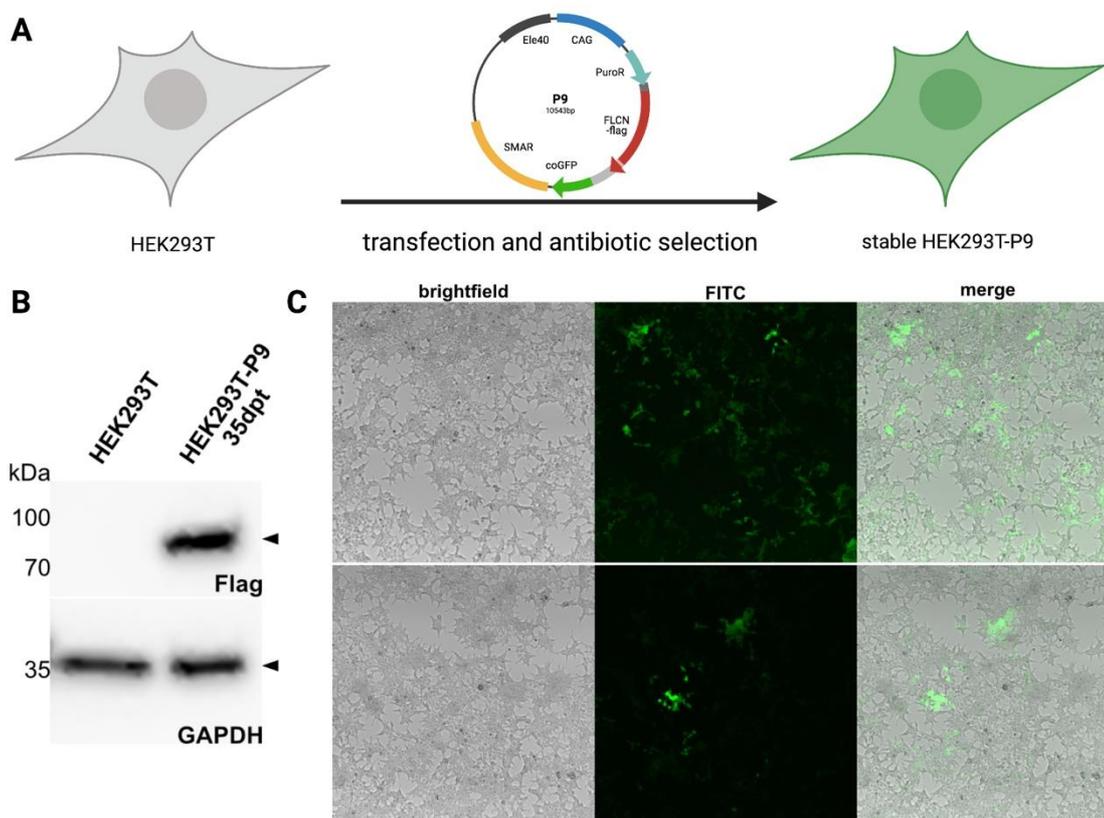
Figure 8: Stable HEK293T-P9 cells.
A) Overview schematic of the generation of HEK293T-P9 cells. B) Western Blot WT and HEK293T-P9 cells 35dpt probed for FLCN-flag and GAPDH as loading control. C) Fluorescence microscopy picture of selection resistant stable HEK293T-P9 cells 35dpt displaying patchy GFP expression.

8.2.2 Isolation of an S/MAR expressing cell population using FACS sorting

Accordingly, we developed a new approach which avoided the use of puromycin selection and utilised the properties of GFP as a reporter gene and a fluorescent selection marker. Vectors simply encoding the gene of interest (FLCN) and GFP were used in the following experiments, providing the possibility to directly observe expression of the S/MAR DNA vector, and by inference FLCN expression while providing indications of the homogeneity of expression in the established cell line.

8.2.2.1 Establishment of a FACS purification protocol in HEK293T cells

An initial evaluation of the feasibility of making drug-free stable cells with this new generation of FLCN vectors was performed in HEK293T cells with GFP (pCAG, P17) and FLCN-GFP (P14, P15) S/MAR DNA vectors (Figure 9A). Viable, GFP expressing (GFP+) cells were sorted and further expanded and cultured 1, 17 and 44dpt via FACS sorting (Figure 9B). A first significant drop of GFP+ population size was observed within the first two weeks after plasmid delivery with only 1-3% GFP+ population 17dpt. Subsequently to this sorting, another major expression drop was observed until 44dpt, with 27-28% GFP+ cells in the FLCN expressing cell lines and 11-12% GFP+ cells in the GFP vector-transfected cell lines. In line with observations of S/MAR establishment using drug selection, where S/MAR DNA vector cells lines are assumed stable after approximately 30 days, cell lines maintained a 98% purity of GFP expression after the sorting 44dpt until at least the monitored 76dpt. Both, GFP and the presence or absence of FLCN was confirmed in the expected samples via qPCR 62dpt (Figure 9C). Together, this data suggests a successful drug-free methodology for establishing stable cell lines with S/MAR DNA vectors.

To provide additional evidence about the episomal state of the S/MAR DNA vectors we performed Rolling Circle Amplification PCR (RCA) and attempted to amplify circular vector DNA from extracted gDNA from the established cell lines. However, while vector amplification in the reaction was visible in the controls when running the digested reaction mix on an agarose gel, the tested samples did not show any bands

(Figure 9D). Further PCR amplification of GFP on the amplified RCA reaction as well as input gDNA was performed subsequently and confirmed GFP DNA with a distinct band in all S/MAR DNA vector samples but not the parental cell line sample and the negative water control (Figure 9E).
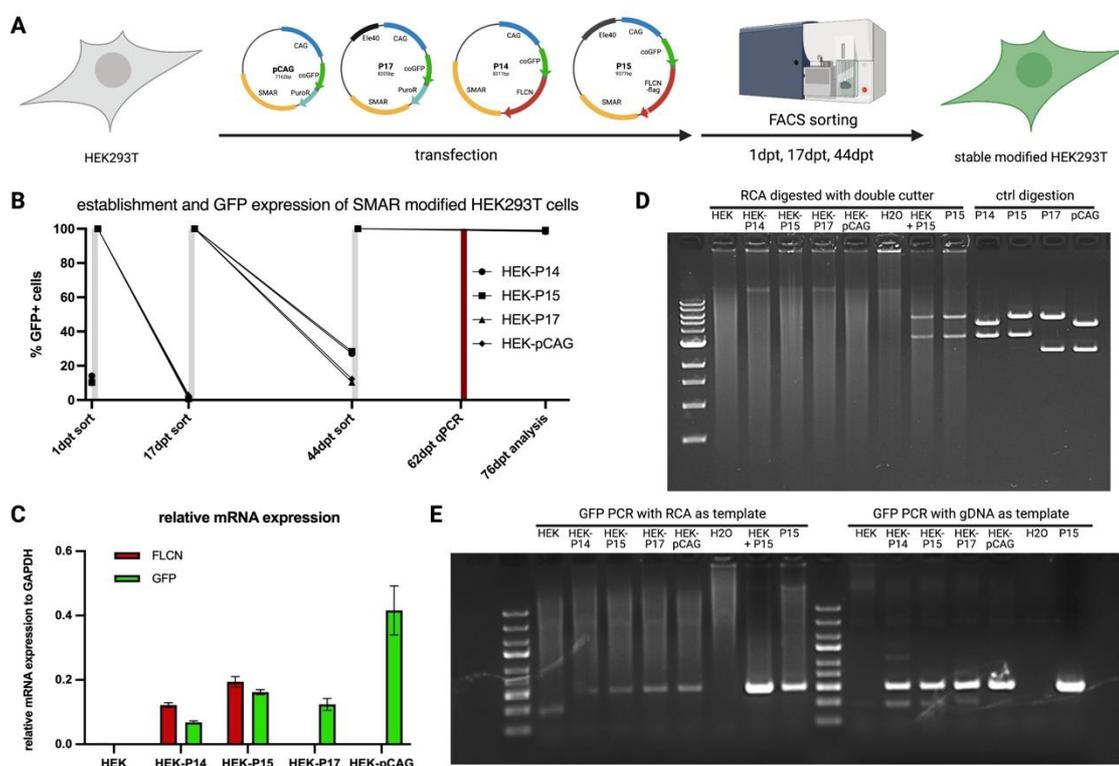

Figure 9: Drug-free establishment of stable HEK293T-S/MAR cells.
A) Overview schematic of the generation of stable HEK293T-S/MAR cells using the GFP vectors pCAG and P17 and the FLCN-GFP vectors P14 and P15. Vectors are depicted in detail and described in Figure 5 chapter 8.1. GFP+ cell population was sorted and expanded 1, 17 and 44dpt. B) Data of % of GFP+ cell population during FACS sorting and analysis of the established four different cell lines plotted in a time-dependent manner. C) qPCR of the different cell lines in technical replicates (samples taken 62dpt) confirming FLCN and GFP levels relative to GAPDH. D) Rolling circle amplification (RCA) on gDNA of established cell lines isolated 57dpt. E) PCR on RCA (left) and gDNA (right) of samples from D).

8.2.2.2 Folliculin restoration in the Birt-Hogg-Dubé syndrome cell line UOK257 using FACS sorting

As a next step, the establishment via FACS sorting was applied to the BHD cell line UOK257. As depicted in Figure 10A, UOK257 cells were transfected either with the plasmid vectors (pS/MAR) P13, P14, P15 and P17 or the Nanovectors™ (nS/MAR) N2, N4, N6 and nCAG. Similar to the previous experiment, viable GFP+ cells were FACS purified and cultured further for 6 and 32dpt for pS/MAR transfected cells, and 15dpt for nS/MAR transfected cells.

RESULTS

Subsequent culturing and routine FACS analysis until 83dpt (pS/MAR) or 66dpt (nS/MAR) displayed loss of GFP+ population in the FLCN-restored pS/MAR cell lines UOK257-P13 and UOK257-P14 (Figure 10B) and in the UOK257-N2 nS/MAR cell line (Figure 10C).
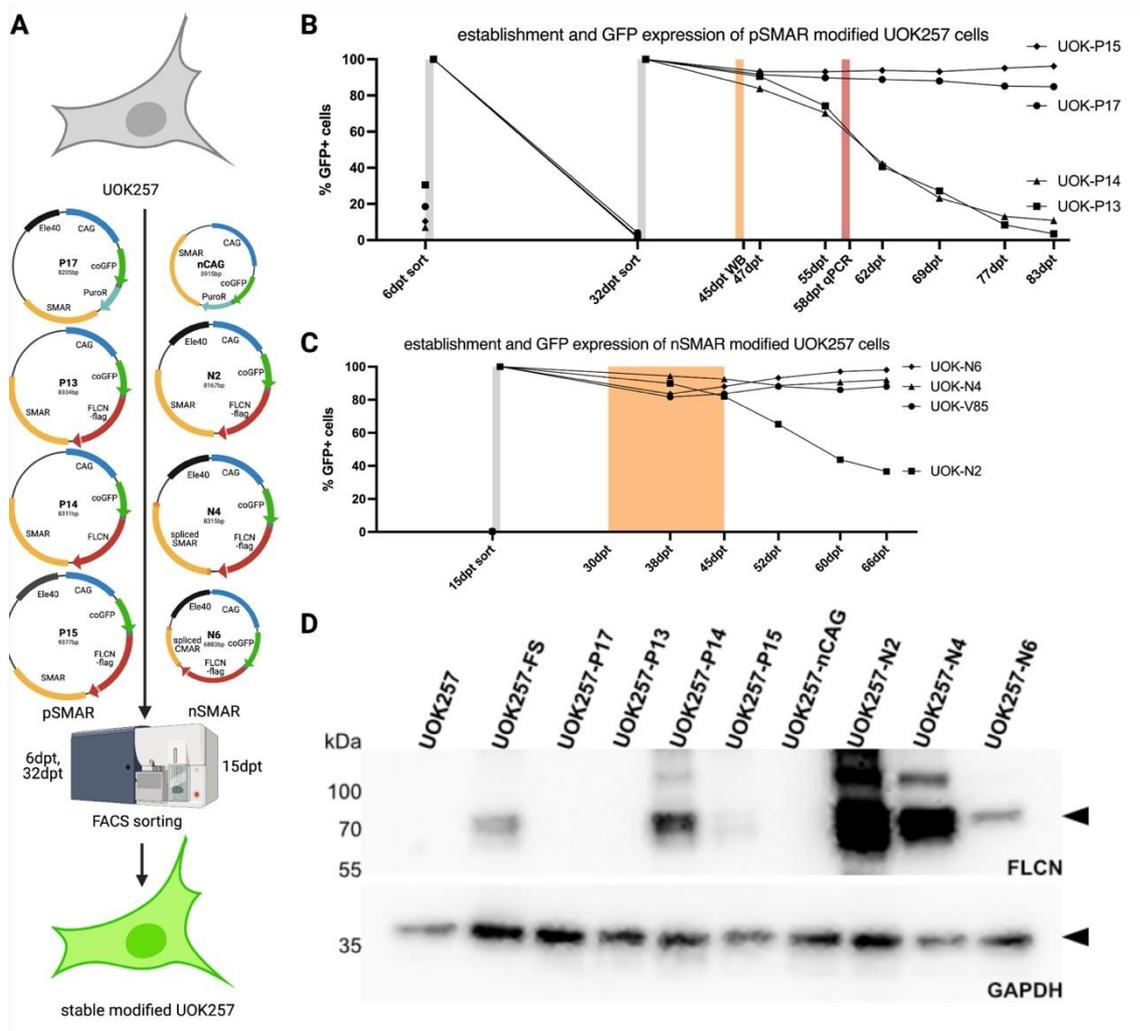

Figure 10: Drug-free establishment of stable UOK257-S/MAR cells.
A) Overview schematic of the generation of stable UOK257-S/MAR cells using the GFP vectors pS/MAR DNA vectors P17, P13, P14 and P15 and the nS/MAR DNA vectors nCAG, N2, N4 and N6. Vectors are depicted in detail and described in Figure 5 chapter 8.1. GFP expressing cells were sorted 6 and 32dpt for pS/MAR transfected cells and 15dpt for nS/MAR transfected cells. B) Data of % of GFP+ cell population during FACS sorting and analysis of the pS/MAR established cell lines plotted in a time-dependent manner. C) Data of % of GFP+ cell population during FACS sorting and analysis of the nS/MAR established cell lines plotted in a time-dependent manner. D) Western Blot of the different established cell lines showing FLCN protein levels and GAPDH as loading control. Samples taken as labelled in orange in B) and C).

Western Blot samples were taken when most cells expressed GFP (45dpt (pS/MAR) or between 30-45dpt (nS/MAR)) and probed for FLCN and GAPDH. While the nS/MAR cell lines UOK257-N2 and UOK257-N4 presented very strong FLCN expression, in general, FLCN expression was confirmed in all FLCN-S/MAR DNA vector cell lines except for UOK257-P13.

Further qPCR validation from 58dpt of the confirmed pS/MAR cell lines UOK257-P14, -P15 and -P17 as well as parental and UOK257-FS cells as control further demonstrated no FLCN expression in UOK257-P15 compared to parental FLCN deficient UOK257 cells (Figure 11A). Together with the respective weak FLCN protein band in the Western Blot from UOK257-P15 samples 13 days previously suggest continuous loss of transgene expression. However, at the same time of apparently losing FLCN transgene expression, the UOK257-P15 cell line is the only FLCN encoding S/MAR-UOK257 cell line which maintains GFP expression in FACS analysis throughout the observed time of 83 days of culturing. While transgene expression of GFP and FLCN was very inconsistent in the established cell lines, also qPCR for different FLCN pathways did not result in conclusive alterations in expression levels. This was even the case when comparing the controls and the early timepoints of the UOK257-P14 cell line, where both GFP and FLCN expression was confirmed.

Additional assays such as southern blot and RCA did not provide detection of bands and thus an indication that the S/MAR DNA vectors episomal entity could be recovered (data not shown). Finally, PCR on gDNA was performed (Figure 11B). While observing distinct bands for the internal control unlike for drug-free established HEK293T cells (Figure 9E), PCR for GFP only provided a distinct band in the positive control of parental gDNA with plasmid DNA, while parental gDNA alone resulted in some unspecific bands. Notably, the second positive control of gDNA from cells 2dpt displayed similar unspecific bands, suggesting that these bands can be easily obtained with the gDNA of UOK257 and the absence of a high amount of template. In line with the FACS GFP expression, faint bands at the right size for GFP could be observed both for UOK257-P15 and P17 cells. It should be noted, that the obtained GFP bands were inconsistent upon different PCR reactions as visible in the different patterns of the two GFP agarose gels. An additional PCR amplifying the GFP-FLCN part of the S/MAR DNA vectors resulted in no band in any of the samples, while providing a distinctive band in the positive control. The inconclusiveness of the data despite applying a drug-free establishment method emphasises the unstable characteristics of the UOK257 cancer cell type.
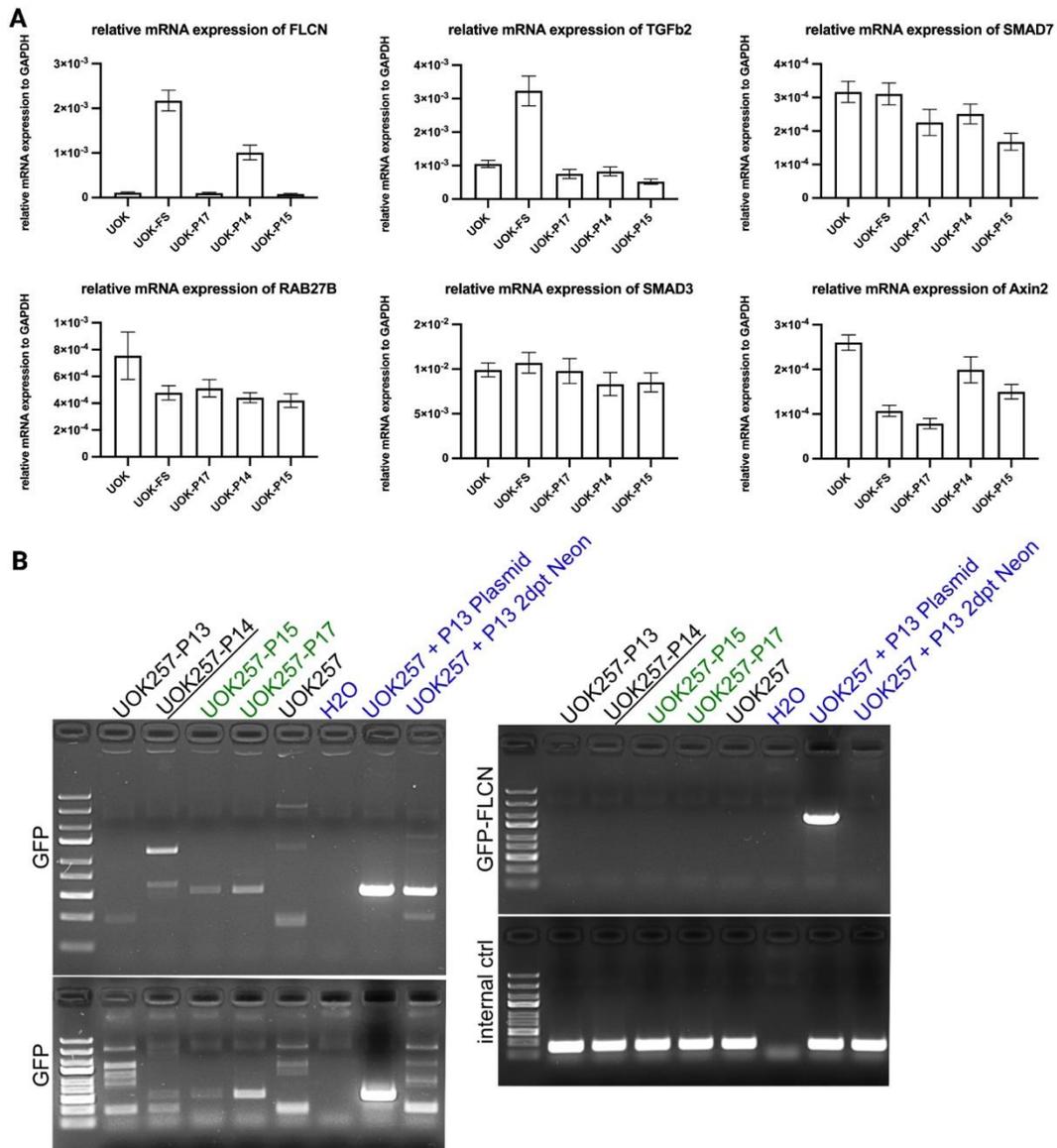
Figure 11: Analysis of FACS established pS/MAR UOK257 cells.
A) qPCR analysis of technical replicates (samples taken 58dpt) for FLCN and genes important in signalling pathways which are known to be affected by FLCN (Axin2, SMAD3, SMAD7, TGFβ2, RAB27B) relative to GAPDH. B) PCR on gDNA amplifying GFP, GFP-FLCN or an internal control gene.

8.3   Genetic modification and simultaneous reprogramming of normal human dermal fibroblasts

After verifying the functionality of the FLCN S/MAR DNA vectors in combination with either puromycin selection or FACS sorting we aimed for a more sophisticated cellular model to investigate the importance of FLCN for tumorigenesis in kidney cancer. An ideal gene therapy vector for BHD could be applied as preventive gene therapy before the endogenous functional FLCN is lost and tumorigenesis occurs. To enable the modelling of different cell types and more sophisticated cellular models, we focused on the application of these FLCN S/MAR DNA vectors on human iPSC cells and their progeny.

In previous work we demonstrated that our S/MAR DNA vectors and specifically pCAG and nCAG, which the FLCN vectors were based on, can be used to successfully modify mouse pluripotent stem cells without altering their pluripotency and while maintaining transgene expression throughout reprogramming and differentiation. They can even be used to generate transgenic mice (Roig Merino et al., in press). Together with Dr Alicia Roig Merino we obtained the first data utilising these GFP-S/MAR DNA vectors in hiPSCs. Together with EBNA reprogramming vectors, the pCAG S/MAR DNA vector was delivered to normal human dermal fibroblasts (NHDFs) (Figure 12A). Successful delivery was confirmed by GFP expression 1dpt. To continue reprogramming, transfected cells were transferred to feeder cells and hiPSC feeder media 8dpt. Upon 25-30dpt distinct hiPSC colonies emerged, including GFP expressing hiPSC cells.

A GFP expressing hiPSC colony was expanded and cultured further. Besides validating pluripotency markers (data not shown, published in (Roig Merino, 2018)), random differentiation was performed to verify the differentiation capabilities and persistent transgene expression of S/MAR modified hiPSCs in line with observations with mouse pluripotent cells. GFP expressing iPSCs were clumped and cultured in suspension to form embryoid bodies (EBs) and subsequently plated on gelatinized plates to allow attachment and outgrowth of differentiated cells (Figure 12B). Fluorescence microscopy pictures provide evidence for maintained GFP expression throughout this

differentiation process as well as in the differentiated progenies. The ability to differentiate into cells of all three germ layers was confirmed via IF staining for markers of Ectoderm, Endoderm and Mesoderm respectively.
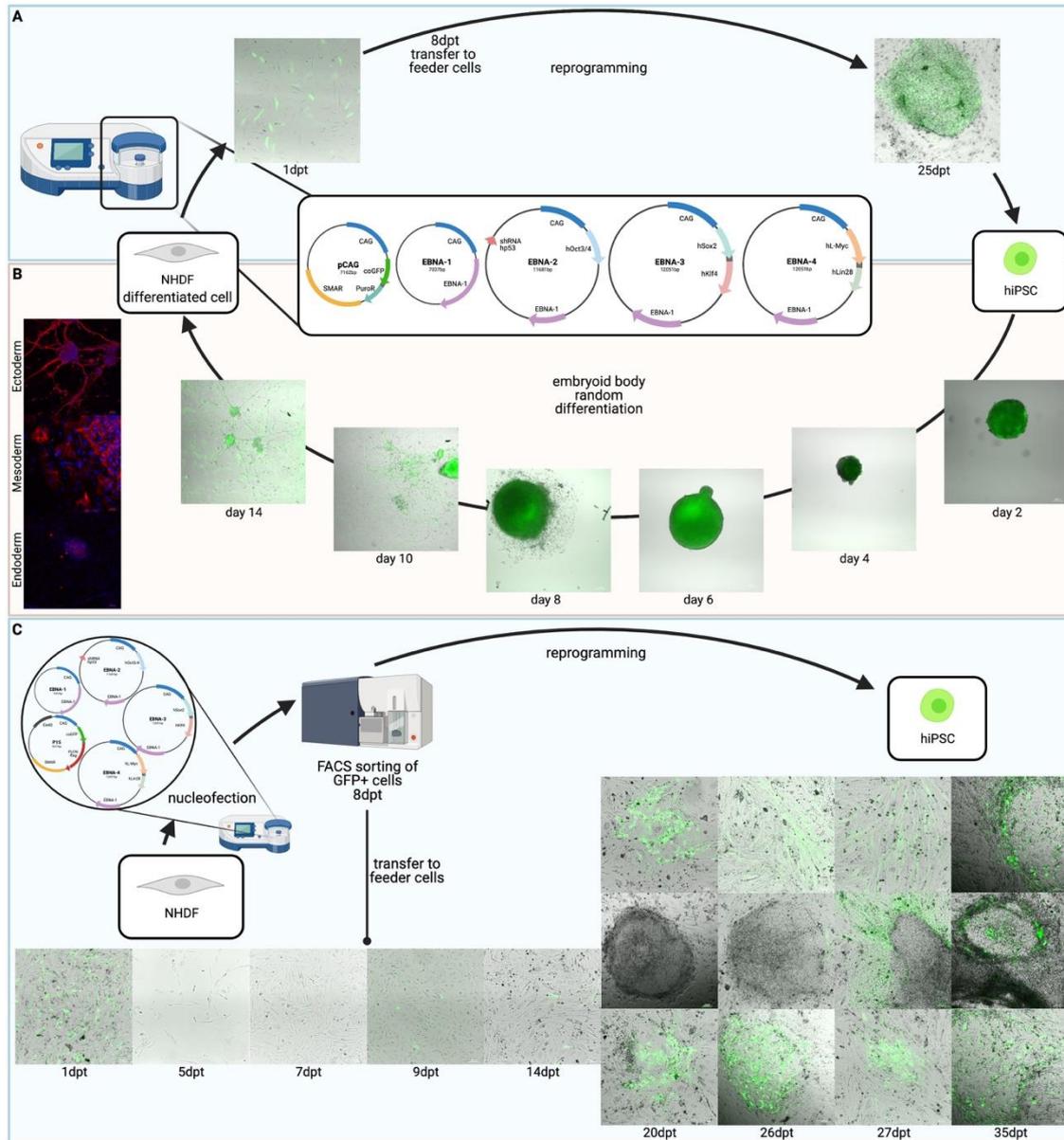

Figure 12: Reprogramming of NHDFs and simultaneous genetic modification with S/MAR DNA vectors.
A) Delivery S/MAR pCAG vector and EBNA reprogramming vectors via Amaxa II nucleofection to normal human dermal fibroblasts (NHDFs). S/MAR DNA vectors are depicted in detail and described in Figure 5 chapter 8.1. EBNA reprogramming vectors are depicted in detail and described in Figure 3 chapter 5.8.3. Schematic of reprogramming to hiPSCs with fluorescence microscopy pictures are depicted. B) pCAG expressing hiPSC cells were subjected to random differentiation via embryoid body formation. Fluorescence microscopy pictures of GFP expressing embryoid bodies and resulting differentiated progenies are shown. Immunofluorescence pictures of staining differentiated cells for germ layer markers Ectoderm (b3Tub), Mesoderm (aSMA) and Endoderm (FoxA2) are shown to the left. C) Schematic of delivery of the FLCN-S/MAR DNA vector P15 together with EBNA reprogramming vectors. P15 vector expressing cells were sorted 8dpt and plated on feeder cells in stem cell conditions to continue reprogramming. Fluorescence microscopy pictures demonstrate GFP expression and formation of hiPSC colonies without GFP expression.

Based on this initial work in hiPSCs, a similar establishment protocol with the FLCN
expressing S/MAR DNA vector P15 was attempted (Figure 12C). After delivery of EBNA
and P15 vectors to NHDFs, S/MAR DNA vector expressing cells were sorted and plated
on feeder cells in stem cell conditions for further reprogramming. Thus, making sure
that only human reprogramming cells were further cultured which had expressed the
P15 and enhancing chances of P15-expressing stem cells. Distinct hiPSC colonies were
obtained between 20-35dpt. However, observation with a fluorescence microscope
showed predominantly GFP expression in fibroblast-like cells or partially
reprogrammed colonies while no GFP expression was observed in fully
reprogrammed colonies.

8.4   Optimisation of reprogramming, cell source and culturing conditions

After the preliminary experiments suggested, that S/MAR DNA vectors can be used for the modification of human pluripotent stem cells as in mouse ones, in a next step the reprogramming protocol, the cell source as well as hiPSC culturing were optimised to enable an optimised process and the direct modification of hiPSCs.

8.4.1   Comparison of reprogramming vector delivery

With reprogramming being a very inefficient process, we hypothesised that increasing transfection efficiency of the reprogramming vectors could improve efficiency. Routinely, the Amaxa II nucleofector with the NHDF kit was used for EBNA vector delivery to NHDFs. To test our hypothesis, first, the transfection efficiency with pCAG or nCAG was compared when using the neon electroporator or the Amaxa II nucleofector (Figure 13A, C). Both GFP plasmids resulted in similar viabilities (~95%), but the neon electroporator achieved a higher transfection efficiency reaching between 65-67% compared to 25-30% GFP expressing cells with the Amaxa II nucleofection (Figure 13C left axis).

To directly compare transfection efficiency effect on reprogramming efficiency, aliquots of the same cells were transfected at the same time with the EBNA reprogramming vectors with these two delivery systems. Reprogramming was performed utilising the established protocol where cells were plated in a specific count onto feeder cells 8dpt and then the media was switched to stem cell media. AP staining was performed 30dpt which allowed the manual counting of fully reprogrammed hiPSC colonies, where each colony is assumed to arise from one single fibroblast (Figure 13B). The reprogramming efficiency was calculated based on the iPSC colony count in respect to plated cells on day 8. Interestingly, despite achieving a better transfection efficiency, reprogramming efficiency was higher with the Amaxa II (0.086%) device than with the neon (0.062%) (Figure 13C right axis), suggesting that delivery efficiency alone is not the key factor for successful reprogramming.
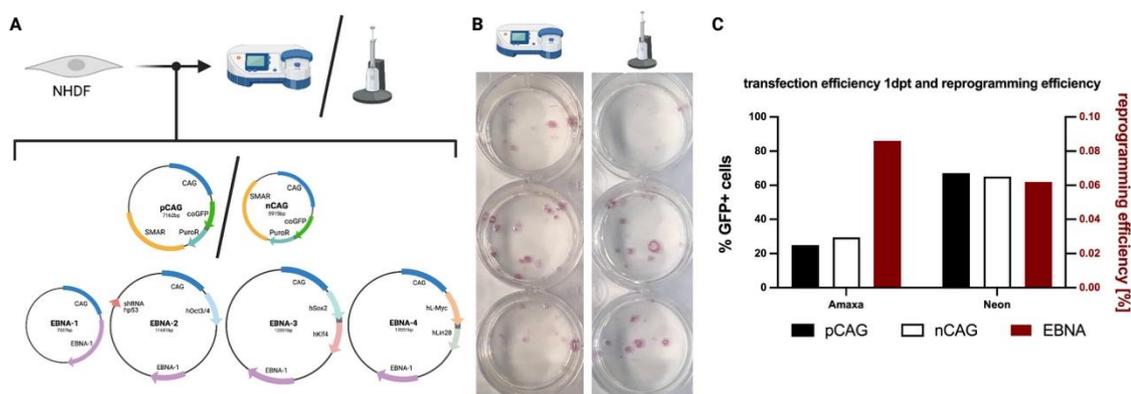
Figure 13: Transfection and reprogramming efficiency comparison of NHDFs transfected with the Amaxa II nucleofector or the neon electroporator.
A) Schematic of the different plasmids (pCAG, nCAG for transfection efficiency, EBNA vectors for reprogramming) delivered by either the Amaxa II nucleofector or the neon electroporator. S/MAR DNA vectors are depicted in detail and described in Figure 5 chapter 8.1. EBNA reprogramming vectors are depicted in detail and described in Figure 3 chapter 5.8.3. B) Pictures of AP staining of iPSC colonies emerging 30dpt with either the Amaxa II nucleofector or the neon electroporator. C) Bar diagram showing the transfection efficiency (left axis) and reprogramming efficiency (right axis) for the different vector delivery methods.

8.4.2 Urinary derived cells as an alternative cell source for hiPSCs

Our established protocol to obtain hiPSCs required NHDFs, the delivery of EBNA reprogramming vectors via the Amaxa II nucleofector, and a switch to hiPSC feeder-dependent conditions 8 days after nucleofection. Fibroblast derived iPSCs (FiPSCs) usually emerged around day 25-35 (Figure 14A). As a next step, we tested readily and non-invasively available cells isolated from urine as an alternative cell source to fibroblasts. In initial attempts, urine from healthy donors was centrifuged, washed with PBS and the pellets were resuspended and plated on gelatinised cell culture plates (Figure 14B). As culturing medium, a mixture of DMEM/F12 with the following supplements adapted from (T. Zhou et al., 2012; T. Zhou et al., 2011) were used: 10% FCS, 1% p/s, 1% amphotericin B, 10ng/ml epidermal growth factor, 5ug/ml insulin, 0.05ug/ml epinephrine, 36ng/ul hydrocortisone, 5ug/ml transferrin, 4pg/ml triode-L-thyronine. Primary cell colonies were obtained 7-14 days after plating, expanded, and used for reprogramming.
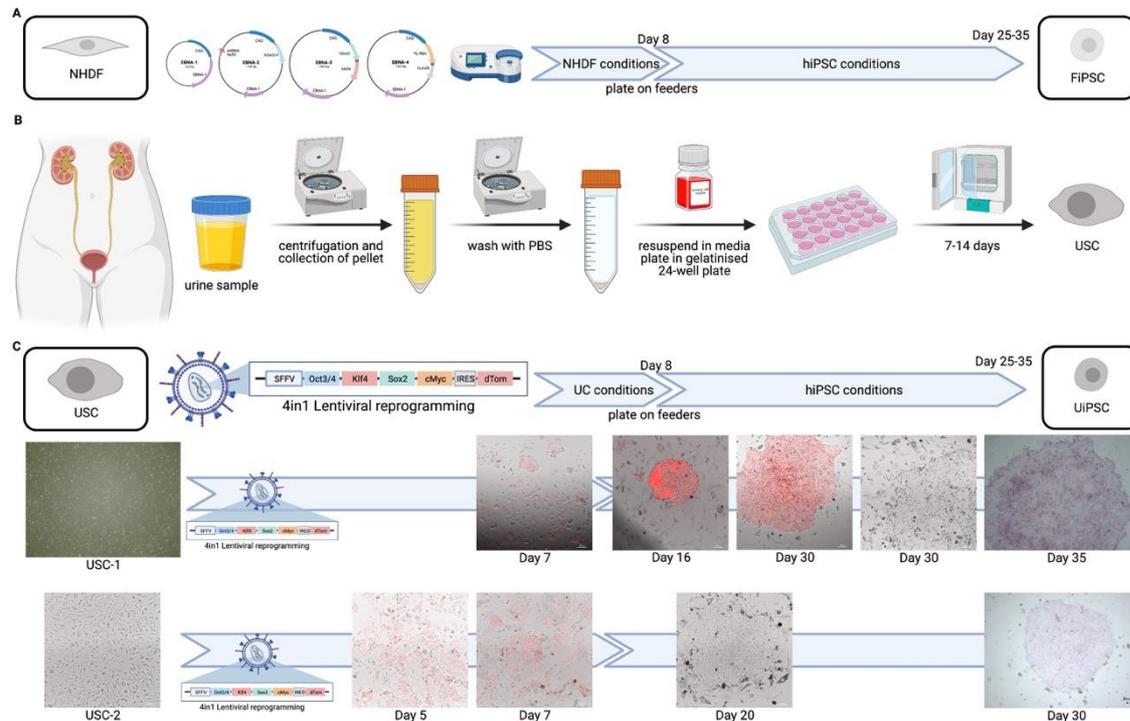
RESULTS

Figure 14: Cells isolated from the urine as cell source for hiPSCs.
A) Schematic of established reprogramming protocol for NHDFs using EBNA reprogramming vectors delivered by the Amaxa II nucleofector and feeder-dependent reprogramming obtaining fibroblast derived iPSCs (FiPSCs). EBNA reprogramming vectors are depicted in detail and described in Figure 3 chapter 5.8.3. B) Schematic of the process to isolate and culture cells from the urine C) Overview of reprogramming of urinary cells from two different donors using a 4-in-1 Lentiviral reprogramming vector. The Lentiviral 4in1 reprogramming vector is depicted in detail and described in Figure 2 chapter 5.8.1. Successful transduction was confirmed with a dTomato reporter gene. Fully reprogrammed urinary cell derived iPSCs (UiPSCs) were confirmed with AP staining.

