Mechanisms of clinical variability:

Gain vs. loss of function

X-chromosomal reactivation underlying female resilience

Mechanismen der klinischen Variabilität:

Funktionsgewinn gegen Funktionsverlust

X-chromosomale Reaktivierung als Grundlage der weiblichen Resilienz

Dissertation

Zur Erlangung des Grades

Doktor der Naturwissenschaften

Am Fachbereich Biologie

Der Johannes Gutenberg-Universität Mainz

Stephan Käseberg

geb. am 26.02.1990 in Attendorn

Mainz, März 2021

1. Berichterstatter: Prof. Dr. [redacted]

2. Berichterstatter: Prof. Dr. [redacted]

Tag der mündlichen Prüfung: 19.01.2023

Table of contents

Zusammenfassung

Das Opitz BBB/G Syndrom (OS) ist eine monogenetische Erkrankung die hauptsächlich die ventrale Mittellinie betrifft. Typische Symptome sind Hypertelorismus, Hypospadie und Entwicklungsverzögerungen. OS wird durch Mutationen im X-chromosomalen Gen *MID1* verursacht und hat eine deutlich höhere Prävalenz in Männern als in Frauen. MID1 ist eine Ubiquitin Ligase, die in vielen verschiedenen Signalwegen eine Rolle spielt. Es wird angenommen, dass weibliche Mutationsträger aufgrund der zufälligen X-Inaktivierung, die zur Expression des wildtypischen Proteins in ca. 50% aller Zellen führt, weniger schwer betroffen sind als männliche. Die zufällige X-inaktivierung findet während der weiblichen Embryonalentwicklung statt und ist ein Mechanismus zur Dosiskompensation. Es wurden verschiedene Gene beschrieben, die im Menschen der X-inaktivierung entfliehen, entweder ubiquitär oder zellartspezifisch.

Wir haben induzierte pluripotente Stammzellen (iPSCs) generiert, die unterschiedliche Mutationen im *MID1* Gen tragen. Mittels CRISPR/Cas9 Genomeditierung wurden Mutationen am N-terminalen Ende von MID1 in männliche Kontrollzellen eingefügt. Außerdem wurden patienten-spezifische männliche iPSCs mit einer hemizygoten Mutation im C-terminus des MID1 Proteins generiert. Weibliche iPSCs mit derselben Mutation, heterozygot, wurden auf ihren X-Inaktivierungsstatus hin untersucht. iPSC-Klone die entweder nur das wildtypische oder nur das mutante MID1 Protein exprimieren wurden identifiziert, kloniert und so als isogene Kontrollen für weitere Experimente verwendet. Männliche und weibliche iPSCs wurden in cerebrale Organoide (in Kooperation mit Prof. Dr. [redacted] und Prof. Dr. [redacted]) oder 2-dimensional wachsende Neuronen differenziert.

Die cerebralen Organoide, die aus den männlichen iPSCs mit den N-terminalen Mutationen im *MID1* Gen generiert wurden, entwickelten einen Hyperdorsalisierungsphänotyp, der mittels Färbung von TTR, einem Marker des Plexus choroideus nachgewiesen wurde. In weiteren Experimenten konnte gezeigt werden, dass der SHH Signalweg in neuronalen Vorläuferzellen gestört ist, vermutlich verursacht durch N-terminal trunkierte MID1 Proteine, die durch die Verwendung von alternativen Translationsstarts im ersten kodierenden Exon von *MID1* entstehen. Wir konnten zeigen, dass diese Proteine ihre Fähigkeit mit den Mikrotubuli zu assoziieren behalten. Dies legt die Hypothese nahe, dass es sich um einen „gain-of-function“ Mechanismus handelt. Da keine OS Patienten mit N-terminalen MID1 Mutationen bekannt sind, vermuten wir, dass diese „gain-of-function“ Mutationen im frühen Embryonalstadium letal sind.

Im Gegensatz dazu zeigten die cerebralen Organoide, die aus den patienten-spezifischen iPSCs mit C-terminaler Mutation generiert wurden, einen anderen, milderen Phänotyp. Die neuronalen Vorläuferzellen in den Organoiden konnten den Zellzyklus nicht verlassen, was zu einer Akkumulation von PAX6 positiven Zellen führte und in der Folge zu einer Reduktion von Neuronen. Dieser Phänotyp entspricht wahrscheinlich dem allgemein akzeptierten „loss-of-function“ Phänotyp, wie er auch in OS Patienten auftritt. Ein Vergleich der cerebralen Organoide, die aus weiblichen iPSC-Klonen generiert wurden, die die *MID1* Mutation auf dem aktiv exprimierenden X-Chromosom tragen, mit solchen, die aus hemizygoten männlichen iPSCs mit derselben Mutation generiert wurden, zeigte einen deutlich milderen Phänotypen in den weiblichen als in den männlichen Organoiden. In weiteren Experimenten konnten wir zeigen, dass wildtypisches MID1 in den weiblichen Patientenzellen differenzierungsabhängig in neuronalen Zellen vom inaktiven X-Chromosom reaktiviert wird. Darüber hinaus haben wir weitere Gene identifiziert, die ebenfalls reaktiviert werden. Der Großteil dieser Gene spielt eine Rolle in der neuronalen Entwicklung oder wirkt als epigenetische Modulatoren. Wir vermuten, dass die gehirnspezifische Reaktivierung dieser Gene eine schützende Wirkung während der weiblichen Gehirnentwicklung hat und zur höheren Plastizität im weiblichen Gehirn beiträgt.

Abstract

Opitz BBB/G syndrome (OS) is a monogenic disorder that mainly affects ventral midline structures. Typical symptoms include hypertelorism, hypospadias and developmental delay. OS is caused by mutations in the X-chromosomal gene *MID1* and presents with a clear gender bias towards males. MID1 is a ubiquitin ligase that is involved in several signalling pathways. Female mutation carriers are thought to be less severely affected compared to males because of random X-chromosomal inactivation (XCI) that results in approximately 50% of their cells expressing the wildtype protein. XCI occurs in female embryos during early development and is a dosage compensatory mechanism. Several genes are known to escape XCI in humans either ubiquitously or in a tissue-specific manner.

We have generated induced pluripotent stem cells (iPSCs) carrying different mutations in *MID1*. Using CRISPR/Cas9 genome editing, mutations at the N-terminal end of MID1 were introduced in male wildtype iPSCs. Furthermore, patient-specific male iPSCs with a hemizygous mutation at the C-terminal end of MID1 were generated. Female iPSCs carrying the same mutation heterozygously were screened regarding their X-inactivation. iPSC-clones expressing either wildtype or mutant MID1 exclusively were identified, cloned and used as isogenic controls for further experiments. Male and female iPSCs were differentiated into cerebral organoids (in collaboration with Prof. Dr. Benedikt Berninger and Prof. Dr. [redacted]) and 2-dimensional neuronal cultures.

The cerebral organoids generated from the male iPSCs carrying N-terminal mutations in MID1 developed a hyperdorsalization phenotype shown by immunofluorescent staining of TTR, a marker of the choroid plexus. Further experiments revealed an altered SHH signalling in neuronal progenitor cells (NPCs) probably caused by N-terminal truncated MID1 proteins that are translated from alternative translation initiation sites in the first coding exon of *MID1*. We were able to show that these proteins do not lose their ability to bind to microtubules. This suggests the hypothesis of a gain-of-function mechanism being responsible for the phenotype. Since no OS patient carrying an N-terminal mutation in MID1 has been reported, we think that these gain-of-function mutations are lethal during early embryonic development.

In contrast, the cerebral organoids generated from the patient-specific iPSCs carrying a C-terminal MID1 mutation developed a different, milder phenotype. The neuronal progenitors failed to exit the cell cycle, leading to an accumulation of PAX6 positive cells and following, a reduced number of neuronal cells. This phenotype is likely to resemble the commonly accepted loss-of-function phenotype present in OS patients. Comparison of the cerebral organoids

generated from the female iPSC-clones carrying the *MID1* mutation on the active X-chromosome, with those generated from the hemizygous male iPSCs carrying the same mutation, revealed a milder phenotype in the female cerebral organoids. Further experiments showed that expression of wildtype MID1 from the inactive X-chromosome is being reactivated in the female patient-specific cells upon differentiation into neuronal cells. We were also able to identify more genes that escape XCI in neuronal cells. The majority of these genes is involved in neurodevelopmental processes and epigenetic modulation. We think that the brain-specific reactivation of these genes is a protective factor during female brain development and causes a higher plasticity in the female brain.

1 Introduction

1.1 Induced pluripotent stem cells as a model for disease

Human induced pluripotent stem cells (iPSCs) were first introduced in 2007 by a working group around Shinya Yamanaka (Takahashi et al., 2007). They were able to generate pluripotent stem cells from adult fibroblasts that resembled embryonic stem cells in DNA methylation, gene expression and protein expression by retroviral transfection of four genes that are now known as the Yamanaka factors: ***K****LF4*, *c-****M****YC*, ***O****CT4*, ***S****OX2* (KMOS). Although, each of these genes has been shown to be replaceable by other genes or small molecules, the KMOS transcription factors remain the most abundantly used combination for reprogramming cells (Guo & Chen, 2015).

KLF4 is part of the C2H2-type zinc-finger krueppel family of transcription factors and plays an important role in cell proliferation and differentiation by controlling the G1- to S-phase transition during the cell cycle. It can also induce apoptosis, as it is a known downregulator of the apoptotic protein p53. The zinc-finger protein is also a potential tumor suppressor gene and plays an important role in the formation of the skin barrier function (Ghaleb & Yang, 2017; *KLF4 - Krueppel-like Factor 4 - Uniprot*, n.d.; *KLF4 Gene - GeneCards*, n.d.).

c-MYC is a nuclear-phosphoprotein and as a transcription factor involved in cell cycle progression, apoptosis and cellular transformation. It is often upregulated in cancers and is considered a proto-oncogene making it the most frequently criticized gene of the Yamanaka factors. Although it was shown that c-MYC is not crucial for successfully reprogramming cells, it is still being used as it improves reprogramming efficiencies significantly. In embryonic stem cells (ESCs) it controls self-renewal (*C-Myc - C-Myc Protein - Uniprot*, n.d.; *MYC Gene - GeneCards*, n.d.; Miller et al., 2012).

OCT4 is a POU-type transcription factor with a highly conserved octamer-binding domain important for embryonic development and stem cell pluripotency. Dysregulation of OCT4 can induce tumor development (*POU5F1 - POU Domain, Class 5, Transcription Factor 1 - Uniprot*, n.d.; *POU5F1 Gene - GeneCards*, n.d.; G. Shi & Jin, 2010; Zeineddine et al., 2014). It was shown that overexpression of OCT4 alone is sufficient to induce reprogramming in certain cell types (J. B. Kim et al., 2009; Tsai et al., 2011), by contrast, reprogramming using the Yamanaka factors except OCT4, results in iPSCs with less deteriorating epigenetic changes and thus of higher quality (Velychko et al., 2019).

SOX2 is part of the SRY-related HMG-box family of transcription factors. By forming a trimeric complex with OCT4 and DNA, it controls expression of genes important for embryonic development and cell fate determination. It is mediating neural stem cell self-renewal and might function as a switch by counteracting genes inducing neuronal differentiation (Schaefer & Lengerke, 2020; *SOX2 - Transcription Factor SOX-2 - Uniprot*, n.d.; *SOX2 Gene - GeneCards*, n.d.; S. Zhang, 2014).

Before the development of iPSCs several other cell types were used as models for neuronal diseases that all showed different limitations. One of the first disease models were lymphoblastoid cell lines that were generated from lymphocytes of patients by transfection with the Epstein-Barr Virus *in vitro* (Omi et al., 2017). These cells can easily be generated in a patient-specific manner but have clear limitations when it comes to modelling diseases not related to blood, especially neuronal diseases as lymphocytes are derived from the mesoderm, while the brain is derived from the ectoderm (Jack et al., 2014). Another cell type very common for modelling neuronal diseases are fibroblasts. These can also be easily generated in a patient-specific manner, although a skin biopsy, that is necessary for the generation of the fibroblasts is more invasive than just taking blood (Vangipuram et al., 2013). Fibroblasts are also mesodermal cells but often show expression of genes associated with neuronal cells (which are ectodermal cells), though at lower levels (Auburger et al., 2012). Both of these cell types were and still are commonly used in labs for modelling neuronal and non-neuronal diseases. When ESCs were first isolated from human pre-implantation embryos in 1998, they were expected to be the future gold standard in modelling diseases (Thomson, 1998). ESCs are pluripotent and can be differentiated into every cell type except the placenta. Protocols for neural cells (S. C. Zhang et al., 2001), cardiomyocytes (He et al., 2003) and other cell types were soon available, enabling research on the cell type potentially affected by a disease. Unlike fibroblasts and lymphoblastoids, ESCs are usually not available in a patient-specific manner, making precise genome editing necessary, to induce mutations associated with a disease. And although this is possible, ESCs are still lacking the genetic and epigenetic background of the patients (Rowe & Daley, 2019). To reach both, pluripotency and patient-specificity, iPSCs are needed. They can be differentiated into every cell type of the body (except placenta) like ESCs and they are generated from patient-specific somatic cells.

1.1.1 Embryonic stem cells versus induced pluripotent stem cells

Apart from their pluripotency and ability to be differentiated into every cell type except for placenta, ESCs and iPSCs share many characteristics (Figure 1). Both show a highly similar

chromatin methylation pattern, promoter and histone demethylation (Brouwer et al., 2016). Also, both cells types show similar global DNA methylation pattern, although a few marks are specific to iPSCs. It is hypothesized that these differences point towards an incomplete reprogramming of iPSCs (Lister et al., 2011). The comparable epigenetic status leads to the typical transcriptome and proteome that can be found in stem cells and is used for confirming iPSCs as stem cells after reprogramming. When culturing iPSCs and ESCs, both show a typical morphology of monolayer cells growing as colonies with clear borders and big nuclei that make up 90% of the total cell volume. Also, the doubling time is in a comparable range and does not change over time; both can be expanded indefinitely due to telomerase activity (Takahashi et al., 2007; Takahashi & Yamanaka, 2006).

To prove potency and differentiability of iPSCs and ESCs, the ability of both cell types to form embryoid bodies (EBs) is tested, a three-dimensional aggregate of pluripotent stem cells that is used in a first step of many protocols describing differentiation of stem cells. Another way to prove pluripotency is the injection of cells into immunosuppressed mice where only stem cells are able to form teratomas, tumors consisting of cells of all three germ layers (Brouwer et al., 2016).

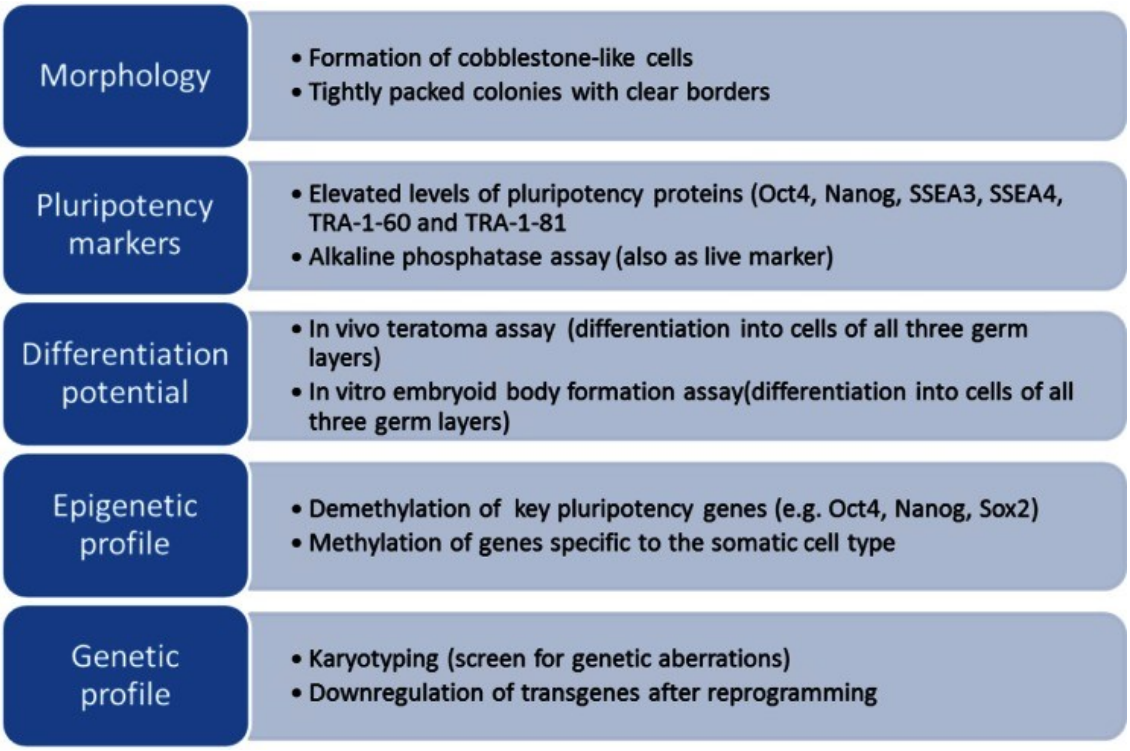


Figure 1 Factors used for characterizing reprogrammed iPSCs. ESCs and iPSCs commonly share different characteristics that are assessed for characterization and quality control (Brouwer et al., 2016).

With all the similarities that iPSCs and ESCs share, two substantial differences exist. First, the already mentioned patient-specific character of iPSCs, and second, the ethical concerns that arise when using ESCs. ESCs are generated by disrupting a pre-implantation embryo that is able to form life, while iPSCs are generated from somatic cells of a donor with almost no physical injury.

1.1.2 Different Reprogramming methods

The first protocols for generating human iPSCs in 2007 were all based on viruses integrating the reprogramming factors into the genome of the host cell and by this potentially disrupting other genes (Takahashi et al., 2007; Yu et al., 2007). Over the past years several non-integrative viral or non-viral based reprogramming methods were developed that reprogram cells with the same or better efficiency as the first integrative viral systems (Brouwer et al., 2016). These include e.g. sendaiviral transfection, plasmid-based reprogramming or protein-based reprogramming. To overcome the use of integration-based reprogramming was the first step to lower the potential of iPSCs to proliferate uncontrollable in a cancer like manner. The next step is the adjustment and control of exogenous and endogenous reprogramming factor expression (Okano et al., 2013), as all of the Yamanaka factors are in some way associated with cancer (see 1.1). The use of other genes or a combination of other genes might be one possibility, but many of the genes known to be able to induce reprogramming can in some way be connected to cancer.

Another important advance that was made since the first description of iPSCs is the use of different cell types as starting material for reprogramming. The first protocols all used fibroblasts for reprogramming, that can only be obtained by a small skin biopsy under local anaesthesia leaving a visible scar. Over the recent years protocols successfully using keratinocytes (Yamanaka, 2010), PBMCs (Peripheral Blood Mononuclear Cells) from blood (Staerk et al., 2010) and renal epithelial cells from urine (Zhou et al., 2012) for reprogramming were published (Yamanaka, 2010).

1.1.3 Differentiation of induced pluripotent stem cells

With the first successful isolation of ESCs in 1998 researchers started to develop protocols to differentiate stem cells into the desired cell type. First protocols were soon available for neurons, cardiomyocytes, pancreatic β-cells and blood cells (see 1.1), all with the long-term perspective of an allogeneic cell therapy (Rowe & Daley, 2019). When iPSCs became available many of these protocols were adapted and used for patient-specific disease modelling avoiding the need for gene editing as in ESCs. Using this approach, scientists were able to identify

phenotypes associated with different diseases including e.g. spinal muscular atrophy (SMA), a type of neuromuscular disorder resulting in the loss of motor neurons. When differentiating iPSCs from patients suffering from SMA into motor neurons, a significant reduction of motor neurons was observed in comparison to wildtype cells (Ebert et al., 2009). The observed phenotype was stably reproducible and eventually led to the several discoveries of small molecules rescuing the observed phenotype on the cellular level (Lee et al., 2012; Yanhong Shi et al., 2017). Unfortunately, all of these molecules failed to progress to clinical trials (Rowe & Daley, 2019).

Newer research now aims to generate three-dimensional organoids resembling the cellular mixture and heterogeneity found in human organs and thus improving the readout and translation to clinical trials. Such three-dimensional organoids are for example cerebral organoids containing functional neurons (Lancaster et al., 2013; Lancaster & Knoblich, 2014). The cells show a cortical self-organization with regions expressing markers specific for forebrain and hindbrain. Spontaneous electrical activity can be observed in the neurons. To better resemble human body function, organoids can be transplanted into mice, creating human mouse chimeras. The human organoids are vascularized in the mouse and interact with the body via hormones, the blood system, the nervous system and the immune system. The next years will show how much this improves the outcome and potential to translate the results to clinical trials (Rowe & Daley, 2019).

1.1.4 Isogenic controls

The patient-specific character favours iPSCs over ESCs, nevertheless, this does not mean that genome editing is not required. While ESCs are typically edited to induce a certain patient-specific mutation (Rais et al., 2013), patients' mutations in iPSCs are specifically repaired to achieve isogenic controls (Yanhong Shi et al., 2017). Studies have shown that iPSCs generated from unrelated healthy patients differ significantly in DNA methylation, gene expression and their efficiency to be differentiated. A key study by Kilpinen et al. analysed 711 iPSC lines derived from 301 healthy donors by the "Human Induced Pluripotent Stem Cells Initiative" to identify differences caused by the genetic background, but also differences caused by the reprogramming procedure. They found 5-46% of the variation in iPSC phenotypes to be caused by genetic background differences (Kilpinen et al., 2017). To model diseases *in vitro*, it is therefore inevitable to use iPSC lines as controls that are genetically as similar as possible to disease lines. Over the past years several precise genome editing tools have been published, with CRISPR/Cas being the most widely used.

1.2 CRISPR/Cas genome editing

CRISPR/Cas genome editing originates from a prokaryotic adaptive immune system that protects bacteria against macrophages (Barrangou et al., 2007) and was first proposed as a tool for genome editing in 2012 (Jinek et al., 2012). The CRISPR/Cas system requires two main components, the Cas9 protein and the sgRNA (single guide RNA) (Figure 2). The Cas9 protein is the active part of the system that cuts the DNA at the desired position. There are several different Cas9 proteins available with different characteristics. While the "classical" Cas9 needs a specific PAM-sequence (protospacer adjacent motif) of NGG and induces a double strand break, there are Cas9 proteins from other bacteria available that recognize alternative PAM-sequences, increasing the number of possible loci for genome editing (Waddington et al., 2016). There are also artificially mutated Cas9 proteins available that do not induce a double strand break but just a single strand nick or are completely inactive and fused to a repressor or activator to modulate gene expression (Cong et al., 2013; Ledford, 2016). The second key component is the sgRNA that consists of two subdomains, the crRNA (crisprRNA) and the tracrRNA (trans-activating crRNA). The tracrRNA forms a secondary structure that interacts with the Cas9 enzyme and facilitates sgRNA and Cas9 interaction while the crRNA contains the 20 bp sequence specific for the gene locus that is targeted (Jinek et al., 2012). Together, the Cas9 protein and the sgRNA form an RNA-protein complex that is able to cut DNA at a specific locus.

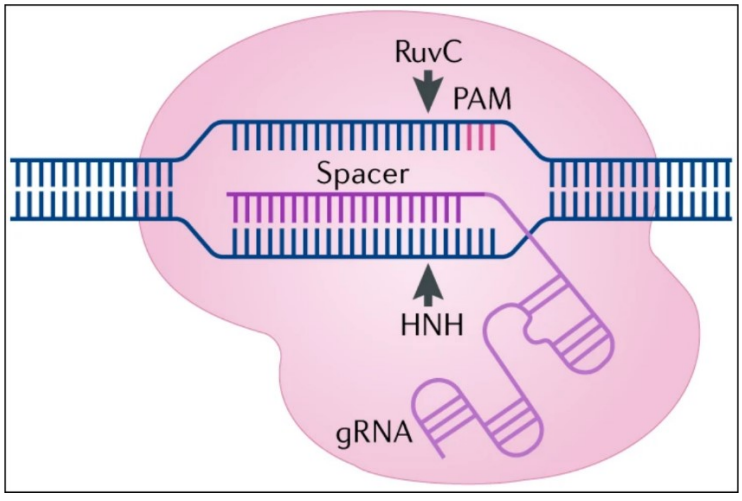


Figure 2 CRISPR/Cas genome editing. The crRNA of the gRNA recognizes the target sequence (blue) and facilitates Cas9 completion. Cas9 induces a double-strand break via its active domains three bases adjacent to the PAM sequence (Pickar-Oliver & Gersbach, 2019).

Over the past years many labs and companies have developed their own systems for quick and efficient sgRNA design, cloning and transfection together with an optimized Cas protein. The

first protocols published used a plasmid-based transfection of sgRNA and Cas, either on a single plasmid or on two individual plasmids, relying on the cellular transcription and translation system. Newer publications are also using pre-synthesized sgRNAs and Cas proteins that do not require any cellular modifications after the transfection. After induction of a double strand break, the cell has two different options to repair the DNA damage, either using the Non-homologous end-joining (NHEJ) pathway, or using the homology directed repair (HDR) pathway (ben Jehuda et al., 2018).

1.2.1 Non-homologous end-joining

The NHEJ pathway is an unspecific DNA repair mechanism that is error prone and often leads to small insertions or deletions but does not require any region of homology (Figure 3). After induction of a double strand break, several proteins are involved in recognizing and repairing (Guirouilh-Barbat et al., 2004; Iyama & Wilson, 2013). In a first step the two DNA ends are bound by DNA protein kinases that hold them in proximity and recruit other enzymes required for the next steps. Mutations typically occur during the end processing, where enzymes remove nucleotides if the two ends are not compatible or lacking the 3'hydroxyl or 5'phosphate overhangs needed for ligation. The final step is the ligation of the two DNA ends by the DNA ligase IV complex.

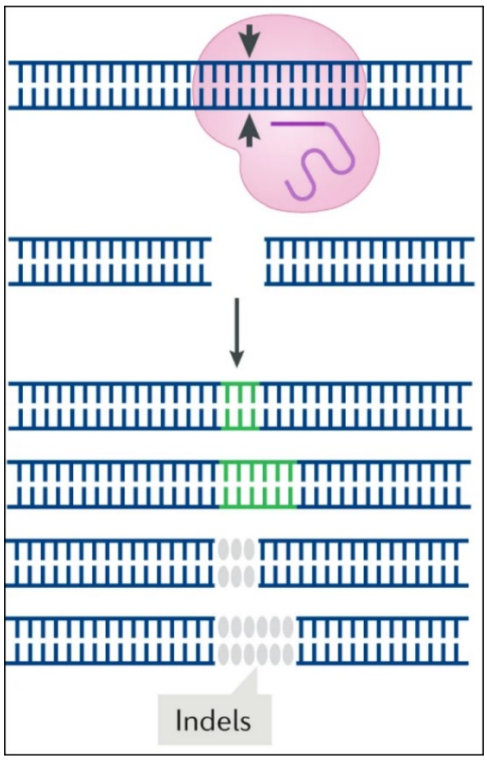


Figure 3 NHEJ mediated DNA repair after CRISPR/Cas. After induction of a double strand break, the NHEJ pathway is used to ligate the two DNA strands frequently causing an Indel mutation (Pickar-Oliver & Gersbach, 2019).

1.2.2 Homology directed repair

HDR is the precise way of the cell to repair a double-strand break but, unlike NHEJ, requires a template sequence. This template sequence can either come from the sister chromosome or can be co-transfected with the sgRNA and the Cas protein to induce specific mutations or repair a mutation (Figure 4). The first protein on site is the MRN complex binding both DNA ends and holding them in proximity (Iyama & Wilson, 2013). The complex then recruits enzymes that produce short 3' overhangs of single-stranded DNA by trimming the 5' ends. The RPA protein then binds the single-stranded DNA and searches with the help of other proteins for a similar sequence. When a similar sequence (from the template) is found, they form a displacement loop allowing a DNA polymerase to synthetize new DNA by extending the 3' end. The structure is then called a Holiday junction and the cell can either continue with the double-strand break repair (DSBR) or the synthesis-dependent strand annealing (SDSA) pathway. DSBR leads to the second 3' end also forming a Holiday junction and then commonly to a crossing over. The SDSA leads to a repair without crossing over as the newly synthesized overhang leaves the Holiday junction and is then able to anneal to the damaged 3' overhang remaining. A ligase then finishes the pathway by closing all remaining single strand gaps.

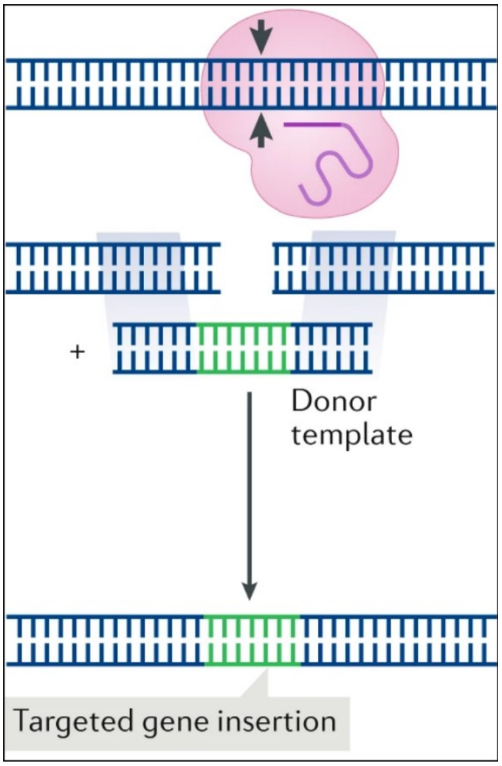


Figure 4 HDR mediated DNA repair after CRISPR/Cas. After induction of a double strand break, the HDR pathway is used to ligate the two DNA strands using a donor template (Pickar-Oliver & Gersbach, 2019).

1.3 Opitz BBB/G syndrome

Opitz BBB/G syndrome (OS) is a congenital disorder with a prevalence of 1-9:100.000 (OMIM: 300.000, ORPHA: 2745). It is characterized by a broad variety of different symptoms mainly affecting the midline. Most typically patients suffer from hypertelorism and hypospadias, while cleft lip/palate, laryngotracheoesophageal abnormalities, imperforate anus and cardiac defects occur only in a smaller fraction of patients (Figure 5). Of note, about one third of patients suffer from developmental delay (de Falco et al., 2003; So et al., 2005). Typically, males show more symptoms with a more severe manifestation, while females often only present with hypertelorism.

Figure 5 Patient with OS. The hypertelorism, telecanthi, mild entropion of lower eyelid, high nasal bridge and surgically corrected cleft lip are evident (Quaderi et al., 1997).

OS was already described in the 1960s as two different disorders, the BBB syndrome and the G syndrome, but further studies in the 1990s suggested they are the same disorder (Quaderi et al., 1997; Robin et al., 1995). Segregation analysis pointed early towards an X-linked monogenic disorder (Robin et al., 1995), and the *MID1* gene on Xp22 causative for OS was identified in 1997 (Quaderi et al., 1997). A second form of OS maps to chromosome 22 (Robin et al., 1995).

1.3.1 *MID1* gene and protein

The *MID1* gene consists of nine constitutive coding exons. It is located on Xp22.2 and spans ~400 kb. Transcription of *MID1* is controlled by at least five promoter regions, generating transcripts with alternative 5' untranslated regions. While some of these are tissue-specific, others show ubiquitous expression (Landry & Mager, 2002; Quaderi et al., 1997; Winter et al.,

2016). The gene encodes for the 667 aa long MID1 protein consisting of six domains with different functions (Figure 6). The RING-finger domain is the most N-terminal part of the protein followed by two B-Boxes. Both are characterized by Cys-His motifs and play an important role in protein-protein interactions. The Coiled-coil domain is promoting homodimerization of MID1 and heterodimerization of MID1 with its sister protein MID2. The two C-terminal domains fibronectin type III (FNIII) and B30.2 are also protein-protein interaction domains. The C-terminal end of MID1 is responsible for microtubule association that allows transportation of MID1 along the microtubules (Winter et al., 2016). The functional MID1 protein is an E3 ubiquitin ligase regulating other proteins via ubiquitination, including PP2Ac, Fused (FU) and PAX6 (see also 1.3.3, 1.3.4 and 1.3.5).

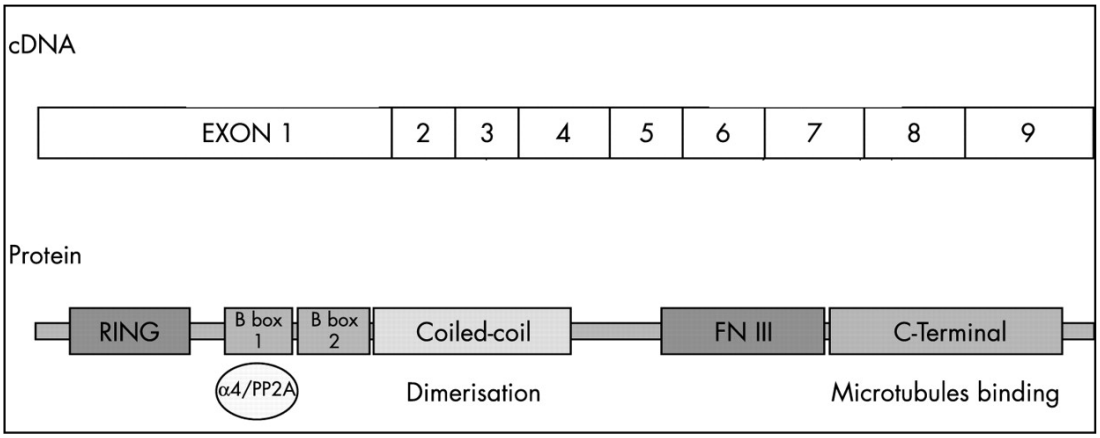


Figure 6 MID1 cDNA and protein. The cDNA consists of nine coding exons. The protein has six domains either important for protein-protein interaction or dimerization (Pinson et al., 2004).

1.3.2 Mutations in *MID1*

To date many different mutations have been reported for *MID1* with no clear genotype-phenotype correlation (Figure 7). Even patients with the same mutation can show a different manifestation of symptoms. Most of the mutations are leading to an early stop codon or are point mutations causing a single amino acid change. These can be, together with small in frame deletions or duplications and splice-site mutations, found over the whole gene except for the region coding for the N-terminal RING-finger. In addition, there are patients with full deletions or duplications of single exons and even a deletion of the whole gene has been reported. The full gene deletion and the reported nonsense-mediated mRNA decay (NMD) in some patients point towards a loss-of-function mechanism responsible for the OS phenotype (Winter et al., 2016).

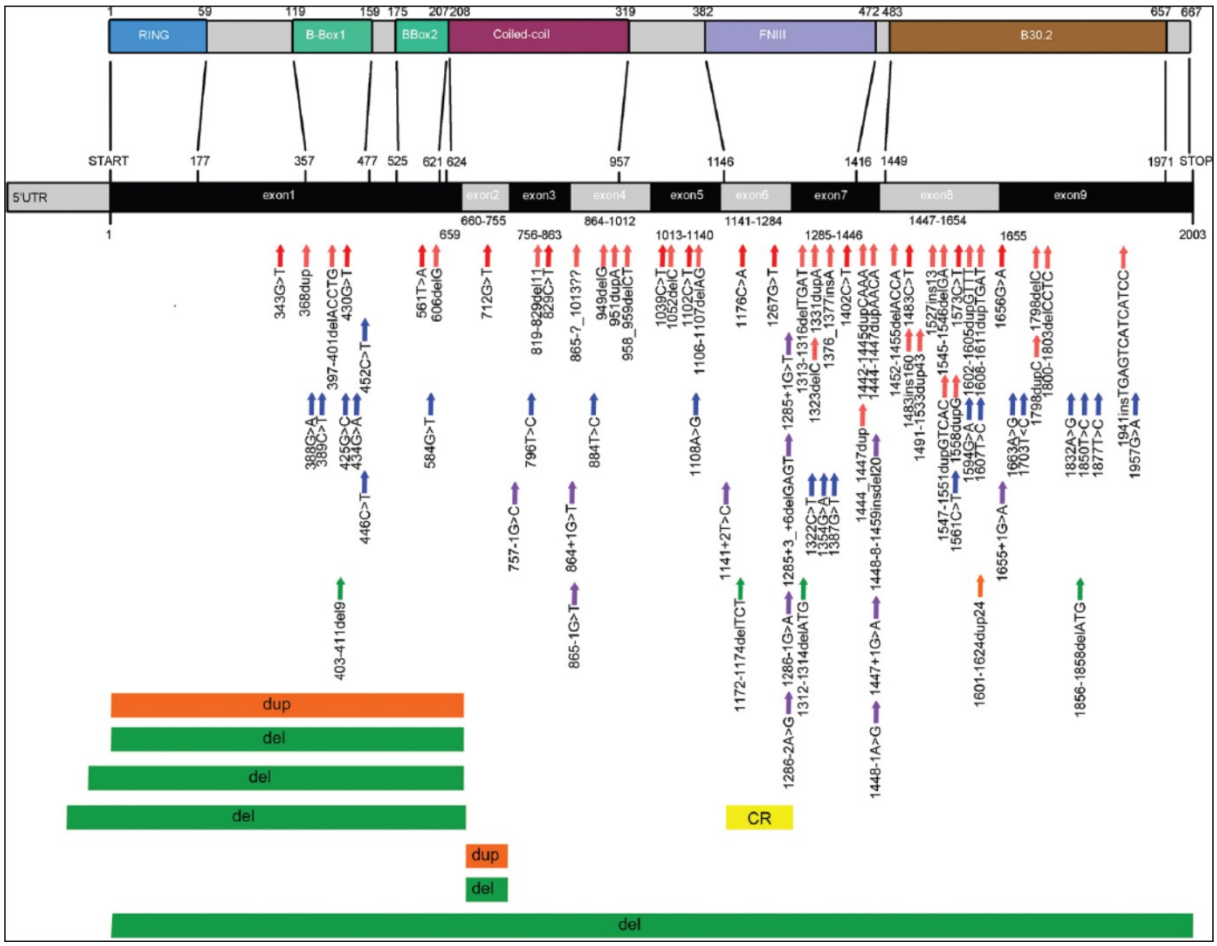


Figure 7 MID1 mRNA and protein with mutations that have been reported in patients. Red arrow: mutations that lead to an early stop codon; blue arrow: mutations that lead to amino acid changes; orange bars/arrows: duplications; green bars/arrows: insertions; violet arrows: mutations affecting splicing; yellow bar: complex rearrangement (Winter et al., 2016).

1.3.3 MID1 and mTOR

The first protein identified to be a target of MID1 was the catalytic subunit of protein phosphatase 2A (PP2Ac) that is polyubiquitinated by MID1 and thus marked for degradation by the proteasome (Figure 8). Mutated MID1 is not able to bind PP2Ac as it aggregates in the cytoplasm if translated at all, leading to elevated PP2Ac levels and hypophosphorylation of microtubule associated proteins (Trockenbacher et al., 2001). Both, elevated PP2Ac and hypophosphorylation of microtubule associated proteins are thought to contribute to the OS pathogenesis.

The mammalian target of rapamycin (mTOR) protein and signalling pathway is regulated by intra- and extracellular signals serving as a key regulator of cell metabolism, growth and proliferation. Upstream signals include insulin and growth factors, as well as cellular energy and nutrient levels. The mTOR pathway has been found to be dysregulated e.g. in diabetes,

obesity and certain cancers (Laplante & Sabatini, 2009). The mTOR protein is the central part of two protein complexes: mTORC1 and mTORC2. While mTORC1 is regulating translation and serving as an energy sensor via 4EBP1 and p70S6K, mTORC2 is more important for cytoskeleton organization and regulates cell survival and proliferation via AKT (Saxton & Sabatini, 2017).

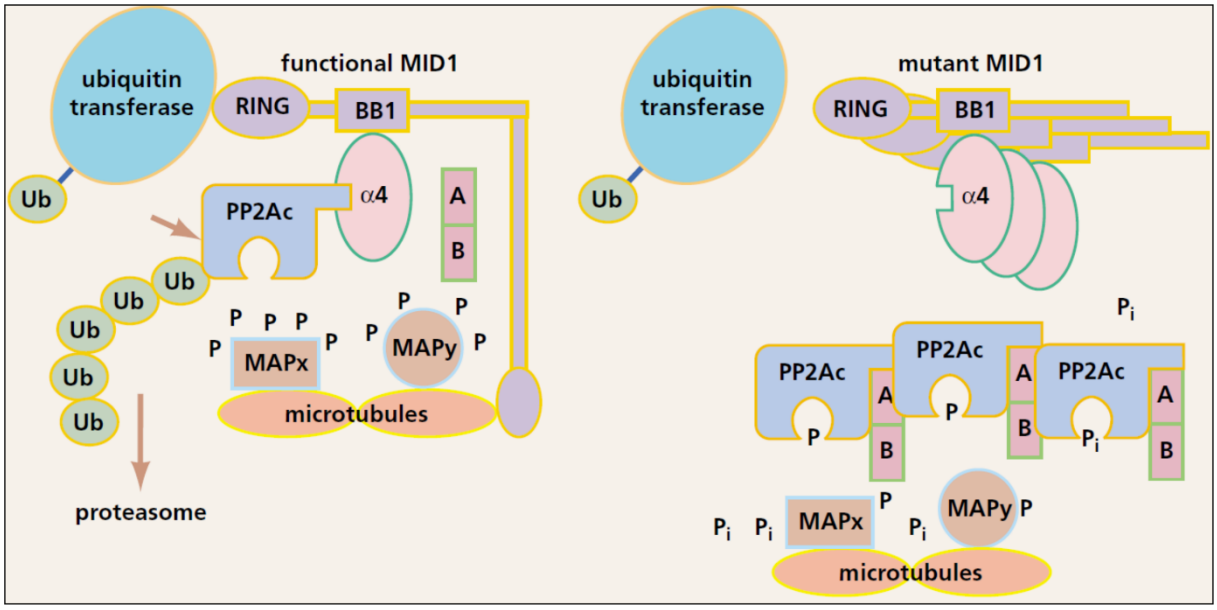


Figure 8 MID1 protein function. Left: Wildtype MID1 binds to the microtubules with its C-terminal end. The B-Box1 subunit binds to α4 that mediates binding of PP2Ac by replacing subunits a and b. PP2Ac is then polyubiquitinated by a ubiquitin transferase binding to the N-terminal RING-finger domain. The polyubiquitinated PP2Ac is degraded in the proteasome. Furthermore, MID1 binding leads to hyperphosphorylation of microtubule-associated proteins. Right: C-terminally mutated MID1 aggregates in the cytoplasm and cannot induce polyubiquitination of PP2Ac. PP2Ac accumulates and microtubule associated proteins are hypophosphorylated (Trockenbacher et al., 2001).

MID1 and PP2Ac are both known regulators of mTOR and mTORC1 signalling (Figure 9) and dysregulation of mTOR might play role in the pathogenesis of OS (Liu et al., 2011). PP2Ac is a counterplayer of mTOR signalling by dephosphorylating the translation regulatory proteins 4EBP1 and p70S6K and thus inhibiting mTOR dependent translation. MID1 on the other hand is a positive regulator of mTOR signalling by inducing translation of the mTOR activator PDPK-1. Additionally, mTOR can also trigger the formation of the PP2Ac-α4-MID1 complex, thus inhibiting PP2Ac activity (Aranda-Orgille et al., 2011). In OS patients this fine-tuned regulation is disturbed. A truncated or mutated MID1 protein leads to a reduction of mTOR signalling and an impaired protein synthesis by accumulation of PP2Ac and reduced PDPK-1 translation.

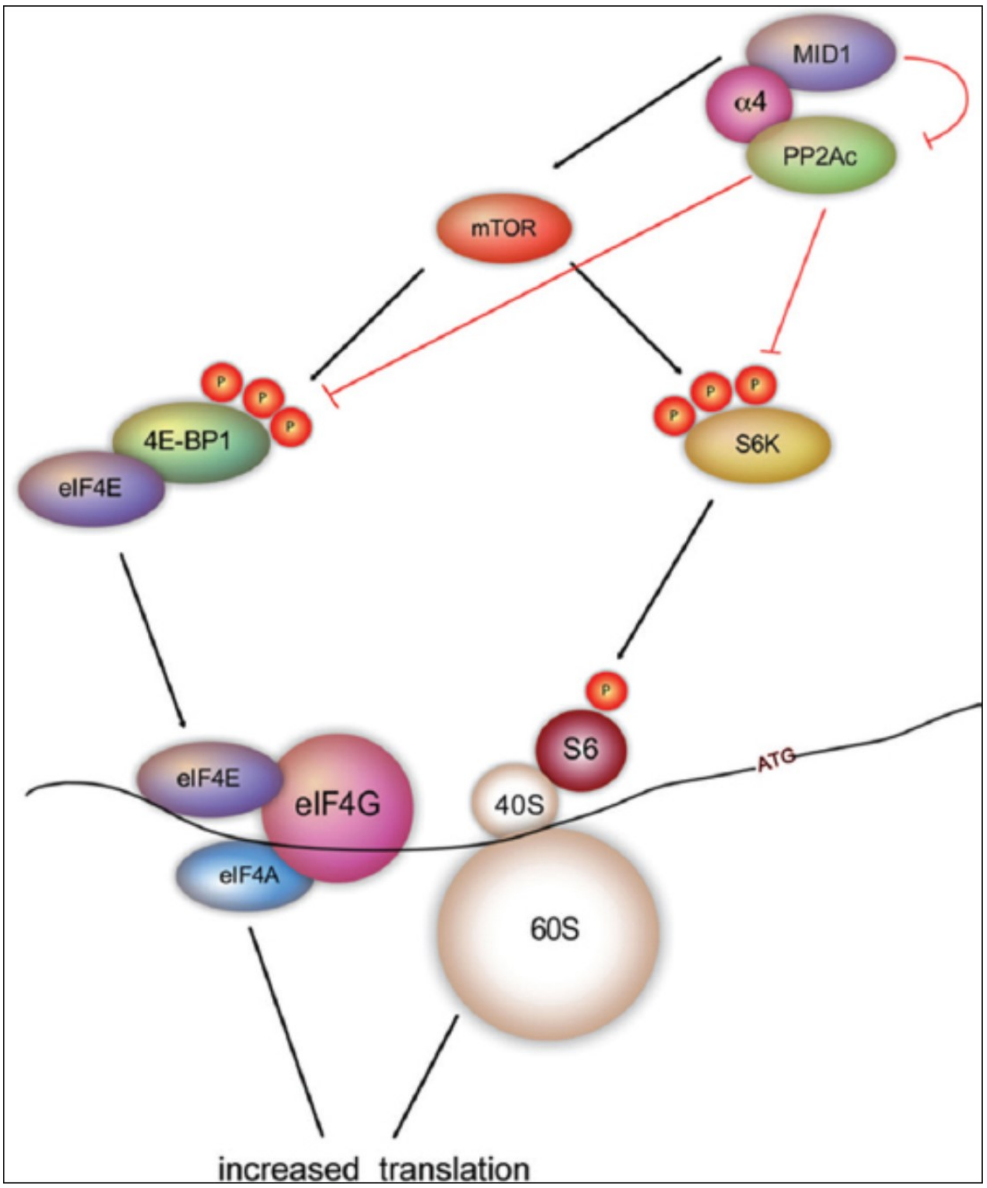


Figure 9 MID1 and mTOR signalling. MID1 is a positive regulator of mTOR signalling by activating mTOR dependent protein translation and inhibiting PP2Ac dependent inhibition. PP2Ac is a negative regulator of mTOR signalling by dephosphorylation of p70S6K and 4EBP1 (Winter et al., 2016).

1.3.4 MID1 and SHH

Sonic Hedgehog (SHH) together with Indian (IHH) and Desert Hedgehog (DHH) forms the mammalian Hedgehog signalling pathway that is highly conserved in all bilaterians and plays an important role during embryogenesis. There, it forms a concentration gradient important for midline, limb and neural tube formation. In the developed body, SHH influences proliferation of adult stem cells and has been found to be active in certain types of cancer. By binding to its receptor Patched1 (PTCH1), SHH activates Smoothened (SMO). SMO can then inhibit cleavage of GLI1-3 and thus activate expression of target genes (Figure 10). In the absence of

SHH, PTCH1 inhibits SMO. and GLI1-3 is cleaved, and target gene expression is inhibited (Murone et al., 1999).

