Deciphering the role of orphan nuclear receptor GCNF in germ layer specification and early neural induction by utilizing CRISPR/Cas9

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Hohen Medizinischen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität

Bonn

<u>Nils</u> Christian Braun

aus Köln

2022

Angefertigt mit der Genehmigung

der Medizinischen Fakultät der Universität Bonn

1.   Gutachter: Prof. Dr. Oliver Brüstle

2.   Gutachter: Prof. Dr. Hubert Schorle

Tag der Mündlichen Prüfung: 21.12.2021

Aus dem Institut für Rekonstruktive Neurobiologie

Direktor: Prof. Dr. Oliver Brüstle

Table of contents

List of abbreviations

A	Adenine
AAV	Adeno-associated virus
AF-1	Activation function 1
AF-2	Activation function 2
AM	Active Motifs
AnP	Antarctic phosphatase
APS	Ammonium persulfate
ATP	Adenosine triphosphate
BCA	Bicinchoninic acid assay
BF1	Brain factor 1
bFGF	Basic fibroblast growth factor
bGH	Bovine growth hormone polyadenylation sequence
C	Cytosine
c-MYC	Avian myelocytomatosis virus oncogene cellular homolog
Cas	CRISPR-associated
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
CO_2	Carbondioxide
COUP-TFII	Chicken ovalbumin upstream promoter transcription factor 2
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	Crispr-RNA
CYP26	Cytochrome P450 26A1
DAPI	4',6-diamidino-2-phenylindole
DBD	DNA binding domain
DEPC	Diethyl pyrocarbonate
DMEM	Dulbecco's Modified Eagle's Medium
DMEM/F12	Dulbecco's Modified Eagle's Medium/Nutritient Mixture F-12
DMSO	Dimethyl sulfoxide
DNA	Desoxyribonucleic acid
DNMT	DNA methyltransferase

dNTP	Desoxyribonucleosid triphosphate
DR0	Direct repeat with zero spacer
DSB	Double strand break
dsDNA	Double stranded DNA
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
eGFP	Enhanced Green fluorescent protein
EMSA	Electrophoretic mobility shift assay
En1	Engrailed 1
EOMES	Eomesodermin
FCS	Fetal calf serum
FDA	Food and Drug Administration
FEZF1	FEZ family zinc finger 1/Forebrain embryonic zinc finger-like protein 1
FEZF2	FEZ family zinc finger 2/Forebrain embryonic zinc finger-like protein 2
FGF2	Fibroblast growth factor 2
FGF8	Fibroblast growth factor 8
FOXG1	Forkhead box G1
FST	Follistatin
G	Guanine
GCNF	Germ Cell Nuclear Factor
HAL	Homology Arm left
HAND1	Heart and neural crest derivatives expressed 1
HAR	Homology Arm right
HCl	Hydrogen chloride
HDR	homology directed repair
HEK293	Human embryonic kidney 293
hESC	Human embryonic stem cell
hiPSC	Human induced pluripotent stem cell
HNH	Endonuclease domain that contains histidine and asparagine
hPSC	Human pluripotent stem cell
HRP	Horseradish peroxidase
iap	E. coli alkaline phosphatase

ICC	Immunocytochemistry
ICM	Inner cell mass
IgG	Immunglobulin G
kDa	Kilo Dalton
KLF4	Kruppel-like factor 4
KRAB	Kruppel-associated box of Kox1
KSR	Knockout serum replacement
LAAP	L-ascorbic acid 2-phosphate
lacZ	β-galactosidase
LBD	Ligand binding domain
lncRNA	long non-coding RNA
lt-NES	Long-term self-renewing neuroepithelial-like stem cells
loxP	Locus of X-over P1
MBD	Methyl-CpG-binding domain
mESC	murine embryonic stem cell
$MgCl_2$	Magnesium chloride
miRNA	micro RNA
MIT	Massachusetts Institute of Technology
mRNA	messenger RNA
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NaSe	Sodium selenite
NCoR1	Nuclear receptor co-repressor 1
NE	Non-edited
NEAA	Non-essential amino acids
NFIB	Nuclear factor 1 B-type
NGS	Next generation sequencing
NHEJ	Non-homologous end joining
NPC	neural precursor cell
NR6A1	Nuclear Receptor Subfamily 6, Group A, Member 1
NRIP1	Nuclear receptor-interacting protein 1
NSC	neural stem cell

NT2/D1	NTERA-2 cl.D1
NTD	N-terminal Domain
OCT4	Octamer-binding transcription factor 4
ORF	Open reading frame
Ori	Origin of replication
OSKM	OCT4, SOX2, KLF4, c-MYC
PAM	Protospacer adjacent motif
PAX2	Paired box 2
PAX3	Paired box 3
PAX5	Paired box 5
PAX6	Paired box 6
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PGK	Phosphoglycerate kinase 1 promoter
PNK	Polynucleotide Kinase
polyA	polyadenylation sequence
PuroR	Puromycin resistance
PVDF	Polyvinylidene fluoride
qRT-PCR	Quantitative Real-Time PCR
RA	Retinoic acid
RE	Response element
RI	ROCK-inhibitor Y-27632
RIPA	Radioimmunoprecipitation assay buffer
RNA	Ribonucleic acid
RNP	Ribonucleoprotein complex
rpm	Rounds per minutes
RSPO3	R-Spondin 3
RuvC	Endonuclease domain found in an E. coli DNA repair protein
RXR	Retinoid X receptor
SDS	Sodium dodecyl sulfate
sgRNA	Single-guide RNA

smNPC	Small molecule neural precursor cell
SMRT	Silencing mediator for retinoid or thyroid-hormone receptors
SOX2	SRY-box transcription factor 2
SPRY1	Protein sprouty homolog 1
SRSR	Short regularly spaced repeats
ssODN	Single-stranded oligonucleotide
Str8	Stimulated by retinoic acid gene 8
SYBR	SYBR Green I
T	Thymine
T2A	Self cleaving peptide T2A
TALEN	Transcription-like effector nucease
TAT	Transactivator of Transcription
TBR1	T-box, brain, 1
TBS-T	Tris-buffered saline with Tween20
TDGF1	Teratocarcinoma-derived growth factor 1 = CRIPTO
TEMED	Tetramethylethylenediamine
TGFβ1	Transforming growth factor β1
tracrRNA	Transactivating crRNA
TREP	Tandem Repeats
TRIF	Transiently retinoid-induced factor
Tris	Tris(hydroxymethyl)aminomethane
TRPM8	Transient receptor potential cation channel subfamily M member 8
UTR	Untranslated region
VP16	Herpes simplex virus protein vmw65
VPR	VP64-p65-Rta
WNT	Wingless-related integration site
ZFN	Zinc finger nuclease
ZO-1	Zonula occludens-1

1. Introduction

1.1 The discovery of pluripotent stem cells revolutionized regenerative medicine

Stem cells are defined as being capable of both self-renewal as well as giving rise to differentiated progeny that may carry out a defined function in an association of cells or an organ system. Pluripotent stem cells (PSCs) are able to form cells from all three germ layers, ectoderm, mesoderm and endoderm (De Los Angeles et al., 2015), thereby giving rise to all cells of the human body. A major breakthrough in regenerative medicine, was the establishment of murine embryonic stem cells (mESCs) from the inner cell mass (ICM) of a mouse blastocyst (Evans and Kaufman, 1981).

By studying mESCs, researchers not only learned how to generate many different cell types (Keller, 1995) but also how to genetically modify mESCs in the cell culture dish and reintroduce the modified cells into blastocysts to create a mouse deficient for a desired gene (Kuehn et al., 1987) or carrying a desired genetic modification (Nandi et al., 1988). The work on genetically modified mice revolutionized biomedicine and allowed valuable insights into processes of development as well as into mechanisms of disease (Watase and Zoghbi, 2003). Despite the groundbreaking research on mESCs, it still needed more than a decade of research until human embryonic stem cell lines (hESCs) were established from the ICM of a human blastocyst (Thomson et al., 1998). The generation of hESCs holds the promise to not only learn about a disease in a model system but also to create therapies to replace cells that were irreversibly damaged due to injury or disease. The ethical concerns raised by the generation and research with hESCs could be overcome by the discovery of induced pluripotent stem cells (iPSCs). These cells are generated by introducing four embryonic genes OCT4, SOX2, KLF4 and c-MYC into terminally differentiated somatic cells and reprogram them to an embryonic stem cell-like cell type (Takahashi and Yamanaka, 2006). Soon after murine iPSCs had been described (Takahashi and Yamanaka, 2006), it became also possible to reprogram a variety of human somatic cells to hiPSCs, including fibroblasts and blood (Seki et al., 2010; Takahashi et al., 2007; Yu et al., 2007). In terms of cell replacement therapy, cells generated from patient hiPSCs have the advantage of being immunologically compatible with the donor tissue, whereas there may be the need for immunosuppression in hESC-

derived cell replacement therapies (Fig. 1). Although cell replacement therapies for clinical translation are still in their infancy, some of them are already at the level of primate studies (Kikuchi et al., 2017; Liu et al., 2018) or even at the level of clinical trials (Deinsberger et al., 2020; Mandai et al., 2017). Generating hiPSCs from healthy donors and differentiating them site by site with hiPSCs from a patient suffering from a specific disease contributed to the field of disease modelling. In many cases this methodology helped to decipher how a disease unfolds its detrimental effect on the molecular level. Disease modelling could thereby also facilitate search for pharmacological substances to treat a specific aspect of a pathology (Fig. 1). However, to assess phenotypes of disease it is often crucial to obtain a specific cell type from human pluripotent stem cells (hPSCs). Until today, scientists generated a roadmap for differentiation into a variety of cell types, with which it becomes increasingly possible to generate a desired cell type (Tabar and Studer, 2014).
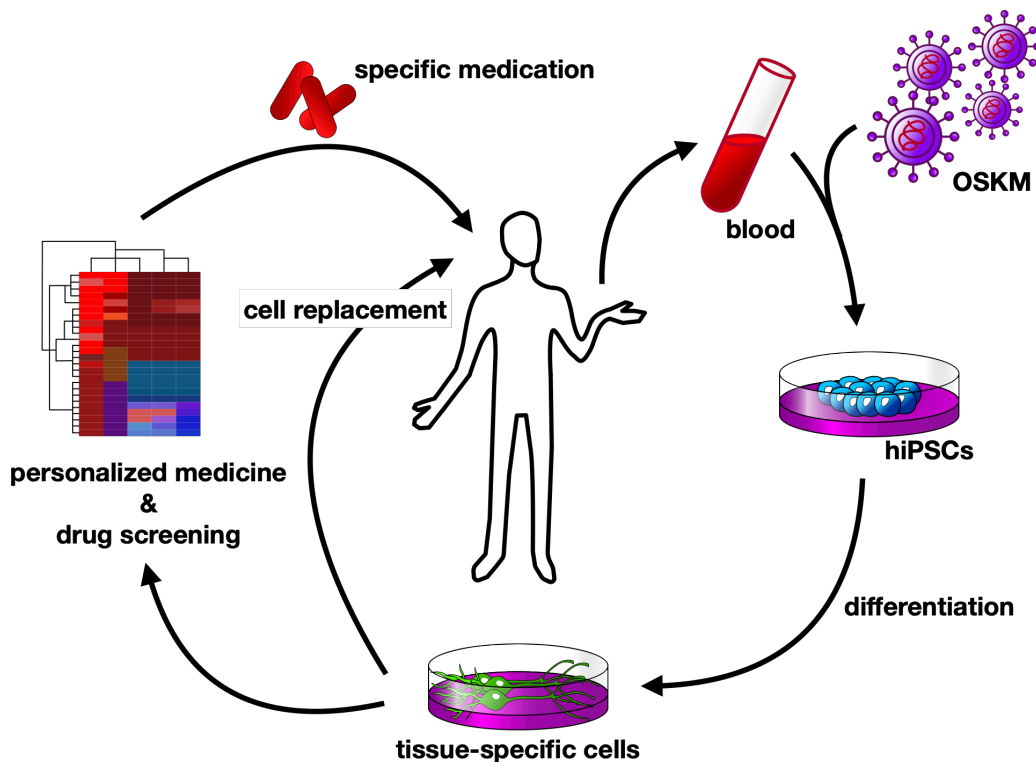
Fig. 1: Stem cells have revolutionized regenerative medicine. The generation of hiPSCs from blood with the four Yamanaka factors (OSKM), packaged into non-integrating vectors, like Sendai viruses, makes it possible to readily obtain pluripotent stem cells. HiPSCs can be differentiated into cells that are hard or impossible to access, like neuronal subtypes. Neurons that had been generated that way may be utilized in a cell replacement therapy or for personalized medicine and drug screening that may yield specific medication for a certain disease.

1.2   Human pluripotent stem cells to study the development of the nervous system

As hPSCs resemble a very early stage of human development they can be utilized to obtain information about embryogenesis. This is especially interesting for questions that address organs like the central nervous system where model organisms like the mouse lack key features of human development and where it is impossible to obtain human tissue due to ethical concerns. HPSCs can be expanded indefinitely without losing pluripotency when kept in self-renewing conditions (Takahashi et al., 2007; Thomson et al., 1998). Once hPSCs enter the process of differentiation they undergo a large variety of changes on morphologic, transcriptomic and epigenetic level (Han et al., 2018). These changes are reminiscences of the stages it takes a unicellular zygote to generate an embryo. After fertilization, the zygote undergoes several cleavages and thereby forms the ball-shaped morula in a process called compaction, that is accompanied by zygotic genome activation (Jukam et al., 2017). The morula develops into the fluid-filled blastocyst which is composed of two parts: The outer surface of the blastocyst consists of cells of the trophoblast, which forms the placenta after implantation (Shahbazi and Zernicka-Goetz, 2018). At the interior of the blastocyst lies the ICM (Pedersen et al., 1986). The ICM develops into the hypoblast and the epiblast (Roode et al., 2012). In the process of gastrulation, the epiblast develops further into the three germ layers, ectoderm, mesoderm and endoderm (Rossant and Tam, 2017). The ectoderm will differentiate into the brain and gives also rise to the epidermis (Murry and Keller, 2008). In the presence of morphogens like NOGGIN, FOLLISTATIN and CHORDIN, a part of the ectoderm differentiates into the neuroectoderm, which is the earliest stage of brain development (Conti and Cattaneo, 2010; Pevny et al., 1998). In the process of neurulation, the neuroectoderm subdivides into the neural plate and the neural crest. The cells of the neural crest are formed where neural plate and epidermis adjoin and they give rise to a large variety of cells including neuronal and non-neuronal cells (Green et al., 2015). The neural plate instead folds to the neural groove which then closes to form the neural tube. It is possible to assign different neurodevelopmental stages to *in vitro* counterparts in the form of characterized stem cell populations (Conti and Cattaneo, 2010) (Fig. 2 A). Starting from hPSCs (Fig. 2 B) these populations of neural stem or precursor cells (NSCs or NPCs respectively) have been reported to resemble different developmental stages. Early

neuroepithelial stem cells are characterized by the formation of rosette structures that are supposed to be an *in vitro* reminiscence of the polarized cells of the neural plate and neural tube (Elkabetz et al., 2008), a theory which is supported by the presence of zonula occludens (ZO-1) expression at the center of the rosette structures (Elkabetz et al., 2008; Koch et al., 2009) (Fig. 2 C). The first neural differentiation paradigms of hPSCs relied on undirected differentiation and subsequent extraction of rosette structures (Elkabetz et al., 2008; Koch et al., 2009). However, more recent approaches rely on inhibition of TGFβ- as well as BMP-mediated SMAD signaling (dual-SMAD inhibition) and allow the neuralization of hPSCs with high efficiency (Chambers et al., 2009). When the early neuroepithelial stem cells differentiate to neural stem cells *in vitro* they can either form clustered colonies of cells when exposed to WNT signaling, like small molecule neural precursor cells (smNPCs) (Reinhardt et al., 2013) (Fig. 2 D) or smaller rosette structures when maintained as lt-NES with EGF and FGF2 (Koch et al., 2009) (Fig. 2 E). A population of rosette forming NSCs, similar to lt-NES, had been generated directly from fetal brain tissue, adding evidence to the hypothesis of cultured NSCs being *in vitro* correlates of an *in vivo* counterpart (Tailor et al., 2013). Upon growth withdrawal, NSCs further differentiate and generate mature neurons (Fig. 2 F). *In vivo*, the neural tube structure already resembles the basic architecture of the adult brain, with the most anterior part of the neural tube generating the telencephalon, and the following parts from anterior to posterior generating the diencephalon, mesencephalon, cerebellum, hindbrain and spinal cord (Stiles and Jernigan, 2010). The central cavity enclosed by the neural tube will develop into the cerebrospinal fluid filled ventricular system (Stiles and Jernigan, 2010). In the next weeks and months of development, these structures become more and more mature and neurons gradually adopt their final identity. Whereas neurons can be generated from stable NSC intermediates, like smNPCs or lt-NES, they typically exhibit posterization which might be a problem when more anterior neuronal cell types are required (Conti and Cattaneo, 2010). Therefore, many recent approaches focus on generation of specific neural subtypes, like cortical excitatory neurons or dopaminergic neurons, directly from hPSCs (Kriks et al., 2011; Shi et al., 2012). Usually, these paradigms utilize dual-SMAD inhibition to neuralize a population of hPSCs and then expose them with specific stimuli to guide them into the desired cell type, without stabilizing transiently emerging NSCs.
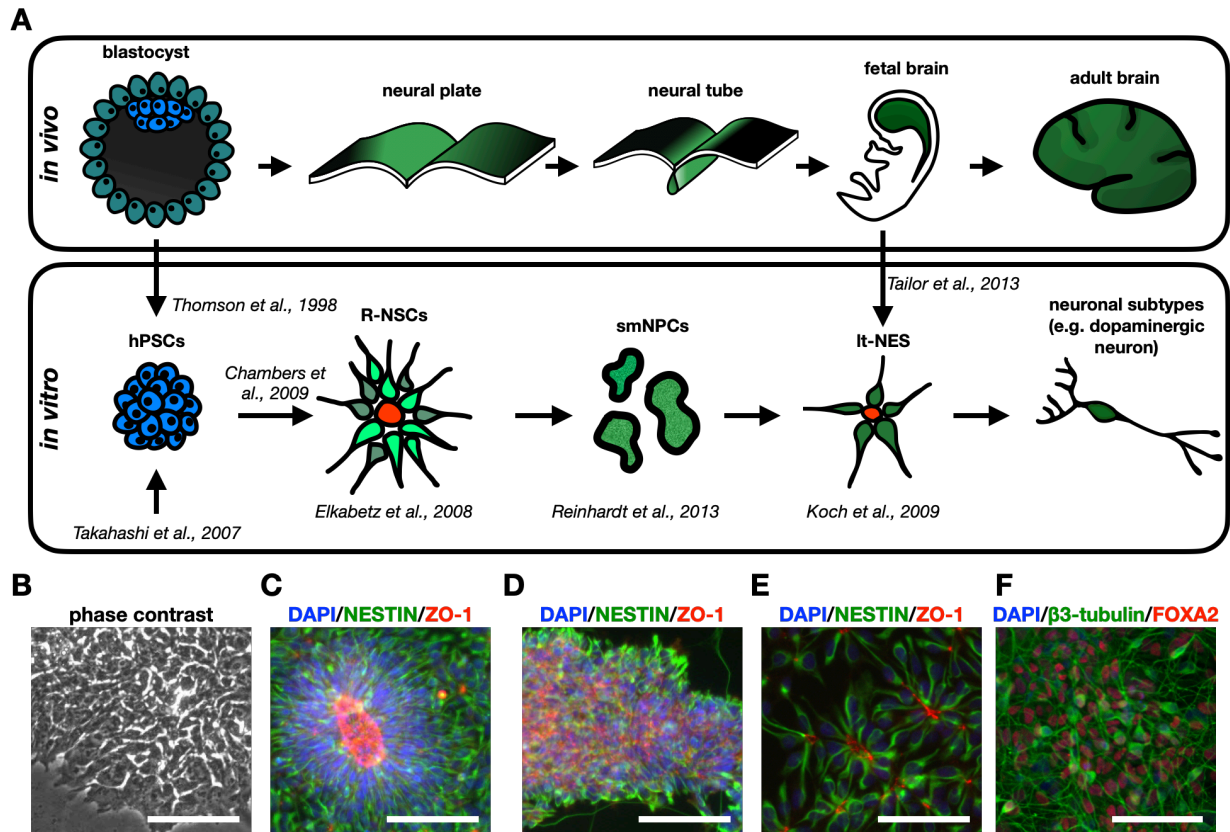
Fig. 2: Different stages of neurodevelopment have *in vitro* counterparts. **(A)** Stages of human neurodevelopment and their *in vitro* counterparts. Illustration is adapted from Conti and Cattaneo (2010). **(B-F)** Microscopic images of the cell types with its decreasing developmental potential and increasing differentiation. **(B)** HPSCs exhibit the highest potential as they can give rise to almost all cell types of the human body, including cells of the central nervous system. **(C)** Once hPSCs enter the neural lineage, neural stem cells are generated that are usually NESTIN-positive. The spontaneous rosette formation is a property that is believed to be reminiscent of the neural tube. ZO-1 expression at the center of the rosette structures is a marker for this polarization. **(D)** Neural stem cells can either form clusters when kept under constant WNT signaling (smNPCs) or **(E)** exhibit rosette formation when kept under EGF and FGF2 (lt-NES). **(F)** Finally, neurons are generated that can be further divided into specific subtypes, e.g. dopaminergic neurons in the midbrain that are positive for β3-tubulin and FOXA2. Scale bar = 100 µm.

1.3 GCNF during embryonic and neural development

In the last decades researchers gained a deeper understanding of the molecular level of brain development. It requires an extraordinary orchestration of transcription factors, miRNAs, secreted proteins, DNA methylation and histone modification to generate such

a complex organ as the human brain. Classical transcription factors and nuclear receptors are crucial for guiding neural differentiation as they interact with specific motifs on the DNA, thereby repressing or activating downstream genes. Nuclear receptors form a diverse family which had been systematized by the Nuclear Receptors Nomenclature Committee in 1999 (Nuclear Receptors Nomenclature Committee, 1999). They are involved in many essential biological processes including inflammation, metabolism, reproduction and development (Glass and Ogawa, 2006; Jeong and Mangelsdorf, 2009) and are especially interesting for clinical applications because they bind to DNA in order to transmit a signal after sensing a signaling molecule. Indeed, around 13 % of FDA-approved drugs act on nuclear receptors (Overington et al., 2006). Many nuclear receptors are well known for eliciting effects of clinically and endocrinologically well-known hormones like thyroids, retinoids, estrogen, testosterone, and the glucocorticoids (Tata, 2002). Stem cell maintenance and differentiation is also modulated by nuclear receptors, by participating in gene networks that mediate self-renewal or shift the balance towards differentiation (Jeong and Mangelsdorf, 2009; Mullen et al., 2007).

1.3.1 GCNF belongs to the orphan nuclear receptors and binds to specific response elements

A very interesting candidate gene that may be involved in the modulation of stem cell self-renewal and differentiation is Germ Cell Nuclear Factor (GCNF, gene symbol NR6A1). GCNF belongs to the orphan nuclear receptors because no ligand has yet been found (Wang and Cooney, 2013). Human GCNF is mapped to chromosome 9 (Agoulnik et al., 1998) with the recent Ensembl coordinates being Chromosome 9: 124,517,275 – 124,771,311 reverse strand GRCh38:CM000671.2 in the region of q33.3 (Yates et al., 2020). Human GCNF shares a high degree of homology with murine GCNF with an overall of 98.7 % similarity (Süsens and Borgmeyer, 1996). Whereas there is only few information about GCNFs function in comparison to its well-known relatives, it shares the basic structure with all members of the nuclear receptor family (Fig. 3 A) (Greschik and Schule, 1998; Tata, 2002). The GCNF DNA-binding domain (DBD) contains two Zinc-finger motifs and is evolutionary strongly conserved (Wang and Cooney, 2013). Indeed, whereas the overall amino acid sequence of GCNF shares 32 – 34 % with the retinoid X receptor

(RXR), the GCNF DBD shares up to 61 % similarity with the RXR DBD, including a P-box sequence that is crucial for DNA binding (Hirose et al., 1995). The DBD mediates the interaction of GCNF with its specific response element (RE) on the DNA (Fig. 3 B). This interaction has recently been resolved by a crystal structure study (Weikum et al., 2016). The function of GCNF relies on a functional DBD as GCNF is not capable of serving its role in embryonic development without DNA binding. The loss of a functional DBD phenocopies the full knockout of GCNF (Lan et al., 2002). Due to the similarity of the GCNF DBD to the RXR DBD it was suggested early that their RE should share common features. Electrophoretic mobility shift assay (EMSA) identified GCNF binding to either an extended half-site TCAAGGTCA or a direct repeat with zero spacer (DR0) AGGTCAAGGTCA (Chen et al., 1994; Yan et al., 1997). Whereas some nuclear receptors depend on heterodimerization with RXR, EMSA suggests that GCNF binds DNA as a homodimer (Borgmeyer, 1997). Upon RA treatment, some experiments suggested that GCNF might also bind to its DR0 RE as part of a larger TRIF complex (transiently retinoid-induced factor) (Gu et al., 2005). The ligand-binding domain (LBD) is another common feature of nuclear receptors (Huang et al., 2010; Kumar and Thompson, 1999). Whereas, the GCNF DBD is related to the RXR DBD, the GCNF LBD shares 26 % similarity with the COUP-TFII LBD (Hirose et al., 1995; Süsens and Borgmeyer, 1996; Weikum et al., 2016). Generally, the nuclear receptor LBD consists of 12 α-helices that serve in binding a lipophilic ligand, where the helix 12 harbors an AF-2 activation function, that mediates activation of a gene upon ligand-binding. However, in the helix 12 core region of GCNF a glutamine is substituted with a lysine, thereby rendering this region of GCNF different from any other nuclear receptor (Greschik et al., 1999; Kapelle et al., 1997; Yan and Jetten, 2000). More recent studies indicate that the LBD harbors a β-sheet which was shown to interact with transcriptional corepressors like NCoR1 and SMRT (Weikum et al., 2016) (Fig. 3 C) and a Gal4-GCNF LBD assay indicates that GCNF interacts with DNA methyltransferases (Gu et al., 2011) (Fig. 3 D). The repressive activity of GCNF has also been reported in several functional cell culture studies (Cooney et al., 1998; Fuhrmann et al., 2001; Hummelke and Cooney, 2004). Interestingly, a conditional ablation of the GCNF LBD does not disrupt its function in female ovarian germ cells indicating that it is not required for its function as a transcriptional repressor (Okumura et al., 2013). The hinge

region lies in between the DBD and LBD and was also described to interact with corepressors (Yan et al., 1997).

A

GCNF protein domains

B

GCNF response element
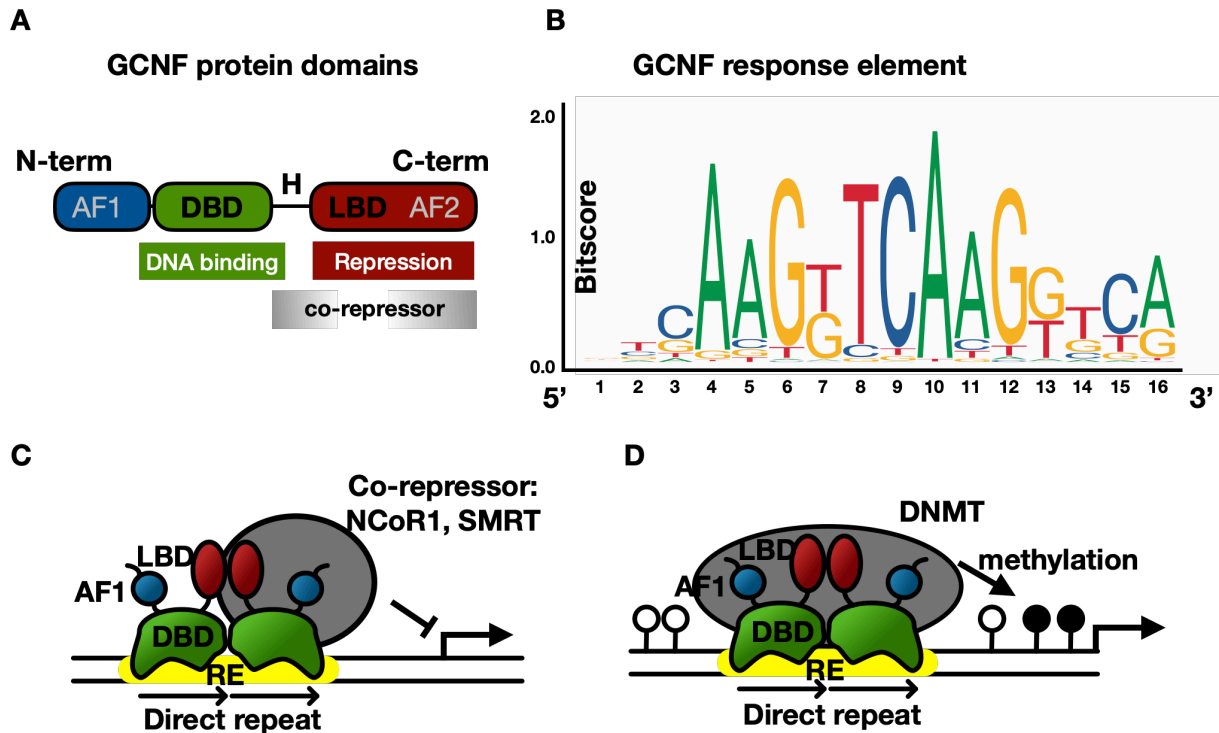
Fig. 3: GCNF belongs to the orphan nuclear receptors and binds to a response element (RE) on the DNA to mediate gene regulation. **(A)** The GCNF protein consists of different domains. The DNA-binding-domain (DBD) mediates the interaction and binding to the respective RE on the DNA. Whereas other nuclear receptors usually have an N- and C-terminal activation function, the AF1 seems to be absent in GCNF and the AF2 amino acid sequence is different from the consensus sequence (AFs marked in grey). Ligand-binding domain (LBD), Hinge region (H). **(B)** Depiction of the GCNF RE as a Bitscore model, generated by the JASPAR online tool (Fornes et al., 2020). The frequency of each nucleotide is indicated by the size of the letters. **(C)** Upon binding to its RE on the DNA, GCNF acts as a repressor of transcription by recruiting co-repressors like NCoR1 and SMRT. **(D)** GCNF is also reported to recruit DNA methyltransferases (DNMT), that mediate silencing of target genes by inducing DNA methylation. Schemes in C and D are adapted from Gu et al. (2011).

1.3.2 GCNF plays an important role during embryonic and neural development

As the first full-length GCNF open reading frame (ORF) was obtained from a mouse testicular cDNA library, its implication in gonad development was hypothesized (Chen et al., 1994). Indeed, GCNF is indispensable for gametogenesis and GCNF knockout mice

show impaired spermatogenesis and oogenesis (Sabour et al., 2014). Similarly, *in vitro*, GCNF-deficient mESCs do not properly differentiate into germ cells when treated with respective developmental cues (Sabour et al., 2014). In the adult organism, GCNF only remains expressed in the gametes (Katz et al., 1997). Interestingly, the gametes are also the only place where persistent OCT4 expression is observed (Schöler et al., 1989b; Yeom et al., 1996; Yoshimizu et al., 1999). OCT4 is also highly expressed in the ICM of the blastocyst stage and is one of the four Yamanaka factors that mediate reprogramming of somatic cells into iPSCs (Takahashi and Yamanaka, 2006). This renders OCT4 being a key component of the pluripotency network (Schöler et al., 1989a; Schöler et al., 1989b; Takahashi and Yamanaka, 2006). The presence of DR0 elements in the OCT4 promoter suggested that GCNF might negatively regulate OCT4 when PSCs exit pluripotency and enter differentiation. First evidence indicating GCNF-mediated regulation of OCT4 came from retinoic acid (RA) induced differentiation of the murine embryonic carcinoma cell-line P19. In that context, endogenous GCNF as well as exogenously overexpressed GCNF is capable of silencing OCT4 expression by binding to its RE in the OCT4 promoter, whereas GCNF fused to a VP16 transactivator domain upregulates OCT4 (Fuhrmann et al., 2001). GCNF establishes OCT4 silencing upon RA-induced differentiation by recruiting MBD2 and MBD3 (methyl-CpG binding domain protein) as well as DNA methyltransferases to the OCT4 promoter, which is subsequently methylated (Gu et al., 2011). Indeed, it was also shown that GCNF represses OCT4 and modulates gene expression in hESCs (Wang et al., 2016). Interestingly, partially reprogrammed iPSCs do not differ phenotypically from regular iPSCs, but are not capable of proper differentiation. The impairment in differentiation was found to be mediated by inaccurate demethylation of the GCNF promoter (Wang et al., 2013). The question remains why GCNF is downregulating OCT4 on the one hand, but is expressed in the adult reproductive organs along with OCT4. Apparently, in germ cells GCNF downregulates OCT4 to facilitate induction of Str8 (stimulated by retinoic acid gene 8) which in turn induces meiosis in primordial germ cells (Sabour et al., 2014). A similar mechanism might explain the role of GCNF in placental development (Morasso et al., 1999). As the silencing of OCT4 is one of the key processes that mediates differentiation and is indispensable for gastrulation, the question arises whether GCNF only serves in exiting pluripotency or whether it is additionally involved in mediating entry into a specific germ layer. Among the first clues that GCNF might also be

involved in neural differentiation is the observation that a murine embryonic carcinoma cell line upregulates GCNF expression upon RA-induced neural differentiation (Bauer et al., 1997). In Xenopus laevis, GCNF seems to be involved in anterior-posterior patterning (David et al., 1998). The loss of GCNF results in the expression of the RA degrading enzyme CYP26, which leads to a net decrease of RA and in turn to the expansion of anterior structures and a more posterior shifted midbrain-hindbrain boundary (Barreto et al., 2003a). Additionally, the use of morpholinos that downregulate GCNF expression in Xenopus oocytes result in defective neural tube closure and affect integrin levels (Barreto et al., 2003b). In situ hybridization data indicates GCNF expression in the embryonic mouse brain (Bauer et al., 1997). A knockout of GCNF in the mouse results in embryonic lethality at day E10.5 due to disrupted heart development (Chung et al., 2001), but is also characterized by malformation of the neural tube, disturbed construction of the anterior neural ridge as well as the midbrain-hindbrain organizer region (Akamatsu et al., 2009; Chung et al., 2006). In mammals, GCNF seems to be especially involved in forebrain patterning. In the mouse embryo, loss of GCNF leads to decreased levels of key components of forebrain development like BF1, FGF8, SPRY1 (Chung et al., 2006). Additionally, also the midbrain-hindbrain-boundary is not set up in the regular way indicated by a loss of PAX2, EN1 and WNT1 as well as disturbed development of the cranial nerves (Chung et al., 2006). There is also evidence that GCNF promotes the development of primitive NSCs, that still have characteristics of PSCs, in comparison to definitive NSCs, which are restricted to the neural lineage (Akamatsu et al., 2009). Interestingly, it has also been shown that GCNF binds to DR0 elements in the CRIPTO (gene symbol TDGF1) promoter (Hentschke et al., 2006). CRIPTO is known to be part of the pluripotency network (Fiorenzano et al., 2016), and additionally is crucial for early mesoderm specification as well as heart development (Xu et al., 1998). It is especially worth noting that a CRIPTO ablation in the mouse embryo enhances neuronal differentiation (Parisi et al., 2003). Therefore, CRIPTO downregulation by GCNF might mediate both exit of pluripotency as well as favoring ectoderm towards mesoderm specification. Taken together, GCNF serves as an important factor in embryogenesis. Besides its well described mediation of pluripotency exit, a large amount of evidence is pointing to an additional role in neural development. Given the high amount of evolutionary

conservation of GCNF, it is very likely to serve a similar purpose in human neural development as described in mouse and Xenopus models.

1.4 Genome editing is a powerful tool for both clinical translation and basic research

In the second half of the twentieth century scientists suggested DNA being at the core of inheritance (Watson and Crick, 1974). In this model the sequence of the nucleotides adenine (A), thymine (T), guanine (G) and cytosine (C) encodes information within the DNA. DNA serves as the information storage molecule in the form of discrete sections called genes. Genes determine the structure of proteins in the process of protein biosynthesis. Whereas DNA itself is an inert molecule, proteins are the effectors of biological activity and maintain all of the cell's vital functions. Distortion of DNA at specific loci causes an aberrant protein to be expressed. The analysis of the resulting phenotype after mutation of a protein coding gene enables researchers to understand and characterize the function of the according protein. In lower organisms, like drosophila, introduction of unspecific DNA damage by chemical mutagens was sufficient to uncover important processes of development (Nusslein-Volhard and Wieschaus, 1980), which also contributed to the understanding of human disease (Bier, 2005). However, this method could not be utilized for higher organisms like the mouse, because it is not feasible to create mouse mutants in that high numbers (Capecchi, 2001). Whereas homologous recombination was previously only known to occur in cells that undergo meiosis (Creighton and McClintock, 1931; Hunter, 2015), it was discovered to be also present in somatic cells (Folger et al., 1982). Site-directed mutagenesis that relied on homologous recombination proofed to be sufficient to insert exogenous DNA sequences at desired locations in human carcinoma cell lines (Smithies et al., 1985) as well as mESCs (Thomas and Capecchi, 1987). MESCs that had been edited were utilized to generate mice deficient for a specific gene (Thomas and Capecchi, 1990), and the concept of the knockout mouse became famous for deciphering the role of many genes involved in human health and disease (Watase and Zoghbi, 2003). Whereas techniques were developed to recover the rare events of a desired modification with antibiotic selection (Mansour et al., 1988) the overall efficiency can be as low as one successfully targeted clone per one million electroporated cells (Capecchi, 1989; Thompson et al., 1989).

1.4.1 Introducing double strand breaks at specific locations in the genome results in enhanced site-directed mutagenesis

Experiments with rare-cutting endonucleases (meganucleases) like the homing endonuclease I-SceI from yeast provided evidence that the efficiency of modifying a gene can be enhanced by the introduction of double strand breaks (DSBs) (Rouet et al., 1994a). The specificity of meganucleases to one specific site is mediated by long restriction sites of up to 18 bp (Belfort and Bonocora, 2014; Jasin, 1996) (Fig. 4 A) and opened the possibility to obtain further understanding of the cellular repair machinery. In the mouse genome a DSB at a specific location resulted in an enhanced homologous recombination frequency of up to 1000 fold (Rouet et al., 1994b). Additionally, two major repair mechanisms could be distinguished that cells utilize to fix a DSB: The first of these processes was later termed homology-directed repair (HDR) in the genome editing field and describes the high accuracy repair of DNA utilizing homologous recombination. Opposed to that, in the second process, non-homologous end joining (NHEJ), the proximal and distal sites of the DNA are joined and ligated. This process might serve well in repairing DNA, but is also prone to introduce errors and mutations at the respective location in the genome (Jasin and Rothstein, 2013; Rouet et al., 1994b). Even though a variety of meganucleases where described, their large recognition sites had to be inserted into the cell line first, before making use of the DSB-mediated enhancement of homologous recombination (Belfort and Bonocora, 2014).