When applying our established NHDF reprogramming protocol to USCs, the transfection efficiency with the Amaxa II was very low and no successfully reprogrammed urinary derived iPSCs (UiPSCs) were obtained after delivery of EBNA vectors. Therefore, different transfection methods for these cells were tested as well as a Lentiviral 4in1 reprogramming vector. While being routinely used for reprogramming of mouse cells, it encodes mouse reprogramming factors codon optimised for human cells. We were able to successfully transduce urinary cells with this reprogramming Lentivirus as confirmed by dTomato reporter gene expression (Figure 14C). After continuing the reprogramming in line with the established reprogramming protocol, morphologically distinct hiPSC colonies emerged. Completed reprogramming was monitored by the loss of dTomato expression in these hiPSC colonies upon inactivity of the SFFV promoter of the Lentiviral vector in pluripotent cells and was further confirmed with AP staining. This reprogramming experiment was attempted and successfully performed with cells isolated from two healthy donors.

Stem cell colonies from the second donor were expanded and used in the following experiments.

8.4.3 Optimisation of hiPSC culturing

Until this stage, human iPSC culturing was limited to feeder-dependent reprogramming and passaging by transferring scraped cell clumps on fresh feeder cells. This not only meant the co-culturing of mouse feeder cells with human iPSCs but also resulted in inconsistent colony sizes, unreliable passaging and made it impossible to deliver DNA specifically to the iPSCs. In a first approach, the generation of puro-resistant feeder cells was attempted to aid drug selection of transfected cells. However, to establish reliable protocols, the optimising of hiPSC culturing conditions was needed to enable direct modification of hiPSCs.

Therefore, different commercial feeder-free stem cell media, matrixes and passaging reagents were tested with the feeder-dependent reprogrammed NHDF and urinary cell-derived iPSCs (FiPSCs and UiPSCs). An initial test using the DMEM/F12 based feeder-dependent stem cell media as feeder-conditioned media on either vitronectin or geltrex as feeder-replacement caused differentiation of all plated stem cells. The combination of vitronectin with the commercial feeder-free TeSR-E8 media resulted in many differentiated cells. When switching to StemFit Basic02 media on vitronectin some hiPSC colonies with distinct morphology were obtained, while still leading to substantial differentiation (Figure 15A). However, the combination of this media with iMatrix Laminin-511 enabled the outgrowth of predominantly hiPSC colonies (Figure 15B). With this media-matrix combination, both passaging as clumps as well as single cells was feasible. Detachment in clumps was achieved by 1min incubation of ReLesR, subsequent aspiration and further 5-12min incubation at RT before cells were detached by shaking at 1200rpm (Figure 15B, C). While these clumps enabled passaging without pre-coating of some cell culture plates, reliable plating densities were difficult to obtain. Single-cell passaging was tried with StemPro accutase as well as TrypLE in combination with the presence of ROCK inhibitor for 24h after plating.

RESULTS

Especially accutase provided reproducible and reliable single-cell plating with consistent survival and plating density.
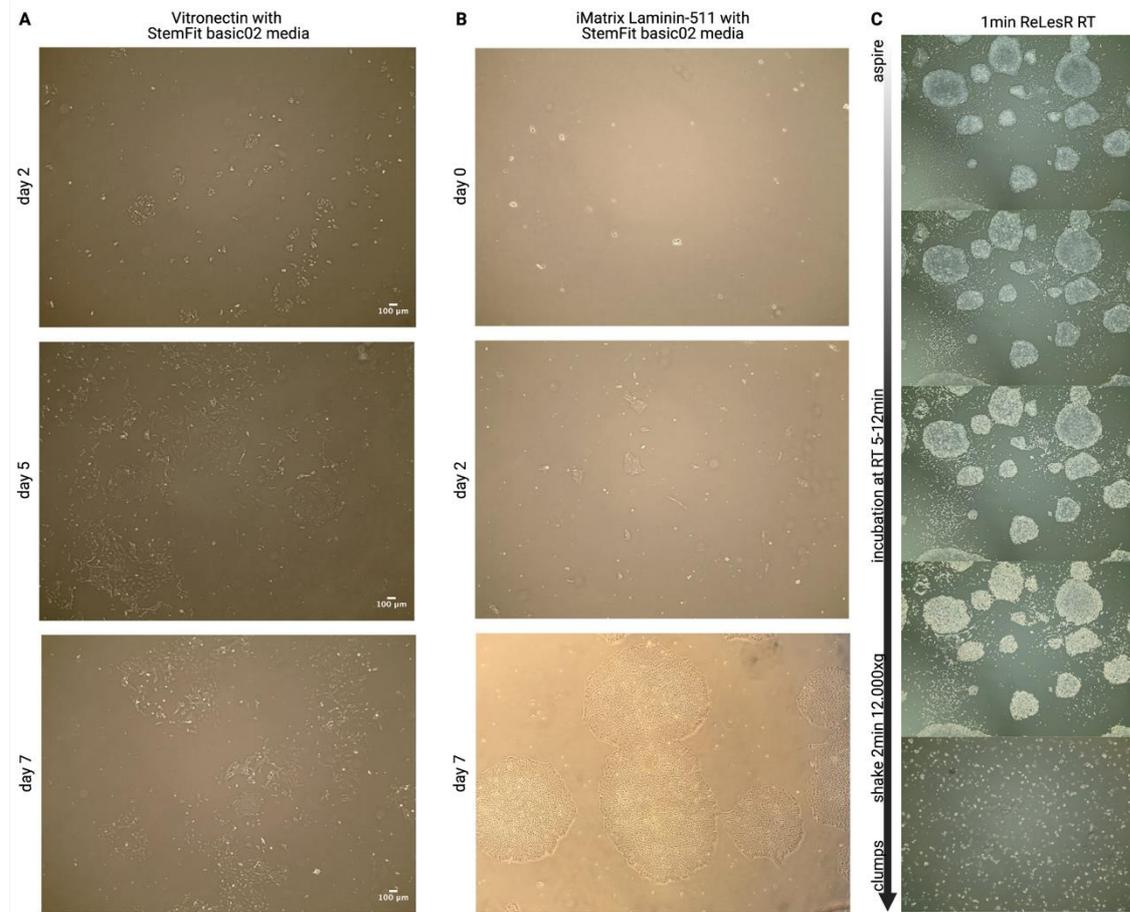

Figure 15: Overview of different tested feeder-free hiPSC culturing and passaging conditions.
A) Schematic of the process to isolate and culture cells from the urine. B) Schematic of established reprogramming protocol for NHDFs using EBNA reprogramming vectors delivered by the Amaxa II nucleofector and feeder-dependent reprogramming. C) Overview of reprogramming of urinary cells from two different donors using a 4in1 Lentiviral reprogramming vector. Successful transduction was confirmed with a dTomato reporter gene.

Despite following the manufacturer's described protocols, the premixing of iMatrix-Laminin511 with the cell suspension instead of precoating of culture-ware proved inconsistent with different plates and especially not successful when using glass-bottom plates. Thus, 1h precoating was routinely performed for single-cell culture. A switch to the cheaper iMatrix Laminin-511 silk during the project did not affect stem cell cultures.

To allow cryopreservation of stem cells, different freezing media as well as handling protocols were tested. The best stem cell outgrowth was achieved using the StemFit Basic02 stem cell media supplemented with 10% DMSO. Despite the required precoating of plates and the use of ROCK inhibitor (ROCKi) after thawing, very gentle

handling after thawing severely improved viability and outgrowth. Thereby, special attention was put on pipetting thawed cells with pipettes with wide tips, e.g., 5ml serological plastic pipettes instead of 1000ul pipette tips. Also, to avoid osmotic shock, fresh media was added dropwise to the thawed cell suspension and cell pellet.

RESULTS

8.5 Genetic modification of hiPSCs using GFP-S/MAR DNA vectors

8.5.1 Delivery of S/MAR DNA vectors to hiPSCs

In preliminary attempts, hiPSCs handled on feeder cells were scraped off, detached as single cells, nucleofected with the Amaxa II device and plated on puro resistant feeder cells to enable selection for S/MAR expressing cells. This handling caused stress to the cells and resulted in poor viability and colony outgrowth.

The optimised cell culturing conditions in a feeder-free system allowed for the first time the specific delivery of the S/MAR DNA vectors to only hiPSCs. Due to the nature of hiPSCs, the delivery of foreign genetic material is not straightforward and needed optimisation.

TrypLE single-cell plating and transfection of GFP-S/MAR DNA vectors with the Amaxa 4D nucleofector in the P3 buffer and CB-150 program resulted in very poor efficiency and viability, with remaining viable cells dying upon puromycin selection two days post transfection.

The transfection reagent lipofectamine stem provided the most reliable, efficient, and least toxic reagent to effectively transfect hiPSCs (Figure 16). The transfection efficiency of cells maintained in the feeder-free culturing media StemFit Basic02 and cells handled in optimem for 6h during transfection was compared. Importantly, the StemFit Basic02 media did not interfere with the transfection further simplifying the transfection protocol. Initially, mainly hiPSCs passaged in clumps were transfected with this reagent due to increased cell death upon single-cell transfection. However, this way plating density was inconsistent and cell numbers did not suffice to perform FACS analysis to get reliable transfection efficiency data.

Towards the end of the project, the DNA vector lab entered into a collaborative partnership with Maxcyte, and their electroporation device was available to evaluate its performance electroporating hiPSCs. An initial trial of the optimisation programs opt 6, 8 and 9 with single-cell hiPSCs resulted in transfection efficiency between

60-85% while maintaining 85-95% viability with the pCAG S/MAR DNA vector (data not shown), suggesting a promising option for future experiments.
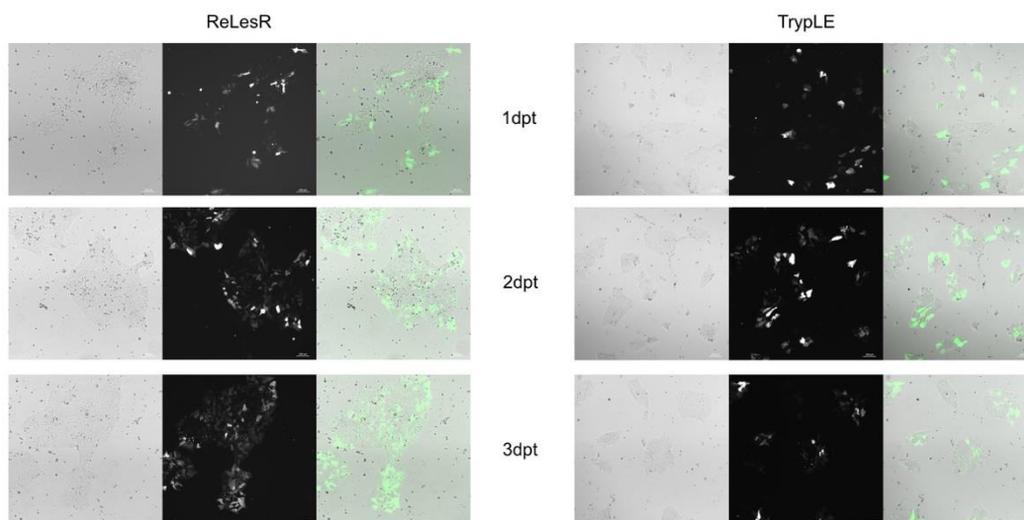

Figure 16: Transfection of clump and single-cell passaged hiPSCs with lipofectamine stem and a GFP-S/MAR DNA vector.
Feeder-free cultured fibroblast derived hiPSCs were either plated as clumps with ReLesR (left) or as single cells with TrypLE (right) in 24-well plates. The next day, cells were transfected with 2ug lipofectamine stem and 500ng of a GFP-S/MAR DNA vector. Successful transfection was monitored one, two- and three-days post transfection (dpt) under a fluorescence microscope.

8.5.2 Establishment of stable GFP-S/MAR-hiPSCs using puromycin selection

Combining the optimised culturing conditions and the ability to deliver S/MAR DNA vectors directly to hiPSCs we aimed to generate stable GFP expressing S/MAR-hiPSCs as a next step.

Initially, feeder-reprogrammed FiPSCs which were subsequently transitioned to a feeder-free culture were utilised (Figure 17A). The GFP-p2a-PuroR encoding S/MAR DNA vector P17 was delivered using lipofectamine stem after clump passaging and puromycin selection was applied subsequently for 3-4 weeks (Figure 17B, C). During the selection time, puromycin was withdrawn during passaging and when cell death was severe to aid outgrowth of cells resulting in stable GFP-expressing hiPSCs which for the first time were directly modified with S/MAR DNA vectors at the hiPSC stage. GFP expression and outgrowth of cells during selection and after selection are shown in Figure 17C.
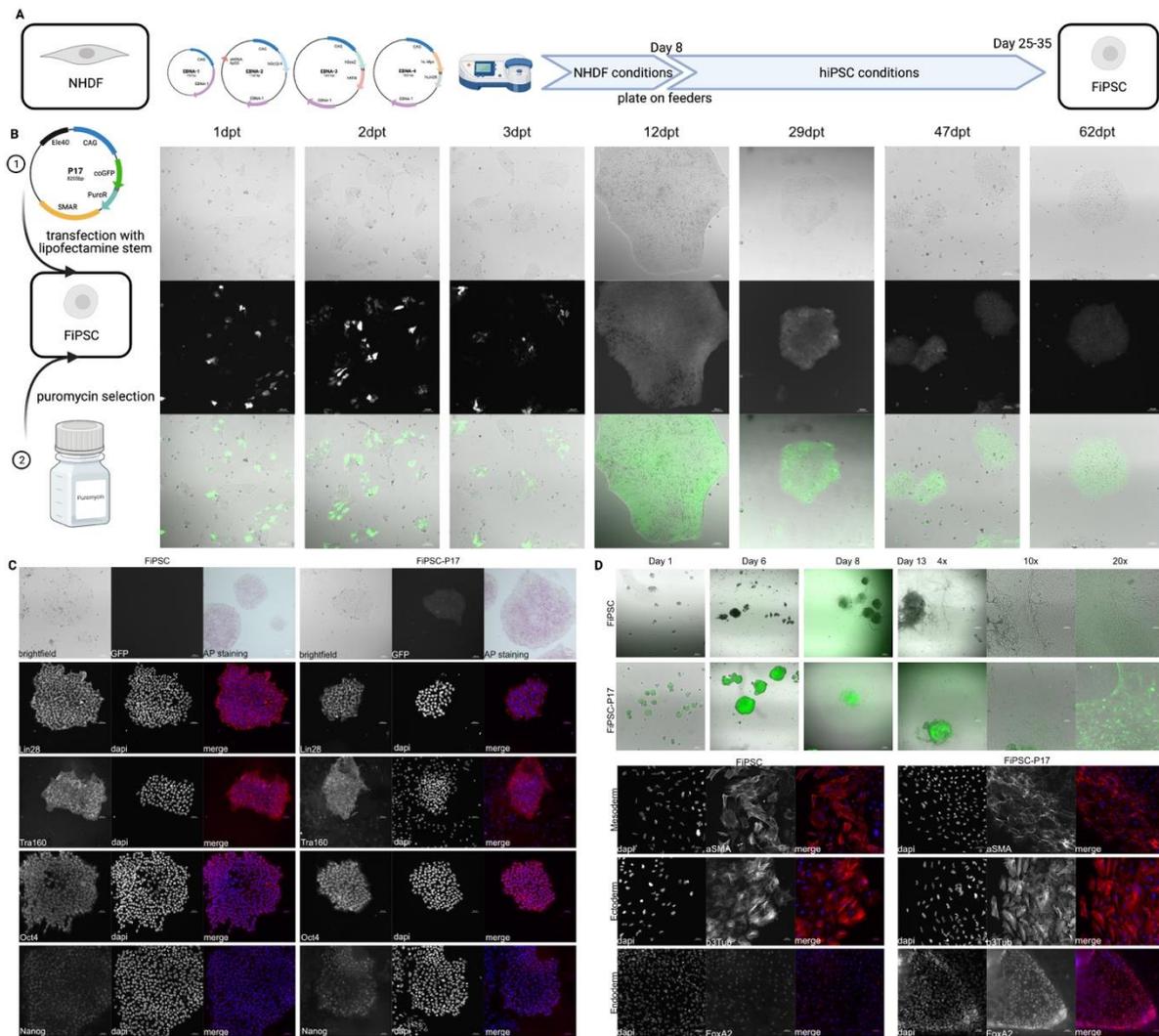
RESULTS

Figure 17: Establishment of P17 modified FiPSCs using puromycin selection and characterisation of parental and stable cell line.
A) Schematic of derivation of utilised FiPSCs: NHDFs were reprogrammed on feeder cells using EBNA vectors. EBNA reprogramming vectors are depicted in detail and described in Figure 3 chapter 5.8.3. B) GFP-p2a-PuroR encoding P17 S/MAR DNA vector was transfected to FiPSCs and puromycin selection was performed subsequently during the first 3-4 weeks of culturing when cells were growing normally. Fluorescence microscopy pictures 1-62dpt show outgrowth of GFP expressing hiPSC colonies and maintenance of GFP expression after drug selection. S/MAR DNA vectors are depicted in detail and described in Figure 5 chapter 8.1. C) AP staining as well as immunofluorescence pictures of stainings for the pluripotency markers Lin28, Tra160, Oct4 and Nanog confirmed expression in both parental (left) and S/MAR DNA vector modified FiPSCs (right). D) Parental and S/MAR modified FiPSCs were subjected to random differentiation via EB formation. Outgrowth of GFP expressing progenies upon attachment of EBs was verified by fluorescence microscopy pictures (top). Markers of the three germ layers Mesoderm (aSMA), Ectoderm (b3Tub) and Endoderm (FoxA2) were confirmed in these differentiated cells for both cell lines (bottom). IF antibody stainings are shown in red with blue DAPI counterstaining.

In line with previous experience with hiPSCs modified simultaneously to reprogramming, pluripotency marker expression (alkaline phosphatase, Lin28, Tra160, Oct4, Nanog) was confirmed for both the parental as well as the GFP-expressing FiPSCs suggesting no change in stem cell characteristics (Figure 17C). To further confirm this, cells were subjected to EB formation and random differentiation. Also, in this case, GFP expression was maintained in differentiated progenies of FiPSC-P17 cells which

grew out from the EBs (Figure 17D). The differentiated progenies of both parental and FiPSC-P17 cells were stained for germ layer markers, confirming the ability of both cell lines to differentiate into cells of the three germ layers.

We aimed to reproduce the successful establishment of directly modified hiPSCs with GFP-S/MAR DNA vectors. The setting was further elaborated by modifying not only the EBNA, feeder-reprogrammed FiPSCs, but also Lentiviral, feeder-reprogrammed UiPSCs (Figure 18A). Both cell lines were this time transfected using pCAG or nCAG and subjected to puromycin selection as described before (Figure 18B). We were able to obtain stable GFP-expressing hiPSC lines with both pCAG and nCAG S/MAR DNA vectors and from the different sources without observing any differences in overall behaviour.

Again, pluripotency characteristics of both parental and all established S/MAR iPSC lines were confirmed via AP and IF stainings. To preserve endogenous GFP, we modified the established IF protocol. Pluripotency marker staining with DAPI counterstaining and endogenous GFP was performed and suggested no difference in the tested cell lines (Figure 18C).

The previously performed random differentiation via EB formation was not very reliable and consistent. Thus, the capability of the different hiPSC lines to differentiate into cells of the three germ layers was tested by subjecting the cells to a guided trilineage differentiation utilising a commercial kit (Miltenyi). Differentiated cells were then stained with respective markers for the three germ layers while again maintaining endogenous GFP (Figure 18D). The successful differentiation supports that for both hiPSC types the S/MAR modification does not inhibit trilineage differentiation. Furthermore, maintenance of endogenous GFP during the staining shows more clearly than before the persistent GFP expression and thus S/MAR expression maintenance and persistency upon the 7-day differentiation process.

First attempts to further analyse the status of the S/MAR DNA vectors were conducted. An initial southern blot, PCR on gDNA as well as RCA of circular S/MAR DNA

RESULTS

was not successful (data not shown). However, together the data provides first proof, that the direct modification of hiPSCs from different sources is possible. In line with expectations and previous knowledge, the S/MAR DNA vectors did not alter their stem cell characteristics. Data performed with the pCAG and nCAG-S/MAR DNA vectors was published together with the work in mouse stem cells with shared first authorship (Roig Merino et al., in press).
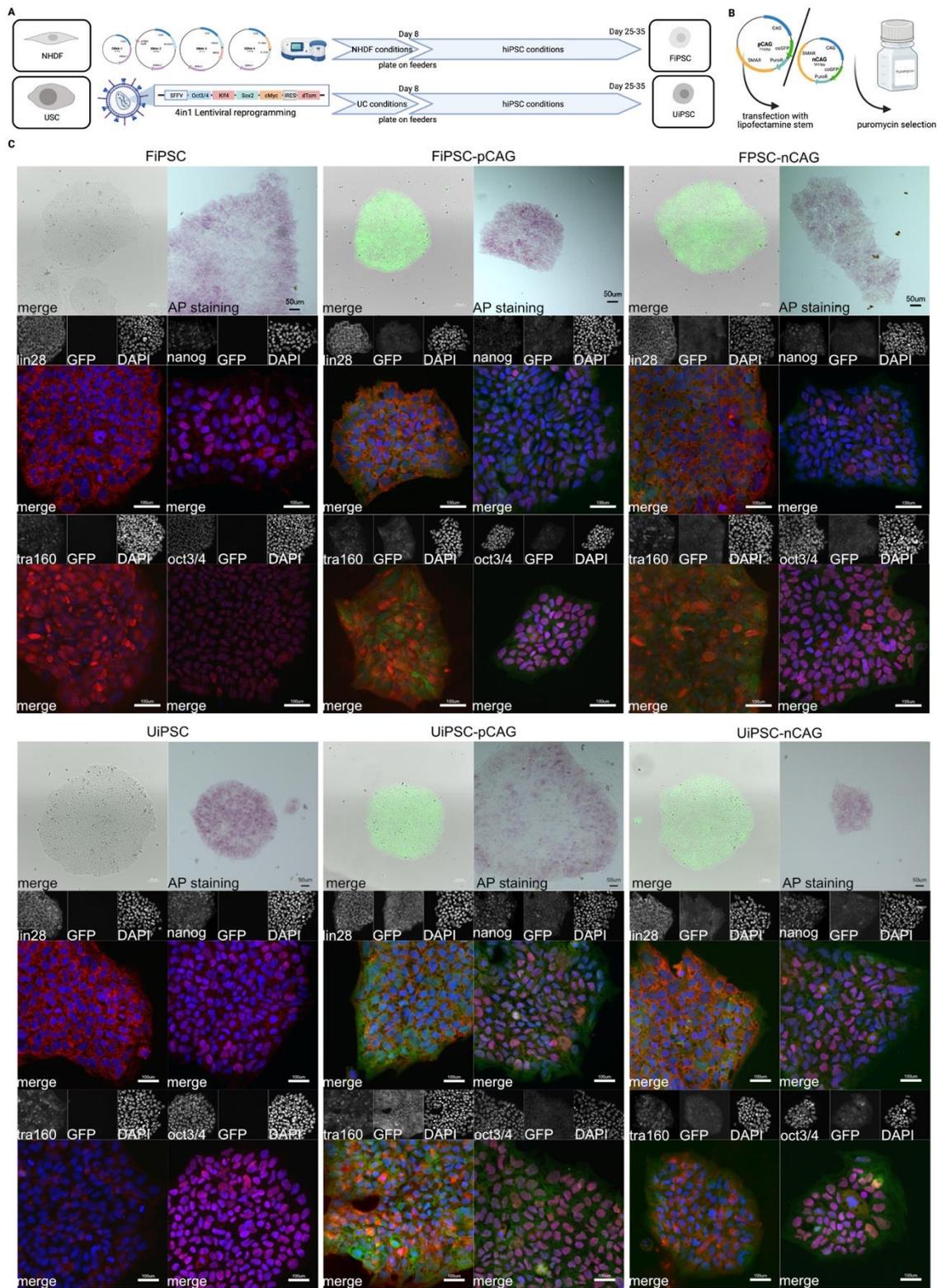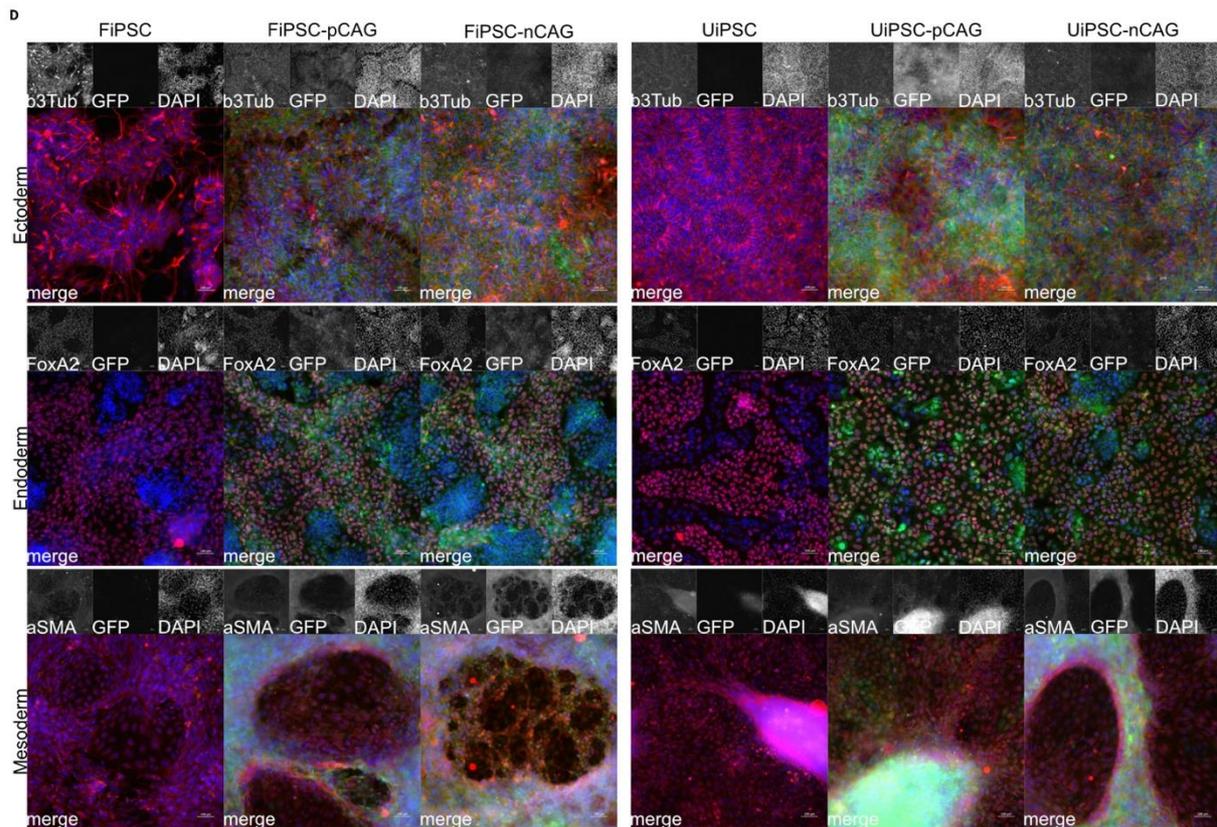
Figure 18: Establishment of pCAG and nCAG modified FiPSCs and UiPSCs using puromycin selection and characterisation of parental and stable cell line.
A) Schematic of derivation of utilised iPSCs: NHDFs or USCs were reprogrammed on feeder cells using EBNA vectors. EBNA reprogramming vectors are depicted in detail and described in Figure 3 chapter 5.8.3. B) GFP-p2a-PuroR encoding pCAG or nCAG S/MAR DNA vectors were transfected to the two hiPSC lines and puromycin selection was performed subsequently during the first 3-4 weeks of culturing when cells were growing normal. S/MAR DNA vectors are depicted in detail and described in Figure 5 chapter 8.1. C) Fluorescence microscopy pictures show GFP expression in established hiPSC colonies and no expression in parental cell lines. AP staining as well as immunofluorescence pictures of stainings for the pluripotency markers Lin28, Tra160, Oct4 and Nanog confirmed expression in both parental (left) and S/MAR DNA vector modified iPSCs (right). D) Parental and S/MAR modified iPSCs were subjected to trilineage differentiation media (Miltenyi). Immunofluorescence staining confirmed successful differentiation into cells of all germ layers (red, Mesoderm (aSMA), Ectoderm (b3Tub) and Endoderm (FoxA2)) while endogenous GFP was maintained throughout the differentiation and the staining protocol (green).

8.5.3   Establishment of stable S/MAR-hiPSCs using FACS sorting

Based on the S/MAR establishment experience with cell lines, we aimed to also handle the hiPSCs as gently as possible and establish stable S/MAR expressing cell lines in a drug-free setting. A first attempt with the initially performed clump plating for transfection resulted in a fast loss of GFP transgene expression within the first week of culture. Notably, with clump culture passaging some GFP positive cells remained attached to the old plate while more of the not expressing cells were passaged and further cultured. In combination with cell culture optimisation as described before in

chapter 8.4.3, single-cell passaging was routinely established and facilitated a protocol
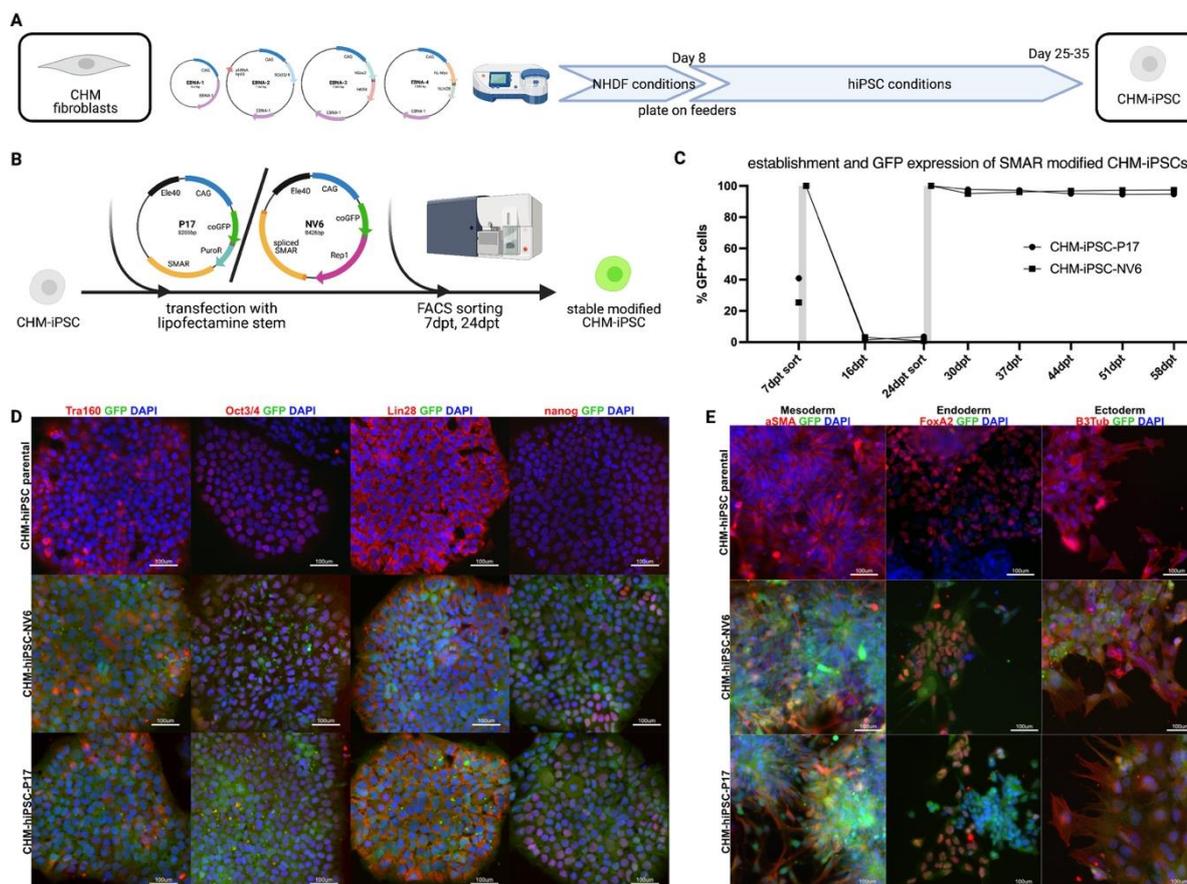
for successful FACS sorting of hiPSCs.

Figure 19: First drug-free establishment of S/MAR modified hiPSCs.
A) Choroideremia (CHM) patient derived fibroblasts were reprogrammed on feeders using EBNA reprogramming vectors. EBNA reprogramming vectors are depicted in detail and described in Figure 3 chapter 5.8.3. B) CHM-iPSCs were transfected with the P17 GFP-S/MAR DNA vector or the Nanovector™ NV6, encoding the functional Rep1 gene (credits: Patrick V. Almeida). Single, alive, GFP expressing cells were then sorted 7dpt and 24dpt and further cultured to obtain a pure, S/MAR DNA vector expressing hiPSC population. S/MAR DNA vectors are depicted in detail and described in Figure 5 chapter 8.1. C) FACS data overview of the population size of GFP+ alive, single hiPSCs (in %) is depicted during the establishment and FACS analysis beyond the sorting. D) Immunofluorescence staining for the pluripotency markers Lin28, Tra160, Oct4 and Nanog confirmed expression in parental (top) and both S/MAR DNA vector modified iPSCs. E) Immunofluorescence staining confirmed successful guided differentiation into cells of all germ layers (red, Mesoderm (aSMA), Ectoderm (b3Tub) and Endoderm (FoxA2) while endogenous GFP was maintained throughout the differentiation and the staining protocol (green).

In collaboration with Prof. Mariya Moosajee's lab at the Crick Institute, London,

fibroblasts from Choroideremia (CHM) patients were obtained and reprogrammed in

our lab with the established protocol using EBNA vectors and feeder-dependent

culturing (Figure 19A). The obtained CHM-iPSCs were then transfected with the GFP

S/MAR DNA vector P17 or the Nanovector™ NV6. The latter had been designed and

cloned based on P17 and previous Rep1 vectors (Dr Alicia Roig Merino) by Dr Patrick V.

Almeida and synthesised by Nature Technologies (NTC). It encodes the genetic

information for GFP-p2a-Rep1, with defective Rep1 being the responsible gene for the CHM disease.

Seven and 24dpt, alive, single, GFP expressing CHM-iPSCs were FACS sorted and further expanded and cultured (Figure 19B). Because of the low cell number, transfection efficiency was not measured 24 or 48h after transfection, but a low GFP expression of 20-40% was obtained 7dpt. Intermediate FACS analysis showed a severe drop in expression within the next nine days (16dpt) to 1-3% GFP+ population which showed little change until the next sorting 24dpt. The GFP expression was further monitored until 58dpt and stayed at a high level of 95-98% GFP expressing alive single cells suggesting a successful drug-free establishment of S/MAR DNA vectors in these CHM-iPSCs (Figure 19C). As with the previously established hiPSC-S/MAR and parental cell lines, pluripotency marker expression with maintained endogenous GFP was confirmed (Figure 19D). Also, all cells were subjected to trilineage differentiation and the capacity to differentiate while maintaining GFP expression was confirmed by IF staining of germ layer markers (Figure 19E). The reprogramming and establishment of these cells were performed together with Dr Patrick V. Almeida and the master's thesis student Bianca Maria Carrara.

8.6   Optimised generation of S/MAR modified urinary derived iPSCs

Our results so far established that S/MAR DNA vectors are a suitable tool for the genetic modification of hiPSCs from different sources. The combination of our optimised vectors for gene therapy and their utilisation in hiPSCs opens the door towards their application for genetically modified cell therapies. Thus, we further elaborated the implemented protocols and aimed to use clinically relevant conditions such as animal- and xeno-free culturing. Specifically, we focused on the potential application in BHD disease and the generation of advanced BHD cell models.

8.6.1   Transition to feeder-free reprogramming

Usually, the primary cell source for iPSCs, particularly primary fibroblasts, are a limited resource. With our current reprogramming approach, 500.000 cells were nucleofected with EBNA vectors and cultured. On day 8 however, only 30.000 cells were plated per well of a 6wp on hiPSC culturing conditions while the remaining majority of nucleofected cells were not further cultured. To save resources as well as time and simplify the obtention of low passage reprogrammed hiPSCs we tested if we could freeze the remaining cells from day 8 and successfully continue reprogramming at a later timepoint (Figure 20A).