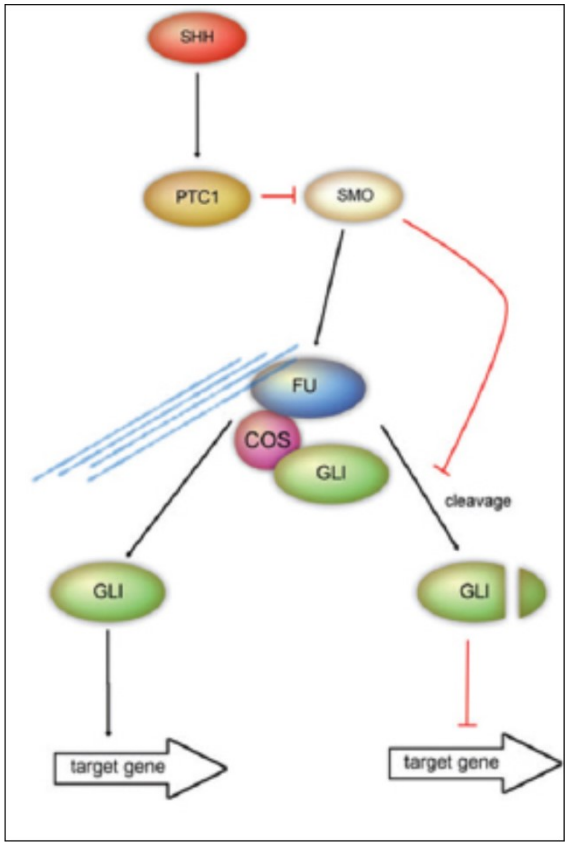


Figure 10 SHH signalling pathway. SHH binds to its receptor PTCH1 stopping the inhibition of SMO. SMO can then inhibit GLI cleavage and activate target gene expression. In the absence of SHH, SMO is inhibited by PTCH1 and GLI is cleaved, thus inhibiting target gene expression (Winter et al., 2016).

Two studies in 2008 and 2014 showed that MID1 and PP2Ac regulate nuclear localization of GLI3 probably via ubiquitination and cleavage of FU in cancer cell lines. FU is part of the SHH pathway and associates with microtubules. The MID1-PP2Ac protein complex polyubiquitinates FU through its E3 ubiquitin ligase activity, inducing a cleavage leading to a short and a long fragment of FU. The long fragment, lacking the N-terminal kinase domain, is then promoting formation of the full-length GLI3 protein, that translocates to the nucleus and induces target gene expression (Krauß et al., 2008; Schweiger et al., 2014).

1.3.5 MID1 and PAX6

PAX6 is a highly conserved key transcription factor during eye and central nervous system development (van Heyningen & Williamson, 2002). Mutations in PAX6 have been shown to cause malformations of the eye in Drosophila melanogaster as well as in the mouse and in humans (Washington et al., 2009). PAX6 is a known repressor of SHH expression in the

developing mouse brain, e.g. in the diencephalon, where loss of PAX6 causes patterning defects. A feedback loop for PAX6 and SHH has been suggested (Caballero et al., 2014).

A study in Xenopus showed that Mid1 regulates Pax6 posttranslational through polyubiquitination and subsequent proteasomal degradation. Mid1 expression is in this context driven by Shh to form a sharp boundary that is important for cell type definition and proper eye development. Mid1 defines the border between the eyestalk and the retina by clearing the Pax6 from the eyestalk region (Pfirrmann et al., 2016).

1.3.6 Mid1 in mouse

The mouse orthologue of MID1 is called Mid1, the protein has a length of 680 aa in comparison to the 667 aa of the human protein. The two proteins show an identity of over 90% when blasting the amino acid sequences. Unlike the human gene, Mid1 is located in the pseudoautosomal boundary (PAR) of the mouse X-chromosome and does not undergo X-inactivation. A Mid1-null mouse model was established in 2010 as a potential disease model for OS (Lancioni et al., 2010). The mouse model offers the opportunity to analyse Mid1 in the brain during different stages of development. The authors found agenesis or hypoplasia of the vermis; a symptom also present in some human OS patients. In concordance with that, they identified deficits in motor coordination as well as procedural and nonassociative learning. Interestingly, the mice did not show any further OS specific phenotype alterations. They have a normal life span, are fertile and show no signs of midline defects observed in human OS patients. The authors suggest that the Mid1 homologue Mid2 might rescue this phenotype in mice. Mid2 shares high similarity with Mid1 including the protein structure and its function as an E3 ligase (Buchner et al., 1999; Short et al., 2002). While both are ubiquitiously expressed in fetal tissue, Mid1 is generally higher expressed than Mid2. On the other hand, Mid2 shows high levels of expression in the heart, where no expression of Mid1 was found pointing towards a tissue-specific expression pattern of the two complementary partners. MID2 can also be found in humans and is located on Xq22. Another study using a different mouse model could confirm that mice carrying Mid1 mutations do not show any midline defects (Lu et al., 2013). Interestingly, they showed that knockout of Mid1 *in vitro* and *in vivo* promotes axon growth and branch formation. They concluded that this is caused by an accumulation of PP2Ac and an impaired phosphorylation of microtubule associated proteins.

In sum, observations from mice show that knock-out of Mid1 alone does only marginally resemble OS symptoms and is therefore a limited model for the disease.

1.4 X-Inactivation

In placental mammals, sex is defined by the gonosomes, while males have one X-chromosome and one Y-chromosome, females have two X-chromosomes. In humans, the Y-chromosome is a relatively small and gene-poor chromosome with an estimated number of ~60 protein coding genes (*Human Genome Assembly GRCh38 - Genome Reference Consortium*, n.d.; *Y[Chr] AND "Homo Sapiens"[Organism] - NCBI*, n.d.). The X-chromosome on the other hand contains ~800 protein coding genes and is about three times the size of the Y-chromosome (*Human Genome Assembly GRCh38 - Genome Reference Consortium*, n.d.; *X[Chr] AND "Homo Sapiens"[Organism] - NCBI*, n.d.). It contains no genes specifically driving female development. It is the pure absence of the Y-chromosome that drives female embryogenesis, as seen in Turner syndrome (Sybert & McCauley, 2004). To compensate for the double gene copy number in females, the cell inactivates one of the two X-chromosomes randomly during embryogenesis (Lyon, 1961).

1.4.1 During embryogenesis

After fusion of sperm and oocyte, the developing human female single cell embryo has two active X-chromosomes, a paternal and a maternal, that both gradually start to express XIST and XACT during the 4-cell stage (Figure 11) and stay active (Vallot et al., 2017). XIST is a long non coding RNA (lncRNA) that is transcribed from the X-chromosome and stays in the nucleus where it activates X-chromosomal inactivation (XCI) (Brown et al., 1991). XACT is also a lncRNA that is expressed from the X-chromosome and covers the active X-chromosome during early development (Vallot et al., 2013). The exact mechanism of XACT is unknown. This phase during early development where both X-chromosomes in female embryos are active although expressing XIST is called preXCI. After implantation of the embryo, random XCI is initiated. XACT expression is stopped completely and XIST expression becomes limited to one of the two X-chromosomes (Ballabio & Willard, 1992; Brown et al., 1991). Which of the two X-chromosomes keeps expressing XIST is a random mechanism in each cell, eventually leading to female mosaic. The XIST remains with the chromosome it is expressed from and induces full XCI by modifying chromosomal organization including Histone modifications and hypermethylation of CpG islands with the help of H3K27me3 (Nora & Heard, 2010). XCI finishes probably around day 30 of pregnancy, depending on the cell lineage and differentiation, but the exact timing remains unclear and is subject of current research (Patrat et al., 2020; Tang et al., 2015; Teklenburg et al., 2012). Once XCI has been established, it is maintained in the cell and all daughter cells with exception for cells of the germline. In primordial germ cells

(PGCs) the inactive X-chromosome is reactivated, starting probably already during early embryonic development and then continuing in an asynchronous manner (Briggs et al., 2015).

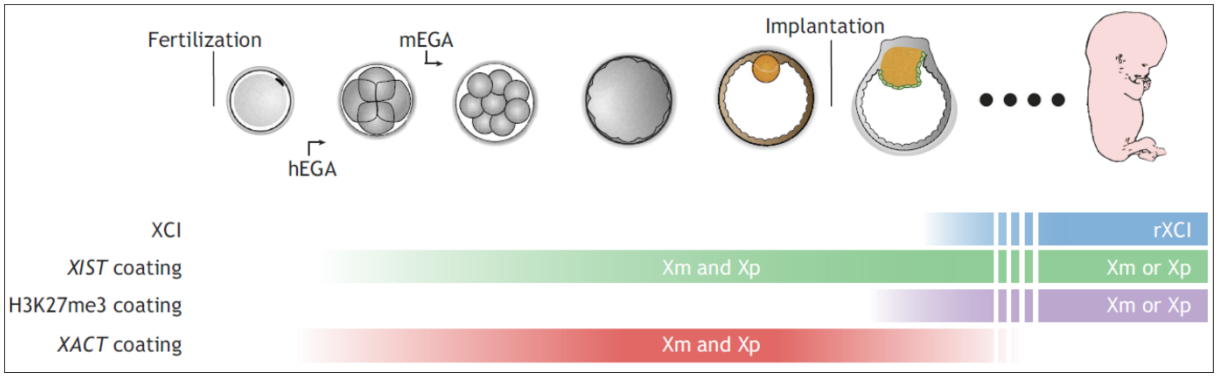


Figure 11 XCI during human embryogenesis. During early development both X-chromosomes express XIST and XACT without inactivating one or the other. After implantation of the embryo, one X-chromosome is randomly silenced by XIST and H3K27me3. (Patrat et al., 2020).

1.4.2 In iPSCs

While mouse iPSCs always show full reactivation of the before inactive X-chromosome (Payer & Lee, 2014), the situation in human iPSCs, but also ESCs, is more complex. Human pluripotent stem cells (PSCs) can be sorted in three different classes regarding their X-inactivation (Geens et al., 2017; Silva et al., 2008). Class I resembles that of mouse iPSCs with two active X-chromosomes (XaXa) and a random XCI during differentiation, resulting in a heterogeneous cell population. Class II PSCs have one active and one inactive X-chromosome (XaXi), show a high expression of XIST (from Xi) and the typical methylation pattern found after XCI. This state of XCI is kept after differentiation, resulting in a clonal cell population regarding X-chromosomal expression. Class III PSCs have no XIST expression and a low methylation level, but do not show a reactivation of both X-chromosomes. They have one active X-chromosome and one eroded X-chromosome (XaXe) expressing some but not all genes from both X-chromosomes. After differentiation this status of XaXe is either kept and even further progresses (Mekhoubad et al., 2012) or XaXi can be achieved independent of XIST (Bar et al., 2019).

In a study analysing over 700 PSCs (male and female, ESCs and iPSCs) via RNAseq (Bar et al., 2019), the authors found that most female iPSC can be assigned into class II, having one active and one inactive X-chromosome. On the other hand, female ESCs are mostly class III, with one active and one eroded X-chromosome, further emphasizing the differences between iPSCs and ESCs. With most of the iPSCs being class II, raises the question, if the initial XCI

from the somatic cell is kept, or if during the reprogramming, there is a timepoint where both X-chromosomes are active and then one is randomly inactivated. Interestingly, there are studies supporting both hypotheses (Barakat et al., 2015; K. Y. Kim et al., 2014; Tchieu et al., 2010). A possible explanation for this are different culturing and reprogramming conditions (Tomoda et al., 2012). Therefore, the state of XCI in iPSCs should always be considered when performing experiments.

1.4.3 Escape genes

Interestingly, not all genes of the inactivate X-chromosome are being silenced. Some keep being expressed from both X-chromosomes, the so-called escape genes (Lyon, 1961). The exact mechanism and the reason for genes escaping XCI is unknown, but there are studies trying to unveil these questions. By inserting a known escape gene at several positions on the Xi, it was shown that escape genes might have intrinsic factors allowing them to escape XCI (Li & Carrel, 2008). Other studies suggest that non-coding regions play a role in escaping XCI. While Alu-elements are enriched in regions escaping XCI, LINE-elements are enriched in regions subject to XCI (Bailey et al., 2000; Cotton, Chen, et al., 2013; Tannan et al., 2013). The evolutionary reasons for genes escaping XCI remain unknown, it is thought though, that a remaining homology to genes on the Y-chromosome, a female advantage or the pure escape of a flanking gene might play a role (Bellott et al., 2014; Horvath et al., 2013; Posynick & Brown, 2019).

Although there are differences and exceptions, there are some general statements that can be made about escape genes:

(a) While in the mouse, there are only ~3% of genes escaping XCI, it is estimated that in humans 15-25% of X-chromosomal genes escape XCI, either in all or in certain tissues (Balaton et al., 2015; Berletch et al., 2010).

(b) The bigger part of these genes is located on the p-arm, the evolutionary younger and less conserved arm of the X-chromosome (Carrel & Willard, 2005).

(c) The expression level of a gene from Xi is lower than that from Xa (Balaton & Brown, 2016; Tukiainen et al., 2017).

(d) Clinical and phenotypical diversity observed in patients with X-chromosomal monogenetic disorders might result from escapee expression (Berletch et al., 2010).

(e) A high heterogeneity of escape genes can be found in different cell types, individuals and experimental settings (Balaton et al., 2015; Balaton & Brown, 2016; Tukiainen et al., 2017).

(f) Only a few escape genes exhibit a consistent expression across a wide range of cell types and conditions (Balaton & Brown, 2016; Tukiainen et al., 2017).

1.4.4 Tissue specificity of escape genes

A study in 2017 analysed publicly available RNAseq data from the Genotype-Tissue Expression (GTEx) project for allele-specific expression of X-chromosomal genes in different human tissues (Tukiainen et al., 2017). They found ~15% of X-chromosomal genes to escape XCI in general and ~15% of genes to escape XCI in a tissue specific manner confirming previous studies of others (Figure 12) (Carrel & Willard, 2005; Cotton, Ge, et al., 2013).

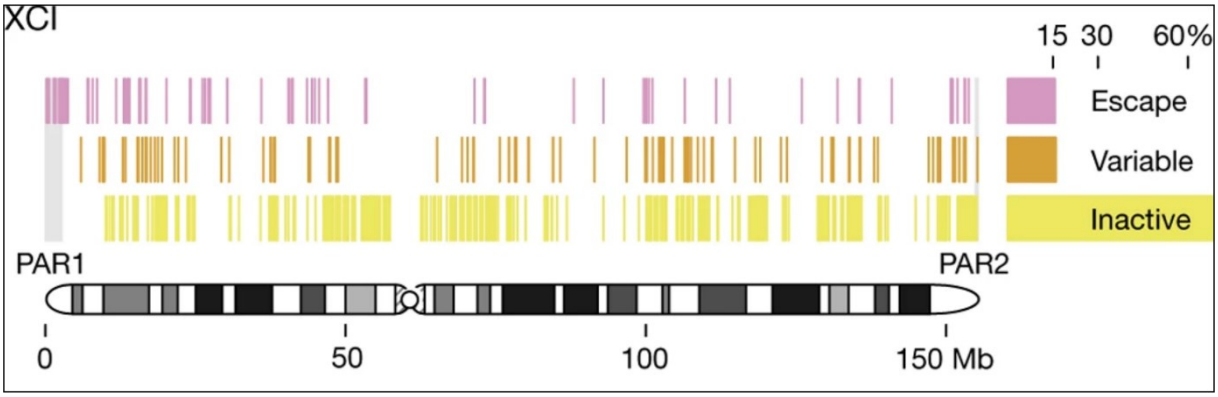


Figure 12 Genes escaping XCI. Pink lines mark genes generally escaping XCI. Orange lines mark genes escaping XCI in certain tissues. Yellow lines mark genes subject to XCI. All in their relative position on the X-chromosome (Tukiainen et al., 2017).

Further research on tissue specific escape from XCI is limited by the tissues available. Directed differentiation of iPSCs might be a suitable tool to overcome this limitation, especially for neuronal tissue that is hardly available.

1.5 Aim

Many studies have described OS in patients with all its phenotypical differences. However, the mechanisms leading to the development of OS pathology and underlying the clinical variability remain largely unknown. Molecular pathways of MID1 were analysed using cell lines or fibroblasts as a primary, patient-specific source. Also mouse models were generated, but reached their limitations in modelling OS probably due to the differences between man and mouse. A patient-specific cellular model based on state-of-the-art reprogramming methodology was supposed to close this gap and shed light into the mechanistic basis of OS.

To achieve this, a workflow including isolation of patient-specific fibroblasts, reprogramming of fibroblasts into iPSCs, generation of isogenic controls and differentiation into neurons was established.

In a second step, the generated cells (iPSCs, NPCs, neurons) and cerebral organoids (in collaboration with Prof. Dr. [redacted] and Prof. Dr. [redacted]) were analysed to identify OS-specific phenotypes and to characterize differences between N- and C-terminal MID1 mutations.

Finally, in clones of female cells expressing mutant MID1 either from the active or the inactive X-chromosome, X-inactivation pattern of MID1 and other X-chromosomal genes in iPSCs, NPCs and neurons was studied.

2 Material and Methods

2.1 Material

2.1.1 Equipment

Table 1 Equipment used during this work. Standard laboratory equipment (e.g. centrifuges) is not included.

Name	Manufacturer	Model
4D-Nucleofector	Lonza	AAF-1002B + AAF-1002X
BD FACSAria™III Cell Sorter	BD	
ChemiDoc™ MP Imaging System	Bio-Rad	170-8280
Nanodrop™ One^{C} Spectralphotometer	ThermoFisher Scientific	ND-ONEC-W
NextSeq	Illumina	NextSeq500
PyroMark Q96 instrument	Qiagen	9001525
Revolve Microscope	Echo	Revolve
StepOnePlus™ System	ThermoFisher Scientific	4376600

2.1.1 Chemicals and Media

Table 2 Chemicals used during this work.

Name	Manufacturer
Acrylamid	Carl Roth
Agar	Carl Roth
Agarose	AppliChem
Ampicillin	AppliChem
APS	Sigma
Boraic acid	Carl Roth
Bromphenol blue	Carl Roth
BSA	Carl Roth
$CaCl_2$	Carl Roth
EDTA	AppliChem
Ethidiumbromide	Carl Roth
Glycerin	Carl Roth
Glycin	Carl Roth
Hepes	Carl Roth
HEPES	Carl Roth
Isopropanol	Carl Roth
Kanamycin	AppliChem
Milk powder	Carl Roth
Na_2HPO_4	Carl Roth
NaCl	Carl Roth
NaOH	Carl Roth
PFA	Carl Roth
PhosStop	Roche
Protease inhibitor	Roche
Proteinase K	AppliChem
SDS	Carl Roth
Streptavidin-Sepharose beads	GE Healthcare
TEMED	Invitrogen
Tris	Carl Roth
Triton X 100	Carl Roth
Trytpone	Carl Roth
Tween20	Carl Roth
Urea	Carl Roth
Vectashield	Vector Laboratories
Yeast extract	Carl Roth
β-mercaptoethanol	Carl Roth

Table 3 Cell culture media and compounds used during this work. Special compositions are given in the method part.

Name	Manufacturer	Ordering Number
Accutase	ThermoFisher Scientific	A1110501
Advanced DMEM	Gibco LifeTechnologies	12634
B27+VitA-supplement	Gibco LifeTechnologies	17504001
B27-supplement	Gibco LifeTechnologies	12587010
bFGF	Gibco LifeTechnologies	PHG0264
CloneR™ 10X Cloning Supplement	Stemcell Technologies	5889
Collagenase IV	Gibco LifeTechnologies	17104-019
DMEM	Gibco LifeTechnologies	41966-029
DMEM with GlutaMAX™	Gibco LifeTechnologies	10569-010
DMEM/F-12 GlutaMAX™	Gibco LifeTechnologies	31331-028
DMSO	Carl Roth	175223617
ESC-qualified FBS	Gibco LifeTechnologies	16141-079
FBS	Gibco LifeTechnologies	10270-106
Gelatine	Sigma	61393-100ML
Geltrex	Gibco LifeTechnologies	A1413301
HBSS	Sigma	H6648-1l
IMDM	Gibco LifeTechnologies	12440-053
KOSR	Gibco LifeTechnologies	10828-028
Laminin	Sigma	L2020-1mg
Lipofectamine	Invitrogen	18324-012
Matrigel Matrix	Corning	354277
mTeSR™1	Stemcell Technologies	85851 + 05827
N2-supplement	Gibco LifeTechnologies	17502048
NEAA	Gibco LifeTechnologies	11140-050
Neural Induction Supplement	Gibco LifeTechnologies	A1647801
Neurobasal	Gibco LifeTechnologies	21103049
Opti-MEM	Gibco LifeTechnologies	51702490
PBS	Gibco LifeTechnologies	14190-094
Pen/Strep	Gibco LifeTechnologies	15140-122
Poly-Ornithine	Sigma	P3655-50mg
Rock Inhibitor	Stemcell Technologies	72302
RPMI 1640	Gibco LifeTechnologies	51254487
SB431542	Tocris	1614
TrypLE™Express (Trypsin)	Gibco LifeTechnologies	12604-013

2.1.2 Kits and Enzymes

Table 4 Kits used during this work.

Name	Manufacturer	Ordering number
CytoTune™-iPS 2.0 Sendai Reprogramming Kit	Thermofisher Scientific	A16517
EndoFree Plasmid Maxi Kit	QIAGEN	12362
High Pure PCR Product Purification Kit	Roche	11732676001
High Pure RNA Isolation Kit	Roche	11828665001
NucleoSpin™ Gel and PCR Clean-up	Macherey-Nagel	740609.50
Ovation® SoLo RNA-Seq Library Preparation Kit	NuGEN	0500-96
P3 Primary Cell 4D-Nucleofector™ X Kit L	Lonza	V4XP-3024
PSC Neural Induction Kit	Thermofisher Scientific	A1647801
PyroMark Gold Q96 Reagents	QIAGEN	972804
Western Lightning® Plus-ECL, Enhanced Chemiluminescence Substrate	PerkinElmer	NEL 103001EA

Table 5 Enzymes used during this work.

Name	Manufacturer	Ordering number
AflII	NEB	R0520S
EcoRI	NEB	R3101S
FastStart™ Taq DNA Polymerase	Roche	12032953001
Gibson Assembly® Master Mix	NEB	E2611S
Phusion Polymerase	NEB	M0530S
PrimeScript™ RT Master Mix	TaKaRa	RR036Q
SYBR® Premix Ex Taq™ II	TaKaRa	RR820L
XhoI	NEB	R0146S

2.1.3 Primers and antibodies

Table 6 Primers used during this work. All primers were ordered at Sigma and diluted 1:10 in water for PCR reactions, and 1:20 for sequencing reactions.

Name	Sequence
Cloning of MID1 constructs	random plasmid-specific MID1-specific
MID1_WT_for	ACTGCTCGAGCTCAAGCTTCGGAAACACTGGAGTCAGAACT
MID1_altATG1_for	ACTGCTCGAGCTCAAGCTTCGTCATCACCCTCAGCCAGC
MID1_altATG2_for	ACTGCTCGAGCTCAAGCTTCGACCTCCGCCGAGAAGGTC

MID1_altATG3_for	ACTGCTCGAGCTCAAGCTTCGTGCTTGGAGCATGAGGATG
MID1_altATG4_for	ACTGCTCGAGCTCAAGCTTCGTACTGTGTGACCGATGACC
MID1_WT_rev	ACTGGAATTCTCACGGCAGCTGCTCTGT
DNA	
MID1_Ex1_Surveyor_f	TGCCCTATTTGTCTGGAGCT
MID1_Ex1_Surveyor_r	GTCATGGTGTTGGCGTCAA
RNA	
MID1_Ex8_WT_f	AGTGGCCGGCATTATTGGGAAGTGGT
MID1_Ex9_WT_r	ATGCCCACGCGCCGGAGGT
MID1_Ex9_MUT_r	AGGATGCCCACGCGCCGT
qPCR	
ALDOA_f	TTAGAGAAGATCGGGGACACA
ALDOA_r	GGGCATGGTGCTGGTAGTAG
CDIPT_f	CAGCATGAGTTTGGATGTGG
CDIPT_r	CTGCCTCGGACCACAGAA
DOC2A_f	CACTCTGCACTGTAGCATCCTA
DOC2A_r	TCGTTTTTAGCTTATTGGCCTTA
GAPDH_f	CCACATCGCTCAGACACCAT
GAPDH_r	AAATCCGTTGACTCCGACCTT
KIF22_f	GAACCACCAGGAGACTCTCAA
KIF22_r	GCTGCACTGAACCTGCATAG
KLF4_f	CCCACATGAAGCGACTTCCC
KLF4_r	CAGGTCCAGGAGATCGTTGAA
MAP2_f	GGAGGTGTCTGCAAGGATAGT
MAP2_r	GGTGGAGAAGGAGGCAGATT
MAPK3_f	CCCTAGCCCAGACAGACATG
MAPK3_r	GCACAGTGTCCATTTTCTAACAGT
MID1_Ex1_f	TGTGTGACCGATGACCAGTT
MID1_Ex1-2_r	GTTTTGCTTCAATTTGTCATAGC
MID1_Ex6_f	ACCATATTCACCGGACAAGC
MID1_Ex7_r	GGTTCTGCTTGATGTTGGGTA
MID1_Ex8_f	CAGACACTTGTTCCACACGG
MID1_Ex9_r	GTCACACACCTGAACGCTTC
NANOG_f	AAGGTCCCGGTCAAGAAACAG
NANOG_r	CTTCTGCGTCACACCATTGC
NESTIN_f	CCAGATCGCTCAGGTCCTG
NESTIN_r	AGCTGAGGGAAGTCTTGGAG
OCT4_f	GTGTTCAGCCAAAAGACCATCT
OCT4_r	GGCCTGCATGAGGGTTTCT

PAX6_f	ACCCAAGAGCAAATTGAGGC
PAX6_r	CCATTTGGCCCTTCGATTAGA
POGZ_f	ACACTATCTGCCAGCACTGT
POGZ_r	GCTCACTTTCAAACGCCCAT
PTCH1_f	TCTGGGAAGGGGCGAAATTA
PTCH1_r	CTCCCAGCTGTCCACTTGAT
SEZ6L2_f	TCGTCATCCACTTCTTTGAGG
SEZ6L2_r	ACATAGCTCAGGCTGCGATG
SOX2_f	TGGACAGTTACGCGCACAT
SOX2_r	CGAGTAGGACATGCTGTAGGT
TAU_f	GTGCAAATAGTCTACAAACCAGT
TAU_r	CAATCTTCGACTGGACTCTGT
TUBB3_f	TCGGACTTGCAGCTGGAG
TUBB3_r	CAGGCCTGAAGAGATGTCCA
QUASEP	
CA5B_f	[Btn]TGTGGTAAGGCACAGACA
CA5B_r	ACACTGGGGCAAATATACTGAC
CA5B_s	GCCAGGAAGTACCAGT
MID1_f	AATAACTGGGTGGTGAGACACA
MID1_r	[Btn]AGGCGATAGAGCCGTTAT
MID1_s	GAGCCTGCCCCCCAC
ZNF185_f	AGGCTTATAATGGGCCAAGTTGA
ZNF185_r	[Btn]CAACACATGCCAACATACCTGTAA
ZNF185_s	AGTTTTTCATTTGGTCTTC
Sequencing	
EGFP-C-for	CATGGTCCTGCTGGAGTTCGTG
SV40pA-rev	GAAATTTGTGATGCTATTGC
U6-for	ACTATCATATGCTTACCGTAAC

Table 7 Primary antibodies used during this work.

Protein	Manufacturer	Ordering number
ACTIN	Sigma	A2066-200UL
MAP2	Sigma Aldrich	M4403
MID1 (commercial)	Novus	NBP1-26612
MID1 (homemade)	[redacted]	
NANOG	R&D Systems	AF1997
NESTIN	Merck	MAB5326
PAX6	BioLegend	901301
phosphoS6	Cell Signal	2215S
PP2Ac	Cell Signal	2038S
S6	Cell Signal	2317S
SERPINH1	Sigma-Aldrich	S5950
SOX2	Abcam	ab137385
TAU	Abcam	ab32057
TRA-1-60	Millipore	4360
TUBB3	Sigma Aldrich	T8660-100UL

Table 8 Secondary antibodies used during this work.

Name	Manufacturer	Ordering number
AffiniPure Goat anti-Mouse IgG	Jackson ImmunoResearch	115-035-003
AffiniPure Goat anti-Rabbit IgG	Jackson ImmunoResearch	111-005-003
Goat anti-mouse IgG, Alexa Fluor 488	Invitrogen	A11017
Goat anti-Rabbit IgG, Alexa Fluor 488	Invitrogen	A11008
Goat anti-Rabbit IgG, Alexa Fluor 596	Invitrogen	A11012
Rabbit anti-Goat IgG, Alexa Fluor 568	Invitrogen	A11079
Rabbit anti-Mouse IgG, Alexa Fluor 488	Invitrogen	A11059

2.1.4 Software

Table 9 Software used during this work.

Name	Manufacturer
BioEdit	mBio
Coral Draw	Corel Corporation
Crispr.mit.edu	Zhang Lab
Image Lab	Bio-Rad
Office	Microsoft
Prism	GraphPad
Pyro Q CpG	Qiagen

2.1.5 Plasmids and gRNA constructs

Table 10 Plasmids used during this work.

Name	Manufacturer	Ordering number
Cas9-GFP	[redacted]	
eGFP-C1-MID1-WT	Stephan Käseberg [redacted]	
eGFP-C1-MID1-altATG1		
eGFP-C1-MID1-altATG2		
eGFP-C1-MID1-altATG3		
eGFP-C1-MID1-altATG4		
eGFP-C1-MID1-del4	[redacted]	
gRNA cloning vector	Addgene	41824
pCMV delta R8.2	Addgene	12263
pCMV-VSV-G	Addgene	8454
pMXs-hc-MYC	Addgene	17220
pMXs-hOCT3/4	Addgene	17217
pMXs-hSOX2	Addgene	17218
pMXs-KLF4	Addgene	17219

Table 11 gRNA constructs used for cloning. Black: 5'-overhang required for cloning; red: gRNA sequence; purple: reverse complement sequence of gRNA.

Name	Sequence
MID1_sgRNA1_for	TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGTGAGCCCGTCTAGACCTCGC
MID1_sgRNA1_rev	GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGCGAGGTCTAGACGGGCTCAC
MID1_sgRNA2_for	TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGATGGACTCCACAGACTCGT
MID1_sgRNA2_rev	GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACACGAGTCTGTGGAGTCCATC

2.1.6 Cells

The main experiments of this work were performed with cells originating from three different fibroblast cell lines (Table 12). A male wildtype control cell line (1263/16), a female heterozygous carrier of a *MID1* mutation (16/98) and her son (17/98) carrying the same mutation hemizygously. All were reprogrammed into iPSCs and later differentiated into NPCs and neurons. In a side project, two more female fibroblast cell lines were reprogrammed into iPSCs and differentiated into NPCs and neurons. A female wildtype control cell line (1179/17) and her daughter carrying a microduplication (1180/17).

Table 12 List of cells used during this work.

Patient	Donor age	Cells	Name	Cell type	Sex	Mutation
16/98	~30 years	16/98	Het	Fibroblasts	Female	Heterozygous MID1 c.1801_1804delCTCC
		16/98 M7	OS M7	iPSCs		
		NPC 16/98 M7		NPCs		
		Neuron 16/98 M7		Neurons		
		16/98 M10	Ctrl M10	iPSCs		
		NPC 16/98 M10		NPCs		
		Neuron 16/98 M10		Neurons		
17/98	fetal	17/98	Male OS	Fibroblasts	Male	Hemizygous MID1 c.1801_1804delCTCC
		17/98 M14		iPSCs		
1263/16	26 years	1263/16	WT	Fibroblasts	Male	Wildtype
		1263/16 S1	C1	iPSCs		
		NPC 1263/16 S1		NPCs		
		Neuron 1263/16 S1		Neurons		
		1263/16 S1 MID1 1.7	R2	iPSCs		Hemizygous MID1 c.204_205delAG
		NPC 1263/16 S1 MID1 1.7		NPCs		
		Neuron 1263/16 S1 MID1 1.7		Neurons		

		1263/16 S1 MID1 2.7	R1	iPSCs		Hemizygous MID1 c.136_137insC
		NPC 1263/16 S1 MID1 2.7		NPCs		
		Neuron 1263/16 S1 MID1 2.7		Neurons		
1179/17	~60 years	1179/17	Mother	Fibroblasts	Female	Wildtype
		1179/17 A2/A27		iPSCs		
		NPC 1179/17 A2		NPCs		
		Neuron 1179/17 A2		Neurons		
1180/17	~30 years	1180/17	Daughter	Fibroblasts	Female	Mikroduplikation Chromosome 16p11.2
		1180/17 B8		iPSCs		
		NPC 1180/17 B8		NPCs		
		Neuron 1180/17 B8		Neurons		
		HEK T293	HEK	Human Embryonic Kidney cells	Female	
		HeLa	HeLa	Immortalized cervical cancer cells	Female	

2.2 Methods

All methods were performed following good laboratory practice. All cells were cultured with permission of the donors.

2.2.1 Bacterial Culture

Bacterial cells (*Escherichia coli*) were used to generate and amplify plasmids that were later used for transfection of eukaryotic cells.

2.2.1.1 Cloning of gRNA constructs

For CRISPR/Cas9 genome editing, gene specific gRNA constructs were needed. The 20 bp sequence was designed using an online tool (crispr.mit.edu) that detects all possible gRNAs within a given sequence and rates them regarding their predicted specificity and efficiency. The two highest rated gRNAs were chosen and ordered as oligos.

The first step of gRNA cloning was the annealing of the gRNA oligo with its reverse complement sequence and the extension of the 5' overhangs, to form a double stranded DNA. Therefor, 1 µL of each the forward and the reverse oligo were mixed with 4 µL Phusion buffer, 0,4 µL dNTPs, 13,4 µL water, and 0,2 µL Phusion polymerase (Table 5). The reaction was mixed and incubated in a thermal cycler as shown in Table 13. In parallel, the gRNA cloning vector was linearized by combining 1 µg of vector DNA with 5 µL of CutSmart® buffer, 1 µL AflII restriction enzyme, and water up to 50 µL (all from NEB). The reaction was incubated for 15 min at 37°C.

Table 13 Cycler program used for gRNA annealing and extension.

Temperature	Time
98°C	30 s
72°C	5 min
4°C	forever

After successful linearization and annealing, the double stranded gRNA oligo was cloned into the vector using the "Gibson Assembly MasterMix". 3 µL of vector were mixed with 7 µL of oligo, and 10 µL of the master mix. After incubating at 50°C for 15 min, the ligated vector was transformed into *E. coli* (2.2.1.3).

2.2.1.2 Cloning of MID1 overexpression constructs

MID1 overexpression constructs were used to show how different mutations influence the microtubule binding capacity of MID1. For better visualization, the eGFP-C1 plasmid was

chosen. This plasmid expresses a GFP N-terminally fused to the protein of interest. GFP-MID1 constructs were cloned together with [redacted].

In a first step, the MID1 specific primers (Table 6) were used to amplify the sequences starting at the different ATGs. Therefor, 4 µL of cDNA (5 ng/µL) were mixed with 2,5 µL buffer, 1 µL of each forward and reverse primer, 1 µL dNTPs, 0,5 µL polymerase and filled up with water to a total volume of 25 µL. A separate reaction was performed for each forward primer in combination with the reverse primer. The cDNA was generated as described in 2.2.3.2.2; For amplifying alternative ATG1, cDNA of R2 was used, while for all others, wildtype cDNA was used. The PCR was incubated as shown in Table 14.

Table 14 Cycler program used for MID1 amplification for cloning.

1. Initial denaturation	95°C	2 min
2. Denaturation	95°C	1 min
3. Primer annealing	60°C	30 s
4. Elongation	72°C	4 min
5. Loop	Go to step 2	34X
6. Final elongation	72°C	5 min
7. Store	4°C	Forever

After successful amplification, PCR products were cleaned by removing leftover PCR components by applying the "High Pure PCR Purification Kit" following the manufacturer's instructions.

In a next step, both, the PCR products and the already existent eGFP-C1-MID1-del4 plasmid were enzymatically digested for the following ligation. 30 µL of the purified PCR product were mixed with 4 µL of CutSmart® Buffer, 1 µL of EcoRI and 1 µL of XhoI (all from NEB). For plasmid digestion, 2 µL of the plasmid (1 µg/µL) were mixed with 2 µL of CutSmart® Buffer, 1 µL of EcoRI, 1 µL of XhoI and 14 µL of water. All digestions were incubated at 37°C. After 1 h, the PCR digestions were stored at 4°C, while the plasmid digestion was incubated for another hour at 37°C with an additional 2 µL of CIP to prevent religation.

After gel purification using the "NucleoSpin™ Gel and PCR Clean-up" Kit, the ligation reactions were mixed as shown in Table 15 and incubated at 16°C over night. The following day, ligations were used for transformation.

Table 15 Ligation reactions for cloning of the different GFP-MID1 constructs.

	Linearized Vector [ng]	Digested PCR [ng]	T4 buffer	T4 DNA ligase	H_2O
WT	100	127	2 µL	1 µL	To 20 µL
ATG1		114			
ATG2		106			
ATG3		94			
ATG4		92			

2.2.1.3 Transformation

Cloned constructs were transformed into chemically competent *E. coli* by heat shock. The plasmid was incubated for 30 min on ice after pipetting it into the competent bacteria solution. This was followed by a 90 s incubation at 42°C and a 2 min incubation on ice. Directly thereafter, 250 µL of LB-medium (10 g NaCl, 5 g yeast extract, 10 g tryptone, ad 1 l water) were added and the bacteria was put on a shaker for 30 min at 37°C. The bacterial suspension was then ready for plating on LB-agar plates (10 g NaCl, 5 g yeast extract, 10 g tryptone, 15 g agar, ad 1 l water) with ampicillin or kanamycin, depending on the plasmid. After an overnight incubation at 37°C single colonies were visible on the agar plates.

2.2.1.4 Mini-Prep

To isolate DNA from the bacteria the same buffers as for the Maxi-Prep (2.2.1.5) were used. Single colonies were picked after transformation and expanded in 3 mL LB-medium overnight at 37°C. The following day, 1 mL of the bacteria solution was transferred to a reaction tube and pelleted at 6000 rpm for 5 min. The pellet was then resuspended in 100 µL buffer P1 and the same volume of buffer P2 was added. After incubation for 5 min, the lysis was stopped by adding 100 µL prechilled buffer P3 and vortexing. After incubating for 5 min the solution was centrifuged at 13.500 rpm for 10 min. The supernatant, containing the DNA, was transferred to a fresh tube and 1 mL EtOH (100%) was added for precipitation. The DNA was pelleted at 13.500 rpm for 10 min and washed with 150 µL EtOH (70%) followed by another centrifugation for 5 min at 13.500 rpm. The pellet was then air-dried and finally resuspended in 50 µL water for further analysis. Concentration was measured using the "Nanodrop™ One Spectralphotometer".

2.2.1.5 Maxi-Prep

For Maxi-Prep, the remaining 2 mL of bacteria in LB-medium were further expanded in 200 mL of LB-medium overnight at 37°C. The DNA was isolated using the "EndoFree Plasmid Maxi Kit" following the manufacturer's instructions with small changes. The bacteria was pelleted for 15 min at 6000 g and 4°C. The pellet was resuspended in 10 mL buffer P1 and the same volume of buffer P2 was added. After incubating for 5 min, prechilled buffer P3 was added and everything was mixed by inverting 10 times. The reaction was poured into one of the supplied cartridges and incubated for 10 min before being filtered. The flow through was mixed with 2,5 mL of buffer ER and incubated for 30 min on ice. In the meantime, a supplied column was equilibrated using 10 mL buffer QBT. After the incubation, the solution was filtered using the column. The column filter was washed two times with 30 mL buffer QC, before the DNA was eluted by applying 15 mL buffer QN. The flow through, containing the DNA, was precipitated using 10,5 mL isopropanol (100%) and centrifuged for 1 h at 5000 rpm and 4°C. The pellet was washed with 5 mL EtOH (70%) and again centrifuged as before. Afterwards, the pellet was air-dried and resuspended in a suitable volume of buffer TE for further analysis and transfection into eukaryotic cells. Concentration was measured using the "Nanodrop™ One Spectralphotometer" and adjusted to 1 mg/mL.

2.2.1.6 Sequencing

Plasmids were sent for sequencing after Mini- or Maxi-Prep, to check for correct incorporation of the cloned fragment (gRNA or MID1), or to exclude mutations that sometimes occur during the procedures. 5 µL of the plasmid DNA (80-100 ng/µL) were mixed with 5 µL of a plasmid specific primer (Table 6). The sample was then sent to "Eurofins Genomics" for sequencing. Results were provided online and analysed using the "BioEdit" software.

2.2.2 Cell Culture

All eukaryotic cells were cultured at 37°C in a humidified incubator with 5% CO_2, if not stated otherwise. All procedures were performed under sterile conditions in a laminar flow working bench.

2.2.2.1 Coatings

Cells were routinely cultured on tissue culture graded plastic plates. Certain cell types or procedures required additional coating of the plates to promote adherence and proper growth.

2.2.2.1.1 Gelatine

Gelatine is mostly used for promoting adherence of fibroblasts during critical steps. For Gelatine coating, a 2% stock solution of gelatine was 1:20 diluted with prewarmed PBS and sterile filtered. 1 mL of the solution was pipetted per well of a six well plate and incubated for at least 1 h in a humidified incubator. Afterwards, plates were ready to use.

2.2.2.1.2 Matrigel

Matrigel is an extracellular matrix that was used for culturing stem cells. A manufacturer defined amount of a Matrigel stock solution (dilution factor; LOT specific) was pipetted into 50 mL ice cold DMEM/F-12, thoroughly mixed and incubated for 1 h at 4°C. One well of a six well plate was coated with 2 mL of the solution and was ready for use after at least 1 h incubation at 37°C in a humidified incubator.

2.2.2.1.3 Geltrex

Geltrex has a very similar composition compared to Matrigel and was used during the differentiation of iPSCs into NPCs. The Geltrex stock solution was diluted 1:1000 in ice cold DMEM/F-12, carefully mixed and directly used for coating. 1 mL of the diluted solution was pipetted per well of a six well plate. After incubating for at least 1 h at 37°C in a humidified incubator, plates were ready for use.

2.2.2.1.4 Poly-Ornithine/Laminin

Poly-Ornithine/Laminin coating was used for culturing neuronal stem cells and neurons. A Poly-Ornithine stock solution (500 µg/mL) was diluted (plastic: 1:50; glass: 1:5) in ice cold Borat-Buffer (150mM; pH 8,35). 1,5 mL were pipetted per well of a six well plate and incubated over night at 37°C in a humidified incubator. On day two, wells were washed three times with 2 mL HBSS followed by a three-hour incubation with a 1 µg/mL solution of laminin in HBSS at 4°C. Afterwards plates were frozen at -20°C for long term-storage of up to six months. At least 1 h prior to use, plates were taken out of the freezer and incubated at 37°C in a humidified incubator.

2.2.2.2 Pellets

All cells except neurons were harvested for further experiments 16-18 hours after a full media change during their exponential growth phase. Cells were washed with 2 mL PBS per well of a six well plate, scraped off with 1 mL of fresh PBS and centrifuged in a 15 mL tube for 4 min at 200g. The pellet was resuspended in 1 mL of fresh PBS, transferred to a 1,5 mL tube and

centrifuged at 4500 rpm for 5 min. The supernatant was discarded, and the pellet was either dry frozen at -80°C for DNA and RNA isolation or directly processed for protein isolation.

2.2.2.3 PFA-fixation

For immunofluorescent staining, cells were cultured on coverslips (coated or non-coated) for two to three days after seeding; except for neurons, which were directly differentiated on coverslips. When reaching the desired amount of confluency cells were washed with 1 mL PBS per well of a 12 well plate and fixed with 1 mL of PFA (4% in PBS) for 15-20 min. Cells were then washed three times with 1 mL PBS and either stored at 4°C with 2 mL of fresh PBS or directly used for further experiments.

2.2.2.4 Cell counting

For some experiments exact cell numbers had to be determined. For that, 10 µL of a single cell solution was pipetted on a counting chamber, and cells in all four corner squares were counted. The cell number per mL was calculated using the following formula:

$$cells\ per\ mL = cell\ count * 2500$$

2.2.2.5 HEK and HeLa cell lines

HEK 293T cells were used for production of the reprogramming retroviruses and HeLa cells were used for overexpression of GFP-MID1 constructs.

2.2.2.5.1 Culturing

HEK 293T cells were cultured in IMDM with 10% FBS and 1% Pen/Strep, while HeLa cells required DMEM with 10% FBS and 1% Pen/Strep. Media was only changed during splitting or when mentioned.

2.2.2.5.2 Splitting

Both cell lines were typically split every three to four days at a ratio of 1:10 or 1:20. Therefor, cells were washed with PBS followed by an incubation with trypsin for 5 min at 37°C. Cells were then pipetted off and the desired amount was transferred to a fresh cell culture flask with fresh media.

2.2.2.5.3 Freezing

For freezing, cells were treated as described in 2.2.2.5.2. After trypsin incubation, cells were instead transferred to a 15 mL tube and centrifuged for 4 min at 200 g. The pellet was resuspended in 1 mL of fresh culture media and cell number was determined (2.2.2.4). Cell number was adjusted to ~1 Million/mL and DMSO was added to a final concentration of 10%.

Cells were then transferred to freezing tubes and frozen at -80°C in insulated boxes ensuring a slow temperature drop. After 24 h, cells could be transferred to liquid nitrogen for long term storage.

2.2.2.5.4 Thawing

Frozen cells were thawed by heating them quickly at 37°C in a water bath or in the hand. Cells were transferred to a 15 mL tube containing fresh pre-warmed culture media. After centrifugation for 4 min at 200 g, the pellet was resuspended in fresh, pre-warmed culture media and the cells were transferred to a cell culture flask for culturing.

2.2.2.5.5 Retrovirus production (Lipofectamine transfection)

Lipofectamine was used for HEK cell transfection to generate the retroviruses for reprogramming. On day 1, 3-10 Million cells were seeded on six 15 cm dishes (one plate for KLF4, c-MYC and SOX2, and three plates for OCT4). The following day, cells were transfected with a combination of three different plasmids: 11,25 µg VSV-G, 11,25 µg CMV gag-pol and 7,5 µg of one of the four reprogramming factors (KLF4, c-MYC, SOX2 and OCT4). Therefor, the three plasmids were pipetted in 2 mL of OPTI-MEM and incubated for 5 min. In a second tube, 60 µl of Lipofectamine were mixed with another 2 mL of OPTI-MEM and also incubated for 5 min. Both solutions were then combined and incubated for 15-45 min. After supplying the cells with fresh media, each plate was transfected with 4 mL of the Lipofectamine-DNA solution. Five hours after transfection, the medium was discarded and replaced with fresh culture medium. Two days after transfection, the media -containing the retroviruses- of all plates was collected, mixed and sterile filtered (22 µm). The filtered solution was aliquoted and frozen until use at -80°C.

2.2.2.5.6 Calcium-phosphate transfection

Calcium phosphate was used for HeLa cell transfection with MID1 overexpression constructs. 10.000 cells were seeded per well of a 12 well plate containing a coverslip. 24 h after seeding, 3 µg of plasmid (Table 10) was combined with 10 µL of $CaCl_2$ (2,5 M) and filled up with water to a total volume of 80 µl. While being vortexed, 80 µL of 2x HEPES-Buffer (5,95 g HEPES, 8,18 g NaCl, 750 µL 1 M Na_2HPO_4; ad 500 mL water) was added to the solution. After incubating for 30-45 min the transfection solution was mixed by pipetting and completely used for transfecting a single well of the 12 well plate. Media was completely changed 24 h after transfection and cells were fixed using PFA (2.2.2.3) 48 h after transfection. Coverslips were transferred to slides prepared with 10 µL of mounting medium (Vectashield, 0,5% DAPI). Pictures were taken using the “Echo Revolve” microscope.

2.2.2.6 Fibroblasts

Primary fibroblasts of OS patients or a healthy male control were used during this work (Table 12). While OS patient fibroblasts were already isolated from skin biopsies over 20 years ago with an unknown method, control fibroblasts were isolated during this work.

2.2.2.6.1 Isolation

Fibroblasts were isolated from 4 mm skin punch biopsies following a published protocol (Vangipuram et al., 2013) with small changes. The biopsy was cut in 15-20 equally sized pieces, which were then distributed equally among all six wells of a gelatine coated (2.2.2.1.1) six well plate. Each well contained 800 µl of fibroblast extraction media (FEM: DMEM + 20% FBS + 1% Pen/strep). Cells were monitored daily; every other day 200 µL of FEM was added to compensate for evaporation. After one week, media volume was increased to 2 mL and was changed completely every other day. Fibroblasts migrated out of the biopsy pieces and were transferred to two T75 flasks after 3-4 weeks when reaching confluency using the standard splitting protocol with FEM instead of culture media (2.2.2.6.3). After 2-3 days fibroblasts reached again confluency and were transferred to three T175 flasks using the same method as above. When fibroblasts reached confluency in the T175 flasks, they were again detached using trypsin and cells were diluted to 1 Million per mL after counting (2.2.2.4). DMSO was added to a final concentration of 10% and 1 mL of the cell suspension was frozen per tube as described later (2.2.2.6.4).