Zinc-finger motif containing eukaryotic transcription factors are characterized by a modular structure where each "finger" recognizes a triplet of nucleotides (Pavletich and Pabo, 1991). As DNA binding and cutting domains are separate in the FokI restriction enzyme (Li et al., 1992) fusing its cutting domain with Zinc finger motifs mediated *in vitro* DNA cleavage at a specified position (Kim et al., 1996), especially upon dimerization (Bitinaite et al., 1998) (Fig. 4 B). Zinc Finger Nucleases (ZFNs) were first utilized in Xenopus oocytes (Bibikova et al., 2001) and drosophila (Bibikova et al., 2002), and later transferred to the mouse (Carbery et al., 2010) as well as human cell lines (Urnov et al., 2005). Until now, ZFNs have been used in numerous studies in many different organisms (Urnov et al., 2010).
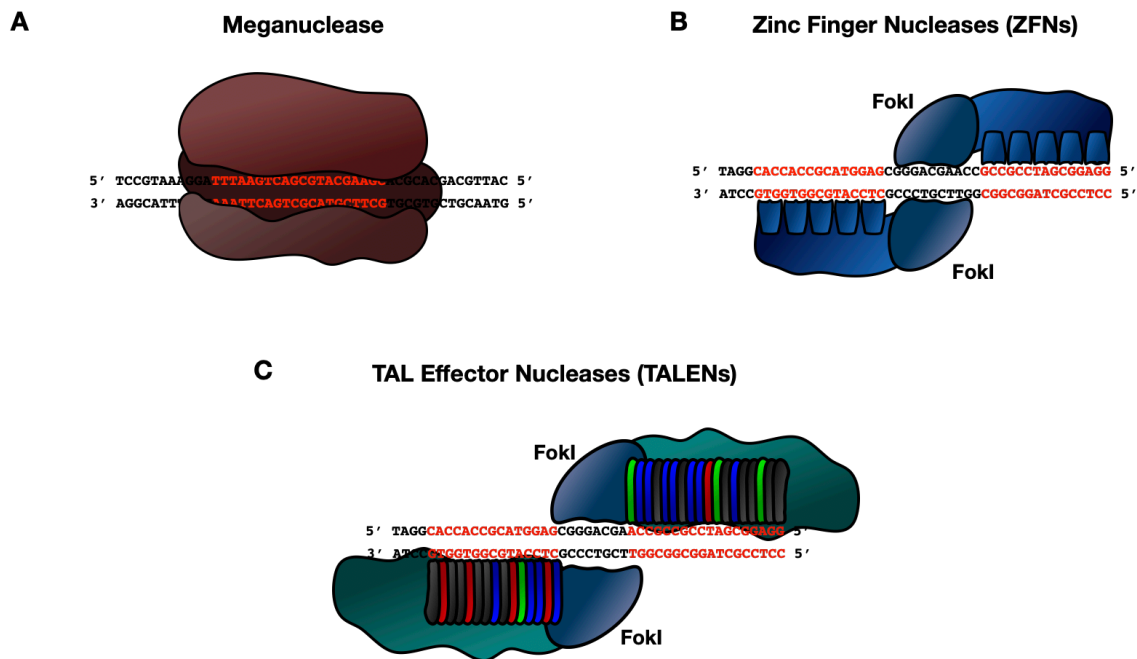
Fig. 4: Genome editing can be enhanced by a variety of endonucleases. **(A)** Meganucleases are restriction enzymes found in yeast and are characterized by long recognition sequences. Meganucleases contributed to the understanding of how a DSB increases homologous recombination. **(B)** Zinc Finger Nucleases (ZFNs) were the first nucleases that could be customized. Each module targets three base pairs. Dimerization of two FokI domains mediate a DSB. **(C)** TAL Effector Nucleases (TALENs) are characterized by modules that target individual base pairs. Although it is difficult to assemble TALEN modules it is more feasible than ZFNs. More information can be found in a comprehensive comparison of ZFNs, TALENs and CRISPR/Cas9 (Kim and Kim, 2014).

In order to infect a variety of plants, Xanthomonas bacteria express Transcription-like effector (TALE) molecules in the plant cell to hijack the hosts transcriptional machinery (Bogdanove et al., 2010). Bioinformatical analysis as well as experimental approaches revealed that two hypervariable amino acids within each repeat can be mapped to one base pair in the binding site of the hosts genome (Boch et al., 2009; Moscou and Bogdanove, 2009). Fusing a customized TALE array to a FokI cutting domain to generate a TAL Effector Nuclease (TALEN) has initially been shown to work in yeast assays (Christian et al., 2010; Li et al., 2011) and after further optimization in human immortal cell lines like K562 and HEK293 cells (Miller et al., 2011) as well as hESCs and hiPSCs (Hockemeyer et al., 2011) (Fig. 4 C). However, whereas ZFNs may result in cytotoxicity,

both ZFNs and TALENs are limited in the sophisticated way to assemble its modular DNA binding region (Kim and Kim, 2014; Ramirez et al., 2008).

1.4.2 CRISPR/Cas systems serve as an adaptive immunity in bacteria and archaea

In 1987 a group of researchers found an obscure repeat region in close proximity of the iap gene in the E. coli K12 strain. Direct repeats with 29 nucleotides were found to be interrupted by spacers of 32 nucleotides (Ishino et al., 1987). In the following years, this repeat region was also found in another chromosomal region of the same E. coli strain as well as in two other gram-negative bacteria, specifically Salmonella typhimurium and Shigella dysenteriae (Nakata et al., 1989). Similar repetitive regions are present in gram-positive bacteria like Mycobacteria (Groenen et al., 1993; Hermans et al., 1991), as well as in archaea (Mojica et al., 1993) where they were studied more systematically in Haloferax mediterrani and Haloferax volcanii and assigned "Tandem Repeats" (TREPs) (Mojica et al., 1995). A comparative study searched for the presence of these structures and characterized common properties in completed microbial genomes. In a huge diversity of species, the structures were usually present in clusters, consisted of partially palindromic, conserved repetitive units interspaced by non-repetitive sequences. The novel repetitive structures were thus re-assigned "short regularly spaced repeats" (SRSR) (Mojica et al., 2000). After another profound analysis and characterization of the SRSRs in over 40 different genomes they were finally assigned Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) (Jansen et al., 2002). In close proximity of the CRISPR loci conserved CRISPR-associated (Cas) genes, cas1 - 4 were found. Due to their basic amino acid sequence it was hypothesized that the cas proteins might be involved in DNA binding or modification (Jansen et al., 2002). Subsequently, the repeats were recognized to be derived from foreign genetic elements, like bacteriophages or plasmids, thereby suggesting a form of prokaryotic immune system (Mojica et al., 2005). Acquisition of new spacers within the CRISPR array mediates immunity against bacteriophages in Streptococcus thermophilus (Barrangou et al., 2007). The spacer sequences within a CRISPR structure was shown to interfere with horizontal gene transfer by plasmids indicating DNA targeting (Marraffini and Sontheimer, 2008). This was accompanied by the recognition of CRISPR giving rise to a pre-crRNA that is processed

into mature crRNAs, as well as description of the respective processing apparatus that is assembled by several cas genes and termed Cascade (Brouns et al., 2008). Conserved motifs that surround the proto-spacer (exogenous DNA that gives rise to a spacer) were hypothesized to prevent auto-immunity. The protospacer adjacent motif (PAM) sequence (e.g. NGG in the Streptococcus pyogenes CRISPR/Cas system), is only present on the invader's DNA, but not in the CRISPR array through which a microbial organism discriminates between an own and a foreign genetic element (Mojica et al., 2009).
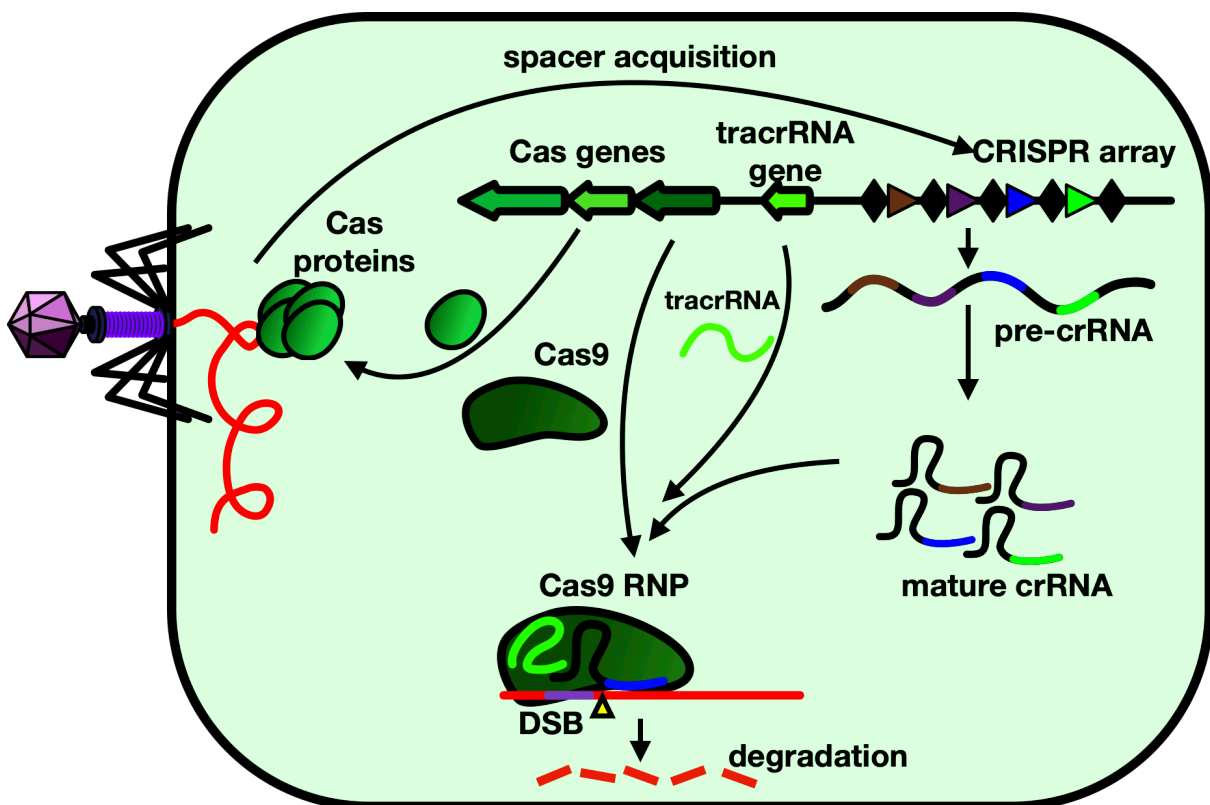
Fig. 5: CRISPR/Cas9 is an ancient adaptive immunity in bacteria and archaea. Cas proteins detect exogenous, viral DNA. In the process of spacer acquisition this machinery incorporates viral sequences with specific features (proto-spacers) as a spacer into a CRISPR array. The CRISPR array gives rise to a pre-crRNA that is processed into mature crRNAs. A tracrRNA gene gives rise to a tracrRNA that associates with a crRNA and the Cas9 protein to the functional Cas9 ribonucleoprotein complex (RNP). The mature Cas9 protein complex detects exogenous viral sequences by Watson-Crick base pairing of the crRNA with the viral sequence and mediates its cleavage by introducing a DSB (yellow triangle) (Wright et al., 2016).

It was found that the maturation of a functional crRNA requires another RNA molecule, the so-called transactivating crRNA (tracrRNA). Interestingly, the sequence of the

tracrRNA shows complementarity to parts of the crRNA thereby suggesting Watson-Crick base pairing between the two molecules (Deltcheva et al., 2011). In 2012, the history of CRISPR research culminated in the description of the CRISPR/Cas9 system. In that system, a crRNA harbors a variable spacer of 20 bp that interacts with its target DNA. In nature the targeted DNA may originate from a virus or a plasmid. The 3' end of the crRNA is derived from the spacer and associates with the tracrRNA. The crRNA/tracrRNA duplex associates with the Cas9 protein to build the functional Cas9 complex (Jinek et al., 2012). It was further shown that the crRNA/tracrRNA/Cas9 ribonucleoprotein complex is capable of cutting DNA templates *in vitro*. The researchers could also show that the same efficiency can be reached by fusing the crRNA and tracrRNA to one single chimeric RNA (Jinek et al., 2012). The knowledge that had been collected about CRISPR made it also possible to assemble the basic principle of CRISPR/Cas9-mediated bacterial immunity (Fig. 5).

1.4.3 The CRISPR/Cas9 system enables versatile and cost-effective genome editing in a variety of cell types

Soon after the uncovering of the RNA-guided endonuclease, several research groups published its application as a tool for genome editing in a variety of cells including hPSCs (Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013). The most widely used system is the Streptococcus pyogenes CRISPR/Cas9 system and encompasses the co-expression of a single-guide RNA (sgRNA) as well as the Cas9 enzyme. The PAM motif in that system is NGG. Upon binding of the sgRNA, Cas9 mediates a blunt cleavage of the DNA three base pairs upstream of the PAM sequence (Jinek et al., 2012) (Fig. 6). This system has been thoroughly characterized and the interaction of its cleavage domains (HNH and RuvC) with the DNA has even been investigated by a crystal structure study (Nishimasu et al., 2014). Modifications to optimize genome editing with the CRISPR/Cas9 system had been reported on different levels. For example, inactivation of one of the two endonuclease domains limits Cas9 cleavage to single strands (so called nicks). Compared to the classical Cas9 endonuclease, nickases have been described to yield significantly lower off-target effects (Ran et al., 2013a).

Fig. 6: The CRISPR/Cas9 system as a tool for genome editing. In the most widely used CRISPR/Cas9 system for genome editing the customizable part of the crRNA is fused with a scaffold RNA that corresponds to the tracrRNA of the bacterial adaptive immunity to a single-guide RNA (sgRNA). Upon Watson-Crick base pairing of the sgRNA to the targeted DNA sequence, the two endonuclease domains HNH and RuvC mediate a DSB. PAM: protospacer adjacent motif in orange, sgRNA sequence in red, scaffold RNA in purple. The DSB site is indicated by two red triangles.

Additionally, CRISPR/Cas9 systems from different bacterial species have been investigated. Promising results have been found with Francisella novicida Cpf1, which acts similar to Cas9 but relies on a crRNA without the need of a tracrRNA, yields sticky ends upon DNA cleavage and exhibits a different PAM specificity (Zetsche et al., 2015). Similarly, Neisseria meningitidis Cas9 (NmCas9) has been investigated and might exhibit a higher specificity due to longer protospacer and PAM sequences (Hou et al., 2013). Staphylococcus aureus Cas9 (SaCas9) turned out to be especially well-suited for gene therapy approaches. Due to its small size it can be packaged into an adeno-associated virus (AAV) along with an sgRNA. A single dose of SaCas9-containing AAVs resulted in efficient gene modification in mice (Ran et al., 2015). SgRNAs have also been subjected to modification. Truncated sgRNAs have been described to significantly reduce off-target activity (Fu et al., 2014) and chemically modified sgRNAs resulted in efficient genome editing of mouse hepatocytes *in vivo* (Yin et al., 2017). Single-stranded oligonucleotides (ssODNs) have already been found to be useful repair templates in ZFN-mediated genome editing and were also successfully used in CRISPR/Cas9-mediated genome editing (Chen et al., 2011; Ran et al., 2013b). This made it possible to investigate single-

nucleotide changes in an isogenic setting of even complex diseases like schizophrenia (Schrode et al., 2019). To further improve ssODN-mediated gene editing, phosphorothioate modifications yielded promising results (Renaud et al., 2016). Another interesting direction of research is a modified Cas9 enzyme, in which both endonuclease domains have been rendered inactive. Upon association with an sgRNA this so-called dead Cas9 (dCas9) has the ability to bind DNA, but will no longer mediate its cleavage. Fusing dCas9 to different co-factors made it possible to guide a functional module to a specific sequence of DNA without damaging it, instead eliciting a desired function like recruitment of repressors or activators. For example, fusing dCas9 to an activation domain like the hybrid tripartite activator VP64-p65-Rta (VPR) (Chavez et al., 2015), or to a repressor module like the Kruppel-associated box of Kox1 (KRAB) domain (Gilbert et al., 2013) leads to transcriptional activation (CRISPRa) or inhibition (CRISPRi) of targeted genes in a desired manner without changing the sequence of nucleotides. As sgRNA pools can easily be generated by oligonucleotide synthesis and incorporated into plasmids for virus production, it is also possible to use CRISPR/Cas9 in screening approaches (Shalem et al., 2014; Wang et al., 2014). Usually, in such an experiment, a pool of viruses is applied to a culture of cells, with each virus encoding a different sgRNA along with the Cas9 enzyme. Each cell of an initially homogenous culture will exhibit a different modification, thereby generating a pool of tens of thousands of cells with different knockouts. This culture can be subjected to a specific read out, like cell survival or the activation of a reporter construct. By purifying genomic DNA from the enriched culture, amplifying the sequences of the sgRNA pool by PCR and subjecting them to NGS, enriched knockouts can be identified by bioinformatical analysis.

As a summary, CRISPR/Cas9 technology and its improvements, were found to be a versatile tool in many fields of biomedicine, and combining CRISPR/Cas9 with hiPSCs is an interesting approach to decipher processes of neurodevelopment.

1.5 Aim of the study

GCNF is indispensable in embryonic development and contributes to germ cell differentiation. Its presence in the early neuroepithelium indicates that it might also be involved in neurodevelopment and neural stem cell formation. Expanding our knowledge of these processes might contribute to a better understanding of brain development, neural stemness and pluripotency. Considering that GCNF belongs to the nuclear receptor family, it might also serve as a target for pharmacological intervention. Small molecule agonists or antagonists of GCNF might be used to activate neural regeneration in cases of stroke or epileptic surgery. This thesis aims at deciphering how GCNF participates in the exit of pluripotency, germ layer specification, and the entry into neural development by utilizing hiPSCs as well as CRISPR/Cas9 technology. To that end several steps had been outlined, which will be addressed within this thesis:

1. *Establishing a CRISPR/Cas9 system to generate hiPSCs deficient for GCNF*

CRISPR/Cas9 offers a cost effective and powerful tool to manipulate genomes of almost any organism. HiPSCs provide the unique possibility to investigate processes of human brain development *in vitro*. Combining these two techniques might yield insights into human brain development without the need of animal experiments and within a reasonable time frame. Therefore, the first aim was the establishment of a CRISPR/Cas9 system to permanently knockout the GCNF gene in hiPSCs.

2. *Characterization of the resulting GCNF-ablated hiPSCs*

Apart from the fact that the procedure of genome editing might act as a noxious stimulus GCNF-ablation itself might interfere with pluripotency or genomic integrity of hiPSCs. To rule out any kind of damage, it is an absolute necessity to thoroughly characterize GCNF-ablated cells. After genotyping we wanted to analyze whether hiPSCs show pluripotency markers, as well as adequate morphology and growth properties. Subsequently, it was a very important part to confirm GCNF ablation on both RNA and protein level. The cells were also subjected to a SNP karyotyping to confirm genomic integrity after genome editing.

3. Impact of GCNF on germ layer specification with a focus on neural development

As GCNF might be involved in early neuroectodermal differentiation it seems plausible to assess the properties of ectodermal differentiation in an undirected differentiation experiment. Therefore, the cells were subjected to an embryoid body-based undirected differentiation experiment to investigate whether there is a bias towards a specific germ layer by qRT-PCR and a TaqMan-based Scorecard assay. One of the key questions to be addressed was whether GCNF-ablated hiPSCs show impaired neurodevelopment. To assess potential deficits in neurodevelopment we wanted to perform a classical dual-SMAD inhibition-based neural induction protocol for 48 h and 96 h. By utilizing a Next Generation Sequencing (NGS) based 3'-mRNA sequencing, we expected to get a first overall picture of how GCNF plays into the orchestration of neuroectodermal differentiation. Subsequent confirmation with qRT-PCR and immunocytochemistry was conducted to consolidate these findings.

2. Materials and methods

2.1 Cell culture

All cell culture work was performed in a laminar flow hood under sterile conditions. Cells were cultivated at 37 °C in a humified incubator with 5 % CO_2. Standard plasticware to cultivate cells were purchased from BD Falcon, Nunc, BD Biosciences, Greiner Bio-One, and PAA. When not stated otherwise cells were centrifuged in a Heraeus Megafuge 1.0R.

2.1.1 Cell lines

All experiments in this work have been performed with the hiPSC line iLB-C-133-bm-s4 which is an established hiPSC line in our institute. This male hiPSC line was generated with Sendai-Virus-mediated reprogramming from blood. All quality checks, like self-renewal and stable pluripotency markers, as well as genomic integrity, assessed by SNP analysis, were passed successfully. The cells were checked for absence of mycoplasma contamination with PCR on a regular base.

2.1.2 Reagents and media for cell culture work

All cell culture reagents (Tab. 1) and specific media for cell maintenance and differentiation experiments (Tab. 3) had been opened or prepared under sterile conditions in a laminar flow hood. Usually media was further supplemented with Penicillin/Streptomycin to reduce the risk of contamination. Small molecules, which were used in cell culture (Tab. 2), were usually obtained as solid powder and were prepared under sterile conditions.

Tab. 1: Media and cell culture solutions

Medium/Reagent	Manufacturer	Medium/Reagent	Manufacturer
2-Mercaptoethanol (50 mM)	Gibco*	Matrigel	Corning Life Science
B27 supplement (50X)	Gibco*	N2 supplement (100x)	GE Healthcare
Collagenase Type IV	Gibco*	Neurobasal	Gibco*
Dimethylsulfoxid (DMSO)	Sigma	Non-essential amino acids (NEAA)	Gibco*
DMEM/F12	Gibco*	PBS	Gibco*
EDTA	PanReac AppliChem	Penicillin/Streptomycin	Gibco*
Geltrex	Gibco*	StemMACS™ iPS-Brew XF	Miltenyi Biotec
KnockOut Serum Replacement (KSR)	Gibco*	StemPro Accutase	Gibco*
KnockOut™ DMEM	Gibco*		
L-Glutamine (200 mM)	Gibco*		

*Gibco is now part of Thermo Fisher Scientific

Tab. 2: Small molecules used in cell culture

Reagent	Manufacturer	stock concentration	solvent
basic fibroblast growth factor (bFGF)	R&D systems**	10 µg/ml	0.1% BSA in PBS
Collagenase IV	Gibco*	1 mg/ml	DMEM/F12
Heparin	Sigma-Aldrich	1 mg/ml	PBS
Insulin	Sigma-Aldrich	5 mg/ml	1% acetic acid
L-Ascorbic-Acid-2-Phosphat (LAAP)	Sigma-Aldrich	64 mg/ml	H_2O
LDN193184	Axon Medchem	2 mM	DMSO
Natrium-Selenit (NaSe)	Sigma-Aldrich	14 mg/ml (working stock: 14 µg/ml)	H_2O
Puromycin	Enzo Life Sciences	10 mg/ml	H_2O
ROCK-inhibitor Y-27632	BioTechne	10 mM	H_2O
SB431542	Axon Medchem	50 mM	DMSO
TAT-Cre	Merck Millipore	1500 units	
Transferrin	Sigma-Aldrich	10 mg/ml	H_2O
Transforming growth factor b1 (TGFβ1)	PeproTech	2 µg/ml	0.1% BSA in PBS

All stocks were stored at -20 °C.
* Gibco and Invitrogen are now part of Thermo Fisher Scientific.
** R&D systems is now part of BioTechne.

Tab. 3: Cell culture media for hiPSCs and differentiation paradigms

N2B27	
48.25%	DMEM/F12
48.25%	Neurobasal
0.5%	N2 supplement
1%	L-Glutamine
1%	B27 supplement
1%	Penicillin/Streptomycin

embryoid body medium	
79%	KnockOut™ DMEM
1%	Pen/Strep
20%	KSR
1%	NEAA
0.5%	L-Glutamine

hiPSC freezing medium	
90%	KSR
10%	DMSO

E8 medium	
99%	DMEM/F12
64 µg/ml	L-Ascorbic-Acid-2-Phosphat (LAAP)
14 n/ml	Natrium-Selenit (NaSe)
20 µg/ml	Insulin
10 µg/ml	Transferrin
1%	Penicillin/Streptomycin
100 ng/ml	Heparin
2 ng/ml	Transforming growth factor b1 (TGFβ1)
10 ng/ml	basic fibroblast growth factor (bFGF)

2.1.3 Coating

Either Matrigel or Geltrex were thawed overnight at 4 °C on ice and diluted in cold DMEM/F12 in a ratio of 1:30 or 1:60. Plastic dishes were coated with an appropriate volume of diluted Geltrex or Matrigel according to dish size. The covered plates were then stored at 4 °C until usage.

2.1.4 Cryopreservation and thawing of cells

To establish a cell bank, hiPSC cultures were grown until a confluency of ~ 90 %. Cells were incubated with Accutase for 10 minutes at 37 °C and washed with DMEM/F12. After centrifugation for 5 min at 1200 rpm, the cell pellet was resuspended in ice-cold hiPSC freezing medium (Tab. 3) and transferred to cryotubes. The cell suspension was frozen at -80 °C in CoolCell™ cryoboxes for at least 24 h before transfer to a -150 °C freezer for long-term storage. In order to thaw the cell suspension, the cryotube was warmed up at 37 °C in a water bath for 1 – 2 min until only a frozen clump remained. The cells were then transferred to a falcon tube containing DMEM/F12, followed by centrifugation for 5 min at

1200 rpm. The cell pellet was resuspended in 10 µM ROCK-inhibitor Y-27632 (RI) containing StemMACS™ iPS-Brew or E8 medium (Tab. 3).

2.1.5   Maintenance of human induced pluripotent stem cells

Proliferating hiPSCs were cultivated feeder-free on Geltrex- or Matrigel-coated 6-well plates with StemMACS™ iPS-Brew or E8 medium until they reached a confluency of ~ 90 %. The cells were either passaged with EDTA/PBS or Accutase. For the first method, cells were incubated 3 min with EDTA/PBS before dissociated colonies were transferred to freshly prepared Geltrex- or Matrigel-coated 6-well plates, using a 1000 µl pipette. When the cells were passaged as single cells, they were incubated with Accutase for 10 min at 37 °C before washing them with DMEM/F12, transferred to a 15 ml falcon tube and centrifuged for 5 min at 1200 rpm. The cell pellet was resuspended in 10 µM RI containing StemMACS™ iPS-Brew or E8 medium and plated on freshly prepared Geltrex- or Matrigel-coated 6-well plates.

2.1.6   Undirected embryoid body-based differentiation experiment

For the embryoid body-based undirected differentiation experiment hiPSCs were expanded in StemMACS™ iPS-Brew until they reached confluency. After incubation with 1 - 1.2 mg/ml Collagenase IV for 45 – 60 min, hiPSCs detached in a continuous layer. The layer of cells was carefully transferred to a falcon tube and centrifuged for 1 min at 800 rpm. The cells were resuspended in StemMACS™ iPS-Brew and transferred to uncoated petri dishes. Within the next 48 h, the cells formed aggregates that are commonly known as embryoid bodies. The embryoid bodies were collected, centrifuged at 300 rpm for 1 min and resuspended in embryoid body medium (Tab. 3) and transferred to fresh petri dishes. The cell aggregates were differentiated for 2 weeks in embryoid body medium before RNA was extracted from pooled embryoid bodies per cell line.

2.1.7   Neural induction experiment with dual-SMAD inhibition

For the neural induction experiment, 300,000 cells of non-edited and GCNF-ablated hiPSCs were seeded on 6-well plates for RNA extraction. Additionally, hiPSCs were

seeded on coverslips for immunocytochemistry. The cells were grown until they reached a confluency of around 90 % in StemMACS™ iPS-Brew before the medium was switched to neural induction medium, which consists of N2B27 medium (Tab. 3) supplemented with LDN193184 (400 nM) and SB431542 (10 µM) (Tab. 2), two small molecule inhibitors of SMAD signaling. The cells were cultivated for 48 h and 96 h before cells were subjected to RNA extraction using the Trifast reagent or fixed with paraformaldehyde for immunocytochemistry.

2.2   Cloning

Genome editing in hiPSCs encompasses two major tasks: Firstly, the generation of a Cas9 vector with a customized sgRNA to cause a DSB at a specific site in the genome. Secondly, the generation of a gene targeting vector that mediates homologous recombination at the site of the DSB to mediate incorporation of exogenous DNA at the respective locus. To perform the sgRNA customization of the Cas9 enzyme a widely used CRISPR/Cas9 vector (pSpCas9(BB)-2A-Puro (PX459) V2.0 [Addgene plasmid #62988]) was purchased from Addgene and amplified in E. coli. The small spacer sequence was exchanged with oligonucleotides corresponding to a specific site in the human endogenous GCNF locus. The sequences were determined with the CHOPCHOP web tool (https://chopchop.cbu.uib.no). To generate gene targeting vectors for both GCNF-ablation as well as GCNF-tagging, the human reference genome from Ensembl (http://www.ensembl.org) was utilized to retrieve the consensus DNA sequence of the GCNF gene. Homology arms were PCR amplified from human genomic DNA and cloned upstream and downstream of a PGK driven puromycin resistance cassette. In gene targeting vectors that delivered a tag, the corresponding sequence was cloned between the left homology arm and the puromycin resistance cassette. HiPSCs were then nucleofected with gene targeting vectors, along with respective sgRNA/Cas9 plasmids. All Sanger sequencing results are presented by the SnapGene® Viewer software (from GSL Biotech; available at snapgene.com).

2.2.1 Duplexing of oligonucleotides and sgRNA sequence customization of the Cas9 plasmid

The sgRNA customization of the Cas9 plasmid was performed as described previously (Ran et al., 2013b). Briefly, two oligonucleotides were duplexed to yield a short DNA fragment that encodes the sgRNA and exhibits overhangs to fit in the digested BbsI restriction site of pSpCas9(BB)-2A-Puro (PX459) V2.0 (addgene plasmid #62988). Generally, the standardized protocol to program the Cas9 plasmid encompassed the following steps: First, the two complementary oligonucleotides are duplexed by mixing equal amounts of the forward and reverse oligonucleotides (100 µM) and adding Polynucleotide kinase (PNK) as well as a PNK buffer. For efficient phosphorylation it is important that the buffer contains ATP. This might either be contained in the buffer or added separately. Subsequently, the mixture is subjected to the duplexing program in a thermocycler (Tab. 4). The resulting phosphorylated oligonucleotide duplex is diluted 1:200 and 1 µl of the dilution is added to a ligation reaction (16 °C overnight incubation) that includes the linearized and dephosphorylated pX459 vector (Tab. 4). The plasmid is then utilized to transform competent bacteria. The correct incorporation is finally confirmed by Sanger sequencing (Seqlab).

Tab. 4: Standardized duplexing procedure to program the Cas9 enzyme

Annealing and phosphorylation	
1 µl	Oligo FOR (100 µM)
1 µl	Oligo REV (100 µM)
1 µl	10X T4 Ligation Buffer (NEB)
0.5 µl	T4 PNK (NEB)
6.5 µl	H₂O

Thermocycler program for duplex annealing		
Step	Temperature	Time
1	37 °C	30 min
2	95 °C	5 min
3	25 °C	ramp down 5°C/min

Ligation in the pX459 vector	
50 ng	BbsI-digested plasmid
1 µl	Phosphorylated and annealed oligo duplex (1:200 dilution)
1 µl	T4 Ligation buffer (10X)
1 µl	T4 Ligase
to 10 µl	H₂O

These procedures were performed according to the recommendations of the Feng Zhang laboratory.

2.2.2 Generation of DNA fragments with PCR

Several PCR programs were utilized in this study. The standard PCR protocol, which was performed to obtain DNA constructs, was based on Q5 polymerase (Tab. 5) utilizing a standard PCR program (Tab. 6). For fragments that originated from difficult templates, like genomic DNA, a touchdown PCR program was utilized (Tab. 7). In a first round, a gradual decrease of annealing temperature enhances specific primer binding so that the specific PCR product is enriched, whereas the second round serves in amplifying the enriched DNA at a lower annealing temperature. The same touchdown PCR protocol utilized for obtaining fragments from genomic DNA was also used to perform genotyping for single cell-derived hiPSCs, that have been subjected to CRISPR/Cas9-mediated genome editing (see 2.3.2). The cloning of the right homology arm for the gene targeting vector was difficult and a different strategy had to be used. A larger fragment encompassing the intronic region that preceded and followed the first exon was amplified using the Phusion polymerase (Tab. 8, Tab. 9). This fragment served as a template for another PCR reaction, by which the right homology arm could be obtained (Tab. 10).

Tab. 5: 25 µl Q5 polymerase chain reaction

Volume	Final	Reagent
5 µl	1X	5X Q5 Reaction Buffer
2.5 µl	1 mM	10 mM dNTPs
0.33 µl	0.132 µM	10 µM Primermix
5 µl	1X	5X Q5 High GC Enhancer
0.25 µl	0.02 U/µl	Q5 polymerase
variable	< 1 µg	template
Up to 25 µl		H_2O

usually the mix was prepared with 10% of excess.

Tab. 6: Cycling conditions for Q5 standard polymerase chain reaction

Step		Temperature	Time
Initial denaturation		98 °C	1-5 min
Denaturation	35x	98 °C	30 s
Annealing		59 °C*	30 s*
Elongation		72 °C	60 s
Final Elongation		72 °C	5 min
Pause		10 °C	-

*primer dependent

Tab. 7: Cycling conditions for Q5 polymerase touchdown PCR

Step		Temperature	Time
Initial denaturation		98 °C	5 min
Denaturation	6x*	98 °C	60 s
Annealing		63 °C	30 s
Elongation		72 °C	5 min
Denaturation	35x**	98 °C	60 s
Annealing		57 °C	30 s
Elongation		72 °C	5 min
Final elongation		72 °C	10 min
Pause		10 °C	-

*with $\Delta T = -1$
**with $\Delta T = 0$

Tab. 8: 50 µl Phusion polymerase chain reaction

Volume	Final	Reagent
10 µl	1X	5X Phusion GC Buffer
1 µl	200 µM	10 mM dNTP
2.5 µl	500 nM	10 µM primer for
2.5 µl	500 nM	10 µM primer rev
variable	<1 µg	template DNA
1.5 µl	3%	DMSO
0.5 µl	1.0 unit	Phusion polymerase
Up to 50 µl	-	H_2O

Reaction performed according to the NEB protocol.

Tab. 9: Cycling conditions for Phusion polymerase larger fragments

Step		Temperature	Time
Initial denaturation		98 °C	5 min
Denaturation	45x	98 °C	60 s
Annealing		60 °C*	30 s*
Elongation		72 °C	220 s
Final Elongation		72 °C	5 min
pause		10 °C	-

*primer dependent

Tab. 10: Cycling conditions for Phusion polymerase right homology arm

Step		Temperature	Time
Initial denaturation		98 °C	5 min
Denaturation	35x	98 °C	60 s
Annealing		72 °C*	30 s*
Elongation		72 °C	120 s
Final Elongation		72 °C	5 min
pause		10 °C	-

*primer dependent

2.2.3 Restriction digestion, dephosphorylation and ligation of DNA fragments

To perform a restriction digest, the appropriate amount of DNA was diluted in H_2O and 1 µl of the specific restriction endonuclease was mixed with the respective 10X buffer. After incubation at 37 °C for 1 h, the digested fragment was separated on an agarose gel. The obtained band was excised and the linearized DNA was subjected to gel purification. The concentration of the purified DNA was estimated with a NanoDrop spectrophotometer (Thermo Fisher Scientific). Plasmid DNA, that served as a backbone, was further treated with 1 µl of Antarctic phosphatase (AnP) (and a respective amount of 10X AnP buffer) to prevent re-ligation without incorporation of the insert. The DNA was exposed to AnP for 30 min at 37 °C. Subsequently, the enzyme was heat inactivated by incubating it for 5 – 10 min at 80 °C. The linearized, dephosphorylated backbone was stored at -20 °C. Ligation was performed by incubation of the respective DNA fragments with 1 µl of T4 DNA ligase in the appropriate ligation buffer overnight at 16 °C.

2.2.4 Gibson Assembly to build gene targeting vectors

Gibson assembly was performed by utilizing the NEBuilder HiFi DNA Assembly. The NEBuilder online tool (https://nebuilder.neb.com/) was used to determine the overhangs best suited for the Gibson Assembly reaction. Several steps were performed to obtain a gene targeting vector for the exon 1 of the GCNF gene. First, the AAVS1-TRE3G-EGFP

(Addgene plasmid #52343) was linearized using the restriction enzymes SbfI and NotI, thereby leaving a backbone fragment that harbors an Ampicillin resistance gene, as well as the ori region. To obtain the sequence of the homology arms, the respective genomic region was searched in the Ensembl genome browser (Yates et al., 2020) and homology arms were cloned from genomic DNA using a touchdown PCR (see 2.2.2). The puromycin resistance cassette was assembled from the following sources: The PGK promoter was obtained from pLVX-pTight-Puro, a widely-used construct for lentiviral delivery of transgenes. The bGH polyA sequence as well as the puromycin resistance gene was amplified from the AAVS1-TRE3G-EGFP plasmid. The PCR fragments were obtained by using a classical PCR (Tab. 6). The Gibson Assembly was performed following the manufacturer's instructions. Usually, 0.05 pmol per fragment were mixed in a volume of 10 µl. The NEBuilder HiFi DNA Assembly Mastermix was added in a ratio of 1:1 to a final volume of 20 µl. The mix was vortexed and centrifuged followed by an incubation at 50 °C for 1 h in a thermocycler. 1 – 5 µl of the assembly preparation was then used to transform competent bacteria. Single cell-derived clones were amplified and analyzed either by restriction digestion or colony PCR. Promising candidate clones were further amplified and subjected to Sanger sequencing. A similar strategy was used for cloning the gene targeting vectors for C-terminal tagging of the endogenous GCNF gene. A spacer, that harbors BbsI restriction sites, was included between the left homology arm and the loxP-flanked puromycin cassette. This allows for an efficient exchange of C-terminal tags.

2.2.5 Transformation of competent bacteria

Transformation of bacteria was done by standardized procedures and recipes (Tab. 11, Tab. 12, Tab. 13). Briefly, competent bacteria (either house-made or NEB DH5α E. coli) were thawed on ice. 1 – 5 µl of ligation preparation was added, bacteria were gently mixed and incubated on ice for 30 min. In the next step, a heat shock was performed at 42 °C for 45 s. Subsequently, cells were incubated on ice for 2 min and 500 µl prewarmed LB was added. After incubation at 37 °C for 1 h in a heating block while shaking at around 300 rpm the bacteria suspension was seeded on ampicillin containing Agar plates. After an overnight incubation at 37 °C bacteria colonies were readily visible. Single colonies were picked for further amplification in LB medium containing 100 mg/ml ampicillin. To

perform a miniprep 5 – 6 ml of LB medium was inoculated and incubated at 37 °C overnight. The following day, the bacteria suspension was lysed and DNA was prepared using the peqGOLD plasmid miniprep Kit. The DNA concentration of the miniprep preparation was estimated by NanoDrop. To check for the correct integration of an insert either a plasmid digest or a colony PCR was performed. Finally, the correct sequences of chosen clones were confirmed by Sanger sequencing. To establish long term storage of plasmids, glycerol stocks were prepared by mixing 300 µl of bacteria solution with 300 µl of autoclaved glycerol. Glycerol stocks were stored at – 80 °C. Usually, once the desired sequence of a plasmid was verified, a higher volume of bacterial cultures was prepared and plasmid DNA was purified by utilizing a PureLink™ HiPure Plasmid Maxiprep Kit or Midiprep Kit.