Since cells were cultured in hiPSC conditions without special supplements from day 8 on, we hypothesised that also feeder-free hiPSC conditions could enable successful reprogramming. We tested our feeder-dependent next to three different feeder-free conditions using either iMatrix Laminin-511 or Geltrex as a coating agent and StemFit Basic02 or TeSR-E8 as feeder-free culturing media (Figure 20B). NHDFs nucleofected with EBNA reprogramming vectors frozen on day 8 of reprogramming were used for this experiment. Positive AP staining on day 28 confirms that reprogramming after thawing was successful in all conditions. FiPSCs reprogrammed in the iMatrix Laminin-511 and StemFit Basic02 media conditions were successfully expanded and cultured, further emphasising the feasibility of these reprogramming conditions for an animal- and xeno-free generation of hiPSCs.
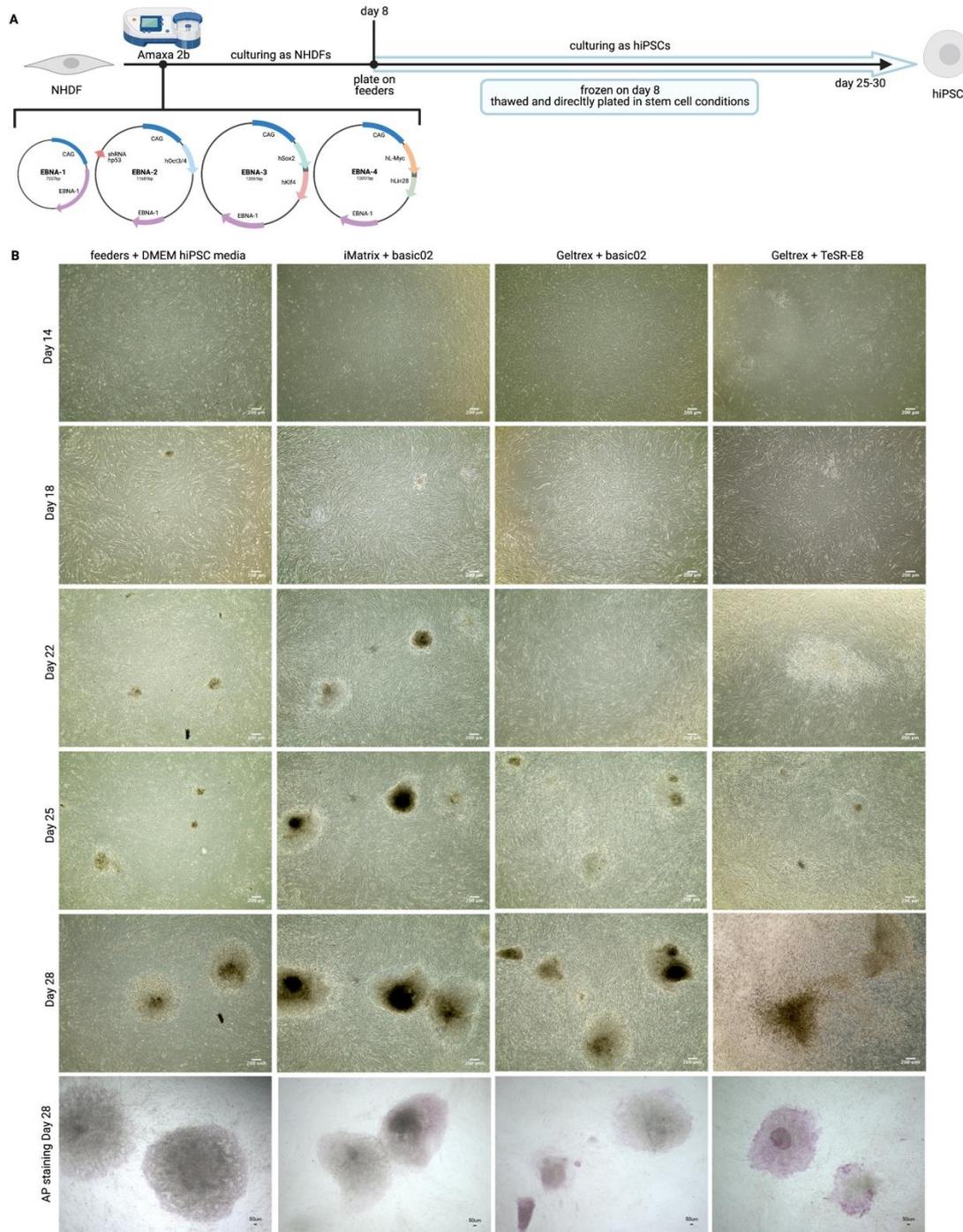
Figure 20: Feeder-free reprogramming establishment and optional freezing.
A) Schematical overview: EBNA reprogramming vectors were nucleofected via the Amaxa II. Cells were frozen on day 8 and reprogramming in stem cell culturing conditions was continued upon thawing. EBNA reprogramming vectors are depicted in detail and described in Figure 3 chapter 5.8.3. B) Brightfield pictures monitoring the reprogramming of the thawed NHDFs on day 14, 18, 22, 25 and 28 in culture after EBNA vector delivery. Different stem cell culture conditions were applied: 1) established feeder cells + DMEM hiPSC media, 2) iMatrix Laminin-511 + StemFit Basic02, 3) geltrex + StemFit Basic02, 4) geltrex + TeSR-E8. AP staining confirms successful reprogramming in all conditions.

8.6.2 hiPSCs derived from urinary cells using episomal vectors in feeder-free conditions

After demonstrating that we could successfully isolate and reprogram USCs from two donors we next aimed to further standardise the isolation conditions and performed 18 more isolation attempts from a total of 8 healthy donors with different culturing conditions. Isolation was considered successful if cell colonies were visible within two weeks after isolation. An overview is shown in Figure 21A.

First, we investigated whether gelatine coating aids the attachment of cells in the urine with our previously used homemade mixture of DMEM/F12 with the described supplements from chapter 8.4.2 and different FCS concentrations. While we did not obtain USC colonies in uncoated plates from three isolations, gelatine coating seemed to facilitate cell attachment and colony outgrowth: seven out of eight isolations were successful with 10% FCS in the media as well as the single isolation performed with only 5% FCS in the media.

We next tried the commercially available renal epithelial cell growth basal medium (REBM, Lonza). Here, an aliquot of the basal media was freshly supplemented with the respective recommended amount of REGM SingleQuots supplements. Like the homemade USC media, this kit contained hEGF, insulin, hydrocortisone, gentamicin and amphotericin, transferrin, triiodothyronine, and epinephrine. FCS was included for a final concentration of only 0.5%. Surprisingly, direct culturing of urine pellet in this media did not lead to USC outgrowth in two attempts.

Finally, we followed the isolation protocol as published in (Mulder et al.) using a mixture of both media: Washed urine pellets were resuspended in DMEM/F12 media supplemented with the recommended amount of REGM SingleQuots supplements and 10% FCS. After four days of culturing media was switched to the fully supplemented REBM media and thus FCS levels were dropped severely. With this approach, in line with the publication, all five isolation attempts were successful and isolated USCs could easily be expanded.

RESULTS

Representative brightfield pictures of obtained colonies 10 days after isolation are shown in Figure 21B. As reported, in most isolations two distinct cell types were obtained and could be distinguished by morphology and growth behaviour, as shown representatively in the expanded USCs in Figure 21C (Shi & Cheung, 2021; T. Zhou et al., 2012).
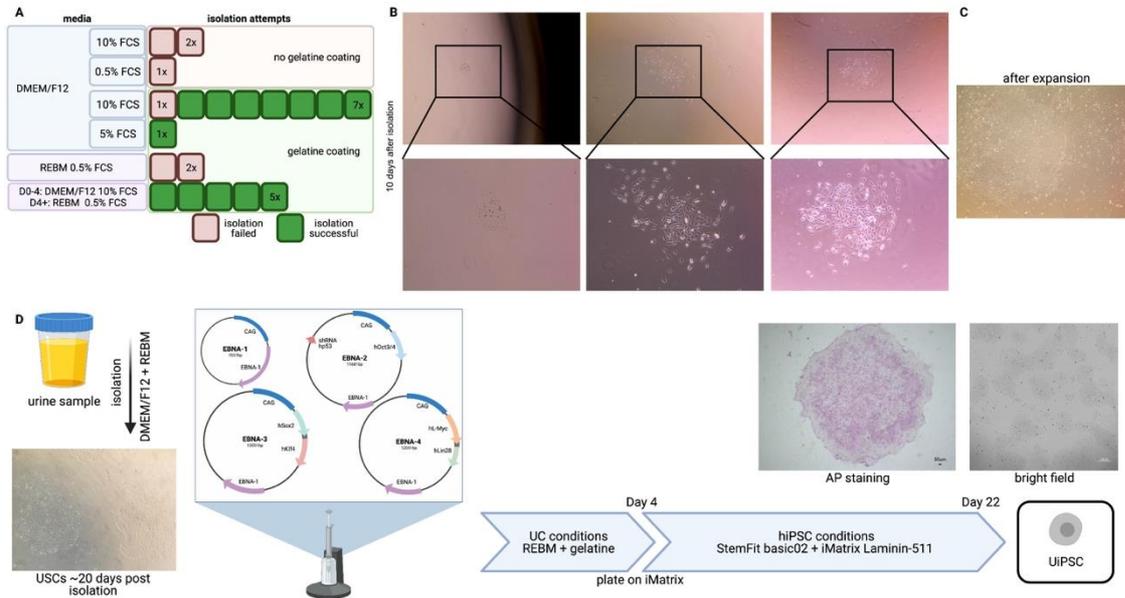

Figure 21: Optimised USC isolation and feeder-free episomal reprogramming.
A) Overview of successful and failed USC isolations with different media and coating conditions. B) Brightfield pictures of USCs colonies 10 days after isolation and C) after expansion. Two cell types can be distinguished by their morphology. D) Schematical overview of optimised USC isolation and reprogramming: USCs were isolated with DMEM/F12 + REBM medium. After ~20 days of expansion EBNA vectors were delivered by neon electroporation and feeder-free reprogramming with StemFit Basic02 and iMatrix Laminin-511 was performed. Picture of positive AP staining (left) as well as brightfield picture of obtained feeder-free UiPSCs (right) are shown. EBNA reprogramming vectors are depicted in detail and described in Figure 3 chapter 5.8.3.

With the previously obtained USCs which were isolated with the homemade DMEM/F12 media, efficient DNA vector delivery for reprogramming was not successful and instead Lentiviral reprogramming was performed. Further optimisations using the neon electroporator resulted in poor efficiency and viability.

Interestingly, in the publication where USCs were isolated with the combination of DMEM/F12 and REBM media, these cells were successfully reprogrammed with EBNA reprogramming vectors using the neon electroporator (Mulder et al., 2020). We followed this protocol with USCs from two different donors which were isolated with the two media. This time, high viability, as well as efficiency, was achieved using a GFP vector (data not shown). We were now able to deliver the episomal EBNA

reprogramming vectors to the two different USCs and subjected the cells to the previously tested feeder-free reprogramming conditions using iMatrix Laminin-511 and StemFit Basic02 (Figure 21D). Positive AP staining and outgrowth of distinct hiPSC colonies, devoid of differentiating areas, were obtained.

Together, an optimised protocol for successful and reproducible isolation, episomal and xeno-free reprogramming of cells from the urine was established. We further focused on these optimally derived UiPSCs from two donors (donor 6 and 7, U6iPSC and U7iPSC, respectively) in the following chapters.

8.6.3   Comparison of transfection efficiency based on vector type and cell line

We next used these episomal and xeno-free generated fibroblast- and USC-derived iPSCs to further optimise and quantify the DNA vector delivery. As reported before in chapter 8.5.1, transfection with lipofectamine stem via clump plating resulted in inconsistent plating densities, which also affected the viability and transfection efficiency. We thus thrived to optimise a lipofection protocol to transfect single-cell plated hiPSCs. To increase survival, transfection in combination with ROCKi was tested but proved to inhibit transfection. Upon further optimisation and increasing of initial plating density cell death was decreased and reliable transfection with increased and consistent cell number was achieved.

We transfected FiPSCs, U6iPSCs and U7iPSCs with the GFP-S/MAR DNA vectors P17, pCAG and nCAG and our tested FLCN-S/MAR DNA vectors P13, P14 and P15. To enable better comparison the molar ratio of the different vectors was calculated, and the same number of vector molecules was utilised for transfection. Transfection efficiency was measured by GFP expressing, alive, single cells 2dpt via FACS analysis. The data from three independent biological replicates were obtained (Figure 22A).
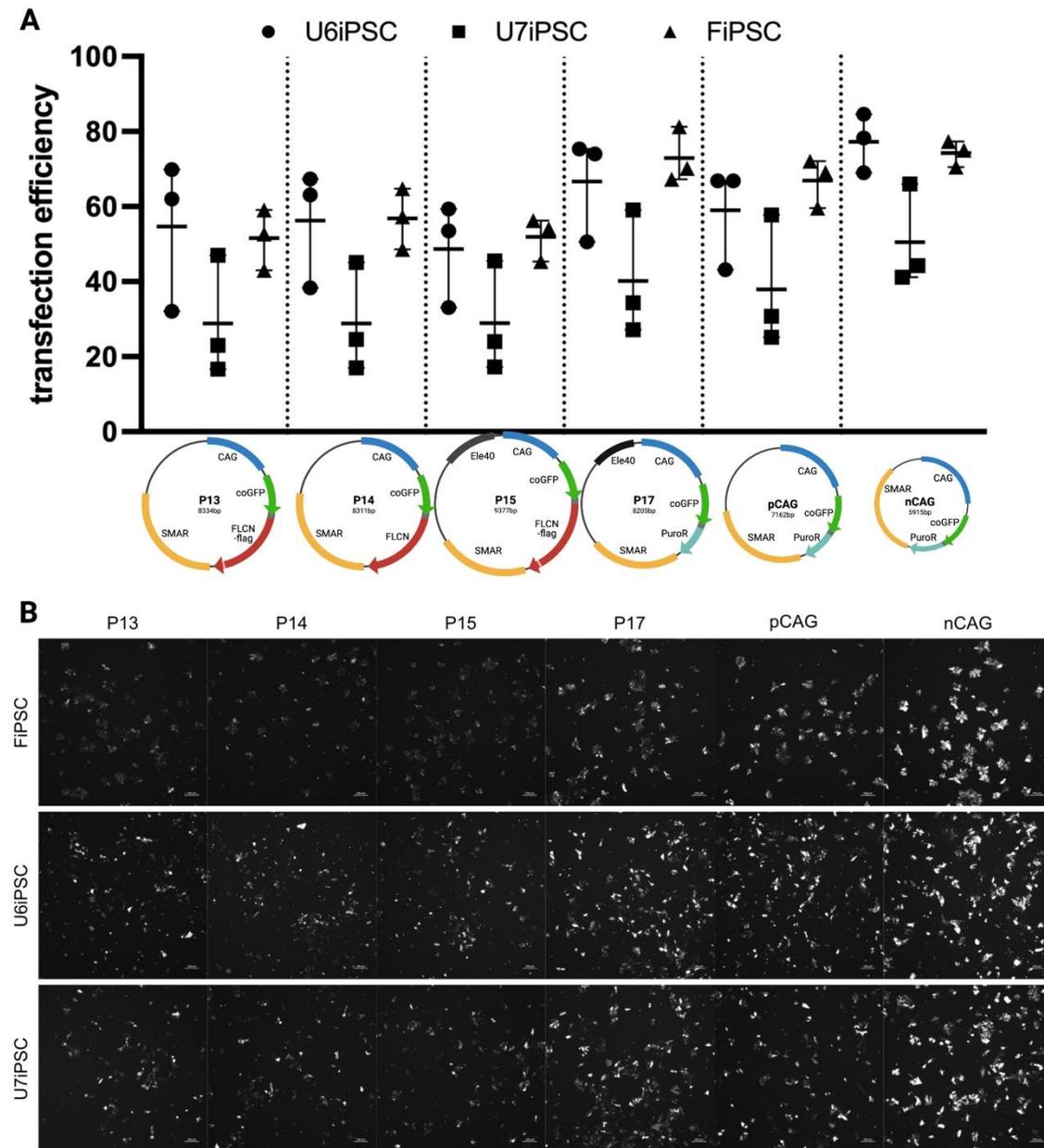
RESULTS

Figure 22: Transfection efficiency comparison of GFP and FLCN S/MAR DNA vectors in optimally derived FiPSCs and UiPSCs.
A) Transfection efficiency by GFP expression of single, alive cells in FACS analysis of three biological replicates in the three different cell lines U6iPSC, U7iPSC and FiPSCs are shown. Cells were transfected using lipofectamine stem with a variety of FLCN expressing S/MAR DNA vectors (P13, P14 and P15, left) and GFP control S/MAR DNA vectors (P17, pCAG and nCAG, right). B) Representative fluorescence microscopy pictures of GFP expression 2dpt.

A transfection efficiency within a range of 20-80% was achieved. Despite aiming for consistent conditions, the efficiency varied profusely between the replicates (up to 40%). However, within each cell line, a tendency of higher efficiency upon transfection of GFP vectors (right) compared to FLCN expressing vectors (left) was detectable. As expected, despite the molecular correction the Nanovector™ nCAG performed better than the plasmid vectors. More distinctively, with U7iPSC consistently showing the

smallest GFP+ population, transfection efficiency seemed to be influenced strongly by the used cell line suggesting donor specificity. The efficiency trends are further accentuated in the representative fluorescence microscopy pictures in Figure 22B. Notably, with U6iPSCs and FiPSCs performing similarly, the data did not show any disadvantage of utilising USC-derived hiPSCs.

8.6.4 Establishment and characterisation of GFP- and Folliculin-overexpressing S/MAR DNA vector modified hiPSCs

We continued to focus on the USC-derived cell lines. Stable GFP, as well as FLCN S/MAR DNA vector expressing cell lines were established with our drug-free approach using FACS sorting. Cell lines with all three FLCN vectors (P13, P14 and P15) as well as two of the three GFP vectors (P17, pCAG or nCAG) were established once for U6iPSCs and two independent times for U7iPSCs (Figure 23A, B, C). Data performed with the pCAG and nCAG-S/MAR DNA vectors was published together with the work in mouse stem cells with shared first authorship (Roig Merino et al., in press).

First, the transfected U7iPSCs and U6iPSCs were sorted 11dpt and 27dpt. Subsequently, the GFP expression was monitored via FACS analysis every 5 days after passaging until 153dpt (Figure 23A, B). As previously seen, the GFP expressing population dropped after an initial transfection efficiency between 43-83% down to 1-10% on day 11. The obtained expanded GFP+ cell population again dropped GFP expression, however, this time some cell lines mostly lost GFP expression while others remained GFP expression in up to 60% of alive single cells.

Interestingly, the long-term monitoring revealed a loss in GFP expression after establishment in half of the established cell lines. Respective cell lines are highlighted in red (FLCN S/MAR DNA vector) or green (GFP S/MAR DNA vector) for easier visualisation. Notably, this loss in transgene expression occurred independently in both cell lines and with all vectors (Figure 23A, B).

RESULTS

A third establishment experiment, again utilising U7iPSCs was performed (Figure 23C). This time, the GFP expressing cells were sorted and further cultured 6, 12, 28 and 44dpt. The establishment FACS data further supports previous data from cancer cell line and the hiPSC drug-free establishments: at 6dpt GFP expression was maintained in many cells, while 12dpt only a very small percentage of cells expressed GFP. In line with the previous establishments of UiPSCs with sorting done at 11dpt and 27dpt, some of these cells lost GFP expression, but 40-98% of cells kept expressing GFP at 28dpt. This suggests a significant impact of even one day during these first 30 days of culturing and seems to be independent of the utilised vector. Subsequent FACS monitoring of the latter established cells demonstrated almost pure transgene expressing cell populations in all five cell lines until 170dpt, with only the P13-U7iPSC cell line slightly losing GFP expression after 120dpt.

In summary, the reason for loss of transgene expression does not seem to be related to vector type or cell donor. However, the latter data from Figure 23C suggests that more sorting timepoints during the establishment might result in better expression stability in the S/MAR modified cells. It should be noted that during the first sorting the sorting gates for GFP expression were not very stringently chosen, since the GFP+ population had a wide GFP intensity distribution. However, the GFP expressing population was very distinct and stringently gated during the second sorting. We considered the possibility that the second sorting might have had doublets or included few not-GFP-expressing cells, which then overgrew the GFP expressing population upon further culture.

We investigated this by sorting the five cell lines which lost GFP expression from the establishment with two sortings 98dpt (Figure 23A+B, highlighted in green or red). This time, for further analysis both the GFP expressing (GFP+) and not expressing (GFP-) cell populations were sorted and maintained. The data of these five cell lines before sorting and FACS monitoring from the GFP+ cells after sorting are shown in Figure 23D. Despite a stringent sorting 98dpt, two out of the five cell lines continued to further lose GFP expression.

To investigate if the cells which are not expressing the S/MAR DNA vector have a growth advantage over S/MAR expressing cells, we compared the cell growth of the sorted GFP+ and GFP- cells with a live growth assay using the xCelligence machine. Cells were plated in the required gold-coated plate in technical triplicates and cultured in the routine culturing scheme for five days. The data from the exponential growth rate between the media changes (24h and 96h after plating) was selected and baseline corrected (Figure 23E). The sorted populations of each of the five cell lines are depicted in green (GFP+) and black (GFP-) for easier visualisation in separate graphs. The mean of the technical triplicates indicates a higher growth rate of the GFP- populations of all five cell lines. However, contrary to expectations, the two cell lines U6iPSC-P14 and U7iPSC-pCAG, which show a loss in GFP expression after sorting show only a slight growth advantage of GFP- population while U6iPSC-P15 and U6iPSC-P17 demonstrate a strong growth difference of the two populations but showed maintenance of GFP expression after sorting. Together, this suggests that the growth advantage of not expressing cells versus S/MAR DNA vector expressing cells seems to be not the main driver of loss of transgene expressing cell populations.

Finally, we investigated whether the S/MAR DNA vector was detectable not only in the GFP+ but also in the GFP- cell population. Since the GFP- population was determined to not express the S/MAR DNA vectors 98dpt but were part of the GFP expressing cell population sorted 27dpt, we assumed that these cells must have received the S/MAR DNA vector but might have silenced its expression. The gDNA of both cell populations from the five cell lines as well as the parental U6iPSC line was isolated and used as a template to PCR amplify the GFP-FLCN (left) or GFP (right) sequence (Figure 23F). The utilisation of different primer concentrations (top and bottom) allowed the detection of the GFP-FLCN sequence in all FLCN-S/MAR transfected cell lines in both, the GFP- and the GFP+ population and distinct negative and positive controls. Detection of the GFP sequence was not easily visible in all samples but was again obtained from gDNA samples of the GFP+ as well as the GFP- cell populations.
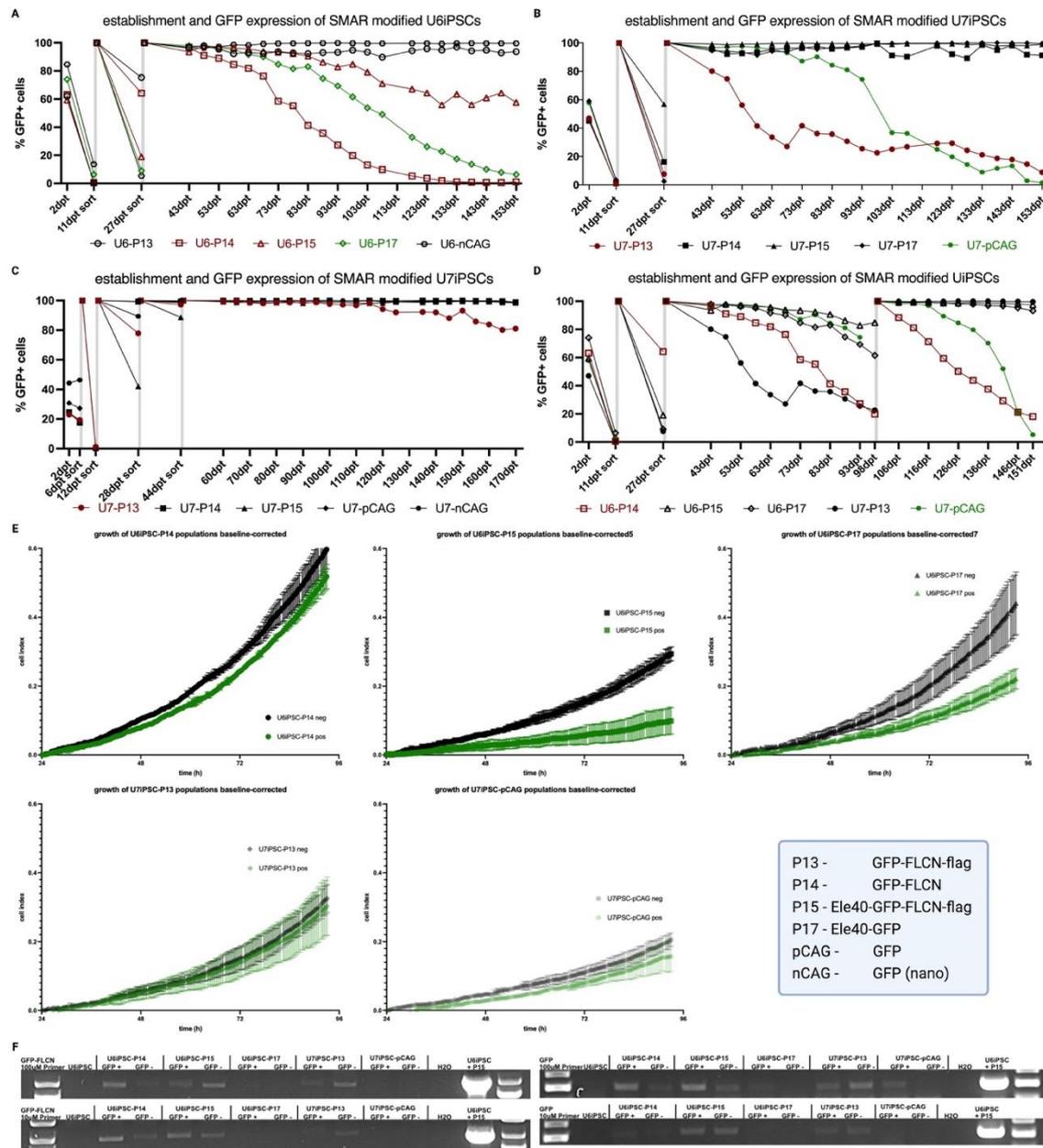
Figure 23: Establishment and stable modification of FLCN and GFP S/MAR DNA vectors in UiPSCs.
A-D) FACS data of GFP expressing, single alive cells during sorting and analysis are shown. Cell lines which are losing GFP expression are highlighted in green (GFP-S/MAR DNA vector) or red (FLCN-S/MAR DNA vector). The blue box highlights the features of the different used S/MAR DNA vectors. A) Data of U6iPSCs transfected with P13, P14, P15, P17 and nCAG vectors. B) Data of U7iPSCs transfected with P13, P14, P15, P17 and pCAG vectors. C) Data of U7iPSCs transfected with P13, P14, P15, pCAG and nCAG. D) Data of the five highlighted cell lines from A) and B) are shown until a third sorting 98dpt. From 98dpt, the data of the sorted and further maintained GFP expressing population is shown. E) Data of growth analysis with the xCelligence machine is shown. Only the exponential growth between 24h and 96h after plating was chosen and baseline corrected. The mean and standard deviation of technical replicates is depicted. F) Agarose gel of PCR on gDNA of the GFP+ and GFP- population of cells from D) after sorting 98dpt. Left: Primers enclosing the GFP-FLCN sequence were used. Right: Primers enclosing the GFP sequence were used. Either 100uM (top) or 10uM (bottom) primer concentration (top) was used.

We moved on to thoroughly characterise the established cell lines. Besides the FACS monitored GFP expression, qPCR further validated the presence of GFP mRNA in all stable cell lines but not in the parental lines (Figure 24A, B). Intron-spanning primers for FLCN further confirmed also FLCN overexpression in the P13, P14 and P15 cell lines and only weak expression in parental lines. Additionally, a primer set with one primer binding at the FLAG-tag sequence specifically detected vector expression in the FLCN-flag encoding P13 and P15 vectors. The intensity of transgene expression varied without an obvious pattern. Data from the U6iPSC and U7iPSC cell lines which were established with two FACS sortings are shown (Figure 23A, B).

Positive AP staining and expression of the pluripotency markers Lin28, Nanog, Tra160 and Oct3 expression was confirmed for the parental and for all established cell lines. Representative IF pictures for GFP-S/MAR DNA vector cell lines and FLCN-S/MAR DNA vector cell lines are shown in Figure 24C. All established cell lines were subjected to trilineage differentiation and in line with previous results subsequent IF staining confirmed successful differentiation into cells of the three germ layers while maintaining GFP expression. Again, representative pictures for GFP-S/MAR DNA vector cell lines and FLCN-S/MAR DNA vector cell lines are shown in Figure 24D.

We next aimed to obtain quantitative data for the maintenance of transgene expression upon differentiation. Thus, for the selected cell lines U6iPSC-P15, U7iPSC-P15, U7iPSC-P17, U7iPSC-pCAG and U7iPSC-nCAG also a larger number of cells were differentiated with the trilineage differentiation kit. After 7 days of differentiation, differentiated cells were harvested GFP expression of alive, single cells were analysed. The long-term culture FACS data showed a loss in GFP expression over time in some established cell lines. To verify that the loss of expression is solely due to differentiation, control hiPSCs that were not subjected to differentiation but routinely cultured were analysed together with the differentiated sample.
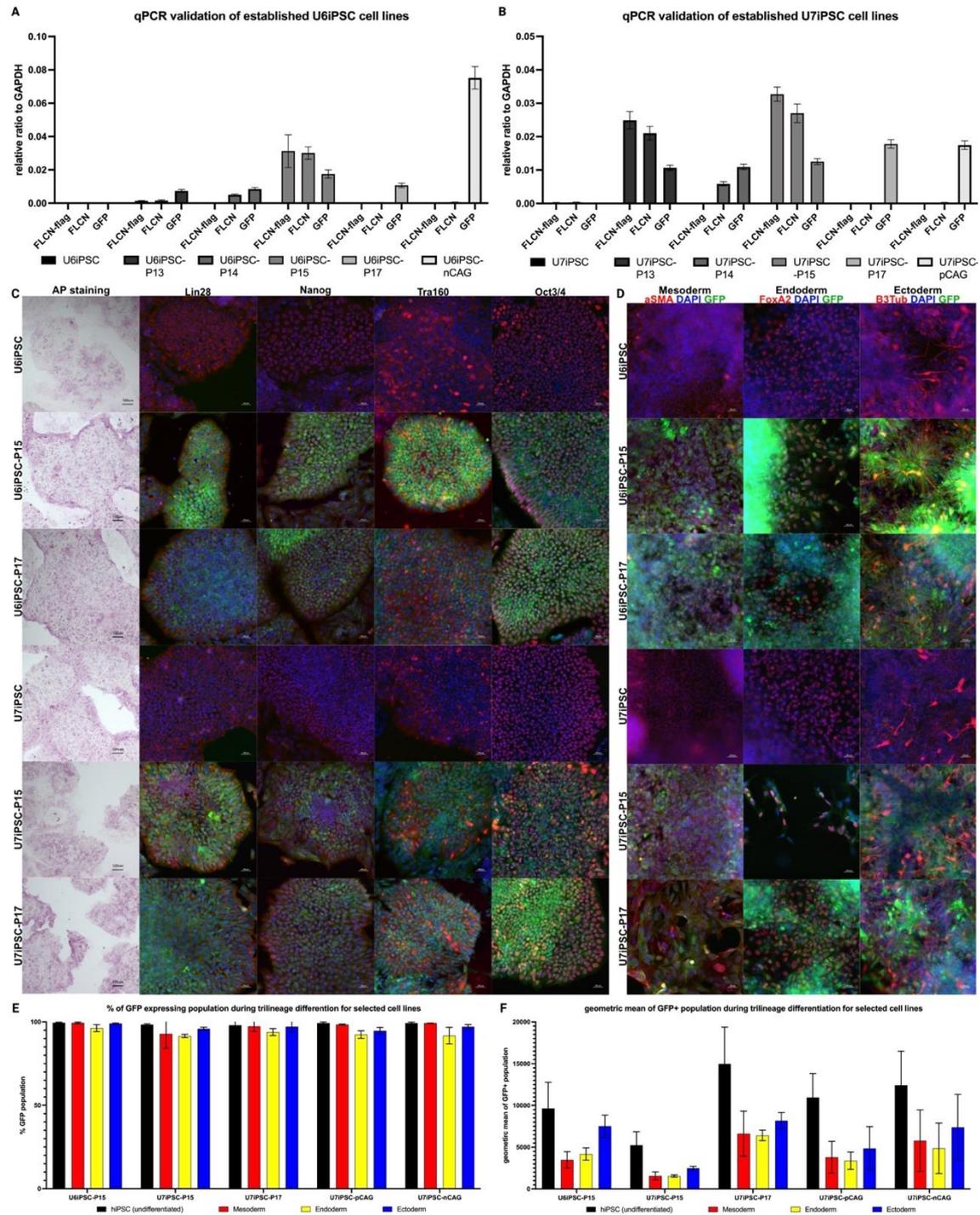
Figure 24: Characterisation of stably S/MAR modified UiPSCs.
A+B) qPCR validation of expression of FLCN-flag, FLCN and GFP relative to GAPDH in the established stable U6iPSC (A) or U7iPSC (B) cell lines and parental cell line, technical replicates (n=3) are shown. C) Validation of positive AP staining and IF pluripotency marker expression (red) with maintained endogenous GFP (green) and DAPI counterstaining (blue). Only pictures of exemplary cell lines are shown. D) IF staining of germ layer marker expression (red), endogenous GFP (green) and DAPI (blue) counterstain after trilineage differentiation. Only pictures of exemplary cell lines are shown. E+F) Trilineage differentiation was performed in biological triplicates with selected cell lines. GFP+ population percentage of alive, single cells from FACS analysis after trilineage differentiation and control undifferentiated cells (E) or geometric mean of these GFP+ populations (F) are presented.

The experiment was performed in biological triplicates. The lowest obtained GFP expressing population count was 90%, meaning a maximum loss of transgene expressing cells of 10% upon the seven days of differentiation (Figure 24E). The strongest loss in GFP expressing cells was visible in endoderm differentiation. A more severe impact of differentiation was visible when looking at the geometric mean of the GFP+ population (Figure 24F). In all cases, the GFP-S/MAR cell lines showed a higher mean fluorescent intensity (MFI) than the FLCN-S/MAR cell lines. While most cells maintained GFP expression, the MFI of these cells decreases profusely by 22-70% when normalised to the undifferentiated control data. This decrease was visible in all differentiated cell lines independent of cell line donor, utilised vector or differentiated germ layer.

RESULTS

8.7 Generation of Folliculin knock-out UiPSCs

After having optimised protocols for the generation of hiPSCs and the use of GFP- and FLCN-S/MAR DNA vectors in these cells in place, we aimed to take a further step towards generating a cell model for BHD and a S/MAR DNA vector gene therapy.

We made use of the CRISPR/Cas9 technology to generate FLCN KO UiPSCs to further investigate the importance of FLCN and the effect of FLCN restoration using our S/MAR DNA vector platform. We designed and cloned gRNAs binding the first protein-coding exon of FLCN into a px458-RFP Crispr/Cas9 vector and transfected both U6iPSCs and U7iPSCs (Figure 25A). Three days after transfection, RFP expressing iPSCs were sorted and plated as single cells in 96 well plates. Upon outgrowth, only wells with one stem cell colony were expanded and cultured further. When cells were expanded to a 12 well plate format, gDNA of the obtained Crispr vector received single-cell clones were isolated. A 352bp long sequence enclosing the gRNA binding site was amplified by PCR, the product was analysed on an agarose gel and the resulting DNA band cut, purified, and sent for sequencing. Figure 25B shows the binding site of the designed gRNA around the start of the protein translation sequence (red) and the predicted dsDNA cutting site at the codons for the fourth amino acid of FLCN (highlighted in blue).

Out of 15 sequenced single-cell clones we picked the single-cell clones 1, 2, 3 and 7 from U6iPSC and 1, 2 and 4 from U7iPSC (U6.1iPSC, U6.2iPSC, U6.3iPSC, U6.7iPSC, U7.1iPSC, U7.2iPSC, U7.4iPSC, respectively) (Figure 25C). Their sequencing results in comparison to the WT sequence showed insertion of an A at the predicted cutting site in the cell lines U6.1iPSC, U6.3iPSC, U6.7iPSC and U7.4iPSC. U7.1iPSC and U7.2iPSC showed a deletion of a C or several insertions deletions. All these mutations were predicted to lead to a frameshift and stop codon within the first amino acids of the FLCN protein. Importantly, sequencing of U6.2iPSC demonstrated the WT sequence. This cell line was thus used in the following as a WT control to exclude that the transfection, single-cell sorting and expansion affected the cells.