2.2.2.6.2 Culturing

Fibroblasts were routinely cultured in IMDM containing 15% FBS and 1% Pen/Strep. As already mentioned, FEM was used during fibroblast isolation and was replaced by normal culture media after the first freezing.

2.2.2.6.3 Splitting

Fibroblasts were split when reaching 80-90% confluency at a ratio of 1:3, typically every three to four days. For splitting, cells were washed with PBS and incubated with trypsin for 5 min at 37°C. The reaction was stopped by adding fresh culture medium and cells were mixed by pipetting to ensure a single cell solution before pipetting them in a fresh flask.

2.2.2.6.4 Freezing

For freezing, fibroblasts were treated as described in 2.2.2.6.3. After stopping the reaction with fresh media cells were counted and diluted to 1 Million per mL. DMSO was added to a final

concentration of 10% and cells were frozen in aliquots of 1 mL in freezing tubes as described above (2.2.2.5.3)

2.2.2.6.5 Thawing

Fibroblasts were thawed using the same protocol as above (2.2.2.5.4), using fibroblast culture medium.

2.2.2.7 Induced Pluripotent Stem Cells (iPSCs)

iPSCs were generated by reprogramming fibroblasts via viral transfection of the four reprogramming factors (KLF4, c-MYC, OCT4 and SOX2). While control fibroblasts were reprogrammed as a part of this work using Sendaiviruses, the OS fibroblasts were already reprogrammed before this work started by Dr. Eva Weis using Retroviruses. Following, both methods used for reprogramming will be stated for completeness.

2.2.2.7.1 Reprogramming (Sendaiviral)

Male control fibroblasts were reprogrammed using commercially available Sendaiviruses that encode for the four reprogramming factors ("CytoTune™-iPS 2.0 Sendai Reprogramming Kit"). Fibroblasts were reprogrammed following the manufacturer's instructions with small changes using the Feeder-Dependent approach.

On day -2 fibroblasts were gradiently seeded as duplicates on a six well plate (10-35 x 10^4 cells/well) replacing the standard fibroblast medium by reprogramming fibroblast medium (RFM: DMEM, 10% ESC-qualified FBS, 1% NEAA, 0,1% β-mercaptoethanol). On day 0 sendaiviral transduction was performed; a well showing 30-60% confluency was chosen. Exact cell number was determined by harvesting one well (2.2.2.6.3; 2.2.2.4) and using the duplicate for reprogramming. The required virus amount was calculated using the given formula and is shown in Table 16:

$$Volume\ of\ Virus\ [\mu L] = \frac{MOI\ x\ number\ of\ cells}{titer\ of\ virus\ x\ 10^{-3}}$$

Table 16 Virus needed for reprogramming 220.000 male control fibroblasts (1263/16).

Virus	MOI	Titer	Volume [µL]
KOS	5	150.000.000	14,7
c-MYC	5	150.000.000	14,7
KLF4	3	150.000.000	8,8

The calculated volumes were all added to 1 mL of RFM and within 5 min the media was pipetted on the chosen well, replacing the old medium. 24 h after transduction, the media was changed completely and replaced by fresh RFM. Cells were then fed every other day for the next six days. On day 5 feeder cells (MEFs) were prepared for the next step of the reprogramming. MEFs were thawed following the protocol previously described (2.2.2.5.4) using RFM. Cells were counted (2.2.2.4) and 2,5 x 10^5 cells were seeded per well of a six well plate; using a full six well plate for each fibroblast cell line that was reprogrammed. On day 7 after transduction, fibroblasts were ready to be transferred onto the MEFs. Fibroblasts were harvested (2.2.2.6.3), counted (2.2.2.4) and seeded gradiently (5, 10, 20, 40, 80 x 10^3 and rest) onto the prepared MEFs using RFM. On day 8 media was changed to iPSC medium (DMEM/F-12, 20% KOSR, 1% NEAA, 0,1% β-mercaptoethanol, 1% Pen/Strep, 0,04% bFGF) and from then replaced every day, increasing the volume over time to compensate for cell growth. About 21-28 days after transduction, iPSC colonies reached an appropriate size and were clearly visible. Approximately 50 colonies were picked by scraping and sucking with a 200 µL pipette under a microscope placed inside the working bench. Each colony was transferred to a single well of a Matrigel coated (2.2.2.1.2) 12 well plate with mTeSR1. Medium was changed every other day and colonies were monitored regularly. After a maximum of 10 days well grown colonies were split on a single well of a Matrigel coated (2.2.2.1.2) six well plate and were then treated as regular iPSCs.

2.2.2.7.2 Reprogramming (Retroviral)

OS patient fibroblasts were reprogrammed using self-made retroviruses (2.2.2.5.5) by Dr. Eva Weis in collaboration with Prof. Dr. Beate Winner in Erlangen. Fibroblasts were transduced by Spinfection.

On day -1 100.000 fibroblasts (2.2.2.6.3; 2.2.2.4) were seeded per well of a Gelatine coated (2.2.2.1.1) six well plate using standard fibroblast medium without Pen/Strep. On day 0 spinfection of the fibroblasts started. In the morning, an appropriate amount of the previously produced retrovirus solution (2.2.2.5.5) was supplemented with a final concentration of 8 µg/mL Polybrene and 5 mL were pipetted on each well to be reprogrammed. Cells were then centrifuged for 1 h at 800 g. The retrovirus solution was replaced by fresh media afterwards. This spinfection was repeated in the evening of the same day and the morning of the next day, each time using a fresh aliquot of the retrovirus solution with Polybrene. On day 2 (24 h after last spinfection), cells were harvested (2.2.2.6.3) and seeded on previously prepared 10 cm MEF plates; one well per plate. MEFs were prepared as described above (2.2.2.7.1) using

standard fibroblast medium. On day 3 medium was changed to iPSC medium adding 0,1% SB431542 additionally. Medium was changed every other day and cells were monitored regularly. All following steps were equal to the Sendaiviral reprogramming protocol (2.2.2.7.1).

2.2.2.7.3 Culturing

iPSCs were routinely cultured in mTeSR1 with Pen/Strep on Matrigel coated (2.2.2.1.2) six well plates. Medium was changed every day.

2.2.2.7.4 Splitting

iPSCs were typically split every 4-6 days using PBS/EDTA (500 mL PBS, 0,9 g NaCl, 250 µL 1 M EDTA) a self-made, enzyme-free splitting reagent at a ratio of 1:6 to 1:12. Cells were washed twice with 1 mL of PBS/EDTA and then incubated with 1 mL for 2 min. The PBS/EDTA was removed and 1 mL of mTeSR1 was added. Cells were scraped and resuspended with ~2,5 mL of additional mTeSR1 to break up big clumps and get colonies of about 50-500 cells. The desired amount of cell solution was then transferred to a fresh Matrigel coated (2.2.2.1.2) six well plate and incubated for 2 days before changing the medium.

For certain steps, a single cell solution of iPSCs was required. Therefor, iPSCs were washed with PBS and incubated with 1 mL trypsin for 5-6 min. ~2,5 mL PBS was added to the well and resuspended harshly to achieve a single cell solution. Cells were counted and the desired cell number (2.2.2.4) was pelleted at 200 g for 5 min for the following experiment.

2.2.2.7.5 Picking

As already described above, picking single colonies is a technique required during reprogramming, but also during standard iPSC culture and differentiation. A single iPSC colony was scraped and sucked using a 200 µL pipette and transferred to a new well of a Matrigel coated (2.2.2.1.2) plate.

2.2.2.7.6 Freezing

For freezing, iPSCs were treated with PBS/EDTA as described (2.2.2.7.4). Instead of using mTeSR1, the reaction was stopped using DMEM/F-12. After scraping, cells were pelleted at 200 g for 4 min. The pellet was resuspended in iPS-freezing medium (mTeSR1, 20% KOSR, 10% DMSO) until big clumps were deleted. The solution was transferred to freezing tubes and frozen at -80°C as described (2.2.2.5.3).

2.2.2.7.7 Thawing

iPSCs were thawed as described above (2.2.2.5.4). After thawing, cells were transferred to a tube containing DMEM/F-12, and the pelleted cells were resuspended in mTeSR1 before seeding them on a Matrigel coated plate (2.2.2.1.2).

2.2.2.7.8 Cleaning (removing differentiated cells)

During standard culture, but especially after reprogramming or electroporation, iPSCs sometimes differentiated spontaneously. As differentiated cells usually grow faster than stem cells, they would have soon overgrown the iPSCs in the same well. Therefore, spontaneously differentiated cells were removed manually whenever they occurred.

With a deformed Pasteur pipette, differentiated colonies were removed while monitoring the well through a microscope placed inside the working bench. After removing all visible differentiated colonies, the cells were washed with DMEM/F-12 and fed with mTeSR1.

2.2.2.7.9 Genome editing

For CRISPR/Cas9 genome editing iPSCs were electroporated using the Lonza 4D-Nucleofector™ X Unit. After single cell solution was achieved, 800.000 cells per transfection reaction were pelleted (2.2.2.7.4). Cells were resuspended in 100 µL electroporation buffer P3 immediately adding 2,5 µg of each the Cas9 and gRNA plasmid (Figure 13). Cell-plasmid solution was transferred to the electroporation cuvette and electroporated using the program CB-150 of the Nucleofector. After electroporation 100 µL of RPMI was pipetted into the cuvette and incubated for 10 min at 37°C. As a last step, the whole volume was transferred to a fresh well of a Matrigel coated (2.2.2.1.2) six well plate with 2 mL of mTeSR1 supplemented with ROCK-inhibitor (10 µM).

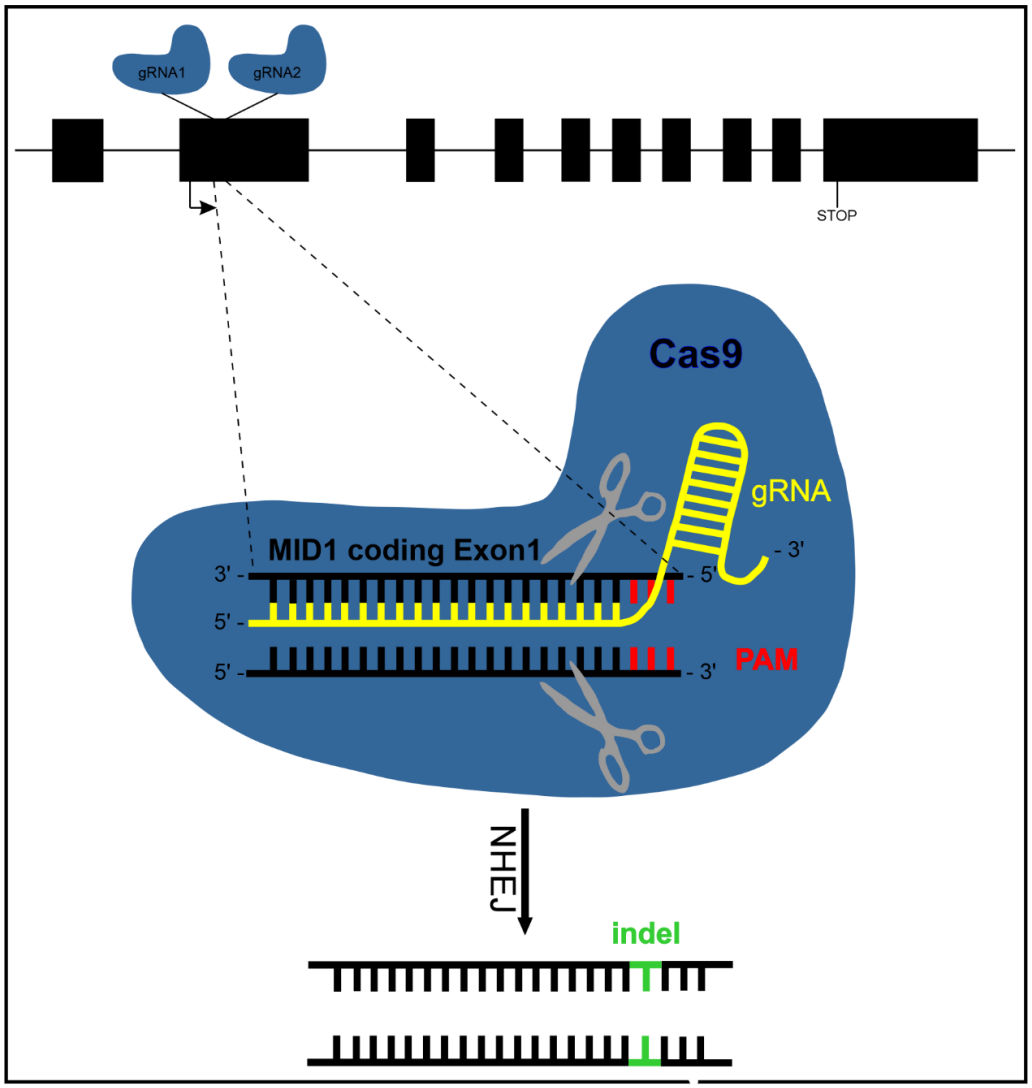


Figure 13 CRISPR strategy to induce frameshift mutations in the first coding exon of MID1. iPSCs were electroporated with a plasmid encoding the Cas9 protein and a plasmid encoding either gRNA1 or gRNA2. The induced DSB was sought to be repaired by NHEJ.

On day 1 after the electroporation, media was changed completely to mTeSR1 without ROCK-inhibitor. On day 2 after the electroporation GFP-positive cells were FACS-sorted as single cells into a Matrigel coated (2.2.2.1.2) 96 well plate (mTeSR1 + 10% CloneR). After obtaining a single cell solution (2.2.2.7.4), iPSCs were resuspended in 300 µL mTeSR1 supplemented with 10% CloneR and sorted at the FLOW Cytometry Core Facility by Alexander Hohberger using a "FACSAria" flow cytometer. Remaining cells, after sorting one cell per well, were discarded. 48 h after electroporation media was completely changed and replaced by 100 µL of mTeSR1 supplemented with 10% CloneR. After 24 h another 25 µL of mTeSR1 + 10% CloneR was added to each well. On day 4 after sorting, media was completely replaced by mTeSR1 and cells were fed every other day. Cells were monitored and first colonies were usually visible

after seven days. When covering at least 50% of the well surface (day 10-14), one well of the 96 well plate was splitted into one well of a Matrigel coated (2.2.2.1.2) 12 well plate using PBS/EDTA (2.2.2.7.4). From here cells were expanded following the standard iPSC culture.

2.2.2.8 Neuronal Precursor Cells (NPCs)

iPSCs were differentiated into NPCs using two different methods. The differentiation with selfmade buffers was first established but showed to be very inconsistent and error prone later. The commercial Kit for differentiation gave more consistent results and was quicker.

2.2.2.8.1 Differentiation (selfmade buffers)

The first step of differentiating iPSC into NPCs was the generation of Embryoid Bodies (EBs). iPSCs were washed with DMEM/F-12 and incubated with Collagenase IV for 20 min, followed by three wash steps with DMEM/F-12. After adding 1 mL mTeSR1, cells were scraped and resuspended to break up cell clumps. The cell solution was transferred to one well of an Ultra-Low attachment plate and incubated for 24 h. Medium was then changed to Neuronal Medium (NM: DMEM/F-12, 1% N2-supplement, 2% B27-supplement, 1% Pen/Strep) with full media changes every other day. On day 7 after starting the differentiation, EBs were seeded on one well of a Poly-Ornithine/Laminin coated six well plate (2.2.2.1.4). Media was changed every other day. This step was the generation of Neuronal Rosettes (NRs) from EBs. On day 7 after starting NR formation rosette like structures were visible inside the attached EBs and were picked (2.2.2.7.5) for establishing a proliferative NPC line from NRs. Picked EBs were transferred to a fresh well of a Poly-Ornithine/Laminin coated six well plate (2.2.2.1.4). On day 2 after picking, media was completely changed and from here in addition supplemented with 0,1% FGF2 (NM+FGF2). On day 3-4 after picking, cells were splitted using trypsin (2.2.2.8.4) and a proliferative NPC line was established (2.2.2.8.3).

2.2.2.8.2 Differentiation (commercial Kit)

As a second option, iPSCs were differentiated into NPCs using the "PSC Neural Induction Kit". On day 0 iPSCs were gradiently split (200, 300, 400, 500, 750 x10^3 and rest) using PBS/EDTA (2.2.2.7.4). After 24 h one well with a confluency of 15-25% was chosen for differentiation. Media was changed to Neural Induction Medium (NIM: Neurobasal, 2% Neural Induction Supplement, 1% Pen/Strep), renewed with increasing volumes every other day to compensate for cell growth. On day 7 of neural induction, cells were confluent and ready to be used for establishing a proliferative NPC line. Cells were washed with PBS and detached with Accutase for 6 min at 37°C. After adding PBS cells were resuspended to break up cell clumps and filtered through a 100 µm strainer. Cells were then pelleted at 300 g for 4 min and washed with PBS

before centrifuging again with the same conditions. The pellet was then resuspended in Neural Expansion Medium (NEM: 49% Neurobasal, 49% Advanced DMEM, 2% Neural Induction Supplement, 1% Pen/Strep) supplemented with ROCK-inhibitor (5 µM) and seeded (~500x10^3 cells per well) on a Geltrex coated (2.2.2.1.3) six well plate. After 24 h, medium was replaced by NEM and cells were monitored daily with full media changes every other day. When reaching confluency cells were split using trypsin (2.2.2.8.4) and were then cultured as a proliferative NPC line (2.2.2.8.3).

2.2.2.8.3 Culturing

NPCs were routinely cultured on Poly-Ornithine/Laminin coated six well plates (2.2.2.1.4) in NM+FGF2 with full media changes every other day.

2.2.2.8.4 Splitting

NPCs were split when reaching confluency, typically after 5-7 days, using trypsin. Cells were washed with PBS and incubated with trypsin for 5 min at 37°C. To stop the reaction, NM supplemented with 15% KOSR (NM+KOSR) was added to the cells. NPCs were thoroughly resuspended and centrifuged for 5 min at 200 g. The pellet was resuspended in NM+FGF2 and the cells were counted (2.2.2.4). ~500.000 cells were seeded per well of a six well plate.

2.2.2.8.5 Freezing

For freezing, NPCs were treated as described (2.2.2.8.4); NPC freezing medium (NM, 10% KOSR, 10% DMSO) was used for resuspending the pellet after centrifugation. Cells were counted (2.2.2.4), and 1 mL containing ~1 Million cells was frozen per freezing tube at -80°C as previously described (2.2.2.5.3).

2.2.2.8.6 Thawing

NPCs were thawed as described above (2.2.2.5.4). After thawing, cells were transferred to a tube containing NM+KOSR, and the pelleted cells were resuspended in NM+FGF2 before seeding them on a Poly-Ornithine/Laminin coated six well plate (2.2.2.1.4).

2.2.2.9 Neurons

Neurons were generated by differentiation of NPCs.

2.2.2.9.1 Differentiation and Culturing

For differentiation, NPCs were harvested as described (2.2.2.8.4) and seeded at very low cell numbers. Typically, 50-100 thousand cells were seeded per well of a Poly-Ornithine/Laminin coated six well plate (2.2.2.1.4). 24 h after seeding, cells were washed twice with PBS and media was replaced by NM+VitA (DMEM/F-12, 1% N2-supplement, 2% B27+VitA-

supplement, 1% Pen/Strep). From this point on, the CO_2 level in the incubator was increased to 8%. Every 3-4 days, fresh media was added to the cells without removing the old. When reaching a total volume of 10-12 mL per well, 20-50% of the media was removed every time before adding fresh media. The amount of media exchanged was depending on how much of the media was used (yellow colour). Using this feeding scheme, NPCs were differentiated into neurons for up to 6 weeks.

2.2.2.10 Brain Organoids

Brain organoids (cerebral organoids) were generated in collaboration with Prof. Dr. [redacted] and Prof. Dr. [redacted] in their lab in Munich. The lab work was done by Dr. [redacted] and [redacted]. They generated brain organoids from the iPSCs shown in this work and followed the protocol published by Lancaster et al. (Lancaster & Knoblich, 2014) with small changes. Some of the results from this work will be shown to support the results of the present study, but the protocol will not be given in detail.

2.2.3 Molecular Methods

As already described, harvested cells were pelleted (2.2.2.2) and used for DNA, RNA and Protein isolation.

2.2.3.1 DNA Analysis

DNA analysis was performed to confirm DNA stability after experiments known to be prone for DNA mutations, and to identify newly induced mutations after CRISPR/Cas9 genome editing.

2.2.3.1.1 Karyotyping

Karyotyping was performed to exclude chromosomal aberrations in iPSCs after reprogramming and CRISPR/Cas9 genome editing. For this, living cells were needed instead of already pelleted cells. Karyotyping was performed in the diagnostics' lab of the "Institut für Humangenetik" of the University Medical Center Mainz by technicians (mainly Denise Seyler). Therefore, the protocol is not given in full detail.

iPSCs grown under standard conditions were blocked from exiting the metaphase stage of mitosis by adding colcemid to the cells. After an incubation, cells were further processed in several steps, finally resulting in microscopic slides with metaphase chromosomes ready for analysis.

2.2.3.1.2 DNA isolation

DNA was isolated from cell pellets (2.2.2.2) by using self-made buffers. The pellet was resuspended in 200 µL SE-Buffer (0,439 g NaCl, 0,841 g Na-EDTA ad 100 mL water) before 20 µL of SDS (10%) and 20 µL of Proteinase K (15 mg/mL) were added. After incubating at 37°C over night, 80 µL of 6 M NaCl were added and mixed by vortexing for 15 s. The suspension was centrifuged at 6000 rpm for 15 min and the supernatant was transferred to a fresh tube, while the pellet was discarded. For precipitating the DNA, 600 µL of ethanol (100%) was added and the tube was centrifuged at 14.000 rpm for 20 min. The resulting pellet was washed twice with 1000 µL ethanol (70%), each wash followed by a centrifugation of 5 min at 14.000 rpm. After air drying, the pellet was dissolved in 20 µL H_2O. The DNA concentration was measured using the "Nanodrop™ One Spectralphotometer". Concentration was adjusted to 25 ng/µL

2.2.3.1.3 PCR amplification

Isolated DNA was amplified using a standard PCR protocol. 50-100 ng of DNA were mixed with 2,5 µL buffer, 1 µL of each primer (Table 6), 0,5 µL dNTPs, 0,2 µL polymerase and filled up with water to a total volume of 25 µL. The DNA was amplified using a cycler program dependent on the primer and the expected product length (Table 17). Successful amplification was checked by separating 5 µL of the PCR on a 1,5% agarose gel with EtBr.

Table 17 Standard Cycler program used for DNA amplification. X stands for the primer specific melting temperature and Y for the product length specific elongation time.

1. Initial denaturation	95°C	2 min
2. Denaturation	95°C	30 s
3. Primer annealing	X°C	30 s
4. Elongation	72°C	Y s
5. Loop	Go to step 2	34X
6. Final elongation	72°C	5 min
7. Store	4°C	Forever

2.2.3.1.4 Sequencing

Amplified DNA was further analysed by sequencing after removing leftover PCR components by applying the "High Pure PCR Purification Kit" following the manufacturer's instructions. Purified DNA was then prepared for sequencing by combining 5 µL of the PCR product with 5 µL of the primer. The sample was then sent to "Eurofins Genomics" for sequencing. Results were provided online and analysed using the "BioEdit" software.

2.2.3.1.5 Whole Exome Sequencing (WES)

Whole Exome Sequencing was performed to exclude off-target mutations after CRISPR/Cas9 genome editing. This method was performed in the diagnostics lab of the "Institut für Humangenetik" of the University Medical Center Mainz by technicians. Therefore, the protocol is not given in full detail.

After library preparation, DNA was sequenced using the Illumina NEXTSeq 550. Data was processed and analysed by [redacted] and [redacted].

2.2.3.2 RNA Analysis

RNA analysis was performed to identify expression differences between cells in a qualitative and quantitative way.

2.2.3.2.1 RNA isolation

RNA was isolated using the "High Pure RNA Isolation Kit" following the manufacturer's instructions for "Isolation of Total RNA from Cultured Cells". The pellet was resuspended in 200 µL PBS and 400 µL of lysis/binding buffer was added. After vortexing the tube for 15 s, the solution was filtered through one of the supplied filter tubes at 8000 g for 30 s. The flowthrough was discarded and the filter was incubated for 15 min at RT with 10 µL of DNAse I diluted in 90 µL of DNAse incubation buffer. The filter was then washed using 500 µL of wash buffer 1 and centrifuged at 8000 g for 30 s. The same step was repeated using wash buffer 2. In a last wash step 200 µL of wash buffer 2 were pipetted on the filter and centrifuged for 2 min at 13.000 g. The filter was then transferred to a fresh tube and the RNA was eluted using 30 µL elution buffer. The RNA concentration was measured using the "Nanodrop™ One Spectralphotometer".

2.2.3.2.2 cDNA synthesis

For further analysis, RNA was transcribed in cDNA using the "PrimeScript™ RT Master Mix". 500 ng of RNA were filled up with water to a final volume of 8 µL and mixed with 2 µL of the master mix. The solution was then incubated as shown in Table 18, and afterwards diluted with water to a final concentration of 5 ng/µL.

Table 18 Cycler program used for cDNA synthesis.

Temperature	Time
37°C	15 min
85°C	5 s
4°C	forever

2.2.3.2.3 allele-specific RT PCR

The allele-specific RT PCR was used to identify female iPSC clones that exclusively express the wildtype or the mutant (4 bp deletion) MID1, and to check the stability of the expression during differentiation. For each cDNA sample (2.2.3.2.2), two reactions were performed, one with the wildtype reverse primer, and one with the mutant reverse primer (Table 6).

4 µL of diluted cDNA were mixed with 2,5 µL buffer, 5 µL GC-rich, 1 µL of the forward primer and 1 µL of the allele-specific reverse primer, 0,5 µL dNTPs, 10,8 µL water, and 0,2 µL polymerase. The cDNA was amplified using the cycler program depicted in Table 19. The PCR was analysed by separation on a 1,5% agarose gel with EtBr.

Table 19 Cycler program used for the allele-specific RT-PCR.

1. Initial denaturation	95°C	2 min
2. Denaturation	95°C	30 s
3. Primer annealing	70°C	30 s
4. Elongation	72°C	40 s
5. Loop	Go to step 2	34X
6. Final elongation	72°C	5 min
7. Store	4°C	Forever

2.2.3.2.4 Quantification of Allele-specific expression by Pyrosequencing (QUASEP)

QUASEP-assays were performed to quantitatively measure the allele-specific expression of genes. In a first step, the previously generated cDNA (2.2.3.2.2) was amplified by PCR, with one of the primers being biotinylated (Table 6). For a single reaction, 8 µL of cDNA was mixed with 5 µL of Buffer, 1 µL of dNTPs, 2,5 µL of each primer, 30,6 µL of water, and 0,4 µL of polymerase and amplified in a PCR cycler (Table 20). Each sample was measured as a triplicate to correct for technical errors.

Table 20 Cycler program used for amplifying QUASEP-PCRs.

1. Initial denaturation	95°C	2 min
2. Denaturation	95°C	30 s
3. Primer annealing	60°C	30 s
4. Elongation	72°C	40 s
5. Loop	Go to step 2	34X
6. Final elongation	72°C	5 min
7. Store	4°C	Forever

Successful amplification was checked by separating 5 µL of the PCR reaction on a 1,5% agarose gel with EtBr. The remaining volume was then used for the Pyrosequencing following the standard protocol recommended when using the "PyroMark Gold Q96 Reagents". The remaining 45 µL of the PCR reaction were mixed with 20 µL binding buffer and 4 µL Streptavidin-Sepharose beads in one well of a 96 well PCR plate and vortexed until needed. In a pyrosequencing plate, 2 µL of the sequencing primer (Table 6) was mixed with 40 µL of annealing buffer. Using the "Vacuum Prep Table" of the pyrosequencer, the sample mixture of the PCR was aspirated and successively incubated for 5 s each in 70% ethanol, denaturation buffer (0,2 M NaOH), and wash buffer (10 mM Tris-acetate, pH 7,4). The now purified single-stranded PCR product was released into the pyrosequencing plate and incubated for 2 min at 80°C. In the meantime, a sequencing cartridge was prepared with all four nucleotides, the enzyme, and the substrate mix for sequencing, with the volumes calculated by the "Pyro Q CpG" software based on the sequence to analyse. The pyrosequencing plate and the cartridge were then placed in the sequencer and the run was started. The results were analysed using the same software as above and displayed as percentage values.

2.2.3.2.5 RT-qPCR

RT-qPCR is a method to quantitatively measure DNA amplification in real time while performing a PCR. 4 µL of the previously generated cDNA (2.2.3.2.2) was mixed with 7,5 µL of "TB Green® Premix Ex Taq^{TM} II", 2,5 µL of water, and 1 µL of a gene specific primer pair (Table 6) per well of a 96 well plate. Every sample was measured as a triplicate to correct for technical errors. The prepared 96 well plate was then loaded in the "StepOnePlus System", and the PCR was run (Table 21). Data was analysed using the supplied software, Microsoft Excel and GraphPad Prism based on the ΔΔCt-method with GAPDH as the reference gene

Table 21 Cycler program used for RT-qPCR experiments.

1. Initial denaturation	95°C	15 min
2. Denaturation	95°C	15 s
3. Primer annealing	60°C	30 s
4. Elongation and data collection	72°C	40 s
5. Loop	Go to step 2	44X
6. Melt curve		

2.2.3.2.6 RNAseq

RNAseq was performed to identify candidate genes for further studies on X-chromosomal reactivation during neuronal differentiation.

The library for RNAseq experiments was prepared from 5 ng of total RNA using the "Ovation® Solo RNA-Seq Library Preparation Kit" following the manufacturer's instructions without changes.

The finished library was denatured for the sequencing by mixing 5 µL of the 4 nM library with 5 µL of NaOH (0,2 M). After incubating for 5 min, 5 µL of Tris buffer were added (200 nM Tris-HCl, pH 7). Finally, 985 µL of prechilled HT1 were added. 3 pM of the resulting 20 pM denatured library were then loaded on a "NextSeq 500/550 High Output Cartridge". The sequencing run was performed as a paired-end run with 2 times 76 cycles and an expected output of 50 Million reads per sample for allele-specific analysis.

Data was processed and analysed by [redacted] and [redacted].

2.2.3.3 Protein Analysis

Protein analysis was performed to identify expression differences between cells in a qualitative and quantitative way.

2.2.3.3.1 Immunofluorescent staining

Immunofluorescent staining was used to qualitatively identify protein expression in cells. Previously PFA-fixed (2.2.2.3) cells were first incubated with 1 mL of blocking buffer (PBS, 5% BSA, 0,3% Triton) for 30 min. This was followed by an overnight incubation at 4°C with the primary antibody (Table 7) diluted in blocking buffer. On the second day, cells were washed three times for 10 min with wash buffer 1 (PBS, 0,1% Triton), followed by an incubation for 1 h in the dark with the secondary antibody (Table 8) diluted in blocking buffer. This was again followed by three wash steps with wash buffer 2 (PBS, 0,3% Triton). In a last step, coverslips were transferred to slides prepared with 10 µL of mounting medium (Vectashield, 0,5% DAPI). Pictures were taken using the "Echo Revolve" microscope.

2.2.3.3.2 Protein isolation

Protein was isolated from cell pellets using Magic Mix (48% urea, 15 mM Tris pH 7,5, 8,7% Glycerin, 1% SDS, 143 mM β-mercaptoethanol) containing protease and phosphatase inhibitors. The pellet was resuspended in a suitable amount of Magic Mix and the solution was transferred to a QIAshredder. After centrifugation at 12.000 rpm for 2 min, the solution was pipetted into a fresh tube and frozen at -80°C until needed. The concentration was not measured.

2.2.3.3.3 SDS gel-electrophoresis and Western Blot

SDS gel-electrophoresis was used to separate proteins by their size, that were previously isolated from cells (2.2.3.3.2). The separating gel was prepared and overlayed with the collecting gel once it was polymerized (Table 22). The isolated protein was prepared for gel-electrophoresis by mixing 10 µL of the protein solution with 5 µL of loading dye (Magic Mix with bromphenol blue) and incubating at 95°C for 5 min. The denatured protein solution was then, together with a size marker, loaded on the gel and separated at 199 V for ~50 min. The gel was then used for blotting on a PVDF-membrane by stacking filter paper, the membrane, the gel, and again filter paper in the blotting chamber using the "Trans Blot Turbo Transfer Pack". The blot was run at 1 A for 30 min. The membrane, now containing the separated proteins, was afterwards incubated with blocking buffer (PBS, 0,1% Tween, 5% milk powder/BSA) for 1 h on a shaker. The blocking buffer was removed, and the membrane was then incubated overnight with the primary antibody (Table 7) diluted in blocking buffer. On the second day, the membrane was washed three times for 10 min with PBS-T (PBS, 0,1% Tween)

and then incubated for 1 h with the secondary antibody (Table 8) diluted in blocking buffer. After three more washing steps, the membrane was exposed using the “Western Lightning Plus-ECL, Enhanced Chemiluminescence Substrate” and the “ChemiDoc Imaging System”. Images were prepared and analysed using the “Image Lab” software.

Table 22 Composition of separating and stacking gel for SDS gel-electrophoresis.

Gel	Acrylamid (30%)	H_2O	Tris buffer	SDS (10%)	APS (10%)	TEMED
Separating	1,7 mL	1,9 mL	1,3 mL (1,5 M; pH 8,8)	50 µL	50 µL	2 µL
Stacking	330 µL	1,4 mL	250 µL (0,5 M; pH 6,8)	20 µL	20 µL	2 µL

3 Results

3.1 Cell type confirmation and characterization

Before performing any further experiments, all cells were characterized to confirm their cellular identity. Please refer to chapter 2.1.6 and Table 12 for detailed information about the used cells and how they were labelled.

3.1.1 Fibroblasts

As already mentioned, three different sets of fibroblasts were used. The male and female OS patient-specific fibroblasts were isolated already over 20 years ago with an unknown method, while the wildtype control fibroblasts were freshly isolated from a skin biopsy as a part of this thesis.

3.1.1.1 Morphology

All three fibroblast cell lines showed an elongated morphology and grew as monolayers (Figure 14). The growth rate was steady, with the male OS fibroblasts growing slightly faster than the two others.

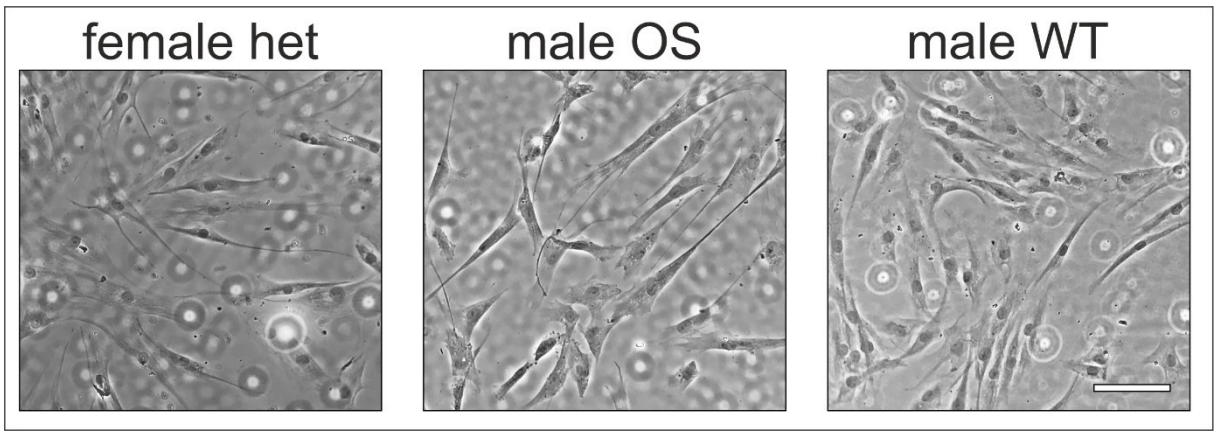


Figure 14 Morphology of fibroblasts. White light pictures of fibroblasts show an elongated morphology. Scale bar: 100 µM.

3.1.1.2 Immunofluorescent staining

To prove their cellular identity, immunofluorescent staining was performed. All three fibroblasts showed expression of the collagen protein SERPINH1, located in the cytoplasm (Figure 15).

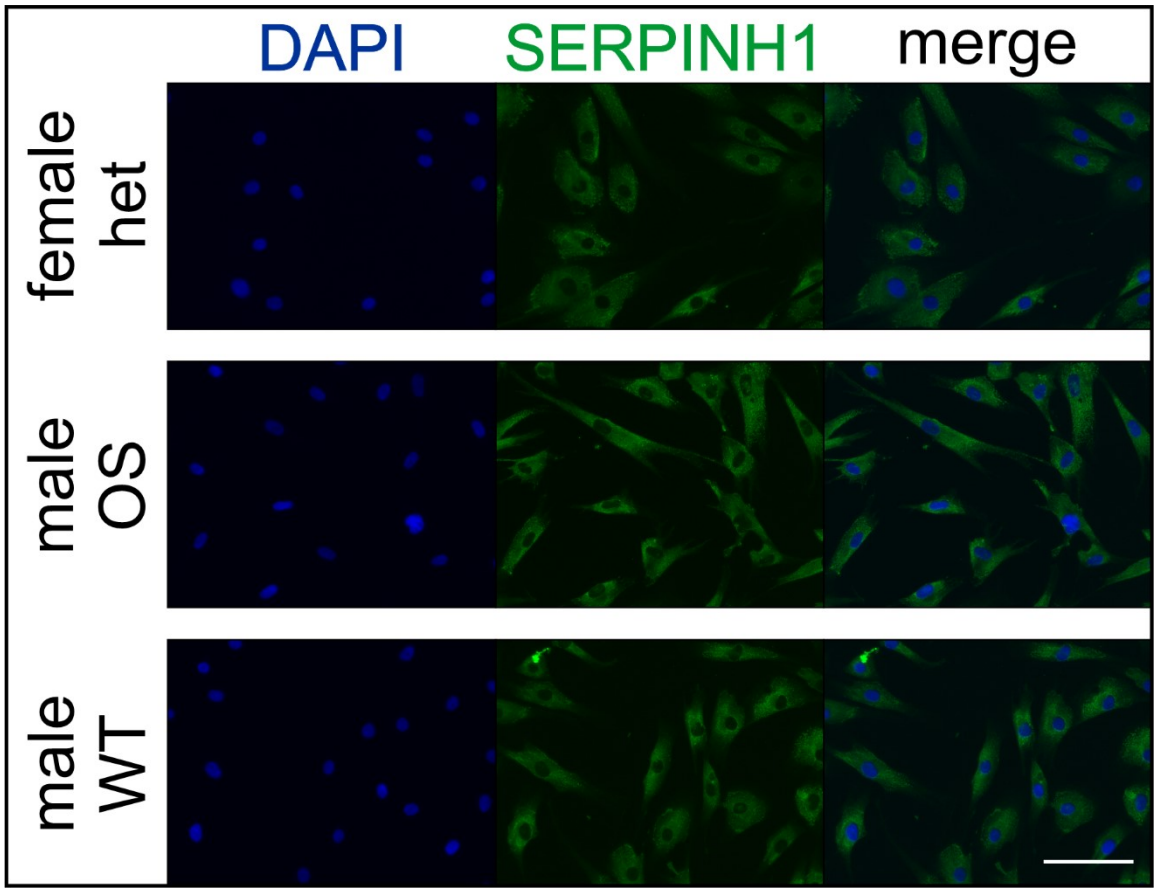


Figure 15 Immunofluorescent staining of the fibroblasts. Fibroblasts show expression of the collagen protein SERPINH1. Scale bar: 50 µM.

3.1.2 iPSCs

Fully characterized fibroblasts were used for reprogramming into iPSCs. Three to four weeks after viral transfection (retro- or sendaiviral), colonies started to emerge from the fibroblast culture and were expanded. The following results show the characterization of the reprogrammed cells C1, ctrl M10 and OS M7 and the cell lines R1 and R2 generated from C1 by CRISPR/Cas genome editing (see also chapter 2.2.2.7.9 and 3.3.2).

3.1.2.1 Morphology

All iPSCs used were steadily growing as densely packed colonies with sharp edges and as monolayers (Figure 16). Single cells within colonies showed a round morphology and the nucleus accounted for approximately 90% of the whole cell volume.

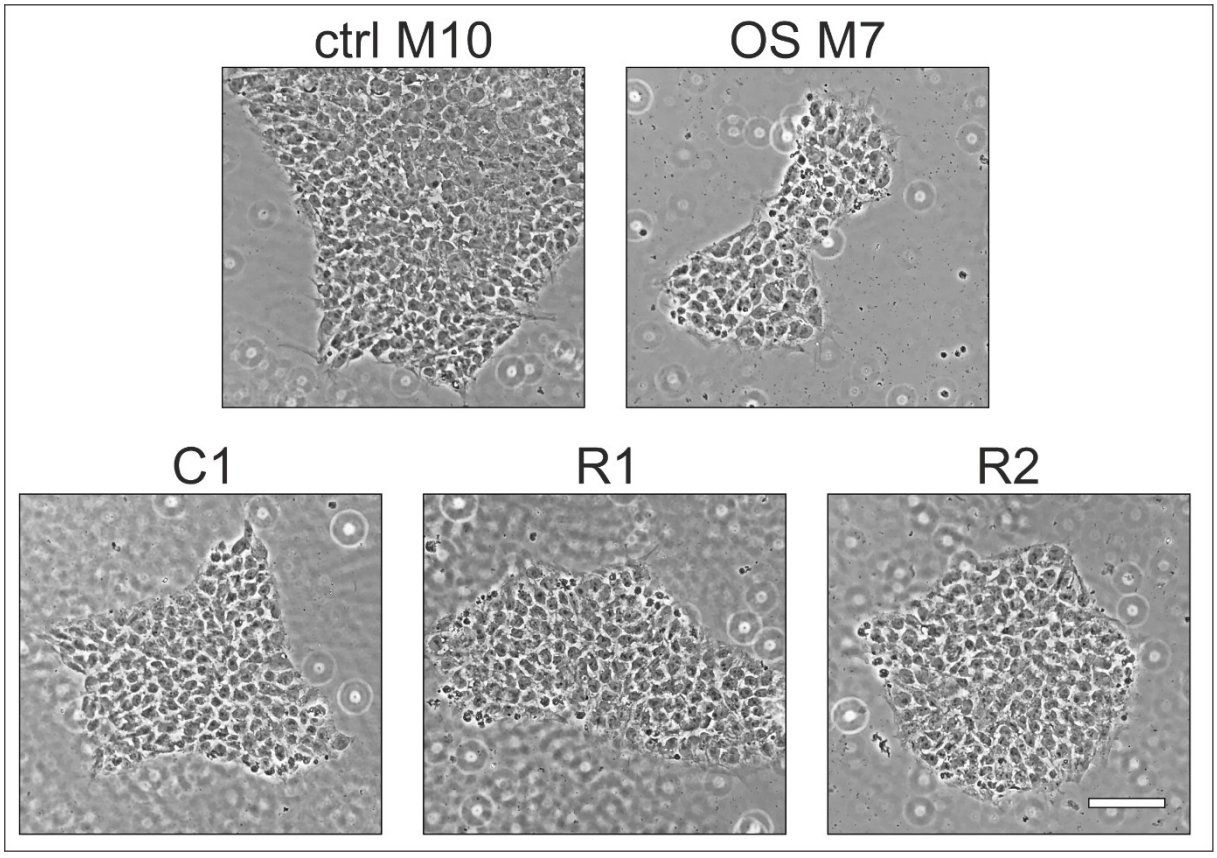


Figure 16 Morphology of iPSCs. White light pictures of iPSCs show densely packed colonies growing as monolayers. Single cells appear round with big nuclei. Scale bar: 100 µM.

3.1.2.2 Karyogram

To exclude chromosomal re-arrangements that might have occured during reprogramming or genome editing, all iPSCs were karyotyped (Figure 17). The two iPSC clones derived from the female heterozygous fibroblasts showed a normal female karyotype of 46,XX, while the iPSC clone derived from the male wildtype fibroblasts showed a normal male karyotype of 46,XY. The two iPSC clones derived from the C1 iPSCs after genome editing also showed a normal male karyotype of 46,XY. None of the used iPSCs showed any kind of chromosomal abnormality at any timepoint.

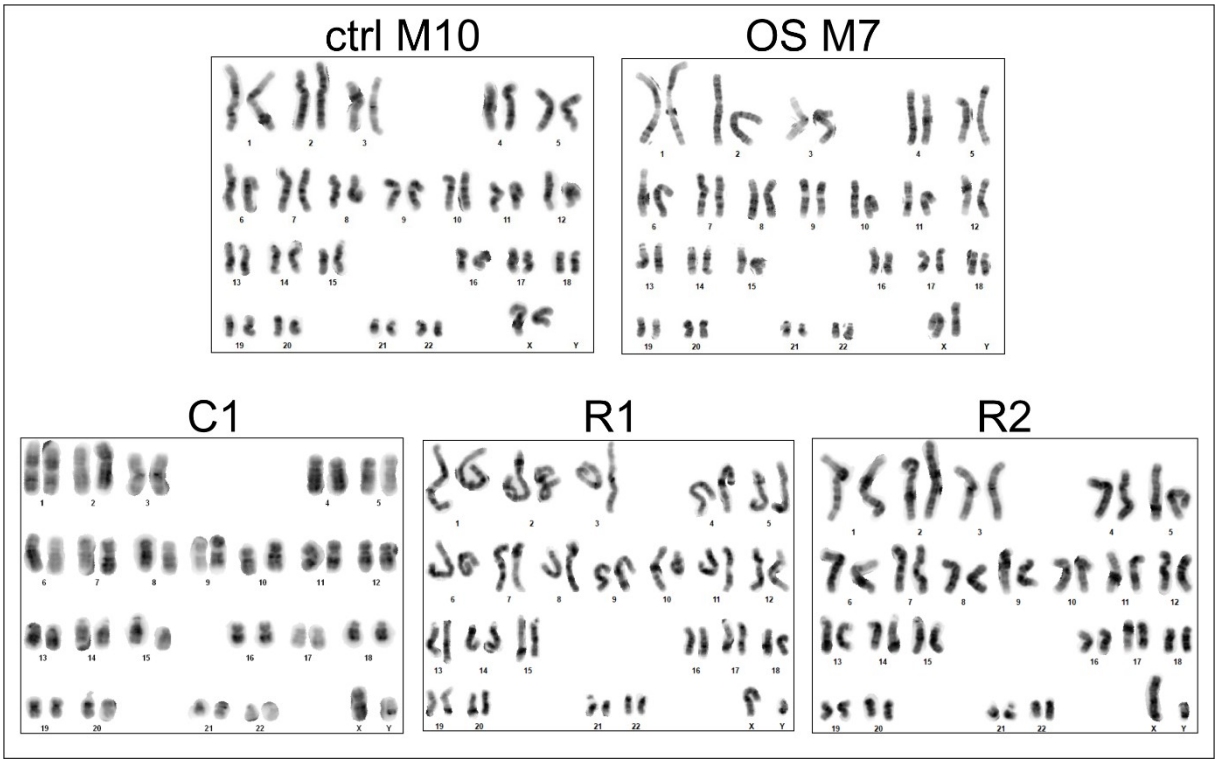


Figure 17 Karyogramm of the iPSCs. C1, R1 and R2 iPSCs showed a normal male karyotype of 46,XY by G-banding analysis, while ctrl M10 and OS M7 iPSCs showed a normal female karyotype of 46,XX by G-banding analysis.

3.1.2.3 Pluripotency (RT-qPCR)

To prove the pluripotency of the used iPSCs, RT-qPCR was performed to quantify the mRNA expression of the known pluripotency marker genes *KLF4*, *NANOG*, *OCT4*, and *SOX2* (Figure 18). Measured values were normalized against *GAPDH* and plotted relative to the fibroblasts they were derived from. All iPSCs showed a similar expression profile of the four genes. *KLF4* is downregulated three to five times on a log2 scale, while the other three genes are upregulated. *NANOG* and *OCT4* are about 10 times upregulated on a log2 scale, while *SOX2* is about 15 times upregulated on the same scale. RT-qPCR showed that the used iPSCs express the pluripotency marker genes at similar levels, different from the original fibroblasts.