Tab. 11: Reagents and recipes for DNA amplification in bacteria

LB Medium:	
10 g	Tryptone
5 g	Yeast extract
10 g	NaCl
1 ml	NaOH solution (1 M)
Fill H$_2$O to 1000 ml.	
After Autoclaving store it at 4 °C	

Ampicillin
100 mg/ml Ampicillin sodium salt in H$_2$O

Bacteria glycerol stock
Mix 300 µl Bacteria solution with 300 µl glycerol, store at – 80 °C

LB Agar plates
15 g Agar have to be dissolved in 1000 ml of LB and subjected to autoclaving.
For agar plate preparation, heat Agar until it is completely dissolved. When the solution reaches 50 °C, add Ampicillin to a final concentration of 100 µg/ml.
Pour the solution into sterile petri dishes and let it solidify at RT. Store the prepared plates in sealed plastic bags at 4 °C.

All components purchased from Sigma-Aldrich or Roth.

Tab. 12: Solutions for agarose gel electrophoresis

50x TAE buffer	
242 g	Tris
100 ml	EDTA (0.5 M sol., pH 8.5)
57 ml	Acetic Acid (100%)
Add H$_2$O to 1000 ml.	

All components were purchased from Sigma-Aldrich.

6x DNA loading buffer	
3 ml	Glycerol
0.25 mg	Bromophenol blue
0.25 mg	Xylene cyanol
Add H$_2$O to 10 ml.	

Tab. 13: Reagents and kits for cloning

Reagent/Kit	Manufacturer	Reagent/Kit	Manufacturer
Agarose	Peqlab**	Plasmid miniprep Kit, peqGOLD	Peqlab**
Antarctic Phosphatase (AnP)	NEB	PureLink™ HiPure Plasmid Maxiprep Kit	Invitrogen*
DNA ladders (100 bp & 1 kb)	Both ladders from NEB and Peqlab**	PureLink™ HiPure Plasmid Midiprep Kit	Invitrogen*
DNeasy Blood & Tissue Kit	Qiagen	Q5 High-Fidelity DNA Polymerase	NEB
dNTPs	Peqlab**	Restriction enzymes	NEB
Gel extraction Kit, peqGOLD	Peqlab**	T4 ligase	NEB
Maxwell®RSC Blood DNA Kit	Promega	T4 Polynucleotidkinase	NEB and Thermo Fisher Scientific
NEB DH5a	NEB	Taq DNA Polyermase	Invitrogen*
NEB stable	NEB		
NEBuilder HiFi DNA Assembly Master Mix	NEB		
Phusion High-Fidelity PCR Kit	NEB		

*Invitrogen is now part of Thermo Fisher Scientific.
**Peqlab is now part of VWR.

2.2.6 Insert confirmation

Insert confirmation was either done by a test restriction digestion or by a colony PCR. Generally, a single test digestion reaction consisted of ~ 2 µl of purified plasmid DNA obtained by a miniprep, 0.5 µl of a specific restriction enzyme suited to test for the presence of the correct sequence, 2 µl of a respective buffer, and H_2O filled to a total reaction volume of 20 µl. The reaction mix was incubated at 37 °C for 1 h, before 6X loading buffer was added and the mixture was loaded on a gel. Colonies that showed the correct restriction fragment pattern was subjected to further analysis. To perform a colony PCR, colonies were picked and thoroughly diluted in 20 µl of water. Depending on how many clones were subjected to analysis, a Taq-polymerase containing master mix was prepared that included primers that were present in both backbone and insert (Tab. 14). 19 µl of master mix and 1 µl of bacterial solution were mixed and a specific PCR program was applied (Tab. 15). Subsequently, the PCR reactions were mixed with 4 µl of 6X loading buffer and run on a gel. Positive colonies were subjected to further analysis.

Tab. 14: Mastermix for one colony PCR reaction

Volume	Reagent
2 µl	10X PCR Buffer
0.6 µl	50 mM MgCl$_2$
0.4 µl	10 mM dNTPs
13.9 µl	H$_2$O
0.1 µl	Taq polymerase
2 µl	Primermix 10 µM
1 µl	Bacterial solution

Tab. 15: Cycling conditions for colony PCR

Step		Temperature	Time
Initial denaturation to break bacterial cell membranes		95 °C	5 min
Denaturation		95 °C	30 s
Annealing	35x	60 °C	30 s
Elongation		72 °C	30 s
Final elongation		72 °C	10 min
Pause		10 °C	-

2.3   Genome editing with CRISPR/Cas9

Genome editing relies on transferring a plasmid encoding for Cas9 and a customizable sgRNA along with a gene targeting vector to a cell line. The cloning of these components is discussed in the last section (see 2.2), whereas this section covers the use of the CRISPR/Cas9 constructs in cell culture.

2.3.1   Nucleofection and selection

The nucleofections in this study were performed with the Lonza 2D Nucleofector. HiPSCs were incubated with RI containing StemMACS™ iPS-Brew for approximately 60 min. During this incubation time, the nucleofection reagents were mixed. The basis of the nucleofection mix consists of 82 µl of Lonza Nucleofection solution and 18 µl of Lonza supplement 1. Two plasmids were added to the nucleofection mix. The first plasmid encodes the Cas9 enzyme, and a puroymycin N-acetyltransferase, that mediates transient resistance against puromycin. Additionally, it encodes an U6-driven, customizable sgRNA sequence. The second plasmid is the gene targeting vector that mediates incorporation of exogenous DNA into the genome of nucleofected hiPSCs. A pmaxCloning™ plasmid may be added to monitor the nucleofection efficiency. The pmaxCloning™ plasmid does not integrate, is only transiently active and is diluted out of the culture after around 2 days. The RI-incubated cells were dissociated with Accutase for 15 min at 37 °C, harvested and diluted in DMEM/F12. Around 2.5 mio cells were

transferred to a 2 ml Eppendorf tube and centrifuged for 4 min at 1200 rpm. The hiPSCs were carefully resuspended in the nucleofection mix. The nucleofection mix was transferred to an electroporation cuvette, which was subjected to program B-23 of the Lonza 2D Nucleofector. This program has been optimized to nucleofect hiPSCs. After that, hiPSCs are resuspended in antibiotic free StemMACS™ iPS-Brew supplemented with 10 µM RI. Finally, cells were plated on fresh Geltrex- or Matrigel coated 6-well plates. A medium change with RI-supplemented StemMACS™ iPS-Brew was done 24 h later. Pen/Strep was either added to the cell culture medium after 24 h or at 48 h of nucleofection. 48 h after the cells had been subjected to nucleofection the medium was switched to StemMACS™ iPS-Brew supplemented with RI and 0.5 µg/ml puromycin, which was maintained for 7 days. During the selection for edited cells, puromycin-resistant clones that had undergone successful editing events emerged. After 7 days of selection, polyclonal confluent cultures were obtained, which could be used for preliminary experiments, although it is recommended to subject the cells to single cell cloning in order to derive monoclonal cell lines.

2.3.2   Genotyping and SNP analysis of monoclonal genome edited cell line

To derive single cell clones, the cells were seeded on Geltrex-coated 10 cm dishes in single cell density in StemMACS™ iPS-Brew, supplemented with RI. The cells were switched to non-RI containing medium, when groups of 20 – 100 cells appeared that were presumably derived from a single cell. 24 h after RI withdrawal, the loose cells gathered in small colonies. When the colonies were grown to an appropriate size, they were picked with a pipette and transferred to single wells of a 48 well plate in RI-containing StemMACS™ iPS-Brew medium. Upon further expansion, genomic DNA of single cell-derived clones was either won by lysing cells with Quick extract (Biozym) or by preparation of high-quality DNA. For Quick extract DNA preparation, the cell suspension was diluted in Quick extract reagent and treated 5 min at 65 °C and 10 min at 95 °C. To obtain high-quality genomic DNA, cells were collected and DNA was extracted with the DNeasy Blood & Tissue Kit (Qiagen) or a Maxwell® RSC Instrument. The prepared DNA was subjected to genotyping, utilizing a touchdown PCR (Tab. 7). Cells that showed the aimed genotype were expanded and cryo vials were frozen as backups. High-quality DNA was handed

over to the genomics department where single nucleotide polymorphism (SNP) were assessed by an Infinium OmniExpressExome (Illumina) chip. The obtained data sets were then analyzed to exclude larger genomic aberrations.

2.4 RNA-based expression analysis

To prevent RNA degradation by ubiquitously present RNAses DEPC-containing RNAse-free H_2O, as well as disposable plastic ware and surface RNAse decontaminants (RNAse-ExitusPlus) have been used during this work.

2.4.1 RNA isolation

RNA was isolated with the peqGOLD TriFast (Peqlab – VWR) reagent according to the manufacturer's instruction with small modifications (further reagents see Tab. 16, Tab. 17). The isolated RNA was primarily subjected to reverse transcription and qRT-PCR analysis. One set of RNA samples was also analyzed in a 3'-mRNA sequencing experiment. Briefly, the cells were washed at least one time with PBS to remove dead cells and debris that might interfere with subsequent analysis. The washed cells were lysed with 1 ml of TriFast and transferred to RNAse-free 1.5 ml Eppendorf tubes. The resulting suspension could be stored at - 80 °C before RNA preparation. The thawed samples were mixed with 200 µl of chloroform, incubated for 15 min at RT and after 10 min of centrifugation at 12,000 rpm (Centrifuge 5415R, Eppendorf) three phases were separated within the tube. The aqueous upper phase, that contains the RNA, was transferred to a fresh RNAse-free tube. 500 µl isopropanol and 70 µg/ml glycogen were added to precipitate the RNA, which was followed by an incubation for 1 h on ice or overnight at -20 °C. Subsequently, the RNA was precipitated by centrifugation at 12,000 rpm for 10 min and two washing steps with 800 µl of cold 75 % ethanol in DEPC-H_2O followed. The washed pellet was dried for up to 1 h at RT and then carefully resuspended in 20 µl DEPC-H_2O. To obtain pure RNA samples, the samples were treated with DNAse I (Invitrogen) for 15 min. The DNAse I enzyme was inactivated with EDTA and denatured for 10 min at 65 °C. The generated samples were stored at -80 °C.

Tab. 16: Kits and reagents for RNA extraction

Tab. 17: Recipe for DEPC-H$_2$O

Reagent/Kit	Manufacturer
Chloroform	Roth
DEPC	Sigma-Aldrich
DNAseI Amplification Grade Kit	Invitrogen*
Ethanol for molecular biology	Roth
Glycogen (35 mg/ml)	Peqlab (VWR)
Isopropyl alcohol	Roth
PeqGOLD TriFast	Peqlab (VWR)
RNAase-ExitusPlus	Labomedic

Preparation
Dissolve DEPC to a final concentration of 1 ml/l in H$_2$O
Light-protected overnight incubation while stirring under the hood with lid open
next day autoclaving

*Invitrogen is now part of Thermo Fisher Scientific.

2.4.2 Quantitative RT-PCR analysis

A SYBR Green-based qRT-PCR was utilized to quantify RNA-expression levels. To that end, complementary DNA (cDNA) was generated from the isolated RNA by reverse transcription with the iScript cDNA Synthesis Kit (Bio-Rad) following the manufacturer's instruction. The resulting cDNA was diluted to 150 ng/µl. The PCR was performed in an Eppendorf Realplex Mastercycler with a SYBR Green-based reaction mix (Tab. 18). A master mix was prepared for the reaction that included Taq polymerase as well as the SYBR Green based reaction mix (Tab. 19). After preparing mixes specific for each primer pair and dispensing them on a 96-wellplate (4titude Brooks Life Sciences) 300 ng of cDNA in a volume of 2 µl was added to each reaction (Tab. 19). Technical triplicates were prepared to ensure consistency of the measurements. The PCR itself was performed following a standardized scheme (Tab. 20). The analysis of the resulting data was done by the comparative $\Delta\Delta$Ct method (Livak and Schmittgen, 2001). The mRNA expression levels were normalized to 18S rRNA reference levels. In the qRT-PCR reaction either primers that had been thoroughly characterized in preceding projects were used or a primer testing was performed. The primer testing included a serial dilution assay, subsequent melting curve analysis and gel electrophoresis to confirm product specificity. Primer design was achieved with the Primer3web online tool. MRNA specificity was

achieved by a primer design that yielded PCR products spanning exon-exon junctions. As opposed to genomic DNA, mRNA derived sequences had undergone splicing and therefore are lacking intronic sequences that would interfere with the generation of the product. To account for the 3'-end synthesis bias of the reverse transcription reaction, primers (Tab. 33) were usually designed to bind in the last exons.

Tab. 18: Recipe for 2x qRT-PCR mix (for 10 ml stock)

Component	Manufacturer	Amount
10x PCR buffer (without MgCl$_2$)	Invitrogen	2 ml
MgCl$_2$ Solution (50 mM)	Invitrogen	1.2 ml
dNTPs (each 100 mM)	Peqlab (VWR)	40 µl each
SYBR Green I nucleic acid gel stain (1000x)	Sigma-Aldrich	15 µl
Fluorescein calibration dye (100 µM)	Bio-Rad	2 µl
H$_2$O		6.623 ml

Tab. 19: Reaction set up for Taq qRT-PCR

Component	Amount
2x qRT-PCR mix	10 µl
Taq polymerase	0.12 µl
cDNA (150 – 500 ng/µl)	2 µl
Forward/reverse primer (3 µM each)	0.8 µl
Add H$_2$O to a final volume of 20 µl	

Tab. 20: Cycling conditions for Taq qRT-PCR

Step		Temperature	Time
Initial denaturation		95 °C	3 min
Denaturation	40-45x	95 °C	15 s
Annealing		60 °C	20 s
Elongation		72 °C	30 s
Final elongation		72 °C	10 min
Dissociation curve		55-95 °C	

2.4.3 3'-mRNA sequencing

To perform 3'-mRNA sequencing, RNA samples were diluted to 150 ng/µl and handed over to the Next Generation Sequencing facility. Upon library preparation, the NGS reaction was performed on an Illumina HiSeq 2500 with 10 million single-end reads per sample. After receiving, the FASTQ files were analyzed with R (detailed description of the subsequent analysis, see 2.7.3).

2.5 Western blot analysis

To prepare cell lysates, the medium was removed and the cell culture plate was washed once with PBS. Cold PBS was added and the cells were detached using a cell scraper (Cell lifter, Corning). The cell suspension was collected in a 15 ml Falcon tube on ice. The Falcon tubes were then centrifuged 5 min at 2000 g (Heraeus Megafuge 40R) to pellet the cells. For cell lysis, the pellets were resuspended in RIPA buffer supplemented with protease inhibitor (Tab. 21), transferred to an Eppendorf tube and gently shaken for 30-60 min on ice. The lysate was centrifuged at 12,000 rpm for 15 min at 4 °C (Eppendorf Centrifuge 5415R). The supernatant contained the protein suspension, which was transferred to a fresh Eppendorf tube, frozen on dry ice and stored at -80 °C. To estimate protein concentration a BCA assay was performed (Thermo Fisher Scientific). 40 µg of protein was diluted in fresh Eppendorf tubes and 5xLaemmli buffer (Tab. 21) was added. The mixture was mixed thoroughly and heat-denatured at 95 °C for 10 min. It was then either immediately loaded on a gel or stored at -20 °C.

2.5.1 SDS polyacrylamide gels

For the analysis of proteins, they were separated by their size via SDS polyacrylamide gel electrophoresis (SDS-PAGE). In detail, a Mini-PROTEAN vertical electrophoresis system (Bio-Rad) was utilized. Gels were prepared according to a standard protocol (Tab. 22) and consists of a stacking and a separation gel (10 % SDS). 40 µg of protein per sample were loaded on the gel, next to 10 µl of a size marker (Color Plus Prestained Protein Ladder 10-230 kDa, NEB). Gel electrophoresis was performed in running buffer (Tab. 23) at 100 V for 2 h.

Tab. 21: Reagents for preparing protein lysates

RIPA buffer	
Tris HCl	50 mM (pH 7.5)
EDTA	1 mM
Deoxycholic acid sodium salt	0.5%
NaCl	150 mM
SDS	0.1%
Igepal CA-639	1%
Protease inhibitor cocktail	1:100

5x Laemmli buffer	
Tris HCl	312.5 mM (pH 6.8)
SDS	10%
Glycerol	50%
Bromophenol blue	0.1%
2-Mercaptoethanol (freshly added)	10%

The Halt Protease Inhibitor Cocktail was purchased from Thermo Fisher Scientific, all other reagents were purchased from Roth or Sigma-Aldrich.

Tab. 22: Recipes for self-made SDS polyacrylamide gels

Separation gel (10 ml)	
H_2O	4 ml
Acrylamide Mix (30 %)	3.4 ml
Tris HCl (pH 8.8)	2.6 ml
SDS (10 %)	100 µl
APS (10 %)	100 µl
TEMED	4 µl

Stacking gel (2 ml)	
H_2O	1.15 ml
Acrylamide Mix (30 %)	330 µl
Tris HCl (pH 6.8)	500 µl
SDS (10 %)	20 µl
APS (10 %)	20 µl
TEMED	2 µl

All reagents were purchased from Roth or Sigma-Aldrich.

Tab. 23: Recipes for Western blotting

10x Running buffer (in H_2O)	
Trizma-Base	25 mM
Glycin	193 mM
SDS	0.1% (w/v)

1x Transfer buffer (in H_2O)	
Trizma-Base	2.5 mM
Glycin	19.3 mM
Methanol	20%

Stripping buffer (in H_2O)	
SDS	2% (w/v)
Tris HCl	62.5 mM (pH 6.7)
2-Mercaptoethanol	2 mM (freshly added)

Ponceau staining solution (in H_2O)	
Ponceau S	0.1% (w/v)
Acetic Acid	5 %

TBS-(T) (in H_2O)	
Tris HCl	10 mM (pH 7.5)
NaCl	150 mM (TBS_{50}: 50 mM)
(Tween20)	0.1%

All reagents were purchased from Roth or Sigma-Aldrich.

2.5.2   Protein transfer to PVDF membranes and protein detection

The separated proteins were transferred to a methanol-activated PVDF membrane by using the wet blotting technique with a Mini Trans-Blot Cell (Bio-Rad). Blotting was performed in transfer buffer (Tab. 23) at 70 V for 2 h during which constant stirring and cooling by ice was applied. Subsequently, the successful transfer of proteins was confirmed by a Ponceau staining (Tab. 23) of the membrane. Western blot analysis with antibodies (Tab. 34) was performed with specific conditions (Tab. 24). Blocking of the membrane was performed in 10 % milk powder in TBS-T (Tab. 23) for 1 h at RT. The primary antibody was diluted in TBS-T with 5 % milk powder and depending on the antibody incubated for 2 h at RT or overnight at 4 °C. After three washing steps with TBS-T, the horseradish peroxidase (HRP)-linked secondary antibody was applied in TBS-T and 5 % milk powder for 1 h at RT. After three more washing steps with TBS-T, substrates for the HRP were added to the membrane. Depending on the expected protein concentration, Luminata Classico or Forte Western HRP substrates (Merck Millipore) were used. The proteins were detected by analyzing the chemiluminescence of the HRP-catalyzed reaction of the substrate with a Chemiluminometer (Bio-Rad). Various proteins can be analyzed on the same membrane using antibodies of the same species. To reapply antibodies for different proteins, the remaining antibodies need to be removed by using stripping buffer (Tab. 23), which was supplemented with fresh 2-Mercaptoethanol, and incubation at 55 °C for 15 min. Afterwards, blocking and incubation with respective antibodies could be performed again.

Tab. 24: Condition used for GCNF and β-Actin Western blot analysis

Condition	anti-GCNF antibody	Condition	anti-β-Actin antibody
Blocking	10% milk powder in TBS-T, 1 h at RT	Blocking	10% milk powder in TBS-T, 1 h at RT
Primary antibody solution	anti-GCNF mouse antibody, 1:1000 in 5% milk powder in TBS-T, overnight at 4 °C	Primary antibody solution	anti-β-Actin mouse antibody, 1:5000 in 5% milk powder in TBS-T, 1 h at RT
Washing	TBS-T	Washing	TBS-T
Secondary antibody solution	HRP-goat anti-mouse, 1:1000 in 5% milk powder in TBS-T, 2 h at RT	Secondary antibody solution	HRP-goat anti-mouse, 1:1000 in 5% milk powder in TBS-T, 1 h at RT

2.6   Immunocytochemistry

To perform immunocytochemistry the cells were usually grown on plastic coverslips (Thermo Fisher Scientific). The medium of cultured cells was removed and washed once with PBS, before they were fixed with 4 % PFA for 10 min (ready-to-use PFA, purchased from Thermo Fisher Scientific) or 20 min (in-house made PFA, powder purchased from Sigma-Aldrich, Tab. 25). After PFA was removed and the plates have been washed with PBS, they were stored at 4 °C in PBS containing 0.1 % sodium azide (PBS azide) or were immediately subjected to the staining procedure. Initially, cells were permeabilized with 0.5 % Triton-X-100 in PBS azide for 20 min and then incubated in blocking solution (Tab. 25) for 1 - 2 h at RT. Next, the primary antibodies (Tab. 34) in 5 % FCS blocking solution were applied for either 2 h at RT or overnight at 4 °C. After primary antibody incubation, cells were washed three times with PBS and then incubated with the secondary antibody (Tab. 35) in 5 % FCS blocking solution for 1 – 2 h at RT. Finally, nuclei were stained by DAPI (4',6-diamidino-2-phenylindole, 1:10,000 in PBS) exposure of 2 min. After three more washing steps, the coverslips were transferred to microscope slide and embedded in Mowiol (Tab. 25). Chemical substances used within this protocol were purchased from Sigma-Aldrich or Merck (if not stated otherwise).

Tab. 25: Recipes for immunocytochemistry

PFA 4 % solution
40 g PFA dissolved in 1000 ml PBS
Carefully heat the solution until it is dissolved
Filtering and adjusting to pH 7.4
Aliquot in 50 ml Falcon Tubes and store them at -20 °C

2 N HCl
167 ml HCl to 833 ml H_2O

Blocking solution	
FBS	5-10%
Triton	0.1-0.5%
Sodium azide	1% (w/v)

Mowiol*
2.49 g Mowiol and 6 g Glycerol have to be dissolved in 6 ml H_2O
12 ml of 0.2 M Tris HCl with a pH of 8.5 have to be added
10 min heating to 50 °C
Centrifugation
Aliquoting in 2 ml Eppendorf tubes and storing at -20 °C

*purchased from Roth

2.7 In silico analysis and bioinformatics

2.7.1 Software and online tools

Software that has been used in this study is listed in the following table (Tab. 26).

Tab. 26: Software and online tools

Program/Database	Manufacturer/Source
Affinity Designer Version 1.8.4	Serif
ApE – A plasmid Editor v2.0.61	M. Wayne Davis
CHOPCHOP web tool	https://chopchop.cbu.uib.no (Labun et al., 2016)
clusterProfiler R package	Yu et al., 2012
DESeq2 R package	Love et al., 2014
Endnote X9.3.3	Clarivate Analytics
Ensembl Genome Browser	http://www.ensembl.org/index.html
Illumina Genome Studio V2011.1	Illumina
ImageJ Version 2.0.0-rc-69/1.52p	NIH (Schneider et al., 2012)
JASPAR online tool	Fornes et al., 2020
Keynote Version 10.1 (6913)	Apple
Microsoft Office 2018 for Mac	Microsoft
NEBuilder Assembly tool	https://nebuilder.neb.com/
Numbers Version 10.1 (6913)	Apple
Primer3web Version 4.1.0	https://primer3.ut.ee
PubMed	https://pubmed.ncbi.nlm.nih.gov
Salmon software	Patro et al., 2017
SnapGene® Viewer software	GSL Biotech LLC (https://www.snapgene.com)

2.7.2 Statistical analysis

Most of the experiments were generated as biological replicates that resemble independent experiments. Graphs usually show results as mean + standard deviation (SD), which was computed with Apple Numbers. Statistical significance was analyzed by two-tailed Student's t-test using spreadsheet analysis with Apple Numbers for non-edited control as well as edited experimental conditions. A p-value of ≤ 0.05 was regarded as statistically significant.

2.7.3 Next-Generation Sequencing based 3'-mRNA sequencing

The raw reads containing FASTQ files were provided by the Next Generation Sequencing Core Facility. The downloaded datasets were subjected to transcript abundance estimation by utilizing the Salmon software (Patro et al., 2017). This methodology utilizes an index transcriptome to obtain read quantification. The further analysis was performed by utilizing the well-established R pipeline of the DESeq2 routine (Love et al., 2014). The clusterProfiler R package was used to perform a gene ontology analysis (Yu et al., 2012).

2.8 Supplementary lists

2.8.1 Technical equipment

Tab. 27: Technical equipment

Device	Name	Manufacturer
Electroporator	Amaxa 2D Nucleofector	Lonza
Centrifuge for bacteria pelleting	Sorvall RC 6+ Centrifuge	Thermo Fisher Scientific
Centrifuge for cell culture work	Heraeus Megafuge 1.0R	Thermo Fisher Scientific
Centrifuge	Heraeus Megafuge 40R	Thermo Fisher Scientific
Centrifuge	Centrifuge 5415R	Eppendorf
Chemiluminometer	ChemiDoc	Bio-rad
Fluorescence microscopy inverse	Zeiss AxioObserver Z1	Zeiss
Fluorescence microscopy inverse	Axiovert 200M	Zeiss
Gel documentation	Geldoc2000	Bio-Rad
Gel electrophoresis system used for DNA separation	Agargel	Biometra
Gel electrophoresis system used for Protein separation	Mini-Protean (vertical gel chamber), Mini-Trans-Blot Cell (wet blotting cell)	Bio-Rad
Horizontal flow hood for single cell picking	HERAguard	Thermo Fisher Scientific
Incubator for cell culture	HERAcell	Thermo Fisher Scientific
Light microscope (inverse)	Axiovert 25	Zeiss
Nucleic acid purification platform	Maxwell® RSC Instrument	Promega
Real-Time PCR System	Mastercycler realplex	Eppendorf
Spetrophotometer	NanoDrop 1000	Peqlab (VWR)

Sterile laminar flow hood	HERAsafe	Thermo Fisher Scientific
Thermocycler	T3000 Thermocycler	Biometra

2.8.2 Primers and oligonucleotides for cloning

Tab. 28: Primers and oligonucleotides to program the Cas9 plasmid

Target	Recipient plasmid	Name	Sequence (5'->3'; overhangs for BbsI cutting site in **bold**)
hGCNF Exon 1	pX459*	NR6A1 KOE1 g4 FOR	**CACC** GACGAACCGCCGCCTAGCGG
hGCNF Exon 1	pX459*	NR6A1 KOE1 g4 REV	**AAAC** CCGCTAGGCGGCGGTTCGTC
hGCNF Exon 1	pX459*	NR6A1 C-term g81 FOR	**CACC** GGCAAGGAATGACCTGTTCC
hGCNF Exon 1	pX459*	NR6A1 C-term g81 REV	**AAAC** GGAACAGGTCATTCCTTGCC

*addgene plasmid #62988 pSpCas9(BB)-2A-Puro (PX459) V2.0
(https://www.addgene.org/62988/)

Tab. 29: Primers and oligonucleotides to clone the human GCNF knockout gene targeting vector

Target	Name	Sequence (5'->3')
hGCNF	NR6A1_KOE1_HAL FOR	CCCTCACTAAAGGGACTAG CTAGTCCTGCAGGAAATGA AAGGGGAGGAGGGG
hGCNF	NR6A1_KOE1_HAL REV	GTATGCTATACGAAGTTAT CTAGGCGGCGGTTCGTCC
hGCNF	NR6A1_KOE1_PuroR FOR	CGAACCGCCGCCTAGATAA CTTCGTATAGCATACATTAT ACGAAGTTATCAATTGTACT TAGGATCCC
hGCNF	NR6A1_KOE1_PuroR REV	CTTAGGATCCCTCGAGCGG CCTAGTCAATAATCAATG
hGCNF	NR6A1_KOE1_HAR+mutPAM FOR	TTATTGACTAGGCCGCTCG AGGGATCCTAAGTACAATT GTCTGGTACGCGTATAACT TCGTATAGCATACATTATAC GAAGTTATCGGAGTGGGA GGCGGCGGGGGCTC
hGCNF	NR6A1_KOE1_HAR+mutPAM REV	CGACTCACTATAGGGCGAA TTGAATTTAGCGGCCGCGG CGTTGAGGAAGGTGAGGC GCAG

| hGCNF | NR6A1_KOE1_Intron FOR | AACTGAGGAAAGGGGAAG GC |
| hGCNF | NR6A1_KOE1_Intron REV | GACGAGCGGTAGGAAGGG |

The backbone is derived from Addgene plasmid #52343 AAVS1-TRE3G-EGFP (https://www.addgene.org/52343/)

Tab. 30: Primers and oligonucleotides to clone the human GCNF C-terminal knock-in gene targeting vector

Target	Name	Sequence (5'->3')
hGCNF	NR6AI_HALwoStop+IIS-RS FOR	ACTAAAGGGACTAGCTAGT CCTGCAGGGTTGCACCTG GAGAATTCTG
hGCNF	NR6AI_HALwoStop+IIS-RS REV	ATATCCGCGAAGACCCTTC CTTGCCCACACTGGTC
hGCNF	revPGK-PuroR-bGH+loxP FOR	TGTGGGCAAGGAAGGGTC TTCGCGGATATCATAGAAG ACCTATAACTTCGTATAGC ATACATTATACGAAGTTATC GGCCTAGTCAATAATCAAT GTCGAG
hGCNF	revPGK-PuroR-bGH+loxP REV	CTTAGGATCCCTCGAGGG GTAGGGGAGGCGCTTT
hGCNF	NR6AI_HARmutPAM+MCSlox P FOR	GCCTCCCCTACCCCTCGAG GGATCCTAAGTACAATTGT CTGGTACGCGTATAACTTC GTATAGCATACATTATACG AAGTTATCCTGTTCCAGTC GCCCTCCTCAGGCCAAC
hGCNF	NR6AI_HARmutPAM+MCSlox P REV	CACTATAGGGCGAATTGAA TTTAGCGGCCGCTCAAGTG ATTGCCAGGACTTTGC

The backbone is derived from Addgene plasmid #52343 AAVS1-TRE3G-EGFP (https://www.addgene.org/52343/)

Tab. 31: Primers for Genotyping, insert validation, colony PCR and sequencing

Target	Name	Sequence (5'->3')
hGCNF	NR6A1_KOE1_Intron FOR	AACTGAGGAAAGGGGAAGGC
hGCNF	NR6A1_KOE1_Intron REV	GACGAGCGGTAGGAAGGG
hGCNF	NR6A1_Cterm val FOR	GATATTAAGCAGGCACTCTGTG
hGCNF	NR6A1_Cterm val REV	TTCAGAACATGGCATTCTGCTC
PGK promoter	PGK FOR	AGCACCGCTGAGCAATGG
PGK promoter	PGK REV	ACCAAAGAACGGAGCCGG
Puromycin resistance	PuroR FOR	TTCCTGGCCACCGTCGG
Puromycin resistance	PuroR REV	GAAGAGTTCTTGCAGCTCGG
bGH polyA	bGH FOR	ATCGCATTGTCTGAGTAGGTGT
bGH polyA	bGH REV	ACACCTACTCAGACAATGCGAT
hGCNF	NR6A1_Cterm seq FOR	CAGGACAGAAGTTTTTGGACTA
hGCNF	NR6A1_E1_endo REV	GAAATGAAGCTCGGTGCCCA
U6 promoter	pX459 sgRNA seq	GAGGGCCTATTTCCCATGATT
ccdB element	ccdB REV	GGCGTGTCAATAATATCACTCTG
AM tag	AM tag FOR	CTCTCAGGCTTACTGAACTAGT

Tab. 32: Oligonucleotides for C-terminal tagging gene targeting vector

Origin	Name	Sequence (5'->3')
synthetized	NR6A1 3xFLAG Oligo FOR	GGAAGATTATAAAGATGATGATGATAAAGATTATAAAGATGATGATGATAAAGATTATAAAGATGATGATGATAAATGAACTAGT
synthetized	NR6A1 3xFLAG Oligo REV	TTATACTAGTTCATTTATCATCATCATCTTTATAATCTTTATCATCATCATCTTTATAATCTTTATCATCATCATCTTTATAATC
synthetized	NR6A1-C-AM-TGA FOR	GGAATGCCAAGATCCTCAACGCAAAGGCAACGTGATACTCTCTCAGGCTTACGGGTGCCAAGATCCTCAACGCAAAGGCAACGTGATACTCTCTCAGGCTTACTGAACTAGT
synthetized	NR6A1-C-AM-TGA REV	TTATACTAGTTCAGTAAGCCTGAGAGAGTATCACGTTGCCTTTGCGTTGAGGATCTTGGCACCCGTAAGCCTG

		AGAGAGTATCACGTTGCCTTTGCG TTGAGGATCTTGGCA

2.8.3 Primers for quantitative RT-PCR

Tab. 33: Primers used for mRNA RT-PCR analysis

Target	Primer sequence FOR (5'->3')	Primer sequence REV (5'->3')
18s rRNA	TTCCTTGGACCGGCGCAAG	GCCGCATCGCCGGTCGG
CRIPTO	GCTGCCCAAGAAGTGTTCC	TAGTACGTGCAGACGGTGGT
FEZF2	ACGCTCACTATAATCTCACCCG	TGGAGCTGCGGTTGAACG
FOXG1	TGGGAGATAGGAAAGAGG TGAAAA	GCACCAGGCTGTTGATGCT
GCNF	GAGGCCGGAATAAGAGCATT	CAGGGGAACTGTGGTCACTATC
NRIP1	GTGTTCCATGCCCAGTGAGA	TTGGCTGTGACCTGTGAGAC
OCT4	GTGGAGGAAGCTGACAACAA	ATTCTCCAGGTTGCCTCTCA
PAX3	GCCAATCAACTGATGGCTTT	AGCGGTTGAGGTCTGTGAAC
PAX6	AATAACCTGCCTATGCAACCC	AACTTGAACTGGAACTGACACAC
RSPO3	AGCTGACTGTGATACCTGTTTCA	CCCTTCTTCGTGCATGGACT
WNT4	ATGCTCTGACAACATCGCCT	GCCAGCACGTCTTTACCTCA

2.8.4 Antibodies

Tab. 34: Primary antibodies

Target	Host/Isotype	Source	Catalog #	Dilution*
β-Actin	Mouse IgG monoclonal	Sigma-Aldrich	A1978	1:5000 (WB)
Flag	Mouse IgG monoclonal	Merck KGaA	F1804-200UG	1:500
GCNF	Mouse IgG monoclonal	Perseus Proteomics	PP-H7921	1:1000 (WB)
OCT3/4	Mouse IgG monoclonal	Santa Cruz Biotechnology	sc-5279	1:400
PAX6	Rabbit IgG polyclonal	BioLegend	901301	1:300
SOX2	Rabbit IgG polyclonal	Novus Biologicals	NB110-37235	1:1000
TRA1-60	Mouse IgM monoclonal	EMD Millipore Corp USA	MAB4360	1:300

*Dilution for immunocytochemistry (ICC), if not stated otherwise. WB, Western blot.

Tab. 35: Secondary antibodies

Antibody	Source	Dilution**
HRP-linked goat anti-mouse IgG	Cell Signaling Technology	1:1000 (WB)
Alexa488 goat anti-mouse IgG	Invitrogen*	1:1000
Alexa488 goat anti-rabbit IgG	Invitrogen*	1:1000
Alexa555 goat anti-mouse IgG	Invitrogen*	1:1000
Alexa555 goat anti-rabbit IgG	Invitrogen*	1:1000

*Invitrogen is now part of Thermo Fisher Scientific
**Dilution for immunocytochemistry (ICC), if not stated otherwise. WB, Western blot.

3. Results

3.1 Establishment of GCNF-deficient hiPSCs utilizing CRISPR/Cas9-mediated genome editing

The CRISPR/Cas9 system allows genome editing in cell culture in a short period of time whereas it would have taken months with classical gene targeting techniques. Although its cost efficiency and feasibility are greatly appreciated in the scientific community, some requirements have to be addressed when setting up a genome editing experiment in hiPSCs. The first part of the result section focuses on these technical requirements to establish hiPSC clones deficient for the GCNF gene.

3.1.1 Prediction of sgRNAs targeting the endogenous GCNF gene

For CRISPR/Cas9-mediated genome editing it is of utmost importance to carefully investigate the genomic locus that is desired to be edited. Online databases provide the human reference genome and allowed for a first investigation of the genomic organization of GCNF. Many online algorithms from different laboratories provide sgRNA design by use of bioinformatic algorithms (Liu et al., 2020). These algorithms analyze a specific gene or DNA sequence for the presence of the PAM sequence that is a key requisite for the Cas9 endonuclease domains to introduce a DSB. Additionally, it serves to estimate off-target frequency of a given sgRNA. Due to the user-friendly interface as well as the wide acceptance in the scientific community, the CHOPCHOP online tool (https://chopchop.cbu.uib.no) for sgRNA prediction was chosen in this study to target GCNF (Labun et al., 2016). According to the on-target capacity and off-target risk of a given sgRNA, this algorithm assigns its ranking. The ranking is indicated by a color code for each sgRNA. The sgRNAs that had been assigned red have a bias towards off-targets, the yellow ones are expected to exhibit an intermediate range, whereas the sgRNAs marked as green indicate a low off-target activity. The CHOPCHOP online tool was queried with NR6A1, the gene symbol of GCNF, in order to find an sgRNA that would allow the introduction of a DSB in close proximity of the start codon (Fig. 7 A). The ranking indicated several low off-target sgRNAs, that would allow cutting near the start codon, under the top 10 of predicted sgRNAs (Fig. 7 B). SgRNA 4 was further investigated

because it allows cutting downstream of the start codon, whereas several other high ranked sgRNAs would result in a cut upstream of it. Additionally, sgRNA 4 starts with the DNA base guanine which would allow a strong expression from the U6 promoter (Mali et al., 2013) (Fig. 7 C). Importantly, for sgRNA 4 no off-targets have been found (Fig. 7 B).