We next confirmed the successful knockout (KO) of the six predicted KO cell lines by Western Blot. As expected, FLCN protein was only detected in WT U6iPSC and the predicted WT U6.2iPSC lysates (Figure 25D). Due to the design of the Crispr experiment, we expected the presence of FLCN mRNA with only small alterations as shown in the sequencing data, but no FLCN protein due to an early stop codon upon protein translation. In line with this, contrary to the Western Blot, qPCR analysis of these cell lines did not show a difference in FLCN mRNA expression (data not shown).
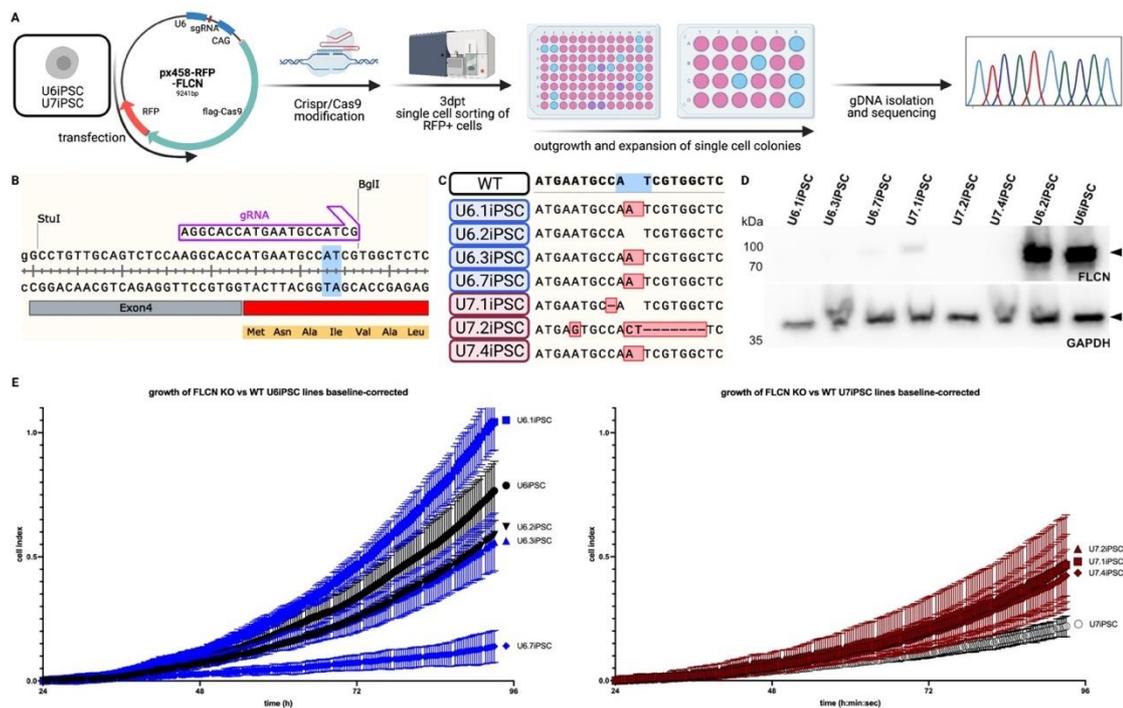

Figure 25: Generation of Crispr/Cas9 FLCN KO UiPSCs.
A) Schematic: U6iPSCs and U7iPSCs were transfected with px458-RFP-FLCN vector, encoding the gRNA against FLCN, Cas9 and RFP as a reporter gene. 3dpt RFP expressing cells were sorted and single cell clones were expanded, their gDNA isolated, the area round the gRNA binding site was PCR amplified and sequenced. B) Binding of the designed gRNA occurred at exon 4 overlapping to the translation start (red) of FLCN. Predicted dsDNA break occurs between the highlighted bases (AT) and the codon for the fourth amino acid of FLCN. C) Sequencing results of WT, four U6iPSC derived single cell clones (clone 1, 2, 3 and 7) and three U7iPSC derived single cell clones (1, 2 and 4). D) Western Blot of lysates from the seven mentioned single cell lines and parental U6iPSCs probed for FLCN and GAPDH as loading control. E). Growth comparison of parental U6iPSCs (left) and U7iPSCs (right) and their respective obtained single cell lines. The xCelligence baseline corrected data of technical triplicates between 24h-96h after plating are shown.

Since FLCN is known to influence cell cycle progression we considered an effect on the growth rate in FLCN KO cells. We compared the obtained KO cell lines (blue and red) and the parental or the control WT cell lines (black) with the xCelligence device (Figure 25E). Baseline corrected data from exponential growth 24h-96h after plating showed a slightly higher growth rate of KO U7iPSCs compared to the parental WT U7iPSC line (right). However, this was not the case for both WT parental U6iPSC and clonal

RESULTS

U6.2iPSC lines when compared to the U6iPSC KO lines. The same y-axis scaling of the two graphs allows the comparison of the two parental WT lines U6iPSC and U7iPSC. Importantly, a bigger growth difference between these WT cell lines is visible than the one observed within their generated KO lines. This suggests that the observed growth rates are mainly an outcome of different clonal behaviour rather than an effect of the FLCN KO.

We further characterised the obtained FLCN KO UiPSC lines and first confirmed sustained expression of hiPSC pluripotency markers (Figure 26A). Next, we were interested in the capability of these cells to differentiate. It was reported that FLCN KO mouse and human ESCs were not able to differentiate, suggesting that FLCN restoration was required at the hiPSC level (Betschinger et al., 2013). A preliminary experiment was performed, where FLCN KO and parental hiPSCs were withdrawn of the stem cell media and their growth factors by changing to DMEM + 10% FCS media for 24, 48 or 72h. The cells were therefore allowed to spontaneously differentiate, and this differentiation capability was measured by AP staining of control cells kept in StemFit Basic02 media and the DMEM cultured cells (Figure 26B). Surprisingly, all tested cell lines behaved similarly and no distinct difference in loss of intensity of AP staining upon growth factor withdrawal was visible between the tested WT and FLCN KO hiPSCs, indicating a continued capability of the FLCN KO iPSCs to differentiate. To further verify this, trilineage differentiation was performed with all FLCN KO iPSC lines. After the seven days of differentiation, successful staining of markers for each germ layer confirmed the ability of these cells to differentiate (Figure 26C).
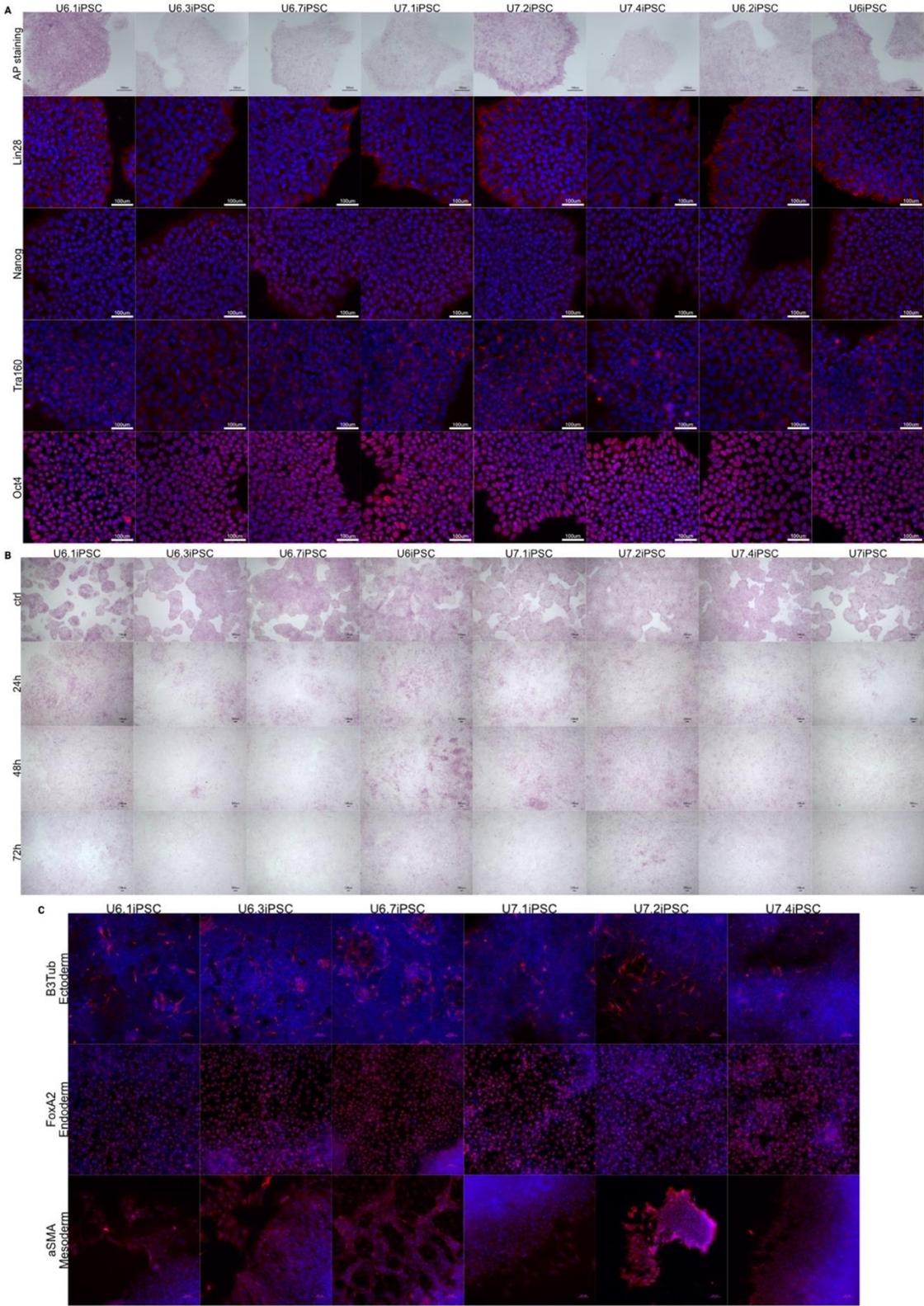
Figure 26: Pluripotency capability of FLCN KO UiPSCs.
A) AP staining and pluripotency marker IF pictured of the obtained FLCN KO hiPSC lines as well as the WT U6.2iPSC line and parental U6iPSCs as control. B) FLCN KO and parental hiPSC lines were subjected to growth factor withdrawal induced differentiation by changing the media to DMEM + 10% FCS for 24, 48 or 72h. AP staining of the different time points as well as control staining is shown. C) IF staining of germ layer markers (red) and DAPI counterstain (blue) after subjecting the FLCN KO hiPSC lines to trilineage differentiation.

RESULTS

8.8 Restoration of Folliculin in Folliculin knock-out UiPSCs

As a next step, we aimed to modify the generated FLCN KO UiPSCs using our S/MAR DNA vector systems. Since the three U6iPSC derived FLCN KO lines all showed the same sequence, only the U6.7iPSC cell line was further used. To investigate potential donor specificity, but also differences in clonal behaviour, as well as possible different off-target effects, all three FLCN KO U7iPSC lines with different FLCN sequence KOs were further used.

First, U7.1iPSC, U7.2iPSC and U7.4iPSCs were established with the GFP S/MAR DNA vector P17 as a control to verify that FLCN KO did not affect the generation of S/MAR modified stable cell lines. Cells were sorted 15, 32 and 52dpt (Figure 27A). While without an earlier sorting point the GFP expressing cell population was only 0.1-0.2% of alive single cells 15dpt, sufficient GFP expressing cells for further outgrowth were obtained. With this late sorting timepoint, the obtained cell population showed GFP expression in at least 93% of alive single cells until the next sorting at 32dpt, suggesting almost full establishment at 15dpt. Sorting at 52dpt confirmed purity of 98-100% of S/MAR expressing cells. All three established cell lines further kept 99.3-100% GFP expression until the next FACS analysis at 108dpt.

Secondly, we aimed to restore FLCN expression in the four FLCN KO iPSC lines with our tested FLCN vector P15 (Figure 27A). With the reported lower transfection efficiency with U7iPSCs, and already some experience with establishment timing, we decided to increase the cell count by performing two transfections. Cells were sorted at the same time, which equalled 12dpt for the first transfection batch and 7dpt for the second. Cells from the same parental cell line were then pooled just before the second sorting at 27 and 32dpt, respectively. As expected, expressing population percentage on 7dpt was higher than at 12dpt, and the mixture of these sorted populations showed retention of GFP expression in at least 80% of the single alive cell population. Long-term analysis 83/88dpt revealed that GFP expression was maintained and only U7.4iPSC-P15 lost GFP expression in approximately 20% of its population.

Successful restoration of FLCN expression in stably established P15 modified FLCN KO

lines and absence of FLCN in unmodified and P17 modified KO lines was confirmed via

Western Blot (Figure 27C, D). For better FLCN expression comparison lysates from

parental WT U7iPSC andU6iPSC lines and their previously generated P15 and P17

modified cell lines U7iPSC-P15, U6iPSC-P15 and U7iPSC-P17 were also loaded. In line

with expectations, strong expression of FLCN was detected in all P15 modified cell

lines, while no band was shown in FLCN KO and FLCN KO-P17 cell lines. The FLCN WT

cell lines U7iPSC, U6iPSC and U7iPSC-P17 demonstrated only a faint band for FLCN
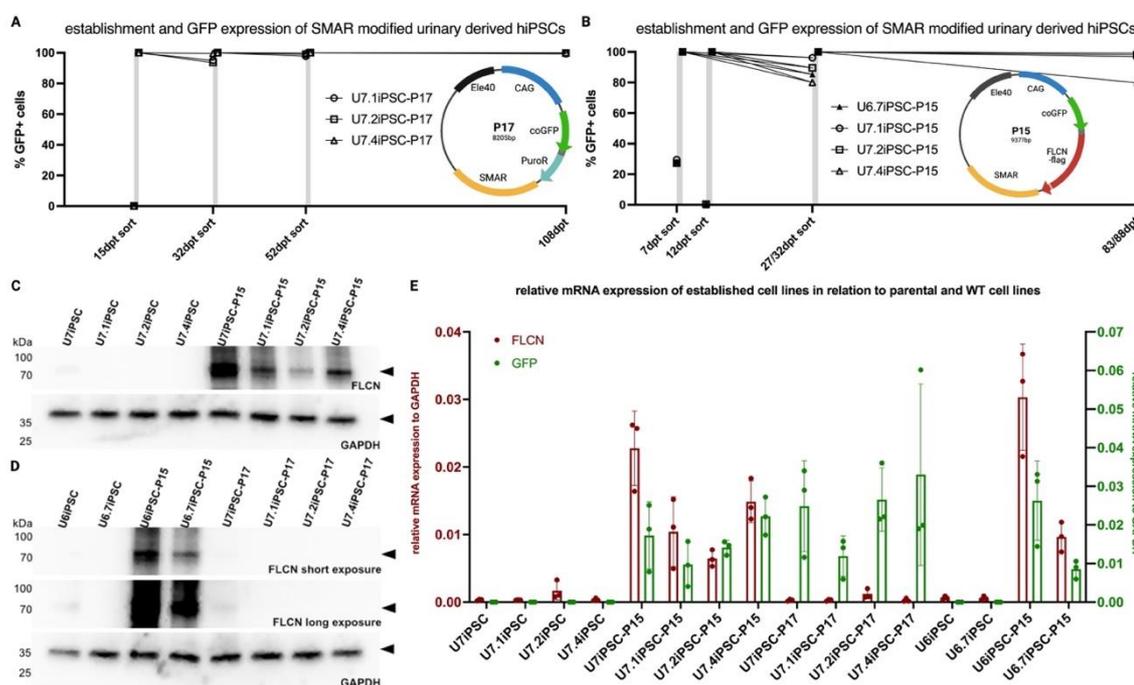
despite prolonged exposure time.

Figure 27: Stable S/MAR modification of FLCN KO UiPSCs and verification of transgene expression.
A+B) FACS data for GFP expressing populations of alive, single cells during establishment and at later timepoints of
S/MAR DNA vector transfected FLCN KO UiPSCs are shown. A) Data for U7iPSC derived FLCN KO lines transfected
with the GFP-S/MAR DNA vector P17. B) Data for the U6.7iPSC and the three U7iPSC derived FLCN KO lines
transfected with the GFP- and FLCN-encoding S/MAR DNA vector. C+D) Western Blot of cell lysates with and
without S/MAR DNA vector establishment probed for FLCN protein expression and GAPDH as loading control. E)
qPCR data showing technical triplicates of FLCN and GFP mRNA presence in the here generated S/MAR established
FLCN KO iPSC lines in comparison to parental U6/U7iPSC lines and the FLCN KO lines without S/MAR modification.

The presence of mRNA of both transgenes, GFP and FLCN was additionally confirmed

by qPCR in the S/MAR modified FLCN KO cell lines in three biological replicates (Figure

27E). Again, expression levels of parental KO and WT cells, as well as modified WT

cells, were tested for comparison. In line with the Western Blot data, GFP mRNA was

detected in all S/MAR modified cell lines. Distinct FLCN expression or overexpression,

respectively, was visible in all P15 modified cells but not in P17 modified or parental
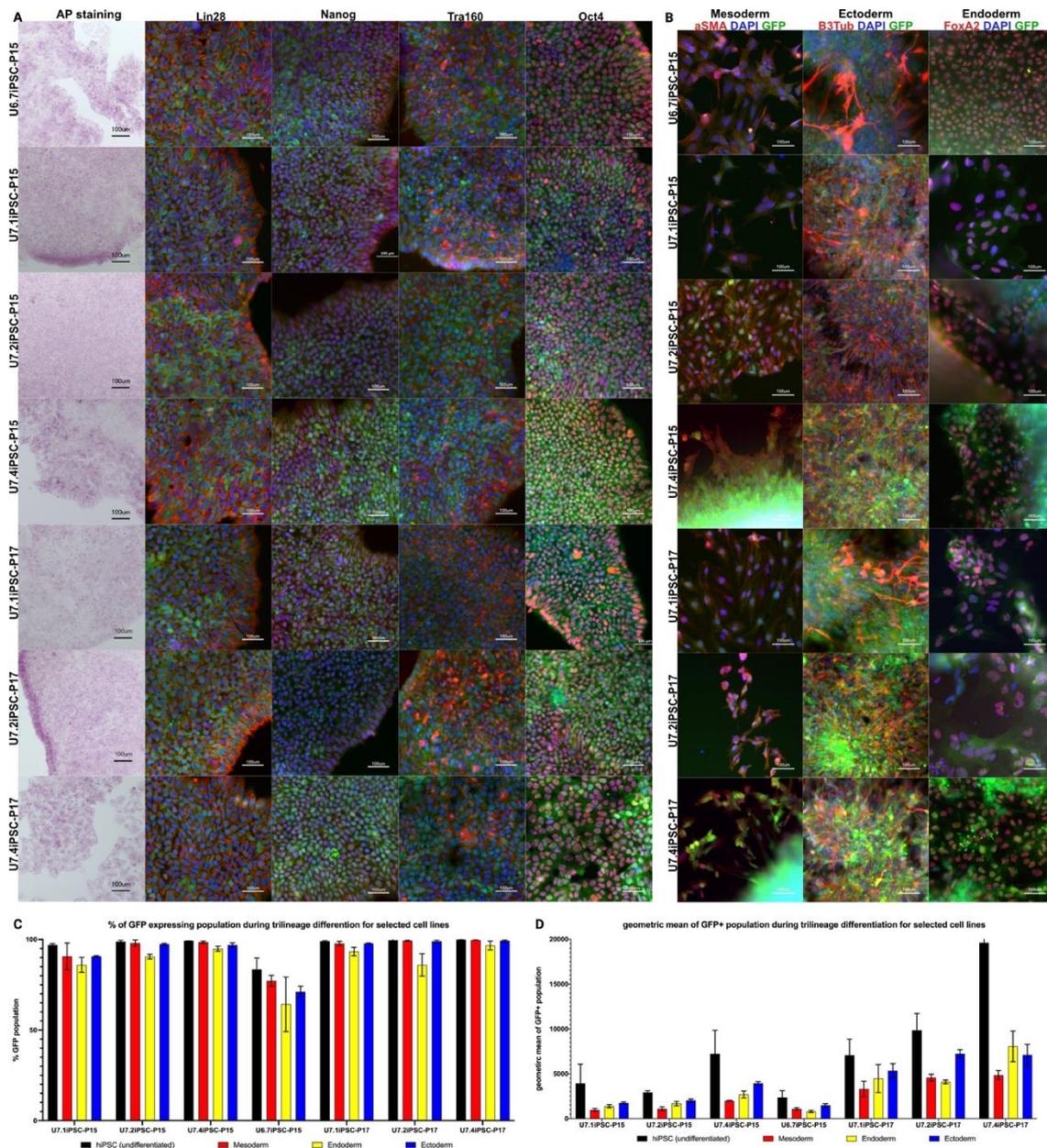
WT or KO cells.

Figure 28: Validation of pluripotency markers and transgene expression during germ layer differentiation in S/MAR modified FLCN KO UiPSCs.

A-D) Data for the stably P15 or P17 modified FLCN KO UiPSCs. A) Validation of positive AP staining and IF pluripotency marker expression (red) with maintained endogenous GFP (green) and DAPI counterstaining (blue). B) IF staining of germ layer marker expression (red), endogenous GFP (green) and DAPI counterstain (blue) after trilineage differentiation. C+D) Trilineage differentiation was performed in biological triplicates. GFP+ population percentage of alive, single cells from FACS analysis after trilineage differentiation and control undifferentiated cells (C) or geometric mean of these GFP+ populations (F) are presented.

As in all previously established cell lines, no alteration in pluripotency marker

expression was found in the S/MAR modified FLCN KO iPSCs (Figure 28A). Also,

successful trilineage differentiation with maintained GFP expression was confirmed by

germ layer marker IF staining (Figure 28B). As before, we furthermore generated quantitative data about the conserved GFP expression upon seven days of trilineage differentiation. As observed for P15 and P17 modified WT U6iPSCs and U7iPSCs, only a slight loss of GFP expressing cells was found upon differentiation in most cell lines (Figure 28C). Again, with up to 10-15% of the population, most of the loss of GFP expressing cells was observed in endoderm differentiated cells. The U6.7iPSC-P15, which showed the previously reported loss in GFP expression upon long-term maintenance also presented the strongest loss of the GFP expressing population upon differentiation. As observed with the S/MAR modified WT-UiPSCs, again a major drop of 25-75% in GFP intensity was observed in the GFP+ cell population upon differentiation compared to undifferentiated controls (Figure 28D).

RESULTS

8.9 Comparison of the expression profiles of different UiPSC lines

Our characterisations of the obtained cell lines so far did not show us any alteration in cell behaviour or properties. To further investigate that we next looked into the expression profiles of the different established cell lines. In collaboration with the lab of Prof. Michael Platten (Immunology and Cancer, DKFZ) and support of Dr Edward Green and Chin Leng Tan (Immunogenomics, DKFZ), single-cell RNA sequencing (RNAseq) was performed. We analysed WT U7iPSCs and FLCN KO U7.1iPSCs, both with and without stable modification using our GFP-S/MAR (P17) or FLCN-GFP-S/MAR (P15) vectors.
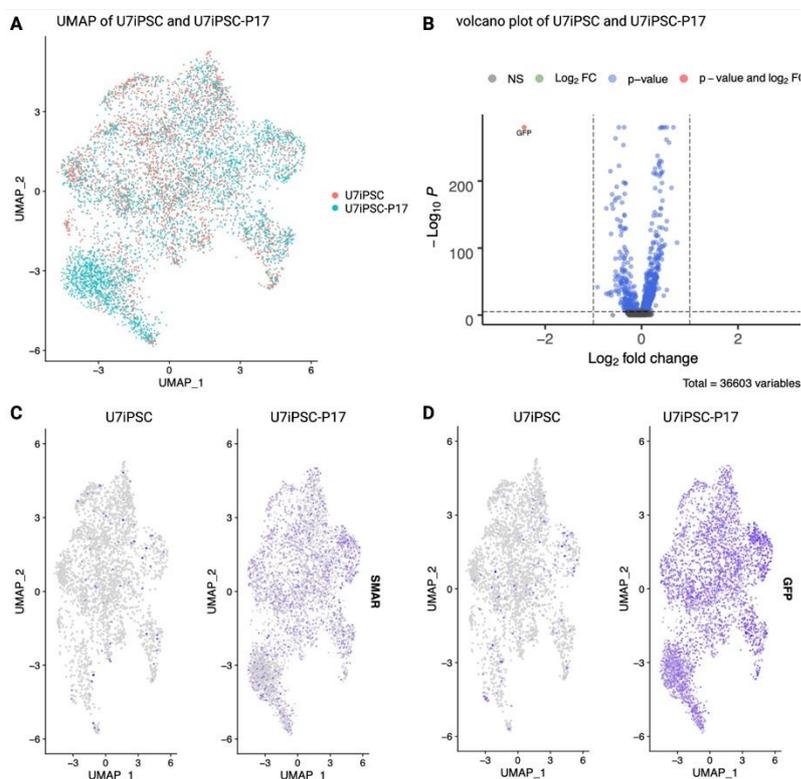
Figure 29: Effect of stable modification with the GFP-S/MAR DNA vector P17 on WT U7iPSCs.
A) UMAP of RNAseq data of U7iPSC and U7iPSC-P17. B) Volcano plot of U7iPSC and U7iPSC-P17 RNAseq data with GFP expression pointed out. C, D) UMAP comparison of U7iPSC and U7iPSC-P17 data in regard to C) S/MAR and D) GFP expression.

In line with previous reports of our group in different cells (Bozza et al., 2021; Bozza et al., 2020; Roig Merino et al., in press), the GFP-S/MAR DNA vector P17 did not demonstrate strong changes in gene expression (Figure 29). The UMAPs of the U7iPSC and stable U7iPSC-P17 cell lines show high levels of similarity with one distinctive cluster at the bottom left for U7iPSC-P17 cells (Figure 29 A). However, volcano plots of these samples demonstrate no major changes in expression beyond a log$_2$ fold change

besides the strong GFP expression (Figure 29 B). Single-cell data further confirmed uniform S/MAR and GFP expression in the U7iPSC-P17 cells without distinct cluster formation (Figure 29 C, D).

We further investigated if the different levels of FLCN expression in our six cell lines had an impact on the pluripotency state of the cell population. First, primed, naïve, and general pluripotency markers as well as a proliferation marker were selected, and clusters of the single-cell RNAseq data were determined Figure 30 A, B, C and D). Figure 30 E shows the pooled data with the determined clusters for primed, naïve, GFP expressing and proliferative cells. Percentages of cells in these clusters are shown separately for each sample in Figure 30 D. As expected, GFP clusters were only observed in S/MAR modified cells. U7iPSC-P17 vector demonstrated the highest percentage of cells clustering into the GFP cluster. In general, the three U7.1iPSC lines demonstrated a slightly smaller percentage of the naïve cell cluster than the three U7iPSC lines. Also, the proliferative cluster was severely reduced in WT U7iPSCs with stable P15 overexpression but increased in all other S/MAR modified cell lines compared to parental cells. The primed cell cluster was distinctively smaller in U7iPSC-P17 cells; however, this could be misleading due to the strong clustering of this sample into the GFP cluster.

Next, a graph-based unsupervised clustering was performed with the data and ten easily distinguishable clusters were chosen based on their distinct differential gene expression. Comparison of the clusters suggested only slight differences in cluster composition of the samples (Figure 30 G). While further investigation needs to be performed to determine the essential pathways of each cluster, again, the GFP cluster was again highly represented in U7iPSC-P17 cells. There were only small changes in cluster composition between WT and KO U7iPSCs. In general, S/MAR DNA vector modification and GFP or FLCN transgene overexpression seemed to have a stronger effect on WT iPSCs than on KO iPSCs, as seen e.g., for proliferation.

RESULTS
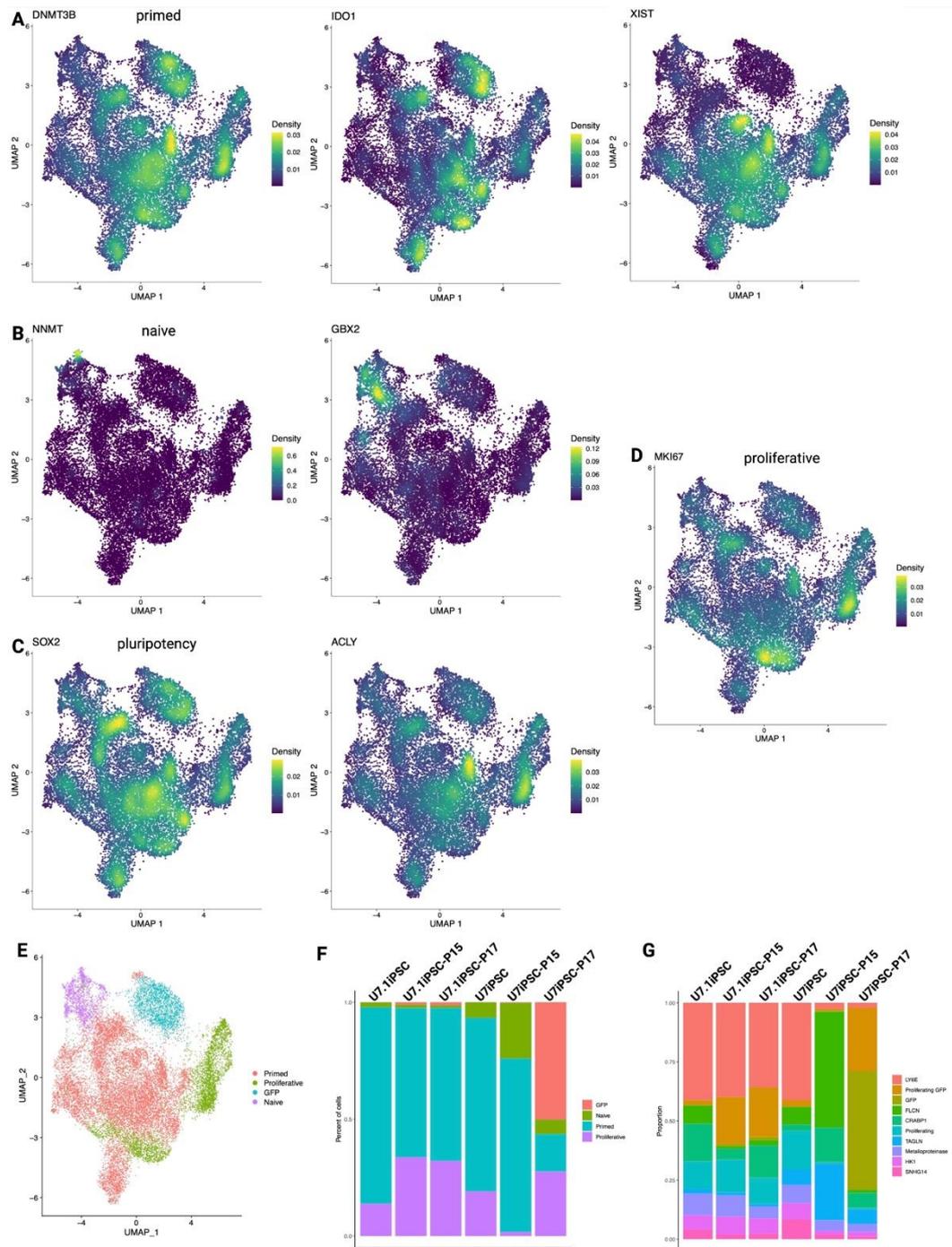
Figure 30: Influence of FLCN expression levels in the different U7iPSC WT, KO and modified cell lines in regard to pluripotency.
A-D): UMAP of RNAseq data of pooled samples WT U7iPSC, U7iPSC-P17, U7iPSC-P15, KO U7.1iPSC, U7.1iPSC-P17 and U7.1iPSC-P15. Expression changes of A) primed stem cell markers DNMT3B, IDO1 and XIST, B) naïve stem cell markers NNMT and GBX2 C) pluripotency markers SOX2 and ACLY, D) proliferation marker MK167 are shown. E) Clustering of pooled expression data into primed, proliferative, GFP and naïve populations based on expression patterns. F) Visualisation of the cluster composition of the different samples. G) unsupervised clustering of the different samples gives an indication of differentially regulated genes.

8.10 Kidney organoids as an advanced model for
Birt-Hogg-Dubé syndrome

So far, we obtained EBNA and feeder free reprogramming derived UiPSCs, generated FLCN KO UiPSCs and stably modified these cell lines with GFP or FLCN-flag-GFP expressing S/MAR DNA vectors. All performed characterisation assays as well as single-cell RNA sequencing indicated similar behaviour and no adverse effect due to the S/MAR DNA vector modification.

While these cell lines can serve as a developmental early model for the study of FLCN signalling pathways, the major health problem for patients diagnosed with BHD is the development of kidney tumours. We now aimed to utilise our optimally obtained cell lines as an advanced cell model by differentiating them into kidney organoids.

By doing so, we were especially interested in evaluating, if FLCN KO triggered tumorigenesis can be studied in this model. Furthermore, these experiments were aimed to verify, that FLCN-encoding S/MAR DNA vectors could be used as preventive gene therapy in the setting of BHD as well as the maintenance of S/MAR DNA vector expression throughout this more elaborate differentiation process.

8.10.1 Implementation of differentiation protocol

In collaboration with Dr Martin Hoogduijn (Erasmus MC, Rotterdam) a kidney organoid differentiation protocol based on the publication of (Garreta et al., 2019) was implemented. Figure 31 A schematically summarises the differentiation protocol with the utilised media and culture conditions as extensively described in chapter 7.2.14. Briefly, CHIR-99021 (CHIR) was used on a monolayer culture to induce posterior primitive streak (PPS), media change to Activin A (ActA), FGF9 and heparin forced intermediate mesoderm (IM) differentiation. Under the shown media conditions, cells were then forced to aggregate in a 96-well v-bottom plate for 2 days and the 3D cell aggregates were transferred on the liquid-air-interface on a transwell plate which caused nephron progenitor cell (NPC) differentiation. In brightfield, visible structures indicate successful renal vesicle (RV) development until finally, nephrogenesis takes place.
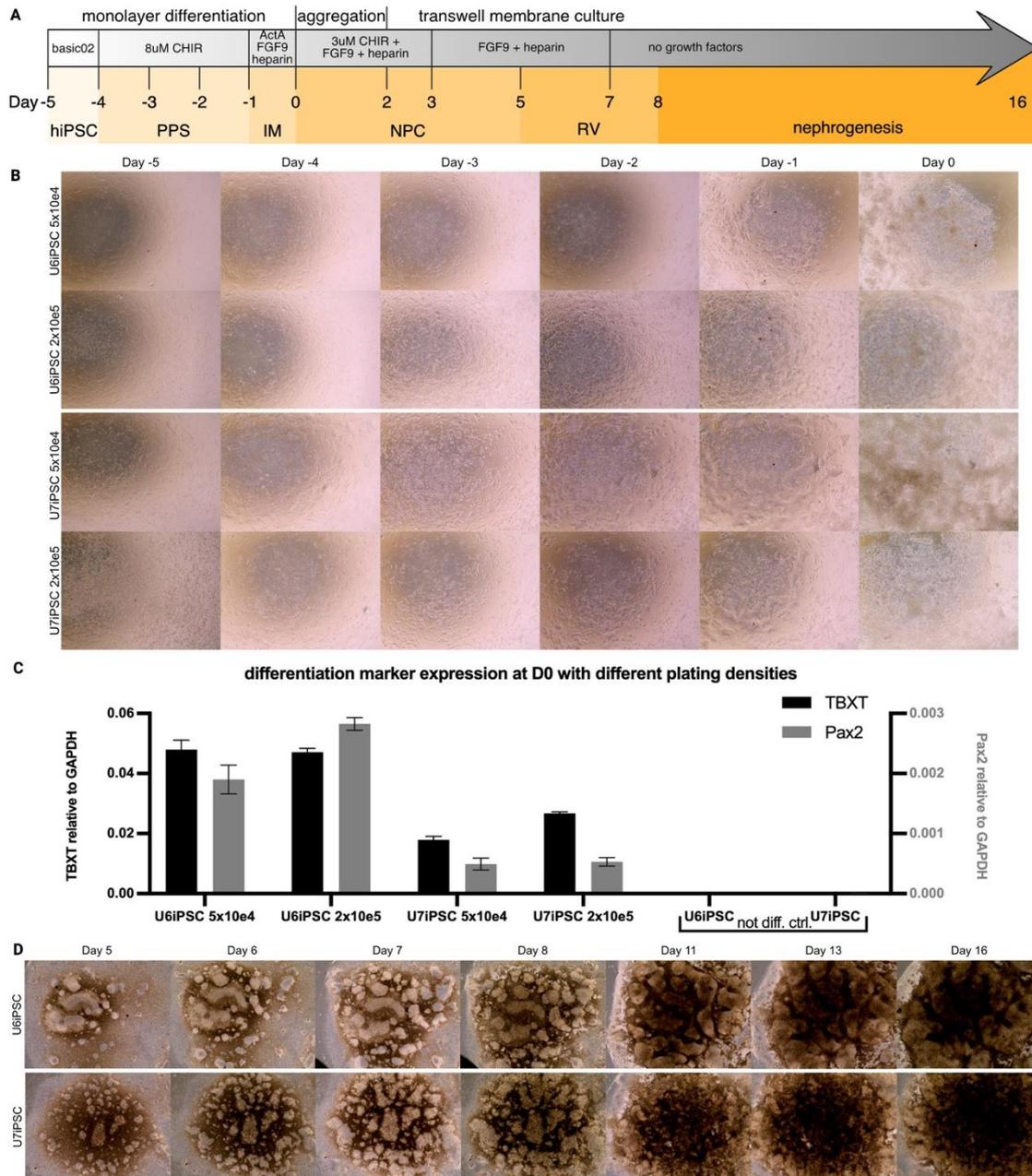
RESULTS

Figure 31: First assessment of kidney organoid differentiation protocol with urinary derived iPSCs.
A) Schematic of kidney organoid differentiation showing timeline with utilised culturing format and media supplements as well as state of the cells. PPS = posterior primitive streak, IM = intermediate mesoderm, NPC = nephron progenitor cells, RV = renal vesicles. B) Brightfield pictures after plating of hiPSCs and during 5 days of monolayer differentiation. C) qPCR of technical replicates of U6iPSCs and U7iPSCs differentiated with different densities (5x10e4 or 2x10e5 cells/well) and undifferentiated parental hiPSCs as control showing TBXT (PPS marker) as well as Pax2 (IM marker) expression of differentiated cells on day 0 of differentiation. D) Representative brightfield pictures of differentiated aggregates cultured on transwell membrane on day 5-16 with visible development of structures.