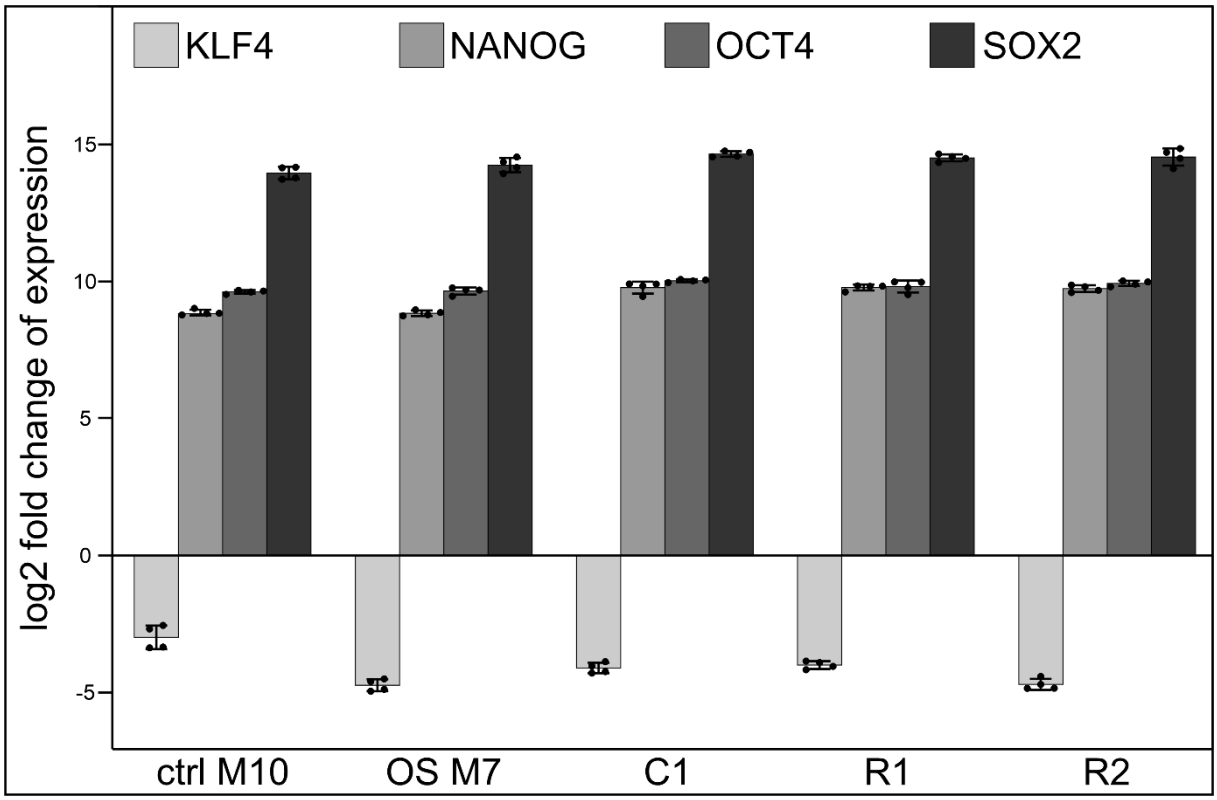


Figure 18 RT-qPCR of the iPSCs. iPSCs showed a decreased mRNA expression of *KLF4* and an increased expression of *NANOG*, *OCT4*, and *SOX2* relative to the original fibroblasts. *GAPDH* was used for normalisation. Values are mean ± SD.

3.1.2.4 Pluripotency (immunofluorescent staining)

To further prove pluripotency of the used iPSCs, immunofluorescent staining of the pluripotency marker proteins NANOG and TRA-1-60 was performed (Figure 19). NANOG is a nuclear transcription factor responsible for self-renewal, while TRA-1-60 is a membrane surface protein exclusively expressed in pluripotent stem cells. All iPSCs showed expression of both proteins. The NANOG staining overlaid with the nuclear counterstaining of DAPI, while the TRA-1-60 staining is more diffuse and localized outside of the nucleus. After already showing expression of pluripotency genes on RNA level, the immunofluorescent staining showed expression of pluripotency genes also on protein level.

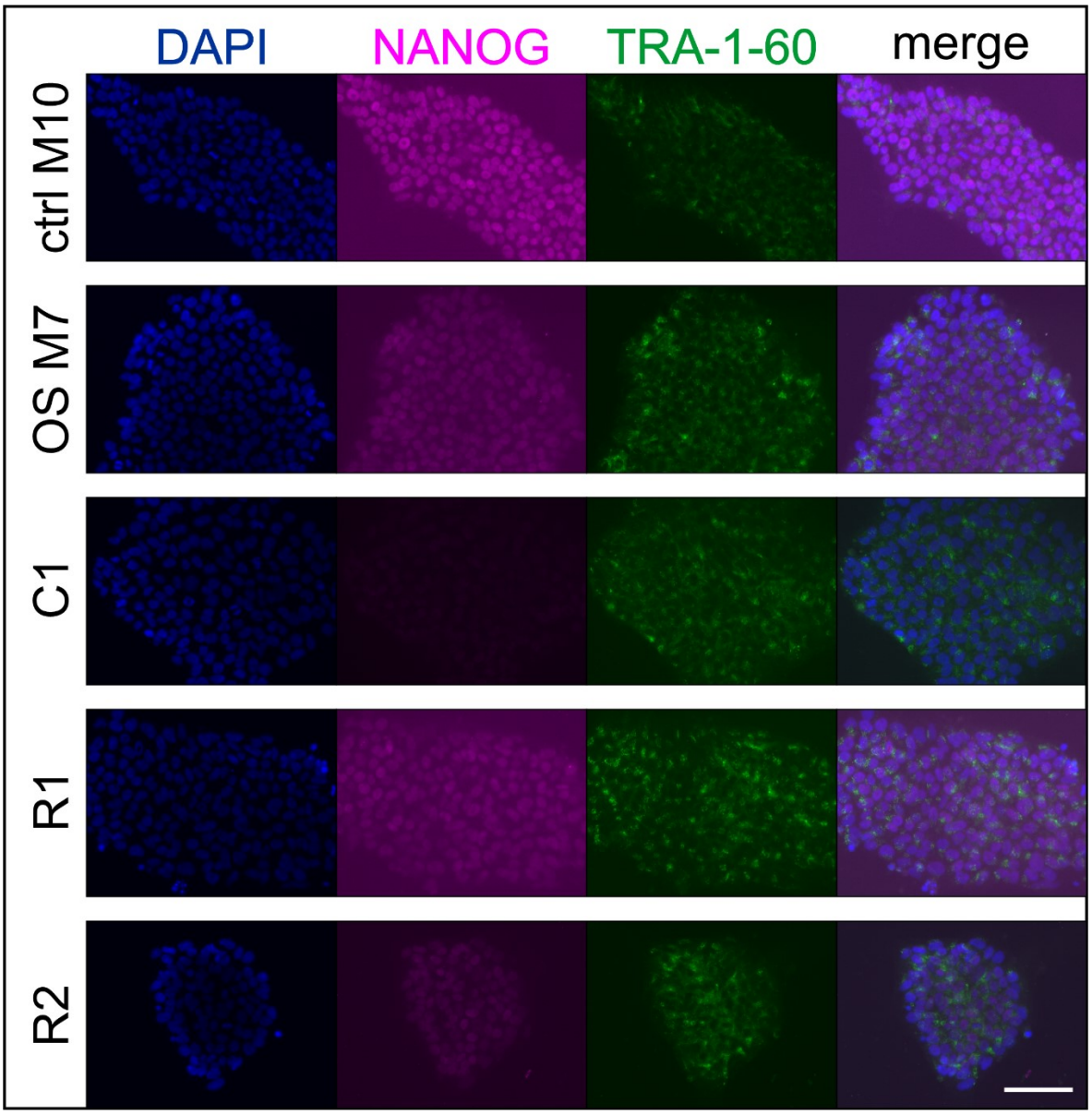


Figure 19 Immunofluorescent staining of the iPSCs. iPSCs showed protein expression of the stem cell markers NANOG and TRA-1-60. Scale bar: 50 µM.

3.1.3 NPCs

Fully characterized iPSCs were differentiated into NPCs. All NPCs were differentiated using the commercial Kit (2.2.2.8.2), while selfmade buffers (2.2.2.8.1) were only used to generate the NPCs for the RNAseq experiment (characterization data not shown).

3.1.3.1 Morphology

NPCs grew as a monolayer culture, showing a slightly elongated cell morphology (Figure 20). Unlike iPSCs, NPCs were split as single cells using trypsin and showed a quick and consistent growth rate. Less than 1% of the cells showed a neuronal like morphology pointing towards spontaneous differentiation.

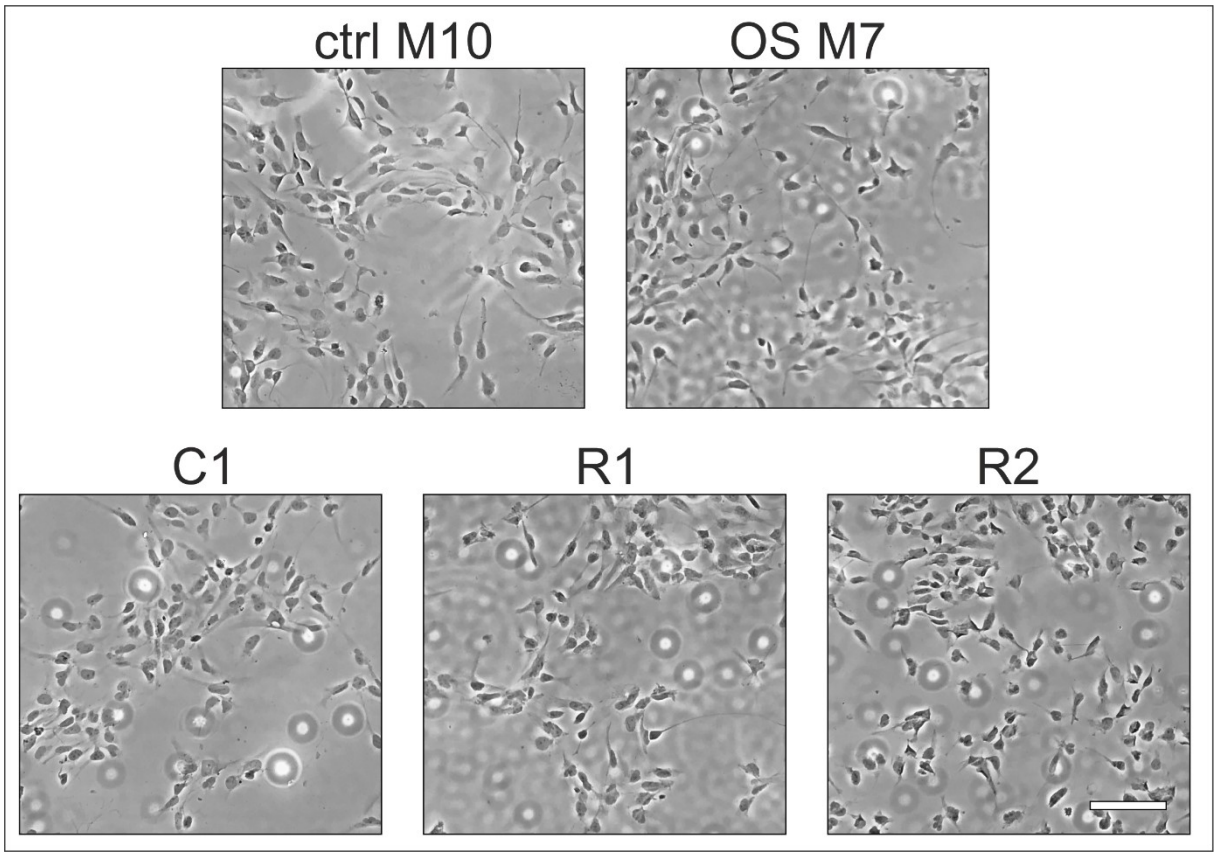


Figure 20 Morphology of NPCs. White light pictures of NPCs show single cells with a slightly elongated morphology growing as a monolayer. Scale bar: 100 µM.

3.1.3.2 RT-qPCR

Like iPSCs, NPC were further characterized by quantification of RNA expression, to prove their correct differentiation (Figure 21). All of the five analysed NPCs showed a similar expression profile of the analysed genes. The pluripotency marker genes were downregulated on a log2 scale between 2-fold (*KLF4*) and 8- to 10-fold (*NANOG* and *OCT4*), while NPC marker genes were upregulated 2-3 fold (*NESTIN*) or even up to 10-fold (*PAX6*). The only gene showing almost no changes in its expression level is *SOX2*, with an upregulation in the NPCs of about 1-fold compared to the respective iPSCs. Since SOX2 is a marker for stem cells including both, iPSCs and NPCs, this small change is not surprising and plausible.

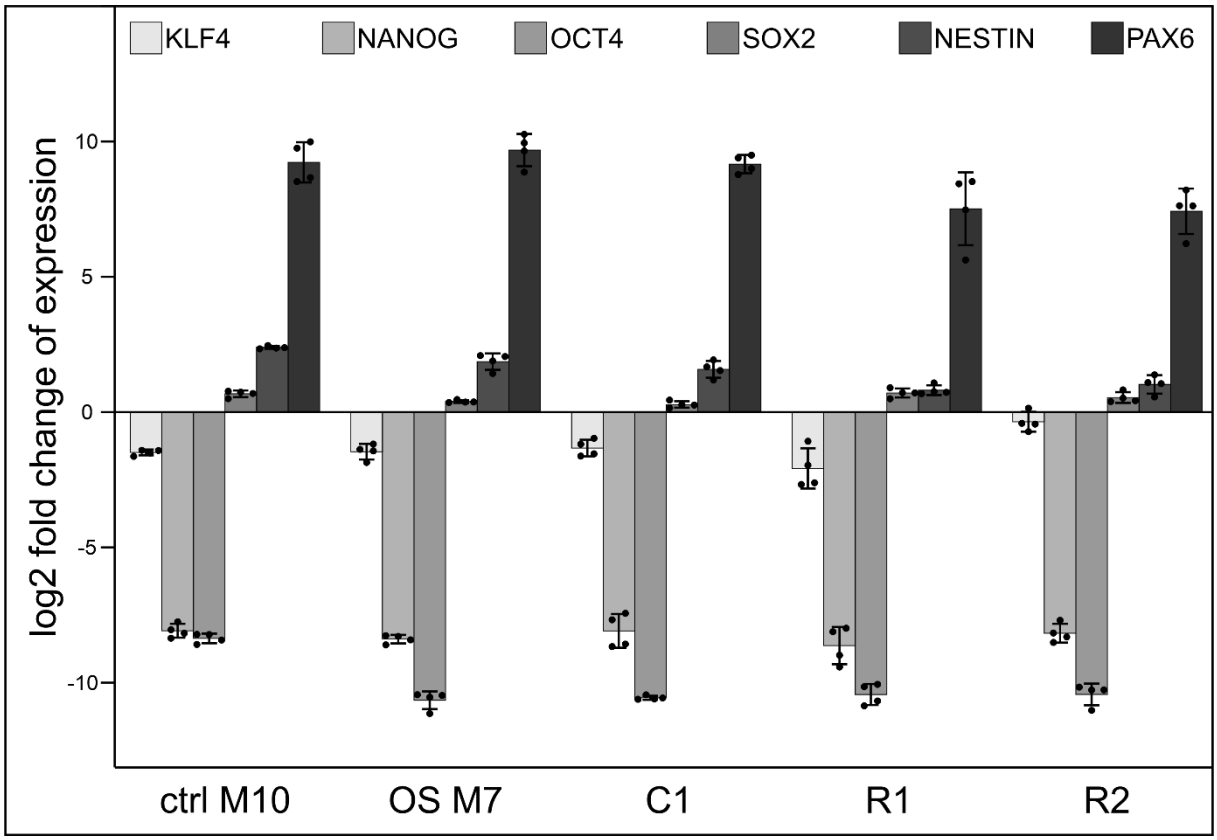


Figure 21 RT-qPCR of the NPCs. NPCs showed a decreased mRNA expression of *KLF4*, *NANOG* and *OCT4*, an almost unchanged expression of *SOX2* and an increased expression of *NESTIN* and *PAX6* relative to the iPSCs they were differentiated from. *GAPDH* was used for normalisation. Values are mean ± SD.

3.1.3.3 Immunofluorescent staining

For further characterization of the NPCs immunofluorescent staining of the neuronal stem cell markers NESTIN, SOX2 and PAX6 was performed (Figure 22, Figure 23). All NPCs showed protein expression of all three markers. The transcription factor SOX2 overlaid with the DAPI counterstaining of the nuclei, while the transcription factor PAX6 showed a slightly diffuse staining of nuclei and cytoplasm. The intermediate filament protein NESTIN was detected in the cytoplasm.

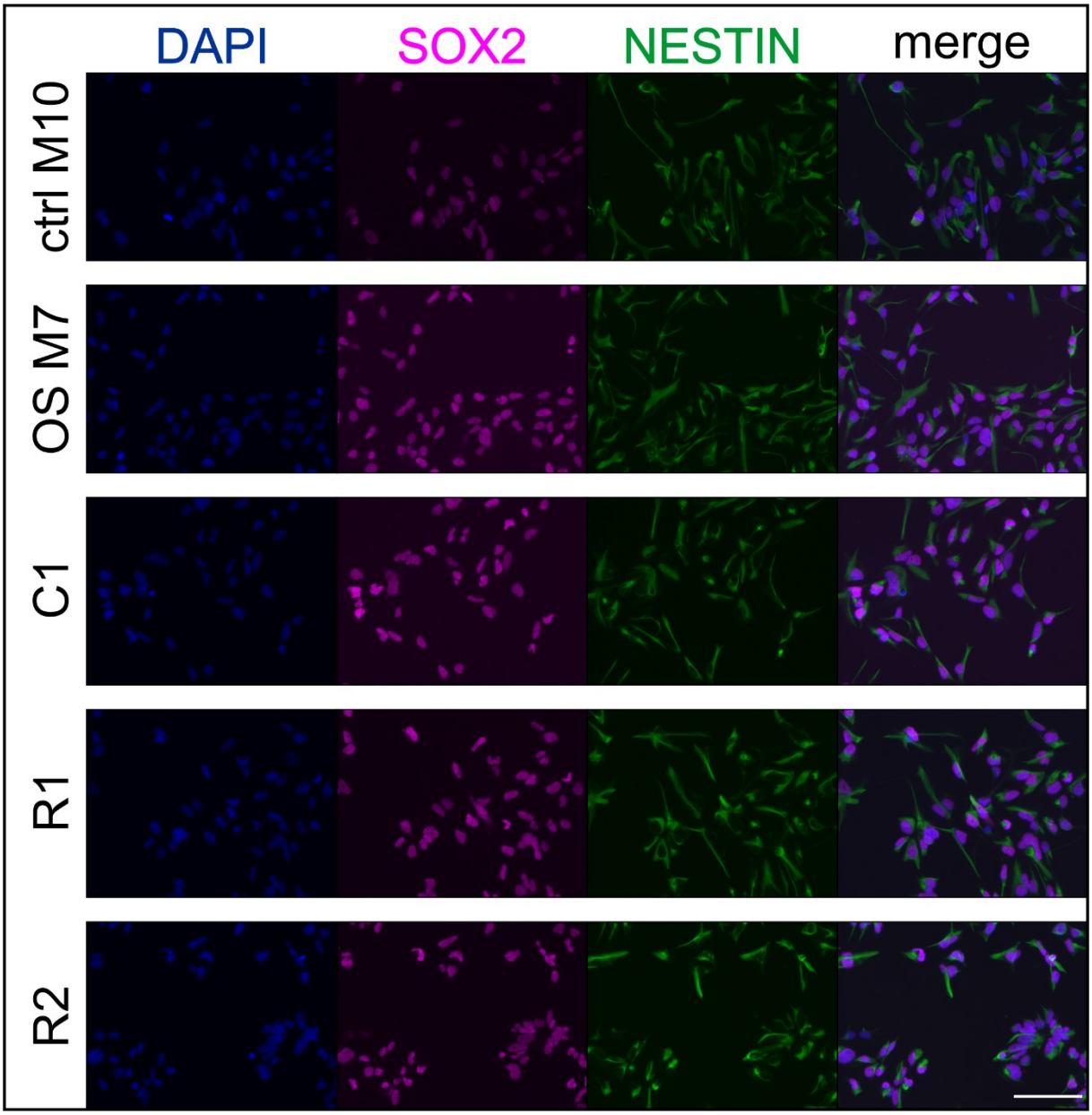


Figure 22 Immunofluorescent staining of NPCs. NPCs showed protein expression of the neuronal stem cell markers NESTIN and SOX2. Scale bar: 50 µM.

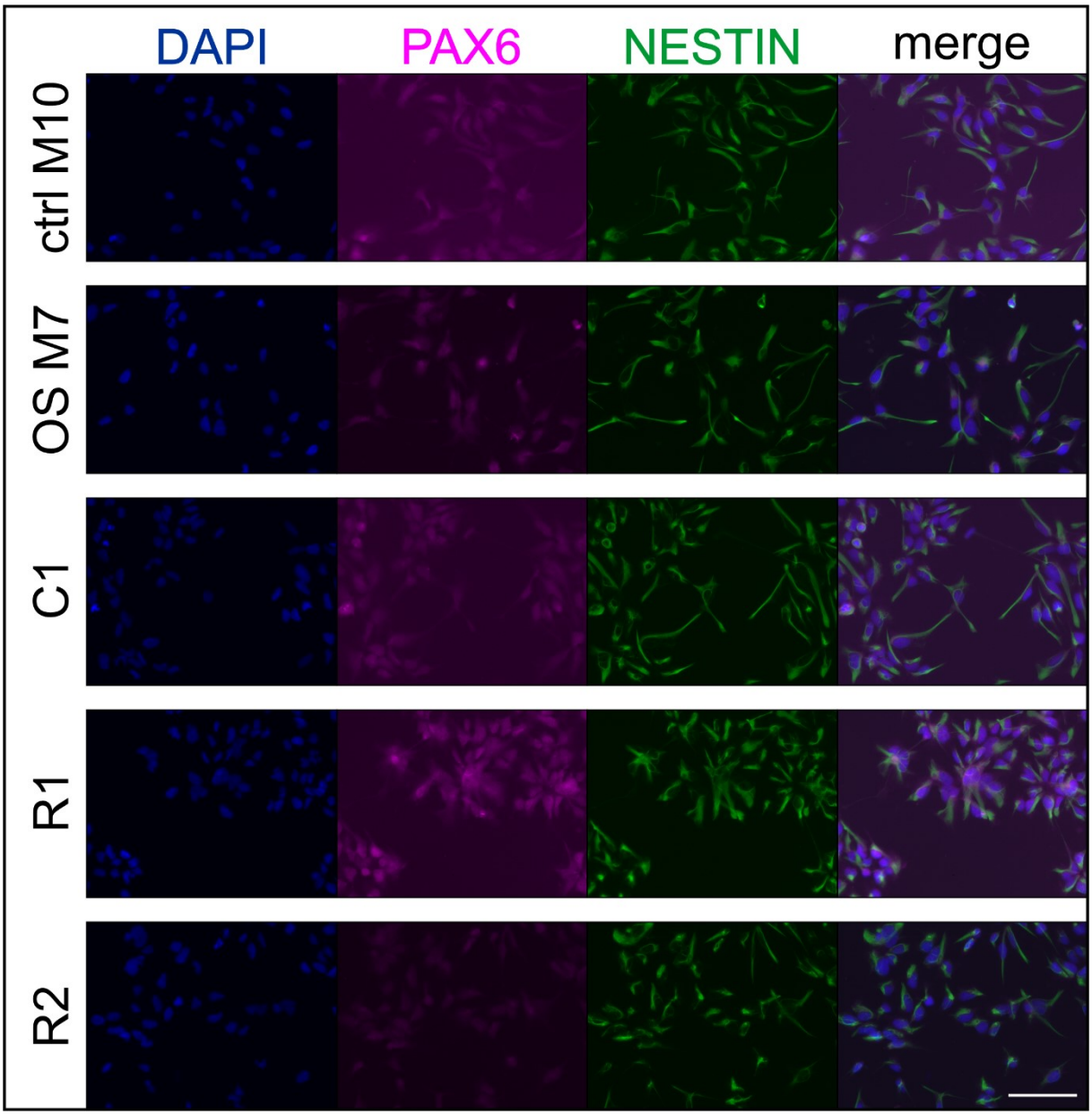


Figure 23 Immunofluorescent staining of NPCs. NPCs showed protein expression of the neuronal stem cell markers NESTIN and PAX6. Scale bar: 50 µM.

3.1.4 Neurons

Fully characterized NPCs were differentiated into neurons by replacing FGF2 in the medium by Vitamin A.

3.1.4.1 Morphology

Over the course of five weeks cell morphology changed considerably. The cells eventually stopped proliferating and showed an elongated morphology with cell bodies surrounded by axonal and dendrite like structures (Figure 24). When reaching three to four weeks of differentiation, the cells started to form spherical clusters growing into all three dimensions.

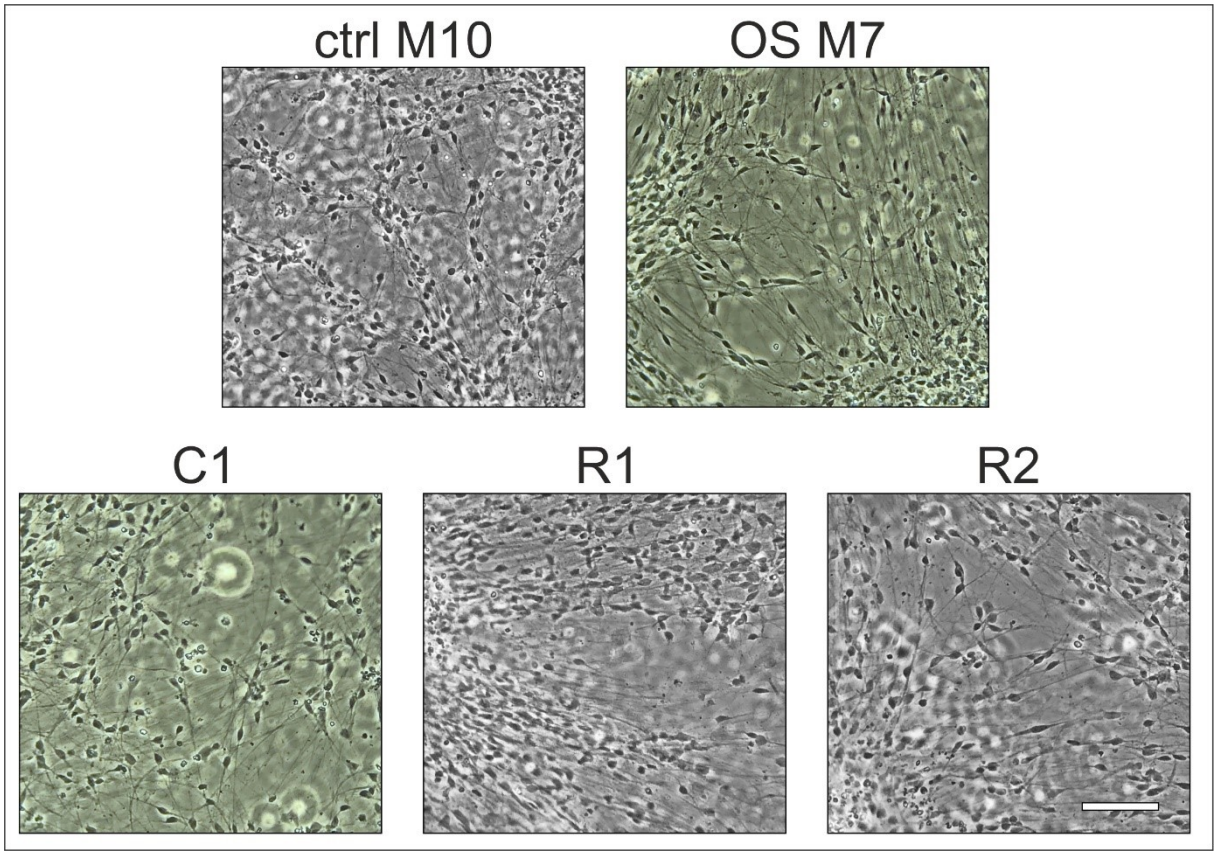


Figure 24 Morphology of neurons. White light pictures of neurons show elongated cells with axonal and dendrite like structures. Cells are clustering and forming aggregates. Scale bar: 100 µM.

3.1.4.2 RT-qPCR

To characterize the neurons, RNA expression was quantified by RT-qPCR (Figure 25). All five neuronal cultures showed a similar expression pattern of the analysed genes. The neuronal stem cell markers (*SOX2*, *NESTIN*, *PAX6*) were downregulated up to 2-fold on a log2 scale, or showed an unchanged expression level compared to the respective NPCs. All three neuronal genes were upregulated in the neuronal cultures compared to the respective NPCs. *TUBB3* and *MAP2* showed an upregulation of 2-3-fold, and *TAU* was upregulated 7-fold, all on a log2 scale.

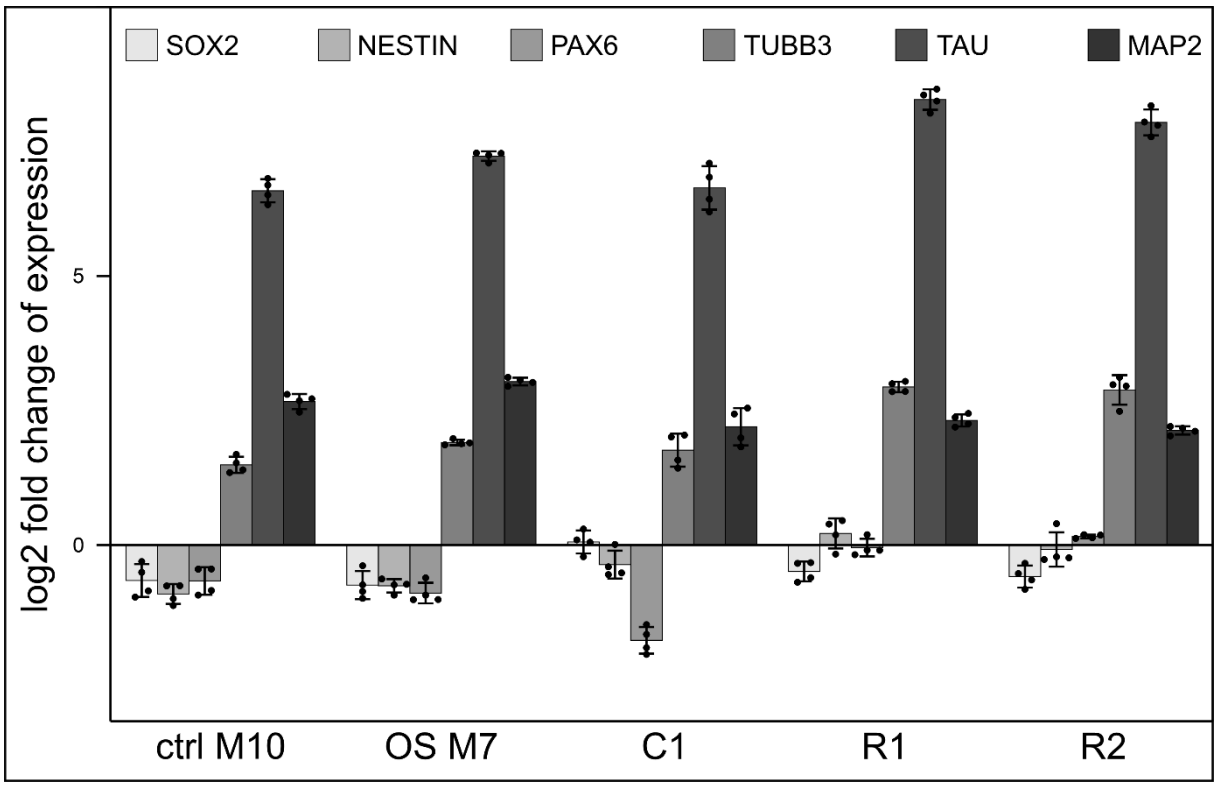


Figure 25 RT-qPCR of the neurons. All neurons showed an increased mRNA expression of the neuronal markers *TUBB3*, *TAU* and *MAP2*. The mRNA expression of the NPC markers differed between the cells from decreased to almost unchanged. *GAPDH* was used for normalisation. Values are mean ± SD.

3.1.4.3 Immunofluorescent staining

For further characterization, all neuronal cultures were analysed by immunofluorescent staining of the neuronal markers TUBB3, TAU and MAP2 (Figure 26, Figure 27). All cultures showed protein expression of the three markers in the axon and dendrite like structures. The axonal marker TAU and the dendrite marker MAP2 both showed a partial overlay with the early neuronal marker TUBB3. Of note, all TAU and MAP2 positive structures also showed expression of TUBB3 but not the other way around.

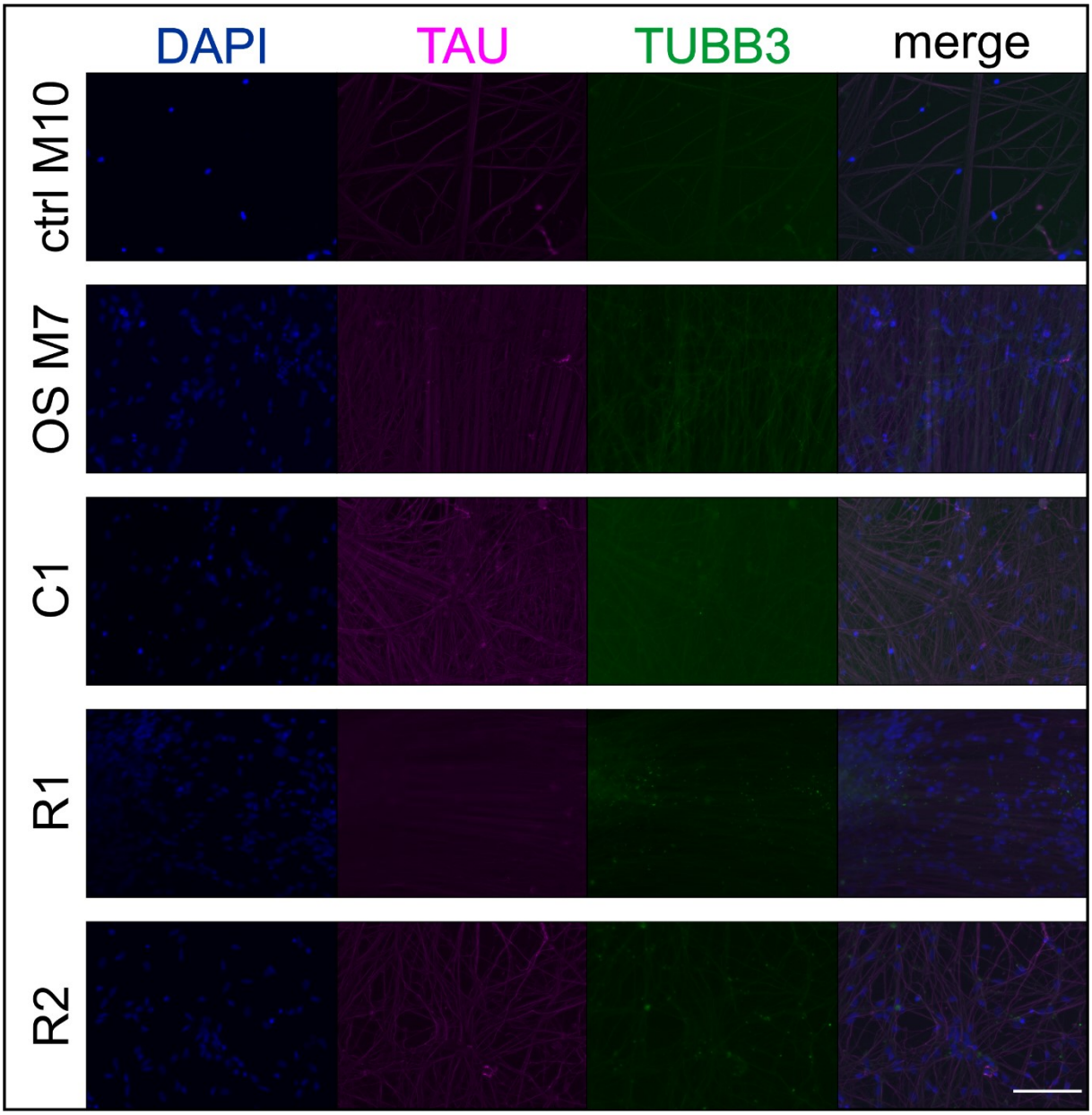


Figure 26 Immunofluorescent staining of the neurons. Neurons showed protein expression of the neuronal markers TUBB3 and TAU. Scale bar: 50 µM.

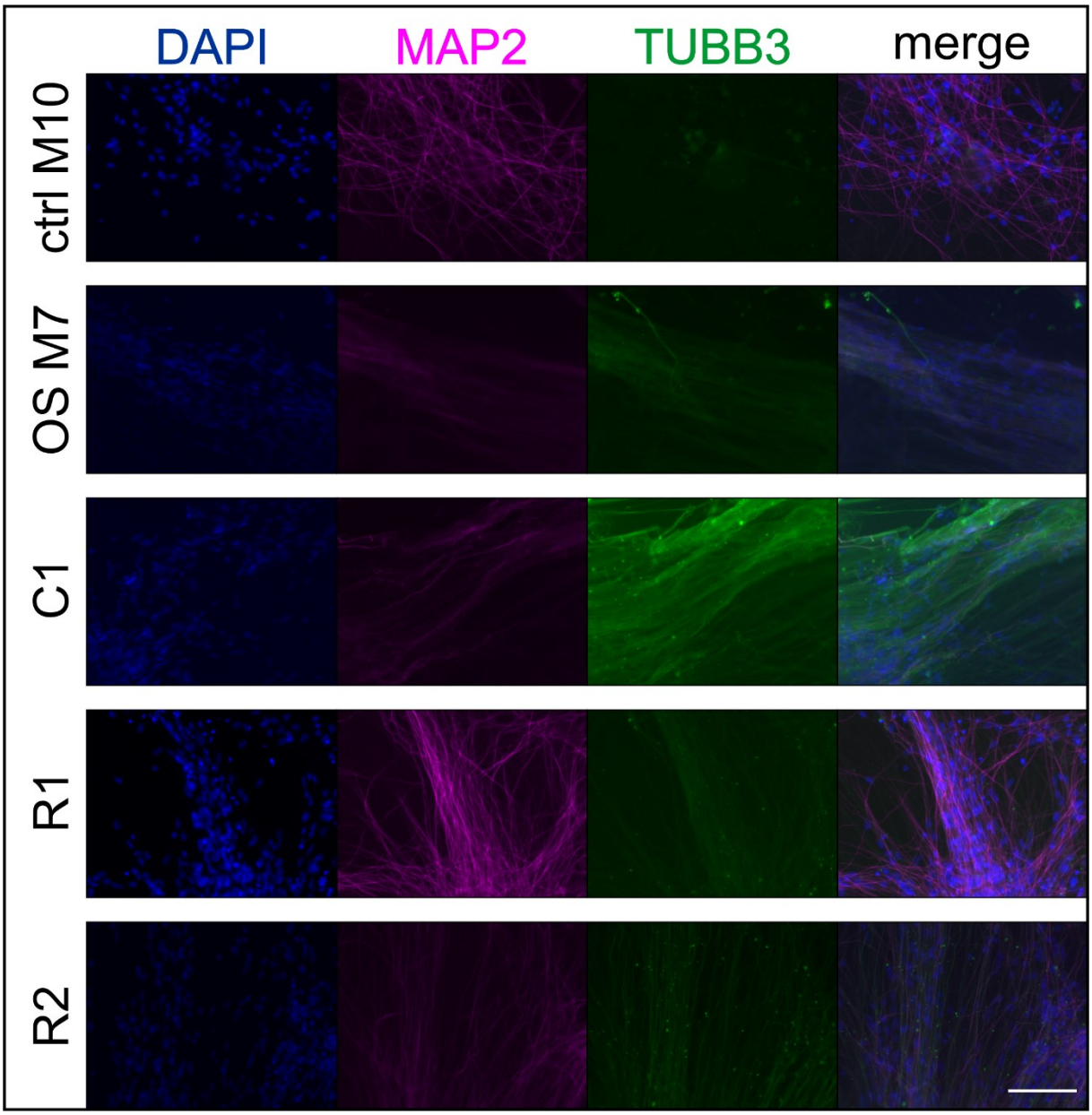


Figure 27 Immunofluorescent staining of the neurons. Neurons showed protein expression of the neuronal markers TUBB3 and MAP2. Scale bar: 50 µM.

3.2 Microduplication of 16p11.2

As a side project, two additional fibroblast cell lines were reprogrammed into iPSCs and then differentiated into NPCs and neurons. The fibroblasts were obtained from two females, a mother (patient number: 1179/17) and her daughter (patient number: 1180/17). The daughter, as well as her brother and father, were previously known to carry a microduplication of ~600 kb on chromosome 16p11.2 (OMIM #614671). Unfortunately, the father was not willing to participate in this study, and the brother was taking medication inhibiting his participation. The cells of the mother were used as a wildtype control.

The family had initially been investigated for connective tissue disease due to an ascending aortic aneurysm and joint laxity in the son. During the diagnostic evaluations a variable spectrum of neuropsychiatric symptoms was revealed in the son, the father and the daughter. The son was diagnosed with developmental delay, learning disability and experienced two psychotic episodes. The father was diagnosed with ADHD and a combined personality disorder. He also received treatment depression and anxiety disorder. The daughter had a psychotic episode at age 18 and has been on antipsychotic treatment. NGS analysis for connective tissue disease did not reveal any mutation, but chromosome microarray analysis identified the previously mentioned microduplication.

3.2.1 Confirmation of cell types

As previously described, all cell types were characterized before performing further experiments.

3.2.1.1 Fibroblasts

Fibroblasts of both mother and daughter were isolated from skin biopsies as previously described (2.2.2.6.1). They showed an elongated morphology and expression of the collagen protein SERPINH1 located in the cytoplasm (Figure 28).

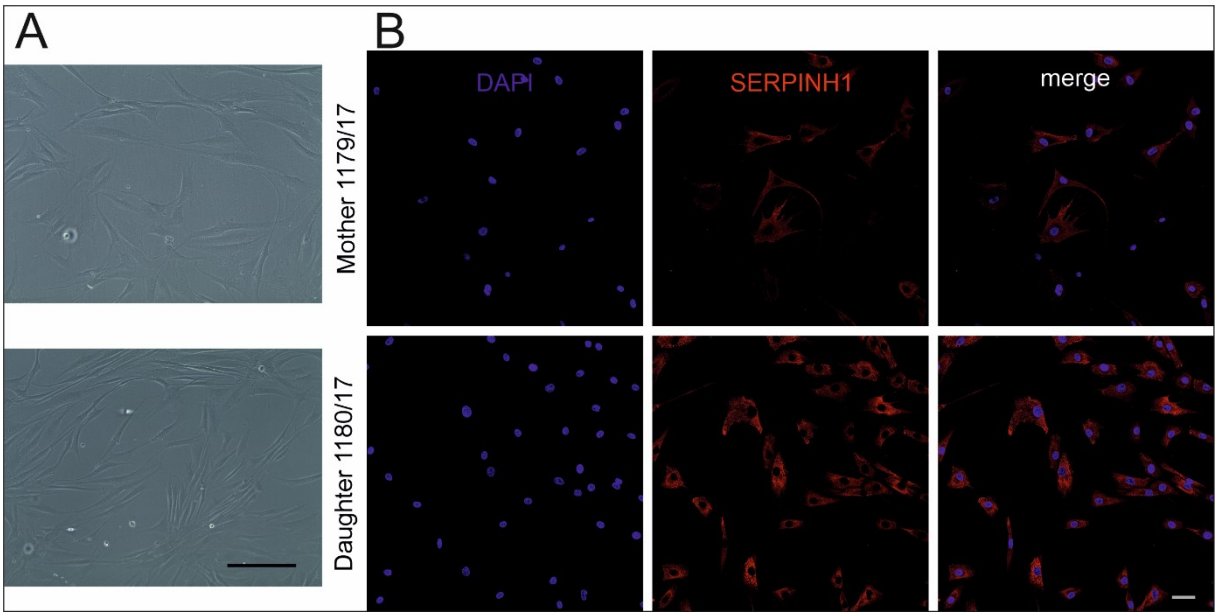


Figure 28 Characterization of 1179/17 and 1180/17 fibroblasts. A: White light pictures of fibroblasts show single cells with a slightly elongated morphology growing as a monolayer. Scale bar: 200 µM. B: Fibroblasts show expression of the collagen protein SERPINH1. Scale bar: 50 µM.

3.2.1.2 iPSCs

Fibroblasts were reprogrammed into iPSCs as previously described (2.2.2.7.1). Two different iPSC clones of the mother (A2 and A27) and one iPSC clone of the daughter (B8) were characterized and later used for gene expression analysis (3.2.2). All three iPSC clones retained a normal female karyotype of 46,XX during the reprogramming. Expression of pluripotency marker genes changed during the reprogramming. As before (3.1.2.3), *KLF4* expression was downregulated 2-fold, while expression of *NANOG*, *OCT4* and *SOX2* was upregulated at least 10-fold, all on a log2 scale. All three clones also showed expression of NANOG and TRA-1-60 on protein level, in combination with the RT-qPCR data a clear evidence that the cells are pluripotent and the reprogramming was successful. Immunofluorescent staining of 1179/17 A2 iPSCs was not performed in the scope of this work but during later experiments in the lab.

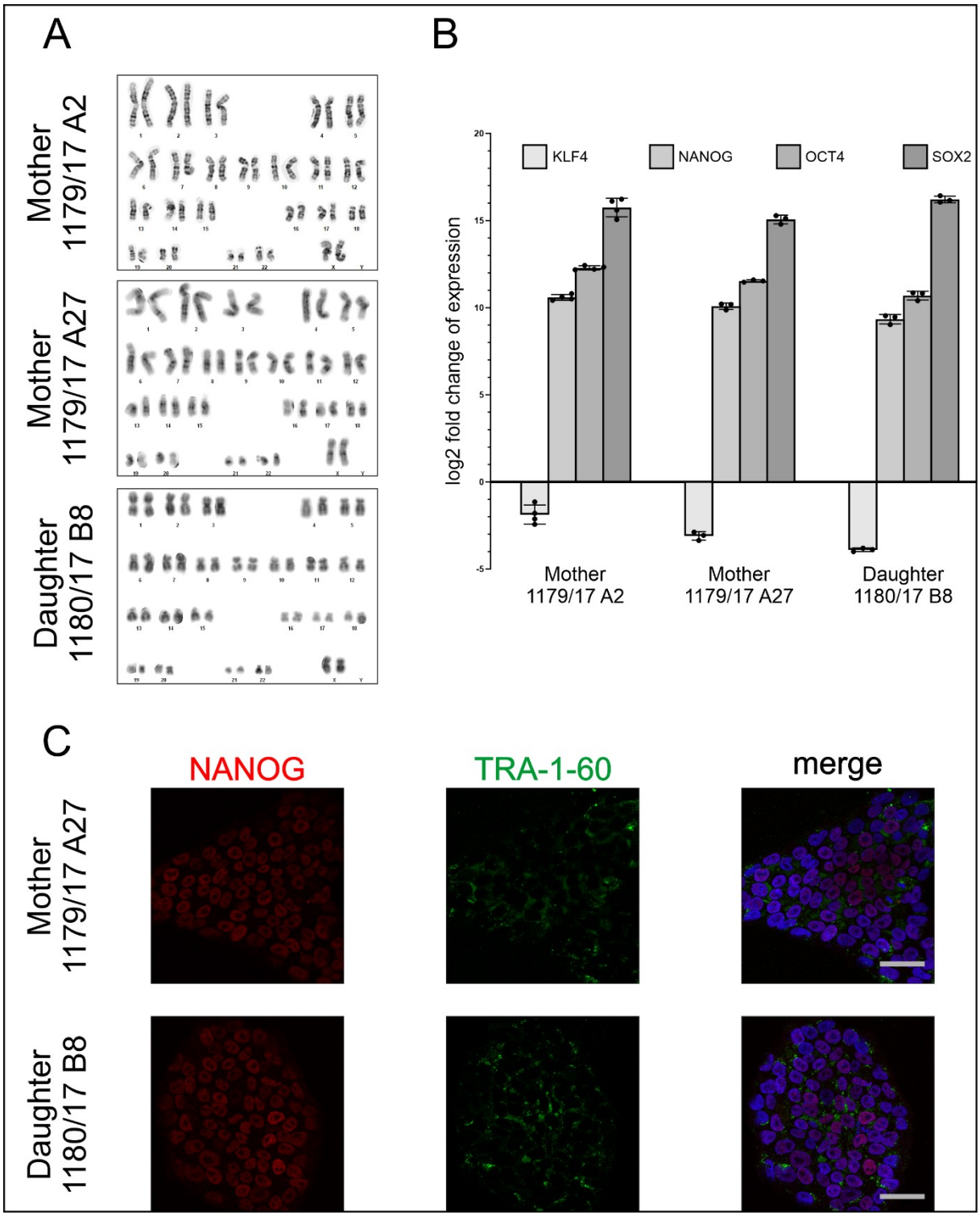


Figure 29 Characterization of 1179/17 and 1180/17 iPSCs. A: iPSCs showed a normal female karyotype of 46,XX by G-banding analysis. B: iPSCs showed a decreased mRNA expression of *KLF4* and an increased expression of *NANOG*, *OCT4*, and *SOX2* relative to the original fibroblasts. *GAPDH* was used for normalisation. Values are mean ± SD. C: Analysed iPSCs showed protein expression of the stem cell markers NANOG and TRA-1-60. Scale bars: 50 µM.