A

B

Ranking	Target sequence	Genomic location	Exon	Strand	GC content (%)	Off-targets		
						0	1	2
1	GTTCGTCCCGCTCCATGCGGTGG	chr9:127533383	1	+	70	0	0	0
2	GCGGTTCGTCCCGCTCCATGCGG	chr9:127533380	1	+	74	0	0	0
3	CGGGACGAACCGCCGCCTAGCGG	chr9:127533370	1	-	78	0	0	0
4	GACGAACCGCCGCCTAGCGGAGG	chr9:127533367	1	-	74	0	0	0
5	CGAACCGCCGCCTAGCGGAGGGG	chr9:127533365	1	-	78	0	0	0
6	GAACCGCCGCCTAGCGGAGGGGG	chr9:127533364	1	-	78	0	0	0

C

Fig. 7: The CHOPCHOP online tool provides possible sgRNAs for a given gene. **(A)** The CHOPCHOP tool (https://chopchop.cbu.uib.no) is an online tool that outputs potential sgRNAs for a specific gene. When queried with GCNF (gene symbol NR6A1) it provided the depicted result. SgRNAs are categorized by a color code, with green showing the best on-target and off-target ratio, whereas sgRNAs that are marked in yellow and red are less well-suited. **(B)** The table indicates the ranking of the provided sgRNAs along with the target sequences, the exact genomic location and potential off-targets. For the highest ranking sgRNAs no off-targets are predicted. SgRNA 4 is marked in green. **(C)** CHOPCHOP output for sgRNA 4 with important components being highlighted. Note the start codon of the GCNF gene (green rectangle), as well as the cutting site and the protospacer-adjacent motif (PAM, blue rectangle), which is required for the Cas9 enzyme to bind the target DNA.

3.1.2 A plasmid-based Cas9 vector allows sgRNA customization to target a specific genomic sequence

Many different methods have been utilized to deliver the genome editing machinery into human cells. Among them is the mRNA-based delivery that relies on chemically modified mRNA (Hashimoto and Takemoto, 2015) as well as ribonucleoprotein complexes (RNP) that are widely appreciated for their efficiency (Chen et al., 2016). More recent systems lower costs by a modular RNP assembly with a customizable synthetic crRNA and a universal synthetic tracrRNA (Vakulskas et al., 2018). Although both mRNA- and RNP-mediated genome editing is advertised with the promise of high efficiency, it is expensive and depends on company supply. The plasmid-based system instead provides maximum flexibility and minimal cost. The most broadly used system originated from the Feng Zhang laboratory from Massachusetts Institute of Technology (MIT), is well established in laboratories throughout the world and can be easily purchased from Addgene (Ran et al., 2013b) (Fig. 8 A). From this series of plasmids, the pX459 plasmid seemed to be especially well-suited for the purpose of this study, as it encodes all necessary components on one plasmid to target and cut a desired gene and additionally allows for puromycin selection. PX459 encodes the Cas9 enzyme followed by a T2A-peptide-linked puromycin N-acetyltransferase that delivers resistance against puromycin. During translation, the T2A-peptide mediates cutting and separation between the Cas9 enzyme and the puromycin resistance (Ran et al., 2013b) (Fig. 8 A). Thus, the pX459 vector makes it possible to select for the delivery of the plasmid into the cells. A separate functional unit of the plasmid is the U6 promoter followed by the sgRNA sequence. The latter consists of a region that can be replaced with a BbsI digestion (Fig. 8 B) and subsequent incorporation of a small oligonucleotide encoding the chosen sgRNA (Fig. 8 C), thereby customizing the sgRNA to bind at the desired location. The customized part is followed by the scaffold sgRNA which resembles the tracrRNA of the bacterial system and mediates association with the Cas9 enzyme. To ensure a strong expression of the sgRNA from the U6 promoter it has to either naturally contain a G at the 5' end or has to be designed that way (Mali et al., 2013). In the case of sgRNA 4 (Fig. 7 B, C) this step can be omitted as the sgRNA already contains a G at the 5' end. The customization of the sgRNA sequence was done as such: The pX459 plasmid was obtained from Addgene and digested with

BbsI, which allows to remove a placeholder oligonucleotide and to replace it with the sequence of sgRNA 4 (Fig. 8 B, C). Successful incorporation of the respective oligonucleotides was confirmed with Sanger sequencing (Fig. 8 D). The resulting plasmid, once delivered to the cells, is supposed to express both Cas9 and the customized sgRNA, thereby navigating the nuclease to introduce a DSB at the desired location.
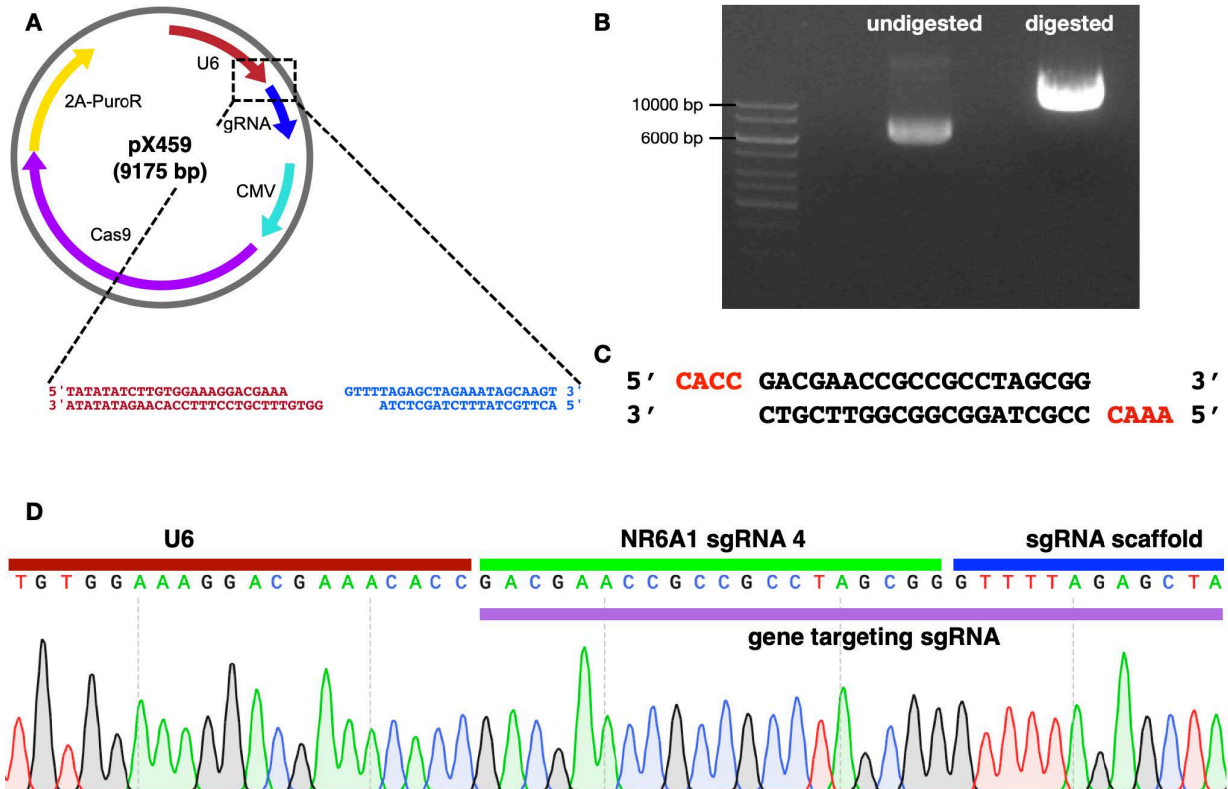
Fig. 8: The all-in-one vector pX459 can easily be customized to target a chosen sequence within the human genome. **(A)** The plasmid map indicates important elements. The sequence that encodes the Cas9 enzyme is followed by a T2A-peptide-linked puromycin resistance. The U6 promoter driven sgRNA can be customized by BbsI digestion and subsequent incorporation of a small oligonucleotide. **(B)** BbsI digestion results in linearization of the plasmid backbone. **(C)** Two duplexed DNA oligonucleotides exhibit overhangs that allow incorporation into the BbsI-digested backbone. **(D)** Sanger sequencing confirms successful customization of the sgRNA sequence. A comprehensive publication describes the procedure in detail (Ran et al., 2013b).

3.1.3 Cloning of a knockout gene targeting vector to introduce a puromycin resistance cassette into the GCNF open reading frame

Another important consideration in a CRISPR/Cas9 approach is, whether to rely solely on NHEJ or the more complex HDR for gene repair. NHEJ has been successfully utilized to generate hiPSCs deficient for a specific gene (Perez et al., 2008). When a cell encounters a DSB at a specific location it may join the resulting ends in an end-to-end fashion in order to repair the damage. This process is fast, efficient and favored in mammalian cells (Sargent et al., 1997). On the other hand, NHEJ is also error-prone and tends to generate insertions or deletions at the respective location (Mao et al., 2008). Upon delivery of the Cas9/sgRNA-containing plasmid into the cells and the introduction of a DSB, NHEJ may introduce indels that might render a cell deficient for a specific gene of interest. On the other hand, upon Cas9/sgRNA delivery, an HDR approach may be utilized to incorporate a desired sequence into a specific site. Even though this process is less efficient it provides the possibility to introduce exogenous DNA into the genome. This may be a fluorescent protein like eGFP or a resistance gene. An HDR-based strategy was followed in this study to introduce a puromycin resistance cassette into the first exon of the GCNF gene (Fig. 9 A, B). This causes a disruption of the GCNF ORF, thereby leading to a GCNF knockout and allows to enrich for cells that incorporated the cassette into the genome by puromycin selection. To that end it was necessary to generate a gene targeting vector by molecular cloning. The goal was also to include loxP sites to make it possible to excise the puromycin resistance gene after editing. The first step was the careful inspection of the endogenous GCNF sequence that was retrieved from the current Ensembl reference genome (Yates et al., 2020) (Fig. 9 A). The desired targeting vector is depicted in a scheme in Fig. 9 B and encompasses a left homology arm (HAL) of 800 bp as well as a right homology arm (HAR) of 800 bp (Fig. 9 C, D). Importantly, a PAM mutation was included in the HAR fragment to prevent cutting of the gene targeting vector by the Cas9 nuclease. The center of the construct harbors a puromycin resistance gene that is driven by a PGK promoter and is terminated by a bGH polyA sequence (Fig. 9 B, C). Notably, whereas cloning of the puromycin cassette and the HAL fragment was readily achieved, cloning of the HAR fragment required optimization and utilization of a different DNA polymerase.

A **human GCNF locus (chromosome 9, reverse strand)**

Exon 1

GCNF ORF start (see B)

B

sgRNA

5'UTR   Exon 1   human GCNF locus

ATG

HAL   PGK   PuroR   HAR   gene targeting vector

C   **D**

PuroR cassette   NR6A1 HAL   NR6A1 HAR   NR6A1 HAR

3000 bp
2000 bp
1500 bp
1000 bp

500 bp

1000 bp
750 bp
500 bp

250 bp
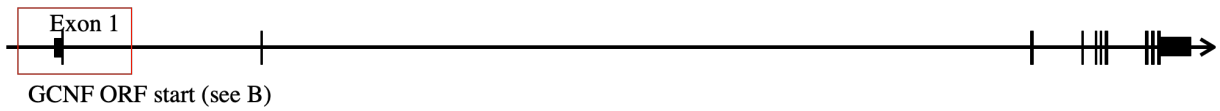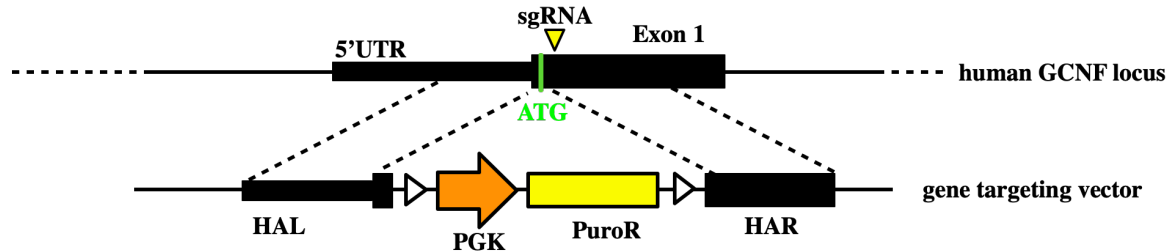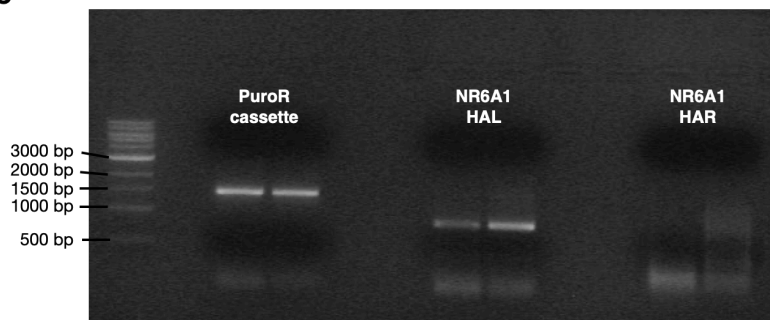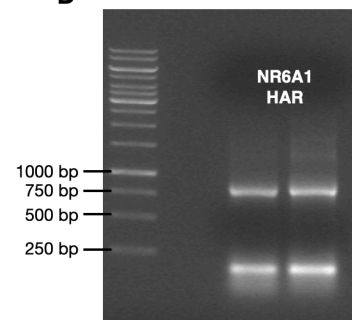
Fig. 9: Construction of a gene targeting vector to disrupt GCNF and introduce a puromycin resistance cassette in hiPSCs. **(A)** GCNF is encoded by the NR6A1 gene on chromosome 9. **(B)** The gene targeting vector encompasses a left homology arm (HAL), a right homology arm (HAR) and a PGK-driven puromycin resistance cassette (PuroR) in between. The start codon (ATG) of the GCNF ORF is indicated in green. The Cas9 cutting site is indicated as a yellow triangle. **(C)** The puromycin resistance cassette (1499 bp) and the HAL fragment (851 bp) was readily amplified by PCR. **(D)** The PCR amplification of the HAR fragment (921 bp) could only be achieved after first amplifying a larger fragment that served as a template for the aimed fragment. It was also needed to utilize a different polymerase (Phusion polymerase). The primers included overhangs for the subsequent Gibson assembly-based NEBuilder reaction. ORF = open reading frame, UTR = untranslated region.

3.1.4   Feasibility test of hiPSC nucleofection

Once, both the Cas9/sgRNA plasmid as well as the gene targeting vector are constructed, several stages have to be accomplished to generate a cell line that is deficient for a given gene (Fig. 10). Whereas lipofection is frequently used when transfecting immortalized cell lines, like HEK293FT cells, it is usually very inefficient in human stem cells (Cao et al., 2010). Therefore, the delivery of the genome editing reagents was performed by

nucleofection, which has previously been described as an efficient way to introduce exogenous DNA into stem cells (Cao et al., 2010).
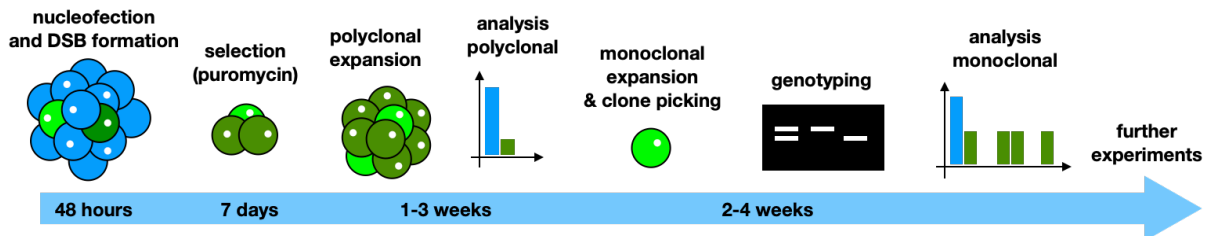
Fig. 10: Timeline to establish an hiPSC line deficient for a specific gene with CRISPR/Cas9. Upon nucleofection of both, the sgRNA/Cas9 plasmid and the gene targeting vector, the cells are plated on fresh plates and cultivated for 48 h to ensure Cas9 expression and DSB formation. Initially, the majority of the cells are non-edited (blue). 48 h post nucleofection, the cells are exposed to puromycin for 7 days to select edited cells, and expanded for 1 – 3 weeks. The selection removes non-edited cells, which do not carry the puromycin resistance cassette, and only cells remain that had undergone heterozygous (dark green) or homozygous editing (bright green). A first analysis can already be done at the polyclonal level. The generation of monoclonal cell lines requires another 2 – 4 weeks and includes single cell seeding, clone picking, genotyping and qRT-PCR analysis.

Nucleofection is an electroporation-based method and was tested with the pmaxCloning™ plasmid which encodes eGFP under control of the strong CMV promoter. The experiment provided a first insight that nucleofection might be capable to efficiently deliver genome editing constructs into hiPSCs (Fig. 11 A, B). Next, nucleofection was utilized to deliver the gene editing constructs that had been generated before to disrupt the endogenous GCNF gene. Therefore, hiPSCs were plated after nucleofection and selected with 0.5 µg/ml puromycin. Upon puromycin treatment visible cell death indicated a robust effect of the antibiotic compound on non-edited cells. After several days of selection, the nucleofected cells could be expanded free of noticeable morphological abnormalities. QRT-PCR was chosen as a first read-out of targeted gene disruption. Compared to the non-edited parental cell line, already in the polyclonal pool of nucleofected and puromycin-selected cells a strong decline in GCNF expression could be detected by qRT-PCR (Supplementary Fig. 1).
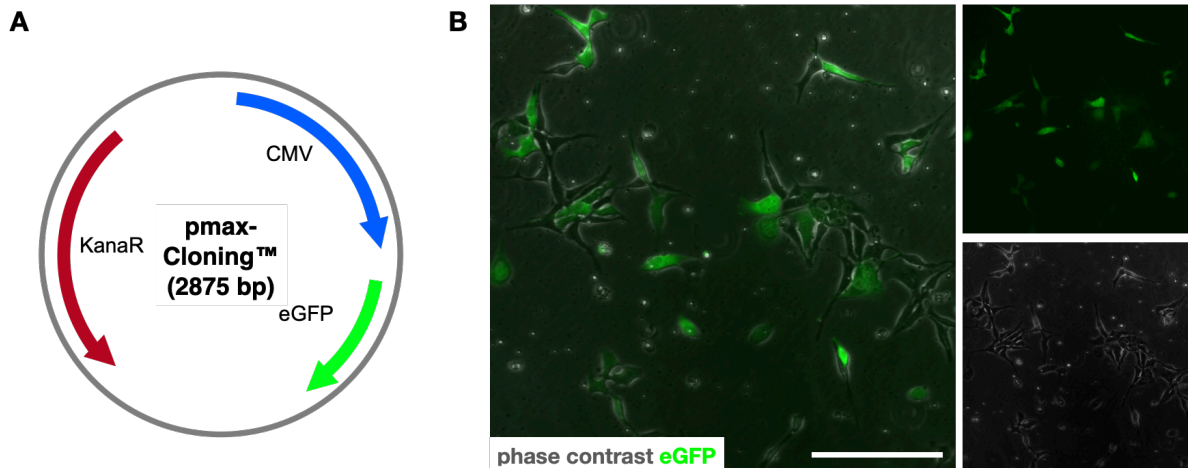
Fig. 11: Strong green fluorescence is induced 24 h after nucleofection of an eGFP-containing plasmid. **(A)** Plasmid map of the control plasmid from Lonza Biosciences (pmaxCloning™). **(B)** EGFP signal 24 h post nucleofection. Scale bar = 200 μm.

3.1.5 Genotyping of single cell-derived hiPSC clones

Encouraged by the promising results on the polyclonal level (Supplementary Fig. 1), single cell-derived clones were generated and subjected to genotyping. To that end, two different PCR reactions were performed for 12 single cell-derived clones (Fig. 12 A). The non-edited parental cell line (NE) was included as a negative control. The first reaction yields a PCR product of 1295 bp and shows the existence of the non-edited endogenous GCNF locus. The second reaction yields a PCR product of 1153 bp and turns positive when the puromycin cassette was successfully integrated. The combination of both reactions makes it possible to assess whether a clone is non-edited (GCNF +/+), whether only one allele had been edited (GCNF +/-, heterozygous editing) or whether both alleles had been edited (GCNF -/-, homozygous editing) (Fig. 12 B). The reaction that included the non-edited parental cell line (NE) gave rise to an endogenous band, but not to an insert-band. Clones 2, 3, 5, 8, 9, 11 and 12 showed both the endogenous band, as well as the insert-band, indicating a heterozygous genotype or clones that contain not only edited, but also non-edited cells (mosaic clones) (Fig. 12 B). Clones 1, 4, 7 and 10 showed only the insert-band, but not the endogenous one, indicating that a homozygous editing occurred (Fig. 12 B). Clone 6 showed an endogenous band, but no insert-band. Due to the fact that clone 6 survived puromycin selection, this indicates a random integration of the plasmid.
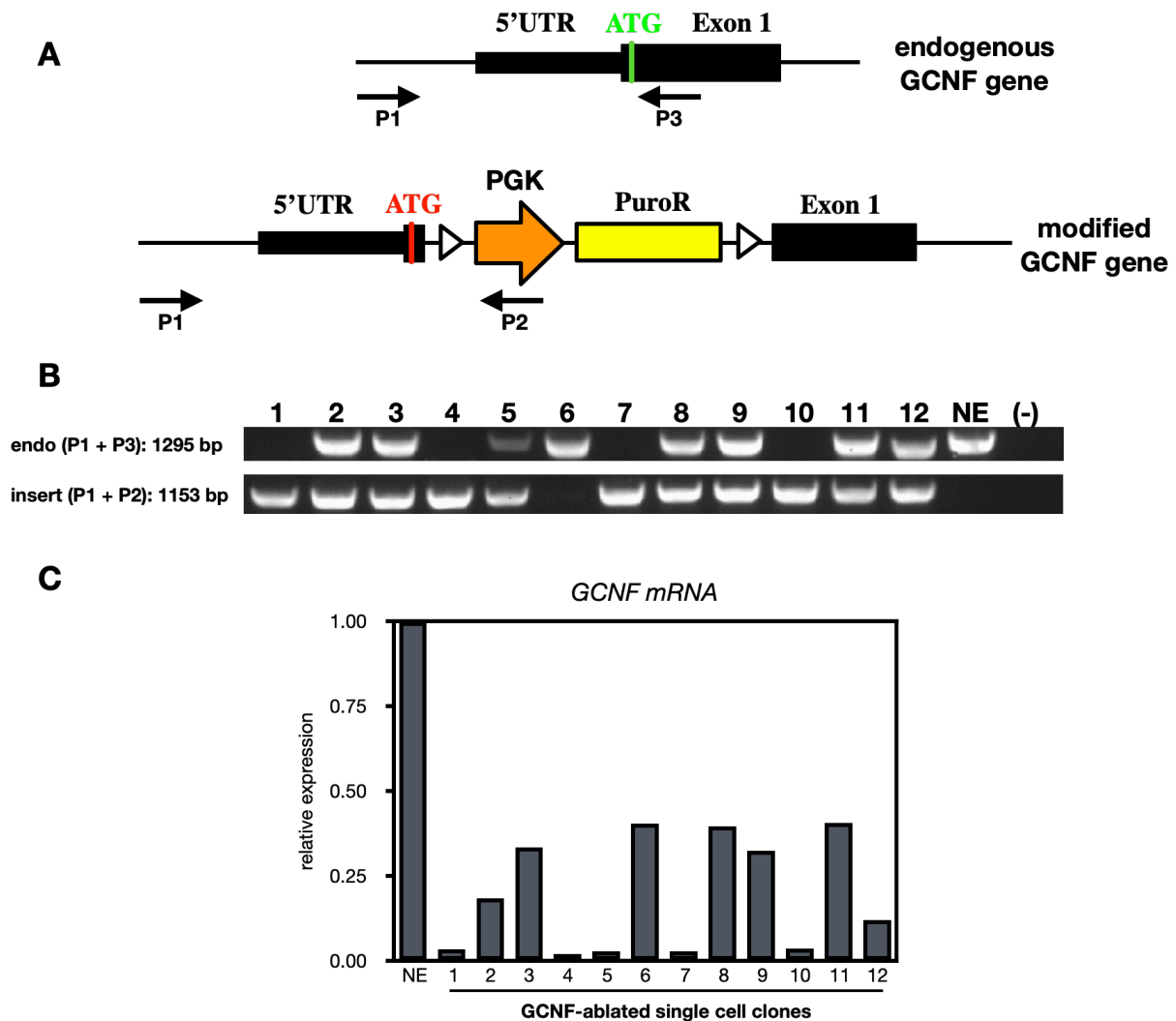
Fig. 12: Genotyping of single cell-derived GCNF-ablated hiPSCs. **(A)** The genotyping relies on two PCR reactions. Primer P1 and P3 generate a 1295 bp band in non-edited alleles. When primer P1 is combined with P2, which binds in the insert, the reaction gives rise to a 1153 bp band that is only present when the puromycin cassette got inserted into the GCNF locus. The functional start of the ORF is marked by a green ATG, the disrupted start of the ORF is marked by a red ATG. **(B)** Presence or absence of the two PCR products on the agarose gel indicate whether a clone exhibits a homozygous loss of GCNF (GCNF -/-, endo-/insert+), a heterozygous loss of GCNF (GCNF +/-, endo+/insert+) or if random integration occurred (GCNF +/+, endo+/insert-). The non-edited, parental cell line (NE, GCNF +/+, endo+/insert-), as well as H_2O (-) was included as a control. The entire agarose gels can be found in Supplementary Fig. 2. **(C)** GCNF expression was also assessed with qRT-PCR to confirm the genotyping results on the transcript level. A reduction of GCNF mRNA indicates heterozygous editing, whereas a depletion of > 95 % indicates homozygous editing. NE = non-edited control. n=1.

To further analyze the cell lines, they were also subjected to qRT-PCR (Fig. 12 C). The clones did either show a strong reduction of GCNF mRNA levels, as compared to the non-edited control, indicating heterozygous editing or an expression close to 0 %, indicating a homozygous editing and a full knockout of the GCNF gene. This was largely aligned with the results obtained from the genotyping reaction, except for clone 5 which showed a strong reduction of GCNF expression close to 0 %, but did show a product indicating a non-edited locus in the genotyping reaction. This may be caused by some non-edited cells contaminating the otherwise homozygous population of edited cells (Fig. 12 B, C). Clone 12 appeared to be heterozygous in the genotyping PCR, but exhibited a fold change reduction of > 85 % in the qRT-PCR, which may also be caused by mosaicism (Fig. 12 B, C).

3.2   GCNF-ablated single cell-derived hiPSC clones pass quality control parameters

Based on the genotyping analyses, 3 presumptive homozygous knockout clones (C1, C4 and C7) were chosen for further characterization. Their insert PCR product was subjected to Sanger sequencing, which is capable of confirming the incorporation of the construct into the endogenous GCNF locus. Indeed, analysis of the resulting sequence indicated a seamless transition of the GCNF 5' upstream region into the region of the left homology arm. This excludes the possibility of the sequence being derived from the gene targeting vector and proofs its exact integration into the desired location in the genome (Fig. 13). The presence of the PGK promoter could also be confirmed in all clones (Supplementary Fig. 3).
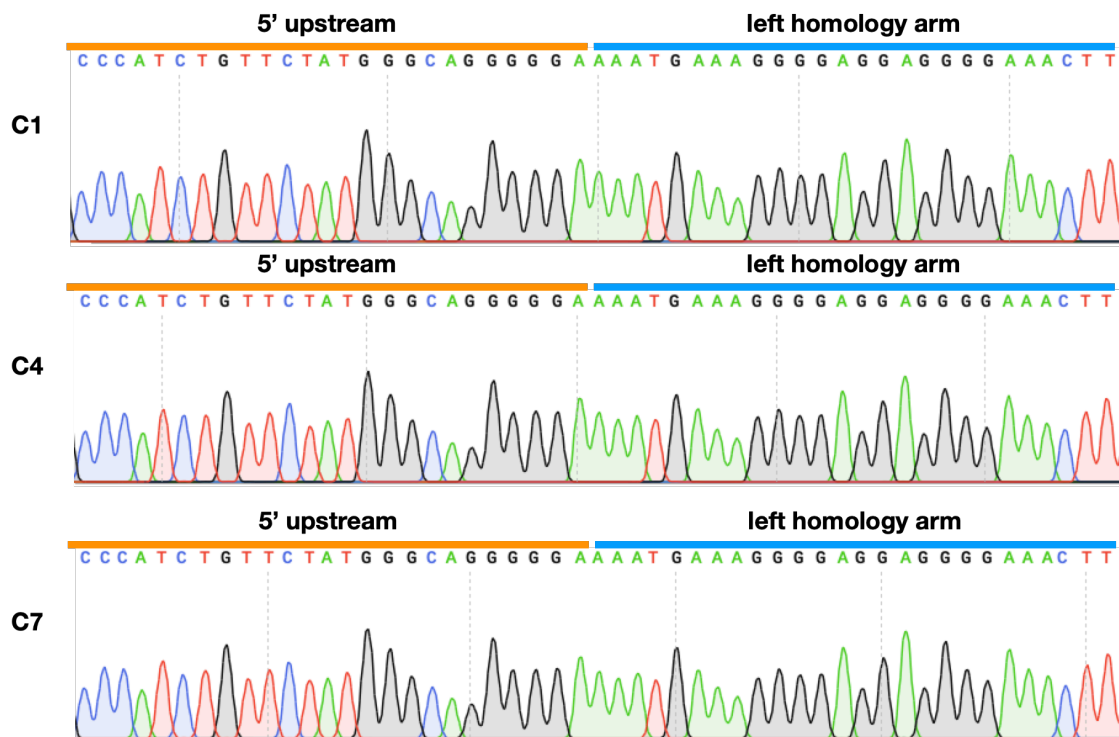
Fig. 13: Sanger sequencing confirmed the incorporation of the puromycin resistance cassette at the desired genomic location in GCNF-ablated hiPSC clones. The PCR product was generated with one primer that binds outside the region of the left homology arm (5' upstream) and one primer that binds inside the insert. The depicted sequence can only be generated when the integration of the puromycin cassette occurred at the aimed genomic location and excludes its derivation from the gene targeting vector. In the same sequencing reaction, the 5' end of the PGK promoter could also be confirmed in all clones (Supplementary Fig. 3).

Non-edited, as well as GCNF-depleted clones were checked for vitality and morphological abnormalities on a regular base in phase contrast microscopy (Fig. 14). No morphological abnormalities or overt differences in proliferation could be detected during cell clone expansion.
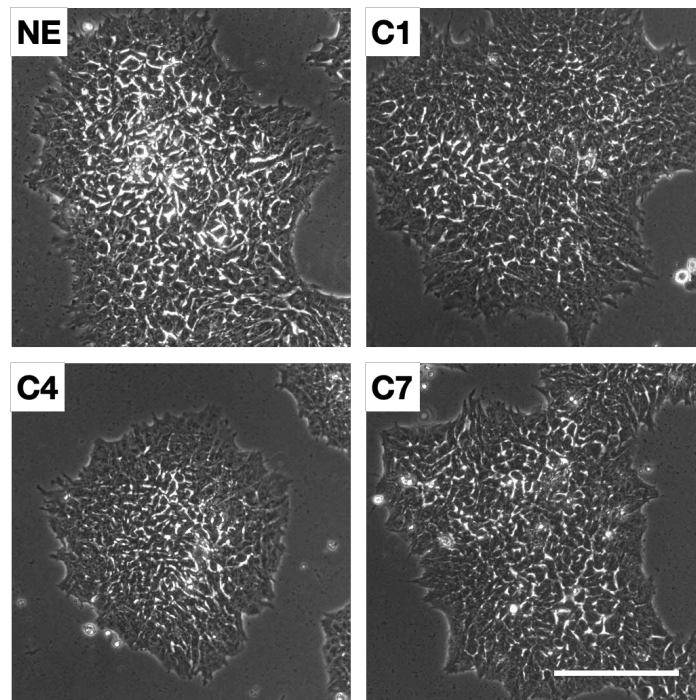
Fig. 14: Phase contrast images of non-edited, as well as GCNF-ablated hiPSC clones. Scale bar = 200 µm.

Immunocytochemistry suggests that the pluripotency markers OCT4 and SOX2 were consistently expressed on a comparable level to the non-edited control hiPSCs in GCNF-ablated clones (Fig. 15 A). To confirm GCNF depletion on mRNA level, RNA of the cells was collected in self-renewing conditions. QRT-PCR revealed a consistent and robust downregulation of GCNF mRNA in GCNF-depleted cells compared to the non-edited parental cell line (Fig. 15 B). Additionally, qRT-PCR was performed for OCT4 and CRIPTO. Both of them are known GCNF target genes and are expected to be upregulated upon GCNF depletion (Hentschke et al., 2006; Wang et al., 2016). However, in self-renewing conditions no obvious upregulation of OCT4 and CRIPTO could be observed, indicating that other factors might compensate for the loss of GCNF in self-renewing conditions (Fig. 15 B). It is important to account for the fact that mRNA expression changes do not necessarily correlate with changes on protein level. Therefore, protein samples were collected in hiPSCs from non-edited controls and GCNF-ablated clones in self-renewing conditions. Western blot analysis showed GCNF depletion in the clones that had undergone the editing procedure (Fig. 15 C, Supplementary Fig. 4). As genome editing relies on introducing a DSB, which is a known cause for genomic aberration (Cullot

et al., 2019; Rayner et al., 2019), genomic DNA was subjected to a SNP analysis. SNP analysis confirmed genomic integrity of the hiPSC lines, and no major aberration could be detected as compared to the non-edited cell line (Supplementary Fig. 5 - 8).
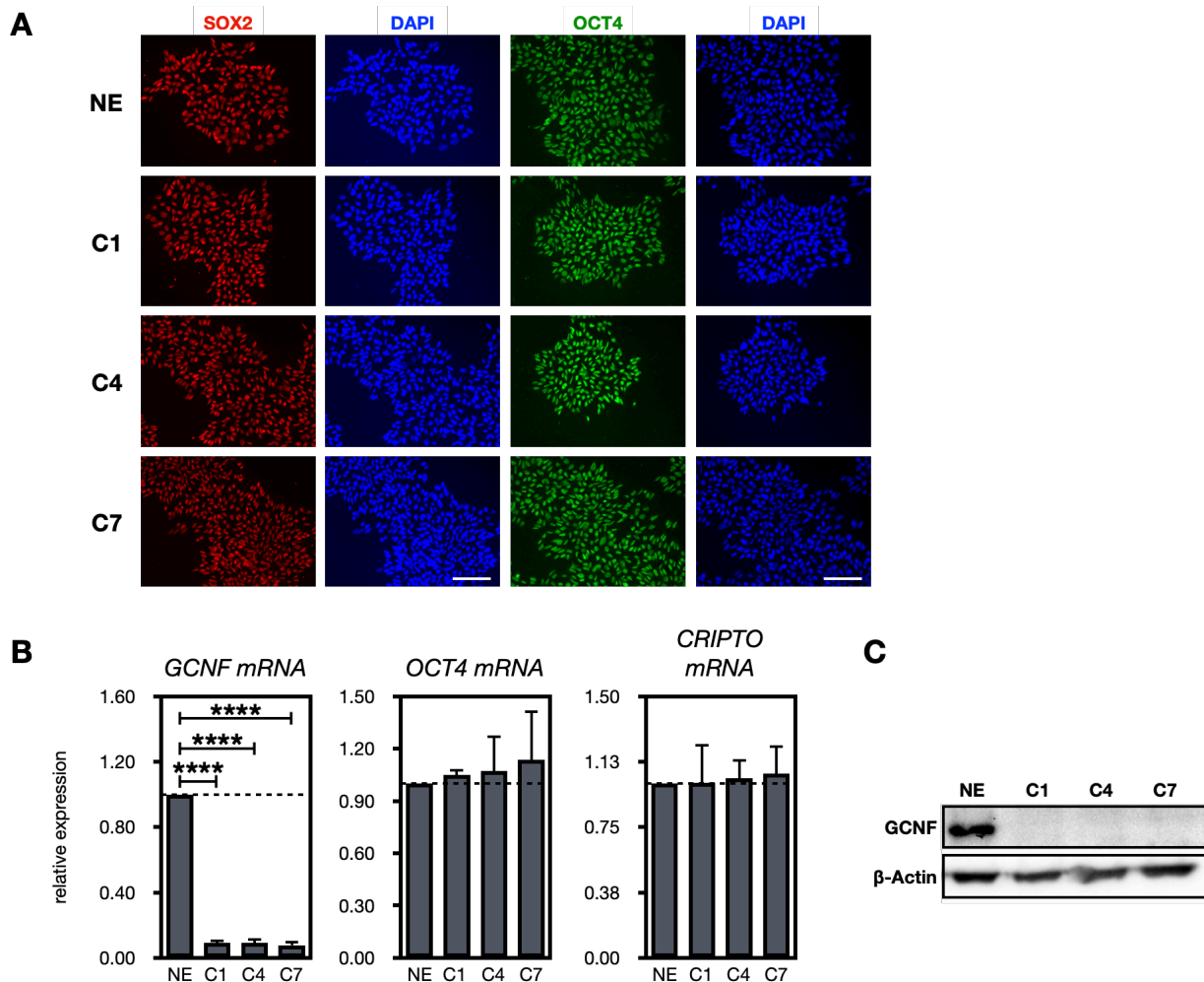
Fig. 15: GCNF-ablated hiPSC clones pass standard quality control and show significant reduction of both GCNF mRNA as well as protein. **(A)** Representative images of immunocytochemistry for SOX2 and OCT4 suggests a low degree of differentiation of the GCNF-ablated clones (C1, C4, C7) as well as the non-edited control (NE) hiPSCs. All stainings were performed in 3 independent biological replicates. Scale bars = 100 µm. **(B)** QRT-PCR analysis shows GCNF down-regulation in the edited clones. Two known GCNF target genes show no signs of upregulation in self-renewing conditions. Data were normalized to 18s rRNA levels and are presented as mean + SD, relative to the expression in non-edited cells (set to 1, dashed line; $n = 3$). ****, $p \leq 0.0001$. **(C)** Western blot analysis indicates prominent reduction of GCNF protein in GCNF-ablated clones and was performed in 3 independent biological replicates from hiPSCs in self-renewing conditions (Supplementary Fig. 4).

3.3 GCNF-deficient hiPSCs show alterations in germ layer specification

Several hiPSC lines deficient for GCNF have been successfully established. The generated cell lines were then employed to study how GCNF mediates exit of pluripotency and whether there is alteration in germ layer specification when hiPSCs are driven towards differentiation. Special emphasis was put on the assessment of alterations in neuroectodermal specification.