As a first step, the plating density in WT U7iPSC and U6iPSC cell lines were optimised, as massive cell death was observed during monolayer differentiation. As exemplarily shown in Figure 31 B, plating density also influenced the number of cells on day 0. However, qPCR on these tested cells on day 0 of the differentiation protocol showed

successful upregulation of TBXT (PPS) and Pax2 (IM) markers in all tested densities (representative densities are shown). With similar expression levels of both U7iPSC and both U6iPSC samples, our data furthermore suggests that intensity of expression was rather dependent on the used cell line than the plating density.

Continuation of the differentiation protocol showed tubular structure formation from day 5 on, indicating successful RV formation and kidney organoid differentiation (exemplary pictures, Figure 31 D). Throughout the culturing on the transwell membrane, intensive cell growth with an accumulation of dead cells where structures seemingly start to fall apart towards day 16 was visible.

8.10.2 Persistent expression of S/MAR DNA vectors throughout kidney organoid differentiation

After the first implementation of the differentiation protocol with WT UiPSCs, we applied these differentiation conditions to the stable P15 or P17 modified U6iPSC and U7iPSC cell lines. Figure 32 shows representative brightfield pictures of structure formation between days 2 to 16 of the differentiation of all six cell lines from three biological replicates. No indication of aberrant differentiation behaviour of S/MAR modified cell lines or FLCN overexpression by S/MAR modification was visible.

As this kidney organoid differentiation represents the most elaborate differentiation protocol performed so far with S/MAR DNA vector modified hiPSCs, we next validated maintenance of S/MAR DNA vector expression throughout differentiation. Figure 33 A shows representative brightfield pictures as well as brightfield and GFP pictures taken with a fluorescence microscope. These pictures further emphasise structures and tubuli formation at days 8 and 11 of differentiation, as well as maintained GFP expression throughout the differentiation.
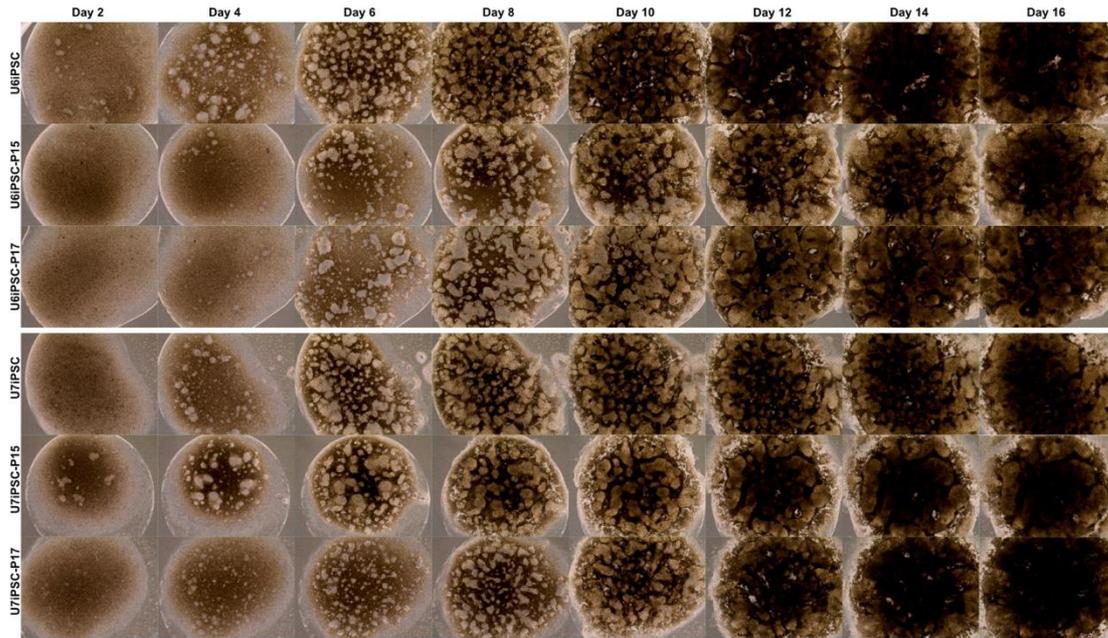
RESULTS

Figure 32: Stable modification with S/MAR GFP or FLCN vectors do not alter kidney organoid differentiation phenotype.
Representative brightfield pictures of structure formation of GFP- or FLCN-S/MAR DNA vector modified UiPSCs on day 2-16.

For quantitative data, we performed FACS analysis of the S/MAR modified cells before, during (day 0) and at the end of the differentiation protocol (day 16). Percentage of GFP expressing alive single cells are depicted in Figure 33 B. To verify, that loss of GFP expressing cells was due to differentiation, and not routine maintenance as observed in some cell lines, maintenance FACS data of utilised cells is included. In two biological replicates, moderate loss of GFP expressing cells during monolayer differentiation is seen, while ~20% loss of GFP expressing cells is seen after 3D differentiation.

A protocol for fixation of day 11-16 kidney organoids for structure analysis with H&E as well as IHC staining was started to be implemented. Optimisations such as the utilisation of a tissue staining dye and fixation at earlier timepoints helped that organoids were embedded in the Histogel block without falling apart and able to be cut. We were only able to obtain a few representative H&E stainings (Figure 33 C), supporting our observations of structure formation in brightfield microscopy.
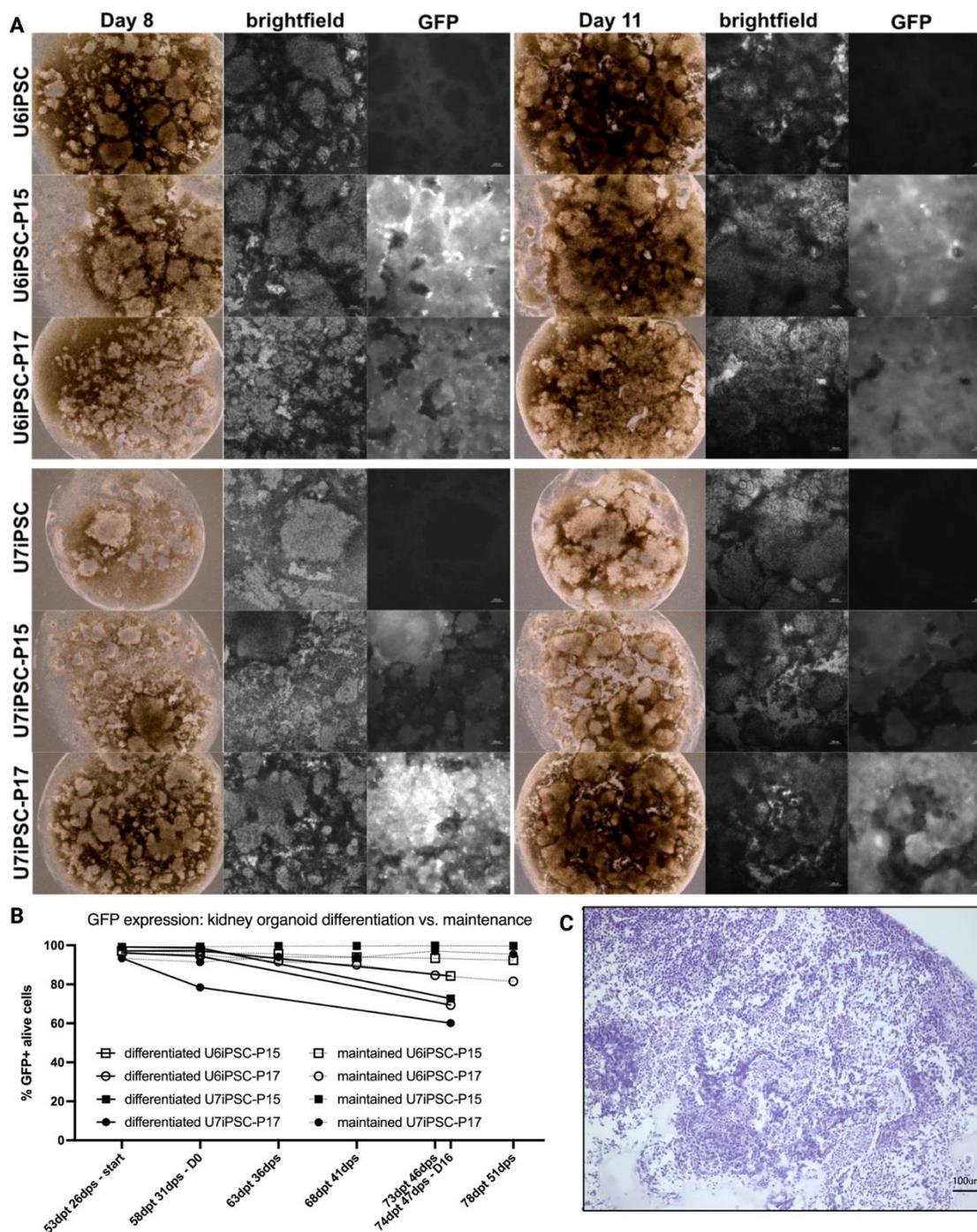
Figure 33: GFP expression upon kidney organoid differentiation and H&E staining.
A) Brightfield as well as brightfield and GFP pictures with a fluorescence microscope of day 8 and day 11 representative kidney organoid structures of different UiPSC lines with and without S/MAR modification. B) FACS data showing % of GFP expressing alive single cells. Data with bold black line shows GFP population of S/MAR modified UiPSCs during kidney organoid differentiation on day of plating, Day 0 and day 16. Data connected with dotted line represents hiPSC undifferentiated maintenance data of used cell lines during the time of experiment. C) H&E staining of a representative kidney organoid is shown.

We next focused on qPCR expression data for S/MAR expression and differentiation

marker confirmation in parental, P15 (FLCN-GFP) or P17 (GFP)-S/MAR modified

U6iPSCs and U7iPSCs.

RESULTS

High GFP, FLCN-flag and FLCN expression levels were observed as expected in S/MAR modified cell lines (Figure 34 A-C). Interestingly, for these genes' expression of day 16 differentiated samples was higher throughout all the tested samples suggesting an increase of transcription of S/MAR DNA vectors upon kidney organoid differentiation. As expected, the PPS marker TBXT was upregulated on day 0, but expression dropped again towards day 16 (Figure 34 D). The IM marker Pax2 was again upregulated at day 0 and showed even stronger expression on day 16 (Figure 34 E). The expression levels for Oct4A as a marker for pluripotency however behaved inconsistent on day 0 but showed increased expression throughout all samples on day 16.
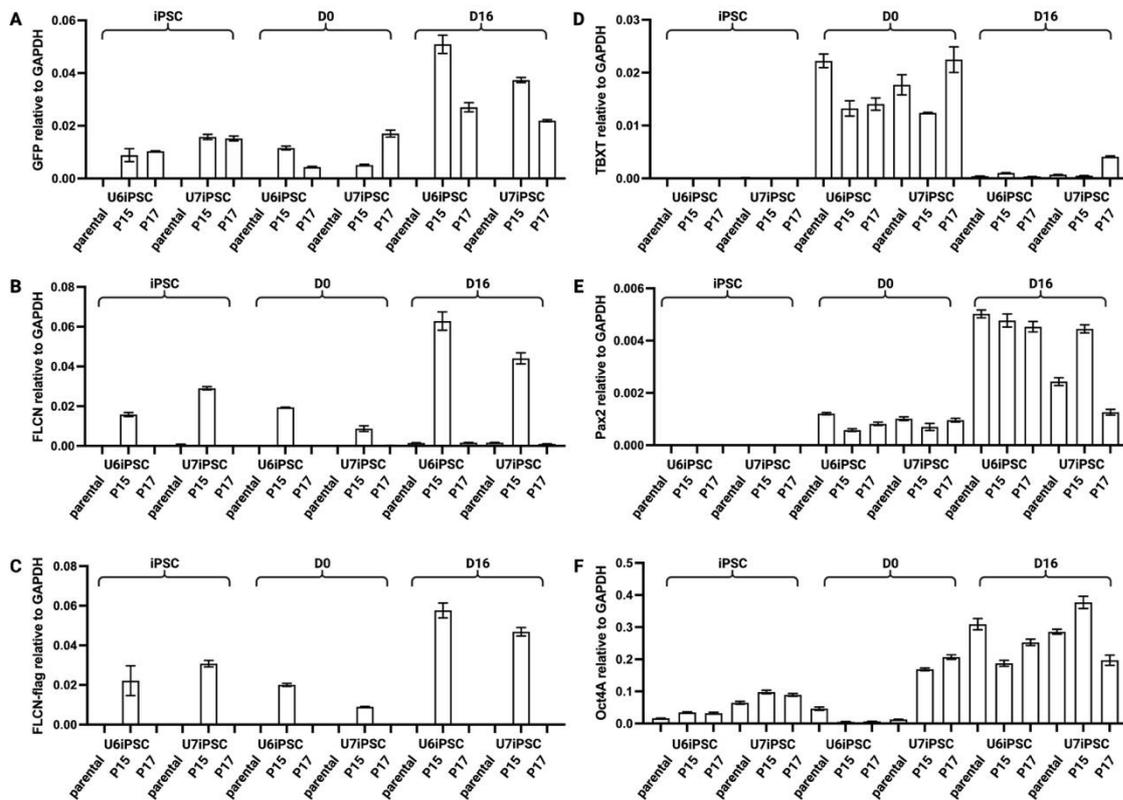
Figure 34: qPCR expression data of S/MAR expression as well as pluripotency and differentiation markers.
A-F) Representative qPCR Data showing technical replicates of parental, P15 (FLCN-GFP) or P17 (GFP)-S/MAR modified UiPSCs. Undifferentiated hiPSC, as well as kidney organoid differentiated samples from day 0 and day 16 of differentiation were checked. A, B, C) GFP, FLCN and FLCN-flag for validation of S/MAR DNA vector expression. D) TBXT as PPS marker. E) Pax2 as IM marker. F) Oct4A as pluripotency marker.

8.10.3 Folliculin knock-out iPSCs are capable to differentiate towards kidney organoids

Against our expectations, our FLCN KO UiPSCs previously were able to differentiate in our trilineage differentiation assay. We next aimed to verify, that these cells were also capable to undergo the more elaborate differentiation towards kidney organoids.
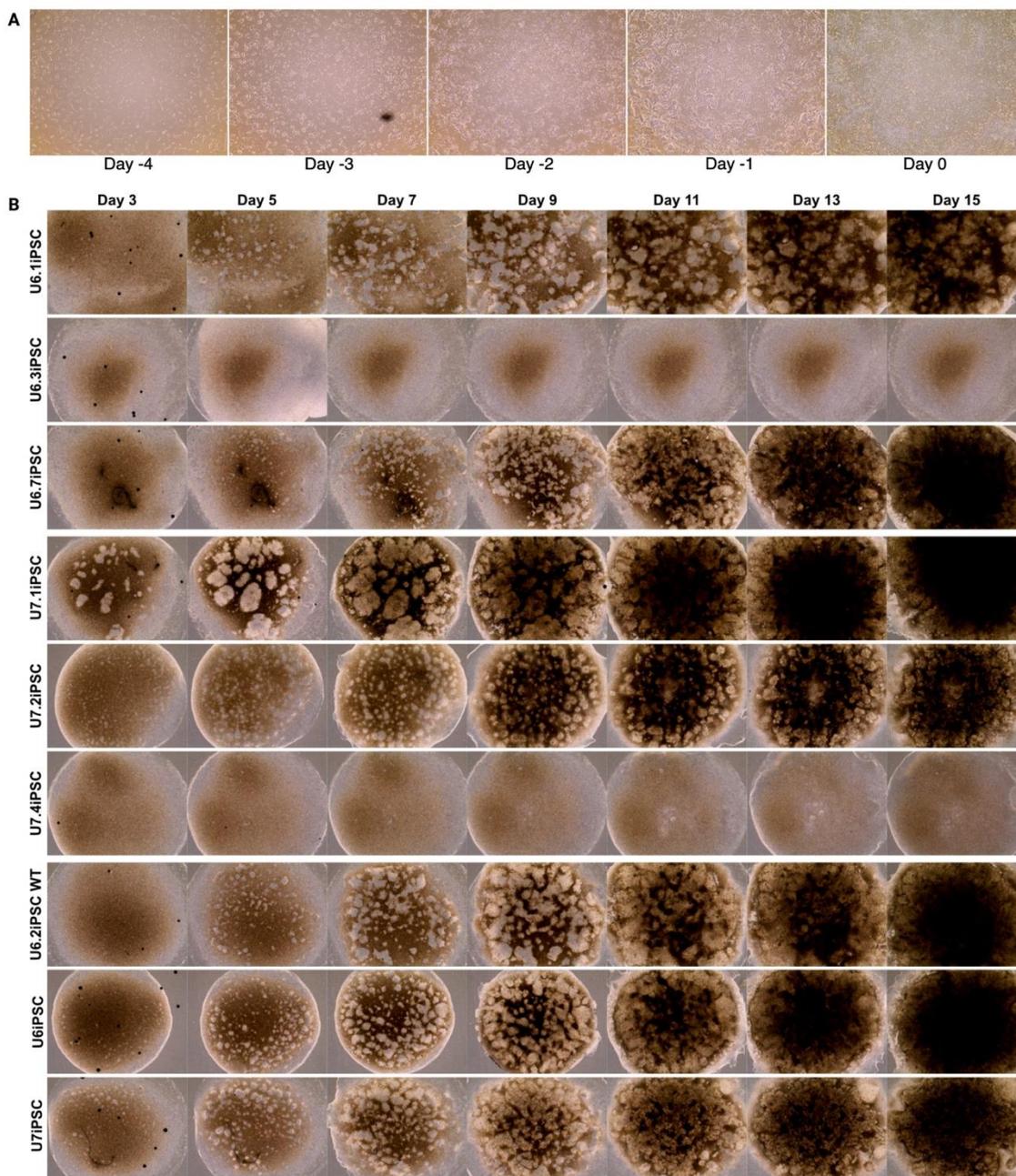
Figure 35: Morphology of kidney organoid differentiation using FLCN KO iPSCs.
A) Representative brightfield pictures of monolayer differentiation of FLCN KO and WT iPSCs. B) Exemplary pictures of representative 3D organoid differentiation and structure formation of FLCN KO and WT iPSCs. Notably, biological replicates showed inconsistency in structure formation or failure of structure formation, where all cell lines were able to form structures at least once.

RESULTS

In line with previous characterisations, our FLCN KO UiPSCs showed no obvious difference in monolayer morphology in comparison to WT iPSCs as exemplarily shown in Figure 35 A. During subsequent 3D differentiation some cell lines failed to produce structures, while others showed only delayed or normal structure formation (Figure 35 B). However, in line with reports from our collaborators, we observed inconsistency in well-structured organoid formation upon biological replicates. While also WT UiPSC lines presented problems with consistency, throughout biological replicates all 6 generated FLCN KO iPSC lines were able to form structures during kidney organoid differentiation at least once. In line with this, we confirmed Pax2 and TBXT upregulation in qPCR assays on day 0 of FLCN KO iPSC differentiations (data not shown).

8.10.4 The presence or absence of Folliculin does not affect kidney organoid differentiation

As last step, we performed replicates of kidney organoid differentiation with a full range of our generated cell lines. We wanted to investigate the effect of absence, restoration, WT or overexpression of FLCN in comparison to our GFP-S/MAR DNA vectors and parental cell lines in direct comparison. Therefore, also FLCN restored stably P15-modified FLCN KO iPSCs were used for these experiments.

While previous experiments were also performed with U6iPSC derived cell lines, we did not observe any difference between these donor lines. Thus, we focused on cell lines established from U7iPSCs for the following experiments.

Figure 36 A shows exemplary brightfield pictures of WT U7iPSCs, as well as the three generated FLCN KO U7iPSCs, and the generated stable P15 (FLCN-Flag-GFP) or P17 (GFP) S/MAR modified lines from one biological replicate. While further replicates demonstrated again inconsistency in structure formation during transwell culturing as listed in Figure 36 B, all cell lines were equally able to form structured kidney organoids at least once.
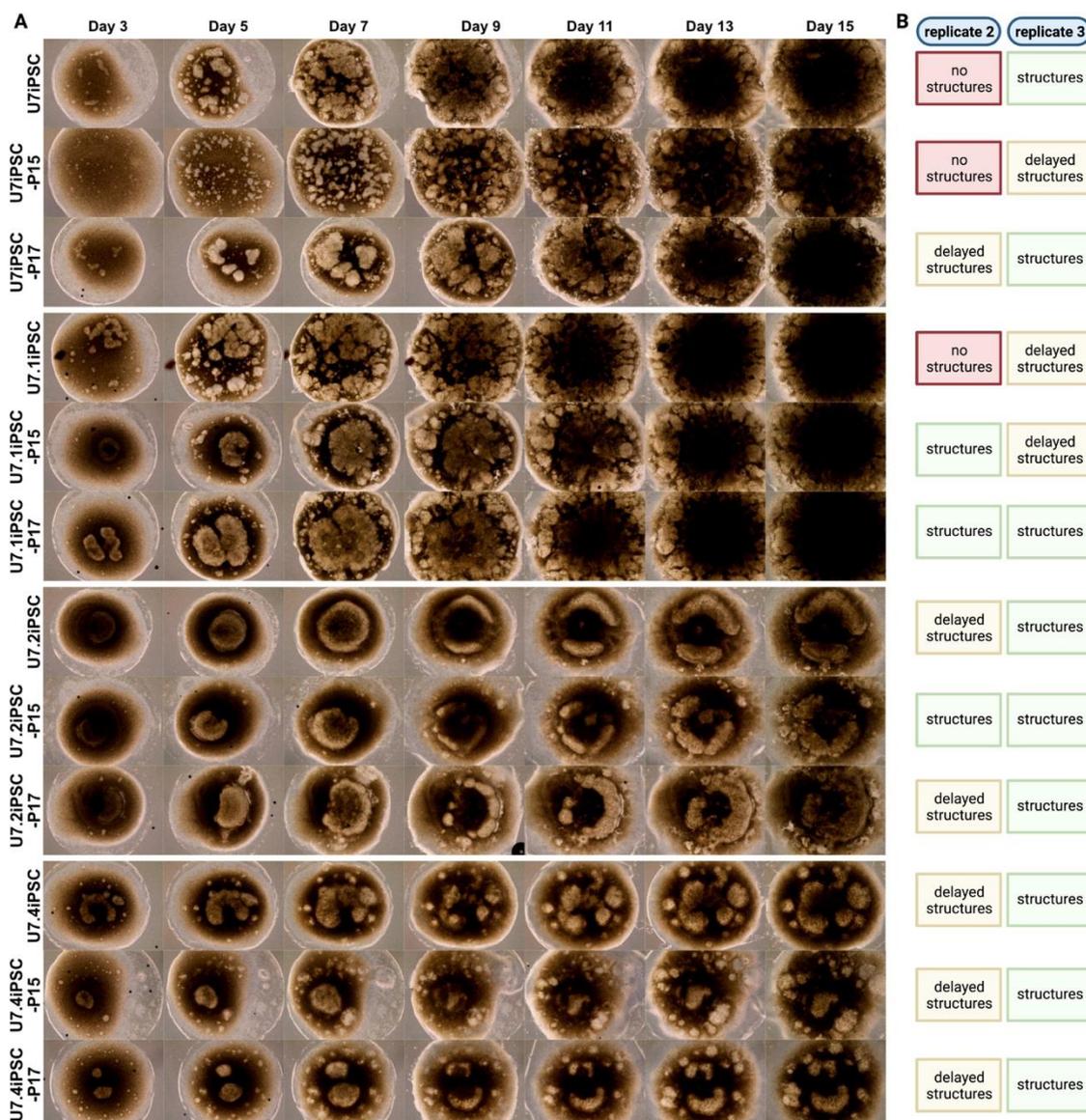
Figure 36: 3D kidney organoid structure development of WT, KO and S/MAR modified WT and KO U7iPSC lines.
A) Exemplary brightfield pictures of structure formation during 3D culturing on day 2-15 of the described variety of U7iPSC derived cell lines from one biological replicate are shown. B) Summary description of successful, delayed or lack of structure formation during kidney organoid differentiation of the same cell lines as in A) from two biological replicates are stated.

We performed FACS analysis for GFP expressing alive single cells as indicators for maintained S/MAR expression on day 0 and day 15 or 13, respectively from two replicates. To exclude loss from maintenance, a sample of the used cells for each differentiation experiment was maintained as hiPSCs and analysed together with the differentiated samples on the respective days.

RESULTS

Consistently throughout all cell lines and replicates, and in line with previous observations only moderate up to 5% loss of GFP+ population was observed after monolayer differentiation on day 0 (Figure 37 A, B). In later stages after 3D differentiation on transwell membranes, more severe loss of GFP+ population was shown. While loss in expression population most of the time varied between 20-30%, in U7iPSC-P15 and in U7iPSC-P17 loss of 50-60% of the expressing population was shown. Notably, these severe losses of GFP expressing cells throughout differentiation did appear inconsistent in different cell lines and did not repeat in the same cell line in biological replicates.
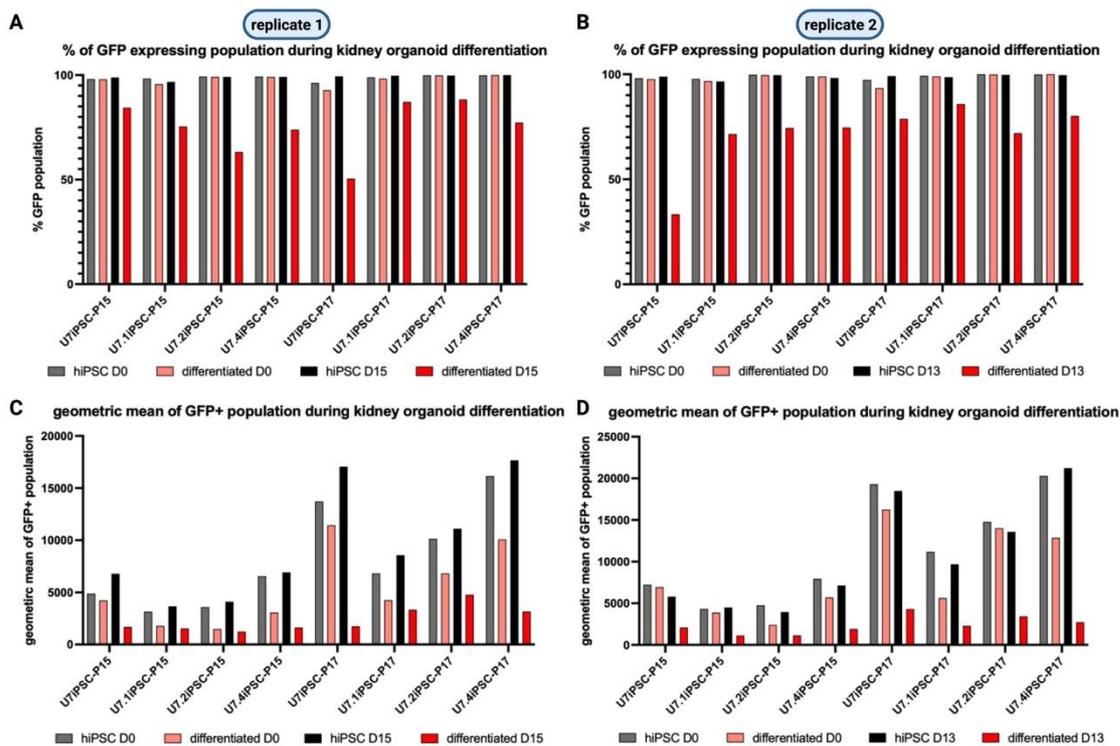
Figure 37: GFP expression maintenance during kidney organoid differentiation.
A, B) Percentage of GFP+ population in alive, single cells from FACS analysis are shown. Samples were taken from undifferentiated maintained hiPSCs at the same timepoint as kidney organoid differentiated samples on Day 0 and Day 15 or Day 13, respectively. C, D) Geometric mean of GFP+ populations from A) and B) are shown. A, C) Data from biological replicate 1. B, D) Data from biological replicate 2.

Besides the loss of GFP expressing cells, the geometric mean of the GFP+ alive single cell population was plotted for both biological replicates (Figure 37 C, D). Small changes in the geometric mean of maintained hiPSCs is most likely due to FACS laser adjustment throughout the experiments. However, data from differentiated cells can be directly compared to maintenance data from the same day (day 0 or day 15/13, respectively). While geometric mean of day 0 samples varied from showing no

difference to a slight loss in geometric mean upon differentiation, all day 15/13 differentiated samples demonstrated strong loss of geometric mean of GFP+ population compared to the respective undifferentiated hiPSC samples.

Subsequent qPCR analysis of samples from day 0 and day 15, 13 or 11, respectively, from three biological replicates was performed. As previously shown, also in these cell lines FLCN-flag, FLCN and GFP overexpression was shown in the respective S/MAR modified cell lines throughout the differentiations (Figure 38 A, B, C). The pluripotency marker nanog was not expressed anymore in almost all differentiated cell lines with strong expression in parental hiPSCs (Figure 38 D). Contrary to this, we observed maintenance of the pluripotency marker Oct4A in many cell lines, but downregulation upon differentiation in all U7.2iPSC and U7.4iPSC derived cell lines (Figure 38 E). Since these S/MAR modified KO cell lines behaved like parental ones this suggests a cell line specific, clonal effect.

As shown before, Pax2 expression was upregulated upon differentiation, with often strong upregulation on late-stage differentiation samples (Figure 38 F). Also, in line with previous observations, TBXT was strongest upregulated on day 0 of differentiation and often less expressed on later stage of differentiation (Figure 38 G). Interestingly, the severe downregulation of TBXT in late-stage differentiation samples compared to day 0 occurred predominantly in U7.2iPSC and U7.4iPSC derived cell lines. We furthermore checked for the podocyte marker PODXL, where iPSC samples showed highest expression levels (Figure 38 H). However, in line with observations on Oct4A and TBXT expression, all U7.2iPSC and U7.4iPSC derived cell lines showed consistently lowest expression of PODXL in differentiated samples.

During the differentiation protocol, the distal tubuli marker ECAD, as well as the proximal tubuli marker Vil1 were strongly downregulated on day 0, and inconsistently again expressed in late-stage differentiated cells (Figure 38 I, J). Finally, the podocyte marker WT1 was only expressed in many late-stage differentiated samples levels (Figure 38 K). Notably, there was again no consistency found in expression throughout cell lines or replicates and no connection to observed structure formation could be

drawn. However, together the data provides a more thorough picture of the obtained structures and most importantly no direct indication of any influence of FLCN absence, presence or overexpression on the differentiation was observed.