3.2.1.3 NPCs

To get a better understanding of the neuronal mechanisms underlying the 16p11.2 microduplication, the iPSCs were differentiated into NPCs. As NPCs of 1179/17 A2 had already been generated, we decided to use this iPSC clone of the mother for the further differentiation and gene expression analysis experiments. We also differentiated the iPSC clone 1180/17 B8 of the daughter into NPCs.

The differentiated NPCs of both iPSC clones showed a decreased expression of the pluripotency markers *NANOG* and *OCT4*, while the stem cell marker *SOX2* remained unchanged (Figure 30 A). Expression of the pluripotency marker *KLF4* was slightly downregulated in the A2 NPCs while being slightly upregulated in the B8 NPCs. The NPC markers *NESTIN* and *PAX6* showed an upregulated expression in both NPC lines. Immunofluorescent staining revealed protein expression of NESTIN, SOX2 and PAX6 in both generated NPC lines (Figure 30 B). Taken together, this data showed that we successfully differentiated the iPSCs of 1179/17 A2 and 1180/17 B8 into NPCs.

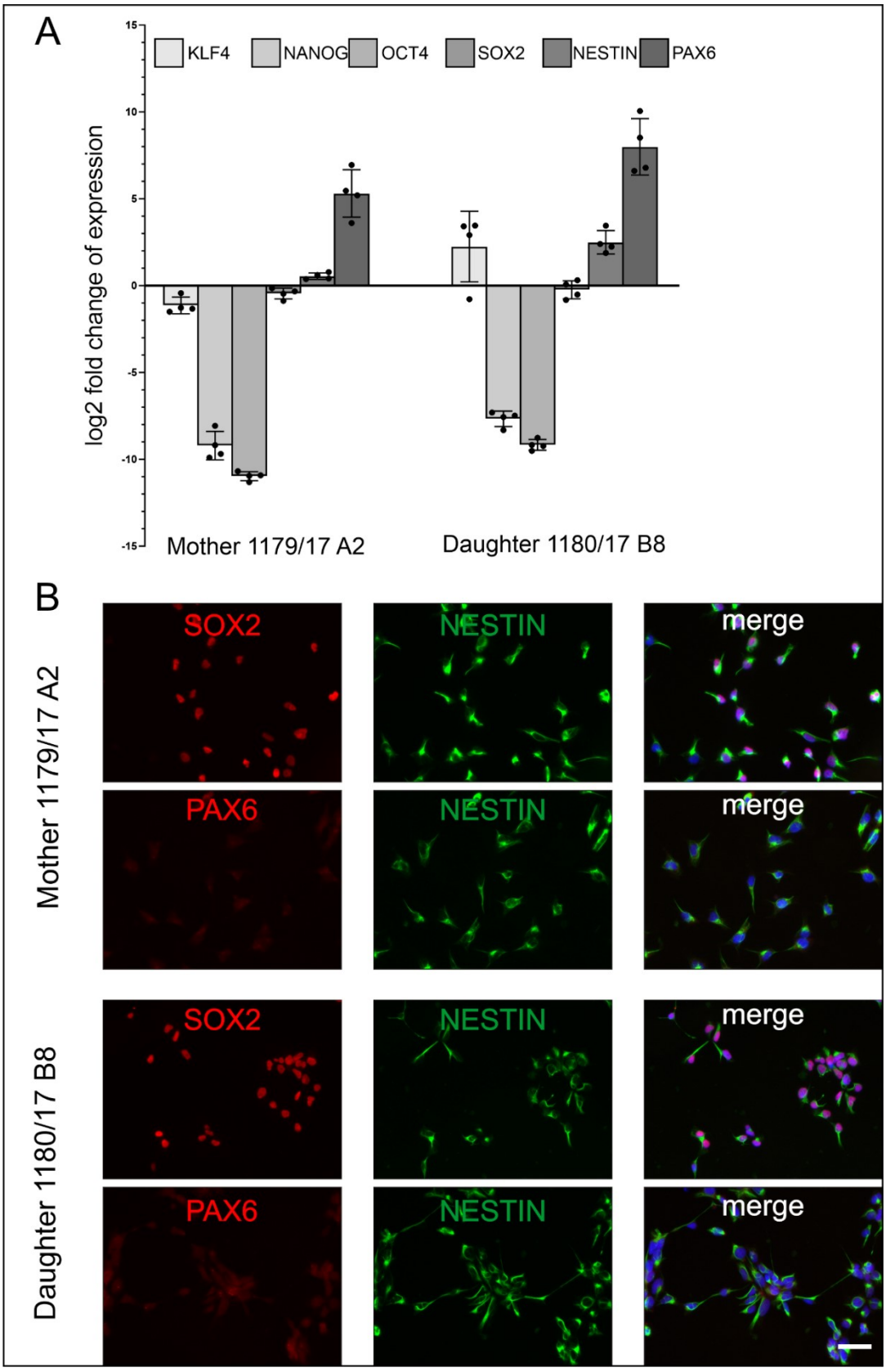


Figure 30 Characterization of 1179/17 and 1180/17 NPCs. A: NPCs showed a decreased mRNA expression of *NANOG* and *OCT4*, an almost unchanged expression of *SOX2* and an increased expression of *NESTIN* and *PAX6* relative to the iPSCs they were differentiated from. *GAPDH* was used for normalisation. Values are mean ± SD. B: NPCs showed protein expression of the neuronal stem cell markers NESTIN, SOX2 and PAX6. Scale bar: 50 µM.

3.2.1.4 Neurons

The previously generated NPCs were further differentiated into neurons. The neurons were characterized using again RT-qPCR and immunofluorescent staining. Interestingly, expression of all three NPC markers (*SOX2*, *NESTIN* and *PAX6*) was increased in the differentiated neurons relative to the NPCs (Figure 31 A). Expression of neuronal markers was also increased, comparable to the previously shown neurons (3.1.3.2). During the 35 days long differentiation of the NPCs into neurons, the cultures of both 1179/17 A2 and 1180/17 B8 showed a continuous growth of NPCs that was not observed when differentiating the NPCs of ctrl M10 or the other previously described cell lines. The cells grew in different layers, mostly with the neurons on top of the NPCs. The immunofluorescent staining eventually showed that these cells differentiated into functional neurons, expressing the neuronal marker TUBB3, the axonal marker TAU and the dendrite marker MAP2 (Figure 31 B).

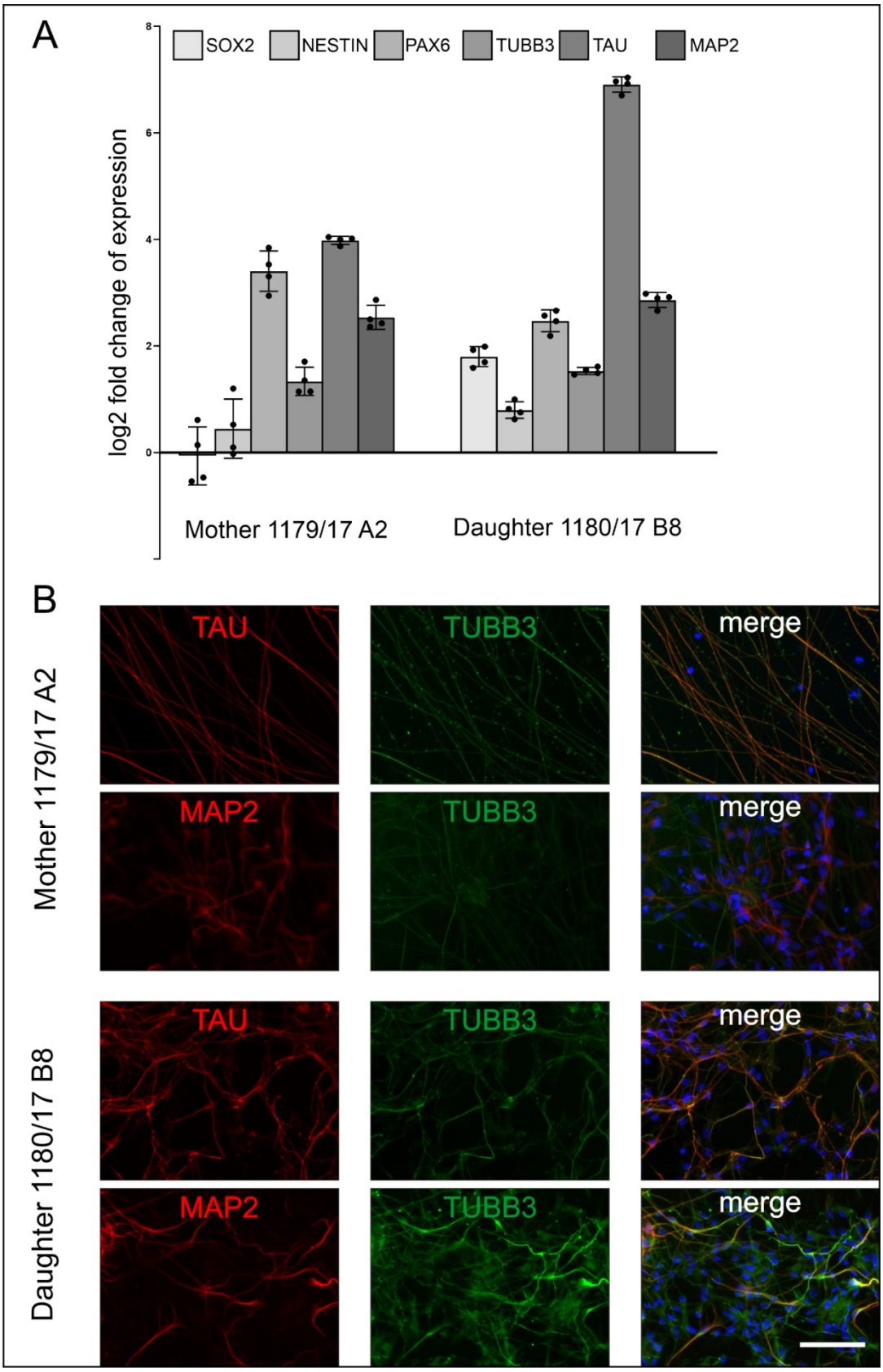


Figure 31 Characterization of 1179/17 and 1180/17 neurons. A: All neurons show an increased mRNA expression of the neuronal markers *TUBB3*, *TAU* and *MAP2*. The mRNA expression of the NPC markers remained unchanged or was increased in relation to the NPCs they were differentiated from. *GAPDH* was used for normalisation. Values are mean ± SD. B: Neurons showed protein expression of the neuronal markers TUBB3, TAU and MAP2. Scale bar: 50 µM.

3.2.2 Gene expression analysis

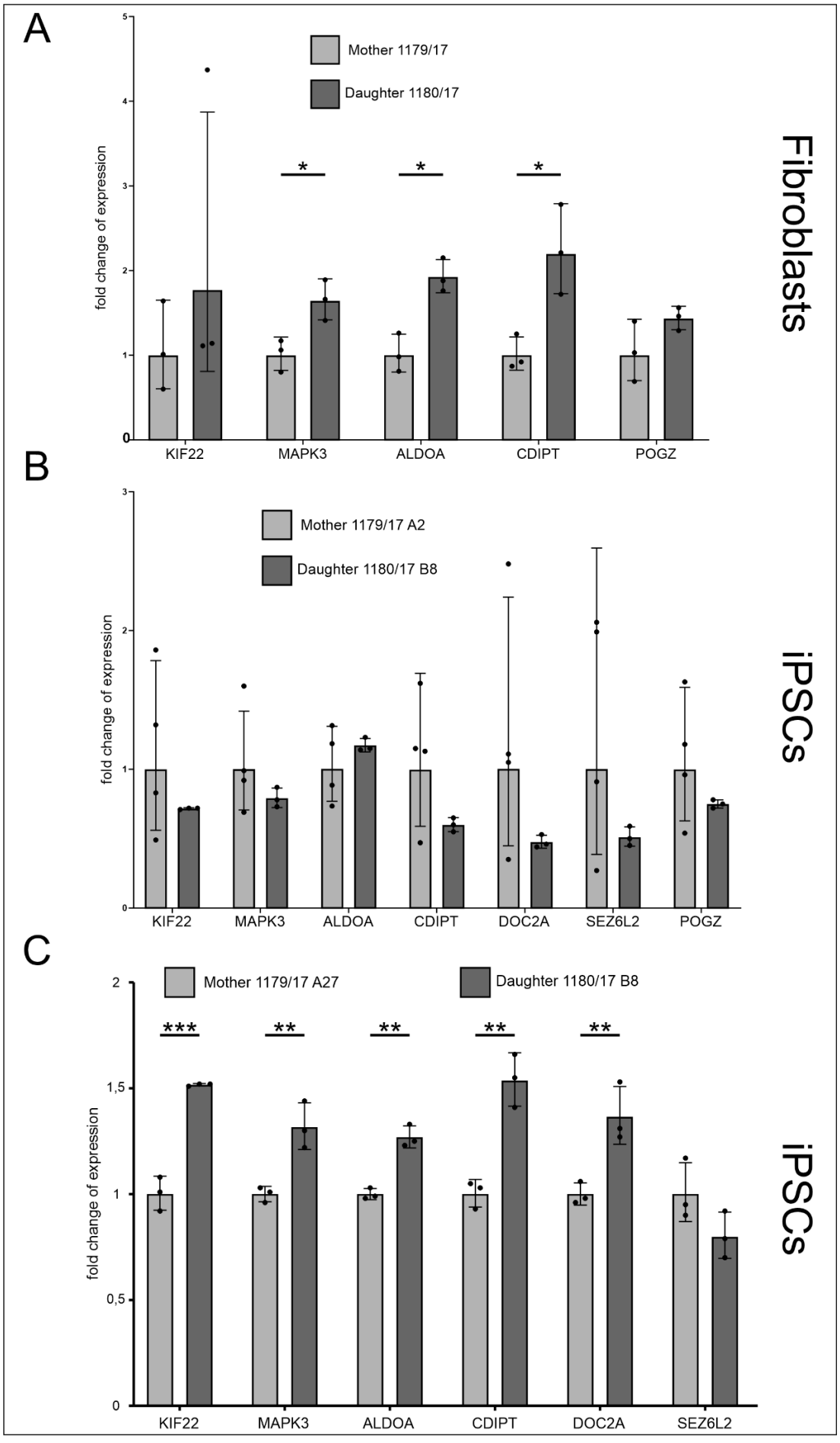
The fully characterized fibroblasts, iPSCs, NPCs and neurons were used for measuring mRNA expression of genes inside the duplicated region of 16p11.2. We chose in total six genes from that region that were predicted to be expressed in all cell types (*Home < Expression Atlas < EMBL-EBI*, n.d.) and considered relevant for neuropsychiatric disorders. In addition, we chose to measure *POGZ* expression; a gene located on chromosome 1 and not connected to the reported microduplication.

As expected, we frequently measured a significantly increased expression of the target genes in the cells of the daughter compared to her mother. In the fibroblasts, expression of *MAPK3*, *ALDOA* and *CDIPT* was significantly increased while expression of *KIF22* and *POGZ* remained unchanged (Figure 32 A). Yet, expression of both genes is slightly increased in the cells of the daughter, but due to a low number of replicates and a high standard deviation their change does not reach statistical significance. *DOC2A* and *SEZ6L2* were not expressed in fibroblasts.

For measuring gene expression in the iPSCs two different iPSC clones of the mother were used as controls. When using the iPSC clone A27 of the mother as a control, gene expression of *KIF22*, *MAPK3*, *ALDOA*, *CDIPT* and *DOC2A* was significantly upregulated in the iPSCs of the daughter, while *SEZ6L2* remained unchanged (Figure 32 C) (*POGZ* expression was not measured). But, when using the iPSC clone A2 of the mother as a control, none of the genes showed a significant change of expression in the iPSCs of the daughter (Figure 32 B). Gene expression in the iPSCs of this clone was extremely variable and the generated results were deemed to be not trustworthy.

Gene expression analysis in the NPCs showed a significant upregulation of *KIF22*, *MAPK3*, *CDIPT* and *SEZ6L2* in the NPCs of the daughter compared to the NPCs of her mother. Of note, *SEZ6L2* was upregulated almost 100-fold. *ALDOA*, *DOC2A* and *POGZ* showed no significant change of expression.

The neurons of the daughter showed a significantly changed expression for all genes measured compared to the neurons generated from her mother. Expression of *KIF22*, *MAPK3*, *CDIPT*, *DOC2A*, *SEZ6L2* and even *POGZ* was significantly increased in neurons of the daughter. Again, *SEZ6L2* showed a very high increase in expression of ~50-fold. Surprisingly, expression of *ALDOA* was significantly decreased in the neurons of the daughter.


A
Mother 1179/17
Daughter 1180/17
fold change of expression
*
*
*
KIF22
MAPK3
ALDOA
CDIPT
POGZ
Fibroblasts
B
Mother 1179/17 A2
Daughter 1180/17 B8
fold change of expression
KIF22
MAPK3
ALDOA
CDIPT
DOC2A
SEZ6L2
POGZ
iPSCs
C
Mother 1179/17 A27
Daughter 1180/17 B8
fold change of expression
***
**
**
**
**
KIF22
MAPK3
ALDOA
CDIPT
DOC2A
SEZ6L2
iPSCs

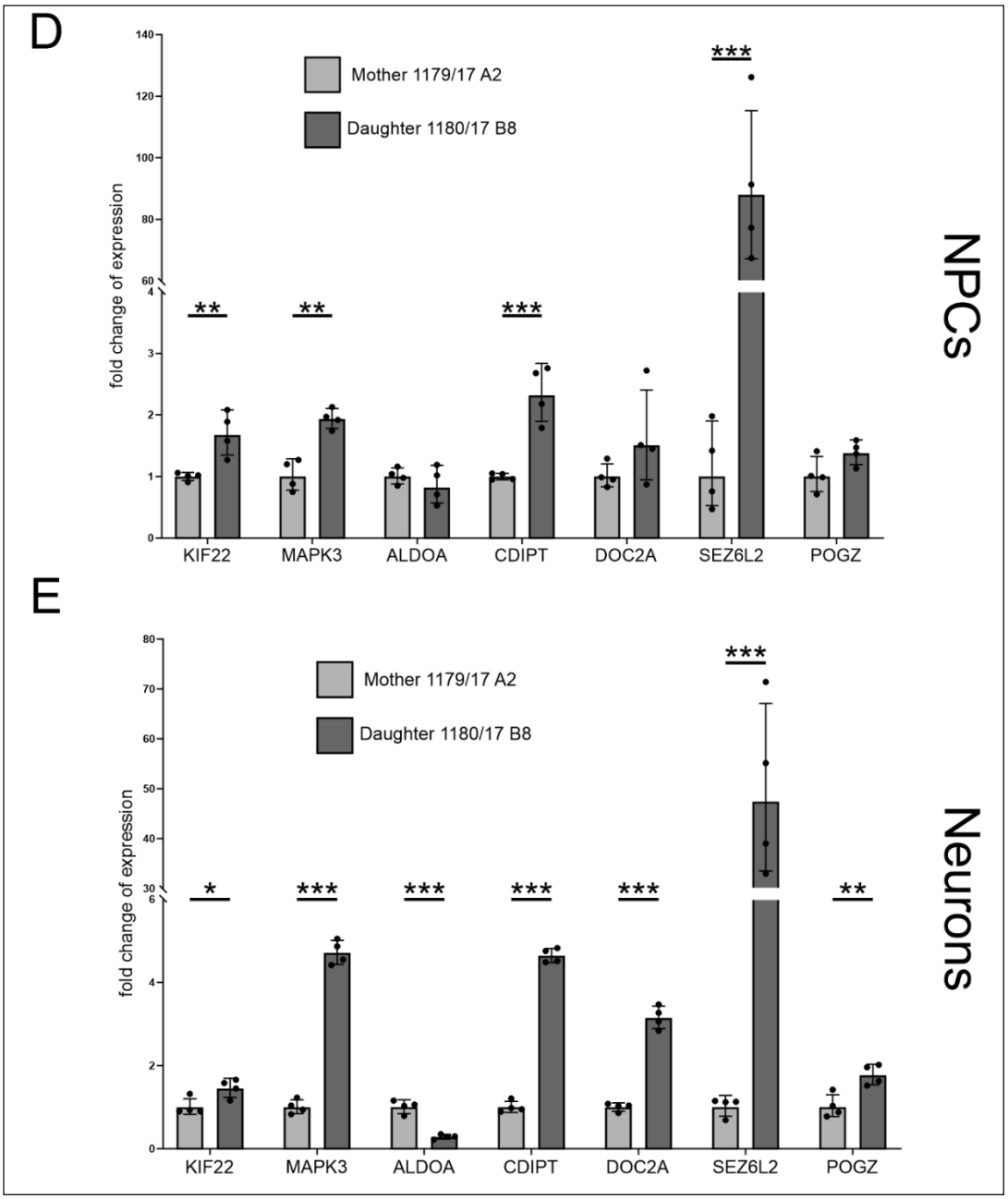


Figure 32 Gene expression analysis of duplicated genes measured by RT-qPCR. A: Fibroblasts. B: iPSCs (A2 and B8). C: iPSCs (A27 and B8). D: NPCs. E: Neurons. GAPDH was used for normalisation. All values are geometric mean ± geometric SD. *P<0,05; **P<0,01; ***P<0,001. Significances were calculated using multiple t-tests. Exact P-values are given in **Table 23**.

Table 23 P-values of multiple t-tests. Multiple t-tests were performed to calculate significant changes in gene expression between different cell types of mother (1179/17) and daughter (1180/17).

	P-value						
	KIF22	MAPK3	ALDOA	CDIPT	DOC2A	SEZ6L2	POGZ
Fibroblasts	0,349708	0,024228	0,010064	0,11844			0,162361
iPSCs (A2+B8)	0,372335	0,319916	0,365069	0,162228	0,184343	0,286400	0,343447
iPSCs (A27+B8)	0,000762	0,005969	0,001107	0,002346	0,007944	0,119213	
NPCs	0,003757	0,002520	0,352425	0,000193	0,154833	0,000014	0,087985
Neurons	0,022891	0,000002	0,000072	<0,000001	0,000002	0,000002	0,008098

3.3 Characterization and Generation of isogenic controls

During reprogramming the cellular DNA can be altered, directly by mutations or indirectly by changes in the methylome. To avoid measuring effects based on these alterations, isogenic controls are required (1.1.4). They were generated using two different approaches. The female heterozygous iPSCs were selected using the expression pattern of X-chromosomal genes, while the male wildtype iPSCs were mutated using CRISPR/Cas9 genome editing.

3.3.1 Cell selection of female heterozygous iPSCs

The female iPSCs were partially characterized during a preceding Master Thesis (Käseberg, 2015) and were known to carry a heterozygous MID1 mutation (c.1801_1804delCTCC) that was already present in the fibroblasts before the reprogramming. In the beginning, the female heterozygous iPSCs were randomly selected and cultured. After reaching a stable state, the cells were screened regarding their *MID1* expression pattern. For this we used two different methods, an allele-specific RT-PCR with primers detecting either the wildtype or the mutant *MID1* mRNA, and a Western Blot with a MID1 antibody detecting only the wildtype protein. Clones expressing the wildtype (ctrl M10) or the mutant (OS M7) allele of *MID1* respectively, were identified (Figure 33).

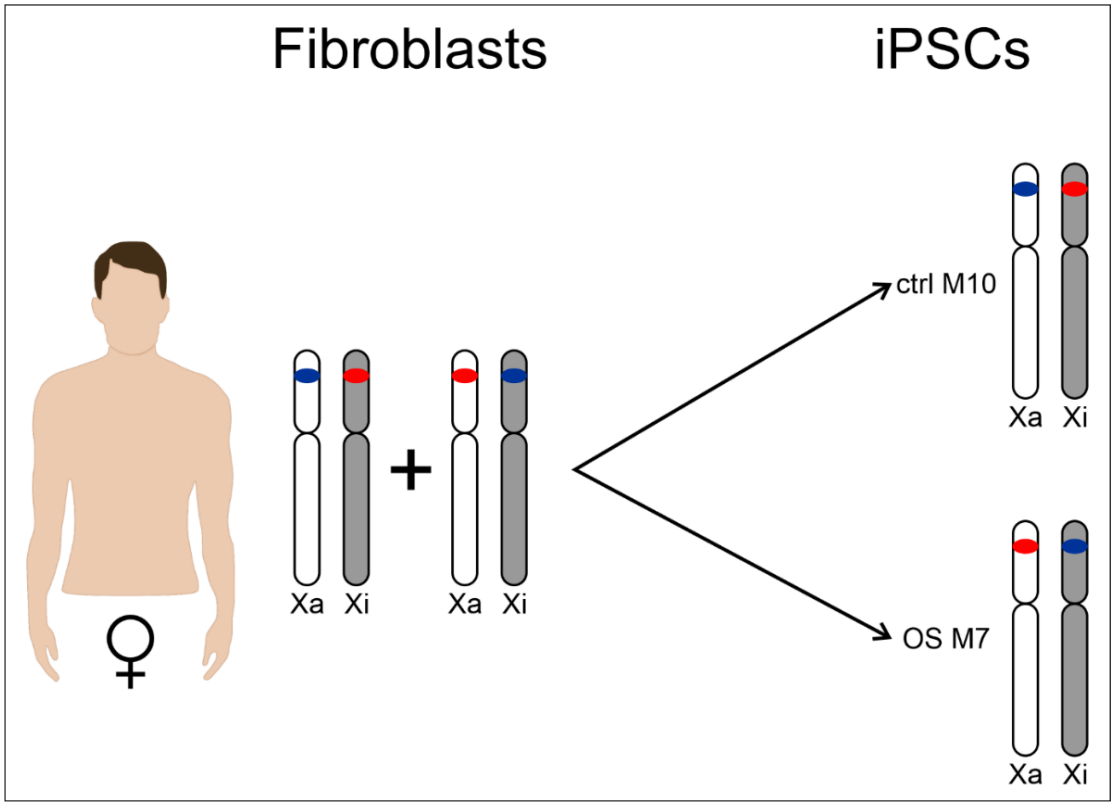


Figure 33 Cell selection of iPSC clones to generate isogenic controls. While female fibroblasts are a mosaic of cells expressing either the wildtype or the mutant allele, iPSC clones show monoallelic expression of one of the alleles.

3.3.1.1 RNA expression of selected female heterozygous iPSCs

An allele-specific RT-PCR showed expression of wildtype *MID1* mRNA in ctrl M10 iPSCs, while mutant *MID1* mRNA was detected in the OS M7 iPSCs (Figure 34). The fibroblasts were used as controls to show the specificity of the PCR. Fibroblasts of the male WT and the male OS can only show a band in one of the two reactions due to the hemizygotic state of the X-chromosome. The female fibroblasts on the other hand, have two X-chromosomes and are known to be heterozygous for the assayed mutation.

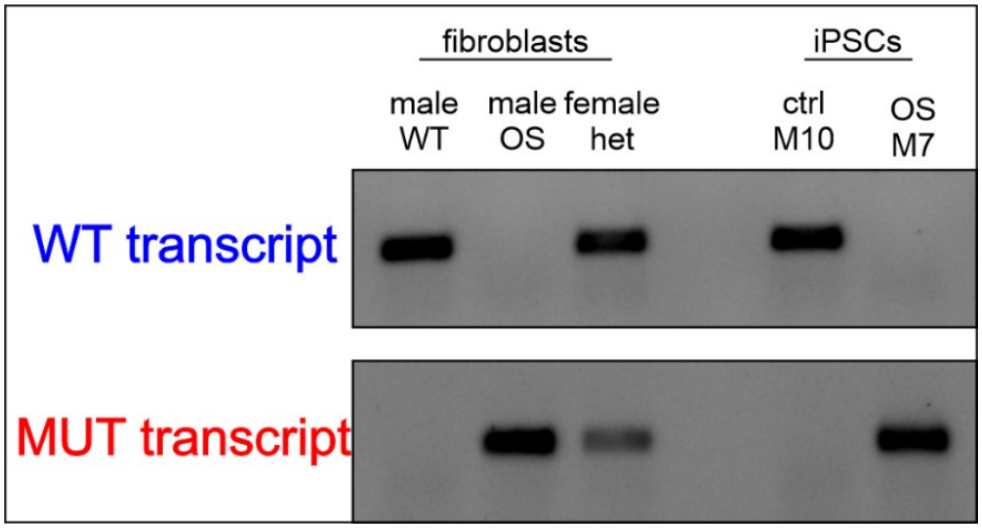


Figure 34 Allele-specific RT-PCR of iPSC clones to identify isogenic controls. Female iPSCs show expression of either the wildtype or the mutant allele of *MID1*. Fibroblasts were used as controls.

3.3.1.2 Protein expression of female heterozygous iPSCs

Monoallelic expression of *MID1* found in the ctrl M10 and OS M7 iPSCs was further confirmed by a Western Blot using an antibody targeting the N-terminal end of MID1 (Figure 35). iPSCs of C1 and the hemizygous OS M14 were used as controls. A band at the size of 75 kDa was detected in all four samples, with the bands in ctrl M10 and C1 being stronger than in the other two samples. While the two bands detected in ctrl M10 and C1 represent the wildtype MID1, the bands detected in OS M7 and OS M14 are presumably unspecific cross reactions. The mutated MID1 mRNA in those cells would result in a C-terminally truncated MID1 protein that is approximately 5 kDa shorter than wildtype MID1, a size difference that is visible on the Western Blot. Though, it remains unclear if the mutated MID1 mRNA is translated at all.

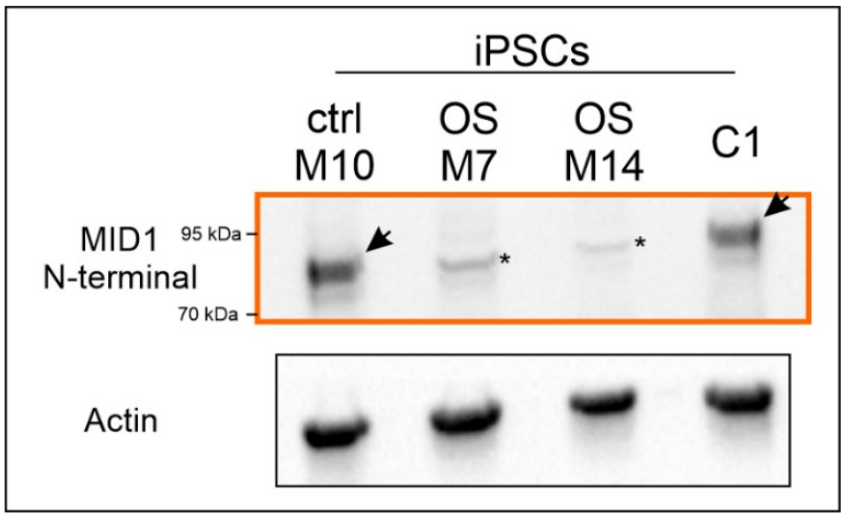


Figure 35 Western Blot with an antibody detecting the N-terminal part of MID1 to identify isogenic controls. C1 and OS M14 iPSCs were used as controls. Actin was used as a loading control. Arrows mark specific bands; asterisks mark unspecific bands.

3.3.1.3 mTOR dysregulation in female iPSCs carrying a heterozygous mutation in the *MID1* gene

For final characterization of the generated isogenic controls ctrl M10 and OS M7, we performed a Western Blot to detect changes in S6 phosphorylation, a well described target of mTOR known to be dysregulated in the fibroblasts of the male OS cell line. In line, OS M7 iPSCs showed a significantly reduced phosphorylation of S6 compared to ctrl M10 iPSCs (Figure 36).

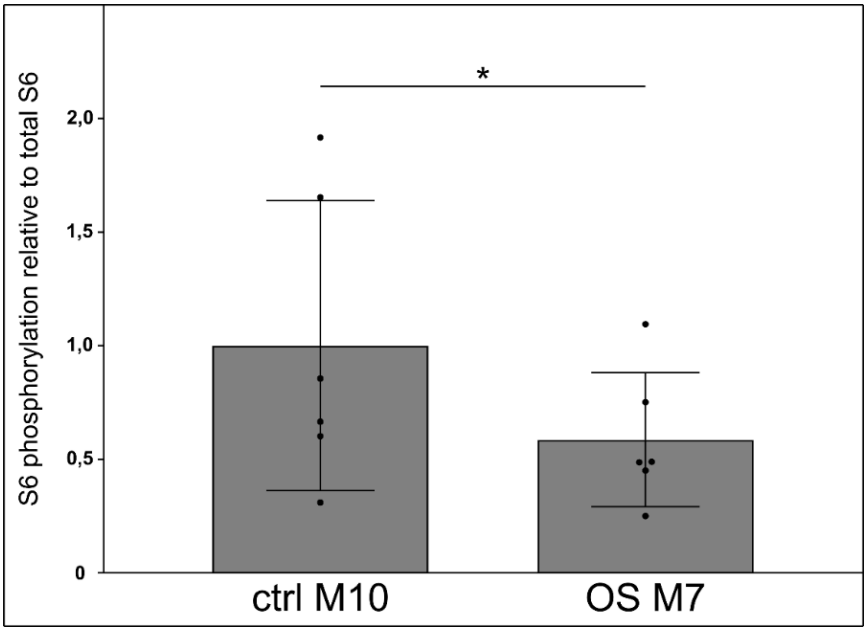


Figure 36 OS M7 iPSCs show a reduced phosphorylation of the mTOR target S6. Western Blots were incubated with antibodies specific for S6 and phosphorylated S6. Band intensities were quantified and relative S6 phosphorylation was calculated. Ratio paired t-test: p=0,0406. Values are mean ± SD.

Taken together, our data shows that we have successfully generated isogenic controls from a female heterozygous carrier of a *MID1* mutation that stably express either the wildtype or the mutant allele, respectively. They also show a dysregulated mTOR signalling that has previously been associated with OS and proves that the cells are indeed a suitable model for OS.

3.3.2 CRISPR/Cas9 of male wildtype iPSCs

The C1 male iPSCs were selected for CRISPR/Cas9 genome editing after reaching a stable state. The iPSCs were electroporated with two different gRNA plasmids, each in combination with a Cas9 plasmid. After FACS sorting and expansion, the cells were screened regarding *MID1* expression and their pluripotent state, as previously described (3.1.2).

3.3.2.1 DNA sequencing of edited male wildtype iPSCs

Sequencing of the first coding exon of *MID1* (Figure 37, Figure 38) revealed a 1 bp insertion in the R1 iPSCs (c.136_137insC) and a 2 bp deletion in the R2 iPSCs (c.204_205delAG). Both mutations are located in the region coding for the RING-Finger subdomain of the MID1 protein and are predicted to cause a frameshift, leading to an early stop codon. Sequencing was performed by [redacted] as a part of her Master Thesis.

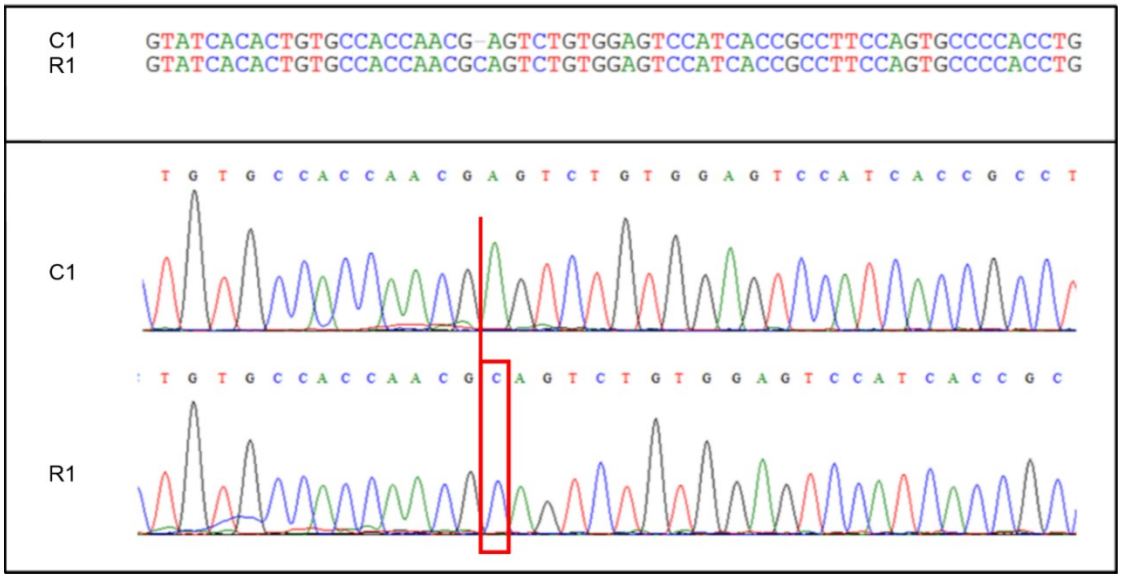


Figure 37 Electropherogram showing the sequencing result of R1 for a chosen area of MID1 coding Exon 1. The red box marks the peak showing the insertion in the R1 iPSCs. C1 DNA shows the wildtype sequence. (Sequencing performed by [redacted]).

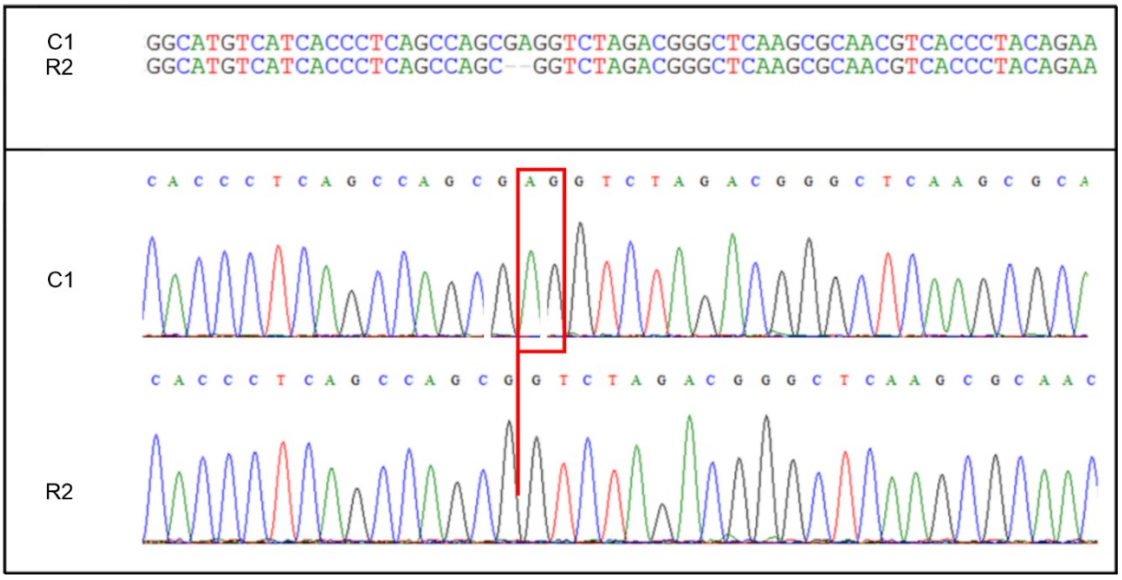


Figure 38 Electropherogram showing the sequencing result of R2 for a chosen area of MID1 coding Exon 1. The red box marks the peaks showing the deletion in R2 iPSCs. C1 DNA shows the wildtype sequence. (Sequencing performed by [redacted]).

3.3.2.2 Off-target analysis of edited male iPSCs

To exclude any mutations induced by the genome editing outside of MID1, we performed Whole Exome Sequencing (WES) with the C1, R1 and R2 iPSCs. All three samples showed a similar number of variants and when checking the predicted off targets (*Guide Design Resources — Zhang Lab*, n.d.) with up to 5 mismatches for the used gRNAs, no differences between the C1, R1 and R2 iPSCs were found.

3.3.2.3 Protein expression of edited male iPSCs

The induced frameshift in the R1 and R2 iPSCs was predicted to cause an early stop codon (p.46GlufsX107 and p.69GlyfsX106 respectively). To check if really no protein was translated, we performed a Western Blot targeting MID1 using two different antibodies. The previously used N-terminal antibody binds to amino acids C-terminal of the frameshift; thus, we expected no specific binding of this antibody in R1 and R2 protein lysates. The C-terminal antibody binds to amino acids coded by the last exon; again, we did not expect any specific binding in R1 and R2 protein lysates.

Using the N-terminal antibody, we were able to detect wildtype MID1 in C1 iPSCs at 75 kDa and an unspecific binding at the same size in R1 and R2 iPSCs (Figure 39) similar to what we found in the female iPSCs of ctrl M10 and OS M7 (3.3.1.2). In the R2 iPSCs an additional band at 69 kDa appeared. When using the C-terminal antibody, the outcome was comparable. We again detected the wildtype MID1 in C1 iPSCs at 75 kDa but not in the R1 and R2 iPSCs. The

additional band at 69 kDa in the R2 iPSCs was again visible, together with more bands between 50 and 70 kDa in all three samples. These additional bands can either be unspecific cross reactions or shorter MID1 isoforms lacking the N-terminal end of the protein.

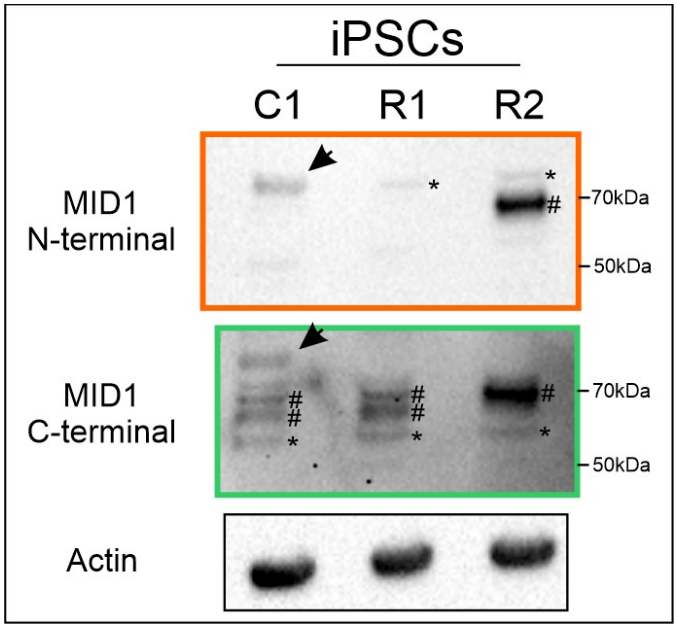


Figure 39 Western Blot with antibodies detecting the N- or C-terminal part of MID1 to confirm the knockout of the gene in R1 and R2 iPSCs. Actin was used as a loading control. Arrows mark specific bands; asterisks mark unspecific bands; diamonds mark potentially shorter MID1 proteins that lack the RING-finger.

3.3.2.4 MID1 alternative ATGs

After finding unexpected bands in the Western Blot of C1, R1 and especially R2 iPSCs using the MID1 antibodies, we had a look at the sequence of the first coding exon of MID1 to find an explanation (Figure 40). In addition to the wildtype ATG (green), we identified three alternative ATGs (purple) downstream that are in frame and were predicted to translate into N-terminally truncated proteins of 64, 58 and 57 kDa respectively. In addition, we identified a fourth alternative ATG (brown) frame that is being rescued by the 2 bp deletion (red) in R2, translating into a protein of 69 kDa. This is also the only alternative ATG that can result in a protein detectable by the N-terminal antibody (binding site between spaces in line 6 and 7).

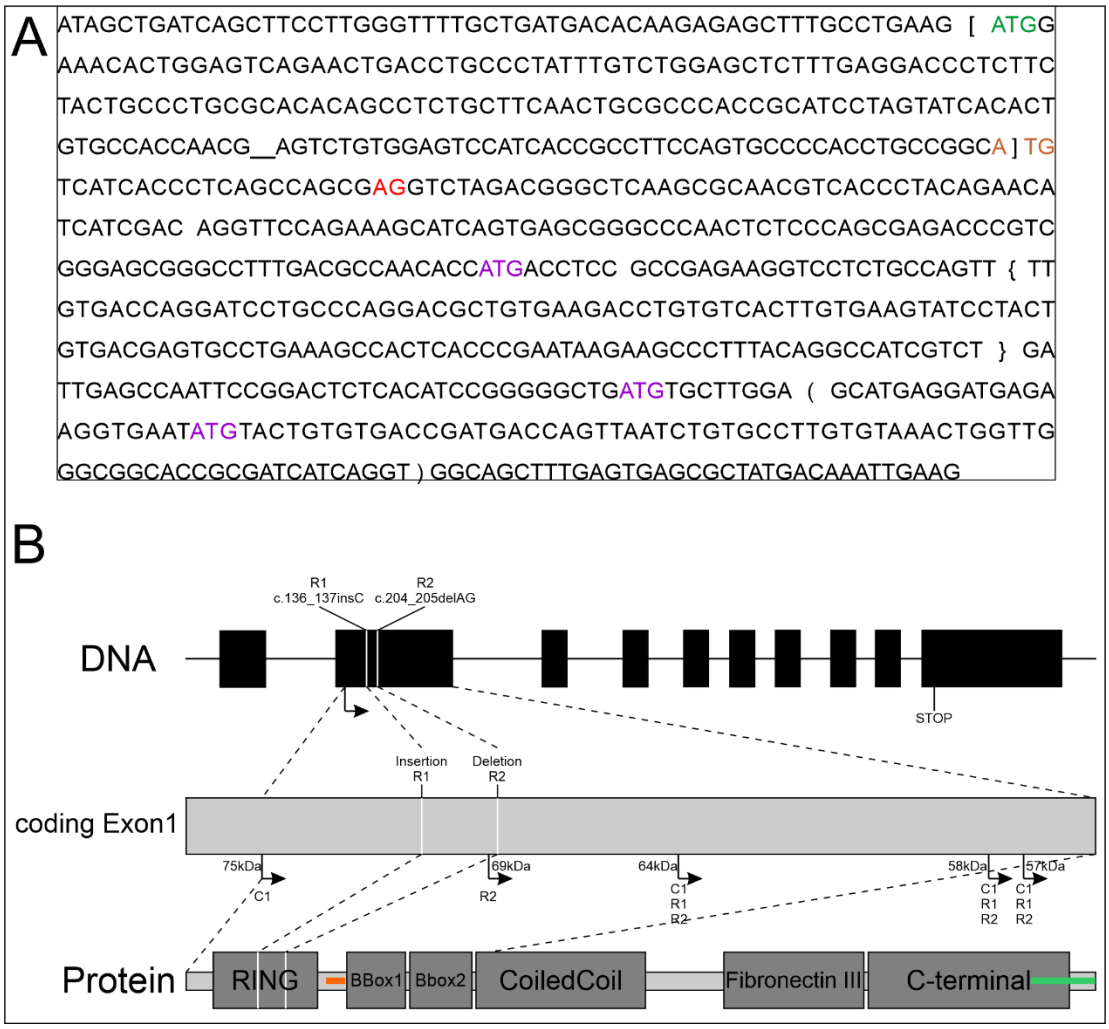


Figure 40 Alternative start sites for translation in the first coding Exon of MID1. A: Sequence representation of alternative ATGs. Green letters mark the original start codon. Purple letters mark alternative start codons within the ORF. Brown letters mark an alternative start codon whose ORF is being rescued by the deletion of R2. Red letters mark the deletion in R2. __ marks the insertion in R1. [ ] marks the RING-finger. { } marks the BBox1. ( ) marks the BBox2. Spaces in line 6 and 7 mark the binding site of the N-terminal MID1 antibody. B: Schematic representation of MID1 DNA, coding Exon1 and protein. Positions of the mutations induced in R1 and R2 are marked. Small arrows indicate identified ATGs. Predicted protein sizes and cell lines that have an open reading frame are depicted next to the corresponding arrows. Orange box indicates the binding position of the N-terminal antibody. Green box indicates the binding position of the C-terminal antibody.