3.3.1 GCNF knockout hiPSCs exhibit impaired germ layer differentiation in an undirected differentiation experiment

The first experiment was aimed at shedding light on the influence of GCNF on germ layer differentiation. Given its importance in the early neural lineage and the alteration in brain development in knockout mice, absence of GCNF might have an influence on germ layer commitment in a human stem cell model (Akamatsu et al., 2009; Chung et al., 2006). Therefore, hiPSCs were detached and transferred to floating culture in a Petri dish (Fig. 16 A). The absence of coating allows the stem cells to assemble into spheroid-shaped embryoid bodies (EBs) in a serum-based medium (Fig. 16 B). The EBs were cultivated for 2 weeks and then subjected to RNA extraction.
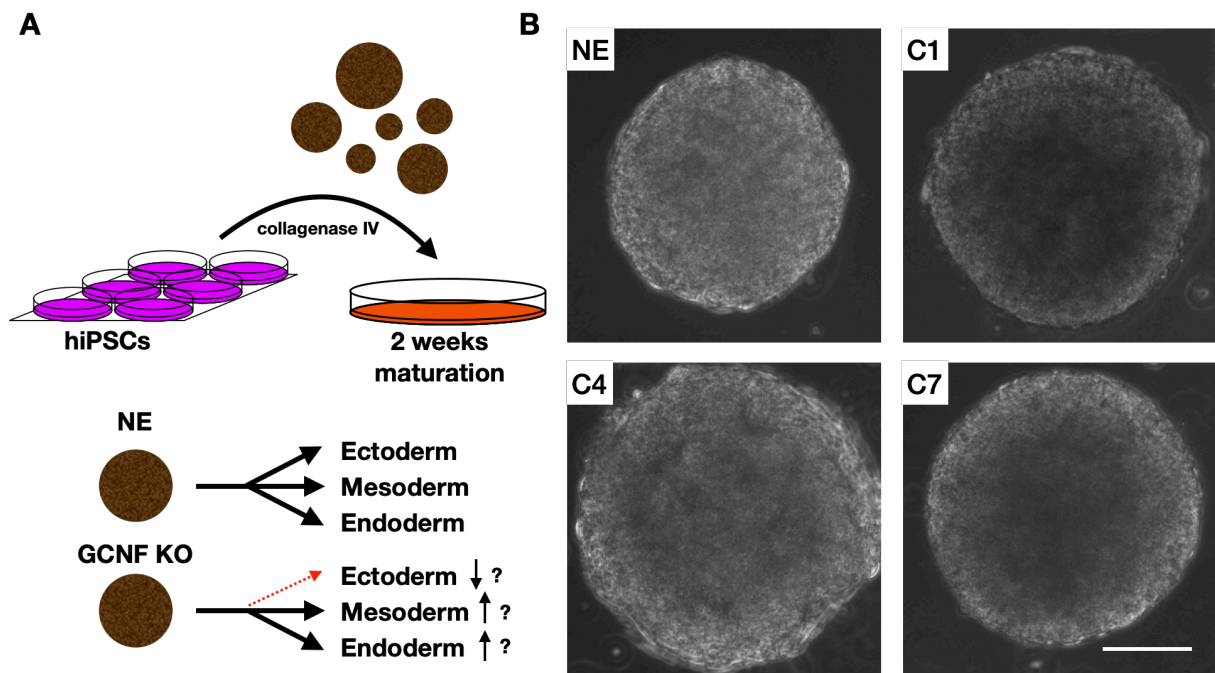
Fig. 16: Embryoid body-based undirected differentiation experiment to assess alterations in germ layer specification. **(A)** HiPSCs were detached from the cell culture dish with collagenase IV and transferred to an uncoated Petri dish in a permissive medium, that allows differentiation into all three germ layers. Comparison of the GCNF-ablated clones with the non-edited control (NE) might uncover differences in germ layer specification. **(B)** Shown are representative images of the spheroid-shaped embryoid bodies (EBs) 48 h after transfer to the EB medium. EBs formed in all clones including non-edited (NE) and GCNF-ablated clones (C1, C4, C7). Scale bar = 100 µm.

As a first step, qRT-PCR was performed for GCNF, OCT4, CRIPTO and PAX6 (Fig. 17 A). The cells exhibited highly significant reduction of GCNF expression in all three GCNF-depleted clones. Analysis of two known GCNF target genes, OCT4 and CRIPTO, showed an upregulation which was partly significant (Fig. 17 A). PAX6 expression was reduced in two clones, although not statistically significant (Fig. 17 A). The non-edited control and two GCNF-deficient clones (C1 and C7) where subjected to a Scorecard assay. The Scorecard is a TaqMan-based method assessing a large set of marker genes at the same time to investigate whether a population of hiPSCs is capable of differentiating into all three germ layers. The analysis of the Scorecard assay with the Thermo Fisher Scientific hPSC Scorecard analysis software yields a score for self-renewal and each germ layer, thereby visualizing whether the hiPSCs are capable to differentiate and to what extent they differentiate into each germ layer (Fig. 17 B). GCNF-depleted clones exhibited a slightly elevated level of the self-renewal score, even though the changes did not reach

statistical significance. In addition, an impairment of the ectodermal score could be observed in the GCNF-depleted clones, that reached statistical significance in C1 (Fig. 17 B).
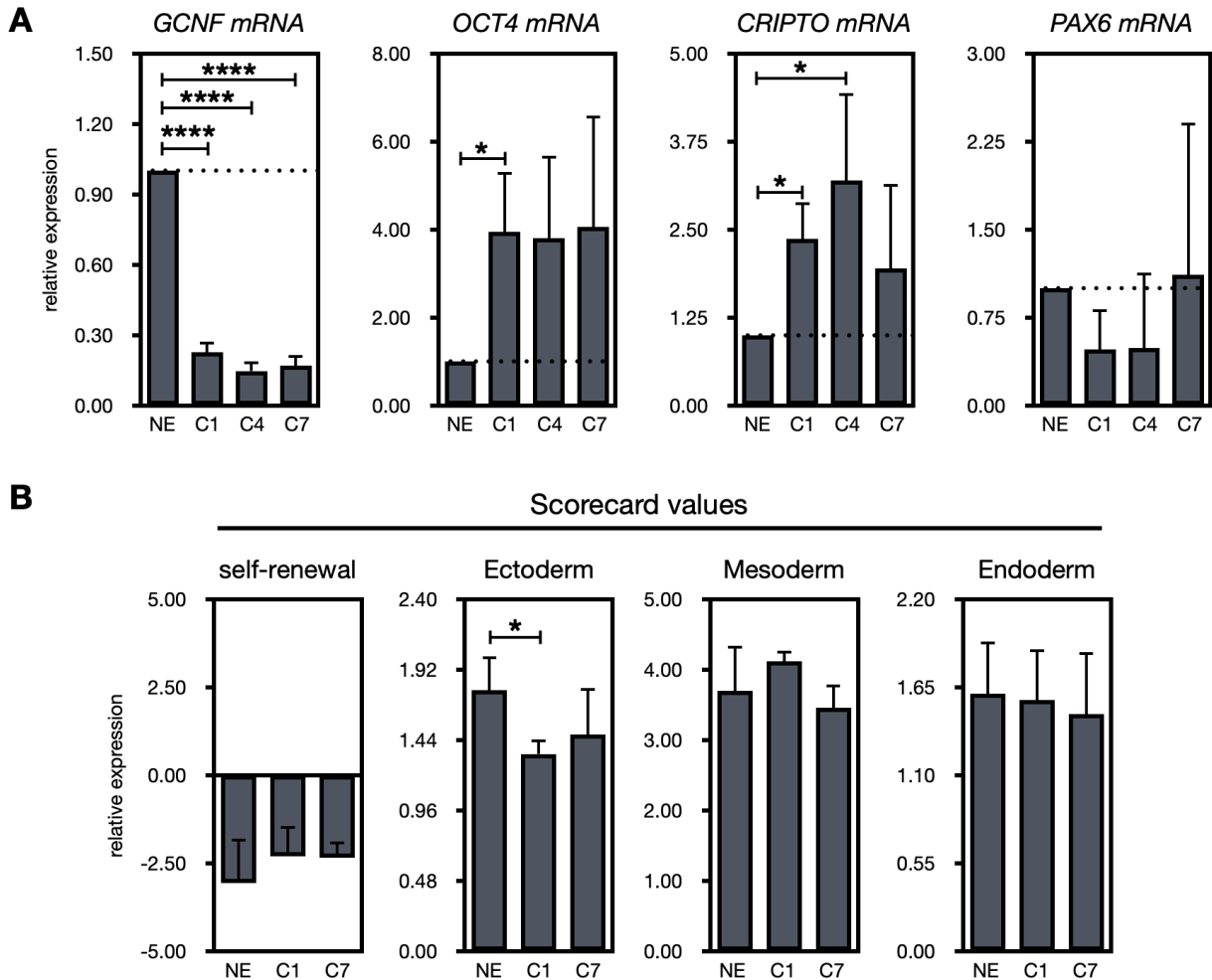
Fig. 17: QRT-PCR and Scorecard values indicate alterations in germ layer specification. **(A)** The mRNA levels of EBs after 14 days of undirected differentiation were analyzed via qRT-PCR and confirmed GCNF depletion in the knockout clones. Its two target genes OCT4 and CRIPTO exhibit an upregulated expression level, although statistical significance is not reached for all clones. A trend indicates that PAX6 expression might be impaired in the knockout clones, although no statistical significance is reached. Data were normalized to 18s rRNA levels and are presented as mean + SD, relative to the expression in non-edited (NE) cells (set to 1, dashed line; n = 3). *, $p \leq 0.05$; ****, $p \leq 0.0001$. **(B)** As compared to the non-edited control (NE), the GCNF-ablated clones show a small increase in the expression of genes that are associated with self-renewal of hiPSCs. The ectoderm score seems to be slightly decreased although this effect is only statistically significant for C1. Data were normalized by the Thermo Fisher Scientific hPSC Scorecard analysis software and are presented as mean + SD. n = 3. *, $p \leq 0.05$.

To account for subtle changes in the undirected differentiation paradigm, statistical significance tests were performed for all 85 marker genes covered by the Scorecard assay. Interestingly, although the analysis yielded only a small list of significantly dysregulated genes, most of them could be assigned to the ectoderm (Fig. 18 A). In total, 5 genes were found to be consistently dysregulated in the ectoderm, whereas only 1 mesodermal marker gene and 3 endodermal marker genes were consistently changed (Fig. 18 B, C). More specifically, whereas the ectodermal marker genes WNT1, PAX3, NOS2 and TRPM8 appeared to be downregulated, PRKCA expression levels increased (Fig. 18 A). HAND1, a mesodermal marker gene as well as the endodermal marker gene EOMES exhibited a consistent upregulation (Fig. 18 B, C). Oppositely, the endodermal marker genes FOXP2 and SST, exhibited a consistent downregulation (Fig. 18 C).
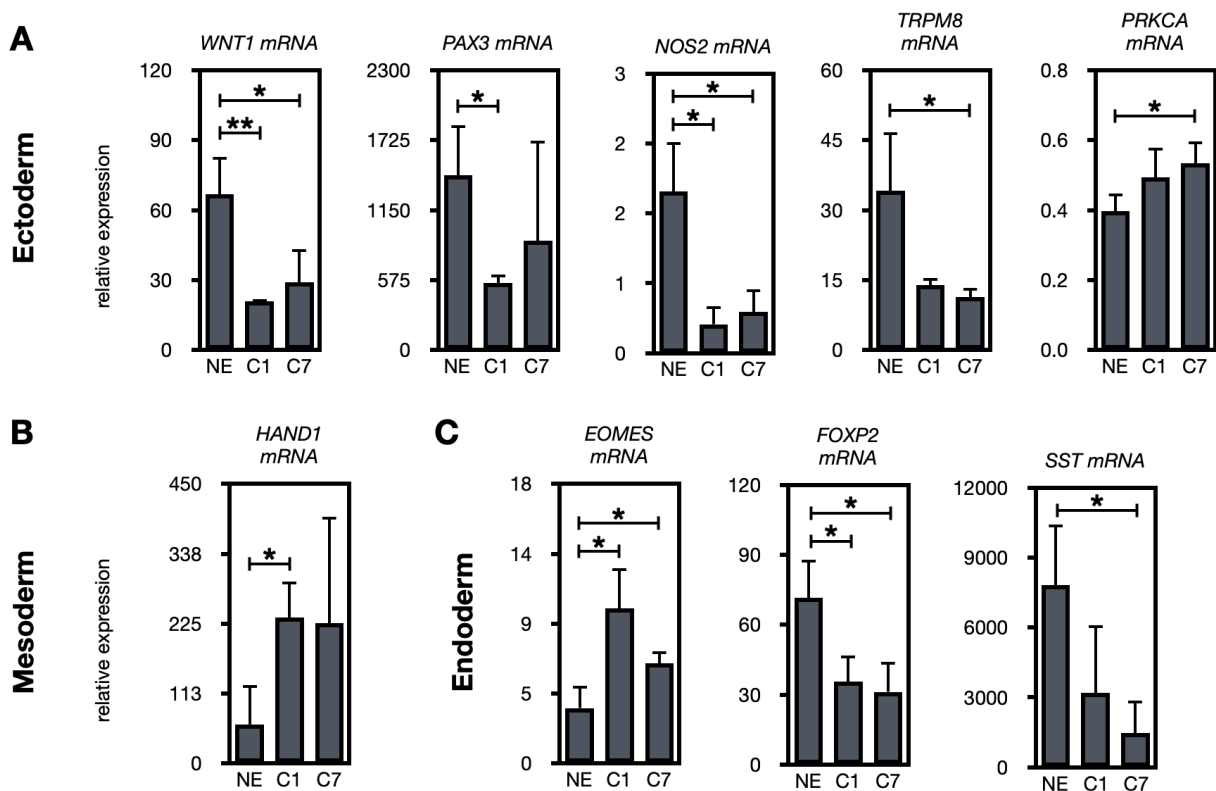
Fig. 18: Germ layer-specific genes that exhibit significant alterations in the Scorecard assay. **(A)** The ectodermal gene expression levels exhibit the most pronounced changes in significantly dysregulated marker genes. **(B, C)** Significantly dysregulated mesodermal and endodermal marker genes. All data were normalized by the Thermo Fisher Scientific hPSC Scorecard analysis software and are presented as mean + SD. n = 3. *, $p \leq 0.05$; **, $p \leq 0.01$.

3.3.2 3'-mRNA sequencing and gene ontology analysis indicate defective neural induction in GCNF-ablated hiPSCs

Given the presumable role of GCNF during early neuroectodermal specification, it would be especially interesting to assess alterations of GCNF-ablated cells in a neural induction paradigm. To that end, non-edited as well as GCNF-deficient hiPSCs were subjected to a directed dual-SMAD inhibition-based neural induction for 48 h and 96 h (Fig. 19). This protocol relies on inhibition of both TGFβ- as well as BMP-mediated SMAD signaling and results in an efficient neuralization of hiPSCs (Chambers et al., 2009).
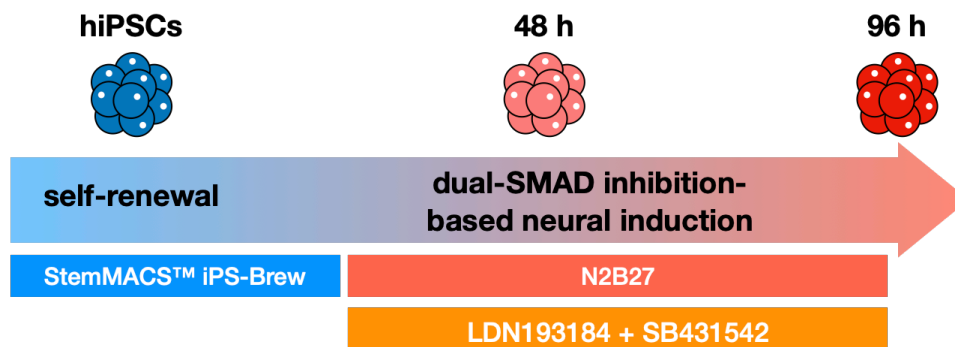
Fig. 19: Dual-SMAD inhibition-based neural induction experiment. The cells were exposed to the dual-SMAD inhibition-based neural induction medium. RNA was collected after 48 h and 96 h. The RNA samples of four independently performed experiments were subjected to qRT-PCR. The RNA samples of one of the four experiments was also subjected to 3'-mRNA sequencing.

The RNA of four experiments was collected on the respective timepoints to perform qRT-PCR analysis. The RNA samples of the 96 h timepoint of one experiment were subjected to 3'-mRNA sequencing. The resulting data sets were analyzed with R. In the analysis, non-edited cells were compared to all three GCNF-deficient clones after 96 h of neural induction with dual-SMAD inhibition. The top differentially expressed genes (DEGs), included both up- and downregulated genes and are displayed as a heatmap in Fig. 20 A. As expected, GCNF (NR6A1) can be found within the top down-regulated genes. Interestingly, GCNF-ablated clones showed a strong decrease in PAX6 transcripts. Additionally, FEZF1 and FEZF2 (forebrain embryonic zinc finger 1 and 2), factors that are important for forebrain development (Shimizu et al., 2010), are strongly downregulated upon GCNF ablation.
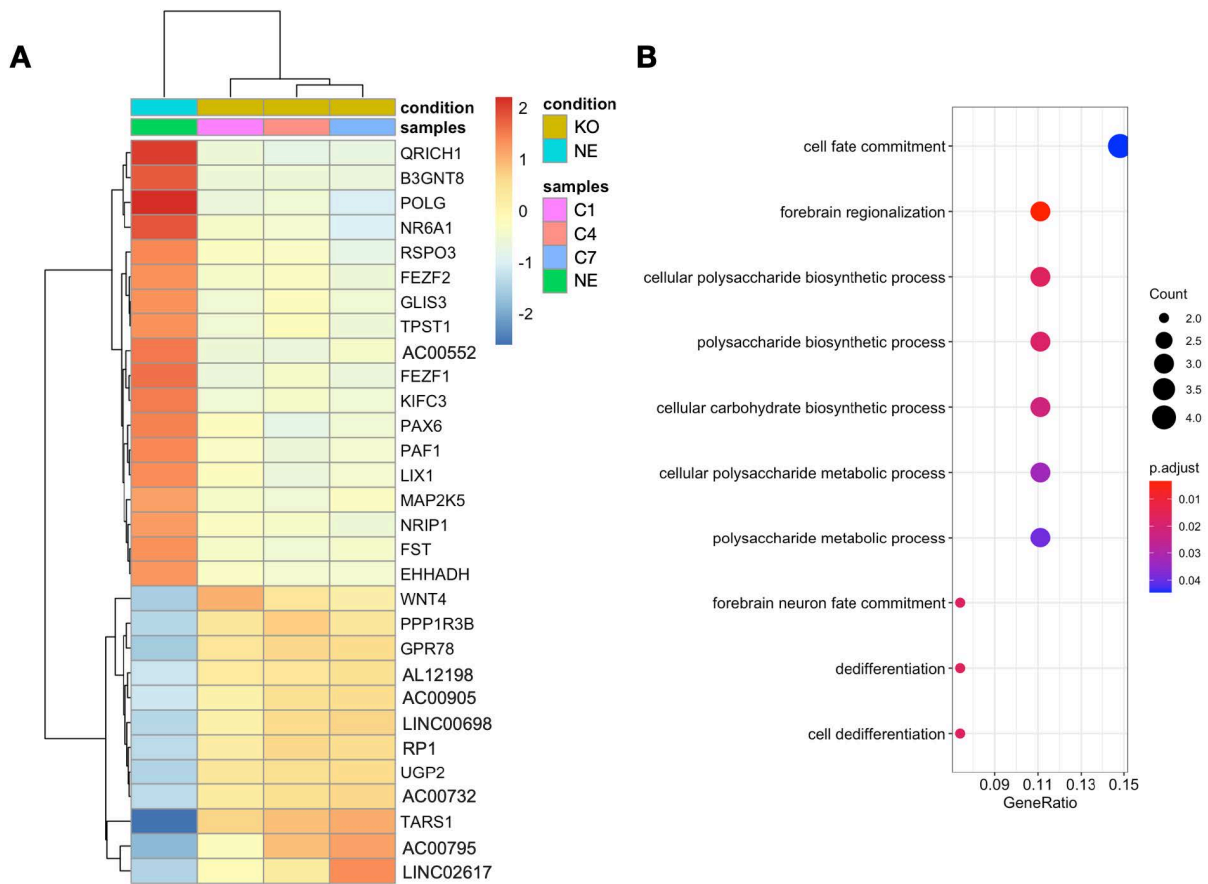
Fig. 20: 3'-mRNA sequencing for analyzing the role of GCNF during neural induction. **(A)** The RNA samples of one of the four experiments were subjected to 3'-mRNA sequencing. In the analysis non-edited hiPSCs (NE) were compared to all GCNF-ablated hiPSCs (C1, C4, C7) after 96 h of dual-SMAD inhibition-based neural induction. Presentation of the top 30 differentially expressed genes (DEGs) between non-edited and all GCNF-ablated clones as a heatmap. Notably, GCNF (NR6A1) is among these genes, thereby confirming the functional knockout. Interestingly, the heatmap shows genes that are implicated in hormonal regulation (e.g. FST, WNT4, NRIP1) as well as genes that are involved in maintenance of neural progenitors (e.g. PAX6, FEZF1, FEZF2). **(B)** Gene Ontology analysis with clusterProfiler (Yu et al., 2012) reveals several enrichment categories, for example 'forebrain regionalization' and 'forebrain neuron fate commitment'.

Also, NRIP1, a coactivator of nuclear receptors (Heim et al., 2007) as well as RSPO3, a mediator of canonical WNT signaling (de Lau et al., 2012) are strongly downregulated. Similarly, FST, which promotes the neuroectodermal fate by inhibiting BMPs (Liu and Niswander, 2005) was found to be down-regulated in the GCNF-deficient cells.

WNT4 was found within the top upregulated genes, a factor that is involved in female gonad development and axon guidance (Jordan et al., 2003; Lyuksyutova et al., 2003).

Gene ontology analysis showed a significant enrichment of several GO terms that point towards the presumable effect of GCNF on brain development. Among the significant GO terms were "forebrain regionalization" and "forebrain neuron commitment" (Fig. 20 B).

3.3.3 Gene expression analysis with quantitative RT-PCR and immunocytochemistry confirms defective neural induction in GCNF-ablated hiPSCs

To confirm the results obtained by the bioinformatical approach, qRT-PCR was performed on genes, that had been identified as interesting candidates to be involved in three different categories that are linked to GCNF function: "pluripotency", "brain development" and "hormonal regulation" (Fig. 21). It also included some additional genes that were hypothesized to exhibit alteration. This was done after 48 h and after 96 h of neural induction to get insights into the temporal dynamic of the GCNF knockout-mediated deficits. In the category of pluripotency, GCNF expression showed the expected strong decrease in all three knockout clones (Fig. 21 A). OCT4 expression showed a non-significant slight increase after 48 h of dual-SMAD inhibition, whereas a significant upregulation of up to 2.34 ± 0.51 fold, compared to the non-edited cell line, was found after 96 h of neural induction (Fig. 21 B). Similarly, CRIPTO showed a non-significant increase upon 48 h of dual-SMAD inhibition, whereas a strong upregulation of up to 6.16 ± 2.35 fold could be observed after 96 h (Fig. 21 C). The upregulation of the two known GCNF target genes OCT4 and CRIPTO is a strong indicator that GCNF loss-of-function indeed impairs the exit of pluripotency. In the category of "brain development" the first genes subjected to qRT-PCR were PAX6 and FEZF2 as they were among the most differentially expressed genes in the 3'-mRNA sequencing experiment (Fig. 20 A). Whereas high variability of PAX6 mRNA levels after 48 h of neural induction indicates that PAX6 is only expressed in a small amount at this timepoint, its expression levels at 96 h were significantly decreased to around 0.25 fold in the edited clones as compared to the non-edited cells (Fig. 21 D). Similarly, FEZF2, a gene that is involved in forebrain differentiation showed a highly significant decrease in mRNA expression levels of around 0.20 fold in the GCNF-ablated clones after 96 h of neural induction. (Fig. 21 E). As FOXG1 was shown to be positively regulated by FEZF2 (Wang et al., 2011) and is impaired in a GCNF knockout mouse model (Chung et al., 2006), it was also included in the qRT-PCR analysis. Indeed, FOXG1 expression exhibited a similar pattern as FEZF2. Whereas upon

48 h of dual-SMAD inhibition only C4 already showed a significant downregulation of FOXG1, 96 h of differentiation resulted in a strong decrease in FOXG1 mRNA levels in the GCNF knockout clones (Fig. 21 F). Notably, PAX3, a known marker of neural crest development, was impaired in the undirected differentiation experiment (Fig. 18 A). Therefore, it was also investigated after 48 h and 96 h in the neural induction experiment. Whereas no cell line expressed PAX3 on a detectable level after 48 h, a strong decline in mRNA levels could be observed in GCNF-ablated clones as compared to non-edited control cells after 96 h of neural induction (Supplementary Fig. 9). The category "hormonal regulation" addressed the fact that GCNF is tightly connected to gonad development and has similar features as other nuclear receptors like the testosterone receptor or the estrogen receptor (Sabour et al., 2014; Tata, 2002; Weikum et al., 2016). Interestingly, WNT4 ranked among the most upregulated genes in the transcriptomic analysis (Fig. 20 A). This was confirmed by qRT-PCR. Whereas no upregulation was observed after 48 h of differentiation, a strong WNT4 upregulation of up to ~12 fold could be observed after 96 h (Fig. 21 G). RSPO3 belongs to the R-spondin family of proteins that are known regulators of WNT signaling (de Lau et al., 2012) and showed a strong decrease already in the 48 h, but more strongly in the 96 h condition (Fig. 21 H). Another interesting candidate in this category was NRIP1, which ranked among the most differentially expressed genes in the transcriptomic analysis (Fig. 20 A). NRIP1 serves as a coreceptor of nuclear receptors and mediates repression of target genes (Heim et al., 2007). Considering that GCNF is described to act as a transcriptional repressor, NRIP1 might be a potential candidate that mediates the repression of GCNF target genes. After 48 h, only C4 showed a significant NRIP1 decrease, whereas after 96 h both C1 and C4 showed a significant downregulation (Fig. 21 I).
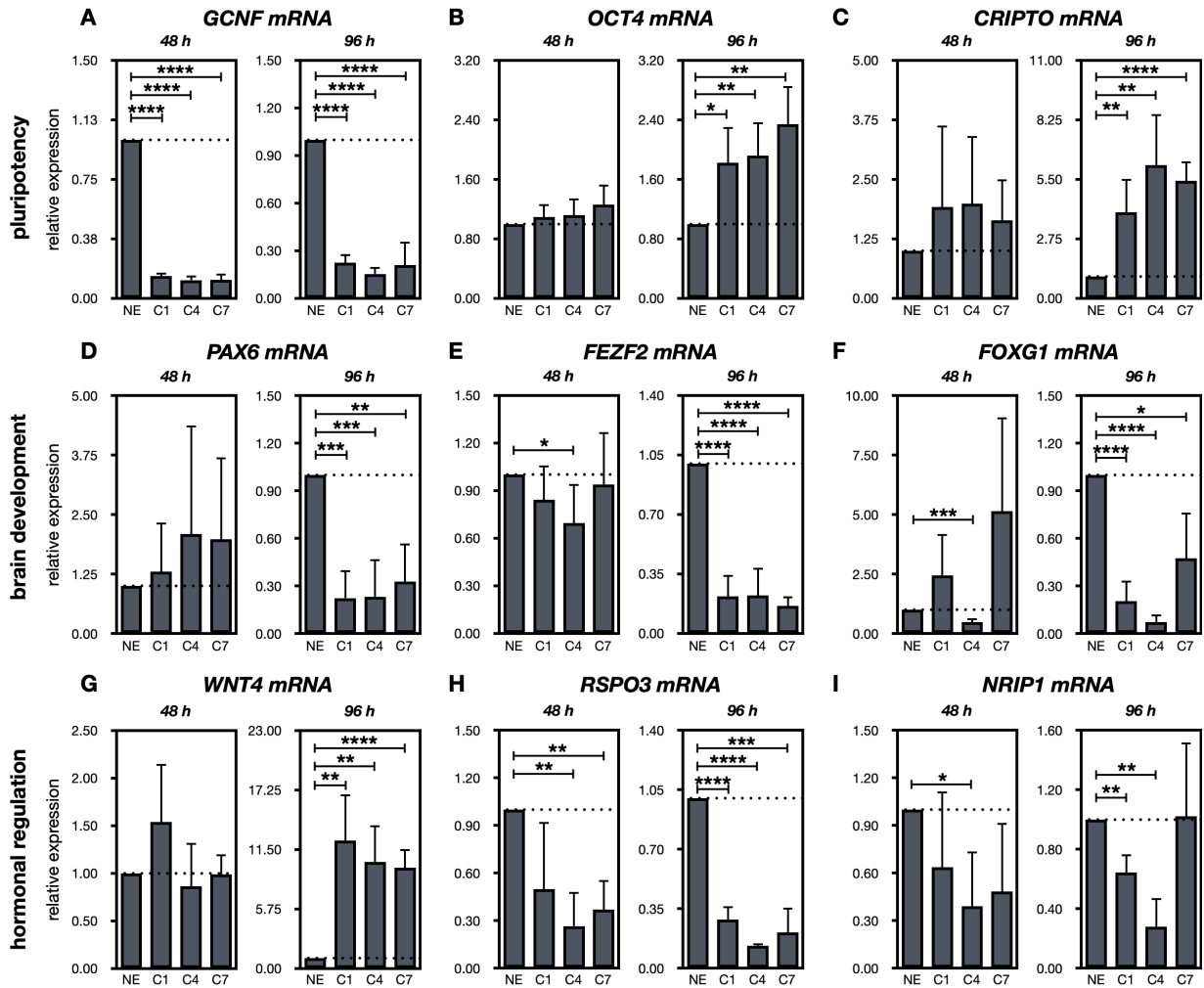
Fig. 21: Genes with significantly altered expression in the 3'-mRNA sequencing were validated by qRT-PCR. These genes were ordered according to their presumable function into three major categories: "pluripotency", "brain development" and "hormonal regulation", all of which have been associated with GCNF function. **(A)** In the category of "pluripotency" GCNF expression was significantly decreased. **(B, C)** Pluripotency-associated genes OCT4 **(B)** and CRIPTO **(C)** were significantly upregulated after 96 h of neural induction. **(D-F)** In the category of "brain development" neural stemness marker PAX6 was significantly decreased **(D)** which was accompanied by a likewise decrease in FEZF2 **(E)** and FOXG1 **(F)**. **(G-I)** The last category addressed "hormonal regulation" and showed a significant upregulation of WNT4 **(G)**, as well as a reduction of RSPO3 **(H)** and NRIP1 **(I)**. All data were normalized to 18s rRNA levels and are presented as mean + SD, relative to the expression in non-edited (NE) cells (set to 1, dashed line; n = 4; RSPO3 and NRIP1 n=3; NRIP1 C7 n=2). *, $p \leq 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$; ****, $p \leq 0.0001$.

Immunocytochemistry for both OCT4 and PAX6, after 96 h of neural induction, confirmed the results on a protein level. A prominent reduction of PAX6 in the GCNF knockout clones as compared to the non-edited control could be observed. On the contrary, reactivity against OCT4 was elevated in the GCNF-ablated clones (Fig. 22).
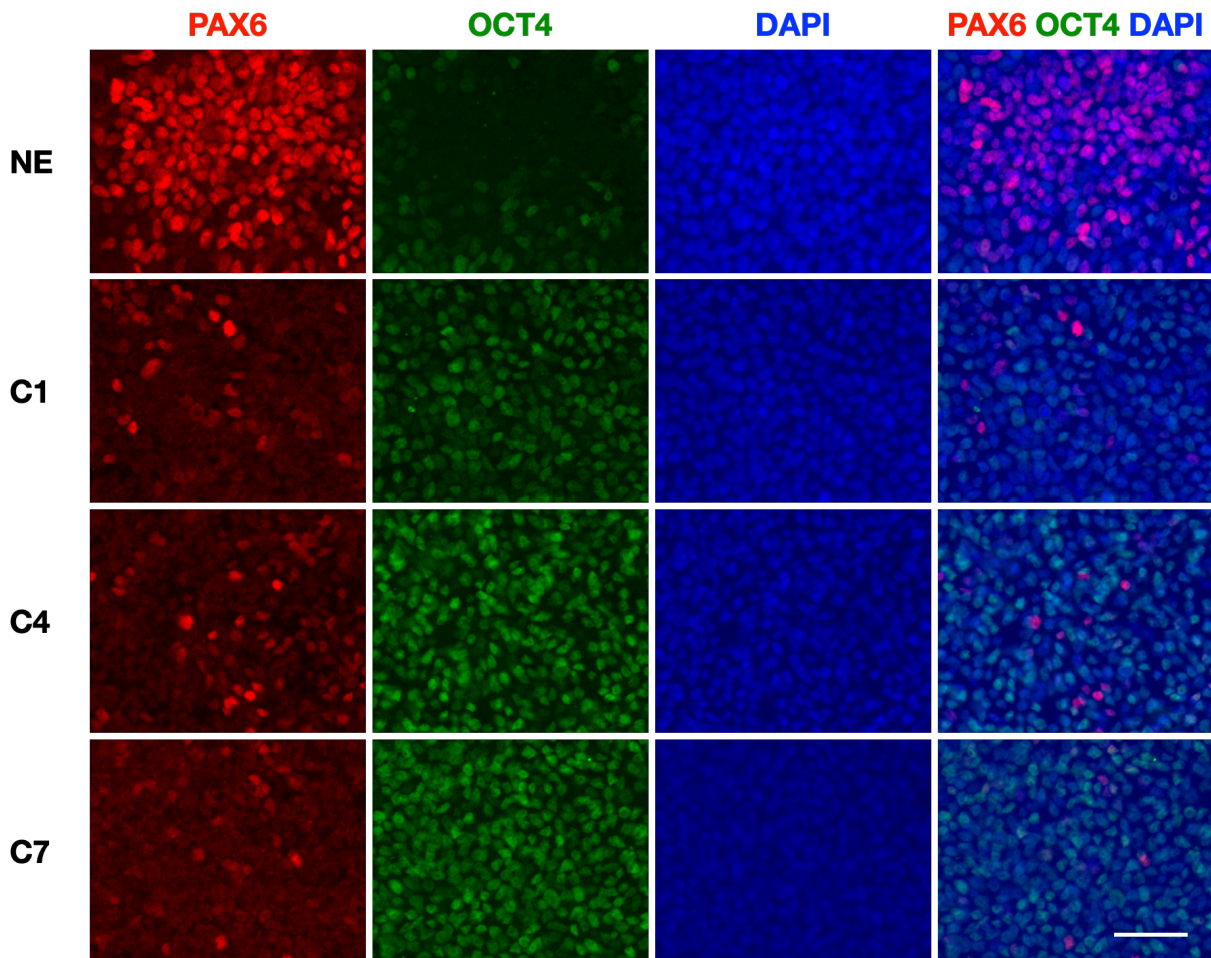
Fig. 22: Immunocytochemistry for OCT4 and PAX6 after 96 h of dual-SMAD inhibition confirms an effect of GCNF ablation on neural induction. GCNF-ablated clones (C1, C4, C7) show higher reactivity of an antibody against OCT4, as compared to the non-edited (NE) control, indicating a decreased capacity to silence the pluripotency network. On the contrary, PAX6 signal shows a prominent reduction in GCNF-ablated clones, suggesting an impairment of neural induction. n=3, exemplary images shown. Scale bar = 50 μm.

3.4 Establishment of a CRISPR/Cas9-mediated C-terminal tagging strategy to assess the GCNF target repertoire with ChIP-seq

During the experiments that were supposed to shed light on how GCNF is involved in neural induction, an important question emerged: are the DEGs that were assessed by 3'-mRNA sequencing, qRT-PCR and immunocytochemistry directly regulated by GCNF and were thus deregulated upon GCNF ablation or were they deregulated as an indirect, secondary effect. To decipher which genes are targeted by GCNF on a genomic level, methods that rely on chromatin immunoprecipitation (ChIP) are frequently utilized. ChIP-experiments rely on fixation of genomic DNA as well as proteins bound to it (Park, 2009). Fragmentation of the DNA followed by an antibody-based pulldown of a specific transcription factor yields DNA fragments that are bound by the respective factor. Subsequent sequencing of the yielded DNA fragments unveils the target repertoire of the transcription factor. The usage of an antibody that targets specifically the transcription factor of interest is crucial for reliable results in a ChIP-seq experiment. However, in our group it proofed to be difficult to utilize GCNF antibodies in ChIP procedures (Klaus et al., in preparation). As the CRISPR/Cas9-mediated GCNF knockout was successfully implemented, the idea was developed to use the same technology to introduce a tag to the C-terminal region of the GCNF gene (Fig. 23 A). This tag might then be visualized with antibodies and used for pulldowns in order to perform ChIP-seq analysis. To that end, the CHOPCHOP sgRNA online tool was utilized to search for an sgRNA that targets the GCNF gene in close proximity of the stop codon (NR6A1 sgRNA 81). Oligonucleotides that code for the sgRNA were ordered and duplexed. Successful incorporation into the pX459 plasmid was confirmed with Sanger sequencing (Fig. 23 B). Additionally, a gene targeting vector was cloned that would allow for targeted recombination upon Cas9-mediated cleavage. The center of the gene targeting vector harbors the desired C-terminal tag, that is followed by a puromycin resistance cassette. The resistance cassette is flanked by loxP sequences, to allow removal of the cassette after successful selection for edited cells. 800 bp of left and right homology arms enclose both the C-terminal tag and the puromycin cassette (Fig. 23 A).
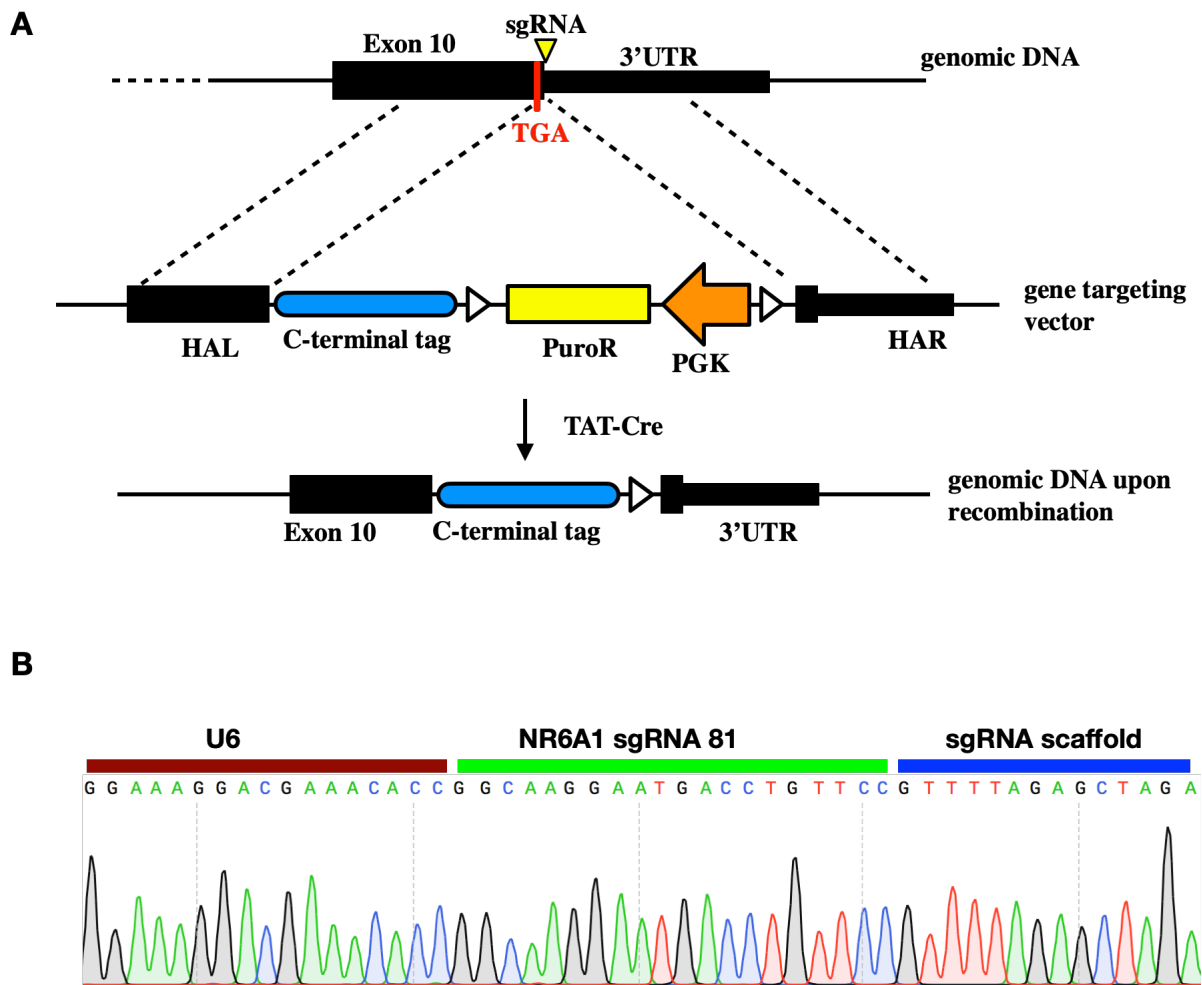
Fig. 23: General strategy for C-terminal tagging of the GCNF gene. **(A)** An sgRNA/Cas9 vector mediates cleavage of the GCNF gene in close proximity of the stop codon (red TGA). Co-transfection of a gene targeting vector provides the template for the incorporation of a C-terminal tag in the endogenous GCNF gene. A loxP-flanked resistance gene allows for transient puromycin selection. The resistance cassette is then removed by Cre recombinase, e.g. by administration of recombinant TAT-Cre. The cutting site of the Cas9 enzyme is indicated by a yellow triangle. **(B)** Sanger sequencing confirms the successful introduction of the sgRNA 81 sequence into the pX459 plasmid that is used to target the C-terminal region of the GCNF gene.