A — FLCN-flag

B — FLCN

C — GFP

D — nanog

E   Oct4A

F   Pax2

G   TBXT

H   PODXL
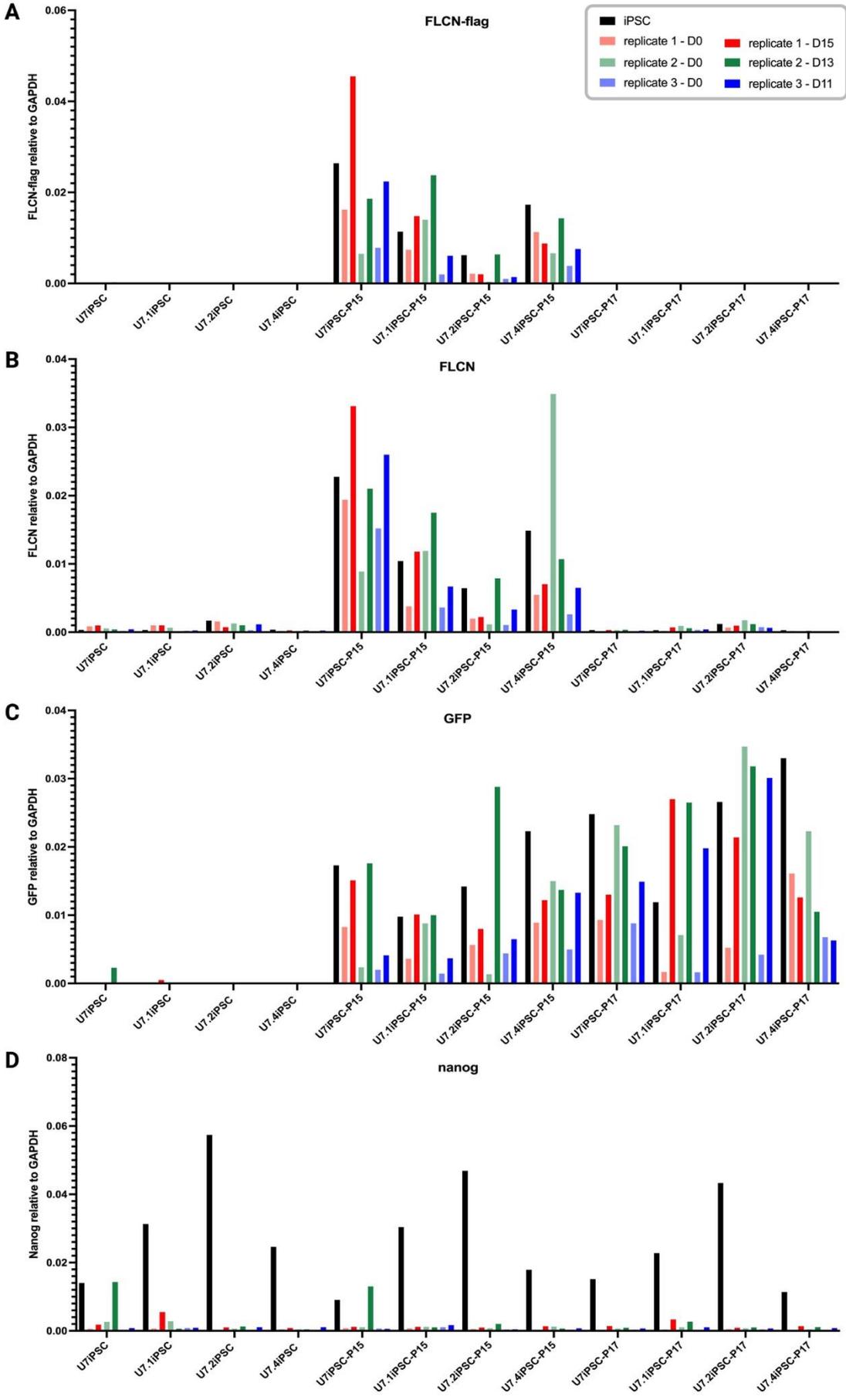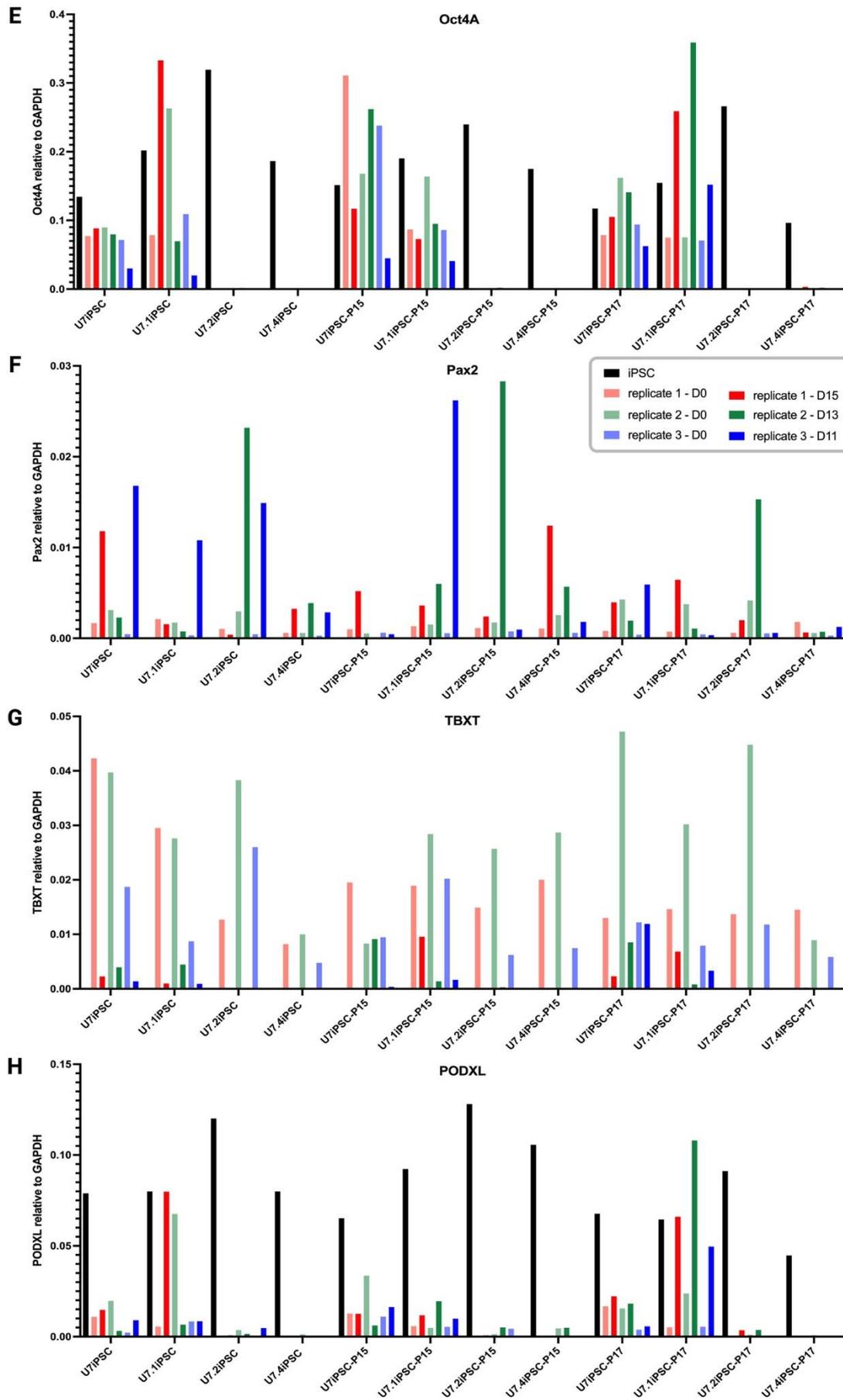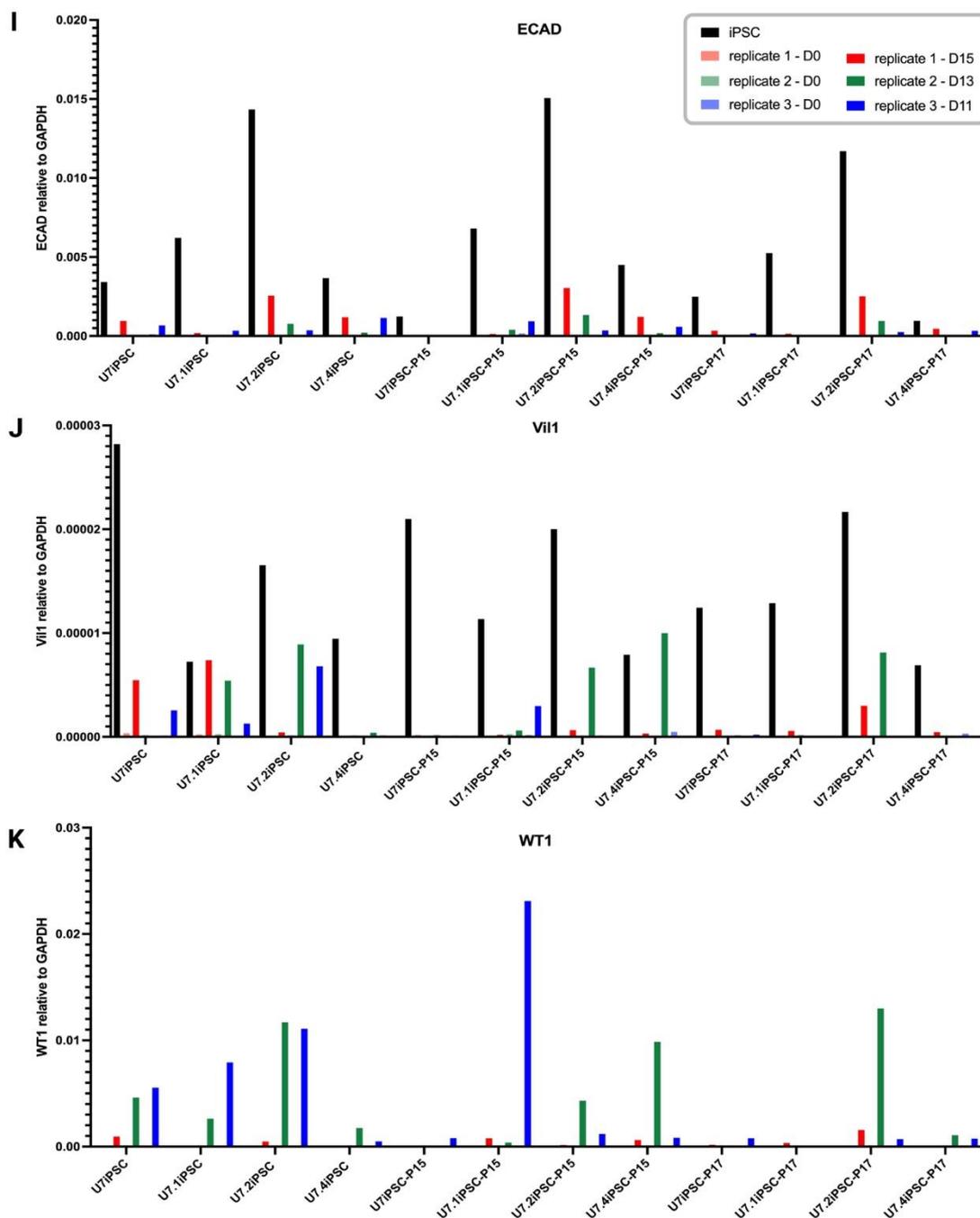
Figure 38: qPCR expression data of S/MAR expression as well as pluripotency and differentiation markers of U7iPSC derived cell lines.

A-K) qPCR data from different U7iPSCs WT, KO and stable P15 (GFP-FLCN)- or P17 (GFP)-S/MAR modified cell lines is shown. Mean of technical replicates for three biological replicates for kidney organoid differentiation at day 0, day 15/13/11 as well as undifferentiated hiPSCs as control is depicted. Description of observed trends are found in section 8.10.4. A, B, C) FLCN-flag, FLCN and GFP for validation of S/MAR DNA vector expression. D, E) Nanog and Oct4A as pluripotency marker. F) Pax2 as IM marker. G) TBXT as PPS marker. H) PODXL as podocyte marker for matured kidney organoids. I) ECAD as distal tubuli marker for matured kidney organoids. J) Vil1 as proximal tubuli marker for matured kidney organoids. K) WT1 as podocyte marker for matured kidney organoids.

DISCUSSION

8.11 Overview

This aim of this thesis was to continue the development and expand the range of applications of the optimised non-viral S/MAR DNA vectors for gene and cell therapy.

We generated a variety of S/MAR DNA vectors encoding not only GFP as a reporter gene but also the functional transgene FLCN. To validate the vectors functionality in cancer cell lines, we were able to show, that S/MAR DNA vector modified cell lines can be stably established via antibiotic selection or by passive purification of expressing cell lines via FACS sorting. In doing so, we generated stable FLCN restored UOK257 cell lines. This cell line was used in a collaborative study to shed light on the involvement of FLCN upon pathogen infection and can be further utilised as a model system to study FLCN pathways with minimal cellular perturbation caused by the vector used for the genetic modification.

We built on our previous experience with mouse stem cells and in preliminary experiments in human stem cells we further established that we could modify human fibroblasts and reprogram them to hiPSCs while maintaining S/MAR DNA vector expression. Pluripotency marker expression, as well as trilineage differentiation was not altered due to genetic modification and S/MAR DNA vector expression was persistent throughout the differentiation process as shown by maintained GFP expression.

For a more efficient and translational utilisation of our vector system for human cell therapy we showed for the first time, that we can directly modify human iPSCs at the pluripotency stage and generate stable cell lines.

To achieve this, we had to implement and establish culturing protocols such as feeder free reprogramming, passaging, single-cell plating and DNA vector delivery. Furthermore, we started a study aimed at utilising cells isolated from the urine for the

non-invasive generation of hiPSCs. We designed the study and obtained ethical approval, optimised urinary cell isolation, and implemented feeder-free episomal reprogramming of these cells.

Successful generation and culturing of these cells combined with optimised delivery of our DNA vectors and our experience in S/MAR DNA vector establishment we generated antibiotic and FACS sorting established stably modified fibroblast, as well as urinary derived hiPSCs. Characterisation of these cell lines was in line with previous experience with our vector system and showed no effect of the genetic modification on pluripotency marker expression or differentiation potential into cells of the three germ layers. Most importantly, despite the severe changes in the chromatin upon differentiation processes, transgene expression was maintained in the majority of the differentiated cells.

We were able to publish this generated data on the direct modification of fibroblast and urinary cell derived hiPSC using GFP-S/MAR DNA vectors with shared first authorship (Roig Merino et al., in press). There, we present data on expression maintenance, pluripotency validation, trilineage differentiation and expression level maintenance upon differentiation together with the work in pluripotent mouse stem cells as well as mouse chimaera formation.

Besides methodological establishments and characterisations with GFP as a reporter gene, we focused also on the functional transgene FLCN in the context of kidney tumour formation in BHD patients. FLCN-overexpression by stable modification of FLCN-S/MAR DNA vectors did not alter the behaviour of the cells.

Furthermore, we generated FLCN KO iPSC lines using the CRISPR/Cas9 system. Against our expectations and reports in FLCN KO ESCs, characterisation of these iPSC lines revealed no difference in stem cell capabilities such as differentiation potential upon FLCN KO. Utilising our S/MAR DNA vectors we stably restored FLCN in the KO cell lines and generated GFP expressing KO cell lines as vector controls.

DISCUSSION

RNA single-cell sequencing was performed on feeder-free, episomally reprogrammed, urinary derived U7iPSC lines. WT, FLCN KO, as well as these cell lines stably modified with the FLCN-flag-GFP (P15) and GFP (P17) S/MAR DNA vectors were analysed. The data supports uniform expression of S/MAR DNA vector upon stable modification and little impact of the vector itself on the host cell emphasising again the important advantage of our system compared to integrating or viral vector systems. No effect of FLCN expression levels on pluripotency state and only minor changes in expression cluster composition were observed.

Since FLCN dependent tumorigenesis in BHD patients occur only in kidneys we further implemented an advanced kidney organoid differentiation protocol which takes up to 21 days. We were able to confirm that all our generated cell lines were able to undergo the differentiation with posterior primitive streak and intermediate mesoderm marker upregulation, analysed several other marker expressions and visible structure formation in brightfield microscopy in 3D aggregates on transwell membranes. This more sophisticated differentiation was again not hindered by S/MAR modification as well as the absence, presence, or overexpression of FLCN. As during germ layer differentiation, S/MAR expression was maintained in most cells throughout the differentiation process as confirmed by FACS and qPCR.

Together our data provides first data for the feasibility to utilise our S/MAR DNA vector system with functional transgenes for the modification of hiPSCs in the context of gene and cell therapy with protocols, that can easily be adjusted for GMP manufacturing.

8.12 Composition of S/MAR DNA vectors generated and utilised in this study

This work builds on previous work from our laboratory on optimising the composition of the S/MAR DNA vectors by Dr Matthias Bozza as well as the application of these vectors in mouse stem cells by Dr Alicia Roig Merino. Combining this, we aimed to provide evidence, that our S/MAR DNA vector platform can also be applied for human iPSCs and verify its suitability for a combined gene and cell therapy with minimal impact on the cells caused by the vector.

We chose the S/MAR DNA vectors pCAG and nCAG as templates for the vector generation for this project since the CAG promoter in general and these vectors in specific were proven to express and be functional in mouse stem cells (Alexopoulou et al., 2008; Liew et al., 2007; J. Liu et al., 2009; Roig Merino, 2018; Roig Merino et al., in press). As an additional feature, some of our vectors included the insulator Ele40, which is intended to shield chromatin-associated repression and maintain transgene expression (Kwaks et al., 2003). Besides GFP as reporter gene, we also focused on FLCN as a functional transgene. For easier discrimination of exogenously introduced and endogenous FLCN, we included an N-terminal Flag-tag in some of our vectors. For later experiments, we focused on the FLCN and GFP vectors P15 and P17, which combine both additional features.

In line with other reports, the CAG promoter supported strong expression in a variety of cell types including the tested cancer cell lines and human iPSCs. However, this promoter consists of GC rich regions, which are known to be causes for silencing events and foreign DNA recognition sites (McLachlan et al., 2000; Y. Zhou et al., 2014). In stem cells, EF1a is often used as a consistent and high expressing promoter and might be a suitable exchange for future studies (Herbst et al., 2012). Preliminary experiments showed even stronger expression with EF1a S/MAR DNA vectors compared to the ones encoding the CAG promoter (data not shown).

DISCUSSION

During the time of this project, the work on our S/MAR DNA vectors and their features continued. Shorter versions of the S/MAR region or spliced versions leading to optimised stability of mRNA were introduced and should be considered in future work but due to time constraints was not yet included in this project.

When utilising functional transgenes in our vectors, we had to compromise on the position of encoded genes. With puromycin establishment and a functional transgene, expressing cells can't be easily monitored during establishment. Functional antibodies for FACS or IF analysis are crucial to evaluate the purity of the cell line when little cell death upon addition of antibiotic drug is observed. With the poorly studied protein FLCN not many well working antibodies were available. IF staining as well as FACS analysis of cells using the FLCN antibody used for WB and antibodies for the Flag tag were tried but remained difficult. Additionally, since FLCN is not a surface protein these analyses were always extensive endpoint assays and were not suitable for easy and fast monitoring of transgene expression. To facilitate this, we kept GFP as reporter gene in the expression cassettes. When trying to arrange PuroR, FLCN and GFP in one expression cassette in our P9 vector, we observed that expression levels of the third transgene were very inconsistent. While the p2a sequence is reported to result in one mRNA equal protein levels, the IRES element was used to connect the third transgene. Here, an additional transcription start is initiated and could have caused the inconsistent GFP expression (Bochkov & Palmenberg, 2006; Leppek et al., 2018). A more optimised arrangement with only 2a connective features as in the Lentiviral 4-in-1 reprogramming vector cassette could be tried in the future (Z. Liu et al., 2017; Y. Wang et al., 2015; Warlich et al., 2011). Future S/MAR DNA vectors could also be designed with two expression cassettes: One expression cassette encodes a reporter transgene such as GFP and is expressed under a constitutively active promoter driving S/MAR maintenance and facilitates establishment. In addition, a tissue-specific or conditionally active promoter with a target gene in its expression cassette could be located on the same vector.

In our attempt to establish protocols that can be easily modified and applied for translational cell therapy approaches, we were interested in an antibiotic free

establishment of our S/MAR DNA vectors. In addition, cancer cell lines are known to be able to acquire drug resistance (Duesberg et al., 2001) and extensive use of antibiotic selection pressure even if it's only during the time of vector establishment could cause changes in the behaviour of the cells (Lanza et al., 2013; Varghese et al., 2017). We thus moved to a passive establishment where we omitted the PuroR and utilised GFP in combination with our functional transgene FLCN as reporter gene to enhance culturing of S/MAR DNA vector expressing cells via FACS sorting.

8.13 S/MAR DNA vector validation in cancer cell lines

First proof of principle experiments of our generated S/MAR DNA vectors for this project were tested in cancer cell lines. HEK293T cells were previously used to demonstrate the functionality of our vectors and again provided successful establishment in this project, both via antibiotic selection as well as FACS purification.

As an alternative cell source and considering our interest in BHD, we further focused on applying or vector system in FLCN deficient cell lines. We aimed to generate FLCN restored cell lines which could be used to study function and mechanism where FLCN is involved while cause little change in the cells by the delivery vector.

We were able to stably restore FLCN in UOK257 cells, and these cells were successfully used to investigate the role of FLCN upon N. gonorrhoeae infection (T. Yang et al., 2020). However, when investigating expression levels of signalling pathways such as TGFβ or WNT, we were not able to obtain consistent results. Cell cycle analysis was attempted, and anchorage-independent cell growth was investigated via a soft agar assay, however, both assays provided no consistent data. Considering the cancerous nature of UOK257 cells these observations suggest that cell properties cannot be fully reversed simply by restoration of FLCN. While our generated cell lines can serve as a useful tool for several studies, it seems like tumorigenesis needs to be investigated in a different cell model and cannot be returned.

8.13.1 Drug-free establishment of S/MAR DNA vectors

Our focus during the experiments with cancer cell lines was the protocol optimisation of S/MAR DNA vector establishment in easy (HEK293T) and more difficult to handle cell lines (UOK257). We aimed to implement protocols with gentle handling and least influence on the host cells and thus focused on a drug-free establishment of S/MAR DNA vectors.

When utilising FACS sorting to enhance the vector expressing cell population we consistently observed similar expression behaviour. During the first 7-14 days after

transfection, a major drop in expressing cell numbers was found. This is in line with observations with transient expressions, where the majority of delivered DNA plasmids get lost and diluted during cell division. We expect only a small number of S/MAR DNA vectors to find a suitable nuclear location in an active area close to a transcription factory. Within the subsequent two weeks after transfection, we often observed another drop in the number of expressing cells. It is reported that active transcription throughout the expression cassette into the S/MAR region is required for the maintenance and replication of the S/MAR DNA vector (Hagedorn et al., 2010; Rupprecht et al., 2010). We therefore suspect, that while some vectors by chance find a place in close proximity to transcription factories in the nucleus, during this second part of establishment some S/MAR DNA vectors lose expression or replication potential since they are not optimally localized. Despite our thorough optimisation of vector composition, of course, silencing events by the host cells also need to be taken into account during the establishment process. In general, we routinely see stable maintenance of S/MAR DNA vector expression after 30 days after transfection for cell lines established with puromycin selection as well as FACS sorted cell lines.

In UOK257 cells, for the nSMAR DNA vectors FACS sorting at day 15 was sufficient to generate stable cell lines (Figure 10), while in HEK293T cells with pSMAR DNA vectors after the 17dpt sorting 17dpt still a high percentage of expressing cells were lost within the next days (Figure 9). Besides the trend of generally seeing stable vector expression maintenance after 30 days after transfection, this observation could indicate slight differences in establishment in different cell lines presumably caused by different population doubling times or advantages of the Nanovectors™ in the later stage of establishment. However, more experiments with thorough comparison in different cell lines and similar sorting timepoints need to be done to verify these observations.

When establishing S/MAR modified UOK257 cells with a variety of plasmid and Nanovectors™, against our expectations one out of four nSMAR DNA vectors and two out of four pSMAR DNA vectors lost GFP expression over time. Interestingly, the cell lines which lost expression were always ones encoding FLCN in the S/MAR DNA vector, and two out of three did not contain an insulator (Figure 10).

DISCUSSION

In general, the establishment of UOK257 cells was not as straightforward as HEK293T cells. While the UOK257-P15 cell line provides stable GFP expression in FACS analysis throughout the observed 83 days of culturing, a fading protein in WB and no overexpression was detected in qPCR during that time. Also, PCR on gDNA in HEK293T cells showed few unspecific bands (Figure 9 E) but the PCR reactions on the stable UOK257 lines showed several unspecific bands with the GFP primers. GFP-FLCN primers in the UOK257 samples resulted in no bands and only one functional positive control (Figure 11 B). The lack of a band in the 2dpt control can be explained by bad transfection efficiency or lack of sensitivity of the chosen primers. Taken together, despite the drug-free establishment the data emphasises the unstable characteristics of the UOK257 cancer cell type, which might have facilitated possible vector rearrangements.

8.13.2 Episomal maintenance of S/MAR DNA vectors

Several attempts to prove the episomal maintenance of S/MAR DNA vectors in our established cell lines were performed. Plasmid rescue experiments were previously performed successfully in our lab; however, the results are typically inconsistent. The rationale behind this experiment is, that episomally maintained circular plasmid DNA can be isolated together with gDNA, used to transform bacteria, amplified within the bacteria and the obtained plasmid prep should be identical to the initially utilised plasmid for stable cell line establishment. This has been performed before for the isolation of episomal viruses (Bardelli et al., 2017) as well as S/MAR DNA vectors (Hagedorn et al., 2012). We again extensively troubleshot the protocols. For example, gDNA isolation protocol with and without protein precipitation, and with Proteinase K digestion were performed in case the S/MAR DNA vector was tightly anchored to the gDNA via protein interactions. Traditional Hirt extraction protocols as well as isolation with the help of a miniprep column for preferred binding of plasmid DNA were tried. Furthermore, a variety of input gDNA and different bacteria for plasmid transformation was used since S/MAR sequences are repetitive and difficult to amplify (Hagedorn et al., 2012) and the rescued plasmid DNA presumably has a different, mammalian methylation pattern after endogenous replication within the host cell than

the utilised plasmid for initial transfection which was grown in bacteria (Wolff et al., 1992). Despite our efforts, we were not able to establish a reliable plasmid rescue protocol for our cell lines.

Besides plasmid rescue, RCAs were attempted several times. Here, the capability of a special polymerase to specifically amplify circular DNA is used to amplify episomal DNA from gDNA extractions. While most polymerases are not able to amplify the utilised CAG promoter, we were able to show amplification of our CAG containing S/MAR DNA vectors upon high amounts of background gDNA. However, we were not able to amplify our vectors in gDNA samples from established HEK293T lines (Figure 9 D). When calculating the required input of gDNA based on 1-2 S/MAR DNA vector copies per cell in established cell lines to reach the sensitivity limit of the assay, gDNA input of hundreds ug gDNA are suggested. We therefore speculate, that despite functional controls, we were not able to add the required amount of template into the reaction to reach the detection limit of the assay. However, we were able to amplify parts of the S/MAR DNA vector from these gDNA samples in routine PCR reactions.

Despite not being able to validate episomal maintenance of our vectors in the established cell lines, we can refer to several reports of the episomal nature of S/MAR DNA vectors (Bozza et al., 2021; Bozza et al., 2020; Hagedorn et al., 2012). Most importantly and most relevant for this project, pCAG and nCAG previously have been shown to be episomally maintained in mouse stem cells (Roig Merino, 2018; Roig Merino et al., in press).In this publication, also successful chimaera formation with S/MAR DNA vector modified mESCs were performed. While S/MAR DNA vector expression was maintained throughout in vivo differentiation, the S/MAR DNA vectors were not detected after spermatogenesis and in F1 generation, highly supporting the episomal nature of our vectors.

8.14 Generation of S/MAR modified hiPSCs during reprogramming

After verifying the generated S/MAR DNA vectors we moved on to applying these vectors in human iPSCs. Previously, S/MAR DNA vectors has been successfully applied in mouse ESCs and iPSCs. Here, despite the significant chromosomal restructuring during these processes, stable modification, as well as sustained expression throughout reprogramming and differentiation, was shown (Roig Merino, 2018; Roig Merino et al., in press). In preliminary experiments in human stem cells, GFP encoding S/MAR DNA vectors were transiently delivered to hESCs and microarray analysis confirmed little impact of the vector in the host cells. Furthermore, S/MAR DNA vectors encoding reprogramming factors were successfully used to reprogram patient-derived fibroblasts.

As a next step towards the application of our vector system in human iPSCs, together with Dr Alicia Roig Merino we simultaneously delivered EBNA reprogramming vectors and S/MAR DNA vectors to commercial human fibroblasts (Figure 12). For this, the delivery of five plasmids at the same time was necessary. To enhance reprogramming of S/MAR modified cells, for the FLCN-S/MAR DNA vector-transfected cells, GFP expressing cells were sorted on day 8 and thus only cells that expressed the S/MAR DNA vectors were continued for reprogramming. While cell morphology changes at the beginning of the reprogramming seemed normal, towards later stages of the reprogramming protocol, distinct hiPSC colonies that did not express GFP emerged, while GFP expressing cells seemed to persist in an intermediate state or simply expand at a fibroblast-like morphology. Considering the sorting of S/MAR expressing cells, this suggests that silencing of the vector during reprogramming enhanced reprogramming efficiencies, while maintained expression seemed to hinder efficient reprogramming. We were able to obtain pCAG expressing hiPSCs, while one attempt of generating FLCN overexpressing hiPSCs using the S/MAR DNA vector P15 did not result in successfully reprogrammed iPSC colonies. More experiments are needed to investigate if FLCN overexpression negatively influenced the reprogramming, or if any S/MAR DNA vector expression negatively affects reprogramming efficiency.

Once obtained, in line with previous findings in mouse iPSCs, the GFP-S/MAR modified iPSCs persisted pluripotency marker expression like the parental cells. We also showed that all cell lines were able to differentiate into cells of the three germ layers, while GFP as a transgene was persistently expressed. This is essential for a possible application of our vector system in cell therapy.

8.15 Optimisation of reprogramming, cell source and culturing conditions

8.15.1 Optimisation of reprogramming

We hypothesised, that increased efficiency of the delivery of the reprogramming vectors should result in increased reprogramming efficiency. Against our expectations, the less efficient delivery using the Amaxa 2D nucleofector resulted in higher reprogramming efficiency (chapter 8.4.1). When looking at the numbers, for the neon electroporation the double number of cells was used than for the Amaxa 2D nucleofection, while the same amount of DNA was applied. Carefully chosen expression level amounts of reprogramming factors and their balance between the factors are known to be important for successful reprogramming (Gonzalez et al., 2011; Ng et al., 2020). Our unexpected results might be caused by a different level of expression of the reprogramming vectors based on the delivery of fewer plasmids per cell. Further experiments need to be performed to establish this. Due to time constraints, this was not further investigated, but the established reprogramming with the Amaxa 2D delivery was continued.

8.15.2 Urine cells as alternative cell source for hiPSCs

Besides reprogramming vector delivery, we also investigated alternative cell sources for the generation of hiPSC. Urinary derived cells have recently been reported to be efficiently used to generate iPSCs from an easily obtainable, non-invasive source (Falzarano & Ferlini, 2019; Pavathuparambil Abdul Manaph et al., 2018). This is an important aspect when it comes to patients' wellbeing and especially when potential donors are underaged. Also, it has been shown that these urine cells are originated from the kidney. While iPSCs are very close to ESCs, it is reported that they maintain episomal memory of their cell of origin which can influence differentiation potential (K. Kim et al., 2010; Siller et al., 2016). With our interest in BHD and its caused kidney cancer the kidney origin of the iPSC source might be advantageous. A study plan for the work with human urine material was developed and ethical approval was obtained from the ethics committee of the medical faculty of the University Heidelberg (study

S-550/2019). First experiments on isolation and reprogramming of USCs were promising and encouraged us to continue working on this.

8.15.3 The importance of hiPSC culturing conditions

When it comes to consistency and reproducibility in experiments and especially when heading towards treating a patient, it is important to have a system in place which has consistent media compositions, is xeno- and animal-free and can possibly be easily adjusted to the use of GMP reagents (Zakrzewski et al., 2019). We thus strived to move from mouse feeder dependent cell growth and unreliable manual picking towards a feeder-, xeno- and animal-free system. After testing several matrix-media combinations, iMatrix Laminin-511 and StemFit Basic02 media provided best results for routine hiPSC culturing (chapter 8.4.3). With the feeder-independent nature of this combination and a manufactured GMP version of the media, it should be possible to move protocols established in this fashion from the current conditions towards a clinically relevant protocol.

DISCUSSION

8.16 Direct modification of hiPSCs with S/MAR DNA vectors

After optimisation of cell culture conditions to feeder-free iPSC culturing, we also facilitated the direct modification of hiPSCs without the interference of another cell type. Both reprogramming and S/MAR DNA vector establishment are inefficient and time-consuming procedures. While it is for example reported for Fanconi Anaemia, that the FancA gene is essential for successful reprogramming (Muller et al., 2012), if lack of the gene of interest does not interfere or even enhance reprogramming it might be better to perform these steps separately. The direct modification of hiPSCs opens a way to use off-the-shelf, validated stocks of iPSCs and increase efficiency by ruling out reprogramming. To make this feasible, first several steps had to be optimised.

While stem cells are by nature refractory to genetic modification, by optimising the culturing conditions along with testing different delivery methods we were able to deliver our S/MAR DNA vectors to our feeder-free hiPSCs (chapter 8.5.2). The use of puromycin enabled us to generate stable P17, pCAG and nCAG modified cell lines. There was no obvious difference or advantage between the performance of the different vectors and the cell lines did not show any alterations due to the modification. As in the iPSC lines that were modified during reprogramming, pluripotency characteristics were maintained and trilineage differentiation showed persistent GFP expression in these cell lines. Data performed with the pCAG and nCAG-S/MAR DNA vectors was published together with the work in mouse stem cells (Roig Merino et al., in press).

Instead of the previously performed random differentiation via EB formation we implemented a more reliable commercial trilineage differentiation. With a distinct germ layer per well we were already able to distinguish by brightfield microscopy neuronal rosettes in ectoderm samples and distinct tubular structures in endoderm samples. Furthermore, we improved the IF protocol to be able to record GFP expression in differentiated tissues simultaneous to germ layer marker staining. We found that the combination of acetone/methanol fixation denatured the GFP barrel

form and thus its fluorescence ability. However, fixation using PFA, handling in the dark and the usage of cell culture grade PBS with respective reliable pH enabled the maintenance of GFP throughout the staining procedure.

8.16.1 Drug-free establishment of S/MAR modified hiPSCs

Having established protocols for hiPSC reprogramming, culturing, and DNA delivery, we next applied our experience from cancer cell lines and implemented a drug-free establishment method with FACS sorting.

In general, the observations of expression loss during the first 30 days were similarly in hiPSCs as seen in cancer cell lines. However, hiPSCs are very sensitive and sorting of too few cells resulted often in poor survival of cells. Timings for sortings needed to be balanced between a too early sorting with not enough expressing cells and extensive culturing of not-expressing cells with the usage of expensive media and matrix. We found that close monitoring of cell growth behaviour and time point of extensive expression loss was important to determine best sorting points for each cell line. The first experiment with patient-derived CHM-iPSCs suggests already stable expression of S/MAR DNA vectors in cells 16dpt since there is little fluctuation between the GFP+ population count in FACS analysis on day 16 and day 24 with stable maintenance after sorting on day 24 (Figure 19, chapter 8.5.3). Over the course of our project, however, similar to what was done with cancer cell lines two sorting timepoints, one around day 10-14 and one around day 30 after transfection seemed effective and efficient regarding the discussed cell count and culturing expenses.

8.17 Optimised generation of hiPSCs and S/MAR modification

8.17.1 Feeder-free reprogramming

We completed our attempt to generate cells that were animal- and xeno-free handled by testing a feeder-free reprogramming protocol (chapter 8.6.1). This experiment was combined with a test for optimised usage of time and cell source since the latter is often a precious material when working with human samples. Leftover cells from day 8 of reprogramming were, thawed, reprogramming was continued, and we were able to generate distinct iPSC colonies with this shortened time window. This suggests the possibility of saving cell resources as well as time for repeated reprogramming experiments. However, it must be noted that we saw severely decrease reprogramming efficiency in NHDFs that were stored for more than one year in liquid nitrogen. We expect also decreased reprogramming efficiency of cells frozen within reprogramming when stored for an extended period of time. Thorough comparison studies of feeder-free, feeder-dependent and frozen cell reprogramming needs to be performed to make a statement about absolute efficiencies.

8.17.2 Optimisation of urinary derived iPSC generation

After promising preliminary generations of Lentivirus reprogrammed feeder-dependent derived UiPSCs (chapter 8.4.2) we continued optimising the isolation and established episomal, feeder-free EBNA reprogramming based on the publication by (Mulder et al., 2020).

Isolation efficiency comparison proved to be challenging since many factors need to be considered. In other studies, often efficiency was based on a certain volume of obtained urine. However, one could argue that at a certain point more urine might not flush out more cells and viable cells are just diluted. In general, our experiments showed an increased occurrence of squamous cells in samples from female donors. This also led to differently sized cell pellets and distribution of one sample into smaller aliquots for comparison had to be performed.

When comparing different media conditions, it is important to exclude donor specificity. It was not certain if the gender of the donor influences isolation efficiency. Also, when using the same donor subsequent samples might mean that fewer cells are flushed out in the later sample. In line with this, it has been reported that there are more cells in the morning urine. While we never experienced problems with contaminations, morning urine might contain higher contamination potential. We also experience extensive cell count after several evenings of alcohol consumption of one donor, but failure in isolation suggesting that mostly dead cells were flushed out.

We attempted media comparison for USC isolation to the best of our knowledge but in general saw high success rates. Cells solely isolated in DMEM/F12 sometimes expanded into well sized colonies but failed to further expand. This did not occur when isolating the cells with the combination media based on the publication of (Mulder et al., 2020).

In line with our approach of optimally generated cells with least alterations in the host cell, we aimed towards a viral free, episomal reprogramming protocol. This time, we successfully delivered EBNA reprogramming vectors to the USCs and a reprogramming efficiency between 0.05-0.08%, similarly to that of fibroblasts, was obtained. However, the experiment was performed with freezing in between due to Covid-19 lockdowns of the laboratory and different reprogramming vector delivery methods were used for the different cells. Therefore, a thorough comparison still needs to be done to verify differences in reprogramming efficiency and speed of reprogramming.

8.17.3 Transfection efficiency of hiPSCs from different sources including the functional transgene Folliculin

After obtaining fibroblast and urinary cell-derived iPSCs in a completely feeder-free manner further optimised the delivery of our DNA vectors to these cells. Optimised cell seeding, single cell plating with the use of ROCKi (Watanabe et al., 2007) and the testing of different lipofectamine stem reagent and DNA amounts allowed us to reproducibly transfect a high number of cells with reliable high cell viability.

During the handling of the iPSCs and the transfection procedures, special emphasis was put on keeping the conditions as consistent and equal as possible. A strict passaging schedule of five days with constant cell seeding and media changes was followed, and cells were handled with as little stress as possible. Despite these efforts and attempts for consistency, we still obtained high variability in the transfection efficiencies of three biological replicates for each of the three cell lines (Figure 22). Routinely, ROCKi was removed between 20-30h after plating. While we did not observe any impact on cell growth when keeping within this time period in routine maintenance the exact timing and affiliated addition of transfection mix could be further investigated in regard to transfection efficiency.

Besides high variability, more importantly, donor specificity was observed to be an important factor for transfection efficiency. Both tested UiPSC lines demonstrated a distinct transfection efficiency pattern which further suggests that not cell type but rather donor or cell line had the highest effect on transfection efficiency. It would be helpful to investigate further replicates and iPSCs from different donors or sources or cells from the same donor but different cell sources to further emphasise this observation.

Previously, the functional transgene Rep1 was transfected and successfully established in CHM-iPSCs in a drug-free manner. Building on this preliminary experiment, we now for the first time applied our cloned and tested FLCN-S/MAR DNA vectors in hiPSCs. Direct comparison of FLCN- and GFP-S/MAR DNA vectors showed higher transfection efficiency with GFP-S/MAR DNA vectors throughout all the used cell lines. Considering the profound functional role of FLCN in cellular pathways compared to GFP this finding was expected. The obtained transfection efficiency with FLCN S/MAR DNA vectors still provided high enough percentages to be used for further establishment.

8.17.4 Drug-free GFP- and FLCN-S/MAR DNA vector establishment

The establishment of FLCN- and GFP-S/MAR DNA vectors in the feeder-free generated UiPSCs was performed with the help of FACS sorting. Data performed with the pCAG and nCAG-S/MAR DNA vectors was published together with the work in mouse stem

cells (Roig Merino et al., in press). Sorting on 27 or 28dpt showed a high variability of maintenance of S/MAR DNA vector expression (Figure 23). In general, FLCN-S/MAR DNA vectors seemed to be maintained at a higher percentage at the later stage of establishment. However, cell growth and expression numbers varied strongly within samples from the beginning, and to facilitate establishment the maximal amount of GFP expressing cells of each sample were sorted and continued to be cultured. The different starting cell numbers after each sorting could have resulted in altered growth or maintenance and in return could explain the variation in expression patterns.