In a next step, we used three different online tools to predict the probability of the identified ATGs to be used. As the likelihood of ATGs used as translational start sites depends on many factors, including the genotype, we simulated usage of the ATGs under all three genotypes (Table 24). All three tools gave comparable results for the different ATGs. The wildtype ATG achieved medium scores, showing that it does not require a high score for an ATG to be used. The second ATG frame of which gets rescued by the deletion induced in R2, showed overall

high results only with the rescuing mutation. The third ATG is assigned the highest scores on average, while ATG four and five are assigned the lowest scores on average.

Taken together, this shows that it is not only the wildtype ATG that is likely to be used, but also other ATGs within the ORF, further supporting our hypothesis of N-terminally truncated MID1 proteins.

Table 24 ATG usage scores in order of appearance. Three different online tools (*NetStart 1.0 Prediction Server*, n.d.; *Prediction of Translation Initiation ATG*, n.d.; *Prediction of Translation Initiation Site*, n.d.) were used to determine the quality of the different ATGs in the first coding exon of MID1. ATGPR: Reliability score 0-1; NetStart: 0-1000; DNATIS Miner: 0-1000.

	ATGPR			NetStart			DNATIS Miner		
ATG	C1	R1	R2	C1	R1	R2	C1	R1	R2
1 (WT)	0,47	0,33	0,41	640			408		
2	0,04	0,04	0,63	579	491	793	139	79	853
3	0,47	0,79	0,47	798			724		
4	0,04			355			594		
5	0,04			677			280		

3.3.3 MID1 mRNA stability in wildtype and mutant cells

The previous results of the generated isogenic cell lines, ctrl M10 and OS M7 on the one and especially C1, R1 and R2 on the other hand, point towards a more complex regulation of *MID1* translation as previously known. To confirm that in both cases MID1 mRNA is translated and does not undergo NMD the *MID1* mRNA was measured by RT-qPCR (Figure 41). We used three different primer pairs amplifying different regions of the *MID1* mRNA. All five iPSC lines that were measured showed a similar *MID1* expression profile, neither the OS M7, nor the R1 and R2 iPSCs showed any signs of a reduced *MID1* expression. We also measured *MID1* expression in NPCs differentiated from the five iPS lines and again none of the cell lines showed a significantly reduced *MID1* expression. Interestingly, the female NPCs showed an elevated *MID1* expression compared to the male NPCs.

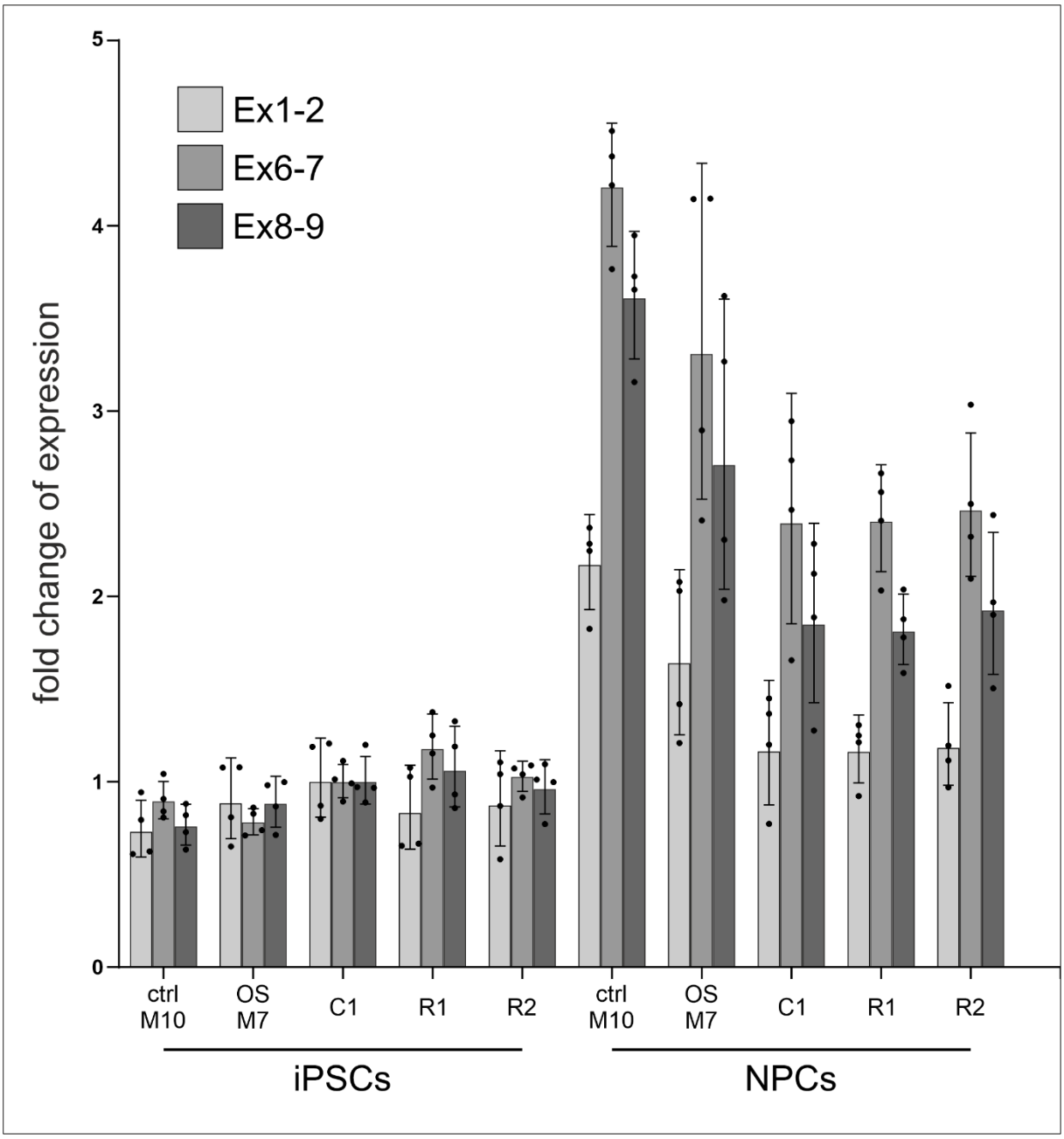


Figure 41 MID1 mRNA is not degraded in iPSCs and NPCs. RT-qPCR for different Exons of *MID1* showed, that mRNA levels do not differ between isogenic controls independent of their cell type. NPCs showed and overall increased expression of *MID1* mRNA compared to iPSCs. Values are geometric mean ± geometric SD.

3.3.4 MID1 wildtype and mutant protein localisation

As all our previous data point towards MID1 mRNA being translated independently of the mutations of our specific cell lines, we generated constructs overexpressing MID1 with an N-terminal GFP tag. In addition to the wildtype MID1, that was already present in the lab, we also generated plasmids expressing N-terminally truncated MID1 starting at the different ATGs we previously identified to be potentially used. We also used previously generated plasmids expressing MID1 carrying the 4 bp deletion found in the male OS and the female heterozygous patients, as well as two plasmids expressing N-terminally truncated MID1 lacking the RING-finger or the RING-finger and both BBoxes (Figure 42 A).

As previously shown (Schweiger et al., 1999), overexpressed wildtype MID1 showed a clear microtubule binding. Interestingly, the same was observed for all N-terminally truncated MID1 proteins except for the protein lacking the RING-finger and both BBoxes (Figure 42 B+C). Overexpression of this protein showed no clear microtubule binding but a diffuse distribution in the cytoplasm. The C-terminally truncated MID1 protein with the 4 bp deletion did also not show a microtubule binding, but, as previously reported (Schweiger et al., 1999), formed aggregates in the cytoplasm. Of note, the two MID1 isoforms showing no microtubule binding are both caused by patient-specific mutations, while the mutated isoforms showing microtubule binding have never been found in patients. The 4 bp deletion is deeply researched in this thesis and the patients reported with this mutation showed a clear OS phenotype (Liu et al., 2011). The deletion of the RING-finger and both BBoxes is highly comparable to a full deletion of the first coding exon of MID1, a mutation that was found in OS patients and has previously been described in the literature (Winter et al., 2016).

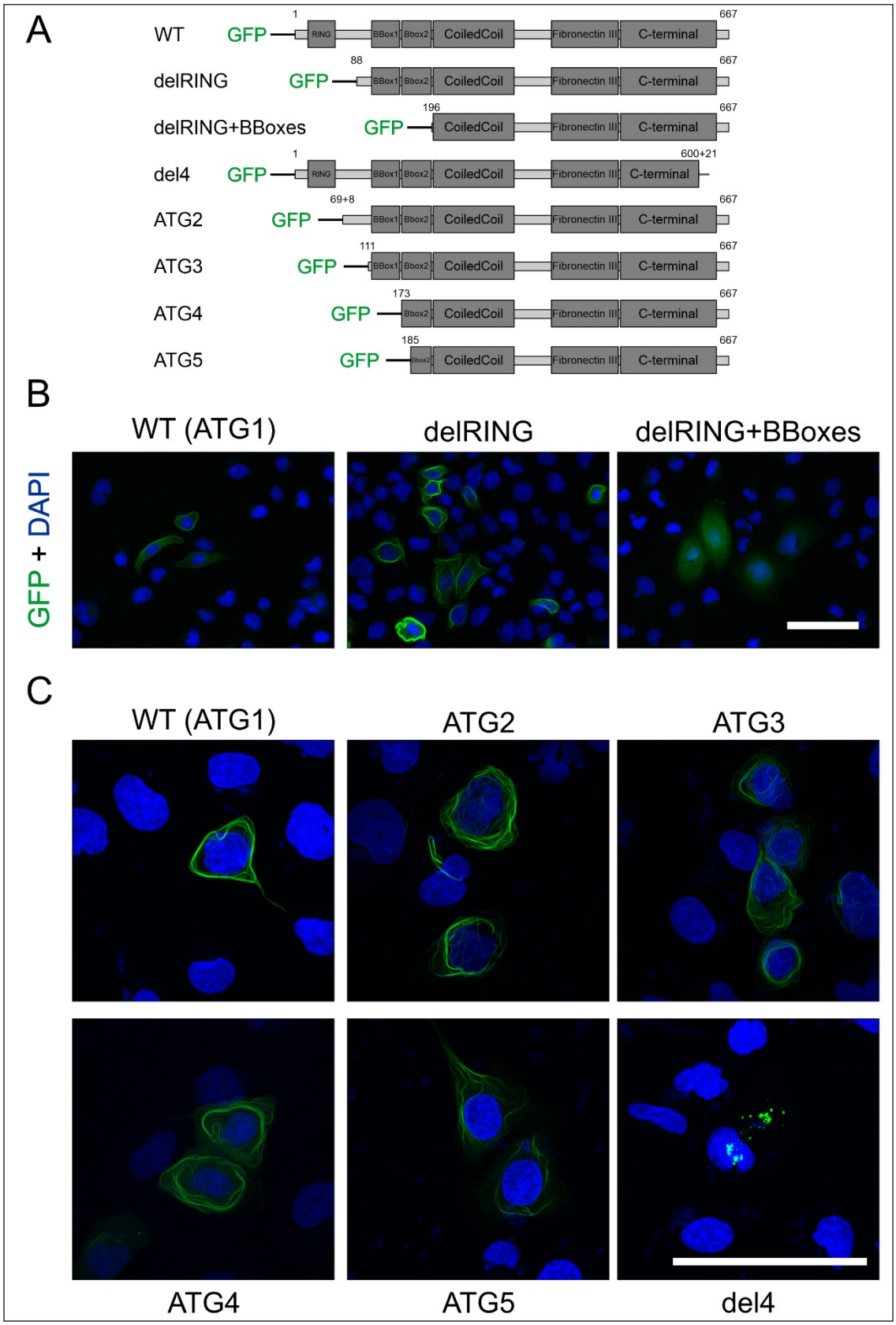


Figure 42 Cellular localization of wildtype and truncated MID1 proteins. A: Different MID1 proteins were overexpressed in HeLa cells using the eGFP-C1 vector with an N-terminally fused GFP. Numbers indicate the first and last amino present. B: Immunofluorescent images of HeLa cells after overexpression of GFP-MID1 WT, delRING and delRING+BBoxes. C: Immunofluorescent images of HeLa cells after overexpression of GFP-MID1 WT, ATG2-5 and del4. Scale bars: 100 µm.

3.4 Cerebral organoids

In collaboration with Prof. Dr. [redacted], Prof. Dr. [redacted], Dr. [redacted] and [redacted], cerebral organoids of the female and male iPSCs were generated. These data were generated by Dr. [redacted] and [redacted] and will be shown for completeness and to explain why further experiments were performed as a part of this thesis.

3.4.1 Female organoids (M-Lines)

In addition to the previously mentioned iPSC clones of 16/98 (ctrl M10 and OS M7) (Table 12), two other iPSC clones of 16/98 were used for generating cerebral organoids. OS M19, like OS M7 a clone that expresses the mutant MID1 from the active X-chromosome and ctrl M16, that like ctrl M10 expresses wildtype MID1. Furthermore, a totally unrelated control cell line was used (ctrl nr). The cerebral organoids generated from the ctrl iPSCs (nr, M10, M16) showed a significantly smaller ventricular area in comparison to the cerebral organoids generated from the OS iPSCs (M7, M19), while the total number of ventricles per organoid was unchanged (Figure 43). This phenotype can either be caused by an increased cell size or an increase in the number of cells of the ventricular zone in OS cerebral organoids.

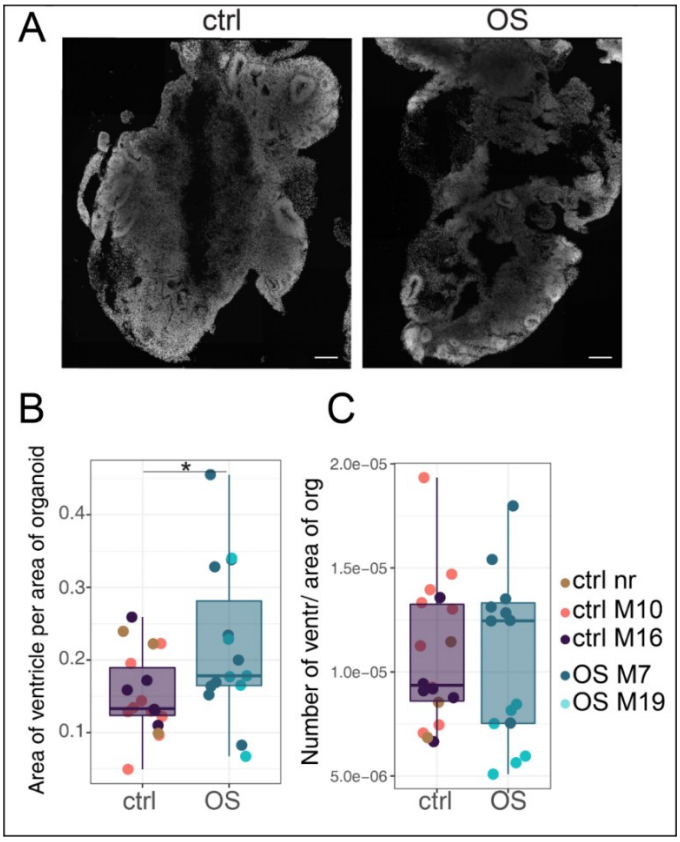


Figure 43 Overview of M-Line cerebral organoids. A Representative cryostat cross sections of cerebral organoids generated from ctrl and OS iPSCs of 16/98. B OS cerebral organoids show a significantly increased area of ventricles in comparison to ctrl organoids. C The total number of ventricles is unchanged. Values are mean ± SD. Scale bars: 50 µM.

As the ventricular zones of cerebral organoids are comprised of SOX2 and PAX6 positive neuronal stem cells, cerebral organoids were cut in slices and stained for PAX6 and the neuronal marker TUBB3. Counting of PAX6 positive cells revealed a significantly increased number of PAX6 positive cells in OS cerebral organoids (Figure 44 A + B). These results were further supported by measuring the thickness of the ventricular zones and the neuronal layer. In line, OS cerebral organoids showed an increased thickness of the ventricular zone and a reduced thickness of the neuronal layer (Figure 44 C + D). All together this data suggests that the OS cerebral organoids are comprised of an increased number of neuronal stem cells and a reduced number of differentiated neurons. Further experiments revealed that this phenotype is caused by a failure of the neuronal stem cells to exit the cell cycle (R. Menon, 2019).

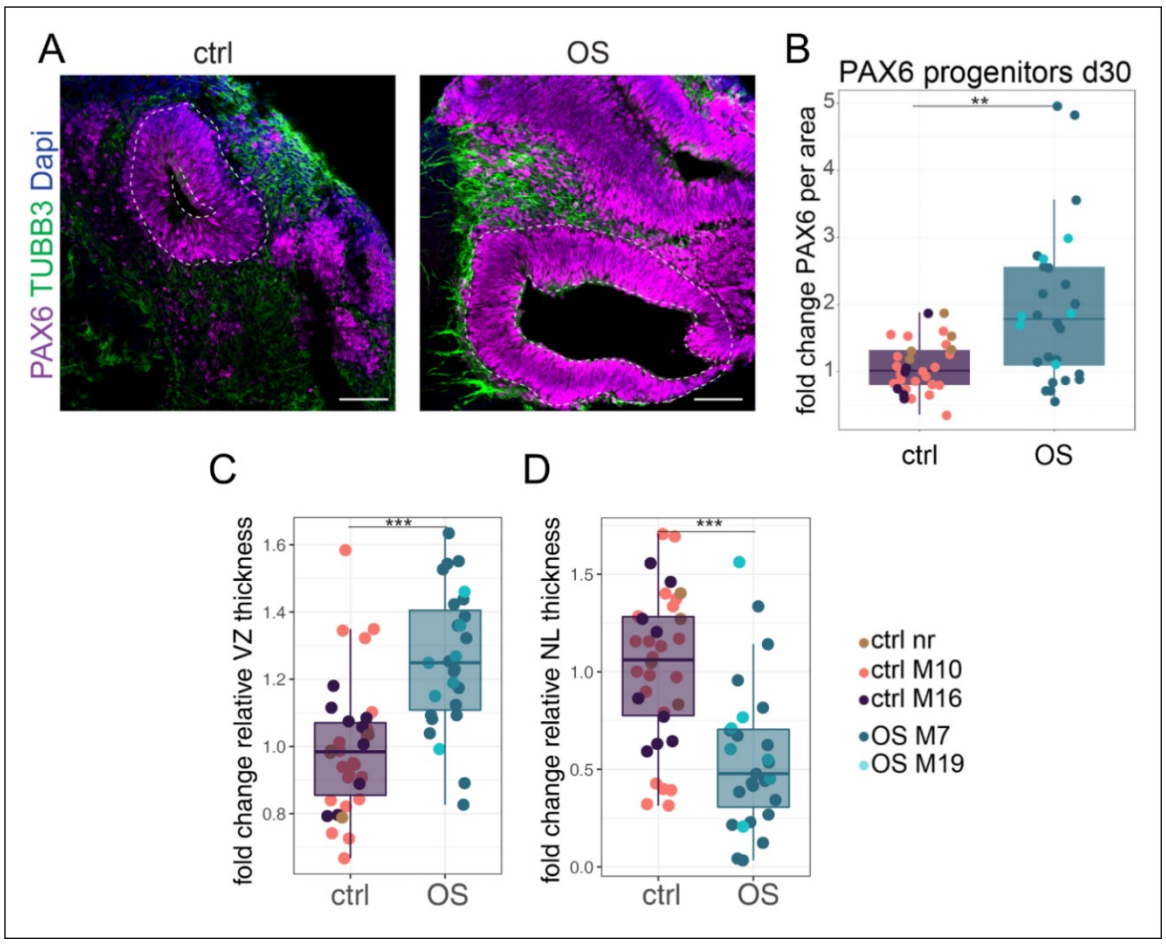


Figure 44 Differences in PAX6 expression of cerebral organoids generated from ctrl M10/M16 and OS M7/M19. A Representative cryostat cross sections of cerebral organoids generated from ctrl and OS iPSCs of 16/98 after immunofluorescent staining against neuronal marker TUBB3 and neuronal stem cell marker PAX6. B OS organoids show a significantly increased area of PAX6 positive cells in comparison to ctrl organoids. C OS organoids show an increased thickness of the ventricular zone in comparison to ctrl organoids. D OS organoids show a reduced thickness of the neuronal layer in comparison to ctrl organoids. Values are mean ± SD. Scale bars: 50 µM.

Interestingly, when comparing the number of PAX6 positive cells in ctrl and OS cerebral organoids of 16/98 with OS cerebral organoids of a male patient (17/98), a gradual phenotype was observed. The female OS organoids showed a reduced number of PAX6 positive cells in comparison to the male OS cerebral organoids but an increase number of PAX6 positive cells compared to the female ctrl cerebral organoids (Figure 45 A + B). This intermediate phenotype might point towards a rescue by the wildtype allele of the female OS cerebral organoids.

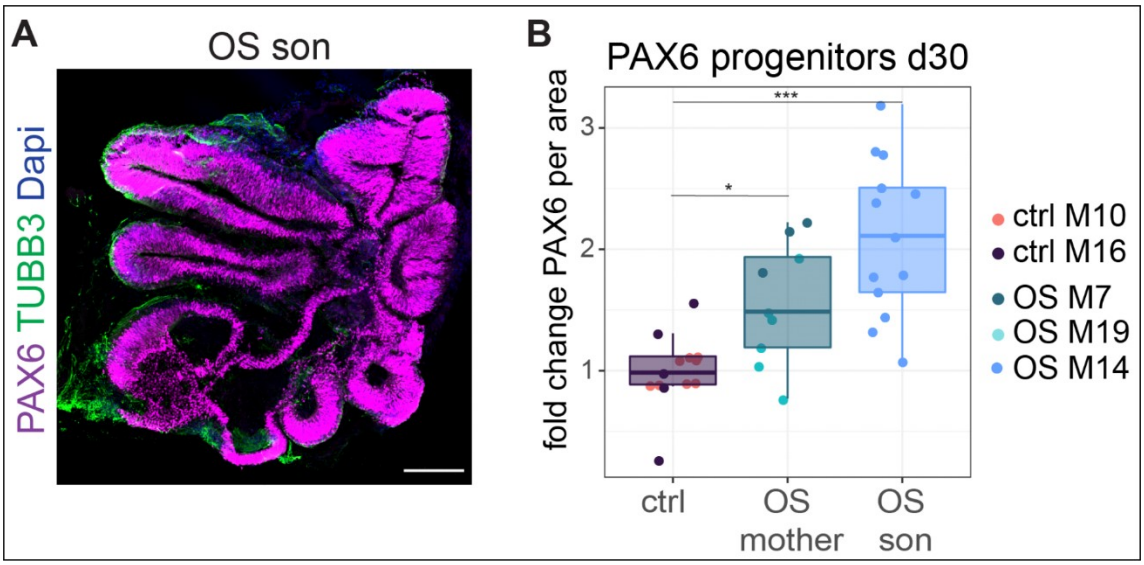


Figure 45 Differences in PAX6 expression in male OS organoids. A Representative cryostat cross section of cerebral organoids generated from OS iPSCs of 17/98 after immunofluorescent staining against neuronal marker TUBB3 and neuronal stem cell marker PAX6. B OS organoids of the son (17/98) show a significantly increased area of PAX6 positive cells in comparison to ctrl organoids and an increased number area of PAX6 positive cells in comparison to OS organoids of the mother. Values are mean ± SD. Scale bar: 50 µM.

3.4.2 Male organoids

After the full characterization of the R1 and R2 iPSCs, the iPSCs were used for generating cerebral organoids and analysed using the same approach as used for the M-lines. While the C1 cerebral organoids developed normally and were comparable to the ctrl M10 organoids, the ones generated from the R1 and R2 iPSCs showed a completely different morphology in comparison to the female and male OS organoids (Figure 46 A). Both showed a reduced number of ventricles and a reduced area of the ventricular zones in comparison to C1 cerebral organoids in a gradual way, with R1 having a more severe phenotype than R2 (Figure 46 B + C). This, together with the previously observed alternative ATGs being used for translation, points towards a differing mechanism underlying phenotypes associated to N-terminally and C-terminally mutated MID1.

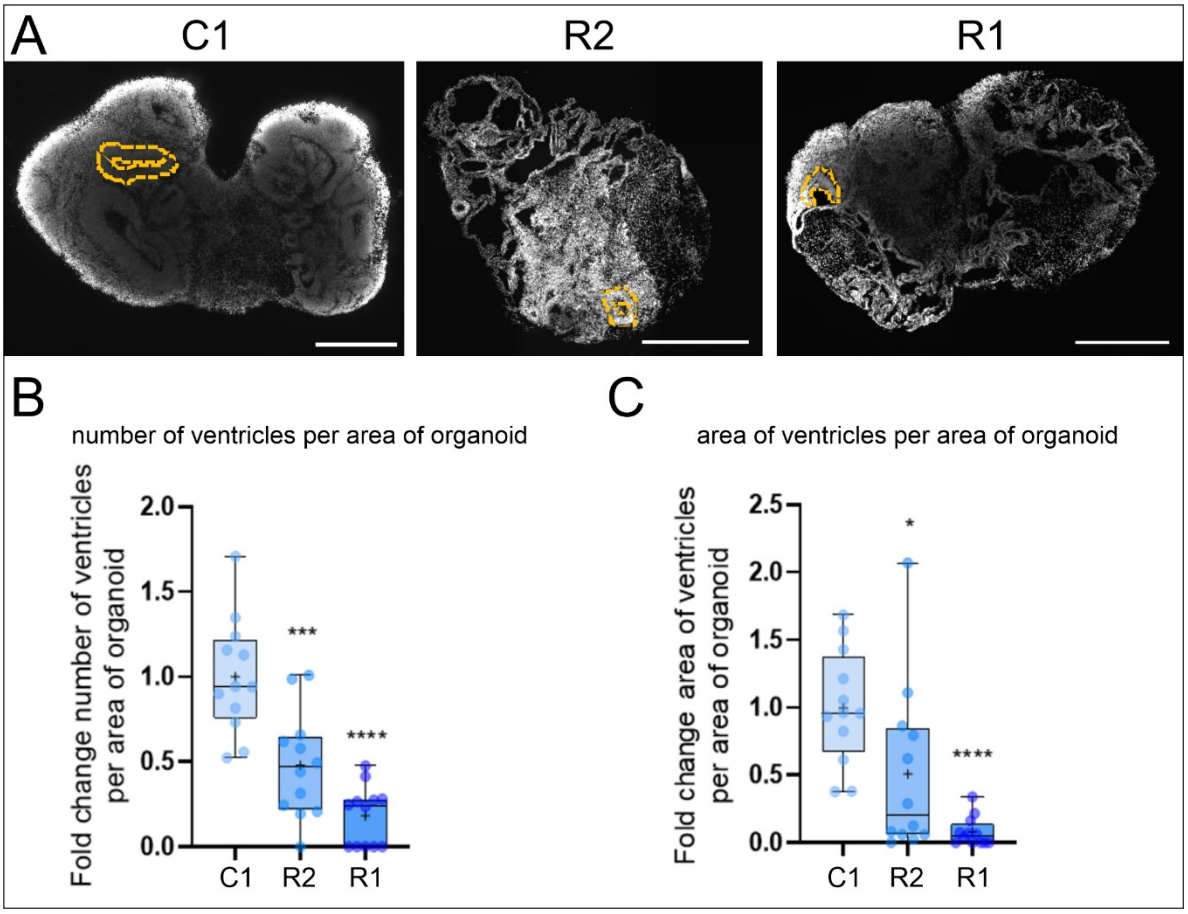


Figure 46 Overview of S-Line cerebral organoids. A Representative cryostat cross sections of cerebral organoids generated from S1, R2 and R1 iPSCs. B R1 and R2 cerebral organoids show a significantly reduced number of ventricles in comparison to C1 organoids. C R1 and R2 cerebral organoids show a significantly reduced area of ventricles in comparison to C1 organoids. Values are mean ± SD. Scale bars: 500 µM.

In a next step, slices of the cerebral organoids were stained for the neuronal markers TUBB3 or MAP2 and the neuronal stem cell markers PAX6 or SOX2, as before. In concordance with the previous results, C1 cerebral organoids showed a typical staining of ventricular zones positive for PAX6 or SOX2 surrounded by neuronal layer regions positive for TUBB3 or MAP2. R1 and R2 cerebral organoids on the other hand, showed only few ventricular regions positive for PAX6 and SOX2 and only parts of the remaining organoid stained positive for TUBB3 and MAP 2 (Figure 47). Big areas in R1 and R2 cerebral organoids showed no clear staining of either the neuronal stem cell markers or the neuronal markers. Again, a gradual phenotype between C1, R2 and R1 was observed.

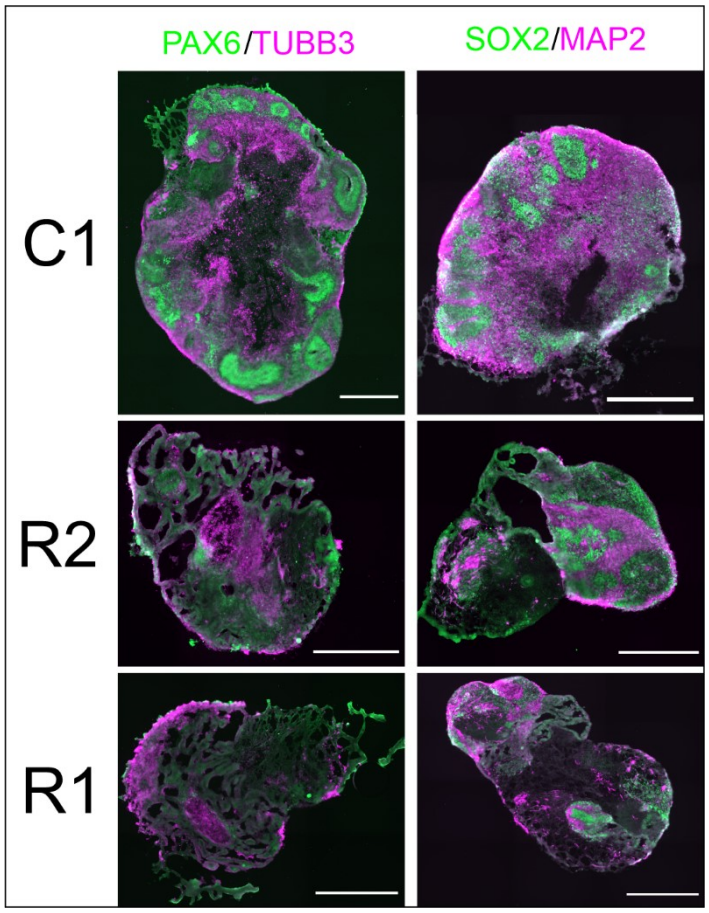


Figure 47 Differences in neuronal (stem cell) marker expression in S-line cerebral organoids. Representative cryostat cross sections of cerebral organoids generated from C1, R1 and R2 iPSCs after immunofluorescent staining against the neuronal markers TUBB3 or MAP2 and the neuronal stem cell markers PAX6 or SOX2. Values are mean ± SD. Scale bars: 500 µM.

The morphology of the unstained parts in the R1 and R2 cerebral organoids resembled that of the choroid plexus, a tissue producing the cerebrospinal fluid in the brain and also found in cerebral organoids, though at very low levels. Slices of the cerebral organoids of C1, R1 and R2 were then stained for the choroid plexus marker TTR. Quantification of the expression revealed an increased area of R1 and R2 cerebral organoids staining positive for TTR in comparison to C1 organoids (Figure 48). During brain development, the choroid plexus develops in the dorsal part of the brain induced by a low concentration of SHH. As SHH-signalling has been found to be regulated by MID1, we hypothesized that this hyperdorsalization phenotype is caused by an altered SHH signalling because of the mutations induced in R1 and R2 iPSCs.

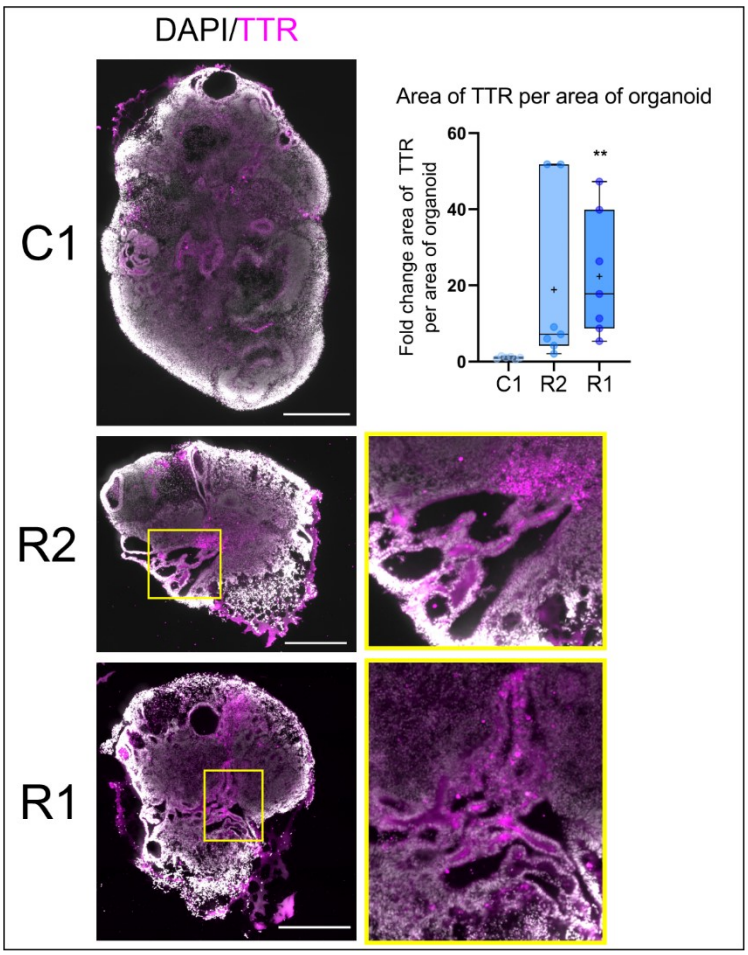


Figure 48 Hyperdorsalization of S-line cerebral organoids. Representative cryostat cross sections of cerebral organoids generated from C1, R1 and R2 iPSCs after immunofluorescent staining against the choroid plexus marker TTR. Quantification of TTR positive regions showed a significantly increased area of TTR in R1 and R2 cerebral organoids. Values are mean ± SD. Scale bars: 500 µM.

The following data has been generated by the candidate and is again part of the thesis.

3.4.3 PTCH1 mRNA expression

To confirm that the increased amount of choroid plexus tissue in the cerebral organoids of R1 and R2 is indeed caused by an altered SHH signalling we generated NPCs and neurons of C1, R1 and R2 iPSCs to quantify expression of the SHH receptor PTCH1. RT-qPCR showed a significantly decreased expression of *PTCH1* in NPCs of R1 and R2 compared to NPCs of C1, but no changes in the other cell types (Figure 49). The decreased expression of PTCH1 leads to decreased activity of SHH, thus promoting dorsalization of the cells and gives a reasonable explanation for the high expression of TTR in the cerebral organoids. As this can only be observed in the NPCs, but not in the iPSCs or neurons, this seems to be critical only during a certain time window during development. Together with the previous results, this further proves that it is indeed a disturbed SHH signalling that causes the observed phenotype in the cerebral organoids of R1 and R2.

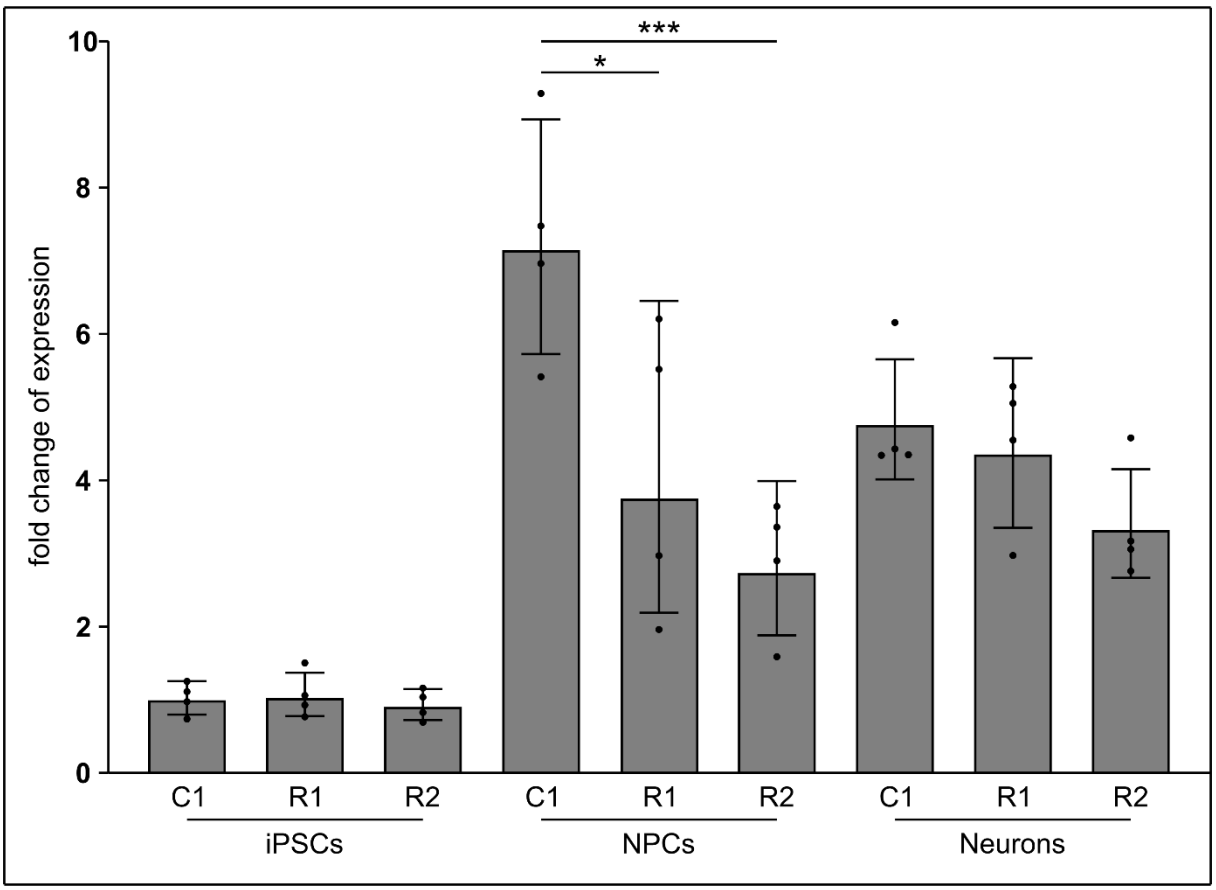


Figure 49 PTCH1 mRNA expression in the different cell types. RT-qPCR of PTCH1 shows significant differences between wildtype and mutant NPCs but not for iPSCs or neurons. Two-way ANOVA: Interaction: F (4, 27) = 2,472, P = 0,0684; Cell type: F (2, 27) = 91,86, P < 0,0001; Genotype: F (2, 27) = 7,328, P = 0,0029; followed by Tukey's multiple comparisons test, *P<0,05, ***P<0,001. Values are geometric mean ± geometric SD.

3.5 X-inactivation changes during differentiation

As previously shown, the cerebral organoids of OS M7 showed an intermediate phenotype between cerebral organoids of ctrl M10 and OS M14. One possibility is that this is caused by a rescue of the inactive wildtype allele. Cerebral organoids as well as NPCs and neurons generated from ctrl M10 and OS M7 were analysed for their *MID1* expression pattern and their overall X-inactivation status.

3.5.1 MID1 expression changes during neuronal differentiation

3.5.1.1 Allele-specific RT-PCR

The allele-specific RT-PCR (Figure 50) showed that, while iPSCs showed a clear monoallelic expression of either wildtype or mutant *MID1*, both, ctrl M10 and OS M7 derived NPCs and neurons expressed wildtype **and** mutant *MID1* mRNA. A similar result was also obtained when analysing the cerebral organoids (Figure 51); 30 and 50 days old cerebral organoids derived from ctrl M10 and OS M7 iPSCs showed expression of both, wildtype and mutant *MID1* mRNA. This suggests that the previously identified monoallelic expression of *MID1* was lost during the neuronal differentiation of the iPSCs and the before inactive allele was reactivated.

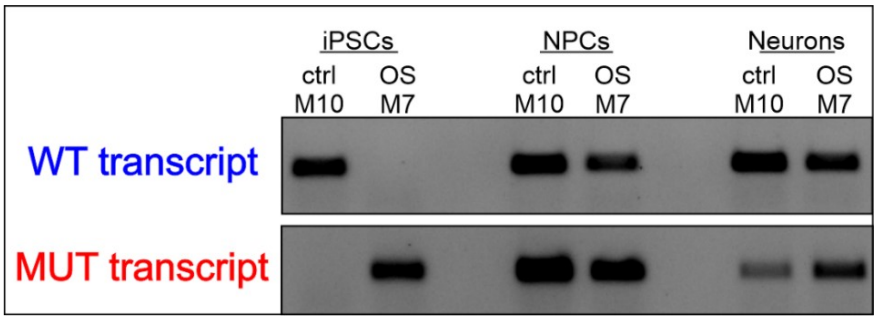


Figure 50 NPCs and neurons show reactivation of the before inactive allele of *MID1*. The allele-specific RT-PCR generates bands for both, the wildtype and the mutant *MID1* transcript in NPCs and neurons, while the expression of the iPSCs is stable.

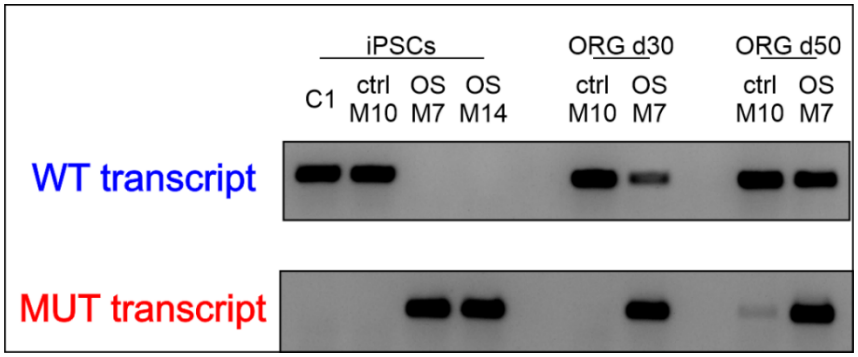


Figure 51 Cerebral organoids show reactivation of the before inactive allele of *MID1*. The allele-specific RT-PCR generates bands for both, the wildtype and the mutant *MID1* mRNA in NPCs and neurons, while the expression in the iPSCs is stable. C1 and OS M14 were used as additional controls.

3.5.1.2 Western Blot

We also repeated the Western Blot to figure out if wildtype MID1 is also detectable on protein level (Figure 52) in the OS M7 derived cells. In concordance with the results of the allele specific RT-PCR, wildtype MID1 protein was found in NPCs and neurons of both ctrl M10 and OS M7, but only in iPSCs of ctrl M10 (Figure 52). This further confirmed the reactivation of the second *MID1* allele during neuronal differentiation.

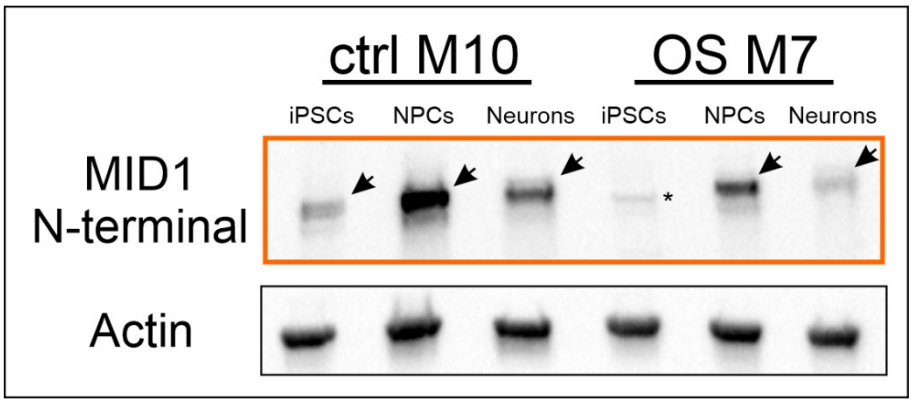


Figure 52 MID1 protein can be detected in ctrl M10 and OS M7 derived NPCs and neurons. Western Blot with an antibody detecting the N-terminal part of MID1. Actin was used as a loading control. Arrows mark specific bands; asterisks mark unspecific bands.

3.5.1.3 QUASEP assay

To quantify wildtype and mutant *MID1* mRNA we designed a QUASEP assay, that allows quantification of allele-specific expression by Pyrosequencing (Figure 53). Fibroblasts were used as controls and showed either monoallelic expression (male fibroblasts) or biallelic expression (female fibroblasts). In line with the results of the allele-specific RT-PCR, iPSCs showed a monoallelic expression of either the wildtype (ctrl M10) or the mutant allele (OS M7). The differentiated NPCs and neurons showed a significant reactivation of the before inactive allele. In contrast to the results of the allele-specific RT-PCR, the cerebral organoids showed no significant reactivation of the before inactive allele. Since cerebral organoids are a mixture of many different cell types, this might point towards a reactivation of MID1 restricted to a certain cell type present in cerebral organoids. The majority of cells in cerebral organoids seem to not reactivate MID1.

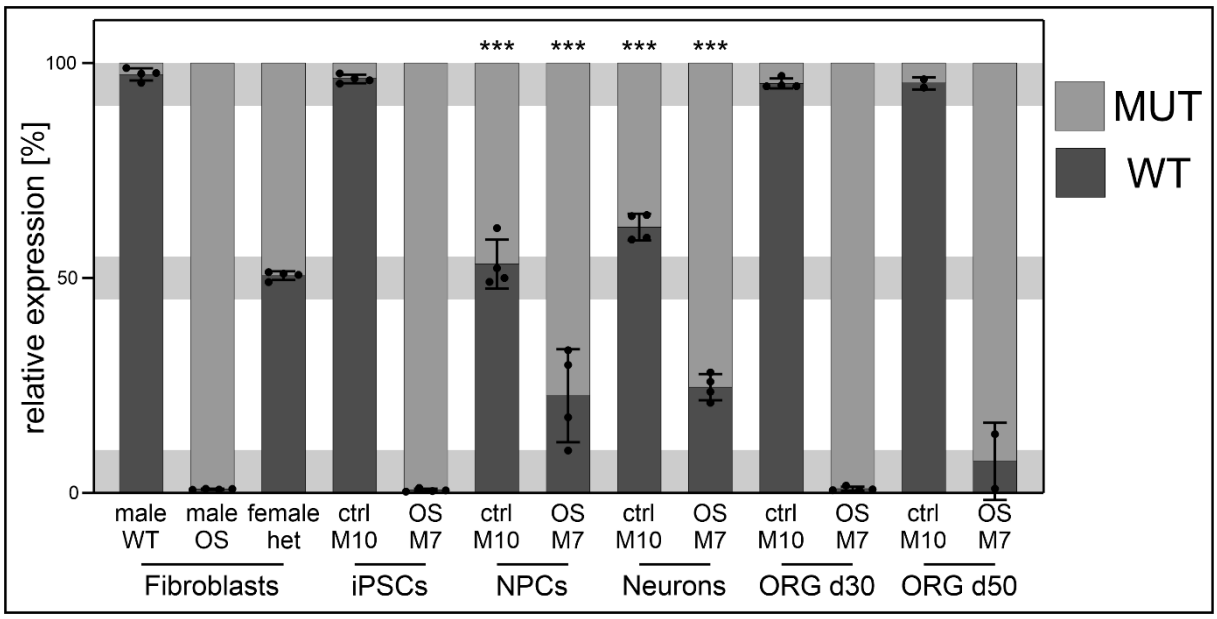


Figure 53 Significant reactivation of the before inactive allele of *MID1* in NPCs and neurons. A QUASEP assay to quantify the amount of reactivation in the NPCs and neurons showed a significant reactivation of the before inactive allele in NPCs and neurons compared to the iPSCs. Fibroblasts were used as controls. Grey background areas on top and bottom mark regions of monoallelic expression; grey background area in middle marks region of equal biallelic expression. One-way ANOVA: F = 327,9, P < 0,0001 followed by Tukey's multiple comparisons test: ***P<0,001. Values are mean ± SD.

3.5.2 CA5B and ZNF185 (QUASEP assay)

To further analyse X-chromosomal expression patterns during neuronal differentiation and to exclude complete erosion of XCI, as it appears in mouse IPSCs, we performed QUASEP assays with two more genes that were previously described (Cantone et al., 2016) to be expressed monoallelic in all tissues (*ZNF185*) or biallelic in all tissues (*CA5B*).