The first C-terminal gene targeting vector was designed to introduce a 3xFLAG tag to the GCNF gene (Fig. 24 A, B). Therefore, oligonucleotides that code for the 3xFLAG tag were ordered and ligated into the gene targeting vector. The plasmid was amplified and hiPSCs were nucleofected with the 3xFLAG tag-containing gene targeting vector as well as the respective Cas9/sgRNA plasmid and subjected to puromycin selection. At the polyclonal level a staining for the 3xFLAG tag was performed in self-renewing conditions.
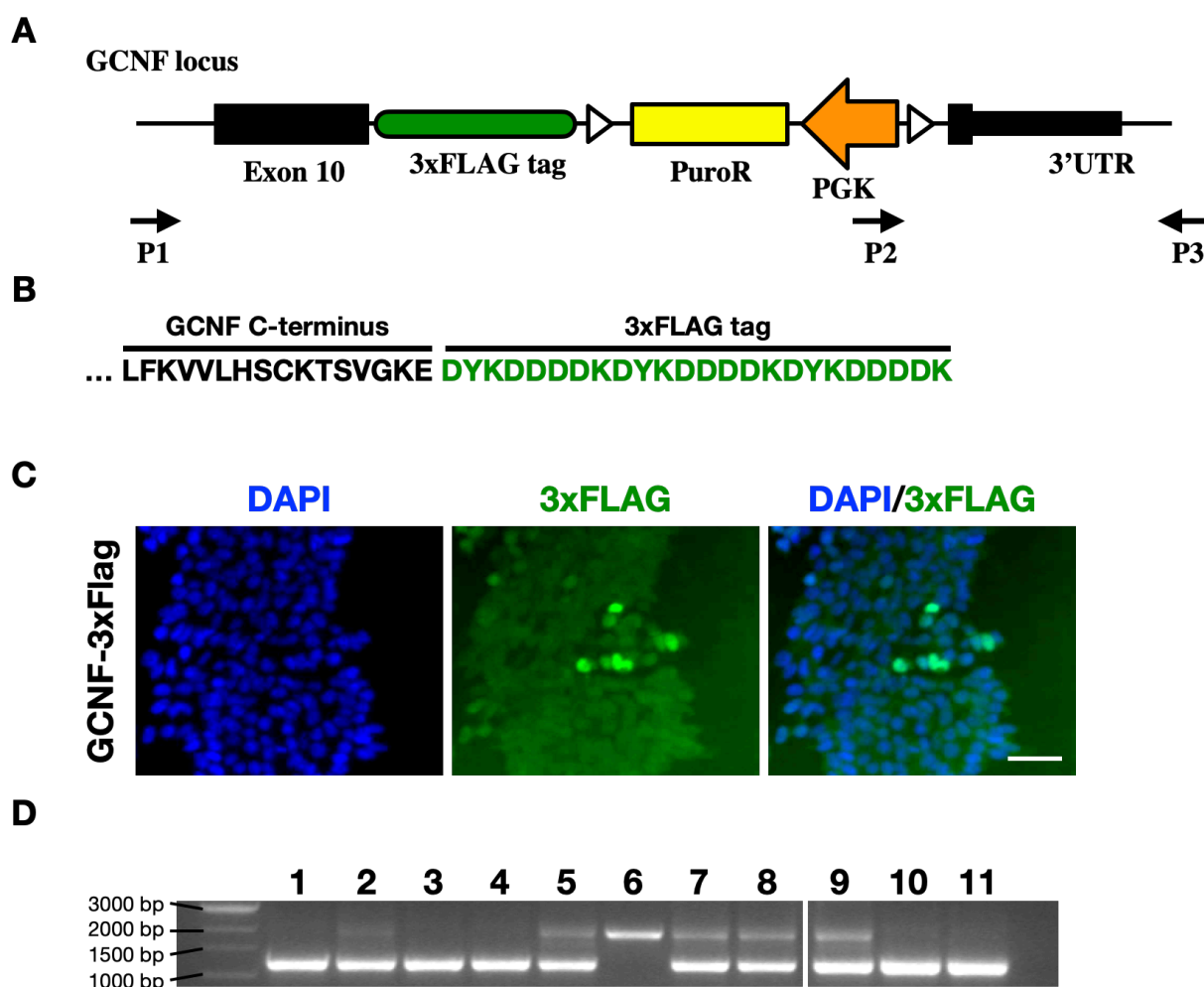
Fig. 24: CRISPR/Cas9-mediated insertion of a 3xFLAG tag in frame with the endogenous GCNF gene. Upon successful Cas9-mediated cutting, homologous recombination mediates insertion of the sequence that encodes for the 3xFLAG tag in frame with the endogenous GCNF gene**.** The sequence is followed by a loxP-flanked puromycin resistance cassette. **(A)** Scheme of the C-terminal GCNF locus upon successful editing. The respective locations of primer binding sites are indicated (P1-3). **(B)** Upon insertion of the sequence at the desired location, the GCNF gene is expressed along with the tag that is fused to its C-terminal region (tag marked in green). **(C)** Staining of the polyclonal, puromycin-selected hiPSCs against the 3xFLAG tag shows strong reactivity in single cells that seem to express the tagged GCNF gene (n=1). Scale bar = 50 µm. **(D)** After picking of single cell-derived clones, a multiplex PCR with all three primers (P1, P2, P3) was performed for genotyping. A 1900 bp band indicates the unaltered genomic location, whereas a 1171 bp band indicates successful editing. Clone 1, 3, 4, 10 and 11 seem to exhibit homozygous editing, clone 2, 5, 7, 8 and 9 are likely to exhibit a heterozygous editing. Clone 6 shows only the endogenous band and thus might exhibit a random integration of the puromycin resistance cassette. The entire agarose gel can be found in Supplementary Fig. 10 A.

Interestingly, the majority of cells did not show reactivity for the 3xFLAG antibody, however some cells did show strong fluorescence (Fig. 24 C). In most of the cells, GCNF protein level may not be high enough to stain at the level of pluripotency, whereas cells that undergo spontaneous differentiation might upregulate GCNF to a detectable level. The C-terminal tagged hiPSCs were subjected to low-density seeding, single cell-derived clones were picked, expanded and subjected to a multiplex genotyping reaction (Fig. 24 D). Genotyping PCR confirmed the presence of heterozygous as well as homozygous editing. However, experiments performed by Frederike Klaus in our group indicated that the ChIP pulldown experiments with the 3xFLAG tag don't yield sufficient amounts of DNA suitable for further analyses.

Since the establishment of ChIP experiments proofed to be difficult, the system was switched to a tag that is utilized by Active Motifs (AM), a company specialized in ChIP-seq strategies. Oligonucleotides encoding for the AM tag were ligated into the C-terminal GCNF targeting vector (Fig. 25 A). After nucleofection of the AM tag-containing gene targeting vector as well as the sgRNA/Cas9 vector into hiPSCs, they were subjected to a transient puromycin selection. Subsequently, the polyclonal pool of cells was exposed to recombinant TAT-Cre protein in order to excise the puromycin resistance gene (Fig. 25 A, B). After TAT-Cre exposure, cells were seeded in single cell density and grown until colonies emerged. These colonies were picked and subjected to a multiplex genotyping reaction. The genotyping results indicated successful excision of the resistance cassette in homozygous and heterozygous hiPSC clones (Fig. 25 C). In homozygous hiPSCs that appeared to have undergone TAT-Cre-mediated recombination, a second PCR was performed to amplify the region that surrounds the AM tag. Sanger sequencing revealed that the entire resistance cassette was successfully excised, leaving only the AM tag inserted in frame with the endogenous GCNF gene followed by a single loxP sequence (usually considered a loxP 'scar') (Fig. 25 D). Thus, the cells were successfully edited to express an AM tag attached to the GCNF C-terminus. These cells are currently used to identify the target gene repertoire of GCNF by utilization of ChIP-seq (Klaus et al., in preparation).
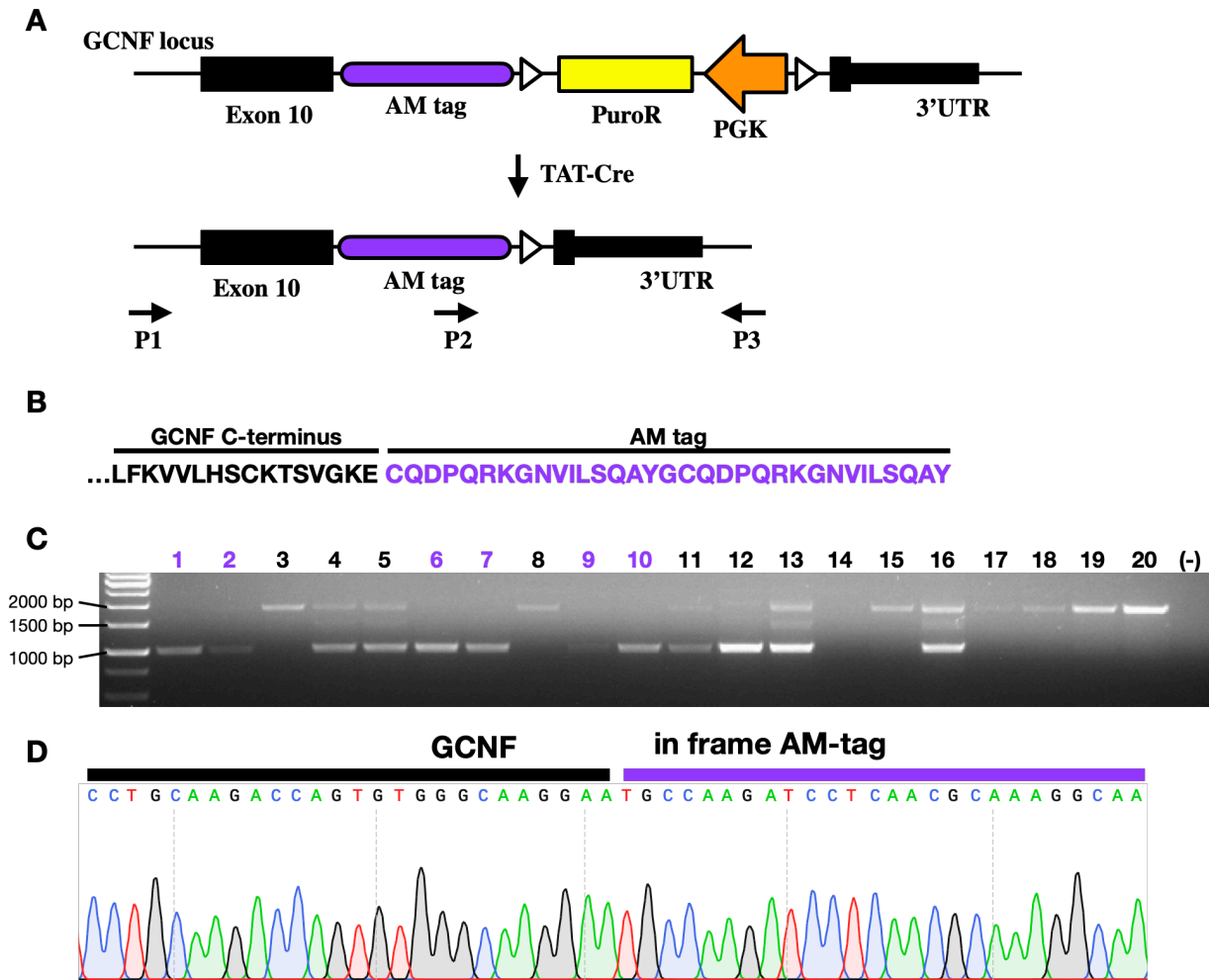
Fig. 25: CRISPR/Cas9-mediated insertion of an AM tag in frame with the endogenous GCNF gene. **(A)** The scheme depicts the TAT-Cre-mediated excision of the puromycin cassette in edited hiPSCs. Primer binding sites are indicated (P1-3). **(B)** Upon insertion of the sequence at the desired location, the GCNF gene is expressed along with the AM tag that is fused to its C-terminal region (tag marked in purple). **(C)** After single cell seeding and picking of colonies, the multiplex genotyping of monoclonal hiPSCs indicates successful heterozygous and homozygous editing as well as excision of the puromycin resistance cassette by recombinant TAT-Cre. Whereas the endogenous band exhibits a size of 1900 bp, the edited band exhibits a size of 1014 bp, when the puromycin cassette was excised. Single cell-derived hiPSCs in which the PCR reaction indicates a successful Cre recombination are marked as purple. The entire agarose gel can be found in Supplementary Fig. 10 B. **(D)** Sanger sequencing confirms the insertion of the AM tag encoding sequence. It did also show the successful excision of the puromycin resistance gene (data not shown).

4. Discussion

The first part of the discussion will focus on the biological question that was addressed in this thesis, namely how GCNF deficiency impacts on early germ layer specification and neural induction of hiPSCs. The second part will discuss implementation as well as potential improvements of the CRISPR/Cas9 system utilized in this study, followed by an outlook to future experiments.

4.1 GCNF-ablated hiPSCs are capable of self-renewal in the pluripotent state but show alterations in early germ layer specification

A major motivation for this work was the fact that a variety of studies suggest GCNF being involved in early germ layer specification and neurodevelopment (Akamatsu et al., 2009; Bauer et al., 1997; Chung et al., 2006; David et al., 1998). However, most of the studies were performed in mice or teratocarcinoma cell lines. As access to human tissue of this stage of development is limited or impossible, human stem cells may be a valuable resource to obtain this information. First of all, it was an important step to characterize the single cell-derived GCNF-ablated hiPSCs at the stage of self-renewal. As GCNF is a well described negative regulator of OCT4, it is interesting, that its disruption did not lead to an upregulation of pluripotency-associated transcription factors at that stage. Instead, OCT4 levels were close to the non-edited level. The same was also true for CRIPTO which is another known target of GCNF (Hentschke et al., 2006). The absence of morphological abnormalities, a regular proliferation and presence of the pluripotency markers OCT4 and SOX2 indicate that loss of GCNF is not significantly interfering with hiPSC self-renewal and the maintenance of the undifferentiated state under *in vitro* conditions. However, it is uncertain whether the GCNF-ablated clones exhibit more subtle changes of pluripotency, a question that can only be addressed by more sophisticated methods like a teratoma-assay, or a deeper analysis of the pluripotency network by bioinformatical approaches. The following experiments aimed to investigate potential alterations in germ layer differentiation and whether these alterations are especially pronounced in the neuroectodermal fate.

4.1.1 GCNF-ablated cells exhibit altered germ layer specification in an EB-based undirected differentiation paradigm

The first attempt to tackle the addressed question was the undirected differentiation of hiPSCs as EBs for 14 days and the subsequent extraction of RNA. QRT-PCR analysis confirmed GCNF ablation on mRNA level. On the other hand, OCT4 and CRIPTO expression proofed to be upregulated. The mRNA levels of the early neural progenitor marker PAX6 were slightly downregulated. Subsequently, a commercial Scorecard assay was utilized in the undirected differentiation experiment to uncover alterations in germ layer specification in differentiating GCNF-ablated hiPSCs. The germ layer scores, which are computed by the Thermo Fisher Scientific hPSC Scorecard analysis software supported the hypothesis of an ectodermal differentiation impairment of the knockout clones, although statistical significance was only reached in the ectoderm score of C1. A closer look at the underlying genes that are assessed by the Scorecard assay proofed to be valuable since most of the significantly changed genes in the experiment were assigned ectodermal genes, among them PAX3 and WNT1. PAX3 is primarily known for being involved in neural crest development and mutations in the gene might cause Waardenburg syndrome, which is characterized by facial dysmorphias as well as deafness (Wang et al., 2008). Additionally, it is associated with defects in neural tube closure (Sudiwala et al., 2019; Zhou and Conway, 2016). WNT1 has been reported to be a key factor in the formation of the isthmic organizer and the developing midbrain (Thomas and Capecchi, 1990) and also appeared to be consistently downregulated in the Scorecard assay. This is in line with previously published data in GCNF knockout mice. In an important study, GCNF knockout mice show a strong downregulation of WNT1 in the region of the isthmic organizer (Chung et al., 2006). Interestingly, PAX2 and PAX5, protein paralogues of PAX3, are impaired in the same study (Chung et al., 2006).

Another interesting candidate that has been consistently downregulated in GCNF-ablated clones appeared to be TRPM8. TRPM8 is widely acknowledged as a Ca^{2+} channel that is activated by menthol and mediates the perception of cold in the peripheral nervous system (Dhaka et al., 2008). Additionally, it has also been found to be present in many regions of the rodent brain (Beukema et al., 2018). Due to the implication of GCNF in gonad development, it is especially interesting, that TRPM8 is highly expressed in both prostate

and prostate cancer where it appears to be upregulated upon androgen exposure (Asuthkar et al., 2015).

Within the mesodermal markers, HAND1, a crucial factor for heart development (McFadden et al., 2005) exhibited higher mRNA levels in the GCNF knockout clones. Its upregulation might be a contributing factor to cardiac failure observed in GCNF knockout mice (Chung et al., 2001). Also, EOMES, a factor that is crucial for endoderm differentiation, appears to be consistently upregulated in GCNF knockout clones in the Scorecard assay. Interestingly, EOMES is reported to repress both pluripotency as well as neuroectodermal specification to promote mesodermal and endodermal specification (Tosic et al., 2019). While EOMES favors endodermal and mesodermal lineages and GCNF might favor the neuroectodermal lineage, both might synergistically repress the pluripotency network.

Despite a reduction of the ectoderm score and alterations in underlying ectodermal genes, the variation in the experimental data was larger than expected. Although the utilized EB-based protocol is advantageous to investigate the capacity of the cells to differentiate into all three germ layers at the same time, the undirected and spontaneous nature of the assay might be a cause of variation. Additionally, the collagenase IV-based generation of EBs could result in size variation as the detached cell layer might break into smaller clusters in an undefined way. To overcome the variability that arises from spontaneous differentiation into all germ layers, non-edited and GCNF-ablated hiPSCs could be differentiated into each germ layer separately by directed differentiation protocols that either mediate differentiation into ectodermal, mesodermal or endodermal fate. Size variation could be approached by the utilization of AggreWells™ in which a defined amount of hiPSCs is seeded in special low attachment wells, which promote the generation of uniformly sized and shaped EBs, and hence may serve in reducing variation.

4.1.2 GCNF-ablated cells exhibit impaired neural induction

Given the crucial role of GCNF in the establishment of the mouse nervous system (Chung et al., 2006), one of the key questions of this study was whether GCNF is also playing such a role in a human stem cell model of early neuroectodermal specification. To address this question, a dual-SMAD inhibition-based neural induction was conducted. It is also

advantageous that such a 2D culture results in a more defined and homogenous condition for a cell population, whereas a 3D culture is characterized by various compartments with heterogeneous cell niches. Thus, a 2D culture might reduce the amount of variation that had been observed in the EB-based experiment. Additionally, the dual-SMAD inhibition drives the entire system into the neural lineage, thereby further reducing the noise that originates from spontaneous and undefined differentiation into multiple germ layers in undirected approaches (Chambers et al., 2009). One important question was, at which time point GCNF elicits its main effect on the early neural lineage entry in differentiating hiPSCs. In several GCNF studies effects are investigated in a time window from 48 h to several days after pluripotent stem cell or teratocarcinoma cell differentiation. Most of the studies describe significant effects after already 48 h of differentiation (Fuhrmann et al., 2001; Wang et al., 2013). A GCNF expression-dependent lacZ reporter indicates strong GCNF upregulation after 72 h of RA-mediated differentiation in mESCs (Lan et al., 2009). In this thesis, time points of 48 h and 96 h of neural induction were chosen because it allows the assessment of a temporal dynamic. The 3'-mRNA sequencing was performed for the 96 h timepoint of one of the four experiments as that timepoint was supposed to allow observation of more pronounced effects of GCNF depletion. Indeed, validation of the transcriptomic analysis with qRT-PCR confirmed the 96 h condition to show the more prominent effects. DEGs were either viewed as potentially linked to pluripotency, hormonal regulation or importance during brain development. Apart from GCNF, the 3'-mRNA sequencing yielded PAX6, FEZF1 and FEZF2 among the top DEGs. This finding was confirmed for both PAX6 and FEZF2 by qRT-PCR. PAX6 is among the most important transcription factors in neurodevelopment and an early indicator of neural fate commitment (Götz et al., 1998; Thakurela et al., 2016). Both FEZF1 and FEZF2 are known to drive NPCs into a telencephalic identity (Shimizu et al., 2010). Importantly, FEZF2 overexpression in mESCs lead to an upregulation of PAX6 and FOXG1 (Wang et al., 2011). FOXG1 is an established marker of neural precursors that are capable of giving rise to cortical neurons (Shi et al., 2012). Additionally, it was shown to be impaired in the forebrain of developing GCNF knockout mouse embryos (Chung et al., 2006). A single-gene query of FOXG1 in the transcriptomic analysis indicated a reduction in the GCNF-ablated clones although the read-counts were very low (data not shown). Therefore, FOXG1 was also assessed in this study despite not being among the top 30 DEGs of the

3'-mRNA sequencing experiment. Indeed, consistent with previous data, FOXG1 expression was decreased in GCNF-ablated cells subjected to the neural induction paradigm. The impairment of FOXG1 expression might be a result of FEZF2 downregulation. Notably, ChIP-seq experiments indicated FEZF2 being regulated by FOXG1-binding to the FEZF2 promoter and specific enhancer regions (Eckler et al., 2014). The combinatorial depletion of transcription factors, that are keys for early brain patterning as well as neurogenesis, clearly indicate that GCNF is playing an important role in these processes. Thus, it might be, that GCNF, apart from mediating exit of pluripotency, also activates a neural induction or even a telencephalic gene expression program. This hypothesis could be investigated by a time series analysis during a cortical differentiation. A very promising tool allows the conditional ablation of a gene in hiPSCs (Chen et al., 2015). A conditional ablation of GCNF at various timepoints of a cortical differentiation could clarify at which steps GCNF elicits its main function. The role of GCNF as an important factor of neural progenitors is further supported by an important study on epigenetic remodeling during *in vitro* neural differentiation from hPSCs to neurons. In this study, GCNF was among the top DEGs between the pluripotent stem cell stage and the neuroepithelial stem cell stage (Ziller et al., 2015). Additionally, the study emphasizes the role of PAX6 and FOXG1 in the neuroectodermal commitment. Whereas Ziller et al. (2015) place GCNF as an equal part of the neural core network, it may rather be, that it acts on top of these factors. In such a model GCNF would mediate exit of pluripotency and then either directly or indirectly activate factors that drive neuroectoderm differentiation. An interesting tool that could help to clarify that relationship would be a GCNF reporter. Cells with a high activity of the GCNF reporter could be compared to cells with a low expression of the GCNF reporter in early stages of neuroectodermal specification. An enrichment of neural transcription factors in cells with a higher GCNF reporter activity would indicate that GCNF is directly or indirectly activating genes that drive the cells into the neuroectoderm. This approach might be combined with a close monitoring of neuroectodermal transcription factors during differentiation from hiPSCs to early neuroectodermal progenitors.

Whereas the experiments that were performed in this thesis indicate that the neuroectodermal lineage specification is indeed altered it is important to consider that GCNF deficiency results in a delayed exit of pluripotency, as indicated by upregulated

levels of OCT4 and CRIPTO. Therefore, deficits in neuroectoderm specification may not be a direct effect of GCNF deficiency but rather originate from persistent OCT4 expression and as a consequence, activation of the pluripotency network. Only further experimentation is capable to clarify whether deficiencies in neural induction are an indirect effect of delayed silencing of pluripotency, or whether loss of GCNF is the primary cause of a disrupted neural induction. To validate the effect of GCNF on neural induction it would be interesting to perform an experiment, in which key components of the pluripotency network, like NANOG or OCT4 are silenced by siRNAs. If the deficits in neural induction are still present in the GCNF-ablated hiPSCs, it would be a strong indicator of GCNF being a positive regulator of early neural induction.

It is also worth noting, that there is growing evidence that factors that mediate pluripotency are also involved in lineage entry. In an important study, OCT4 was found to suppress the neural lineage, and proofed to be crucial for mesendodermal differentiation, whereas SOX2, another pluripotency factor, mediates the transition of pluripotent cells into the neuroectoderm (Thomson et al., 2011). By directly repressing OCT4, but not SOX2, GCNF might shift the balance between OCT4 and SOX2 towards the latter and therefore simultaneously initiate the exit of pluripotency as well as the entry into the neural lineage. In such a model the two processes differentiation and lineage selection would be strongly connected.

4.1.3 GCNF-ablated cells show alterations in WNT signaling

A direct link between FEZF2 expression and WNT signaling in mESCs has been shown since mESC differentiation for 4 days, followed by FEZF2 overexpression for 48 h results in significant downregulation of WNT1, WNT3a and WNT8b, whereas WNT4 shows a non-significant upregulation (Wang et al., 2011). Adversely, in the neural induction experiment of this thesis, GCNF ablation leads to a highly significant WNT4 upregulation. In the undirected differentiation experiment, GCNF-ablated clones show an impaired WNT1 expression, which is also true for FEZF2 overexpressing mESCs (Wang et al., 2011). Thus, in mESCs, FEZF2 gain-of-function positively regulates WNT4 and negatively regulates WNT1, whereas GCNF ablation in hiPSCs leads to FEZF2 downregulation accompanied by WNT4 upregulation in the neural induction paradigm and WNT1

decrease in the undirected differentiation paradigm. This might indicate that FEZF2 acquired a different role in WNT regulation in human neurodevelopment as compared to the mouse system. It is well-known that WNT molecules play a crucial role in brain development at almost all stages (Mulligan and Cheyette, 2012). WNT1 acts primarily through canonical WNT signaling (Niehrs, 2012) and is required for proper formation of the midbrain (Thomas and Capecchi, 1990). Oppositely, WNT4 is a crucial factor secreted by the floor plate of the spinal cord to guide commissural axons to its anterior target (Lyuksyutova et al., 2003; Mulligan and Cheyette, 2012). WNT4 is associated with the non-canonical WNT pathway and inhibits canonical WNT signaling by redistribution of β-catenin to the cell membrane (Bernard et al., 2008). Additionally, a WNT4 gain-of-function mouse model results in impaired testicular vasculature (Jordan et al., 2003). Given the fact that GCNF is highly enriched in the testis, GCNF might be an important regulator of WNT4. It would be very interesting to assess whether the WNT4 gene contains GCNF binding sites. This question could be investigated with a luciferase assay. In such an experiment the WNT4 promoter would be cloned upstream of a luciferase gene. If WNT4 is indeed inhibited by functional GCNF it would suggest an elevated level of bioluminescence in the GCNF-ablated clones. In a broader approach this question could also be tackled by a ChIP-based experiment. Endogenous tagging of the GCNF gene with CRISPR/Cas9 and a subsequent ChIP-seq could reveal genes that are regulated by GCNF. The alignment with transcriptomic data would allow to identify genes that are downregulated by GCNF and potentially reveal genes that are upregulated by GCNF.

Interestingly, opposed to WNT4, RSPO3 expression was consistently down-regulated in the GCNF-ablated clones in the bioinformatical as well as the qRT-PCR analysis and might be an interesting part of dysregulated WNT signaling in GCNF-ablated clones. R-spondins are a class of secreted proteins that activate canonical WNT signaling through WNT receptors (Cambier et al., 2014; de Lau et al., 2012).

WNT1 and RSPO3 mediate canonical WNT signaling, whereas WNT4 acts on non-canonical WNT signaling and can therefore be regarded as counteracting. Thus, GCNF ablation may lead to upregulation of WNT4 which shifts the balance towards non-canonical WNT signaling. That may subsequently lead to a loss of RSPO3-mediated canonical WNT signaling. A possible way to gain deeper insights into how GCNF affects WNT signaling would be the utilization of a WNT reporter construct. In an established

system an array of WNT response elements (7xTCF) is cloned upstream of an eGFP gene and thus enables monitoring of canonical WNT signaling (Fuerer and Nusse, 2010). Sorting of cells into populations of high and low WNT activity and subsequent characterization may shed light on how GCNF modulates canonical and non-canonical WNT signaling. It is also noteworthy, that RSPO3 exhibits a strong expression in the dorsal telencephalon (Aoki et al., 2007). RSPO3 is also involved in vascular development (Kazanskaya et al., 2008). As the dual-SMAD inhibition is yielding a homogenous culture of neuralized cells, it would be interesting to investigate RSPO3 levels in the context of a mesodermal differentiation paradigm.

4.1.4   GCNF ablation might cause dysregulation of nuclear receptor co-factors

GCNF belongs to a family of genes, which are important for mediating the response of hormones such as RA or steroids. Therefore, the list of DEGs was especially searched for genes that are involved in hormonal regulation. In the 3'-mRNA sequencing experiment NRIP1 was among the top downregulated DEGs in the GCNF-ablated clones. This was aligned with results from qRT-PCR that also indicated a significant downregulation of NRIP1. Notably, NRIP1 is reported of being highly upregulated upon RA signaling and is reported to mediate repression of ligand-bound nuclear receptors (Kerley et al., 2001). NRIP1 is described to mainly bind to the AF-2 domain, a domain that is present in the LBD of most nuclear receptors, but is not conserved in GCNF (Greschik et al., 1999; Kerley et al., 2001; Weikum et al., 2016). Kerley et al. (2001) promote a model in which NRIP1 limits RA signaling. The same lab differentiated the human EC cell line NT2/D1 with all-trans RA towards the neuronal lineage for 24 h and assessed via microarray analysis how a siRNA against NRIP1 alters gene expression. Interestingly, GCNF is among the top DEGs in this analysis (Heim et al., 2007), although this might be a site effect of an NRIP1 depletion-mediated enhancement of RA signaling. The fact that NRIP1 mediates repression of target genes and also exhibits intrinsic repressive activity may suggest that it is also mediating GCNF target gene repression, despite the alteration in the AF-2 domain. However, in the model proposed by Heim et al., NRIP1 mRNA levels strongly increase upon RA signaling. Subsequently, it negatively regulates a subset of RA-inducible genes, among them being GCNF (Heim et al., 2007). The down-regulation

of NRIP1 in the neural induction experiment of this thesis may thus be a consequence of decreased GCNF-mediated RA signaling. Another study describes NRIP1 being recruited to RA receptor complexes, thereby mediating its repressive activity (Hu et al., 2004). Interestingly, NRIP1 is also expressed in the brain and female mice that lack NRIP1 are infertile, due to a failure to release oocytes at ovulation (Parker et al., 2003). This is in line with the functionality of GCNF in both gonad and brain development. A method to further investigate the relationship of GCNF and its co-factors might be a co-immunoprecipitation experiment which would clarify whether there is indeed an interaction between GCNF and NRIP1. This approach might also yield further binding partners that are involved in GCNF-mediated repressive activity. Additionally, as NRIP1 appears to be tightly connected to RA signaling, RA treatment of the differentiating non-edited and GCNF-ablated hiPSCs may yield further insights into how GCNF and NRIP1 are connected to each other.

4.2 Implementation and potential improvements of CRISPR/Cas9-mediated genome editing in hiPSCs

One of the primary goals of this study was to establish a CRISPR/Cas9 system for mediating genome editing in hiPSCs. A widely-used plasmid-based system was utilized, as it is cost-effective and versatile. Its implementation as a standardized procedure for CRISPR/Cas9-mediated genome editing was successfully achieved. Especially customizing the sgRNA sequence to target a specific genomic location proofed to be robust and makes the technique a powerful "everyday" method for scientists. Although the study yielded insights into the biological questions addressed, it might be improved in a technical way. Whereas cloning of the gene targeting vectors for C-terminal tagging was achieved readily, cloning of the gene targeting vector for gene ablation turned out to be more challenging, as cloning of the right homology arm could only be achieved after several trials and switching of the polymerase. Additionally, it was only partially possible to sequence it, most likely due to a high GC content or secondary structures. The fragment was also smaller than expected which may result from length variation, since it is mostly derived from an intronic region, which usually exhibit a higher degree of polymorphism than exonic regions (Rigau et al., 2019). The high GC content made it also difficult to establish the genotyping PCR. Interestingly, despite the expected lack of the endogenous

PCR product in the knockout clones another smaller PCR product emerged. It is likely that it is an unspecific product originating from another genomic region, but this could only be clarified by gel extraction and subsequent Sanger sequencing. It is also important to note that the genotyping relied on a touchdown PCR that was run for over 40 cycles to exclude the presence of the endogenous allele with higher confidence, which might in turn lead to amplification of unspecific products.

Taking the difficulties into account that arose from cloning the gene targeting vector and establishing a genotyping PCR for the insertion of the puromycin cassette, it might be possible to omit that time-consuming step in future studies by solely relying on NHEJ. Nucleofection of the sgRNA/Cas9 plasmid might already result in a decent number of targeted clones, that can be identified by Sanger sequencing. A highly efficient system that has recently been implemented in our group, relies on delivery of recombinant Cas9 ribonucleoproteincomplexes (RNPs). In such a system a synthetic customized crRNA is assembled with a universal tracrRNA. CrRNA/tracrRNA-loaded RNPs usually result in a high frequency of knockout clones, that can also be identified by Sanger sequencing. To further increase the reliability of the observed phenotypes and to exclude off-target effects, it would be interesting to compare the GCNF-ablated clones to non-edited subclones from the same nucleofection batch. It would also be interesting to repeat the results in hiPSCs from different donors to account for effects that arose from the genomic background. Taking the crucial role of GCNF in germ cells into account as well as the fact that the top DEGs of the neural induction experiment contained sex-specific genes it would be especially interesting to investigate whether there is a phenotypic difference in GCNF knockout clones from male and female hiPSC lines. Sexual dimorphism in early brain development might in fact lead to a differential response to hormonal stimuli and thus alter neural phenotypes in GCNF knockout studies.

Numerous studies emphasized the risk of off-target effects in CRISPR/Cas9-mediated genome editing experiments (Zhang et al., 2015). Although the cell lines that had been generated in this thesis did show overall genomic integrity in the SNP analysis, this does not exclude modifications at a single nucleotide level. Even a random insertion of the gene targeting vector is challenging to detect. Whereas the genotyping PCR can indicate successful integration at the desired location, it does not rule out the possibility of a random integration of the gene targeting vector in an undesired location. The risk of off-

targets could either be improved by a deeper off-target analysis and a more rigorous selection of clones, or by utilization of more specific genome editing tools. For example, a TaqMan-based copy number analysis could be utilized to determine the number of integrated resistance cassettes into the genome. Another possibility would be an NGS-based quality check that would be capable to detect more subtle changes, like single nucleotide modifications. The genome editing machinery might also be engineered in a way to reduce off-targets. This has been achieved by utilizing nickases. These Cas9 variants are deficient for generating a full DSB but instead cleave only one strand of the double helix, resulting in significantly reduced off-target effects (Ran et al., 2013b). Another elegant study utilized a bacterial screen to identify Cas9 variants, that have high on-target, and low off-target activity. This led to the generation of a High-Fidelity Cas9 enzyme (HiFi-Cas9) which can also be purchased as a recombinant protein (Vakulskas et al., 2018). The application of recombinant Cas9 proteins has the additional advantage of controlled temporal exposure of the Cas9 enzyme, which leads to a reduced frequency of off-targets. Notably, especially cells that are known to have damaged DSB repair pathways are prone to off-target effects, whereas hPSCs do not appear to have a strong off-target tendency, probably due to an intact DSB repair mechanism (Smith et al., 2014; Veres et al., 2014).

4.3   Outlook

The most important part of future investigations will be to further characterize the defects that are caused by GCNF ablation. As the neural induction paradigm already identified interesting candidate genes in a very early neurodevelopmental stage, it will be of interest to cultivate GCNF-ablated clones for a longer period of time. Therefore, a 10 day differentiation might be a well-suited starting point. This period of time could even be extended to a full cortical differentiation that might yield interesting insights whether GCNF participates in sophisticated processes like radial glia generation or cortical layering. Especially an organoid or a more standardized spheroid paradigm may address the question, whether GCNF ablation may result in an impaired neurogenesis, or whether it is disturbing the cortical 3D architecture. To identify GCNF expressing niches cortical organoids could be subjected to single cell RNA sequencing. This could help to investigate

whether some regions express GCNF also in later stages of neurodevelopment. To fully understand how GCNF is acting on neurodevelopment, it is also necessary to understand its target gene repertoire. One of the most insightful future experiments could be a ChIP-seq experiment. A very important question that could be addressed with that methodology is whether GCNF is exclusively binding regions of DNA that are upregulated in the knockout. This would confirm its repressive activity. If the ChIP-seq reveals enrichment of downregulated DEGs, it would also suggest a potential function as a transcriptional activator. As there are no working ChIP-grade antibodies available, the strategy of a CRISPR/Cas9-mediated tagging of the endogenous GCNF gene, as described in this thesis, might be a valuable tool. Similarly, it would be interesting to assess the interactome of GCNF with a co-immunoprecipitation experiment. Maybe this could shed light on the binding partners of GCNF and whether NRIP1 is a downstream effector of its repressive activity. Rescue experiments could be done by overexpression of candidate genes, like PAX6 or FEZF2, that exhibit decreased levels in GCNF-ablated clones. Oppositely, it may be interesting to overexpress GCNF in differentiating hiPSCs to investigate whether it alters germ layer specification or increases neuroectodermal specification. Reporter constructs may facilitate investigation of the pathways involved in GCNF-dependent regulation. Especially for prominent signaling pathways like canonical WNT signaling such constructs are readily available and may yield useful insights how GCNF and WNT signaling are connected. It would also be interesting to utilize more advanced CRISPR/Cas9 systems to investigate GCNF function, like dCas9 to target GCNF binding sites in order to mimic the repressive activity of GCNF. By utilizing a CRISPR screen on a GCNF reporter cell line, negative and positive regulators of GCNF levels might be uncovered. To investigate whether GCNF deficiency leads to a general differentiation defect or a neuroectoderm-specific impairment, a conditional knockout of GCNF might be a valuable tool. With a conditional system GCNF could be removed at arbitrary points during neural induction, to investigate whether an impairment in neural differentiation persists even after the cells already entered the neuroectodermal lineage.