8.17.5 Maintenance of S/MAR DNA vector expression

While most cells were successfully established and maintained S/MAR DNA vector expression, five out of 15 established cell lines lost expression during continued culturing. Data performed with the pCAG and nCAG-S/MAR DNA vectors was published together with the work in mouse stem cells (Roig Merino et al., in press). Notably, the establishment performance on 27/28dpt seemed to be independent of long-term expression persistency since the cell lines which lost expression demonstrated a high variety of GFP population at these time points. Also, the observed impact of function transgene expression on transfection efficiency did not recur in long-term maintenance expression since the expression losing cell lines were independent of the expressing S/MAR DNA vector, its encoding transgene and the presence or absence of the insulator.

We hypothesised that the loss in GFP expression indicated uncompleted establishment at the last sorting. To further investigate this, the five cell lines which were losing expression were sorted for GFP+ cells at 98dpt. Subsequent culturing showed that three cell lines persisted GFP expression, while two again dropped GFP expressing population percentage. This indicates that timing was not the only driver of expression loss.

The data further suggests that more sorting timepoints during the establishment time might result in better expression stability in the S/MAR modified cells. It should be

noted that during the first sorting the sorting gates for GFP expression were not very stringently chosen, since the GFP+ population had a wide GFP intensity distribution. However, the GFP expressing population was very distinct and stringently gated during the second sorting. Also, stringent gating at the 98dpt sorting however did not prevent the observed expression loss of two cell lines. However, despite the exclusion of doublets, FACS accuracy issues such as general inaccuracy or sorting of doublets including GFP- cells cannot be completely ruled out.

Since FLCN is involved in cell cycle progression (Kawai et al., 2013) and extensive expression of proteins can affect cell resources, we assumed a growth advantage of GFP- cells over GFP+ cells. XCelligence growth assay analysis of both populations however suggests only a slight advantage and exclude this as the main driver for the observed loss in GFP expression since the respective cell lines show the least effect of growth differences of GFP+ and GFP- populations.

When amplifying the GFP sequence from gDNA of both GFP populations we were able to confirm the presence in GFP+ as well as GFP- populations. Assuming high accuracy of GFP- sorting and low content of potentially remaining GFP+ cells in this sample, we thus concluded that the vector was silenced in GFP- cells and the failure of expression did not result from an overgrowth of cells that were not transfected or lost the vector at an early time point and were sorted with the GFP+ cells by accident. With the taken precautions on consistent handling and passaging of all cell lines simultaneously the cause of this silencing event however was unclear. S/MAR DNA vectors are thought to be lost when transcription is not active throughout the S/MAR feature (Rupprecht et al., 2010). Thus, our finding of the presence of GFP encoding DNA might indicate recently occurring silencing events.

8.17.6 Transgene expression levels in hiPSCs and differentiated progenies

Besides GFP as reporter gene we also confirmed FLCN and GFP overexpression via qPCR in the established, feeder-free obtained S/MAR-UiPSC lines. WT hiPSCs demonstrated weak FLCN expression compared to the strong expression under the

CAG promoter. Interestingly, mRNA levels of the overexpressed transgenes varied strongly between established lines without a clear vector or donor dependency (Figure 24). Similar observations were made in established UOK257 cells. Taken together, our experiments suggest that expression levels cannot be directly controlled but seem to occur by chance and is likely due to the random position of the S/MAR DNA vector in the nucleus and its transcriptional surrounding.

The unaltered characteristics of the iPSCs were shown again by confirmed expression of pluripotency markers in all these cell lines. Also, sustained GFP expression upon successful trilineage differentiation was shown again. FACS data revealed higher geometric mean of the GFP expressing populations in GFP-S/MAR modified cells than in FLCN-S/MAR modified cells. This is in line with expectations of additional stress on cells with the expression of the functional transgene FLCN.

Upon differentiation, a slight loss in GFP expressing cells was observed. However, the cells that maintained GFP expression throughout the process showed substantially reduced geometric mean of GFP fluorescence. This indicates a significant effect of expression pattern changes on the S/MAR DNA vectors and is likely influenced by the location of the established S/MAR DNA vectors and how transcriptionally active their surroundings in the nucleus are before and after differentiation. Data performed with the pCAG and nCAG-S/MAR DNA vectors was published together with the work in mouse stem cells (Roig Merino et al., in press).

8.18 Generation and characterisation of Folliculin knock-out UiPSCs

We had set up a collaboration to obtain BHD patient USCs from the German interdisciplinary special consultancy for BHD patients at LMU Munich. Due to circumstances of the SARS-CoV-2 outbreak, there were not enough patients of this rare disease scheduled to obtain these samples in time for this project. Patient-derived iPSCs hold high promises for the prediction of drug response and disease severity in personalized medicine. However, suitable cell lines for comparison are sometimes missing due to differences in donor age, gender, and epigenetic and genetic background. This can hinder functional studies (Musunuru, 2013). Since its discovery, CRISPR/Cas9 serves as a key tool to generate disease models and has also been successfully applied in hiPSCs (E. J. Kim et al., 2017; Song & Ramakrishna, 2018). It enables scientist to study a disease with healthy and disease cell lines from the same genetic background allowing to solely focus on the effect of the altered gene.

We therefore also made use of this technology and generated our own FLCN KO UiPSC lines for our further investigations (chapter 8.7). We only investigated FLCN KO lines from two different donor and clonal cell lines generated with the same gRNA. Due to time restrictions, the generation of more cell lines with different gRNAs to avoid off-target effects were not possible but would be advisable (Byrne et al., 2014).

The successful KO of FLCN was confirmed by Western Blot. Due to the design of the gRNA to achieve a frameshift instead of removing parts of the FLCN gDNA qPCR assays still detected the presence of FLCN mRNA. Our characterisations of pluripotency markers did not reveal any differences of the FLCN KO UiPSCs compared to the WT UiPSCs. Like reports in BHD$^{-/-}$ mESCs, FLCN KO in hiPSCs seemed to have no effect on growth rate despite the involvement of FLCN in cell cycle (Cash et al., 2011). The authors suggested the special cell cycle machinery in ESCs to cause this observation.

Further analysis concentrated on the differentiation capabilities of the FLCN KO iPSCs. Previous reports show, that homozygous FLCN KO is embryonic lethal in mouse models (Hasumi et al., 2009). This is in line with studies performed with human and mouse

ESCs, where absence of FLCN inhibits cells to differentiate (Betschinger et al., 2013; Mathieu et al., 2019). It was shown that exit of pluripotency was linked to the localisation of the transcription factor Tfe3, and that FLCN acts during the switch from naïve to primed pluripotent states both in human and in mouse cells (Betschinger et al., 2013; Hong et al., 2010; Villegas et al., 2019). However, when performing growth factor withdrawal, unlike the reported FLCN KO mESCs our FLCN KO UiPSCs lost AP expression (Betschinger et al., 2013). Further elaboration of this observation confirmed successful trilineage differentiation of the FLCN KO iPSCs. This finding suggests a primed state of our generated UiPSCs and is in line with reports of obtention of primed hiPSCs upon the use of Yamanaka factors and still inefficient generation of naïve hiPSCs with only recently developed protocols (Bredenkamp et al., 2019).

The phenotypic arrest in naïve pluripotency was reported to be rescued in gene expression RNAseq principal component analysis by Wnt inhibition, however, no differentiation experiments were performed leaving the effect of the inhibition on further cell behaviour upon differentiation unknown (Mathieu et al., 2019). To our knowledge, there has not been any generated FLCN KO hiPSC lines reported which makes our generated cells the currently developmentally earliest possible model to investigate the effect of FLCN on further differentiation processes without required treatment of the cells to enable differentiation. The ability to differentiate these FLCN KO UiPSCs enables the differentiation into a multitude of FLCN KO derived cell models for which differentiation protocols are known but long-term maintenance and thus KO generation would otherwise not be feasible.

DISCUSSION

8.19 Restoration of Folliculin in Folliculin knock-out UiPSCs

To obtain a control cell line to the WT, FLCN overexpressing WT and FLCN KO cells we further established our S/MAR DNA vectors in the generated FLCN KO UiPSCs and restored FLCN with strong expression as well as used GFP-S/MAR DNA vectors as vector control. Again, no effect of restoration of FLCN or GFP-vector expression was found regarding pluripotency or differentiation potential. Again, only minor drop in GFP expressing population with significant loss of geometric mean of GFP intensity was observed upon trilineage differentiation. As with stable S/MAR DNA vector modified WT UiPSCs, endoderm differentiation repeatedly demonstrated the strongest loss in vector expressing population. It would be interesting to also analyse further cell lines from different donors or sources in this regard.

8.19 Restoration of Folliculin in Folliculin knock-out UiPSCs

8.20 Expression comparison of different UiPSC lines

8.20.1 S/MAR DNA vectors have little effect on cell expression

Previous data from our group in cancer cells and primary T-cells suggest little expression changes of the host cell upon S/MAR DNA vector modification while commonly used Lentiviral vectors demonstrate severe changes in expression patterns (Bozza et al., 2021; Bozza et al., 2020). We further showed that this also applies to our CAG-promoter containing S/MAR DNA vectors in transiently transfected hESCs (Roig Merino et al., in press).

In this study, we furthermore conducted single-cell RNAseq of optimally generated UiPSCs with and without S/MAR modification (chapter 8.9). In line with expectations, upon stable expression of the GFP S/MAR DNA vector, WT U7iPSCs showed only minor expression changes below the Log_2 fold despite visible strong overexpression of GFP. We further were able to see a homogeneous expression of S/MAR and GFP in the stable cell line which is in line with observations of a distinct GFP expressing population in FACS analysis.

8.20.2 The impact of presence, absence, and overexpression of Folliculin or GFP-S/MAR DNA vectors in UiPSCs

We further investigated if the different levels of FLCN, GFP and S/MAR expression in our six cell lines had an impact on the pluripotency state of the cell population beyond the pluripotency characterisation that was routinely performed and demonstrated no negative impact.

Genetic markers for primed and naïve pluripotency were based on the human pluripotent stem cell naïve state qPCR array (Stem Cell Technologies). The high cell number in the GFP cluster of the U7iPSC-P17 sample makes it difficult to compare ratios of the other clusters between the samples. However, the GFP cluster overlaps with the general pluripotency cluster and thus change in cell percentage in the GFP cluster might not indicate a change in pluripotency state. Interestingly, this much stronger GFP expression in this sample was not detected in previous qPCR analysis

Figure 27. Additional analysis of differentiation markers, as well as a further breakdown of primed pluripotency markers of the separate samples should be performed and compared but was not feasible due to time constraints. However, the data support that all our tested UiPSCs cells are predominantly primed pluripotent stem cells with an even smaller naïve cell population in FLCN KO cell lines. This is in line with our observation of successful differentiation of FLCN KO cells due to lack of naivety (further discussed before in chapter 8.18). Furthermore, our data support that the undifferentiated pluripotency state is not strongly affected by FLCN, GFP or S/MAR expression.

Additional unsupervised clustering of the data was used to investigate general expression differences in the different samples. As expected, GFP clusters were only observed in S/MAR modified cells. Due to the generation of the FLCN KO via frameshift endogenous FLCN mRNA was detected but could be distinguished from S/MAR DNA vector derived FLCN expression via the Flag tag in the S/MAR DNA vector. Again, only slight differences in cluster composition were observed. For both cluster comparisons, the three FLCN KO based cell lines demonstrate fewer differences in composition than the WT based cell lines. Since the KO cells were generated via single-cell clonal outgrowth, WT cell lines were only manually expanded and clonal cell nature cannot be guaranteed, which might play a role in the observed effect.

Due to time restraints, further analysis and comparison of the samples were not possible. Especially pathway analysis of WT vs KO and restored iPSCs might give further insights into functions of FLCN. However, due to the apparent lack of essential role of FLCN in hiPSCs, further single-cell analysis of differentiated progenies of the generated cell lines might provide more valuable data.

8.21 Kidney organoid differentiation

Our results suggest little effect of FLCN absence or presence in our UiPSCs. The most severe health risk of BHD is the development of kidney cancer. Therefore, we were interested in the effect of FLCN expression in differentiated kidney cells. After verifying our S/MAR DNA vector system for the use in hiPSCs we utilised kidney organoid differentiation with our generated cell lines to implement advanced cellular models to investigate the role of FLCN in kidney cells.

Kidney organoids from paediatric urine samples have been successfully generated recently using a slightly different differentiation protocol (Mulder et al., 2020). We also considered the kidney origin of the cell source for our UiPSCs advantageous for kidney differentiation due to the epigenetic memory that can affect the differentiation capacity towards specific lineages (Shi & Cheung, 2021). In collaboration with the lab of Dr Martin Hoogduijn (Erasmus MC, Rotterdam) we implemented kidney organoid differentiation based on the protocol by (Garreta et al., 2019). Throughout our experiments, we observed inconsistency in differentiation outcome between batches of experiments within the same cell lines. This was also reported by our collaborators. To improve reliability, we made sure to use fresh basal media to avoid acidification and aliquots of reagents to avoid extensive freeze-thaw cycles. Aggregates were also spread on several 96-well v-bottom plates to reduce the time outside the incubator during the transfer on transwell membranes. Due to the necessity of at least 30 days of culturing for S/MAR DNA vector establishment and further extended culturing for the generation of FLCN KO cells our used cell lines had a high passage count, but similar behaviour was seen with low passage cells. While the protocol recommends endpoint analysis at 16 days of 3D culturing, we observed extensive cell death and subsequent disaggregation of our organoids at such a late time point. Taking into account the kidney origin of our cells, and that the initial protocol and the work from our collaborators were performed with fibroblast derived iPSCs this could indicate an effect of the cell origin on the differentiation outcome. A head-to-head differentiation comparison of kidney organoids from our UiPSCs should be performed but was not possible due to time constraints.

Due to the progressing disaggregation of matured organoids fixation and stainings of H&E and IHC were started but not routinely established and need to be implemented further. Instead, focus was put on qPCR analysis. During previously performed trilineage differentiation, only GFP maintenance was analysed via FACS analysis. Expression data of kidney organoid differentiation now showed sometimes reduced GFP expression on day 0 and increased expression at later stages of differentiation in line with continued chromosomal restructuring during differentiation (Figure 38). Furthermore, most changes in GFP expressing populations and the geometric mean of GFP was observed after the second part of differentiation performed in 3D, further emphasising the importance of studying elaborate differentiation protocols with our vector system instead of only trilineage differentiation.

Primers for Oct4A were designed specifically for the isoform linked to pluripotency (X. Wang & Dai, 2010). Upon differentiation of our cell lines, we observed cell line dependent behaviour. All U7.2iPSC and U7.4iPSC based cell lines demonstrated the expected downregulation upon differentiation while WT U7iPSC and U7.1iPSC derived cell lines maintained expression. It has been reported, that EBNA vectors as the ones used for reprogramming are lost at a rate of ~5% per cell cycle (Nanbo et al., 2007). Still, after EBNA reprogramming towards cell therapy approaches it is good practice to screen the obtained iPSC colonies for the absence of EBNA vector expression (Mulder et al., 2020; Zakrzewski et al., 2019). This suggests that maintenance of EBNA reprogramming vectors in fact can occur. Due to insufficiently established culturing conditions, single colony expansion of our hiPSC were generally not performed and only feasible when the KO cell lines were generated. The observation of inconsistently but parental cell line dependent maintenance of Oct4A expression upon differentiation might be an indicator for maintained EBNA reprogramming vectors in the U7iPSC and U7.1iPSC based lines. Further analysis in this regard, such as PCR on gDNA for the EBNA vectors, and its impact on the differentiation needs to be done. Especially, since cell line dependant patterns similar to the Oct4A cell line pattern were observed for TBXT and PODXL. Despite inconsistencies and the need for further analysis, our data provides suggests successful differentiation towards kidney cells.

We were particularly interested to see if FLCN KO triggered tumorigenesis can be studied in this complex cellular model. Against expectations, we did not observe any difference in differentiation potency and structure of all our generated cell lines with varying FLCN expression. Thus, simple demonstration of tumorigenesis as shown with primary gastrointestinal organoids (Li et al., 2014) or polycystic kidney disease in kidney organoids was not achieved (Garreta et al., 2018; Miyoshi et al., 2020). The FLCN induced tumorigenesis in kidney cells suggests an important role of FLCN in adult kidneys. However, there was no increase in endogenous FLCN expression visible upon our kidney organoid differentiations. All currently published kidney organoid protocols are successful at establishing a variety of kidney cell types but are still a big step from generating an organoid that resembles adult kidneys. The protocol we based our work on claims to achieve organoids that resemble human foetal kidneys from the second trimester (Garreta et al., 2019). Thus, future improvements of protocols might allow the study of FLCN dependent tumorigenesis in the dish. Until then, our cell lines and the established protocols can serve as a tool to investigate the involvement of FLCN in early developmental stages.

An attempt to perform single-cell RNAseq in kidney organoid samples was hindered due to time constraints. We believe that such analysis in the future would provide further insights into the developmental state of our organoids, the influence of cell source and the effect of our FLCN-S/MAR DNA vectors. Furthermore, single-cell data can indicate if loss of vector expression appears uniformly in the differentiated kidney cell types or if subgroups demonstrate different behaviours.

9   CONCLUSIONS

In this work we build on the optimised S/MAR DNA vector platform from Dr Matthias Bozza and its application in mouse pluripotent stem cells and elaborated the impact of our work towards translational gene and cell therapy with human iPSCs in the scope of BHD. Briefly, we…

- …generated further S/MAR DNA vectors specifically designed for work in human stem cells

- …moved from GFP as an easy-to-track reporter gene to the functional protein Folliculin and the translational question of the suitability of our vector system as gene therapy vectors for Birt-Hogg-Dubé syndrome

- …verified the expression of vectors in cancer cell lines where we

- …established our S/MAR DNA vectors with the help of antibiotic selection and in a drug-free manner

- …provided proof of principle experiments that S/MAR DNA vectors can be established and maintained simultaneously to reprogramming

- …utilised urinary derived cells as non-invasive, easy to obtain cell source for hiPSCs

- …implemented reprogramming, culturing and transfection protocols to enable the direct modification of hiPSCs from different sources

- …modified hiPSCs with our GFP-S/MAR DNA vectors with antibiotic selection

- …transitioned our protocols towards ones that are xeno- and animal-free and easy to transit to GMP production for future work

- …generated fibroblast and urinary derived hiPSCs with these optimised protocols

- …modified these optimally generated UiPSCs with GFP- and FLCN-S/MAR DNA vectors in a drug-free setting

- …generated Folliculin knock-out cell lines with CRISPR/Cas9

- …characterised the generated KO lines and showed no impact on differentiation capability enabling early developmental investigations on FLCN function in various tissues

- …restored Folliculin expression in the KO lines using our S/MAR DNA vector system

- …demonstrated in line with our work on mouse stem cells that S/MAR DNA vector modification does not alter pluripotency marker expression and differentiation capability in all our established cell lines

- …showed persistent GFP and FLCN expression during germ layer differentiation in all our generated cell lines

- …showed little impact of established GFP- and FLCN-S/MAR DNA vectors on the host cells via single cell RNAseq

- …observed little impact of FLCN absence, presence, or overexpression in hiPSCs via single cell RNAseq

CONCLUSION

- …implemented a kidney organoid protocol

- …demonstrated successful 3D kidney organoid structure development and validated marker expression in all established cell lines

10 OUTLOOK

With this work, we were able to provide proof of principle data that our S/MAR DNA vector platform can be applied for the genetic modification of human iPSCs. Our tool allows this genetic modification to be made with little impact on the cell, enabling studies without major bias caused by the delivery vector. Our S/MAR modified cancer cell lines were successfully used for the investigation of the role of FLCN on *N. gonorrhoeae* survival (T. Yang et al., 2020).

With the major advantages of hiPSCs due to their proliferation and differentiation capabilities, the modification of these cells allows a much wider range of possible applications of our vectors. We build on our previous reports on the generation of S/MAR DNA vectors expressing chimaeras towards human pluripotent cells. This work provides evidence, that elaborate differentiation protocols can be performed with maintained S/MAR DNA vector expression from modified hiPSCs. Besides the study of cellular functions and the generation of disease models, our collaborators (Dr Martin Hoogdujin, Erasmus MC, Rotterdam) for example attempt to utilise our vector system to conditionally express genes important in kidneys for the generation of a more accomplished kidney organoid model. Application in further differentiation protocols like hematopoietic progenitor and NK cell differentiation are planned in our lab.

We emphasised the possible use of cells derived from the urine for the generation of patient-derived iPSCs for disease modelling, drug screens but also in combination with our vector system for gene and cell therapy approaches. This is especially important to emphasise when considering conditions, where predominantly children are affected as in the monogenic disease Fanconi anaemia which we have been working on in our group before (Roig Merino, 2018). Cell sampling proves to be easier, non-invasive and the less ethical concerns allow wider application and benefits for the patients. The focus of this work was laid on the utilisation of these USCs as source for hiPSCs. As outlined in the introduction (chapter 5.9), these cells are reported to already contain some features of adult stemness. Therefore, direct modification of these cells without reprogramming and subsequent differentiation towards e.g., cartilages present

OUTLOOK

promising further possibilities for the use of these cells. Besides saving time and effort for the reprogramming process, the ability to directly modify adult stem cells in general and utilising avoids the elaborate purification process needed for iPSC-based cell therapies to avoid teratoma formation due to remaining undifferentiated cells. Current work of our group therefore also focuses on the genetic modification of HSCs with our S/MAR DNA vector platform.

The establishment of the range of FLCN KO, FLCN overexpressing and restored cell lines can be further utilised for the analysis of the role of FLCN in kidney cells. We generated a developmental early FLCN KO model that is capable to differentiate. Improvements in kidney organoid differentiation protocols could allow researchers to investigate FLCN tumorigenesis in such advanced cell models. To gain more insights into affected pathways and isogeneity of our created cell lines, single-cell RNAseq can shed light on the expression profiles of our different kidney organoids. Besides kidney cells, the differentiation capability of all our generated cells further allows investigation of other cells that are also known to be affected by BHD such as skin and the formation of skin fibrofolliculomas. Also, our collaborators for the obtention of BHD patient samples from Germany's only interdisciplinary BHD outpatient clinic recently reported a possible link of BHD to colorectal cancer development (Sattler et al., 2021). Several protocols on intestinal organoids are reported, where unlike kidney organoids, cultivation and propagation is feasible. Future work could include the investigation of FLCN in these cells with the help of our FLCN-S/MAR DNA vectors.

We further provided proof of principle data that overexpression of FLCN does not induce any adverse effects on the cells. In the scope of Birt-Hogg-Dubé syndrome, we presented preliminary data supporting the feasibility of our vector system to act as preventive gene therapy. Since kidney tumorigenesis occurs upon later age and a loss of function second hit mutation, an additional, third functional copy of FLCN provided by our vectors could rescue the cancer progression. To date, delivery of the vectors to patient kidneys presents a major obstacle to move forward towards possible preventive gene therapy for BHD. In vivo delivery protocols for non-viral vectors in mice or rats via hydrodynamic tail injection or polyethyleneimine were reported

(Boletta et al., 1997; Woodard et al., 2018), and there is promising work performed on utilising lipids, polymers, carbon materials, inorganic structures or local ultrasonic waves for delivery of non-viral vectors to different tissues to enable the treatment of patients (Mohammadinejad et al., 2020; Phillips et al., 2008).

In addition to pathway analysis, disease modelling and developmental studies, our vectors are aimed at being used as therapeutic agents avoiding insertional mutagenesis and having little effect on the host cell. Our S/MAR DNA vectors are currently being developed towards clinical trials for CAR-T-Cell therapy laying the path for GMP manufacturing (Bozza et al., 2021). Our established protocols in hiPSCs allow easy adjustment to clinical-grade protocols setting the groundwork for future applications in gene and cell therapy.

Besides advanced cell models and direct gene delivery of gene therapy vectors, differentiation protocols for transplantation purposes are heavily researched because of organ shortage. Autologous organs grown from patients' cells circumvent organ rejection and compatibility problems. There have already been reports of successful transplantation of differentiated renal progenitors and subsequent restoration of kidney function (Imberti et al., 2015). Improvement of such protocols combined with our presented work suggests a possible future application of our vector system for the genetic modification of patients derived iPSCs, subsequent differentiation of genetically restored organs and transplantation (Hasegawa et al., 2019; Phillips et al., 2008). Ultimately, this would render a valuable therapy approach for BHD.

11 References

Aledo-Serrano, A., Gil-Nagel, A., Isla, J., Mingorance, A., Mendez-Hermida, F., & Hernandez-Alcoceba, R. (2021). Gene therapies and COVID-19 vaccines: a necessary discussion in relation with viral vector-based approaches. *Orphanet J Rare Dis, 16*(1), 316. doi:10.1186/s13023-021-01958-3

Alexopoulou, A. N., Couchman, J. R., & Whiteford, J. R. (2008). The CMV early enhancer/chicken beta actin (CAG) promoter can be used to drive transgene expression during the differentiation of murine embryonic stem cells into vascular progenitors. *BMC Cell Biol, 9*, 2. doi:10.1186/1471-2121-9-2

Arope, S., Harraghy, N., Pjanic, M., & Mermod, N. (2013). Molecular characterization of a human matrix attachment region epigenetic regulator. *PLoS One, 8*(11), e79262. doi:10.1371/journal.pone.0079262

Baba, M., Furihata, M., Hong, S. B., Tessarollo, L., Haines, D. C., Southon, E., . . . Schmidt, L. S. (2008). Kidney-targeted Birt-Hogg-Dube gene inactivation in a mouse model: Erk1/2 and Akt-mTOR activation, cell hyperproliferation, and polycystic kidneys. *J Natl Cancer Inst, 100*(2), 140-154. doi:10.1093/jnci/djm288

Baba, M., Hong, S. B., Sharma, N., Warren, M. B., Nickerson, M. L., Iwamatsu, A., . . . Zbar, B. (2006). Folliculin encoded by the BHD gene interacts with a binding protein, FNIP1, and AMPK, and is involved in AMPK and mTOR signaling. *Proc Natl Acad Sci U S A, 103*(42), 15552-15557. doi:10.1073/pnas.0603781103

Baiker, A., Maercker, C., Piechaczek, C., Schmidt, S. B., Bode, J., Benham, C., & Lipps, H. J. (2000). Mitotic stability of an episomal vector containing a human scaffold/matrix-attached region is provided by association with nuclear matrix. *Nat Cell Biol, 2*(3), 182-184. doi:10.1038/35004061

Bardelli, M., Zarate-Perez, F., Agundez, L., Jolinon, N., Michael Linden, R., Escalante, C. R., & Henckaerts, E. (2017). Analysis of Replicative Intermediates of Adeno-associated Virus through Hirt Extraction and Southern Blotting. *Bio Protoc, 7*(9). doi:10.21769/BioProtoc.2271

Battiwalla, M., & Hematti, P. (2009). Mesenchymal stem cells in hematopoietic stem cell transplantation. *Cytotherapy, 11*(5), 503-515. doi:10.1080/14653240903193806

Betschinger, J., Nichols, J., Dietmann, S., Corrin, P. D., Paddison, P. J., & Smith, A. (2013). Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell, 153*(2), 335-347. doi:10.1016/j.cell.2013.03.012

Bharadwaj, S., Liu, G., Shi, Y., Wu, R., Yang, B., He, T., . . . Zhang, Y. (2013). Multipotential differentiation of human urine-derived stem cells: potential for therapeutic applications in urology. *Stem Cells, 31*(9), 1840-1856. doi:10.1002/stem.1424

Bigger, B. W., Tolmachov, O., Collombet, J. M., Fragkos, M., Palaszewski, I., & Coutelle, C. (2001). An araC-controlled bacterial cre expression system to produce DNA minicircle vectors for nuclear and mitochondrial gene therapy. *J Biol Chem, 276*(25), 23018-23027. doi:10.1074/jbc.M010873200

Bochkov, Y. A., & Palmenberg, A. C. (2006). Translational efficiency of EMCV IRES in bicistronic vectors is dependent upon IRES sequence and gene location. *Biotechniques, 41*(3), 283-284, 286, 288 passim. doi:10.2144/000112243

Boletta, A., Benigni, A., Lutz, J., Remuzzi, G., Soria, M. R., & Monaco, L. (1997). Nonviral gene delivery to the rat kidney with polyethylenimine. *Hum Gene Ther, 8*(10), 1243-1251. doi:10.1089/hum.1997.8.10-1243

Bozza, M. (2017). *The development of a novel S/MAR DNA vector platform for the stable, persistent and safe Genetic Engineering of Dividing Cells.* (Doctor of Natural Sciences Dissertation). Ruperto-Carola University of Heidelberg, Germany,

Bozza, M., De Roia, A., Correia, M. P., Berger, A., Tuch, A., Schmidt, A., . . . Harbottle, R. P. (2021). A nonviral, nonintegrating DNA nanovector platform for the safe, rapid, and persistent manufacture of recombinant T cells. *Sci Adv, 7*(16). doi:10.1126/sciadv.abf1333

Bozza, M., Green, E. W., Espinet, E., De Roia, A., Klein, C., Vogel, V., . . . Harbottle, R. P. (2020). Novel Non-integrating DNA Nano-S/MAR Vectors Restore Gene Function in Isogenic Patient-Derived Pancreatic Tumor Models. *Mol Ther Methods Clin Dev, 17*, 957-968. doi:10.1016/j.omtm.2020.04.017

Bredenkamp, N., Yang, J., Clarke, J., Stirparo, G. G., von Meyenn, F., Dietmann, S., . . . Guo, G. (2019). Wnt Inhibition Facilitates RNA-Mediated Reprogramming of Human Somatic Cells to Naive Pluripotency. *Stem Cell Reports, 13*(6), 1083-1098. doi:10.1016/j.stemcr.2019.10.009

Byrne, S. M., Mali, P., & Church, G. M. (2014). Genome editing in human stem cells. *Methods Enzymol, 546*, 119-138. doi:10.1016/B978-0-12-801185-0.00006-4

Caplan, A. I. (2010). Mesenchymal Stem Cells: The Past, the Present, the Future. *Cartilage, 1*(1), 6-9. doi:10.1177/1947603509354992

Caplan, A. I. (2015). Are All Adult Stem Cells The Same? *Regenerative Engineering and Translational Medicine, 1*(1-4), 4-10. doi:10.1007/s40883-015-0001-4

Cash, T. P., Gruber, J. J., Hartman, T. R., Henske, E. P., & Simon, M. C. (2011). Loss of the Birt-Hogg-Dube tumor suppressor results in apoptotic resistance due to aberrant TGFbeta-mediated transcription. *Oncogene, 30*(22), 2534-2546. doi:10.1038/onc.2010.628

Chan, Y. K., Wang, S. K., Chu, C. J., Copland, D. A., Letizia, A. J., Costa Verdera, H., . . . Church, G. M. (2021). Engineering adeno-associated viral vectors to evade innate immune and inflammatory responses. *Sci Transl Med, 13*(580). doi:10.1126/scitranslmed.abd3438

Chen, B., & Altman, R. B. (2017). Opportunities for developing therapies for rare genetic diseases: focus on gain-of-function and allostery. *Orphanet J Rare Dis, 12*(1), 61. doi:10.1186/s13023-017-0614-4

Chen, L., Li, L., Xing, F., Peng, J., Peng, K., Wang, Y., & Xiang, Z. (2018). Human Urine-Derived Stem Cells: Potential for Cell-Based Therapy of Cartilage Defects. *Stem Cells Int, 2018*, 4686259. doi:10.1155/2018/4686259

Choi, J. Y., Chun, S. Y., Ha, Y. S., Kim, D. H., Kim, J., Song, P. H., . . . Kwon, T. G. (2017). Potency of Human Urine-Derived Stem Cells for Renal Lineage Differentiation. *Tissue Eng Regen Med, 14*(6), 775-785. doi:10.1007/s13770-017-0081-y

References

Concordet, J. P., & Haeussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res, 46*(W1), W242-W245. doi:10.1093/nar/gky354

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., . . . Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science, 339*(6121), 819-823. doi:10.1126/science.1231143

Cyranoski, D. (2020). What CRISPR-baby prison sentences mean for research. *Nature, 577*(7789), 154-155. doi:10.1038/d41586-020-00001-y

Darquet, A. M., Cameron, B., Wils, P., Scherman, D., & Crouzet, J. (1997). A new DNA vehicle for nonviral gene delivery: supercoiled minicircle. *Gene Ther, 4*(12), 1341-1349. doi:10.1038/sj.gt.3300540

Drozd, A. M., Walczak, M. P., Piaskowski, S., Stoczynska-Fidelus, E., Rieske, P., & Grzela, D. P. (2015). Generation of human iPSCs from cells of fibroblastic and epithelial origin by means of the oriP/EBNA-1 episomal reprogramming system. *Stem Cell Res Ther, 6*, 122. doi:10.1186/s13287-015-0112-3

Duesberg, P., Stindl, R., & Hehlmann, R. (2001). Origin of multidrug resistance in cells with and without multidrug resistance genes: chromosome reassortments catalyzed by aneuploidy. *Proc Natl Acad Sci U S A, 98*(20), 11283-11288. doi:10.1073/pnas.201398998

Dunbar, C. E., High, K. A., Joung, J. K., Kohn, D. B., Ozawa, K., & Sadelain, M. (2018). Gene therapy comes of age. *Science, 359*(6372). doi:10.1126/science.aan4672

Falzarano, M. S., & Ferlini, A. (2019). Urinary Stem Cells as Tools to Study Genetic Disease: Overview of the Literature. *J Clin Med, 8*(5). doi:10.3390/jcm8050627

First-Ever CAR T-cell Therapy Approved in U.S. (2017). *Cancer Discov, 7*(10), OF1. doi:10.1158/2159-8290.CD-NB2017-126

Forbes, T. A., Howden, S. E., Lawlor, K., Phipson, B., Maksimovic, J., Hale, L., . . . Little, M. H. (2018). Patient-iPSC-Derived Kidney Organoids Show Functional Validation of a Ciliopathic Renal Phenotype and Reveal Underlying Pathogenetic Mechanisms. *Am J Hum Genet, 102*(5), 816-831. doi:10.1016/j.ajhg.2018.03.014

Freedman, B. S. (2015). Modeling Kidney Disease with iPS Cells. *Biomark Insights, 10*(Suppl 1), 153-169. doi:10.4137/BMI.S20054

Fus-Kujawa, A., Mendrek, B., Trybus, A., Bajdak-Rusinek, K., Stepien, K. L., & Sieron, A. L. (2021). Potential of Induced Pluripotent Stem Cells for Use in Gene Therapy: History, Molecular Bases, and Medical Perspectives. *Biomolecules, 11*(5). doi:10.3390/biom11050699

Garreta, E., Kamm, R. D., Chuva de Sousa Lopes, S. M., Lancaster, M. A., Weiss, R., Trepat, X., . . . Montserrat, N. (2021). Rethinking organoid technology through bioengineering. *Nat Mater, 20*(2), 145-155. doi:10.1038/s41563-020-00804-4

Garreta, E., Montserrat, N., & Belmonte, J. C. I. (2018). Kidney organoids for disease modeling. *Oncotarget, 9*(16), 12552-12553. doi:10.18632/oncotarget.24438

Garreta, E., Prado, P., Tarantino, C., Oria, R., Fanlo, L., Marti, E., . . . Montserrat, N. (2019). Fine tuning the extracellular environment accelerates the derivation of kidney organoids from human pluripotent stem cells. *Nat Mater, 18*(4), 397-405. doi:10.1038/s41563-019-0287-6

Gasiunas, G., & Siksnys, V. (2013). RNA-dependent DNA endonuclease Cas9 of the CRISPR system: Holy Grail of genome editing? *Trends Microbiol, 21*(11), 562-567. doi:10.1016/j.tim.2013.09.001

Glover, D. J., Lipps, H. J., & Jans, D. A. (2005). Towards safe, non-viral therapeutic gene expression in humans. *Nat Rev Genet, 6*(4), 299-310. doi:10.1038/nrg1577

Gonzalez, F., Boue, S., & Izpisua Belmonte, J. C. (2011). Methods for making induced pluripotent stem cells: reprogramming a la carte. *Nat Rev Genet, 12*(4), 231-242. doi:10.1038/nrg2937

Gupta, N., Seyama, K., & McCormack, F. X. (2013). Pulmonary manifestations of Birt-Hogg-Dube syndrome. *Fam Cancer, 12*(3), 387-396. doi:10.1007/s10689-013-9660-9

Gurdon, J. B. (1962). Adult frogs derived from the nuclei of single somatic cells. *Dev Biol, 4*, 256-273. doi:10.1016/0012-1606(62)90043-x

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M. P., Wulffraat, N., Leboulch, P., . . . Cavazzana-Calvo, M. (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science, 302*(5644), 415-419. doi:10.1126/science.1088547

Hagedorn, C., Baiker, A., Postberg, J., Ehrhardt, A., & Lipps, H. J. (2012). Rescue of S/MAR-containing nonviral episomal expression vectors. *Cold Spring Harb Protoc, 2012*(6), 709-711. doi:10.1101/pdb.prot069518