As previously described, *CA5B* showed a biallelic expression in all female cells (fibroblasts, iPSCs, NPCs and neurons), and a monoallelic expression in both male fibroblasts (Figure 54).

ZNF185 on the other hand, showed a monoallelic expression in iPSCs, NPCs and neurons of both ctrl M10 and OS M7 (Figure 55). Due to the mosaic present in female somatic cells, the female fibroblasts showed a biallelic expression.

Taken together, these results show that the expression of a known escape gene and a known non-escape gene is not influenced by the reprogramming of fibroblasts or the neuronal differentiation of iPSCs. This is a further prove that the reactivation of MID1 during neuronal differentiation is controlled by a gene specific mechanism and not a random process.

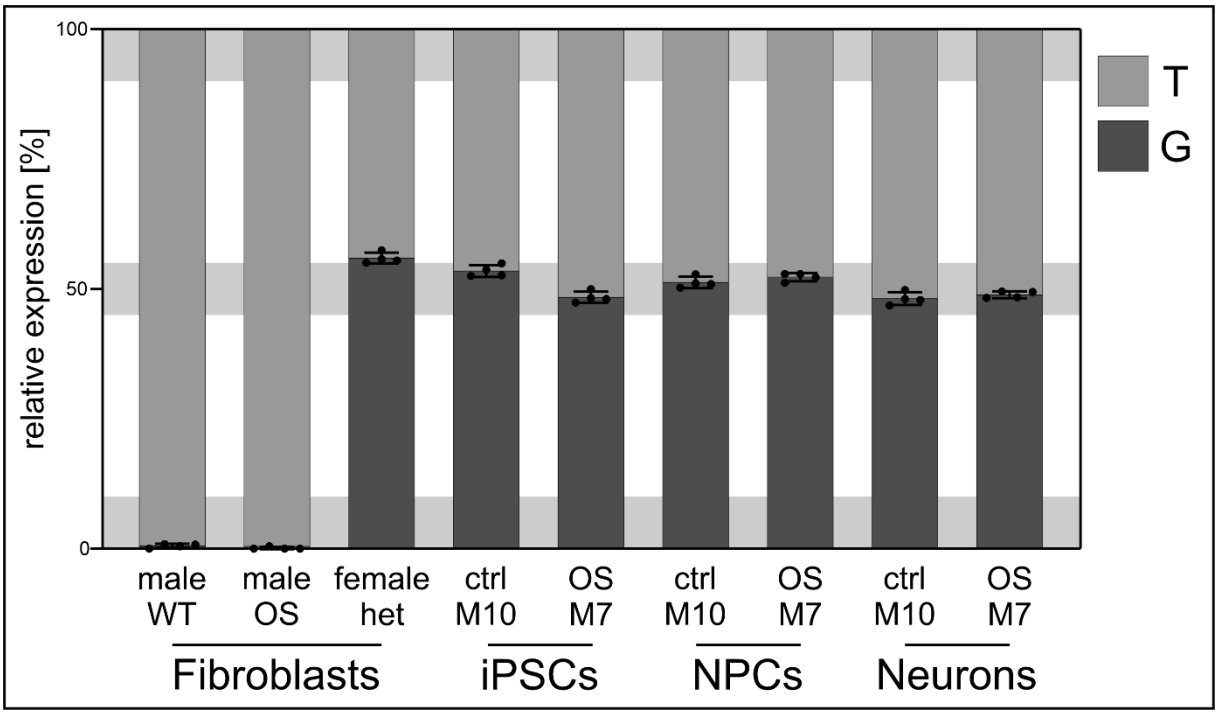


Figure 54 Biallelic expression of *CA5B*. A QUASEP assay detecting a SNP in the *CA5B* gene showed an equal biallelic expression in iPSCs, NPCs and neurons. Fibroblasts were used as controls. Grey background areas on top and bottom mark regions of monoallelic expression; grey background area in middle marks region of equal biallelic expression. Values are mean ± SD.

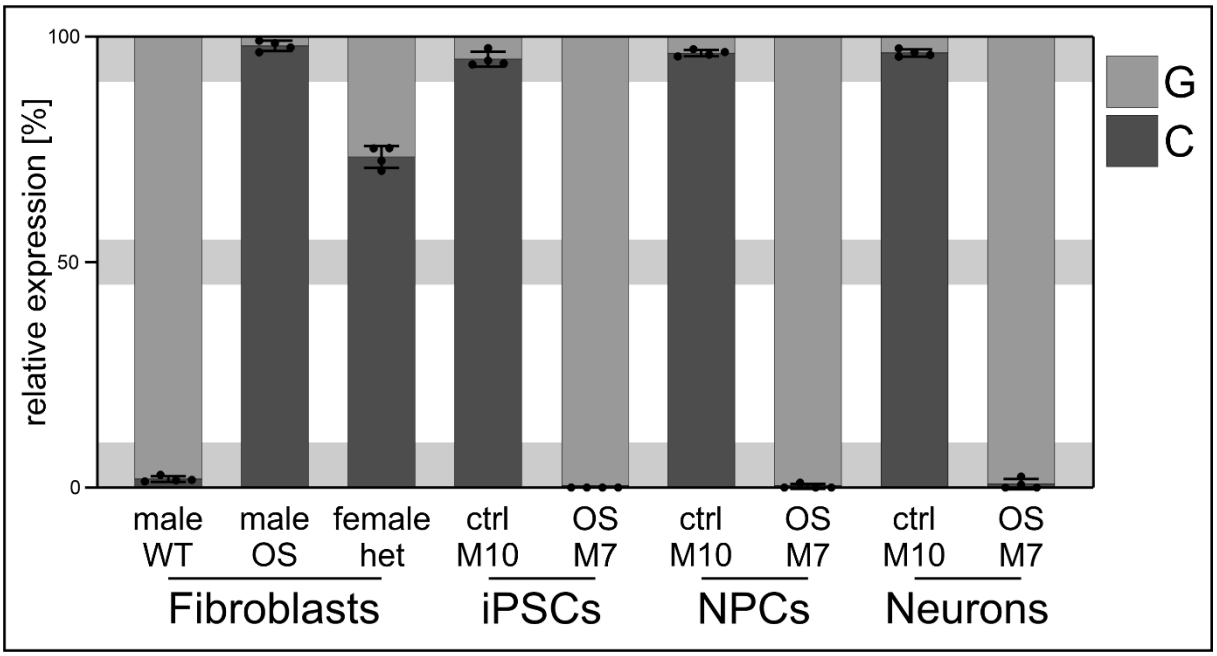


Figure 55 Monoallelic expression of *ZNF185*. A QUASEP assay detecting a SNP in the *ZNF185* gene showed a monoallelic expression in iPSCs, NPCs and neurons. Fibroblasts were used as controls. Grey background areas on top and bottom mark regions of monoallelic expression; grey background area in middle marks region of equal biallelic expression. Values are mean ± SD.

3.5.3 Identification of X-inactivation changes over the whole X-chromosome during neuronal differentiation

Knowing that *MID1* expression changes during neuronal differentiation while other genes like *CA5B* and *ZNF185* showed no change of expression related to the differentiation we tried to address X-in and reactivation in a more global way by performing an allele-specific RNA-seq. We used iPSCs, NPCs and neurons of ctrl M10 and performed a deep sequencing run with the prepared libraries of approximately 50 million reads per sample. The obtained data was processed and analysed with the help of [redacted] and [redacted].

The whole-transcriptome analysis showed clustering of the samples in correspondence with their known cellular identity (Figure 56). iPSCs formed a cluster distant from NPCs and neurons, while NPCs and neurons were closer to each other than to iPSCs. For further analysis, biological replicates D (not shown) and F of the iPSCs were excluded. Both showed a highly different transcriptome compared to the other biological replicates.

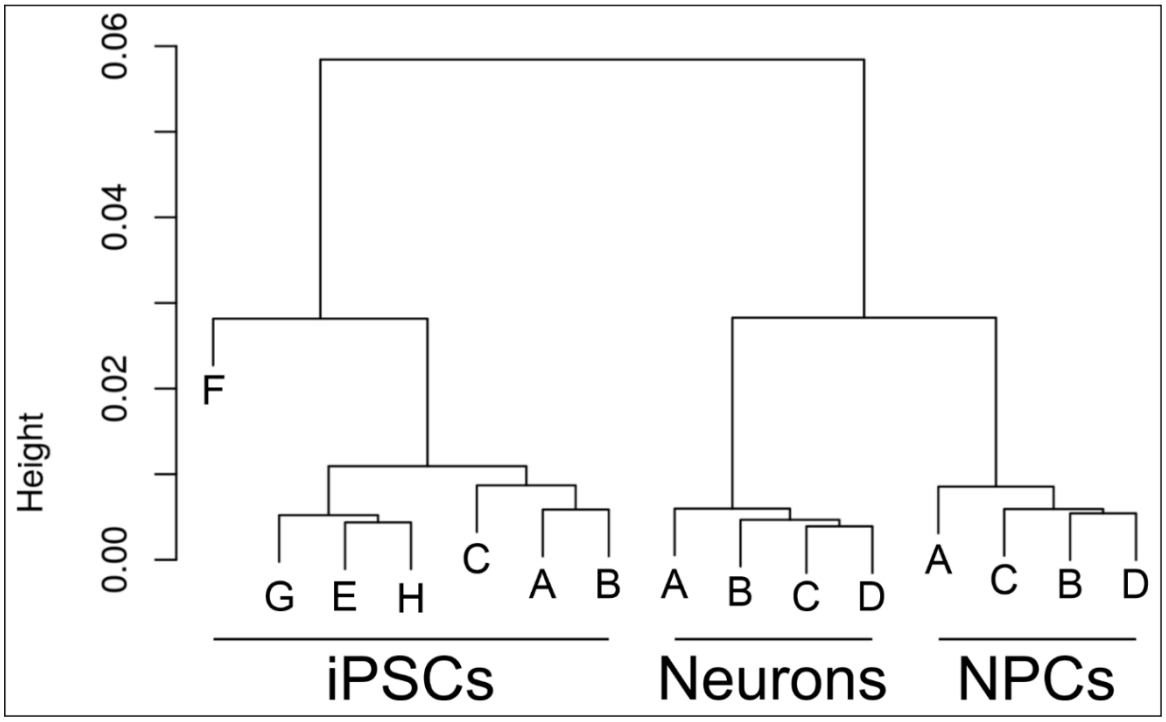


Figure 56 Transcriptome-based clustering of biological replicates used for the RNAseq. The analysed samples form clusters depending on their transcriptome that fit their known cellular identity. iPSC sample D was already excluded before due to sequencing abnormalities; iPSC sample F was excluded for the following analysis.

The remaining biological replicates of the iPSCs, NPCs and neurons were analysed regarding their XCI status. A tool commonly used for analysing WES data was adapted for the allele-specific expression analysis. This tool uses known SNPs distributed over the whole X-

chromosome and visualizes them as white in case of biallelic expression and dark in case of monoallelic expression. The overlay of the different biological replicates of this analysis is shown in Figure 57. We identified four main regions that seem to undergo X-chromosomal reactivation during neuronal differentiation, including Xp22.2, the region where the *MID1* gene is located. Together with the gene specific results obtained via QUASEP assays, this provides further evidence for targeted reactivation of defined X-chromosomal regions during neuronal differentiation.

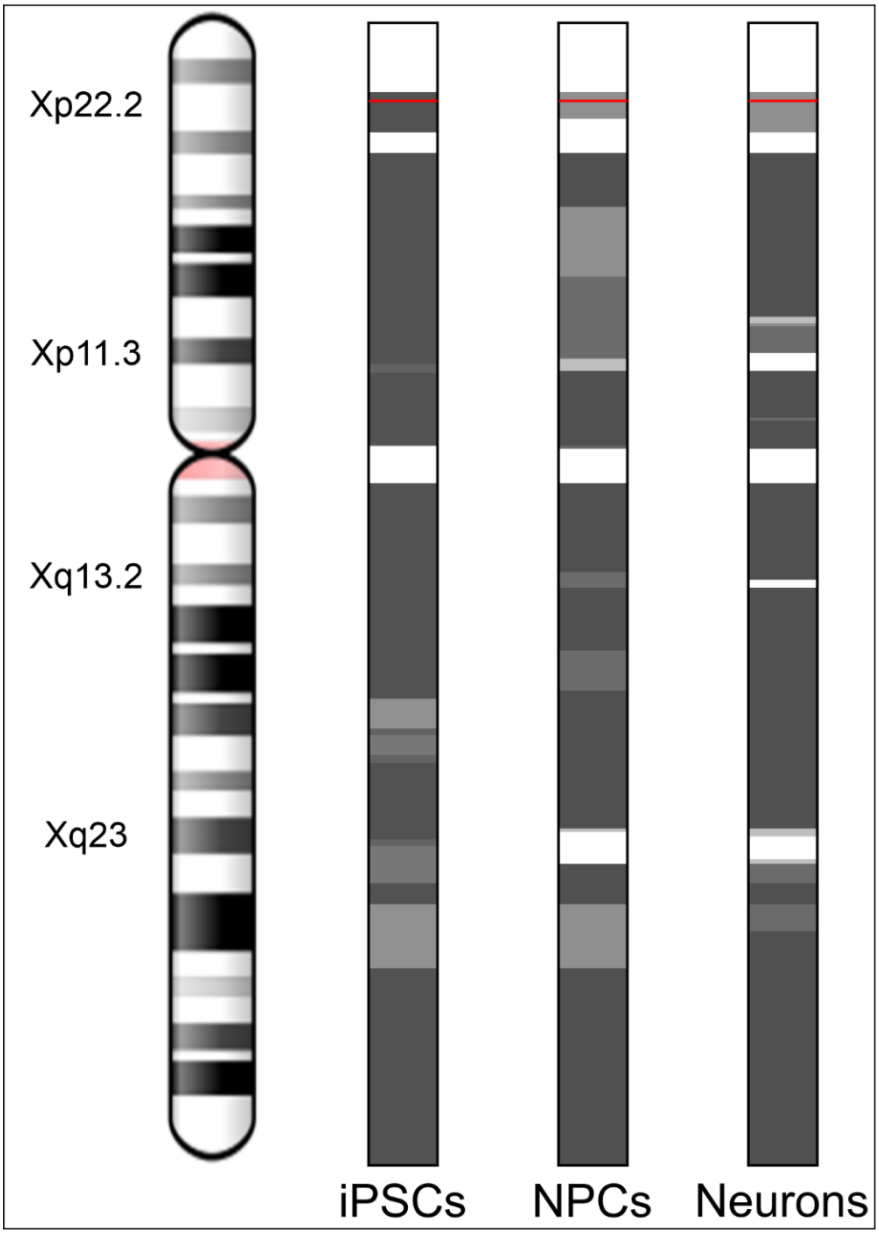


Figure 57 Allele-specific expression of X-chromosomal genes in iPSCs, NPCs and neurons. The transcriptome data was analysed for allele-specific expression of the X-chromosome to identify regions ore genes escaping XCI during neuronal differentiation. Overlap of biological replicates. Dark grey: monoallelic expression; white: biallelic expression; bright grey: intermediate; red line: MID1.

4 Discussion

MID1 is an X-chromosomal gene that is ubiquitously expressed in human tissues and organs. Mutations in *MID1* are known to cause OS in hemizygous males that is typically associated with hypertelorism, hypospadias, cleft-lip/palate, brain anomalies and intellectual disability (Fontanella et al., 2008; Robin et al., 1996). The gene encodes a RING-finger protein involved in embryonic midline formation and PP2Ac degradation (Trockenbacher et al., 2001).

Using primary human fibroblasts, cell lines and animal models, MID1 was shown to interact with mTOR, PAX6 and SHH (Krauß et al., 2008; Liu et al., 2011; Pfirrmann et al., 2016; Schweiger et al., 2014; Winter et al., 2016). This study uses human iPSCs and differentiated neuronal cells, showing a novel gain of function mechanism in cells with N-terminal MID1 mutations and a cell-cycle dependent phenotype of C-terminal MID1 mutations in cerebral organoids. Furthermore and surprisingly, gene-specific X-chromosomal reactivation during neuronal differentiation in cells with C-terminal MID1 mutations was found.

4.1 Characterization of primary and genetically engineered cells

The first aim of this doctoral thesis was to establish a workflow (Figure 58) starting with human skin punch biopsies for the isolation of fibroblasts that are reprogrammed into iPSCs and later differentiated into neuronal cells. Each step was proven by characterization of the corresponding cell type and comparison with the preceding cell type.

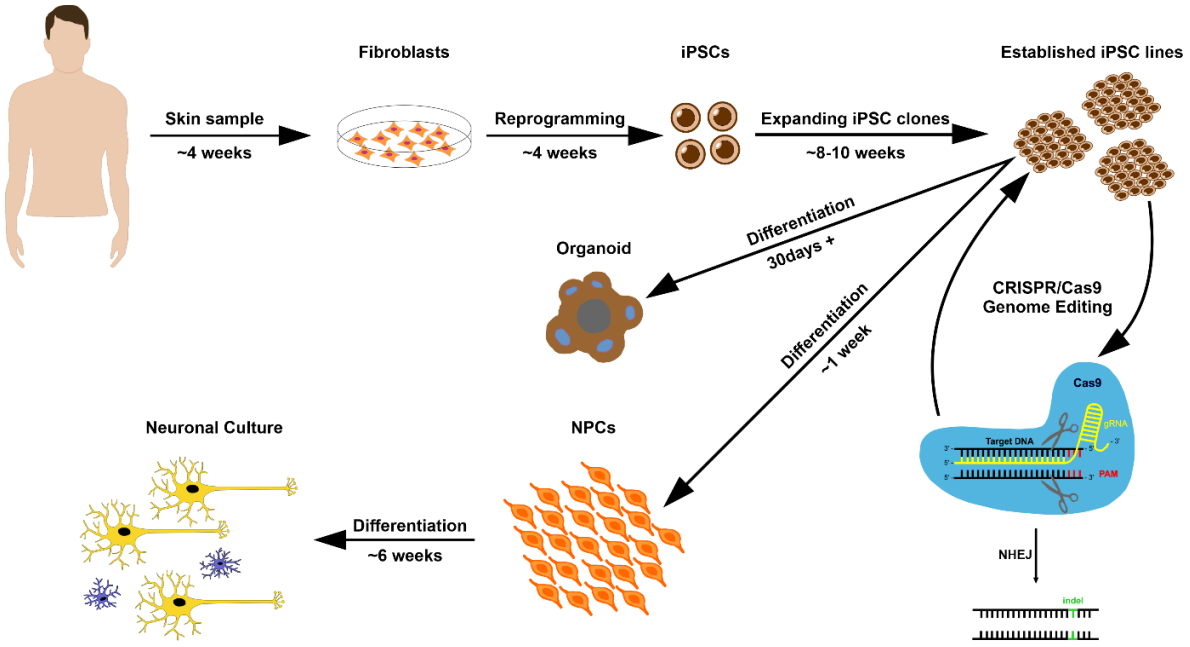


Figure 58 Workflow established during this doctoral thesis. Fibroblasts are isolated from skin punch biopsies and reprogrammed into iPSCs. Fully characterized iPSCs can be used for CRISPR/Cas9 genome editing or differentiated into brain organoids or NPCs. NPCs can be further differentiated into a mixed neuronal culture.

4.1.1 Reprogramming of skin fibroblasts

Reprogramming of human somatic cells into iPSCs is a relatively new technique but already widely used in many labs. The first published protocol for the generation of iPSCs used skin fibroblasts, that were transduced with retroviruses carrying the four reprogramming factors that later got known as the Yamanaka factors (Takahashi et al., 2007). Fibroblasts can be isolated from skin biopsies and grown under standard culturing conditions. Although many more cell types were shown to be capable of reprogramming (Staerk et al., 2010; Yamanaka, 2010; Zhou et al., 2012), fibroblasts remain one of the most commonly used cellular source for reprogramming (Raab et al., 2014).

Skin fibroblasts play an important role in the synthesis of the extracellular matrix and collagen. One of the key proteins involved in collagen synthesis is SERPINH1, a serine proteinase inhibitor of the serpin superfamily, also known as HSP47. It is localized to the endoplasmic reticulum and has been shown to interact with collagen as a molecular chaperon, driving its maturation (Dafforn et al., 2001; Kuroda & Tajima, 2004; *SERPINH1 Gene - GeneCards*, n.d.).

Transduction of the reprogramming factors initiates a process that changes the cellular methylome, transcriptome and proteasome. During this process genes that are responsible for self-renewal and maintaining pluripotency are upregulated while those characteristic for the somatic donor cells are silenced. The morphology of the reprogrammed cells changes, and they are able to maintain pluripotency on their own.

In this study, fibroblasts of different donors were used for reprogramming. All fibroblasts showed expression of SERPINH1, together with a typical morphology that is described as “plump spindle shaped or stellate shaped cells (active fibroblasts) with [a] centrally placed oval or round nucleus” (Ravikanth et al., 2011). During reprogramming, independent of the method used, the cellular morphology changed towards the typical ESC- and iPSC-like morphology. The cells became round, had prominent nucleoli and a high nucleus to cytoplasm ratio. They started to form the typical compact colonies also observed in ESCs, with well-defined and smooth edges and an overall flat appearance (Wakui, 2017). It is well known that copy number variations and chromosomal rearrangements are frequent abnormalities that occur during reprogramming (Hussein et al., 2011; Sobol et al., 2015). Most of our cells maintained a normal karyotype of 46,XX or 46,XY respectively throughout the whole study. Similar to previous publications (Altieri et al., 2019) we found elevated expression levels of pluripotency markers by RT-qPCR with the exception of *KLF4*, which was downregulated in comparison to the original fibroblasts. KLF4 is a transcription factor that is involved in skin barrier formation and

thus relatively high expressed in skin fibroblasts (Jaubert et al., 2003; Segre et al., 1999). Another proof of successful reprogramming that is commonly used is the detection of proteins characteristic for iPSCs using immunofluorescent staining (Havlicek et al., 2014). In addition, we also performed routinely mycoplasma testing and all fibroblasts were tested negative for HIV, HBV and HCV in the "Zentrallabor der Universitätsmedizin Mainz". As in this study, it is typically the combination of several methods that is considered a proof for the successful reprogramming and generation of iPSCs (Altieri et al., 2019).

Nevertheless, it must be mentioned that in the scope of this work, additional experiments to further asses the quality of our iPSCs were not performed. A common tool to show the pluripotent state of the generated iPSCs, for example, is the differentiation into all three germ layers, which can either be done by a teratoma assay, where immunosuppressed mice are being injected with iPSCs, or by a differentiation assay where iPSCs are spontaneously differentiated *in vitro* (Baghbaderani et al., 2016; Havlicek et al., 2014; Takahashi et al., 2007). [redacted] in the lab has just recently established a directed differentiation of iPSCs into all three germ layers in our lab showing that our iPSCs are indeed able to differentiate into cells of all three germ layers, further proving their pluripotency.

4.1.2 Dual SMAD based differentiation of induced pluripotent stem cells

Dual SMAD based differentiation of human iPSCs into neuronal stem cells was first published in 2009, eliminating the need for embryoid body formation, co-culture of feeder cells or selective conditions during culturing (Chambers et al., 2009). This method relies on blocking the signalling pathways of BMP and TGFB that both belong to the TGFβ superfamily and play an important role during embryonic development. Extracellular TGFβ signalling activates or represses intracellular SMAD proteins that drive gene expression (Heldin et al., 1997), which gave the method its name. The synergistic action of noggin (blocking BMP signalling) and SB431542 (blocking TGFB signalling) induces a neural conversion of the treated iPSCs (and ESCs) resulting in cells positive for the neuronal progenitor markers PAX6, SOX2 and NESTIN (Kang et al., 2017), while at the same time downregulating expression of the stem cell markers characteristic for iPSCs. The resulting NPCs can be further differentiated into neurons.

In this study, two different dual SMAD based differentiation protocols were used. The first one relies on the generation of embryoid bodies and manual selection of neuronal rosettes. The NPCs generated with this protocol were used for the allele specific RNAseq experiment. The second protocol requires a simple change of media to induce differentiation and showed to be

less prone to failure due to the reduced number of steps required and the no longer required manual selection. We chose to use this method for most of the experiments.

More recent publications often skip the generation of NPCs and use methods to directly differentiate iPSCs into neurons or specific neuronal subtypes (Hong & Do, 2019). Using this approach can save time and avoids the culture of NPCs, but also eliminates the potential to analyse NPCs. In this study, an altered *PTCH1* expression was observed exclusively in NPCs but not in iPSCs and neurons. Directed differentiation of iPSCs into neurons would have missed this phenotype. Other studies also showed that NPCs are a suitable tool to mimic neurodevelopmental disorders and that it does not always require neurons to better understand the molecular mechanisms underlying a disease (Madison et al., 2015; Martin et al., 2020; Mishra et al., 2016). Furthermore, differentiation of NPCs into neurons can give rise of additional information and help to further unveil the complex mechanisms underlying neurodevelopmental disorders.

4.1.3 Vitamin A in neuronal differentiation

Neuronal differentiation is a fine-tuned mechanism led by different signalling molecules and pathways. A molecule critically involved in neuronal differentiation is retinoic acid, a metabolite of vitamin A. It is known to play a role in anterior-posterior patterning of the neural plate and, together with SHH, BMP and FGF, to define the dorsal-ventral axis of the neural tube (Maden, 2007; L. Wilson et al., 2003; Leigh Wilson et al., 2004). Apart from its functions during embryonic development, retinoic acid is also involved in axon regeneration in adults (Maden, 2007; Wong et al., 2006). *In vitro*, retinoic acid induces the differentiation of various types of neurons and glia cells (Maden, 2007). A more recent publication (Tan et al., 2015) has shown that retinoic acid treated (rat) neuronal stem cells (NSCs) show an increased expression of TUBB3 compared to non-treated NSCs. At the same time the treated cells showed a reduction of the astrocyte marker GFAP. This data indicates that retinoic acid "could promote growth of cellular dendrites and neuronal differentiation of NSCs" (Tan et al., 2015).

In this study, NPCs were differentiated into neurons by plating at a certain density and using a culture medium containing vitamin A. Over the course of five weeks the cellular morphology changed towards a neuronal like morphology including dendrites and branching. The cells showed expression of neuronal markers including TAU and MAP2, specific for axons or dendrites respectively, while at the same time, markers specific for NPCs were downregulated. Taken together, these results show that we have successfully generated neurons. However, two more aspects of the generated neurons need to be characterized in future experiments. The exact

neuronal subtype generated remains unknown. By using markers specific for certain neuronal subtypes and brain regions, this can be addressed. Depending on their subtype, the neurons should also show a specific electric activity; using patch-clamp the electrical activity can be measured and characterized. A plausible first guess would be to check the cells for expression of glutamate. Several publications over the past years have used a similar protocol, with retinoic acid as the main driver, to differentiate cells into neurons (Han et al., 2011; Yichen Shi et al., 2012; Siegenthaler et al., 2009). They have further characterized the differentiated neurons as glutamatergic neurons and associated them with a cortical identity. To differentiate NPCs into neurons of other subtypes different sets of small molecules are required often in combination with adjusted cell seeding (Figure 59).

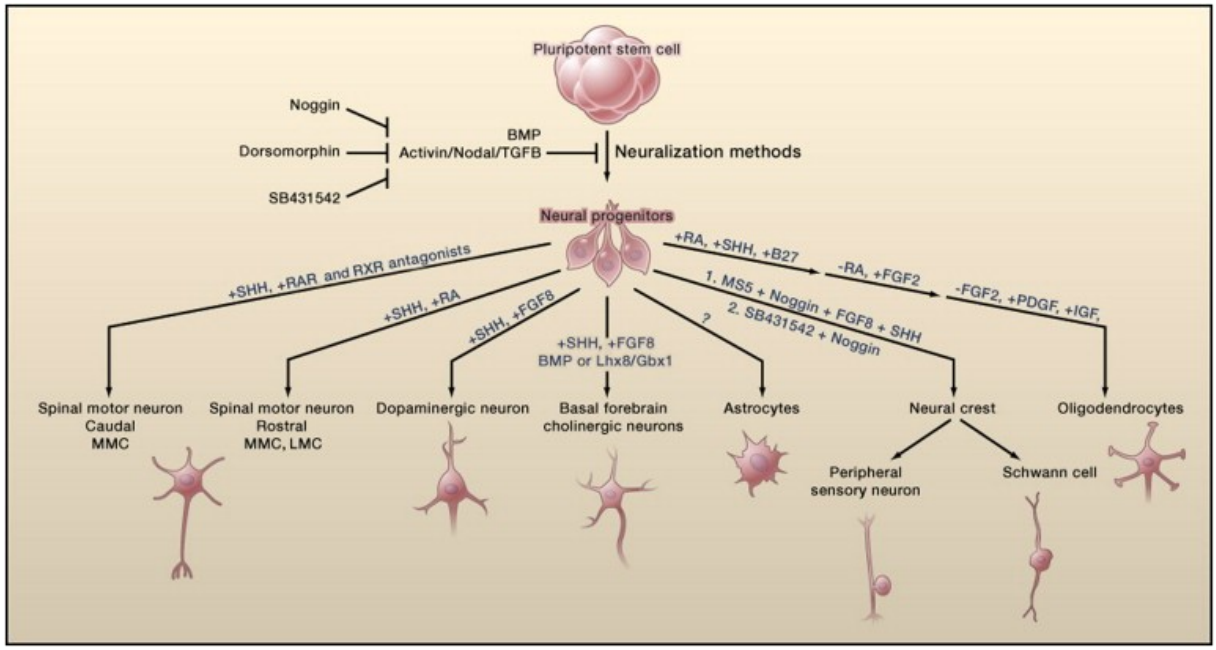


Figure 59 Differentiation into neuronal subtypes. Pluripotent stem cells are differentiated into NPCs via a dual SMAD based protocol. NPCs can be further differentiated into neuronal subtypes using small molecules (Han et al., 2011).

The differentiation into spinal motor neurons, for example, requires a combination of SHH and retinoic acid (Bahmad et al., 2017; Han et al., 2011). This method is often used in amyotrophic lateral sclerosis (ALS) related research. ALS is a neurodegenerative neuromuscular disease that is mainly characterized by progressing muscle weakness. A study in 2008 showed that ALS patient-specific iPSCs can be differentiated into motor neurons and serve as a model to better understand the molecular mechanisms underlying ALS (Dimos et al., 2008).

Other neuronal subtypes of interest are dopaminergic neurons and GABAergic neurons. GABAergic neurons are generated by using SHH as the main driver of differentiation and are of particular interest in Huntington's disease (HD) research (Bahmad et al., 2017; Han et al., 2011). HD is a neurodegenerative disease caused by a CAG-triplet expansion in the *HTT* gene and causes motor, cognitive and psychiatric symptoms that progress over time. In 2010, HD patient-specific iPSCs were differentiated into GABAergic striatal neurons and were predicted to be a suitable model for further research (N. Zhang et al., 2010). Dopaminergic neurons, on the other hand, play an important role in the research of Parkinson's disease (PD) and require a combination of FGF8 and SHH to be generated (Bahmad et al., 2017; Han et al., 2011). PD is a neurodegenerative disease that affects dopaminergic neurons in the brain. The successful differentiation of patient-specific iPSCs into dopaminergic neurons was shown to be a suitable tool for further research on PD (Cooper et al., 2010; Soldner et al., 2009). These commonly researched diseases show that it is essential to choose the correct neuronal subtype differentiation for each disease. Unfortunately, in less commonly researched diseases, the neuronal subtype affected is often unclear.

4.1.4 16p11.2 microduplication

Recurrent microduplications on chromosome 16p11.2 (OMIM #614671) are characterized by a broad variety of neurodevelopmental impairments including intellectual disability, developmental delay, epilepsy and autism among others. In addition, a variable spectrum of psychiatric disorders such as depression, schizophrenia and anxiety can be observed (Bijlsma et al., 2009; Fernandez et al., 2010; Kumar et al., 2008; Marshall et al., 2008; McCarthy et al., 2009; Niarchou et al., 2019; Shinawi et al., 2010; Weiss et al., 2008; Zufferey et al., 2012). All of these symptoms have been shown to be of incomplete penetrance and variable expressivity (Bijlsma et al., 2009). The analysed patient (1180/17) presented with moderate learning disability and had episodes of anxiety disorder and psychosis. Interestingly, her brother and father who both carry the same microduplication on chromosome 16p11.2 also show clear symptoms in line with previous reports (Niarchou et al., 2019). The father has been suffering from serious depression and anxiety disorder for the last 20 years, while the brother showed significant speech delay, moderate learning disability and a generalized anxiety disorder.

The genomic region that is duplicated spans approximately 600 kb and contains ~24 genes (Kumar et al., 2009). Some of these genes have been successfully identified as causative for certain clinical phenotypes observed in carriers, however, molecular effects at the transcriptional level are still not clearly understood (Golzio et al., 2012; Kusenda et al., 2015).

Deshpande and colleagues have investigated cellular phenotypes in human iPSC-derived neurons from 16p11.2 deletion and duplication carriers and have shown reduced neuronal size and dendrite length in neurons from 16p11.2 duplication carriers (Deshpande et al., 2017). They also analysed mRNA expression of duplicated genes and found a consistent upregulation of these genes, especially in NPCs (Figure 60). They analysed a total of 8 genes in iPSCs, NPCs and neurons.

We chose to analyse six genes of the duplicated region of which only two (*ALDOA* and *MAPK3*) were also analysed by Deshpande et al (Deshpande et al., 2017). *ALDOA* and *CDIPT* have been previously reported to have the strongest gene dosage effect (Kusenda et al., 2015). In addition, we chose *KIF22* and *MAPK3* because of their association with joint laxity (another symptom of the brother) (Holm et al., 2011; Horev et al., 2011; McCarthy et al., 2009). *DOC2A* was chosen because it is involved in nervous system development and synaptic transmission (Duncan et al., 2000; Groffen et al., 2006). Finally, *SEZ6L2* was chosen because of its expression pattern that is consistent with the neurodevelopmental basis of autism (Kumar et al., 2009). As a control, we also analysed *POGZ*, a gene located on chromosome 1. Mutations in *POGZ* are known to cause the White-Sutton syndrome that is characterized by intellectual disability and autistic features among many others (OMIM #616364).

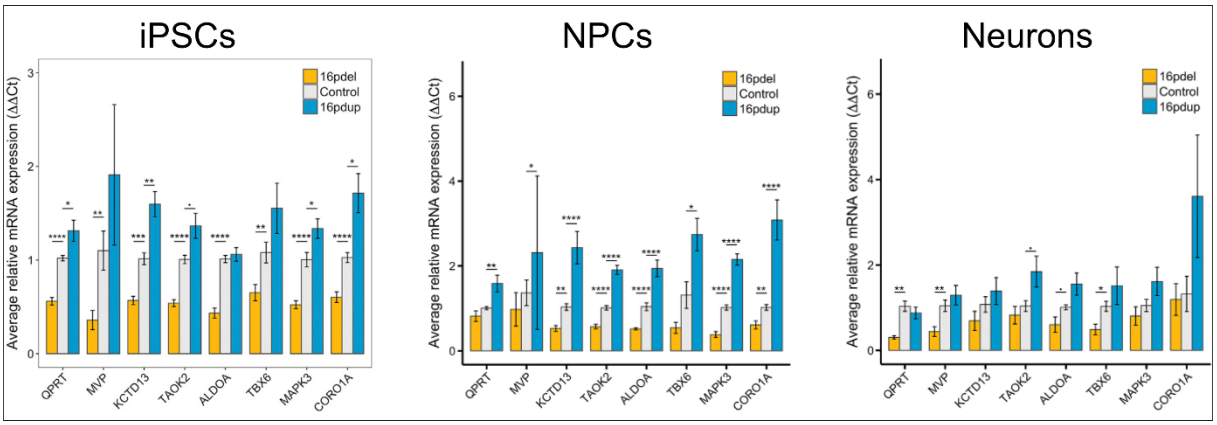


Figure 60 mRNA expression of selected 16p11.2 genes by RT-qPCR. Adapted from (Deshpande et al., 2017).

We found *KIF22*, *MAPK3*, *CDIPT* and *DOC2A* to be stably upregulated by ~1,5-fold in fibroblasts, iPSCs and NPCs and up to ~5-fold in the neurons of the patient (1180/17) compared to her mother (1179/17). As already mentioned, *KIF22* and *MAPK3* are associated with joint laxity, a symptom only found in the brother of the patient, although their expression is upregulated in all analysed cell types. It would be interesting to see if the brother shows a similar

upregulation of these genes. Any different expression could be an explanation for the different phenotypes observed in the patient and her brother. *DOC2A* is involved in nervous system development and synaptic function and has also been shown to be involved in anxiety disorders (Glessner et al., 2010; Hill et al., 2020), a symptom present in the patient. CDIPT is a downstream messenger of many G-coupled receptors that are involved in calcium metabolism, that is of high importance for proper brain function (Bading, 2013; Lykidis et al., 1997).

Two genes of the duplicated region did not show a stable increase of expression in all cell types. Expression of *ALDOA* was upregulated in fibroblasts and iPSCs but not in NPCs and was even seen downregulated in neurons. ALDOA is a glycolytic enzyme involved in fructose metabolism and has two aldolase isozymes (ALDOB and ALDOC). ALDOA and ALDOC are co-expressed in the brain and other nervous tissues (OMIM #103850). Decreased expression of ALDOA in the patient's neurons can probably be compensated by ALDOC. In contrast to *ALDOA*, expression of SEZ6L2 was not downregulated but highly upregulated (50-100-fold) in NPCs and neurons of the patient. SEZ6L2 is not expressed in fibroblasts and no differences in expression can be found in the iPSCs of the patient and her mother. SEZ6L2 has been shown to be involved in neuronal differentiation and neurite outgrowth (Boonen et al., 2016). Furthermore, it is a modulator of AMPA-ADD (adducin) signal transduction and an important scaffolding protein mediating linkage of GluR1 (glutamate receptor 1) and ADD (Yaguchi et al., 2017). While mutations in *SEZ6L2* have been shown to play no major role in the susceptibility to autism spectrum disorders, a dosage-dependent effect is possible (Konyukh et al., 2011; Kumar et al., 2009). Further studies in our patient's cells shall confirm a possible dosage-dependent effect of SEZ6L2 on AMPA-signalling.

Most interesting is the increased expression of *POGZ* that was observed in the neurons of the patient. This indicates that the altered expression of genes of the duplicated region on chromosome 16p11.2 might cause changes in transcription of other genes involved in neuronal development. Further experiments are necessary to identify the exact mechanism causing the elevated *POGZ* expression.

The unexpected expression pattern of *ALDOA* and *SEZ6L2* observed in our patient-derived disease model could point at a strong alternative regulation of expression of these genes resulting in escape of the gene-dosage effect of the duplication in a cell-type specific manner. This could be an explanation for the broad clinical variability that is seen in CNV carriers. A plausible next step would be to analyse the father and the brother of the analysed patient to identify possible differences in gene expression that could explain the different phenotypes

observed. Taken together, this data confirms that the established workflow of generating patient-specific neuronal cells is suitable to mimic a neurodevelopmental disorder and can help to study the mechanisms underlying these diseases.

4.2 Gain- vs loss-of-function mutations of MID1

The second aim of this study was to characterize differences between N- and C-terminal MID1 mutations. While C-terminal mutations occur frequently in OS patients, no N-terminal mutation has ever been described (Winter et al., 2016). The most N-terminal mutation seen in an OS patient is at position 343, leading to a premature stop. This is in close proximity to the B-Box1 subunit starting at position 357 but distant from the RING-finger subunit ending at position 177. The only known mutations involving the RING-finger are full deletions or duplications of the first coding exon, that were found in several patients. We used CRISPR/Cas genome editing to generate two male iPS cell lines carrying a mutation in the RING-finger or in close proximity (position 136 or 204 respectively), both predicted to lead to a frameshift and a premature stop. In addition, fibroblasts of a female heterozygous carrier of a patient-specific 4 bp deletion in the last coding exon of *MID1* (c.1801_1804delCTCC) were reprogrammed into iPSCs. In this case, isogenic controls were generated by clonal selection; as previously described for iPSCs of patients suffering from Rett syndrome (Ananiev et al., 2011) (see also chapter 4.3.1). The N- and the C-terminally mutated iPSCs were then differentiated into NPCs and neurons to analyse cell type specific expression changes and into cerebral organoids to analyse how brain development is influenced.

The three-dimensional differentiation of iPSCs into cerebral organoids was first described in 2013 (Lancaster et al., 2013).Using embryoid bodies and specialized culture media, iPSCs form small brain like structures of approximately 4 mm in diameter. The brain organoids typically consist of neuronal stem cells and neurons of different fates but can also contain retinal tissue or cells of the choroid plexus (Lancaster & Knoblich, 2014). Depending on the iPSC-line and its genetic background, the cerebral organoids favour a more ventral or dorsal differentiation. Using certain drugs, that interfere with SHH signalling and other pathways responsible for the definition of the dorsal-ventral axis, this patterning can be directed towards the wanted identity (Bagley et al., 2017).

Interestingly, the cerebral organoids with the N-terminal MID1 mutation developed a distinct phenotype than the cerebral organoids expressing the C-terminal mutation. The wildtype C1 cerebral organoids developed in a normal way and showed the typical morphology mainly consisting of ventricle-like structures surrounded by neurons. In contrast, the mutant R1 and R2 cerebral organoids showed a significantly reduced number of ventricle-like structures and big areas made up of neither neuronal cells nor neuronal stem cells. These areas stain positive for TTR, a marker of the choroid plexus. This tissue can regularly be found in cerebral organoids

but typically at very low levels (Lancaster et al., 2013). Increased amount of choroid plexus tissue points towards a hyperdorsalization of the organoids as the choroid plexus develops from the dorsal axis of the neural tube (Lun et al., 2015). One of the key proteins during neural tube development is SHH that is expressed as a gradient, with high concentrations in the ventral part and low concentrations in the dorsal part of the neural tube. A decreased expression of SHH and its target genes leads to dorsalization and is a possible explanation for the observed phenotype in the R1 and R2 cerebral organoids. When measuring the expression level of the SHH receptor PTCH1, we indeed found a significant reduction of *PTCH1* expression in the NPCs of R1 and R2. This reduced expression is most likely caused by the mutations induced in *MID1*. As already described (1.3.4), silencing of MID1 stabilizes full-length FU and eventually leads to a decreased expression of SHH target genes and SHH signalling in general (Schweiger et al., 2014; Winter et al., 2016). N-terminal MID1 mutations seem to influence SHH signalling early during development causing a severe phenotype that is likely to be lethal during embryonic development. A recent study (Sasai et al., 2019) summarizing the effects of mutations in the different proteins involved in Shh signalling in mice has shown that mutations in Ptch1 and Shh are lethal during early stages of embryonic mouse development. As SHH signalling is highly conserved between man and mouse these findings can be transferred to human embryonic development. Interestingly, neither the NPCs nor the neurons generated from R1 and R2 iPSCs showed any morphological abnormalities or differed significantly in marker gene expression from the NPCs and neurons of the C1 iPSCs. This shows, that 2D neuronal cultures are a versatile tool for identifying pathways that underlie developmental abnormalities but that it needs the complex arrangement of the brain to mimic a phenotype properly.

The mutations induced in R1 and R2 iPSCs were initially thought to result in a full knockout of *MID1* by causing a frameshift leading to an early stop codon. However, our results suggest that there is a previously unknown gain of function phenotype caused by N-terminal MID1 mutations that differs from the classical loss-of-function phenotype as seen in OS patients. One possibility to explain this are alternative translation start within the first coding exon.

Initiation of translation is a complex process involving several initiation factors that cause the preinitiation complex to scan the mRNA to find the start codon by moving towards its 3’ end (Kozak, 1999). This scanning process has shown to be leaky, leading to the use of alternative start codons, including non-AUG as well as AUG codons up- and down- stream of the “original” (Kozak, 2002). Upstream AUGs outside the ORF typically cause short proteins of unknown function that might play a role in the modulation of mRNA translation (Kozak, 2002;

Meijer & Thomas, 2002; Morris & Geballe, 2000), while those in the ORF result in N-terminally extended versions of the "original" protein. Downstream AUGs typically cause N-terminal truncated versions of the "original" protein that differ in function, stability and activity (Kochetov, 2006). Since the N-terminal part often encodes the cellular localization of a protein, both extended and truncated proteins might play a role in different compartments of the cell. Even the tissue-specific use of alternative start codons is possible (Lukaszewicz et al., 2000).

A recent study has analysed CRISPR-induced gene knockouts in a cell line (Tuladhar et al., 2019). All 13 mutations are commercially available and were thought to be causing a frameshift and an early stop codon. Using Western Blot and RT-PCR, the authors were able to show that only seven mutations were causing full knockouts as anticipated. Of the remaining six mutations, four mutations were shown to cause alternative splicing and two were shown to cause the usage of alternative start codons. Although the induction of Indels in the first coding exon of a gene is the commonly chosen method to induce full gene knockouts, this study shows that the generated cells need to be properly characterized to avoid any unwanted gene expression.

In line with these findings, are the results of the cells carrying an N-terminal MID1 mutation. The R2 cells show an additional band in the Western Blot using two different MID1 specific antibodies. The size of this band fits perfectly to the predicted size of an alternative start codon frame of which gets rescued by the mutation in the R2 cells. The translation prevents the mRNA from undergoing NMD and the newly synthesized protein lacking the RING-finger is likely to be causative for the phenotype observed. Nevertheless, this does not explain the phenotype in R1 cells, where the frame of this alternative start codon is not being rescued by the induced mutation. However, a closer look at the Western Blot using the C-terminal MID1 antibody revealed additional bands in all samples that were shorter compared to the wildtype MID1. The size of these bands fits to three additional alternative start codons within the first coding exon that are 3' of the mutations induced in R1 and R2 and within the classic wildtype frame. The resulting proteins are all lacking the RING-finger and could be causative for the observed phenotype. As these alternative start codons are within the classic wildtype frame, the same bands can also be detected in wildtype cells. This shows that it is not alone the presence of these N-terminally truncated proteins that is causative for the gain of function phenotype but rather the absence of wildtype MID1. MID1 is known to form dimers through its coiled-coil domain, either with itself or with other proteins as MID2 (Short et al., 2002). As the coiled-coil domain

is present in in the N-terminally truncated MID1 isoforms that we identified, it is very likely that there is a heterodimerization of full length MID1 proteins with N-terminally truncated proteins. Sharing one N-terminus in such heterodimers seems to be good enough to maintain the normal MID1 function. This data is a first hint that the balance between full length and truncated MID1 is required during early brain development to fine tune dorsalization and SHH signalling. The exact underlying molecular mechanism that causes the described phenotype in the absence of wildtype MID1, however, remains unknown and needs further experiments to be identified. Because the observed dorsalization effects occur very early during development, we think that N-terminal mutations that are located before the alternative start codons, will lead to very early lethality in embryos. This explains why mutations in this part of the *MID1* gene are not found in patients.

The main difference between the N-terminally truncated MID1 isoforms and the patient-specific C-terminal mutations seems to be the cellular localization of the protein. While overexpression of all four N-terminally truncated MID1 proteins reveals their ability to associate with microtubules, overexpression of patient-specific mutated MID1 confirmed previously published data showing that the microtubule binding ability is lost (Schweiger et al., 1999). Depending on the position of the mutation, the overexpressed patient-specific proteins are either evenly distributed in the cytoplasm or form aggregates, but never associate with microtubules. The lost ability of microtubule binding is the basis for the assumed loss-of-function mechanism that is accepted to be causative for OS. This is further supported by full MID1 gene deletions found in OS patients showing the classical spectrum of symptoms (Cox, 2000; de Falco et al., 2003; So et al., 2005; Winter et al., 2003, 2016).

The loss-of-function is also an explanation for the phenotype observed in the cerebral organoids generated from the female iPSCs of ctrl M10 and OS M7. Both iPSC-lines developed normal cerebral organoids but showed clear differences after staining. The cerebral organoids of OS M7 showed an increased number of $PAX6^{+}$ cells compared to the cerebral organoids of ctrl M10. This is probably caused by the inability of the mutant MID1 to ubiquitinate PAX6 for proteasomal degradation (Pfirrmann et al., 2016), leading to an accumulation of PAX6. Elevated PAX6 prevents the cells from neuronal differentiation and maintains self-renewal and proliferation (Sansom et al., 2009). Over longer differentiation time, this phenotype maintained, but became less apparent (data not shown; (R. Menon, 2019)).

Interestingly, when comparing the cerebral organoids of ctrl M10 and OS M7, both female cell lines generated from a heterozygous mother, with organoids generated from iPSCs of the

hemizygous son (OS M14), we found a significantly stronger phenotype in the male cerebral organoids than compared to the female mutant tissue generated from OS M7. This finding shows that female cells provide protective mechanisms that result in milder phenotypes and let us hypothesize that MID1 might undergo X-chromosomal reactivation during the neuronal differentiation (see chapter 4.3.2).