5. Summary

Nuclear receptors are widely recognized as an important layer of transcriptional regulation and are usually activated by a signaling molecule. Nuclear receptors without a known ligand are designated orphan nuclear receptors. One of those receptors is GCNF. Whereas it is known that GCNF serves an important role in gonad development and its impact on silencing OCT4 upon differentiation of hPSCs is widely appreciated, less is known about its implication in germ layer specification and early neural induction. To tackle this question, the CRISPR/Cas9 technology was used to generate GCNF-deficient hiPSCs. All quality checks like stable pluripotency marker expression, as well as assessment of genomic integrity by SNP analysis, were passed successfully by the GCNF-deficient hiPSC clones. In self-renewing conditions GCNF ablation could be confirmed on mRNA as well as protein level. An undirected embryoid body-based differentiation experiment combined with a TaqMan-based analysis of differentiation potential, pointed towards an impaired ectodermal differentiation during germ layer specification. To investigate the impact of GCNF-deficiency more specifically on neuroectoderm formation, a directed neural induction experiment with GCNF-deficient hiPSC clones was performed and assessed by RNA-seq, qRT-PCR and immunocytochemistry. This approach revealed candidate genes that might be key players in GCNF downstream signaling. While OCT4 and CRIPTO, two known GCNF target genes, were found to be upregulated, the expression of key factors of early neurodevelopment like PAX6, FEZF2 and FOXG1 were significantly impaired upon GCNF ablation. Additionally, upregulation of WNT4 and reduced RSPO3 expression indicate GCNF being involved in WNT signaling. For future ChIP-seq experiments to identify GCNF target genes, CRISPR/Cas9 technology was utilized to generate hiPSC lines with ChIP-compatible tags (FLAG and AM) added to the C-terminus of the endogenous GCNF gene. In summary, CRISPR/Cas9 technology was successfully employed to generate an isogenic set of GCNF-deficient hiPSCs, that revealed a developmental impairment during germ layer specification and early neural induction.

6. List of figures

7. List of tables

8. References

Agoulnik IY, Cho Y, Niederberger C, Kieback DG and Cooney AJ. Cloning, expression analysis and chromosomal localization of the human nuclear receptor gene GCNF. FEBS Letters 1998; 424: 73-78

Akamatsu W, DeVeale B, Okano H, Cooney AJ and van der Kooy D. Suppression of Oct4 by germ cell nuclear factor restricts pluripotency and promotes neural stem cell development in the early neural lineage. J Neurosci 2009; 29: 2113-2124

Aoki M, Mieda M, Ikeda T, Hamada Y, Nakamura H and Okamoto H. R-spondin3 is required for mouse placental development. Dev Biol 2007; 301: 218-226

Asuthkar S, Velpula KK, Elustondo PA, Demirkhanyan L and Zakharian E. TRPM8 channel as a novel molecular target in androgen-regulated prostate cancer cells. Oncotarget 2015; 6: 17221-17236

Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA and Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes. Science 2007; 315: 1709-1712

Barreto G, Borgmeyer U and Dreyer C. The germ cell nuclear factor is required for retinoic acid signaling during Xenopus development. Mechanisms of Development 2003a; 120: 415-428

Barreto G, Reintsch W, Kaufmann C and Dreyer C. The function of Xenopus germ cell nuclear factor (xGCNF) in morphogenetic movements during neurulation. Developmental Biology 2003b; 257: 329-342

Bauer UM, Schneider-Hirsch S, Reinhardt S, Pauly T, Maus A, Wang F, Heiermann R, Rentrop M and Maelicke A. Neuronal cell nuclear factor--a nuclear receptor possibly involved in the control of neurogenesis and neuronal differentiation. Eur J Biochem 1997; 249: 826-837

Belfort M and Bonocora RP. Homing endonucleases: from genetic anomalies to programmable genomic clippers. Methods Mol Biol 2014; 1123: 1-26

Bernard P, Fleming A, Lacombe A, Harley VR and Vilain E. Wnt4 inhibits β-catenin/TCF signalling by redirecting β-catenin to the cell membrane. Biology of the Cell 2008; 100: 167-177

Beukema P, Cecil KL, Peterson E, Mann VR, Matsushita M, Takashima Y, Navlakha S and Barth AL. TrpM8-mediated somatosensation in mouse neocortex. J Comp Neurol 2018; 526: 1444-1456

Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim YG and Chandrasegaran S. Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. Mol Cell Biol 2001; 21: 289-297

Bibikova M, Golic M, Golic KG and Carroll D. Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. . Genetics 2002; 161: 1169-1175

Bier E. Drosophila, the golden bug, emerges as a tool for human genetics. Nat Rev Genet 2005; 6: 9-23

Bitinaite J, Wah DA, Aggarwal AK and Schildkraut I. FokI dimerization is required for DNA cleavage. Proc Natl Acad Sci U S A 1998; 95: 10570-10575

Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A and Bonas U. Breaking the code of DNA binding specificity of TAL-type III effectors. Science 2009; 326: 1509-1512

Bogdanove AJ, Schornack S and Lahaye T. TAL effectors: finding plant genes for disease and defense. Curr Opin Plant Biol 2010; 13: 394-401

Borgmeyer U. Dimeric binding of the mouse germ cell nuclear factor. Eur J Biochem 1997; 244: 120-127

Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, Dickman MJ, Makarova KS, Koonin EV and van der Oost J. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 2008; 321: 960-964

Cambier L, Plate M, Sucov HM and Pashmforoush M. Nkx2-5 regulates cardiac growth through modulation of Wnt signaling by R-spondin3. Development 2014; 141: 2959-2971

Cao F, Xie X, Gollan T, Zhao L, Narsinh K, Lee RJ and Wu JC. Comparison of gene-transfer efficiency in human embryonic stem cells. Mol Imaging Biol 2010; 12: 15-24

Capecchi MR. Altering the genome by homologous recombination. Science 1989; 244: 1288-1292

Capecchi MR. Generating mice with targeted mutations. Nat Med 2001; 7: 1086-1090

Carbery ID, Ji D, Harrington A, Brown V, Weinstein EJ, Liaw L and Cui X. Targeted genome modification in mice using zinc-finger nucleases. Genetics 2010; 186: 451-459

Chambers SM, Fasano CA, Papapetrou EP, Tomishima M, Sadelain M and Studer L. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat Biotechnol 2009; 27: 275-280

Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, E PRI, Lin S, Kiani S, Guzman CD, Wiegand DJ, Ter-Ovanesyan D, Braff JL, Davidsohn N, Housden BE, Perrimon N, Weiss R, Aach J, Collins JJ and Church GM. Highly efficient Cas9-mediated transcriptional programming. Nat Methods 2015; 12: 326-328

Chen F, Cooney AJ, Wang Y, Law SW and O'Malley BW. Cloning of a novel orphan receptor (GCNF) expressed during germ cell development. Mol Endocrinol 1994; 8: 1434-1444

Chen F, Pruett-Miller SM, Huang Y, Gjoka M, Duda K, Taunton J, Collingwood TN, Frodin M and Davis GD. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. Nat Methods 2011; 8: 753-755

Chen S, Lee B, Lee AY, Modzelewski AJ and He L. Highly Efficient Mouse Genome Editing by CRISPR Ribonucleoprotein Electroporation of Zygotes. J Biol Chem 2016; 291: 14457-14467

Chen Y, Cao J, Xiong M, Petersen AJ, Dong Y, Tao Y, Huang CT, Du Z and Zhang SC. Engineering Human Stem Cell Lines with Inducible Gene Knockout using CRISPR/Cas9. Cell Stem Cell 2015; 17: 233-244

Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ and Voytas DF. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 2010; 186: 757-761

Chung AC, Katz D, Pereira FA, Jackson KJ, DeMayo FJ, Cooney AJ and O'Malley BW. Loss of orphan receptor germ cell nuclear factor function results in ectopic development of the tail bud and a novel posterior truncation. Mol Cell Biol 2001; 21: 663-677

Chung AC, Xu X, Niederreither KA and Cooney AJ. Loss of orphan nuclear receptor GCNF function disrupts forebrain development and the establishment of the isthmic organizer. Dev Biol 2006; 293: 13-24

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA and Zhang F. Multiplex genome engineering using CRISPR/Cas systems. Science 2013; 339: 819-823

Conti L and Cattaneo E. Neural stem cell systems: physiological players or in vitro entities? Nat Rev Neurosci 2010; 11: 176-187

Cooney AJ, Hummelke GC, Herman T, Chen F and Jackson KJ. Germ cell nuclear factor is a response element-specific repressor of transcription. Biochem Biophys Res Commun 1998; 245: 94-100

Creighton HB and McClintock B. A Correlation of Cytological and Genetical Crossing-Over in Zea Mays. Proc Natl Acad Sci U S A 1931; 17: 492-497

Cullot G, Boutin J, Toutain J, Prat F, Pennamen P, Rooryck C, Teichmann M, Rousseau E, Lamrissi-Garcia I, Guyonnet-Duperat V, Bibeyran A, Lalanne M, Prouzet-Mauleon V, Turcq B, Ged C, Blouin JM, Richard E, Dabernat S, Moreau-Gaudry F and Bedel A. CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. Nat Commun 2019; 10: 1136

David R, Joos TO and Dreyer C. Anteroposterior patterning and organogenesis of Xenopus laevis require a correct dose of germ cell nuclear factor (xGCNF). Mechanisms of Development 1998; 79: 137-152

de Lau WB, Snel B and Clevers HC. The R-spondin protein family. Genome Biol 2012; 13: 242

De Los Angeles A, Ferrari F, Xi R, Fujiwara Y, Benvenisty N, Deng H, Hochedlinger K, Jaenisch R, Lee S, Leitch HG, Lensch MW, Lujan E, Pei D, Rossant J, Wernig M, Park PJ and Daley GQ. Hallmarks of pluripotency. Nature 2015; 525: 469-478

Deinsberger J, Reisinger D and Weber B. Global trends in clinical trials involving pluripotent stem cells: a systematic multi-database analysis. NPJ Regen Med 2020; 5: 15

Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J and Charpentier E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 2011; 471: 602-607

Dhaka A, Earley TJ, Watson J and Patapoutian A. Visualizing cold spots: TRPM8-expressing sensory neurons and their projections. J Neurosci 2008; 28: 566-575

Eckler MJ, Larkin KA, McKenna WL, Katzman S, Guo C, Roque R, Visel A, Rubenstein JL and Chen B. Multiple conserved regulatory domains promote Fezf2 expression in the developing cerebral cortex. Neural Dev 2014; 9: 6

Elkabetz Y, Panagiotakos G, Al Shamy G, Socci ND, Tabar V and Studer L. Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. Genes Dev 2008; 22: 152-165

Evans MJ and Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature 1981; 292: 154-156

Fiorenzano A, Pascale E, D'Aniello C, Acampora D, Bassalert C, Russo F, Andolfi G, Biffoni M, Francescangeli F, Zeuner A, Angelini C, Chazaud C, Patriarca EJ, Fico A and Minchiotti G. Cripto is essential to capture mouse epiblast stem cell and human embryonic stem cell pluripotency. Nat Commun 2016; 7: 12589

Folger KR, Wong EA, Wahl G and Capecchi MR. Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. Mol Cell Biol 1982; 2: 1372-1387

Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, Modi BP, Correard S, Gheorghe M, Baranasic D, Santana-Garcia W, Tan G, Cheneby J, Ballester B, Parcy F, Sandelin A, Lenhard B, Wasserman WW and Mathelier A. JASPAR 2020: update of the open-access database of transcription factor binding profiles. Nucleic Acids Res 2020; 48: D87-D92

Fu Y, Sander JD, Reyon D, Cascio VM and Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol 2014; 32: 279-284

Fuerer C and Nusse R. Lentiviral vectors to probe and manipulate the Wnt signaling pathway. PLoS One 2010; 5: e9370

Fuhrmann G, Chung ACK, Jackson KJ, Hummelke G, Baniahmad A, Sutter J, Sylvester I, Schöler HR and Cooney AJ. Mouse Germline Restriction of Oct4 Expression by Germ Cell Nuclear Factor. Developmental Cell 2001; 1: 377-387

Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS and Qi LS. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 2013; 154: 442-451

Glass CK and Ogawa S. Combinatorial roles of nuclear receptors in inflammation and immunity. Nat Rev Immunol 2006; 6: 44-55

Götz M, Stoykova A and Gruss P. Pax6 Controls Radial Glia Differentiation in the Cerebral Cortex. Neuron 1998; 21: 1031-1044

Green SA, Simoes-Costa M and Bronner ME. Evolution of vertebrates as viewed from the crest. Nature 2015; 520: 474-482

Greschik H and Schule R. Germ cell nuclear factor: an orphan receptor with unexpected properties. J Mol Med (Berl) 1998; 76: 800-810

Greschik H, Wurtz JM, Hublitz P, Kohler F, Moras D and Schule R. Characterization of the DNA-binding and dimerization properties of the nuclear orphan receptor germ cell nuclear factor. Mol Cell Biol 1999; 19: 690-703

Groenen PM, Bunschoten AE, van Soolingen D and van Embden JD. Nature of DNA polymorphism in the direct repeat cluster of Mycobacterium tuberculosis; application for strain differentiation by a novel typing method. Mol Microbiol 1993; 10: 1057-1065

Gu P, LeMenuet D, Chung ACK, Mancini M, Wheeler DA and Cooney AJ. Orphan Nuclear Receptor GCNF Is Required for the Repression of Pluripotency Genes during Retinoic Acid-Induced Embryonic Stem Cell Differentiation. Molecular and Cellular Biology 2005; 25: 8507-8519

Gu P, Xu X, Le Menuet D, Chung AC and Cooney AJ. Differential recruitment of methyl CpG-binding domain factors and DNA methyltransferases by the orphan receptor germ cell nuclear factor initiates the repression and silencing of Oct4. Stem Cells 2011; 29: 1041-1051

Han X, Chen H, Huang D, Chen H, Fei L, Cheng C, Huang H, Yuan GC and Guo G. Mapping human pluripotent stem cell differentiation pathways using high throughput single-cell RNA-sequencing. Genome Biol 2018; 19: 47

Hashimoto M and Takemoto T. Electroporation enables the efficient mRNA delivery into the mouse zygotes and facilitates CRISPR/Cas9-based genome editing. Sci Rep 2015; 5: 11315

Heim KC, White KA, Deng D, Tomlinson CR, Moore JH, Freemantle SJ and Spinella MJ. Selective repression of retinoic acid target genes by RIP140 during induced tumor cell differentiation of pluripotent human embryonal carcinoma cells. Mol Cancer 2007; 6: 57

Hentschke M, Kurth I, Borgmeyer U and Hubner CA. Germ cell nuclear factor is a repressor of CRIPTO-1 and CRIPTO-3. J Biol Chem 2006; 281: 33497-33504

Hermans PW, van Soolingen D, Bik EM, de Haas PE, Dale JW and van Embden JD. Insertion element IS987 from Mycobacterium bovis BCG is located in a hot-spot integration region for insertion elements in Mycobacterium tuberculosis complex strains. Infect Immun 1991; 59: 2695-2705

Hirose T, O'Brien DA and Jetten AM. RTR: a new member of the nuclear receptor superfamily that is highly expressed in murine testis. Gene 1995; 152: 247-251

Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, Zeitler B, Cherone JM, Meng X, Hinkley SJ, Rebar EJ, Gregory PD, Urnov FD and Jaenisch R. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 2011; 29: 731-734

Hou Z, Zhang Y, Propson NE, Howden SE, Chu LF, Sontheimer EJ and Thomson JA. Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proc Natl Acad Sci U S A 2013; 110: 15644-15649

Hu X, Chen Y, Farooqui M, Thomas MC, Chiang CM and Wei LN. Suppressive effect of receptor-interacting protein 140 on coregulator binding to retinoic acid receptor complexes, histone-modifying enzyme activity, and gene activation. J Biol Chem 2004; 279: 319-325

Huang P, Chandra V and Rastinejad F. Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics. Annu Rev Physiol 2010; 72: 247-272

Hummelke GC and Cooney AJ. Reciprocal regulation of the mouse protamine genes by the orphan nuclear receptor germ cell nuclear factor and CREMtau. Mol Reprod Dev 2004; 68: 394-407

Hunter N. Meiotic Recombination: The Essence of Heredity. Cold Spring Harb Perspect Biol 2015; 7

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

Jansen R, Embden JD, Gaastra W and Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. Mol Microbiol 2002; 43: 1565-1575

Jasin M. Genetic manipulation of genomes with rare-cutting endonucleases. Trends in Genetics 1996; 12: 224-228

Jasin M and Rothstein R. Repair of strand breaks by homologous recombination. Cold Spring Harb Perspect Biol 2013; 5: a012740

Jeong Y and Mangelsdorf DJ. Nuclear receptor regulation of stemness and stem cell differentiation. Exp Mol Med 2009; 41: 525-537

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA and Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 2012; 337: 816-821

Jinek M, East A, Cheng A, Lin S, Ma E and Doudna J. RNA-programmed genome editing in human cells. eLife 2013; 2

Jordan BK, Shen JH, Olaso R, Ingraham HA and Vilain E. Wnt4 overexpression disrupts normal testicular vasculature and inhibits testosterone synthesis by repressing steroidogenic factor 1/beta-catenin synergy. Proc Natl Acad Sci U S A 2003; 100: 10866-10871

Jukam D, Shariati SAM and Skotheim JM. Zygotic Genome Activation in Vertebrates. Dev Cell 2017; 42: 316-332

Kapelle M, Krätzschmar J, Husemann M and Schleuning W-D. cDNA cloning of two closely related forms of human germ cell nuclear factor (GCNF). Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression 1997; 1352: 13-17

Katz D, Niederberger C, Slaughter GR and Cooney AJ. Characterization of germ cell-specific expression of the orphan nuclear receptor, germ cell nuclear factor. Endocrinology 1997; 138: 4364-4372

Kazanskaya O, Ohkawara B, Heroult M, Wu W, Maltry N, Augustin HG and Niehrs C. The Wnt signaling regulator R-spondin 3 promotes angioblast and vascular development. Development 2008; 135: 3655-3664

Keller GM. In vitro differentiation of embryonic stem cells. Current Opinion in Cell Biology 1995; 7: 862-869

Kerley JS, Olsen SL, Freemantle SJ and Spinella MJ. Transcriptional Activation of the Nuclear Receptor Corepressor RIP140 by Retinoic Acid: A Potential Negative-Feedback Regulatory Mechanism. Biochemical and Biophysical Research Communications 2001; 285: 969-975

Kikuchi T, Morizane A, Doi D, Magotani H, Onoe H, Hayashi T, Mizuma H, Takara S, Takahashi R, Inoue H, Morita S, Yamamoto M, Okita K, Nakagawa M, Parmar M and Takahashi J. Human iPS cell-derived dopaminergic neurons function in a primate Parkinson's disease model. Nature 2017; 548: 592-596

Kim H and Kim JS. A guide to genome engineering with programmable nucleases. Nat Rev Genet 2014; 15: 321-334

Kim YG, Cha J and Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A 1996; 93: 1156-1160

Koch P, Opitz T, Steinbeck JA, Ladewig J and Brustle O. A rosette-type, self-renewing human ES cell-derived neural stem cell with potential for in vitro instruction and synaptic integration. Proc Natl Acad Sci U S A 2009; 106: 3225-3230

Kriks S, Shim JW, Piao J, Ganat YM, Wakeman DR, Xie Z, Carrillo-Reid L, Auyeung G, Antonacci C, Buch A, Yang L, Beal MF, Surmeier DJ, Kordower JH, Tabar V and Studer L. Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. Nature 2011; 480: 547-551

Kuehn MR, Bradley A, Robertson EJ and Evans MJ. A potential animal model for Lesch-Nyhan syndrome through introduction of HPRT mutations into mice. Nature 1987; 326: 295-298

Kumar R and Thompson EB. The structure of the nuclear hormone receptors. Steroids 1999; 64: 310-319

Labun K, Montague TG, Gagnon JA, Thyme SB and Valen E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. Nucleic Acids Res 2016; 44: W272-276

Lan Z-J, Xu X, Chung ACK and Cooney AJ. Extra-Germ Cell Expression of Mouse Nuclear Receptor Subfamily 6, Group A, Member 1 (NR6A1)1. Biology of Reproduction 2009; 80: 905-912

Lan ZJ, Chung AC, Xu X, DeMayo FJ and Cooney AJ. The embryonic function of germ cell nuclear factor is dependent on the DNA binding domain. J Biol Chem 2002; 277: 50660-50667

Li L, Wu LP and Chandrasegaran S. Functional domains in Fok I restriction endonuclease. Proc Natl Acad Sci U S A 1992; 89: 4275-4279

Li T, Huang S, Zhao X, Wright DA, Carpenter S, Spalding MH, Weeks DP and Yang B. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Research 2011; 39: 6315-6325

Liu A and Niswander LA. Bone morphogenetic protein signalling and vertebrate nervous system development. Nat Rev Neurosci 2005; 6: 945-954

Liu G, Zhang Y and Zhang T. Computational approaches for effective CRISPR guide RNA design and evaluation. Comput Struct Biotechnol J 2020; 18: 35-44

Liu YW, Chen B, Yang X, Fugate JA, Kalucki FA, Futakuchi-Tsuchida A, Couture L, Vogel KW, Astley CA, Baldessari A, Ogle J, Don CW, Steinberg ZL, Seslar SP, Tuck SA, Tsuchida H, Naumova AV, Dupras SK, Lyu MS, Lee J, Hailey DW, Reinecke H, Pabon L, Fryer BH, MacLellan WR, Thies RS and Murry CE. Human embryonic stem cell-derived cardiomyocytes restore function in infarcted hearts of non-human primates. Nat Biotechnol 2018; 36: 597-605

Livak KJ and Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001; 25: 402-408

Love MI, Huber W and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 2014; 15

Lyuksyutova AI, Lu CC, Milanesio N, King LA, Guo N, Wang Y, Nathans J, Tessier-Lavigne M and Zou Y. Anterior-posterior guidance of commissural axons by Wnt-frizzled signaling. Science 2003; 302: 1984-1988

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE and Church GM. RNA-guided human genome engineering via Cas9. Science 2013; 339: 823-826

Mandai M, Watanabe A, Kurimoto Y, Hirami Y, Morinaga C, Daimon T, Fujihara M, Akimaru H, Sakai N, Shibata Y, Terada M, Nomiya Y, Tanishima S, Nakamura M, Kamao H, Sugita S, Onishi A, Ito T, Fujita K, Kawamata S, Go MJ, Shinohara C, Hata KI, Sawada M, Yamamoto M, Ohta S, Ohara Y, Yoshida K, Kuwahara J, Kitano Y, Amano N, Umekage M, Kitaoka F, Tanaka A, Okada C, Takasu N, Ogawa S, Yamanaka S and Takahashi M. Autologous Induced Stem-Cell-Derived Retinal Cells for Macular Degeneration. N Engl J Med 2017; 376: 1038-1046

Mansour SL, Thomas KR and Capecchi MR. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature 1988; 336: 348-352

Mao Z, Bozzella M, Seluanov A and Gorbunova V. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. Cell Cycle 2008; 7: 2902-2906

Marraffini LA and Sontheimer EJ. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science 2008; 322: 1843-1845

McFadden DG, Barbosa AC, Richardson JA, Schneider MD, Srivastava D and Olson EN. The Hand1 and Hand2 transcription factors regulate expansion of the embryonic cardiac ventricles in a gene dosage-dependent manner. Development 2005; 132: 189-201

Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD and Rebar EJ. A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 2011; 29: 143-148

Mojica FJ, Diez-Villasenor C, Garcia-Martinez J and Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol 2005; 60: 174-182

Mojica FJ, Diez-Villasenor C, Soria E and Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. Mol Microbiol 2000; 36: 244-246

Mojica FJ, Ferrer C, Juez G and Rodriguez-Valera F. Long stretches of short tandem repeats are present in the largest replicons of the Archaea Haloferax mediterranei and Haloferax volcanii and could be involved in replicon partitioning. Mol Microbiol 1995; 17: 85-93

Mojica FJ, Juez G and Rodriguez-Valera F. Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites. Mol Microbiol 1993; 9: 613-621

Mojica FJM, Diez-Villasenor C, Garcia-Martinez J and Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology (Reading) 2009; 155: 733-740

Morasso MI, Grinberg A, Robinson G, Sargent TD and Mahon KA. Placental failure in mice lacking the homeobox gene Dlx3. Proc Natl Acad Sci U S A 1999; 96: 162-167

Moscou MJ and Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science 2009; 326: 1501

Mullen EM, Gu P and Cooney AJ. Nuclear Receptors in Regulation of Mouse ES Cell Pluripotency and Differentiation. PPAR Res 2007; 2007: 61563

Mulligan KA and Cheyette BN. Wnt signaling in vertebrate neural development and function. J Neuroimmune Pharmacol 2012; 7: 774-787

Murry CE and Keller G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. Cell 2008; 132: 661-680

Nakata A, Amemura M and Makino K. Unusual nucleotide arrangement with repeated sequences in the Escherichia coli K-12 chromosome. J Bacteriol 1989; 171: 3553-3556

Nandi AK, Roginski RS, Gregg RG, Smithies O and Skoultchi AI. Regulated expression of genes inserted at the human chromosomal beta-globin locus by homologous recombination. Proc Natl Acad Sci U S A 1988; 85: 3845-3849

Niehrs C. The complex world of WNT receptor signalling. Nat Rev Mol Cell Biol 2012; 13: 767-779

Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F and Nureki O. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell 2014; 156: 935-949

Nuclear Receptors Nomenclature Committee. A Unified Nomenclature System for the Nuclear Receptor Superfamily. Cell 1999; 97: 161-163

Nusslein-Volhard C and Wieschaus E. Mutations affecting segment number and polarity in Drosophila. Nature 1980; 287: 795-801

Okumura LM, Lesch BJ and Page DC. The Ligand Binding Domain of GCNF Is Not Required for Repression of Pluripotency Genes in Mouse Fetal Ovarian Germ Cells. PLoS ONE 2013; 8

Overington JP, Al-Lazikani B and Hopkins AL. How many drug targets are there? Nat Rev Drug Discov 2006; 5: 993-996

Parisi S, D'Andrea D, Lago CT, Adamson ED, Persico MG and Minchiotti G. Nodal-dependent Cripto signaling promotes cardiomyogenesis and redirects the neural fate of embryonic stem cells. J Cell Biol 2003; 163: 303-314

Park PJ. ChIP-seq: advantages and challenges of a maturing technology. Nat Rev Genet 2009; 10: 669-680

Parker M, Leonardsson G, White R, Steel J and Milligan S. Identification of RIP140 as a nuclear receptor cofactor with a role in female reproduction. FEBS Letters 2003; 546: 149-153

Patro R, Duggal G, Love MI, Irizarry RA and Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 2017; 14: 417-419

Pavletich NP and Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science 1991; 252: 809-817

Pedersen RA, Wu K and BaŁakier H. Origin of the inner cell mass in mouse embryos: Cell lineage analysis by microinjection. Developmental Biology 1986; 117: 581-595

Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, Wang N, Lee G, Bartsevich VV, Lee YL, Guschin DY, Rupniewski I, Waite AJ, Carpenito C, Carroll RG, Orange JS, Urnov FD, Rebar EJ, Ando D, Gregory PD, Riley JL, Holmes MC and June CH. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol 2008; 26: 808-816

Pevny LH, Sockanathan S, Placzek M and Lovell-Badge R. A role for SOX1 in neural determination. Development 1998; 125: 1967-1978

Ramirez CL, Foley JE, Wright DA, Muller-Lerch F, Rahman SH, Cornu TI, Winfrey RJ, Sander JD, Fu F, Townsend JA, Cathomen T, Voytas DF and Joung JK. Unexpected failure rates for modular assembly of engineered zinc fingers. Nat Methods 2008; 5: 374-375

Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS, Koonin EV, Sharp PA and Zhang F. In vivo genome editing using Staphylococcus aureus Cas9. Nature 2015; 520: 186-191

Ran FA, Hsu Patrick D, Lin C-Y, Gootenberg Jonathan S, Konermann S, Trevino AE, Scott David A, Inoue A, Matoba S, Zhang Y and Zhang F. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Cell 2013a; 154: 1380-1389

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA and Zhang F. Genome engineering using the CRISPR-Cas9 system. Nature Protocols 2013b; 8: 2281-2308

Rayner E, Durin MA, Thomas R, Moralli D, O'Cathail SM, Tomlinson I, Green CM and Lewis A. CRISPR-Cas9 Causes Chromosomal Instability and Rearrangements in Cancer Cell Lines, Detectable by Cytogenetic Methods. CRISPR J 2019; 2: 406-416

Reinhardt P, Glatza M, Hemmer K, Tsytsyura Y, Thiel CS, Hoing S, Moritz S, Parga JA, Wagner L, Bruder JM, Wu G, Schmid B, Ropke A, Klingauf J, Schwamborn JC, Gasser T, Scholer HR and Sterneckert J. Derivation and expansion using only small molecules of human neural progenitors for neurodegenerative disease modeling. PLoS One 2013; 8: e59252

Renaud JB, Boix C, Charpentier M, De Cian A, Cochennec J, Duvernois-Berthet E, Perrouault L, Tesson L, Edouard J, Thinard R, Cherifi Y, Menoret S, Fontaniere S, de Croze N, Fraichard A, Sohm F, Anegon I, Concordet JP and Giovannangeli C. Improved

Genome Editing Efficiency and Flexibility Using Modified Oligonucleotides with TALEN and CRISPR-Cas9 Nucleases. Cell Rep 2016; 14: 2263-2272

Rigau M, Juan D, Valencia A and Rico D. Intronic CNVs and gene expression variation in human populations. PLoS Genet 2019; 15: e1007902

Roode M, Blair K, Snell P, Elder K, Marchant S, Smith A and Nichols J. Human hypoblast formation is not dependent on FGF signalling. Dev Biol 2012; 361: 358-363

Rossant J and Tam PPL. New Insights into Early Human Development: Lessons for Stem Cell Derivation and Differentiation. Cell Stem Cell 2017; 20: 18-28

Rouet P, Smih F and Jasin M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc Natl Acad Sci U S A 1994a; 91: 6064-6068

Rouet P, Smih F and Jasin M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol Cell Biol 1994b; 14: 8096-8106

Sabour D, Xu X, Chung AC, Le Menuet D, Ko K, Tapia N, Arauzo-Bravo MJ, Gentile L, Greber B, Hubner K, Sebastiano V, Wu G, Scholer HR and Cooney AJ. Germ cell nuclear factor regulates gametogenesis in developing gonads. PLoS One 2014; 9: e103985

Sargent RG, Brenneman MA and Wilson JH. Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination. Mol Cell Biol 1997; 17: 267-277

Schneider CA, Rasband WS and Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 2012; 9: 671-675

Schöler HR, Balling R, Hatzopoulos AK, Suzuki N and Gruss P. Octamer binding proteins confer transcriptional activity in early mouse embryogenesis. The EMBO Journal 1989a; 8: 2551-2557

Schöler HR, Hatzopoulos AK, Balling R, Suzuki N and Gruss P. A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor. The EMBO Journal 1989b; 8: 2543-2550

Schrode N, Ho SM, Yamamuro K, Dobbyn A, Huckins L, Matos MR, Cheng E, Deans PJM, Flaherty E, Barretto N, Topol A, Alganem K, Abadali S, Gregory J, Hoelzli E, Phatnani H, Singh V, Girish D, Aronow B, McCullumsmith R, Hoffman GE, Stahl EA, Morishita H, Sklar P and Brennand KJ. Synergistic effects of common schizophrenia risk variants. Nat Genet 2019; 51: 1475-1485

Seki T, Yuasa S, Oda M, Egashira T, Yae K, Kusumoto D, Nakata H, Tohyama S, Hashimoto H, Kodaira M, Okada Y, Seimiya H, Fusaki N, Hasegawa M and Fukuda K. Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. Cell Stem Cell 2010; 7: 11-14

Shahbazi MN and Zernicka-Goetz M. Deconstructing and reconstructing the mouse and human early embryo. Nat Cell Biol 2018; 20: 878-887

Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, Heckl D, Ebert BL, Root DE, Doench JG and Zhang F. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 2014; 343: 84-87

Shi Y, Kirwan P, Smith J, Robinson HP and Livesey FJ. Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. Nat Neurosci 2012; 15: 477-486, S471

Shimizu T, Nakazawa M, Kani S, Bae YK, Shimizu T, Kageyama R and Hibi M. Zinc finger genes Fezf1 and Fezf2 control neuronal differentiation by repressing Hes5 expression in the forebrain. Development 2010; 137: 1875-1885

Smith C, Gore A, Yan W, Abalde-Atristain L, Li Z, He C, Wang Y, Brodsky Robert A, Zhang K, Cheng L and Ye Z. Whole-Genome Sequencing Analysis Reveals High Specificity of CRISPR/Cas9 and TALEN-Based Genome Editing in Human iPSCs. Cell Stem Cell 2014; 15: 12-13

Smithies O, Gregg RG, Boggs SS, Koralewski MA and Kucherlapati RS. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. Nature 1985; 317: 230-234

Stiles J and Jernigan TL. The basics of brain development. Neuropsychol Rev 2010; 20: 327-348

Sudiwala S, Palmer A, Massa V, Burns AJ, Dunlevy LPE, de Castro SCP, Savery D, Leung KY, Copp AJ and Greene NDE. Cellular mechanisms underlying Pax3-related neural tube defects and their prevention by folic acid. Dis Model Mech 2019; 12

Süsens U and Borgmeyer U. Characterization of the human germ cell nuclear factor gene. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression 1996; 1309: 179-182

Tabar V and Studer L. Pluripotent stem cells in regenerative medicine: challenges and recent progress. Nat Rev Genet 2014; 15: 82-92

Tailor J, Kittappa R, Leto K, Gates M, Borel M, Paulsen O, Spitzer S, Karadottir RT, Rossi F, Falk A and Smith A. Stem cells expanded from the human embryonic hindbrain stably retain regional specification and high neurogenic potency. J Neurosci 2013; 33: 12407-12422

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K and Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007; 131: 861-872

Takahashi K and Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006; 126: 663-676

Tata JR. Signalling through nuclear receptors. Nat Rev Mol Cell Biol 2002; 3: 702-710

Thakurela S, Tiwari N, Schick S, Garding A, Ivanek R, Berninger B and Tiwari VK. Mapping gene regulatory circuitry of Pax6 during neurogenesis. Cell Discov 2016; 2: 15045

Thomas KR and Capecchi MR. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell 1987; 51: 503-512

Thomas KR and Capecchi MR. Targeted disruption of the murine int-1 proto-oncogene resulting in severe abnormalities in midbrain and cerebellar development. Nature 1990; 346: 847-850

Thompson S, Clarke AR, Pow AM, Hooper ML and Melton DW. Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. Cell 1989; 56: 313-321

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS and Jones JM. Embryonic stem cell lines derived from human blastocysts. Science 1998; 282: 1145-1147

Thomson M, Liu SJ, Zou LN, Smith Z, Meissner A and Ramanathan S. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. Cell 2011; 145: 875-889

Tosic J, Kim GJ, Pavlovic M, Schroder CM, Mersiowsky SL, Barg M, Hofherr A, Probst S, Kottgen M, Hein L and Arnold SJ. Eomes and Brachyury control pluripotency exit and germ-layer segregation by changing the chromatin state. Nat Cell Biol 2019; 21: 1518-1531

Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD and Holmes MC. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 2005; 435: 646-651

Urnov FD, Rebar EJ, Holmes MC, Zhang HS and Gregory PD. Genome editing with engineered zinc finger nucleases. Nat Rev Genet 2010; 11: 636-646

Vakulskas CA, Dever DP, Rettig GR, Turk R, Jacobi AM, Collingwood MA, Bode NM, McNeill MS, Yan S, Camarena J, Lee CM, Park SH, Wiebking V, Bak RO, Gomez-Ospina N, Pavel-Dinu M, Sun W, Bao G, Porteus MH and Behlke MA. A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. Nat Med 2018; 24: 1216-1224

Veres A, Gosis BS, Ding Q, Collins R, Ragavendran A, Brand H, Erdin S, Cowan CA, Talkowski ME and Musunuru K. Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. Cell Stem Cell 2014; 15: 27-30

Wang H, Wang X, Xu X, Kyba M and Cooney AJ. Germ Cell Nuclear Factor (GCNF) Represses Oct4 Expression and Globally Modulates Gene Expression in Human Embryonic Stem (hES) Cells. J Biol Chem 2016; 291: 8644-8652

Wang H, Wang X, Xu X, Zwaka TP and Cooney AJ. Epigenetic reprogramming of the germ cell nuclear factor gene is required for proper differentiation of induced pluripotent cells. Stem Cells 2013; 31: 2659-2666

Wang Q and Cooney AJ. Revisiting the role of GCNF in embryonic development. Semin Cell Dev Biol 2013; 24: 679-686

Wang Q, Fang W-H, Krupinski J, Kumar S, Slevin M and Kumar P. Paxgenes in embryogenesis and oncogenesis. Journal of Cellular and Molecular Medicine 2008; 12: 2281-2294

Wang T, Wei JJ, Sabatini DM and Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. Science 2014; 343: 80-84

Wang ZB, Boisvert E, Zhang X, Guo M, Fashoyin A, Du ZW, Zhang SC and Li XJ. Fezf2 regulates telencephalic precursor differentiation from mouse embryonic stem cells. Cereb Cortex 2011; 21: 2177-2186

Watase K and Zoghbi HY. Modelling brain diseases in mice: the challenges of design and analysis. Nat Rev Genet 2003; 4: 296-307

Watson JD and Crick FH. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. J.D. Watson and F.H.C. Crick. Published in Nature, number 4356 April 25, 1953. Nature 1974; 248: 765

Weikum ER, Tuntland ML, Murphy MN and Ortlund EA. A Structural Investigation into Oct4 Regulation by Orphan Nuclear Receptors, Germ Cell Nuclear Factor (GCNF), and Liver Receptor Homolog-1 (LRH-1). J Mol Biol 2016; 428: 4981-4992

Wright AV, Nunez JK and Doudna JA. Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. Cell 2016; 164: 29-44

Xu C, Liguori G, Adamson ED and Persico MG. Specific arrest of cardiogenesis in cultured embryonic stem cells lacking Cripto-1. Dev Biol 1998; 196: 237-247

Yan Z and Jetten AM. Characterization of the repressor function of the nuclear orphan receptor retinoid receptor-related testis-associated receptor/germ cell nuclear factor. J Biol Chem 2000; 275: 35077-35085

Yan ZH, Medvedev A, Hirose T, Gotoh H and Jetten AM. Characterization of the response element and DNA binding properties of the nuclear orphan receptor germ cell nuclear factor/retinoid receptor-related testis-associated receptor. J Biol Chem 1997; 272: 10565-10572

Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, Bhai J, Billis K, Boddu S, Marugan JC, Cummins C, Davidson C, Dodiya K, Fatima R, Gall A, Giron CG, Gil L, Grego T, Haggerty L, Haskell E, Hourlier T, Izuogu OG, Janacek SH, Juettemann T, Kay M, Lavidas I, Le T, Lemos D, Martinez JG, Maurel T, McDowall M, McMahon A, Mohanan S, Moore B, Nuhn M, Oheh DN, Parker A, Parton A, Patricio M, Sakthivel MP, Abdul Salam AI, Schmitt BM, Schuilenburg H, Sheppard D, Sycheva M, Szuba M, Taylor K, Thormann A, Threadgold G, Vullo A, Walts B, Winterbottom A, Zadissa A, Chakiachvili M, Flint B, Frankish A, Hunt SE, G II, Kostadima M, Langridge N, Loveland JE, Martin FJ, Morales J, Mudge JM, Muffato M, Perry E, Ruffier M, Trevanion SJ, Cunningham F, Howe KL, Zerbino DR and Flicek P. Ensembl 2020. Nucleic Acids Res 2020; 48: D682-D688

Yeom YI, Fuhrmann G, Ovitt CE, Brehm A, Ohbo K, Gross M, Hübner K and Schöler HR. Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. . Development 1996; 122: 881-894

Yin H, Song CQ, Suresh S, Wu Q, Walsh S, Rhym LH, Mintzer E, Bolukbasi MF, Zhu LJ, Kauffman K, Mou H, Oberholzer A, Ding J, Kwan SY, Bogorad RL, Zatsepin T, Koteliansky V, Wolfe SA, Xue W, Langer R and Anderson DG. Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing. Nat Biotechnol 2017; 35: 1179-1187

Yoshimizu T, Sugiyama N, De Felice M, Yeom YI, Ohbo K, Masuko K, Obinata M, Abe K, Scholer HR and Matsui Y. Germline-specific expression of the Oct-4/green fluorescent protein (GFP) transgene in mice. Dev Growth Differ 1999; 41: 675-684

Yu G, Wang LG, Han Y and He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012; 16: 284-287

Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin, II and Thomson JA. Induced pluripotent stem cell lines derived from human somatic cells. Science 2007; 318: 1917-1920

Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV and Zhang F. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 2015; 163: 759-771

Zhang XH, Tee LY, Wang XG, Huang QS and Yang SH. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. Mol Ther Nucleic Acids 2015; 4: e264

Zhou HM and Conway SJ. Restricted Pax3 Deletion within the Neural Tube Results in Congenital Hydrocephalus. J Dev Biol 2016; 4

Ziller MJ, Edri R, Yaffe Y, Donaghey J, Pop R, Mallard W, Issner R, Gifford CA, Goren A, Xing J, Gu H, Cachiarelli D, Tsankov A, Epstein C, Rinn JR, Mikkelsen TS, Kohlbacher O, Gnirke A, Bernstein BE, Elkabetz Y and Meissner A. Dissecting neural differentiation regulatory networks through epigenetic footprinting. Nature 2015; 518: 355-359

9. Appendix

9.1 Supplementary figures
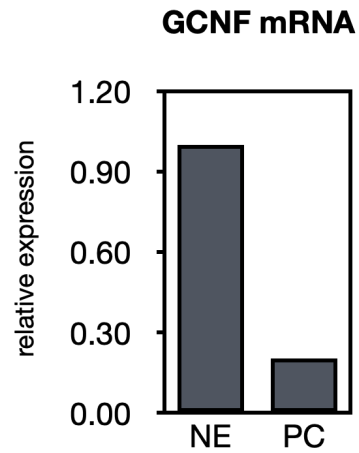
GCNF mRNA

Supplementary Fig. 1: QRT-PCR indicates a prominent reduction of GCNF mRNA in the polyclonal pool of nucleofected hiPSCs. NE indicates non-edited, parental hiPSCs. PC indicates edited polyclonal hiPSCs. n=1.
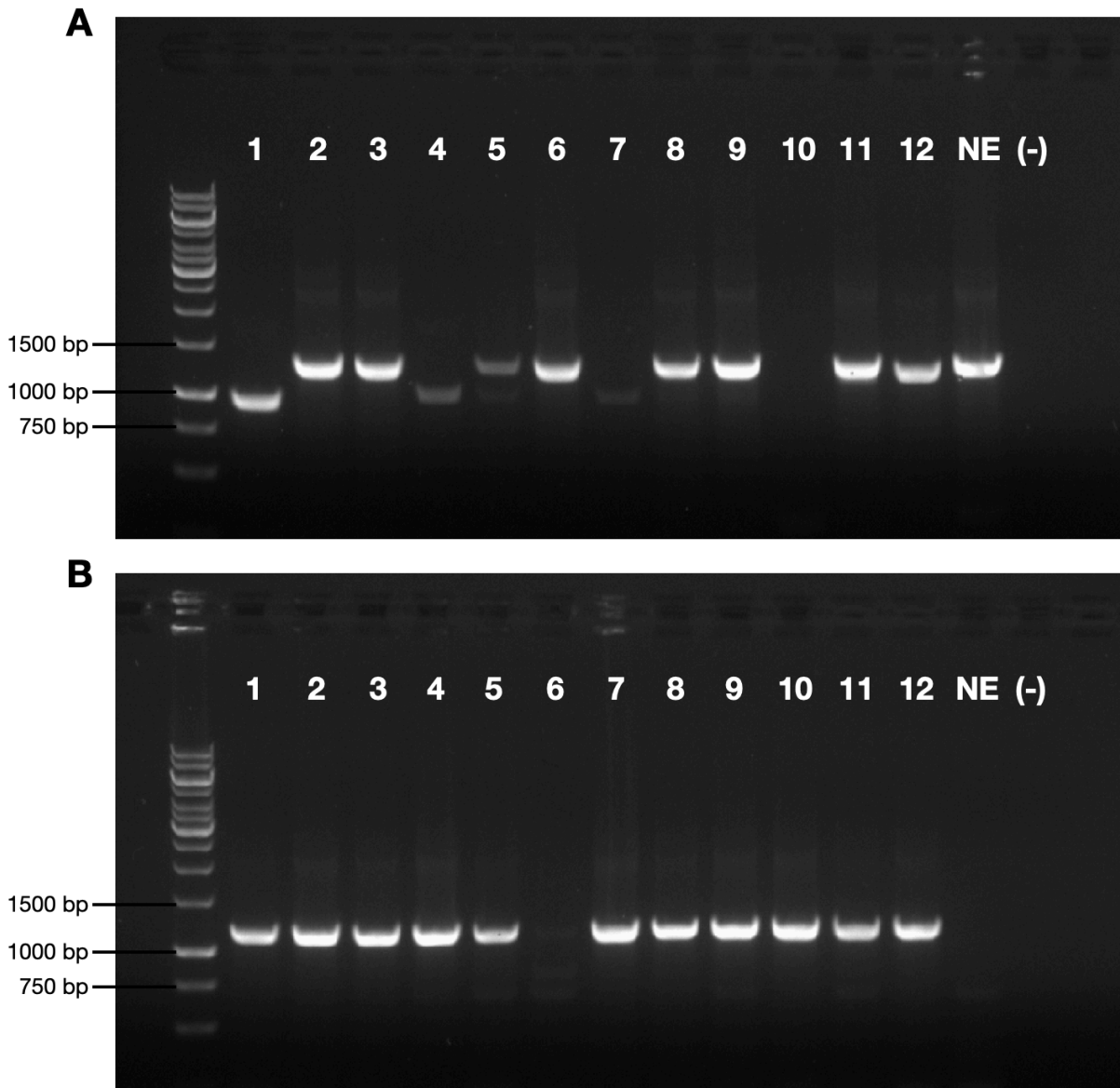
Supplementary Fig. 2: Full agarose gels of the genotyping PCR to confirm genomic alteration of the GCNF locus. **(A)** The first PCR yields a product of 1295 bp when the GCNF locus is intact. This PCR product is generated in the clones 2, 3, 5, 6, 8, 9, 11 and 12, as well as the non-edited control (NE). Notably, a smaller product is amplified in the GCNF knockout clones 1, 4 and 7. This lower product might be an unspecific amplification in the absence of the non-edited template. **(B)** The second PCR reaction yields a 1153 bp product when the puromycin cassette got inserted at the desired genomic location. This reaction yields a PCR product in all clones, except clone 6. The PCR product is also absent in the non-edited control (NE). Sequencing of that product confirmed the incorporation of the exogenous DNA (Fig. 13, Supplementary Fig. 3). On both agarose gels C1 – C12 is followed by the non-edited control (NE) and a water control (-).
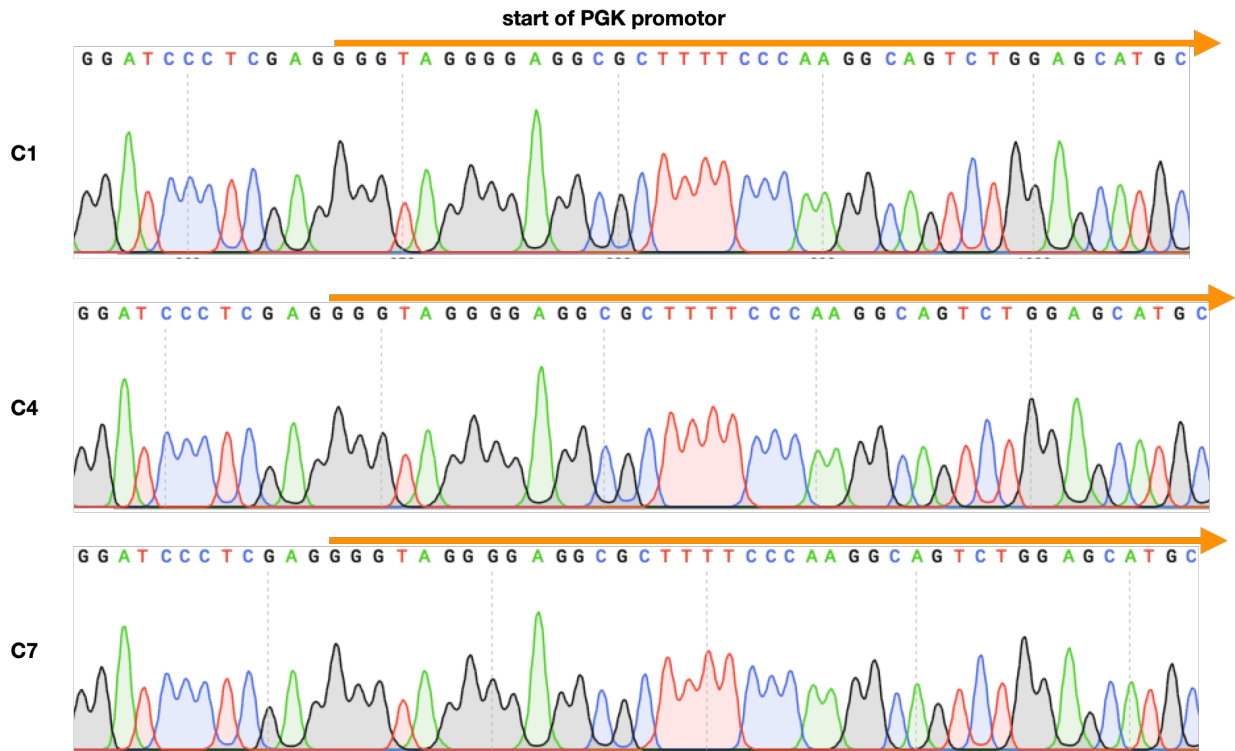
Supplementary Fig. 3: The insert-specific PGK promoter can be found in the Sanger sequencing reactions of the three selected GCNF-ablated clones. The sequence corresponds to the start of the PGK promoter (marked in orange).
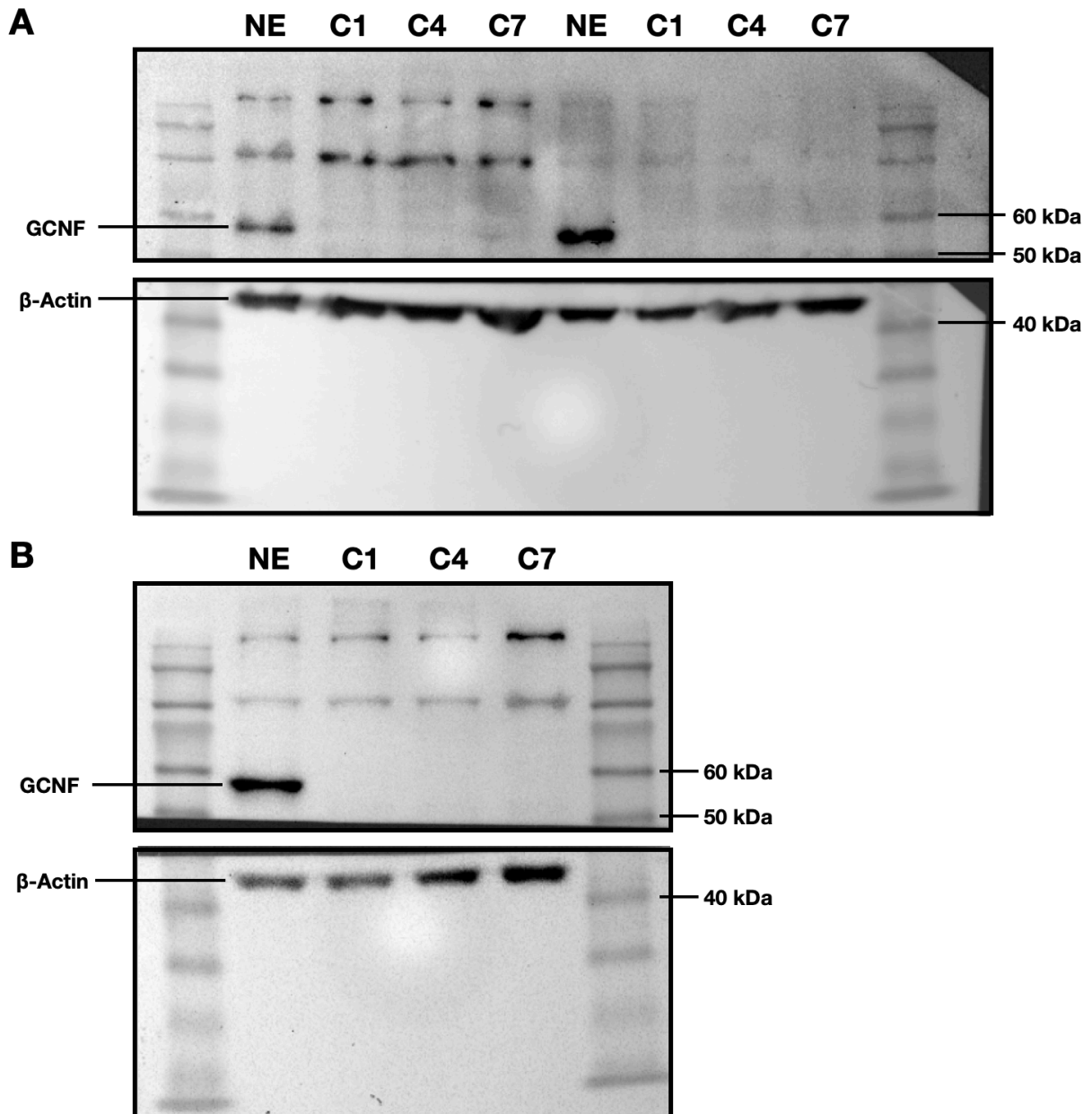
Supplementary Fig. 4: Three biological replicates confirm CRISPR/Cas9-mediated GCNF deficiency of the three selected GCNF-ablated clones in Western blot analysis. **(A)** Two biological replicates were loaded on one membrane. **(B)** The third biological replicate was loaded on a separate membrane. The samples were obtained from hiPSCs under self-renewing conditions. The absence of the GCNF band indicates successful GCNF disruption in all three, single cell-derived, clones (C1, C4, C7), as compared to the non-edited, parental cell line (NE). The membrane was cut in two pieces, the upper part was stained for GCNF, the lower part for β-Actin.
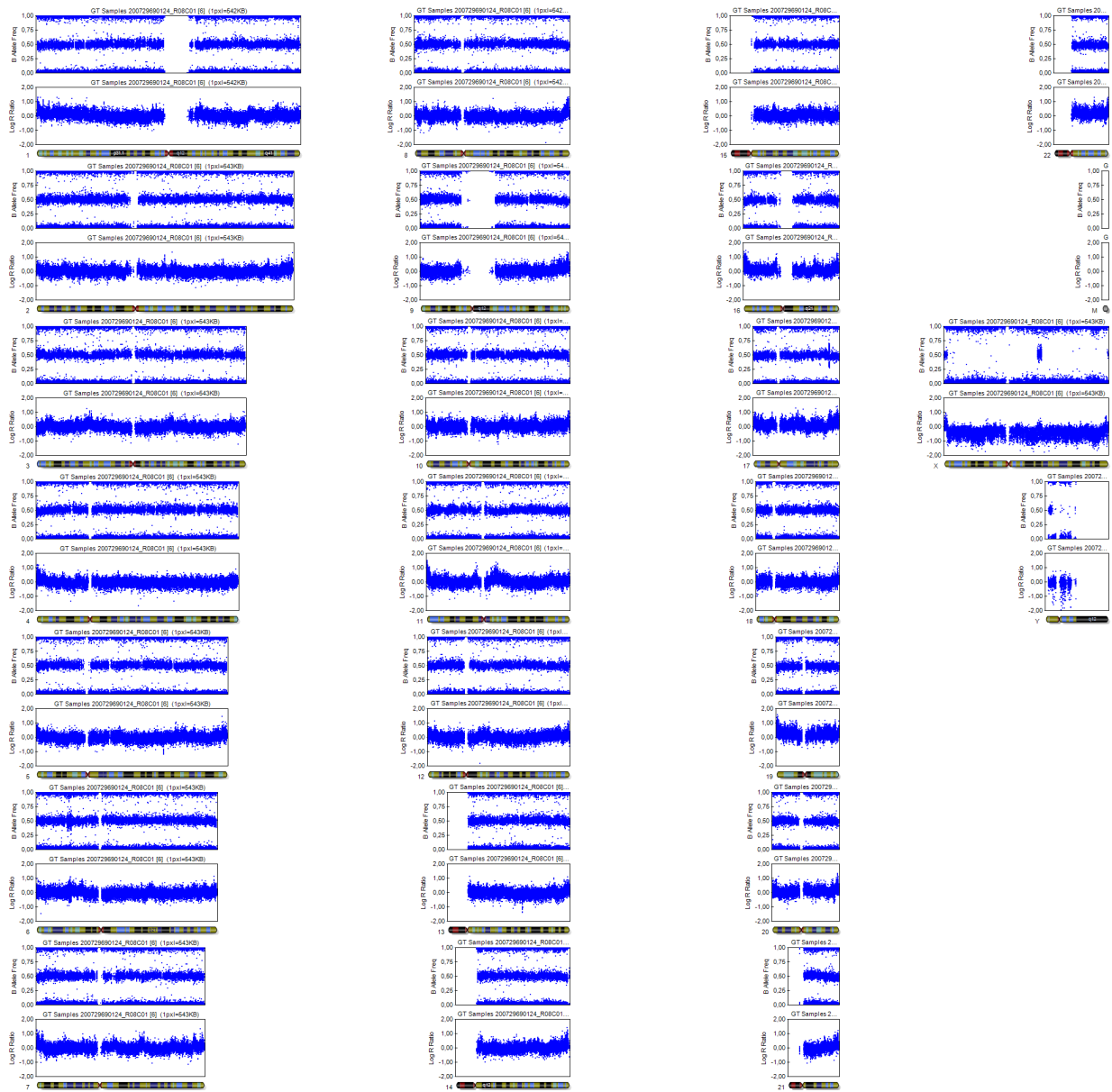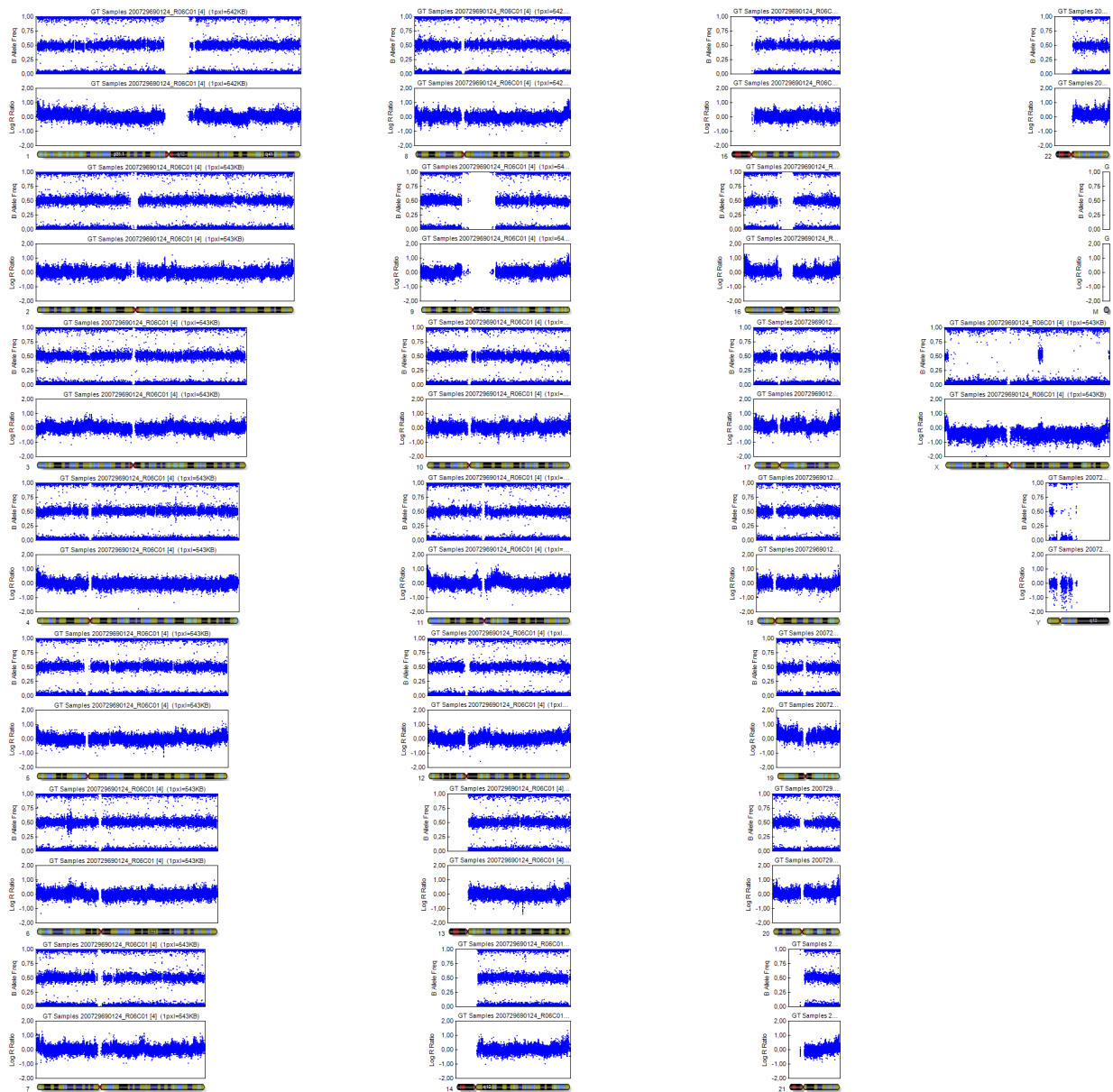
Supplementary Fig. 5: SNP karyotyping of the non-edited (NE) hiPSC line. B allele frequency and log R ratio plots are shown for all chromosomes. No prominent genomic alterations can be detected in the non-edited hiPSC line.
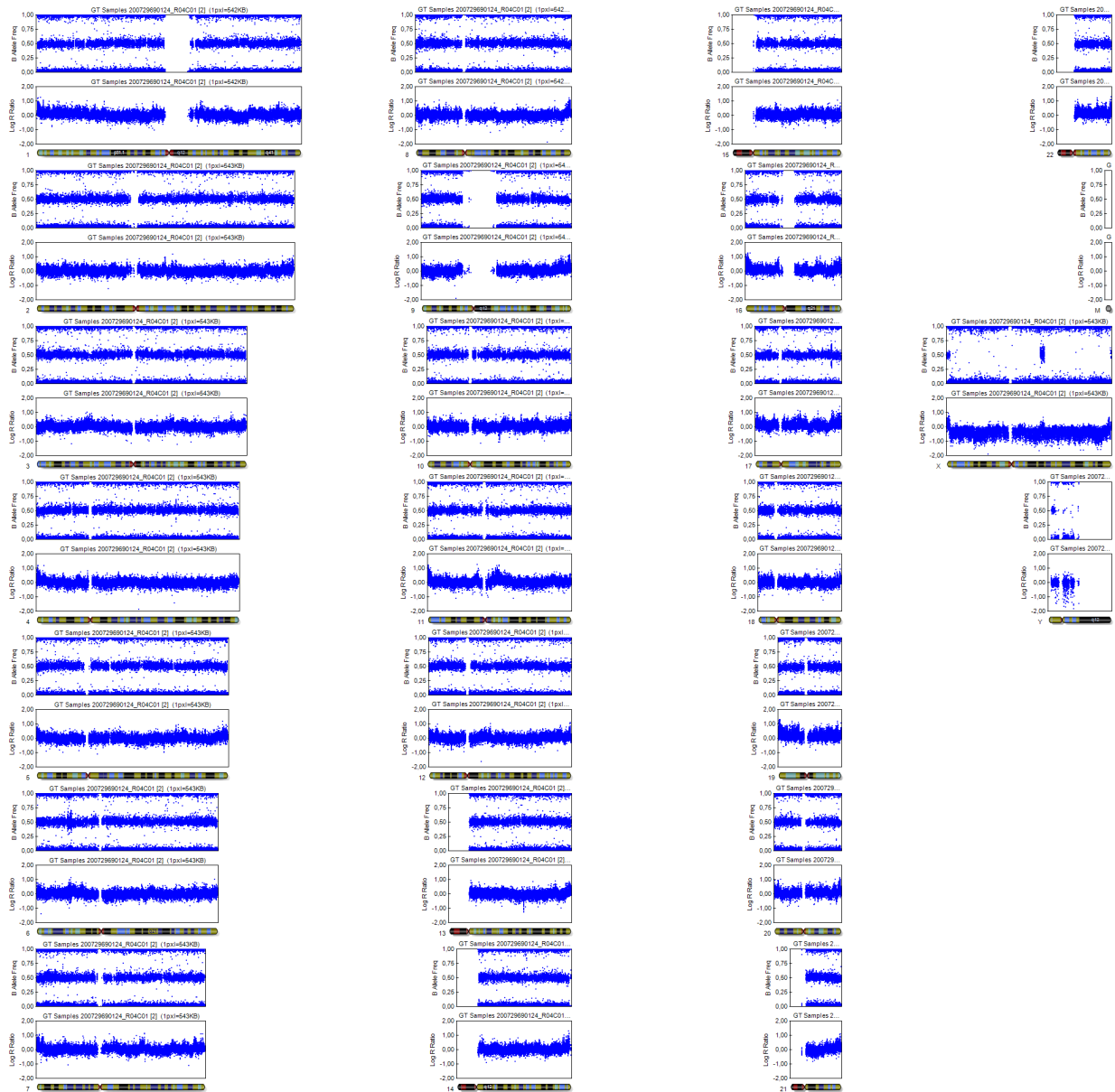
Supplementary Fig. 6: SNP karyotyping of the C1 clone. B allele frequency and log R ratio plots are shown for all chromosomes. No prominent genomic alterations can be detected when compared to the non-edited parental cell line (Supplementary Fig. 5).
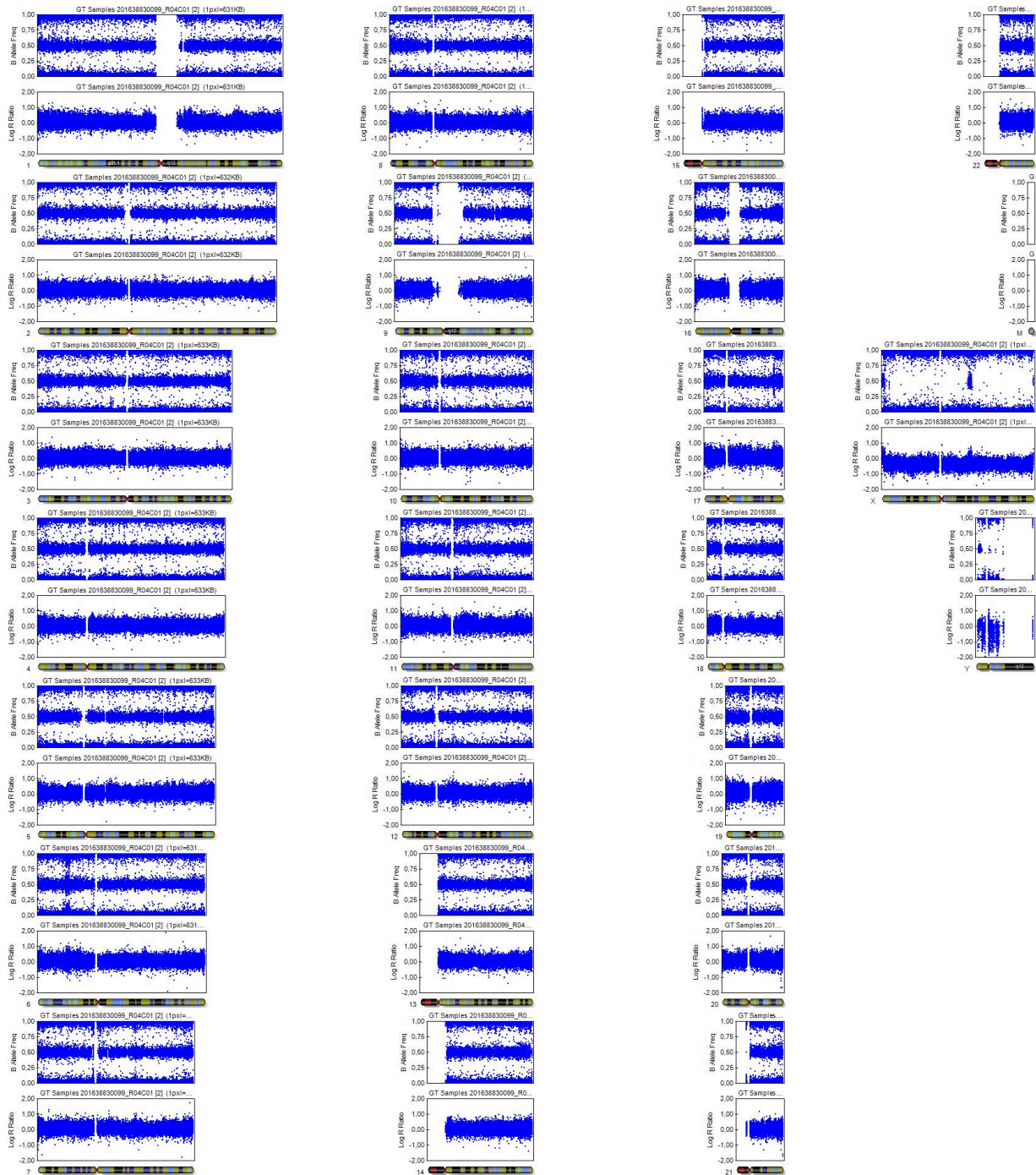
Supplementary Fig. 7: SNP karyotyping of the C4 clone. B allele frequency and log R ratio plots are shown for all chromosomes. No prominent genomic alterations can be detected when compared to the non-edited parental cell line (Supplementary Fig. 5).

Supplementary Fig. 8: SNP karyotyping of the C7 clone. B allele frequency and log R ratio plots are shown for all chromosomes. No prominent genomic alterations can be detected when compared to the non-edited parental cell line (Supplementary Fig. 5).
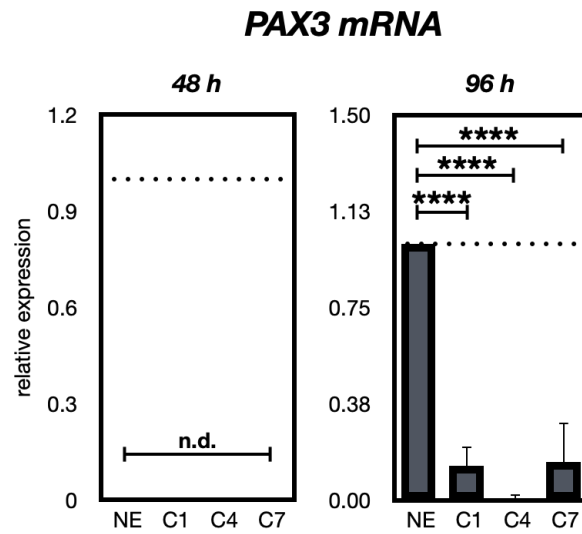
PAX3 mRNA

Supplementary Fig. 9: PAX3 expression is impaired in the GCNF-ablated clones after 96 h of neural induction. Whereas PAX3 mRNA is not detectable after 48 h of neural induction, a prominent reduction of PAX3 mRNA can be observed after 96 h of neural induction in the GCNF-ablated clones when compared to non-edited control (NE). When the signal was below the detection threshold, values were set to 0. Data were normalized to 18s rRNA levels and are presented as mean + SD, relative to the expression in non-edited (NE) cells (set to 1, dashed line; n = 4). ****, p ≤ 0.0001. n.d. = not detectable.

A
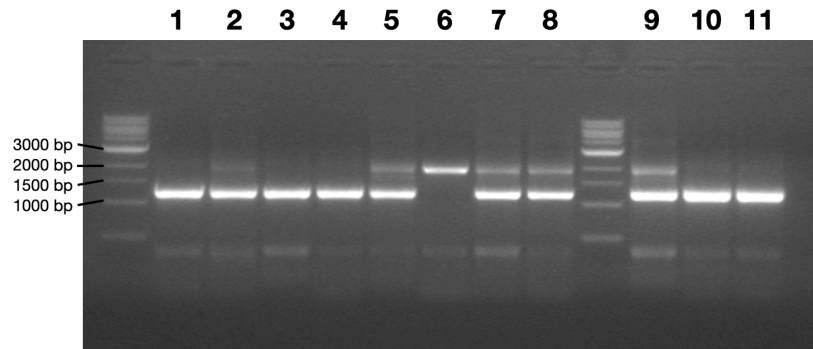
B
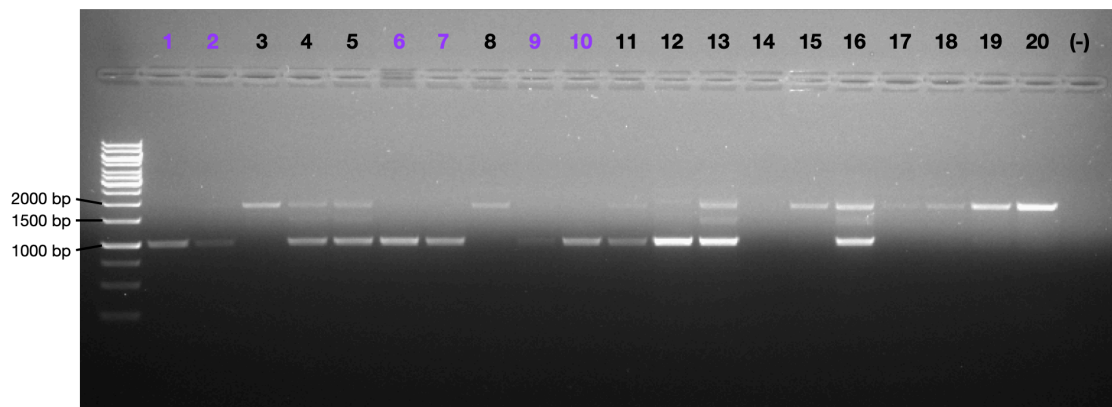
Supplementary Fig. 10: Agarose gels for multiplex PCR reactions for the GCNF tagging approaches. **(A)** The multiplex PCR products for the 11 analyzed clones of the GCNF-3xFLAG hiPSCs were run on a gel. A PCR product of 1900 bp indicates the unaltered genomic location, whereas a PCR product of 1171 bp indicates a successful editing. **(B)** The multiplex PCR products for 20 analyzed GCNF-AM hiPSCs were run on a gel. The unaltered genomic location is indicated by a 1900 bp PCR product, whereas successful incorporation of the exogenous DNA and subsequent Cre-mediated excision is indicated by a PCR product of 1014 bp. Single cell-derived hiPSCs in which the PCR reaction indicates a successful Cre recombination are marked as purple. A water control was included (-).

10. Acknowledgements

First of all, I want to thank Prof. Dr. Oliver Brüstle for giving me the opportunity to write my MD thesis in the Institute of Reconstructive Neurobiology. I am especially grateful that I could learn how to set up the cutting-edge technology CRISPR/Cas9 in stem cells from scratch. It taught me a lot about the scientific method and it was a great experience of participating in a field that already has and will continue to have such a big impact on the pace of biomedicine. I want to thank Dr. Laura Stappert for her supervision and support. Moreover, I want to thank PD Dr. Michael Peitz and PD Dr. Andreas Till for their support and many interesting scientific discussions. I am very grateful to our technicians Monika Veltel and Cornelia Thiele for excellent technical support and a good mood no matter the circumstances. I am also especially grateful to my colleague and friend Frederike Klaus for her great support in the lab and for the interesting scientific and philosophic discussions we had every day in the lab. I also want to thank her for cheering me up in situations that did not work out as good as expected.

I want to thank all members of the Institute of Reconstructive Neurobiology for a nice working atmosphere and all the fun we had in the lab.

Finally, I want to express my deep gratitude to my family and my wife Johanna for all the support in the last years as well as many inspiring ideas and discussions.

11.   List of publications

Braun NC, Klaus F, Veltel M, Stappert L, Peitz M, Brüstle O.
Protracted pluripotency exit of differentiating GCNF-deficient human iPSCs. In review.

Klaus F, Jung K, Schütte S, Veltel M, **Braun NC**, Stappert L, Brüstle O.
The GCNF-BCL11A axis as a regulator of neural stem cell differentiation. In preparation.

Doerr J, Schwarz MK, Wiedermann D, Leinhaas A, Jakobs A, Schloen F, Schwarz I, Diedenhofen M, **Braun NC**, Koch P, Peterson DA, Kubitscheck U, Hoehn M, Brustle O. Whole-brain 3D mapping of human neural transplant innervation. Nat Commun. 2017; 8: 14162. doi: 10.1038/ncomms14162.

Roese-Koerner B, Stappert L, Berger T, **Braun NC**, Veltel M, Jungverdorben J, Evert BO, Peitz M, Borghese L, Brustle O. Reciprocal Regulation between Bifunctional miR-9/9(*) and its Transcriptional Modulator Notch in Human Neural Stem Cell Self-Renewal and Differentiation. Stem Cell Reports. 2016; 7: 207-219. doi: 10.1016/j.stemcr.2016.06.008.