Hagedorn, C., Gogol-Doring, A., Schreiber, S., Epplen, J. T., & Lipps, H. J. (2017). Genome-wide profiling of S/MAR-based replicon contact sites. *Nucleic Acids Res, 45*(13), 7841-7854. doi:10.1093/nar/gkx522

Hagedorn, C., Lipps, H. J., & Rupprecht, S. (2010). The epigenetic regulation of autonomous replicons. *Biomol Concepts, 1*(1), 17-30. doi:10.1515/bmc.2010.009

Hagedorn, C., Schnodt-Fuchs, M., Boehme, P., Abdelrazik, H., Lipps, H. J., & Buning, H. (2017). S/MAR Element Facilitates Episomal Long-Term Persistence of Adeno-Associated Virus Vector Genomes in Proliferating Cells. *Hum Gene Ther, 28*(12), 1169-1179. doi:10.1089/hum.2017.025

Hagedorn, C., Wong, S. P., Harbottle, R., & Lipps, H. J. (2011). Scaffold/matrix attached region-based nonviral episomal vectors. *Hum Gene Ther, 22*(8), 915-923. doi:10.1089/hum.2011.084

Hasegawa, S., Tanaka, T., & Nangaku, M. (2019). Recent advances in renal regeneration. *F1000Res, 8*. doi:10.12688/f1000research.17127.1

Hashizume, T., & Shimizu, N. (2007). Dissection of mammalian replicators by a novel plasmid stability assay. *J Cell Biochem, 101*(3), 552-565. doi:10.1002/jcb.21210

Hasumi, Y., Baba, M., Ajima, R., Hasumi, H., Valera, V. A., Klein, M. E., . . . Linehan, W. M. (2009). Homozygous loss of BHD causes early embryonic lethality and kidney tumor development with activation of mTORC1 and mTORC2. *Proc Natl Acad Sci U S A, 106*(44), 18722-18727. doi:10.1073/pnas.0908853106

He, Q., Wan, C., & Li, G. (2007). Concise review: multipotent mesenchymal stromal cells in blood. *Stem Cells, 25*(1), 69-77. doi:10.1634/stemcells.2006-0335

Herbst, F., Ball, C. R., Tuorto, F., Nowrouzi, A., Wang, W., Zavidij, O., . . . Glimm, H. (2012). Extensive methylation of promoter sequences silences lentiviral transgene expression during stem cell differentiation in vivo. *Mol Ther, 20*(5), 1014-1021. doi:10.1038/mt.2012.46

References

Hofer, M., & Lutolf, M. P. (2021). Engineering organoids. *Nat Rev Mater*, 1-19. doi:10.1038/s41578-021-00279-y

Hombach-Klonisch, S., Panigrahi, S., Rashedi, I., Seifert, A., Alberti, E., Pocar, P., . . . Los, M. (2008). Adult stem cells and their trans-differentiation potential-- perspectives and therapeutic applications. *J Mol Med (Berl), 86*(12), 1301-1314. doi:10.1007/s00109-008-0383-6

Hong, S. B., Oh, H., Valera, V. A., Baba, M., Schmidt, L. S., & Linehan, W. M. (2010). Inactivation of the FLCN tumor suppressor gene induces TFE3 transcriptional activity by increasing its nuclear localization. *PLoS One, 5*(12), e15793. doi:10.1371/journal.pone.0015793

Hoogduijn, M. J., Montserrat, N., van der Laan, L. J. W., Dazzi, F., Perico, N., Kastrup, J., . . . Reinders, M. E. J. (2020). The emergence of regenerative medicine in organ transplantation: 1st European Cell Therapy and Organ Regeneration Section meeting. *Transpl Int*. doi:10.1111/tri.13608

Hudecek, M., & Ivics, Z. (2018). Non-viral therapeutic cell engineering with the Sleeping Beauty transposon system. *Curr Opin Genet Dev, 52*, 100-108. doi:10.1016/j.gde.2018.06.003

Hudon, V., Sabourin, S., Dydensborg, A. B., Kottis, V., Ghazi, A., Paquet, M., . . . Pause, A. (2010). Renal tumour suppressor function of the Birt-Hogg-Dube syndrome gene product folliculin. *J Med Genet, 47*(3), 182-189. doi:10.1136/jmg.2009.072009

Ichida, J. K., Blanchard, J., Lam, K., Son, E. Y., Chung, J. E., Egli, D., . . . Eggan, K. (2009). A small-molecule inhibitor of tgf-Beta signaling replaces sox2 in reprogramming by inducing nanog. *Cell Stem Cell, 5*(5), 491-503. doi:10.1016/j.stem.2009.09.012

Imberti, B., Tomasoni, S., Ciampi, O., Pezzotta, A., Derosas, M., Xinaris, C., . . . Morigi, M. (2015). Renal progenitors derived from human iPSCs engraft and restore function in a mouse model of acute kidney injury. *Sci Rep, 5*, 8826. doi:10.1038/srep08826

Jenke, A. C., Stehle, I. M., Herrmann, F., Eisenberger, T., Baiker, A., Bode, J., . . . Lipps, H. J. (2004). Nuclear scaffold/matrix attached region modules linked to a transcription unit are sufficient for replication and maintenance of a mammalian episome. *Proc Natl Acad Sci U S A, 101*(31), 11322-11327. doi:10.1073/pnas.0401355101

Jung, S., Appleton, E., Ali, M., Church, G. M., & Del Sol, A. (2021). A computer-guided design tool to increase the efficiency of cellular conversions. *Nat Commun, 12*(1), 1659. doi:10.1038/s41467-021-21801-4

Kaeppel, C., Beattie, S. G., Fronza, R., van Logtenstein, R., Salmon, F., Schmidt, S., . . . Schmidt, M. (2013). A largely random AAV integration profile after LPLD gene therapy. *Nat Med, 19*(7), 889-891. doi:10.1038/nm.3230

Kawai, A., Kobayashi, T., & Hino, O. (2013). Folliculin regulates cyclin D1 expression through cis-acting elements in the 3' untranslated region of cyclin D1 mRNA. *Int J Oncol, 42*(5), 1597-1604. doi:10.3892/ijo.2013.1862

Kebriaei, P., Izsvak, Z., Narayanavari, S. A., Singh, H., & Ivics, Z. (2017). Gene Therapy with the Sleeping Beauty Transposon System. *Trends Genet, 33*(11), 852-870. doi:10.1016/j.tig.2017.08.008

Khoo, S. K., Kahnoski, K., Sugimura, J., Petillo, D., Chen, J., Shockley, K., . . . Teh, B. T. (2003). Inactivation of BHD in sporadic renal tumors. *Cancer Res, 63*(15), 4583-4587. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12907635

Kim, E. J., Kang, K. H., & Ju, J. H. (2017). CRISPR-Cas9: a promising tool for gene editing on induced pluripotent stem cells. *Korean J Intern Med, 32*(1), 42-61. doi:10.3904/kjim.2016.198

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., . . . Daley, G. Q. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature, 467*(7313), 285-290. doi:10.1038/nature09342

Kremer, E. J. (2020). Pros and Cons of Adenovirus-Based SARS-CoV-2 Vaccines. *Mol Ther, 28*(11), 2303-2304. doi:10.1016/j.ymthe.2020.10.002

Kwaks, T. H., Barnett, P., Hemrika, W., Siersma, T., Sewalt, R. G., Satijn, D. P., . . . Otte, A. P. (2003). Identification of anti-repressor elements that confer high and stable protein production in mammalian cells. *Nat Biotechnol, 21*(5), 553-558. doi:10.1038/nbt814

Lang, R., Liu, G., Shi, Y., Bharadwaj, S., Leng, X., Zhou, X., . . . Zhang, Y. (2013). Self-renewal and differentiation capacity of urine-derived stem cells after urine preservation for 24 hours. *PLoS One, 8*(1), e53980. doi:10.1371/journal.pone.0053980

Lanza, A. M., Kim, D. S., & Alper, H. S. (2013). Evaluating the influence of selection markers on obtaining selected pools and stable cell lines in human cells. *Biotechnol J, 8*(7), 811-821. doi:10.1002/biot.201200364

Larochelle, A., & Dunbar, C. E. (2013). Hematopoietic stem cell gene therapy:assessing the relevance of preclinical models. *Semin Hematol, 50*(2), 101-130. doi:10.1053/j.seminhematol.2013.03.025

Leppek, K., Das, R., & Barna, M. (2018). Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol, 19*(3), 158-174. doi:10.1038/nrm.2017.103

Leter, E. M., Koopmans, A. K., Gille, J. J., van Os, T. A., Vittoz, G. G., David, E. F., . . . Menko, F. H. (2008). Birt-Hogg-Dube syndrome: clinical and genetic studies of 20 families. *J Invest Dermatol, 128*(1), 45-49. doi:10.1038/sj.jid.5700959

Li, X., Nadauld, L., Ootani, A., Corney, D. C., Pai, R. K., Gevaert, O., . . . Kuo, C. J. (2014). Oncogenic transformation of diverse gastrointestinal tissues in primary organoid culture. *Nat Med, 20*(7), 769-777. doi:10.1038/nm.3585

Liew, C. G., Draper, J. S., Walsh, J., Moore, H., & Andrews, P. W. (2007). Transient and stable transgene expression in human embryonic stem cells. *Stem Cells, 25*(6), 1521-1528. doi:10.1634/stemcells.2006-0634

Little, M. H., & Combes, A. N. (2019). Kidney organoids: accurate models or fortunate accidents. *Genes Dev, 33*(19-20), 1319-1345. doi:10.1101/gad.329573.119

Little, M. H., & Quinlan, C. (2020). Advances in our understanding of genetic kidney disease using kidney organoids. *Pediatr Nephrol, 35*(6), 915-926. doi:10.1007/s00467-019-04259-x

Liu, J., Jones, K. L., Sumer, H., & Verma, P. J. (2009). Stable transgene expression in human embryonic stem cells after simple chemical transfection. *Mol Reprod Dev, 76*(6), 580-586. doi:10.1002/mrd.20983

References

Liu, Z., Chen, O., Wall, J. B. J., Zheng, M., Zhou, Y., Wang, L., . . . Liu, J. (2017). Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci Rep, 7*(1), 2193. doi:10.1038/s41598-017-02460-2

Luijten, M. N., Basten, S. G., Claessens, T., Vernooij, M., Scott, C. L., Janssen, R., . . . van Steensel, M. A. (2013). Birt-Hogg-Dube syndrome is a novel ciliopathy. *Hum Mol Genet, 22*(21), 4383-4397. doi:10.1093/hmg/ddt288

Luke, J., Carnes, A. E., Hodgson, C. P., & Williams, J. A. (2009). Improved antibiotic-free DNA vaccine vectors utilizing a novel RNA based plasmid selection system. *Vaccine, 27*(46), 6454-6459. doi:10.1016/j.vaccine.2009.06.017

Luke, J. M., Carnes, A. E., & Williams, J. A. (2014). Development of antibiotic-free selection system for safer DNA vaccination. *Methods Mol Biol, 1143*, 91-111. doi:10.1007/978-1-4939-0410-5_6

Mack, A. A., Kroboth, S., Rajesh, D., & Wang, W. B. (2011). Generation of induced pluripotent stem cells from CD34+ cells across blood drawn from multiple donors with non-integrating episomal vectors. *PLoS One, 6*(11), e27956. doi:10.1371/journal.pone.0027956

Manaph, N. P. A., Al-Hawwas, M., Bobrovskaya, L., Coates, P. T., & Zhou, X. F. (2018). Correction to: Urine-derived cells for human cell therapy. *Stem Cell Res Ther, 9*(1), 222. doi:10.1186/s13287-018-0974-2

Manzini, S., Vargiolu, A., Stehle, I. M., Bacci, M. L., Cerrito, M. G., Giovannoni, R., . . . Lavitrano, M. (2006). Genetically modified pigs produced with a nonviral episomal vector. *Proc Natl Acad Sci U S A, 103*(47), 17672-17677. doi:10.1073/pnas.0604938103

Mathieu, J., Detraux, D., Kuppers, D., Wang, Y., Cavanaugh, C., Sidhu, S., . . . Ruohola-Baker, H. (2019). Folliculin regulates mTORC1/2 and WNT pathways in early human pluripotency. *Nat Commun, 10*(1), 632. doi:10.1038/s41467-018-08020-0

Mayrhofer, P., Schleef, M., & Jechlinger, W. (2009). Use of minicircle plasmids for gene therapy. *Methods Mol Biol, 542*, 87-104. doi:10.1007/978-1-59745-561-9_4

McLachlan, G., Stevenson, B. J., Davidson, D. J., & Porteous, D. J. (2000). Bacterial DNA is implicated in the inflammatory response to delivery of DNA/DOTAP to mouse lungs. *Gene Ther, 7*(5), 384-392. doi:10.1038/sj.gt.3301097

Miller, A. D. (1992). Human gene therapy comes of age. *Nature, 357*(6378), 455-460. doi:10.1038/357455a0

Mindaye, S. T., Ra, M., Lo Surdo, J., Bauer, S. R., & Alterman, M. A. (2013). Improved proteomic profiling of the cell surface of culture-expanded human bone marrow multipotent stromal cells. *J Proteomics, 78*, 1-14. doi:10.1016/j.jprot.2012.10.028

Mirkovitch, J., Mirault, M. E., & Laemmli, U. K. (1984). Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell, 39*(1), 223-232. doi:10.1016/0092-8674(84)90208-3

Miyoshi, T., Hiratsuka, K., Saiz, E. G., & Morizane, R. (2020). Kidney organoids in translational medicine: Disease modeling and regenerative medicine. *Dev Dyn, 249*(1), 34-45. doi:10.1002/dvdy.22

Mohammadinejad, R., Dehshahri, A., Sagar Madamsetty, V., Zahmatkeshan, M., Tavakol, S., Makvandi, P., . . . Zarrabi, A. (2020). In vivo gene delivery mediated

by non-viral vectors for cancer therapy. *J Control Release, 325*, 249-275. doi:10.1016/j.jconrel.2020.06.038

Morizane, R., & Bonventre, J. V. (2017). Kidney Organoids: A Translational Journey. *Trends Mol Med, 23*(3), 246-263. doi:10.1016/j.molmed.2017.01.001

Mulder, J., Sharmin, S., Chow, T., Rodrigues, D. C., Hildebrandt, M. R., D'Cruz, R., . . . Rosenblum, N. D. (2020). Generation of infant- and pediatric-derived urinary induced pluripotent stem cells competent to form kidney organoids. *Pediatr Res, 87*(4), 647-655. doi:10.1038/s41390-019-0618-y

Muller, L. U., Milsom, M. D., Harris, C. E., Vyas, R., Brumme, K. M., Parmar, K., . . . Williams, D. A. (2012). Overcoming reprogramming resistance of Fanconi anemia cells. *Blood, 119*(23), 5449-5457. doi:10.1182/blood-2012-02-408674

Musunuru, K. (2013). Genome editing of human pluripotent stem cells to generate human cellular disease models. *Dis Model Mech, 6*(4), 896-904. doi:10.1242/dmm.012054

Nanbo, A., Sugden, A., & Sugden, B. (2007). The coupling of synthesis and partitioning of EBV's plasmid replicon is revealed in live cells. *EMBO J, 26*(19), 4252-4262. doi:10.1038/sj.emboj.7601853

Nassor, F., Jarray, R., Biard, D. S. F., Maiza, A., Papy-Garcia, D., Pavoni, S., . . . Yates, F. (2020). Long Term Gene Expression in Human Induced Pluripotent Stem Cells and Cerebral Organoids to Model a Neurodegenerative Disease. *Front Cell Neurosci, 14*, 14. doi:10.3389/fncel.2020.00014

Ng, A. H. M., Khoshakhlagh, P., Rojo Arias, J. E., Pasquini, G., Wang, K., Swiersy, A., . . . Church, G. M. (2020). A comprehensive library of human transcription factors for cell fate engineering. *Nat Biotechnol*. doi:10.1038/s41587-020-0742-6

Nickerson, M. L., Warren, M. B., Toro, J. R., Matrosova, V., Glenn, G., Turner, M. L., . . . Schmidt, L. S. (2002). Mutations in a novel gene lead to kidney tumors, lung wall defects, and benign tumors of the hair follicle in patients with the Birt-Hogg-Dube syndrome. *Cancer Cell, 2*(2), 157-164. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12204536

Nishimura, K., Sano, M., Ohtaka, M., Furuta, B., Umemura, Y., Nakajima, Y., . . . Nakanishi, M. (2011). Development of defective and persistent Sendai virus vector: a unique gene delivery/expression system ideal for cell reprogramming. *J Biol Chem, 286*(6), 4760-4771. doi:10.1074/jbc.M110.183780

O'Neill, A. C., & Ricardo, S. D. (2013). Human kidney cell reprogramming: applications for disease modeling and personalized medicine. *J Am Soc Nephrol, 24*(9), 1347-1356. doi:10.1681/ASN.2012121199

Okita, K., Matsumura, Y., Sato, Y., Okada, A., Morizane, A., Okamoto, S., . . . Yamanaka, S. (2011). A more efficient method to generate integration-free human iPS cells. *Nat Methods, 8*(5), 409-412. doi:10.1038/nmeth.1591

Oliveira Arcolino, F., Tort Piella, A., Papadimitriou, E., Bussolati, B., Antonie, D. J., Murray, P., . . . Levtchenko, E. (2015). Human Urine as a Noninvasive Source of Kidney Cells. *Stem Cells Int, 2015*, 362562. doi:10.1155/2015/362562

Papp, B., & Plath, K. (2011). Reprogramming to pluripotency: stepwise resetting of the epigenetic landscape. *Cell Res, 21*(3), 486-501. doi:10.1038/cr.2011.28

Pavathuparambil Abdul Manaph, N., Al-Hawwas, M., Bobrovskaya, L., Coates, P. T., & Zhou, X. F. (2018). Urine-derived cells for human cell therapy. *Stem Cell Res Ther, 9*(1), 189. doi:10.1186/s13287-018-0932-z

Pavlovich, C. P., Grubb, R. L., 3rd, Hurley, K., Glenn, G. M., Toro, J., Schmidt, L. S., . . . Linehan, W. M. (2005). Evaluation and management of renal tumors in the Birt-Hogg-Dube syndrome. *J Urol, 173*(5), 1482-1486. doi:10.1097/01.ju.0000154629.45832.30

Peng, G. Y., Lin, Y., Li, J. J., Wang, Y., Huang, H. Y., & Shen, Z. Y. (2019). The Application of Induced Pluripotent Stem Cells in Pathogenesis Study and Gene Therapy for Vascular Disorders: Current Progress and Future Challenges. *Stem Cells Int, 2019*, 9613258. doi:10.1155/2019/9613258

Phillips, B., Giannoukakis, N., & Trucco, M. (2008). Renal diseases as targets of gene therapy. *Contrib Nephrol, 159*, 1-12. doi:10.1159/000125552

Piechaczek, C., Fetzer, C., Baiker, A., Bode, J., & Lipps, H. J. (1999). A vector based on the SV40 origin of replication and chromosomal S/MARs replicates episomally in CHO cells. *Nucleic Acids Res, 27*(2), 426-428. doi:10.1093/nar/27.2.426

Poleganov, M. A., Eminli, S., Beissert, T., Herz, S., Moon, J. I., Goldmann, J., . . . Sahin, U. (2015). Efficient Reprogramming of Human Fibroblasts and Blood-Derived Endothelial Progenitor Cells Using Nonmodified RNA for Reprogramming and Immune Evasion. *Hum Gene Ther, 26*(11), 751-766. doi:10.1089/hum.2015.045

Przepiorski, A., Sander, V., Tran, T., Hollywood, J. A., Sorrenson, B., Shih, J. H., . . . Davidson, A. J. (2018). A Simple Bioreactor-Based Method to Generate Kidney Organoids from Pluripotent Stem Cells. *Stem Cell Reports, 11*(2), 470-484. doi:10.1016/j.stemcr.2018.06.018

Rahman, M. S., Wruck, W., Spitzhorn, L. S., Nguyen, L., Bohndorf, M., Martins, S., . . . Adjaye, J. (2020). The FGF, TGFbeta and WNT axis Modulate Self-renewal of Human SIX2(+) Urine Derived Renal Progenitor Cells. *Sci Rep, 10*(1), 739. doi:10.1038/s41598-020-57723-2

Roig Merino, A. (2018). *Genetic Modification of Stem Cells Utilizing S/MAR DNA Vectors.* (Doctor of Natural Sciences Dissertation). Ruperto-Carola University of Heidelberg, Germany,

Roig Merino, A., Urban, M., Bozza, M., Peterson, J., Bullen, L., Bücher-Schäff, M., . . . Harbottle, R. P. (in press). An episomal DNA vector for the persistent genetic modification of pluripotent stem cells and their differentiated progeny. *Stem Cell Reports*.

Rupprecht, S., Hagedorn, C., Seruggia, D., Magnusson, T., Wagner, E., Ogris, M., & Lipps, H. J. (2010). Controlled removal of a nonviral episomal vector from transfected cells. *Gene, 466*(1-2), 36-42. doi:10.1016/j.gene.2010.07.001

Sattler, E. C., Syunyaeva, Z., Reithmair, M., Dempke, W., & Steinlein, O. K. (2021). Colorectal cancer risk in families with Birt-Hogg-Dube syndrome increased. *Eur J Cancer, 151*, 168-174. doi:10.1016/j.ejca.2021.04.013

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., . . . Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods, 9*(7), 676-682. doi:10.1038/nmeth.2019

Schmidt, L. S. (2013). Birt-Hogg-Dube syndrome: from gene discovery to molecularly targeted therapies. *Fam Cancer, 12*(3), 357-364. doi:10.1007/s10689-012-9574-y

Schmidt, L. S., & Linehan, W. M. (2018). FLCN: The causative gene for Birt-Hogg-Dube syndrome. *Gene, 640*, 28-42. doi:10.1016/j.gene.2017.09.044

Seow, Y., & Wood, M. J. (2009). Biological gene delivery vehicles: beyond viral vectors. *Mol Ther, 17*(5), 767-777. doi:10.1038/mt.2009.41

Shi, T., & Cheung, M. (2021). Urine-derived induced pluripotent/neural stem cells for modeling neurological diseases. *Cell Biosci, 11*(1), 85. doi:10.1186/s13578-021-00594-5

Siller, R., Naumovska, E., Mathapati, S., Lycke, M., Greenhough, S., & Sullivan, G. J. (2016). Development of a rapid screen for the endodermal differentiation potential of human pluripotent stem cell lines. *Sci Rep, 6*, 37178. doi:10.1038/srep37178

Song, M., & Ramakrishna, S. (2018). Genome Editing in Stem Cells for Disease Therapeutics. *Mol Biotechnol, 60*(4), 329-338. doi:10.1007/s12033-018-0072-9

Stenler, S., Blomberg, P., & Smith, C. I. (2014). Safety and efficacy of DNA vaccines: plasmids vs. minicircles. *Hum Vaccin Immunother, 10*(5), 1306-1308. doi:10.4161/hv.28077

Sun, Q. L., Zhao, C. P., Chen, S. N., Wang, L., & Wang, T. Y. (2016). Molecular characterization of a human matrix attachment region that improves transgene expression in CHO cells. *Gene, 582*(2), 168-172. doi:10.1016/j.gene.2016.02.009

Szablowska-Gadomska, I., Gorska, A., & Malecki, M. (2013). Induced pluripotent stem cells (iPSc) for gene therapy. *Med Wieku Rozwoj, 17*(3), 191-195. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/24296442

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell, 131*(5), 861-872. doi:10.1016/j.cell.2007.11.019

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. doi:10.1016/j.cell.2006.07.024

Takasato, M., & Little, M. H. (2015). The origin of the mammalian kidney: implications for recreating the kidney in vitro. *Development, 142*(11), 1937-1947. doi:10.1242/dev.104802

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., & Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science, 282*(5391), 1145-1147. doi:10.1126/science.282.5391.1145

Thyagarajan, B., Scheyhing, K., Xue, H., Fontes, A., Chesnut, J., Rao, M., & Lakshmipathy, U. (2009). A single EBV-based vector for stable episomal maintenance and expression of GFP in human embryonic stem cells. *Regen Med, 4*(2), 239-250. doi:10.2217/17460751.4.2.239

Togashi, Y., Kobayashi, T., Momose, S., Ueda, M., Okimoto, K., & Hino, O. (2006). Transgenic rescue from embryonic lethality and renal carcinogenesis in the Nihon rat model by introduction of a wild-type Bhd gene. *Oncogene, 25*(20), 2885-2889. doi:10.1038/sj.onc.1209317

Toro, J. R., Glenn, G., Duray, P., Darling, T., Weirich, G., Zbar, B., . . . Turner, M. L. (1999). Birt-Hogg-Dube syndrome: a novel marker of kidney neoplasia. *Arch Dermatol, 135*(10), 1195-1202. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10522666

References

Toro, J. R., Wei, M. H., Glenn, G. M., Weinreich, M., Toure, O., Vocke, C., . . . Linehan, W. M. (2008). BHD mutations, clinical and molecular genetic investigations of Birt-Hogg-Dube syndrome: a new series of 50 families and a review of published reports. *J Med Genet, 45*(6), 321-331. doi:10.1136/jmg.2007.054304

van Steensel, M. A., Verstraeten, V. L., Frank, J., Kelleners-Smeets, N. W., Poblete-Gutierrez, P., Marcus-Soekarman, D., . . . van Geel, M. (2007). Novel mutations in the BHD gene and absence of loss of heterozygosity in fibrofolliculomas of Birt-Hogg-Dube patients. *J Invest Dermatol, 127*(3), 588-593. doi:10.1038/sj.jid.5700592

Varghese, D. S., Parween, S., Ardah, M. T., Emerald, B. S., & Ansari, S. A. (2017). Effects of Aminoglycoside Antibiotics on Human Embryonic Stem Cell Viability during Differentiation In Vitro. *Stem Cells Int, 2017*, 2451927. doi:10.1155/2017/2451927

Velychko, S., Adachi, K., Kim, K. P., Hou, Y., MacCarthy, C. M., Wu, G., & Scholer, H. R. (2019). Excluding Oct4 from Yamanaka Cocktail Unleashes the Developmental Potential of iPSCs. *Cell Stem Cell, 25*(6), 737-753 e734. doi:10.1016/j.stem.2019.10.002

Viecelli, H. M., Harbottle, R. P., Wong, S. P., Schlegel, A., Chuah, M. K., VandenDriessche, T., . . . Thony, B. (2014). Treatment of phenylketonuria using minicircle-based naked-DNA gene transfer to murine liver. *Hepatology, 60*(3), 1035-1043. doi:10.1002/hep.27104

Villegas, F., Lehalle, D., Mayer, D., Rittirsch, M., Stadler, M. B., Zinner, M., . . . Betschinger, J. (2019). Lysosomal Signaling Licenses Embryonic Stem Cell Differentiation via Inactivation of Tfe3. *Cell Stem Cell, 24*(2), 257-270 e258. doi:10.1016/j.stem.2018.11.021

Vocke, C. D., Yang, Y., Pavlovich, C. P., Schmidt, L. S., Nickerson, M. L., Torres-Cabala, C. A., . . . Linehan, W. M. (2005). High frequency of somatic frameshift BHD gene mutations in Birt-Hogg-Dube-associated renal tumors. *J Natl Cancer Inst, 97*(12), 931-935. doi:10.1093/jnci/dji154

Wagers, A. J., & Weissman, I. L. (2004). Plasticity of adult stem cells. *Cell, 116*(5), 639-648. doi:10.1016/s0092-8674(04)00208-9

Wang, X., & Dai, J. (2010). Concise review: isoforms of OCT4 contribute to the confusing diversity in stem cell biology. *Stem Cells, 28*(5), 885-893. doi:10.1002/stem.419

Wang, Y., Wang, F., Wang, R., Zhao, P., & Xia, Q. (2015). 2A self-cleaving peptide-based multi-gene expression system in the silkworm Bombyx mori. *Sci Rep, 5*, 16273. doi:10.1038/srep16273

Warlich, E., Kuehle, J., Cantz, T., Brugman, M. H., Maetzig, T., Galla, M., . . . Schambach, A. (2011). Lentiviral vector design and imaging approaches to visualize the early stages of cellular reprogramming. *Mol Ther, 19*(4), 782-789. doi:10.1038/mt.2010.314

Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., . . . Sasai, Y. (2007). A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat Biotechnol, 25*(6), 681-686. doi:10.1038/nbt1310

Weiss, D. J., Bertoncello, I., Borok, Z., Kim, C., Panoskaltsis-Mortari, A., Reynolds, S., . . . Prockop, D. J. (2011). Stem cells and cell therapies in lung biology and lung diseases. *Proc Am Thorac Soc, 8*(3), 223-272. doi:10.1513/pats.201012-071DW

XXX

Wold, W. S., & Toth, K. (2013). Adenovirus vectors for gene therapy, vaccination and cancer gene therapy. *Curr Gene Ther, 13*(6), 421-433. doi:10.2174/1566523213666131125095046

Wolff, J. A., Ludtke, J. J., Acsadi, G., Williams, P., & Jani, A. (1992). Long-term persistence of plasmid DNA and foreign gene expression in mouse muscle. *Hum Mol Genet, 1*(6), 363-369. doi:10.1093/hmg/1.6.363

Wong, S. P., & Harbottle, R. P. (2013). Genetic modification of dividing cells using episomally maintained S/MAR DNA vectors. *Mol Ther Nucleic Acids, 2*, e115. doi:10.1038/mtna.2013.40

Woodard, L. E., Welch, R. C., Williams, F. M., Luo, W., Cheng, J., & Wilson, M. H. (2018). Hydrodynamic Renal Pelvis Injection for Non-viral Expression of Proteins in the Kidney. *J Vis Exp*(131). doi:10.3791/56324

Yang, T., Heydarian, M., Kozjak-Pavlovic, V., Urban, M., Harbottle, R. P., & Rudel, T. (2020). Folliculin Controls the Intracellular Survival and Trans-Epithelial Passage of Neisseria gonorrhoeae. *Front Cell Infect Microbiol, 10*, 422. doi:10.3389/fcimb.2020.00422

Yang, Y., Padilla-Nash, H. M., Vira, M. A., Abu-Asab, M. S., Val, D., Worrell, R., . . . Vocke, C. D. (2008). The UOK 257 cell line: a novel model for studies of the human Birt-Hogg-Dube gene pathway. *Cancer Genet Cytogenet, 180*(2), 100-109. doi:10.1016/j.cancergencyto.2007.10.010

Zakrzewski, W., Dobrzynski, M., Szymonowicz, M., & Rybak, Z. (2019). Stem cells: past, present, and future. *Stem Cell Res Ther, 10*(1), 68. doi:10.1186/s13287-019-1165-5

Zhao, Y., Yin, X., Qin, H., Zhu, F., Liu, H., Yang, W., . . . Deng, H. (2008). Two supporting factors greatly improve the efficiency of human iPSC generation. *Cell Stem Cell, 3*(5), 475-479. doi:10.1016/j.stem.2008.10.002

Zhou, T., Benda, C., Dunzinger, S., Huang, Y., Ho, J. C., Yang, J., . . . Esteban, M. A. (2012). Generation of human induced pluripotent stem cells from urine samples. *Nat Protoc, 7*(12), 2080-2089. doi:10.1038/nprot.2012.115

Zhou, T., Benda, C., Duzinger, S., Huang, Y., Li, X., Li, Y., . . . Esteban, M. A. (2011). Generation of induced pluripotent stem cells from urine. *J Am Soc Nephrol, 22*(7), 1221-1228. doi:10.1681/ASN.2011010106

Zhou, Y., Zhang, T., Zhang, Q. K., Jiang, Y., Xu, D. G., Zhang, M., . . . Pan, Q. J. (2014). Unstable expression of transgene is associated with the methylation of CAG promoter in the offspring from the same litter of homozygous transgenic mice. *Mol Biol Rep, 41*(8), 5177-5186. doi:10.1007/s11033-014-3385-1

12 Acknowledgements

This work would not have been possible without the continuous help, support, and encouragement of so many people around me.

I would like to first of all thank Dr Richard Harbottle for taking me into his lab, and guiding, mentoring, and shaping me over the last four years. Thank you for allowing me to be nerdy with my partner in crime and experiments (Ali) and giving me the freedom to develop this project into unforeseen directions like the side road of the urinary stem cell project. Thank you for believing in my work and trusting in me especially in the moments when I didn't. Thank you for all the great memories inside and outside the lab, the training in proper English during lunch, conference trips, pasta competitions, and other gatherings. I'm sorry you didn't get to finish my additional training on increasing my tolerance for spicy food, I will try to work on it and widen the "Urban scale" in the future.

I would also like to thank Prof Dr Stefan Wiemann for his availability as my TAC member and faculty supervisor and the fruitful scientific input and discussions. Furthermore, I thank Prof Dr Lazaro Centanin and Prof Dr Tristan McKay for their prompt availability as chair and fourth member of my defence panel, respectively. Thank you to Dr Mark Eccleston, for his guidance, scientific open-mindedness, and valuable input throughout my PhD as my TAC member. Thank you for connecting me with Dr Martin Hoogduijn who joined my thesis and TAC half the way. A sincere thank you goes to him and his PhD student Steven (Zhaoyu Du) for their continuous valuable input and scientific exchange on kidney organoid differentiation. I would like to further thank Vanessa Vogel and Manuel Reitenberger for their work, help, troubleshooting and efforts on making embedding, fixation, and H&E stainings of these organoids possible.

Thank you to Dr Edward Green for being available for scientific questions and the theoretical and material support on the CRISPR/Cas9 and especially the 10x RNAseq experiments. Here, I would like to further thank Gordon Haltenhof for his help and

guidance during the processing of the sequencing samples, Katharina Linder for the help with the Bioanalyzer and Chin Leng Tan for his efforts and hours spent on analysing the data.

Thank you to the DKFZ FACS core facility for the constant support and reliability that enabled the drug-free establishments.

I thank all the donors for my urine cell study for being open-minded, curious, and willing to donate for science and enabling this work.

Thank you to Hendrik Welsch, Anna Hartley, Lukas Frank and Alexander Sandhof for the great scientific and personal exchange and fun during the joint mentoring of the molecular Cellbiology group of the Heidelberger Life Science lab. Thank you, Tamara Steinfuss, Jessica Devant and Luisa Henkel, for the emotional support and exchange during this adventure of the PhD.

I am deeply grateful for all the members of the lab – the ones who were there from the beginning, the ones who briefly stepped by and the ones who will form the future. These four years were a special time for me that I will never forget.

Thank you Dr Matthias Bozza for teaching me so much about the science and more importantly what matters beyond the lab. It was a pleasure to see you achieve all these milestones and learn from your experience for such a long time.

Dr Alicia Roig Merino, I can't thank you enough for introducing me to the special world of stem cells, for letting me tag along and for guiding me throughout the PhD and way beyond the lab. For always making time and having an open ear, for sharing the knowledge, experience and especially nerdiness and enthusiasm for science and the friendship.

Thank you, Alice de Roia, for the emotional support on long lab days, for always being available to help when experiments went wrong and for joining this ride. It wouldn't have been the same without you.

Thank you, Julia Peterson, for stepping in when I was despairing the dead cell removal in the last big experiment, for the valuable scientific and personal exchange throughout difficult circumstances, and for taking up and building on this work with hiPSCs so the knowledge doesn't get lost.

Acknowledgements

Thank you, Michael Bonadonna, our adopted lab member, for the chats in the corridor, the (theoretical and material) scientific input, but more importantly the times outside of the lab and especially the amazing food and cooking lectures.

Thank you to all the old lab members, Sveta, for teaching me at the very beginning and sharing your knowledge and experience; Patrick Almeida for joining me on the work with human stem cells and the exchange and career advice; Matthias Ehrbar for the time together in the lab, the support with all the little things that eat your time and make you despair and being there.

I thank also to the students that worked with me or with us: Daniel Stumpf, Patrick Wallisch, Tanja Hann and Nicole Codeluppi, but especially Julia Emmenecker and Bianca Maria Carrara for their scientific and personal contributions to this work and the everyday life in and outside the lab.

Four years is a long time. A lot of things are changing – that's how it works. But I'm sure the lab will find its way and the many new members that recently joined and will join soon will make it thrive.

While I spent so much time during the last four years worrying about my stem cells, I had plenty of support from outside that made this adventure possible and helped me to never give up.

Thank you, Nikos, for the past four years together, for taking this huge step to live in Germany, for helping me to enjoy life outside the lab, for withstanding my stress and occupied mind especially during the last 1.5 years and the continuous understanding and support.

Danke Jule, für das nicht-aus-dem-Auge-verlieren in 30 Jahren Freundschaft, dein Verständnis für meinen Weg und die vielen lieben Worte und kleinen Päckchen aus der Heimat, die mir geholfen haben durchzuhalten.

Doch der größte Dank geht an meine Eltern, die mich von klein auf in meiner Neugier, meinem Lernen und all meinen Entscheidungen unterstützt, begleitet und vor allem an mich geglaubt haben. Ohne euch wäre ich heute nicht hier. Danke.