4.3 X-chromosomal characterization

The third aim of this thesis was to study and better understand X-chromosomal inactivation and reactivation of *MID1* and other X-chromosomal genes in iPSCs, NPCs and neurons. The X-chromosome encodes roughly 800 genes of which a high proportion is associated with brain development and function (Inlow & Restifo, 2004). The second X-chromosome in females has been suggested to be a protection factor against neurodevelopmental disorders (Arnold, 2004; Arnold et al., 2016), that affect significantly more males. Females carrying a heterozygous mutation in X-linked genes usually present with approximately 50% of cells expressing the wildtype or the mutant allele, respectively. Nevertheless, this alone is not enough to fully explain the gender bias towards males in the prevalence of neurodevelopmental disorders (Fombonne, 2009).

4.3.1 X-chromosomal inactivation in iPSCs

Fibroblasts of healthy females, like all somatic cells, are typically a mixture of cells that randomly inactivated one or the other X-chromosome during embryonic development; in this study, the female heterozygous fibroblasts showed an expression of *MID1* that was nearly 50% wildtype and mutant, respectively. During reprogramming and upon isolation of single cell clones, clonal cell lines typically showing a completely skewed X-inactivation towards one of the two X-chromosomes can be generated (Tchieu et al., 2010). This also applied for the iPSCs that were generated from the female heterozygous fibroblasts. We were able to identify iPSC-clones that either express only the wildtype *MID1* or only the mutant *MID1*. Of note, the iPSCs not only showed a stable expression of either wildtype or mutant *MID1* on RNA and protein level, but also a dysregulated mTOR signalling as it was previously shown for the male fibroblasts with the 4 bp deletion (Liu et al., 2011).

As already mentioned (1.4.2), human PSCs can be sorted into three classes regarding their X-inactivation (Geens et al., 2017; Silva et al., 2008). Class I resembles that of mouse iPSCs with two stably active X-chromosomes of which one is randomly inactivated during further differentiation, leading to a cellular mosaic. The majority of iPSCs can be assigned to class II. This class is characterized by the X-inactivation pattern that is also found in somatic cells, with one active and one inactive X-chromosome. XCI is controlled by XIST and is typically kept during differentiation. Class III is mainly found in ESCs and is described by one active and one eroded X-chromosome. *XIST* is not expressed, and the overall methylation of the eroded X-chromosome is low. Upon differentiation, this status is either kept and further progresses (Mekhoubad et al., 2012), or full XCI is achieved independent of XIST (Bar et al., 2019).

Although we did not see any sign of erosion in the female iPSCs that were used in this study, they are very likely to be of class III, as there are major arguments against the other two classes. Class I can be excluded as this would have made it impossible to generate iPSC clones that express either mutant or wildtype *MID1*. All iPSC clones of the female heterozygous carrier would have shown expression of both, mutant and wildtype *MID1*. This is further supported by the RNAseq experiment that showed a high degree of monoallelic expression over the whole X-chromosome in the iPSCs. The RNAseq also provides the main argument against class II, as expression of *XIST* was not detected in the iPSCs (data not shown). Combining this information, no expression of XIST and monoallelic expression of *MID1* and the majority of X-chromosomal genes, leaves class III as the only plausible option. A study in 2015 showed that an early step in the erosion of XCI is re-expression of *XACT* from the inactive X-chromosome (Vallot et al., 2015). It would be interesting to see if *XACT* is expressed from both X-chromosomes in the female heterozygous iPSCs, to further prove that they are of class III. On the other hand, a monoallelic expression of *XACT* would not exclude class III but rather give a hint that the iPSCs are in an early stage of erosion of XCI. Other studies concluded that iPSCs, in line with ESCs, are suffering XCI instability with the transition of the class II Xi state towards the class III Xe state (Geens et al., 2017; Mekhoubad et al., 2012; Tchieu et al., 2010). The transition state described in these studies is probably the best fitting for the iPSCs analysed during this study. Recent publications suggest this process of transition as well as the whole class III to be an artifact of PSC-culture that does not appear in other cells types, neither *in vivo*, nor *in vitro*, except for some cancerous cells (Geens et al., 2017; Patrat et al., 2020). It is therefore important to tightly monitor XCI when culturing iPSCs, also after differentiating them.

Surprisingly, the stable monoallelic expression of MID1 was lost during neuronal differentiation.

4.3.2 X-chromosomal reactivation of MID1 in female carriers

Mid1 is known to escape XCI in mice due to its chromosomal position at the boundary of the pseudoautosomal region (Dal Zotto et al., 1998). However, no study has ever described *MID1* to be an escape gene in humans. Mutations in *MID1* are known to cause OS in hemizygous males; female heterozygous carriers typically present with mild symptoms, mainly hypertelorism (So et al., 2005; Winter et al., 2016). It is thought that the female mosaic caused by random XCI during embryonic development is a protection factor against a more severe phenotype. 50% of cells expressing the wildtype MID1 protein are thought to be enough to

rescue the phenotype caused by the 50% of cells expressing a mutant protein. In this study, we identified *MID1* to be a tissue-specific escape gene in neuronal cells, causing an intermediate phenotype in female heterozygous cerebral organoids. This finding further explains why female heterozygous carriers are less severely affected. The reactivation of the wildtype allele in cerebral organoids generated from mutant only expressing iPSCs leads to a milder phenotype compared to cerebral organoids generated from male iPSCs carrying the same mutation hemizygously. This shows that it is not only the female mosaic that is responsible for the less severe phenotype in females but also tissue-specific reactivation of (wildtype) *MID1* that plays a role in the manifestation of the disease. A tissue-specific reactivation of *MID1* during brain development would probably prohibit females from developing any neuropsychiatric symptoms that are connected to OS. Of note, reactivation of the mutant allele in cerebral organoids generated from wildtype only expressing iPSCs did not cause a phenotype significantly different from an unrelated wildtype control cell line, further emphasizing the loss-of-function mechanism underlying OS.

Escape genes show a high degree of heterogeneity, with about half of them escaping XCI in a tissue-specific manner (Tukiainen et al., 2017). Nevertheless, there are some general statements that can be made about escape genes and that apply to most of them (see chapter 1.4.3). *MID1* fulfils almost all criteria of an escape gene, further supporting our findings:

MID1, together with the majority of escape genes, is located on the evolutionary younger and less conserved p-arm of the X-chromosome (Balaton et al., 2015; Berletch et al., 2010).

Expression levels of *MID1* from Xi are typically lower than those from Xa (Balaton & Brown, 2016; Tukiainen et al., 2017).

There is a clinical and phenotypical diversity observed in OS patients and female carriers that probably partially results from escapee expression and is typical for escape genes (Berletch et al., 2010).

Depending on the literature, it is estimated that 15-30% of X-chromosomal genes escape XCI (Balaton et al., 2015; Berletch et al., 2010; Carrel & Willard, 2005; Cotton, Ge, et al., 2013; Tukiainen et al., 2017). Our data adds *MID1* to the list of tissue-specific escape genes, with a biallelic expression in neuronal cells. Human neuronal tissue is hardly available. The major study analysing XCI to identify escape genes, used the publicly available GTEx database (Tukiainen et al., 2017). The GTEx consortium has collected sequencing data from “54 non-diseased tissue sites across nearly 1000 individuals” (*GTEx Portal*, n.d.). A deeper look into

their statistics reveals that roughly 2/3 of their samples come from male donors and brain tissue was only taken from ~200 individuals, depending on the brain region. Knowing these numbers, it is not surprising *MID1* was not identified as an escape gene in earlier studies. There was probably no female donor carrying a mutation or a SNV in the *MID1* gene. This further emphasizes the need for human neuronal models as described in this study.

4.3.3 X-chromosomal reactivation during neuronal differentiation

Apart from the mosaic that is caused by random XCI, females are thought to have so far unknown protective factors, also known as the "female protective model" (Jacquemont et al., 2014). Robinson and colleagues used data of dizygotic twins from different studies and compared autistic traits between male and female probands (Robinson et al., 2013). Their data suggests that the "female sex protects girls from autistic impairments and that girls may require a greater familial etiological load to manifest the phenotype"(Robinson et al., 2013). A study published in 2014 confirmed these findings by analysing an autism spectrum disorder cohort (Jacquemont et al., 2014). They found a significant increase of deleterious SNVs and CNVs in female probands. Interestingly, these deleterious variants were more often maternally inherited. This is a good example that the manifestation of a neurodevelopmental disorder in females does not only depend on a mutation in a disease related gene, but also on the absence or presence of SNVs and CNVs in the whole genome.

Our findings suggest that tissue-specific reactivation of X-chromosomal genes might be another protective factor. Previous studies have already suggested an association between gene expression changes caused by escape from XCI and sexual dimorphism between males and females in aging processes (Berchtold et al., 2008; Davis et al., 2019) and disease vulnerability (Billi et al., 2019; Davis et al., 2020; Forsyth & Anguera, 2021; Hagen et al., 2020; Schurz et al., 2019).

The identification of *MID1* as a novel tissue-specific escape gene let us hypothesize that there might be other previously unknown escape genes that escape XCI only in neuronal cells and were not identified in other studies for the same reasons as *MID1*. A more sophisticated analysis of the RNAseq data did indeed reveal additional genes that escape XCI. The data was analysed as described in a study focussing on tissue-specific escape genes in mice (Berletch et al., 2015). The analysis was performed in collaboration with Prof. Dr. [redacted] and her working group. Using the statistical methods from the previously mentioned publication, we were able to identify a total of 37 genes that escape XCI in the NPCs and/or neurons of ctrl M10. The majority of these genes belong to two big functional classes. They either play a role in

neurodevelopmental processes or they are chromatin modulators. Especially the identification of genes involved in neurodevelopmental processes suggests that the reactivation is based on specific differentiation signals. We think that these signals target genes that are relevant for the correct development of, in this case, neuronal tissue. We hypothesise, differentiation into another tissue, would cause reactivation of a different subset of genes, specific for this tissue. The biallelic expression of genes that are relevant for brain development does not only compensate potential defects caused by mutations. There is also a more general protective effect because of the increased gene expression in comparison to males (Tukiainen et al., 2017). The identification of chromatin modulators expands this effect onto the whole genome. An increased expression of these genes potentially influences gene expression from all chromosomes, leading to a higher plasticity in the female brain and influencing not only X-chromosomal diseases but every disease.

Of the 37 genes that we found to escape XCI, 18 are located on the evolutionary younger p-arm and 17 are located on the more conserved q-arm. As already mentioned, one of the big clusters that is reactivated during neuronal differentiation is located on Xp22.2. Apart from *MID1*, our analysis revealed four more genes (*SHROOM2*, *TMSB4X*, *GPM6B*, *CA5BP1*) in this cluster to be biallelically expressed in NPCs and/or neurons. All of these genes are known to escape XCI (Tukiainen et al., 2017). Interestingly, while *MID1* is associated with OS (Quaderi et al., 1997; Robin et al., 1995), none of the other four genes has been described to be causative for a disease (*OMIM Entry - * 300051 - GLYCOPROTEIN M6B; GPM6B*, n.d.; *OMIM Entry - * 300103 - SHROOM FAMILY MEMBER 2; SHROOM2*, n.d.; *OMIM Entry - * 300159 - THYMOSIN, BETA-4, X CHROMOSOME; TMSB4X*, n.d.; *OMIM Entry - * 300230 - CARBONIC ANHYDRASE VB, MITOCHONDRIAL; CA5B*, n.d.). Except for the pseudogene of CA5B, which has no clear function (*CA5BP1 Gene - GeneCards | CA5BP1 Pseudogene*, n.d.), the genes are all involved in actin cytoskeleton function (*GPM6B Gene - GeneCards | GPM6B Protein | GPM6B Antibody*, n.d.; *SHROOM2 Gene - GeneCards | SHRM2 Protein | SHRM2 Antibody*, n.d.; *TMSB4X Gene - GeneCards | TYB4 Protein | TYB4 Antibody*, n.d.). In the brain, the actin cytoskeleton undergoes constant changes during embryonic development and in adults. It plays an important role in migration, differentiation and maturation of neuronal cells (Kirkcaldie & Dwyer, 2017; Kirkpatrick & Brady, 1999; S. Menon & Gupton, 2016). Disturbances of molecular pathways responsible for actin modulation are therefore likely to be lethal during early development, giving a possible explanation why no mutation in any of these genes has ever been reported.

The only other gene on the p-arm of the X-chromosome that was newly identified to escape XCI apart from *MID* is *GSPT2* (Tukiainen et al., 2017). GSPT2 plays an important role in translational termination by interacting with eRF1 (Hoshino et al., 1998). A study in 2005 found that silencing of GSPT2 does not have a significant effect on cellular function, while reduction of the closely related GSPT1 increases the readthrough and depletes eRF1 levels (Chauvin et al., 2005). Overexpression of GSPT2 rescued these phenotypes partially. This suggests that GSPT2 may function as a backup gene copy of GSPT1 although having at least partially different functions and expression properties (Hoshino et al., 1998).

In contrast to the p-arm, only two genes on the q-arm that we found to escape XCI have previously been reported as escapees (*JPX* and *ZNF711*) (Tukiainen et al., 2017). Among those newly identified is *IGBP1*, also known as *alpha4*, a subunit of PP2A that mediates its binding to MID1 (Trockenbacher et al., 2001). Interestingly, alpha4 only escapes XCI in neurons but not NPCs. The catalytic subunit of PP2A (PP2Ac) is known to be involved in axon formation and elongation (Zhu et al., 2010), and overexpression of PP2Ac in mouse neurons was shown to increase axonal length and branching (Lu et al., 2013). A tightly scheduled degradation of PP2Ac in developed neurons might be necessary to prevent them from developing too many branches and too long axons. Expression of *alpha4* from both X-chromosomes secures the proper degradation of PP2Ac; together with MID1.

Among the genes on the q-arm that were identified to escape XCI, are also five that are part of the Xq13.2 cluster that was previously mentioned. Four of these genes are lncRNAs and are part of the X-inactivation center that regulates XCI: *JPX*, *FTX*, *RP3-368A4.6* and *RP3-368A4.5*. While the exact function of *RP3-368A4.6* and *RP3-368A4.5* is unknown, *JPX* and *FTX* are both known to positively regulate expression of XIST (Chureau et al., 2011; Tian et al., 2010). *JPX* has already been described as an escape gene (Tukiainen et al., 2017), and is known to increase expression of *XIST*. It acts as a counterplayer of *TSIX*, that is transcribed from the active X-chromosome and represses expression of *XIST* (Tian et al., 2010). *FTX* is also known to upregulate expression of *XIST*, most likely by preventing methylation of the *XIST* promoter (Chureau et al., 2011). In the same study, *FTX* was shown to partially escape XCI in the mouse, but this was not reported for humans (Tukiainen et al., 2017). Although both known positive regulators of *XIST* were found to escape XCI in neurons, *XIST* expression was not detected. XCI in the neurons generated during this study remained independent of *XIST*.

Further studies are needed to confirm these preliminary findings. We are currently repeating these experiments using wildtype female cells to exclude that our findings are caused by the

MID1 mutation present in the ctrl M10 cells. [redacted] is performing follow up experiments and tries to unveil the molecular mechanism behind the tissue-specific escape by looking at epigenetic changes that correlate with the RNAseq results. Taken together, these results offer a new possibility for future research on XCI in neuronal cells and provide another protective factor of female brain development.

5 Conclusion

After successfully reprogramming wildtype and patient-specific fibroblasts into iPSCs, we generated neuronal cells, either in 2D culture, or as a 3D culture of cerebral organoids.

Using those cultures, we were able to identify a previously unknown gain-of-function mechanism in cells carrying N-terminally truncating MID1 mutations. The observed phenotype is likely to be lethal during embryonic development, explaining, why such mutations have never been found in patients. Our data gives plausible evidence that there are alternative translation initiation sites in the first coding exon of *MID1*, that are commonly being used, but in the absence of wildtype MID1 are the underlying cause for the phenotype.

Furthermore, we were able to mimic the "classical" OS phenotype, showing that it is an increase of PAX6 that prevents the cells from exiting the cell cycle and to differentiate into neurons that might be causative for the symptoms. We also identified *MID1* to be an escape gene in the female brain, giving an explanation why females are less severely affected of OS. Apart from MID1, we identified gene clusters and single genes on the X-chromosome that escape XCI during neuronal development, pointing towards a more general mechanism.

Taken together, this study established a workflow that is easy to transfer onto every monogenic disorder affecting the brain. The findings on OS are relevant for future research and help to better understand the mechanisms leading to OS.

It needs some more experiments and the generation of additional edited iPSC-lines to properly complete research on these cells and confirm our hypotheses. All of these experiments are currently being performed and we plan to publish them in a peer-reviewed journal as soon as possible.

6 References

Altieri, F., D'Anzi, A., Martello, F., Tardivo, S., Spasari, I., Ferrari, D., Bernardini, L., Lamorte, G., Mazzoccoli, G., Valente, E. M., Vescovi, A. L., & Rosati, J. (2019). Production and characterization of human induced pluripotent stem cells (iPSC) CSSi007-A (4383) from Joubert Syndrome. *Stem Cell Research*, *38*(May), 101480. https://doi.org/10.1016/j.scr.2019.101480

Ananiev, G., Williams, E. C., Li, H., & Chang, Q. (2011). Isogenic pairs of wild type and mutant induced pluripotent stem cell (iPSC) lines from rett syndrome patients as In Vitro disease model. *PLoS ONE*, *6*(9). https://doi.org/10.1371/journal.pone.0025255

Aranda-Orgille, B., Rutschow, D., Zeller, R., Karagiannidis, A. I., Köhler, A., Chen, C., Wilson, T., Krause, S., Roepcke, S., Lilley, D., Schneider, R., & Schweiger, S. (2011). Protein Phosphatase 2A (PP2A)-specific ubiquitin ligase MID1 is a sequence-dependent regulator of translation efficiency controlling 3-Phosphoinositide-dependent Protein Kinase-1 (PDPK-1). *Journal of Biological Chemistry*, *286*(46), 39945–39957. https://doi.org/10.1074/jbc.M111.224451

Arnold, A. P. (2004). Sex chromosomes and brain gender. In *Nature Reviews Neuroscience* (Vol. 5, Issue 9, pp. 701–708). Nature Publishing Group. https://doi.org/10.1038/nrn1494

Arnold, A. P., Reue, K., Eghbali, M., Vilain, E., Chen, X., Ghahramani, N., Itoh, Y., Li, J., Link, J. C., Ngun, T., & Williams-Burris, S. M. (2016). The importance of having two X chromosomes. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 371, Issue 1688). Royal Society of London. https://doi.org/10.1098/rstb.2015.0113

Auburger, G., Klinkenberg, M., Drost, J., Marcus, K., Morales-Gordo, B., Kunz, W. S., Brandt, U., Broccoli, V., Reichmann, H., Gispert, S., & Jendrach, M. (2012). Primary skin fibroblasts as a model of Parkinson's disease. *Molecular Neurobiology*, *46*(1), 20–27. https://doi.org/10.1007/s12035-012-8245-1

Bading, H. (2013). Nuclear calcium signalling in the regulation of brain function. In *Nature Reviews Neuroscience* (Vol. 14, Issue 9, pp. 593–608). Nature Publishing Group. https://doi.org/10.1038/nrn3531

Baghbaderani, B. A., Syama, A., Sivapatham, R., Pei, Y., Mukherjee, O., Fellner, T., Zeng, X., & Rao, M. S. (2016). Detailed Characterization of Human Induced Pluripotent Stem Cells

Manufactured for Therapeutic Applications. *Stem Cell Reviews and Reports*, *12*(4), 394–420. https://doi.org/10.1007/s12015-016-9662-8

Bagley, J. A., Reumann, D., Bian, S., Lévi-Strauss, J., & Knoblich, J. A. (2017). Fused cerebral organoids model interactions between brain regions. *Nature Methods*, *14*(7), 743–751. https://doi.org/10.1038/nmeth.4304

Bahmad, H., Hadadeh, O., Chamaa, F., Cheaito, K., Darwish, B., Makkawi, A. K., & Abou-Kheir, W. (2017). Modeling human neurological and neurodegenerative diseases: From induced pluripotent stem cells to neuronal differentiation and its applications in neurotrauma. In *Frontiers in Molecular Neuroscience* (Vol. 10, p. 50). Frontiers Research Foundation. https://doi.org/10.3389/fnmol.2017.00050

Bailey, J. A., Carrel, L., Chakravarti, A., & Eichler, E. E. (2000). Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: The Lyon repeat hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(12), 6634–6639. https://doi.org/10.1073/pnas.97.12.6634

Balaton, B. P., & Brown, C. J. (2016). Escape Artists of the X Chromosome. In *Trends in Genetics* (Vol. 32, Issue 6, pp. 348–359). Elsevier Ltd. https://doi.org/10.1016/j.tig.2016.03.007

Balaton, B. P., Cotton, A. M., & Brown, C. J. (2015). *Derivation of consensus inactivation status for X-linked genes from genome-wide studies*. https://doi.org/10.1186/s13293-015-0053-7

Ballabio, A., & Willard, H. F. (1992). Mammalian X-chromosome inactivation and the XIST gene. *Current Opinion in Genetics & Development*, *2*(3), 439–447. https://doi.org/10.1016/s0959-437x(05)80155-8

Bar, S., Seaton, L. R., Weissbein, U., Eldar-Geva, T., & Benvenisty, N. (2019). Global Characterization of X Chromosome Inactivation in Human Pluripotent Stem Cells. *Cell Reports*, *27*(1), 20-29.e3. https://doi.org/10.1016/j.celrep.2019.03.019

Barakat, T. S., Ghazvini, M., de Hoon, B., Li, T., Eussen, B., Douben, H., van der Linden, R., van der Stap, N., Boter, M., Laven, J. S., Galjaard, R. J., Grootegoed, J. A., de Klein, A., & Gribnau, J. (2015). Stable X chromosome reactivation in female human induced pluripotent stem cells. *Stem Cell Reports*, *4*(2), 199–208. https://doi.org/10.1016/j.stemcr.2014.12.012

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*(5819), 1709–1712. https://doi.org/10.1126/science.1138140

Bellott, D. W., Hughes, J. F., Skaletsky, H., Brown, L. G., Pyntikova, T., Cho, T. J., Koutseva, N., Zaghlul, S., Graves, T., Rock, S., Kremitzki, C., Fulton, R. S., Dugan, S., Ding, Y., Morton, D., Khan, Z., Lewis, L., Buhay, C., Wang, Q., … Page, D. C. (2014). Mammalian y chromosomes retain widely expressed dosage-sensitive regulators. *Nature*, *508*(7497), 494–499. https://doi.org/10.1038/nature13206

ben Jehuda, R., Shemer, Y., & Binah, O. (2018). Genome Editing in Induced Pluripotent Stem Cells using CRISPR/Cas9. In *Stem Cell Reviews and Reports* (Vol. 14, Issue 3, pp. 323–336). Humana Press Inc. https://doi.org/10.1007/s12015-018-9811-3

Berchtold, N. C., Cribbs, D. H., Coleman, P. D., Rogers, J., Head, E., Kim, R., Beach, T., Miller, C., Troncoso, J., Trojanowski, J. Q., Zielke, H. R., & Cotman, C. W. (2008). Gene expression changes in the course of normal brain aging are sexually dimorphic. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(40), 15605–15610. https://doi.org/10.1073/pnas.0806883105

Berletch, J. B., Ma, W., Yang, F., Shendure, J., Noble, W. S., Disteche, C. M., & Deng, X. (2015). Escape from X Inactivation Varies in Mouse Tissues. *PLoS Genetics*, *11*(3). https://doi.org/10.1371/journal.pgen.1005079

Berletch, J. B., Yang, F., & Disteche, C. M. (2010). Escape from X inactivation in mice and humans. In *Genome Biology* (Vol. 11, Issue 6, p. 213). BioMed Central. https://doi.org/10.1186/gb-2010-11-6-213

Bijlsma, E. K., Gijsbers, A. C. J., Schuurs-Hoeijmakers, J. H. M., van Haeringen, A., Fransen van de Putte, D. E., Anderlid, B. M., Lundin, J., Lapunzina, P., Pérez Jurado, L. A., Delle Chiaie, B., Loeys, B., Menten, B., Oostra, A., Verhelst, H., Amor, D. J., Bruno, D. L., van Essen, A. J., Hordijk, R., Sikkema-Raddatz, B., … Ruivenkamp, C. A. L. (2009). Extending the phenotype of recurrent rearrangements of 16p11.2: Deletions in mentally retarded patients without autism and in normal individuals. *European Journal of Medical Genetics*, *52*(2–3), 77–87. https://doi.org/10.1016/j.ejmg.2009.03.006

Billi, A. C., Kahlenberg, J. M., & Gudjonsson, J. E. (2019). Sex bias in autoimmunity. In *Current opinion in rheumatology* (Vol. 31, Issue 1, pp. 53–61). NLM (Medline). https://doi.org/10.1097/BOR.0000000000000564

Boonen, M., Staudt, C., Gilis, F., Oorschot, V., Klumperman, J., & Jadot, M. (2016). Cathepsin D and its newly identified transport receptor SEZ6L2 can modulate neurite outgrowth. *Journal of Cell Science*, *129*(3), 557–568. https://doi.org/10.1242/jcs.179374

Briggs, S. F., Dominguez, A. A., Chavez, S. L., & Reijo Pera, R. A. (2015). Single-Cell *XIST* Expression in Human Preimplantation Embryos and Newly Reprogrammed Female Induced Pluripotent Stem Cells. *STEM CELLS*, *33*(6), 1771–1781. https://doi.org/10.1002/stem.1992

Brouwer, M., Zhou, H., & Nadif Kasri, N. (2016). Choices for Induction of Pluripotency: Recent Developments in Human Induced Pluripotent Stem Cell Reprogramming Strategies. In *Stem Cell Reviews and Reports* (Vol. 12, Issue 1, pp. 54–72). Humana Press Inc. https://doi.org/10.1007/s12015-015-9622-8

Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., Tonlorenzi, R., & Willard, H. F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature*, *349*(6304), 38–44. https://doi.org/10.1038/349038a0

Buchner, G., Montini, E., Andolfi, G., Quaderi, N., Cainarca, S., Messali, S., Bassi, M. T., Ballabio, A., Meroni, G., & Franco, B. (1999). MID2, a homologue of the Opitz syndrome gene MID1: similarities in subcellular localization and differences in expression during development. In *Human Molecular Genetics* (Vol. 8, Issue 8).

CA5BP1 Gene - GeneCards | CA5BP1 Pseudogene. (n.d.). Retrieved March 3, 2021, from https://www.genecards.org/cgi-bin/carddisp.pl?gene=CA5BP1

Caballero, I. M., Manuel, M. N., Molinek, M., Quintana-Urzainqui, I., Mi, D., Shimogori, T., & Price, D. J. (2014). Cell-Autonomous Repression of Shh by Transcription Factor Pax6 Regulates Diencephalic Patterning by Controlling the Central Diencephalic Organizer. *Cell Reports*, *8*(5), 1405–1418. https://doi.org/10.1016/j.celrep.2014.07.051

Cantone, I., Bagci, H., Dormann, D., Dharmalingam, G., Nesterova, T., Brockdorff, N., Rougeulle, C., Vallot, C., Heard, E., Chaligne, R., Merkenschlager, M., & Fisher, A. G. (2016). Ordered chromatin changes and human X chromosome reactivation by cell fusion-

mediated pluripotent reprogramming. *Nature Communications*, *7*(1), 1–12. https://doi.org/10.1038/ncomms12354

Carrel, L., & Willard, H. F. (2005). X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature*, *434*(7031), 400–404. https://doi.org/10.1038/nature03479

Chambers, S. M., Fasano, C. A., Papapetrou, E. P., Tomishima, M., Sadelain, M., & Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature Biotechnology*, *27*(3), 275–280. https://doi.org/10.1038/nbt.1529

Chauvin, C., Salhi, S., le Goff, C., Viranaicken, W., Diop, D., & Jean-Jean, O. (2005). Involvement of Human Release Factors eRF3a and eRF3b in Translation Termination and Regulation of the Termination Complex Formation. *Molecular and Cellular Biology*, *25*(14), 5801–5811. https://doi.org/10.1128/mcb.25.14.5801-5811.2005

Chureau, C., Chantalat, S., Romito, A., Galvani, A., Duret, L., Avner, P., & Rougeulle, C. (2011). Ftx is a non-coding RNA which affects Xist expression and chromatin structure within the X-inactivation center region. *Human Molecular Genetics*, *20*(4), 705–718. https://doi.org/10.1093/hmg/ddq516

c-myc - C-myc protein - Uniprot. (n.d.). Retrieved April 1, 2020, from https://www.uniprot.org/uniprot/Q6LBK7

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, *339*(6121), 819–823. https://doi.org/10.1126/science.1231143

Cooper, O., Hargus, G., Deleidi, M., Blak, A., Osborn, T., Marlow, E., Lee, K., Levy, A., Perez-Torres, E., Yow, A., & Isacson, O. (2010). Differentiation of human ES and Parkinson's disease iPS cells into ventral midbrain dopaminergic neurons requires a high activity form of SHH, FGF8a and specific regionalization by retinoic acid. *Molecular and Cellular Neuroscience*, *45*(3), 258–266. https://doi.org/10.1016/j.mcn.2010.06.017

Cotton, A. M., Chen, C.-Y., Lam, L. L., Wasserman, W. W., Kobor, M. S., & Brown, C. J. (2013). *Spread of X-chromosome inactivation into autosomal sequences: role for DNA elements, chromatin features and chromosomal domains*. https://doi.org/10.1093/hmg/ddt513

Cotton, A. M., Ge, B., Light, N., Adoue, V., Pastinen, T., & Brown, C. J. (2013). Analysis of expressed SNPs identifies variable extents of expression from the human inactive X chromosome. *Genome Biology*, *14*(11), R122. https://doi.org/10.1186/gb-2013-14-11-r122

Cox, T. C. (2000). New mutations in MID1 provide support for loss of function as the cause of X-linked Opitz syndrome. *Human Molecular Genetics*, *9*(17), 2553–2562. https://doi.org/10.1093/hmg/9.17.2553

Dafforn, T. R., Della, M., & Miller, A. D. (2001). The Molecular Interactions of Heat Shock Protein 47 (Hsp47) and Their Implications for Collagen Biosynthesis. *Journal of Biological Chemistry*, *276*(52), 49310–49319. https://doi.org/10.1074/jbc.M108896200

Dal Zotto, L., Quaderi, N. A., Elliott, R., Lingerfelter, P. A., Carrel, L., Valsecchi, V., Montini, E., Yen, C.-H., Chapman, V., Kalcheva, I., Arrigo, G., Zuffardi, O., Thomas, S., Willard, H. F., Ballabio, A., Disteche, C. M., & Rugarli, E. I. (1998). The mouse Mid1 gene: implications for the pathogenesis of Opitz syndrome and the evolution of the mammalian pseudoautosomal region. In *Human Molecular Genetics* (Vol. 7, Issue 3).

Davis, E. J., Broestl, L., Williams, G., Garay, B. I., Lobach, I., Devidze, N., Kim, D., Bonham, L. W., Miñones-Moyano, E., Moreno, A. J., Wang, D., Chang, K., Abdulai-Saiku, S., Worden, K., Anderson-Bergman, C., Yu, G. Q., White, C. C., Harris, J. A., Miller, B. L., … Dubal, D. B. (2020). A second X chromosome contributes to resilience in a mouse model of Alzheimer's disease. *Science Translational Medicine*, *12*(558). https://doi.org/10.1126/SCITRANSLMED.AAZ5677

Davis, E. J., Lobach, I., & Dubal, D. B. (2019). Female XX sex chromosomes increase survival and extend lifespan in aging mice. *Aging Cell*, *18*(1). https://doi.org/10.1111/acel.12871

de Falco, F., Cainarca, S., Andolfi, G., Ferrentino, R., Berti, C., Rodríguez Criado, G., Rittinger, O., Dennis, N., Odent, S., Rastogi, A., Liebelt, J., Chitayat, D., Winter, R., Jawanda, H., Ballabio, A., Franco, B., & Meroni, G. (2003). X-linked opitz syndrome: Novel mutations in the MID1 gene and redefinition of the clinical spectrum. *American Journal of Medical Genetics*, *120 A*(2), 222–228. https://doi.org/10.1002/ajmg.a.10265

Deshpande, A., Yadav, S., Dao, D. Q., Wu, Z. Y., Hokanson, K. C., Cahill, M. K., Wiita, A. P., Jan, Y. N., Ullian, E. M., & Weiss, L. A. (2017). Cellular Phenotypes in Human iPSC-

Derived Neurons from a Genetic Model of Autism Spectrum Disorder. *Cell Reports*, *21*(10), 2678–2687. https://doi.org/10.1016/j.celrep.2017.11.037

Dimos, J. T., Rodolfa, K. T., Niakan, K. K., Weisenthal, L. M., Mitsumoto, H., Chung, W., Croft, G. F., Saphier, G., Leibel, R., Goland, R., Wichterle, H., Henderson, C. E., & Eggan, K. (2008). Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science*, *321*(5893), 1218–1221. https://doi.org/10.1126/science.1158799

Duncan, R. R., Shipston, M. J., & Chow, R. H. (2000). Double C2 protein. A review. *Biochimie*, *82*(5), 421–426. https://doi.org/10.1016/S0300-9084(00)00214-5

Ebert, A. D., Yu, J., Rose, F. F., Mattis, V. B., Lorson, C. L., Thomson, J. A., & Svendsen, C. N. (2009). Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature*, *457*(7227), 277–280. https://doi.org/10.1038/nature07677

Fernandez, B. A., Roberts, W., Chung, B., Weksberg, R., Meyn, S., Szatmari, P., Joseph-George, A. M., MacKay, S., Whitten, K., Noble, B., Vardy, C., Crosbie, V., Luscombe, S., Tucker, E., Turner, L., Marshall, C. R., & Scherer, S. W. (2010). Phenotypic spectrum associated with de novo and inherited deletions and duplications at 16p11.2 in individuals ascertained for diagnosis of autism spectrum disorder. *Journal of Medical Genetics*, *47*(3), 195–203. https://doi.org/10.1136/jmg.2009.069369

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. In *Pediatric Research* (Vol. 65, Issue 6, pp. 591–598). Nature Publishing Group. https://doi.org/10.1203/PDR.0b013e31819e7203

Fontanella, B., Russolillo, G., & Meroni, G. (2008). MID1 mutations in patients with X-linked Opitz G/BBB syndrome. *Human Mutation*, *29*(5), 584–594. https://doi.org/10.1002/humu.20706

Forsyth, K. S., & Anguera, M. C. (2021). Time to get ill: the intersection of viral infections, sex, and the X chromosome. In *Current Opinion in Physiology* (Vol. 19, pp. 62–72). Elsevier Ltd. https://doi.org/10.1016/j.cophys.2020.09.015

Geens, M., Chuva, S. M., & Lopes, D. S. (2017). X chromosome inactivation in human pluripotent stem cells as a model for human development: back to the drawing board? *Human Reproduction Update*, *23*(5), 520–532. https://doi.org/10.1093/humupd/dmx015

Ghaleb, A. M., & Yang, V. W. (2017). Krüppel-like factor 4 (KLF4): What we currently know. In *Gene* (Vol. 611, pp. 27–37). Elsevier B.V. https://doi.org/10.1016/j.gene.2017.02.025

Glessner, J. T., Reilly, M. P., Kim, C. E., Takahashi, N., Albano, A., Hou, C., Bradfield, J. P., Zhang, H., Sleiman, P. M. A., Flory, J. H., Imielinski, M., Frackelton, E. C., Chiavacci, R., Thomas, K. A., Garris, M., Otieno, F. G., Davidson, M., Weiser, M., Reichenberg, A., … Hakonarson, H. (2010). Strong synaptic transmission impact by copy number variations in schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(23), 10584–10589. https://doi.org/10.1073/pnas.1000274107

Golzio, C., Willer, J., Talkowski, M. E., Oh, E. C., Taniguchi, Y., Jacquemont, S., Reymond, A., Sun, M., Sawa, A., Gusella, J. F., Kamiya, A., Beckmann, J. S., & Katsanis, N. (2012). KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. *Nature*, *485*(7398), 363–367. https://doi.org/10.1038/nature11091

GPM6B Gene - GeneCards | GPM6B Protein | GPM6B Antibody. (n.d.). Retrieved March 3, 2021, from https://www.genecards.org/cgi-bin/carddisp.pl?gene=gpm6b

Groffen, A. J. A., Friedrich, R., Brian, E. C., Ashery, U., & Verhage, M. (2006). DOC2A and DOC2B are sensors for neuronal activity with unique calcium-dependent and kinetic properties. *Journal of Neurochemistry*, *97*(3), 818–833. https://doi.org/10.1111/j.1471-4159.2006.03755.x

GTEx Portal. (n.d.). Retrieved February 22, 2021, from https://www.gtexportal.org/home/

Guide design resources — Zhang Lab. (n.d.). Retrieved November 13, 2020, from https://zlab.bio/guide-design-resources

Guirouilh-Barbat, J., Huck, S., Bertrand, P., Pirzio, L., Desmaze, C., Sabatier, L., & Lopez, B. S. (2004). Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells. *Molecular Cell*, *14*(5), 611–623. https://doi.org/10.1016/j.molcel.2004.05.008

Guo, X. L., & Chen, J. S. (2015). Research on induced pluripotent stem cells and the application in ocular tissues. In *International Journal of Ophthalmology* (Vol. 8, Issue 4, pp. 818–825). International Journal of Ophthalmology (c/o Editorial Office). https://doi.org/10.3980/j.issn.2222-3959.2015.04.31

Hagen, S. H., Henseling, F., Hennesen, J., Savel, H., Delahaye, S., Richert, L., Ziegler, S. M., & Altfeld, M. (2020). Heterogeneous Escape from X Chromosome Inactivation Results in

Sex Differences in Type I IFN Responses at the Single Human pDC Level. *Cell Reports*, *33*(10), 108485. https://doi.org/10.1016/j.celrep.2020.108485

Han, S. S. W., Williams, L. A., & Eggan, K. C. (2011). Constructing and Deconstructing Stem Cell Models of Neurological Disease. In *Neuron* (Vol. 70, Issue 4, pp. 626–644). Cell Press. https://doi.org/10.1016/j.neuron.2011.05.003

Havlicek, S., Kohl, Z., Mishra, H. K., Prots, I., Eberhardt, E., Denguir, N., Wend, H., Plötz, S., Boyer, L., Marchetto, M. C. N., Aigner, S., Sticht, H., Groemer, T. W., Hehr, U., Lampert, A., Schlötzer-schrehardt, U., Winkler, J., Gage, F. H., & Winner, B. (2014). Gene dosage-dependent rescue of HSP neurite defects in SPG4 patients' neurons. *Human Molecular Genetics*, *23*(10), 2527–2541. https://doi.org/10.1093/hmg/ddt644

He, J.-Q., Ma, Y., Lee, Y., Thomson, J. A., & Kamp, T. J. (2003). Human Embryonic Stem Cells Develop Into Multiple Types of Cardiac Myocytes. *Circulation Research*, *93*(1), 32–39. https://doi.org/10.1161/01.RES.0000080317.92718.99

Heldin, C. H., Miyazono, K., & ten Dijke, P. (1997). TGF-β signalling from cell membrane to nucleus through SMAD proteins. In *Nature* (Vol. 390, Issue 6659, pp. 465–471). Nature Publishing Group. https://doi.org/10.1038/37284

Hill, W. D., Weiss, A., Liewald, D. C., Davies, G., Porteous, D. J., Hayward, C., McIntosh, A. M., Gale, C. R., & Deary, I. J. (2020). Genetic contributions to two special factors of neuroticism are associated with affluence, higher intelligence, better health, and longer life. *Molecular Psychiatry*, *25*(11), 3034–3052. https://doi.org/10.1038/s41380-019-0387-3

Holm, T. M., Habashi, J. P., Doyle, J. J., Bedja, D., Chen, Y. C., van Erp, C., Lindsay, M. E., Kim, D., Schoenhoff, F., Cohn, R. D., Loeys, B. L., Thomas, C. J., Patnaik, S., Marugan, J. J., Judge, D. P., & Dietz, H. C. (2011). Noncanonical TGFβ signaling contributes to aortic aneurysm progression in marfan syndrome mice. *Science*, *332*(6027), 358–361. https://doi.org/10.1126/science.1192149

Home < Expression Atlas < EMBL-EBI. (n.d.). Retrieved January 21, 2021, from https://www.ebi.ac.uk/gxa/home

Hong, Y. J., & Do, J. T. (2019). Neural Lineage Differentiation From Pluripotent Stem Cells to Mimic Human Brain Tissues. In *Frontiers in Bioengineering and Biotechnology* (Vol. 7, p. 400). Frontiers Media S.A. https://doi.org/10.3389/fbioe.2019.00400

Horev, G., Ellegood, J., Lerch, J. P., Son, Y. E. E., Muthuswamy, L., Vogel, H., Krieger, A. M., Buja, A., Henkelman, R. M., Wigler, M., & Mills, A. A. (2011). Dosage-dependent phenotypes in models of 16p11.2 lesions found in autism. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(41), 17076–17081. https://doi.org/10.1073/pnas.1114042108

Horvath, L. M., Li, N., & Carrel, L. (2013). Deletion of an X-Inactivation Boundary Disrupts Adjacent Gene Silencing. *PLoS Genetics*, *9*(11), e1003952. https://doi.org/10.1371/journal.pgen.1003952

Hoshino, S. I., Imai, M., Mizutani, M., Kikuchi, Y., Hanaoka, F., Ui, M., & Katada, T. (1998). Molecular cloning of a novel member of the eukaryotic polypeptide chain- releasing factors (eRF): Its identification as eRF3 interacting with eRF1. *Journal of Biological Chemistry*, *273*(35), 22254–22259. https://doi.org/10.1074/jbc.273.35.22254

Human Genome Assembly GRCh38 - Genome Reference Consortium. (n.d.). Retrieved April 24, 2020, from https://www.ncbi.nlm.nih.gov/grc/human/data?asm=GRCh38

Hussein, S. M., Batada, N. N., Vuoristo, S., Ching, R. W., Autio, R., Narvä, E., Ng, S., Sourour, M., Hämälä, R., Olsson, C., Lundin, K., Mikkola, M., Trokovic, R., Peitz, M., Brüstle, O., Bazett-Jones, D. P., Alitalo, K., Lahesmaa, R., Nagy, A., & Otonkoski, T. (2011). Copy number variation and selection during reprogramming to pluripotency. *Nature*, *471*(7336), 58–62. https://doi.org/10.1038/nature09871

Inlow, J. K., & Restifo, L. L. (2004). Molecular and Comparative Genetics of Mental Retardation. *Genetics*, *166*(2), 835–881. https://doi.org/10.1534/genetics.166.2.835

Iyama, T., & Wilson, D. M. (2013). DNA repair mechanisms in dividing and non-dividing cells. *DNA Repair*, *12*(8), 620–636. https://doi.org/10.1016/j.dnarep.2013.04.015

Jack, J., Rotroff, D., & Motsinger-Reif, A. (2014). Lymphoblastoid Cell Lines Models of Drug Response: Successes and Lessons from this Pharmacogenomic Model. *Current Molecular Medicine*, *14*(7), 833–840. https://doi.org/10.2174/1566524014666140811113946

Jacquemont, S., Coe, B. P., Hersch, M., Duyzend, M. H., Krumm, N., Bergmann, S., Beckmann, J. S., Rosenfeld, J. A., & Eichler, E. E. (2014). A higher mutational burden in females supports a “female protective model” in neurodevelopmental disorders. *American Journal of Human Genetics*, *94*(3), 415–425. https://doi.org/10.1016/j.ajhg.2014.02.001

Jaubert, J., Cheng, J., & Segre, J. A. (2003). Ectopic expression of Kruppel like factor 4 (Klf4) accelerates formation of the epidermal permeability barrier. *Development*, *130*(12), 2767–2777. https://doi.org/10.1242/dev.00477

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*(6096), 816–821. https://doi.org/10.1126/science.1225829

Kang, S., Chen, X., Gong, S., Yu, P., Yau, S., Su, Z., Zhou, L., Yu, J., Pan, G., & Shi, L. (2017). Characteristic analyses of a neural differentiation model from iPSC-derived neuron according to morphology, physiology, and global gene expression pattern. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-12452-x

Käseberg, S. (2015). *Etablierung der induzierten pluripotenten Stammzellkultur und Generierung von humanen neuronalen Modellen zur Erforschung neurogenetischer Erkrankungen*.

Kilpinen, H., Goncalves, A., Leha, A., Afzal, V., Alasoo, K., Ashford, S., Bala, S., Bensaddek, D., Casale, F. P., Culley, O. J., Danecek, P., Faulconbridge, A., Harrison, P. W., Kathuria, A., McCarthy, D., McCarthy, S. A., Meleckyte, R., Memari, Y., Moens, N., … Gaffney, D. J. (2017). Common genetic variation drives molecular heterogeneity in human iPSCs. *Nature*, *546*(7658), 370–375. https://doi.org/10.1038/nature22403

Kim, J. B., Sebastiano, V., Wu, G., Araúzo-Bravo, M. J., Sasse, P., Gentile, L., Ko, K., Ruau, D., Ehrich, M., van den Boom, D., Meyer, J., Hübner, K., Bernemann, C., Ortmeier, C., Zenke, M., Fleischmann, B. K., Zaehres, H., & Schöler, H. R. (2009). Oct4-Induced Pluripotency in Adult Neural Stem Cells. *Cell*, *136*(3), 411–419. https://doi.org/10.1016/j.cell.2009.01.023

Kim, K. Y., Hysolli, E., Tanaka, Y., Wang, B., Jung, Y. W., Pan, X., Weissman, S. M., & Park, I. H. (2014). X chromosome of female cells shows dynamic changes in status during human somatic cell reprogramming. *Stem Cell Reports*, *2*(6), 896–909. https://doi.org/10.1016/j.stemcr.2014.04.003

Kirkcaldie, M. T. K., & Dwyer, S. T. (2017). The third wave: Intermediate filaments in the maturing nervous system. In *Molecular and Cellular Neuroscience* (Vol. 84, pp. 68–76). Academic Press Inc. https://doi.org/10.1016/j.mcn.2017.05.010

Kirkpatrick, L. L., & Brady, S. T. (1999). *Molecular Components of the Neuronal Cytoskeleton*. https://www.ncbi.nlm.nih.gov/books/NBK28122/

KLF4 - Krueppel-like factor 4 - Uniprot. (n.d.). Retrieved April 1, 2020, from https://www.uniprot.org/uniprot/O43474

KLF4 Gene - GeneCards. (n.d.). Retrieved April 1, 2020, from https://www.genecards.org/cgi-bin/carddisp.pl?gene=KLF4

Kochetov, A. v. (2006). Alternative translation start sites and their significance for eukaryotic proteomes. In *Molecular Biology* (Vol. 40, Issue 5, pp. 705–712). Springer. https://doi.org/10.1134/S0026893306050049

Konyukh, M., Delorme, R., Chaste, P., Leblond, C., Lemière, N., Nygren, G., Anckarsäter, H., Rastam, M., Ståhlberg, O., Amsellem, F., Gillberg, I. C., Mouren-Simeoni, M. C., Herbrecht, E., Fauchereau, F., Toro, R., Gillberg, C., Leboyer, M., & Bourgeron, T. (2011). Variations of the Candidate SEZ6L2 Gene on Chromosome 16p11.2 in Patients with Autism Spectrum Disorders and in Human Populations. *PLoS ONE*, *6*(3), e17289. https://doi.org/10.1371/journal.pone.0017289

Kozak, M. (1999). Initiation of translation in prokaryotes and eukaryotes. In *Gene* (Vol. 234, Issue 2, pp. 187–208). Elsevier. https://doi.org/10.1016/S0378-1119(99)00210-3

Kozak, M. (2002). Pushing the limits of the scanning mechanism for initiation of translation. In *Gene* (Vol. 299, Issues 1–2, pp. 1–34). Elsevier. https://doi.org/10.1016/S0378-1119(02)01056-9

Krauß, S., Foerster, J., Schneider, R., & Schweiger, S. (2008). Protein phosphatase 2A and rapamycin regulate the nuclear localization and activity of the transcription factor GLI3. *Cancer Research*, *68*(12), 4658–4665. https://doi.org/10.1158/0008-5472.CAN-07-6174

Kumar, R. A., Karamohamed, S., Sudi, J., Conrad, D. F., Brune, C., Badner, J. A., Gilliam, T. C., Nowak, N. J., Cook, E. H., Dobyns, W. B., & Christian, S. L. (2008). Recurrent 16p11.2 microdeletions in autism. *Human Molecular Genetics*, *17*(4), 628–638. https://doi.org/10.1093/hmg/ddm376

Kumar, R. A., Marshall, C. R., Badner, J. A., Babatz, T. D., Mukamel, Z., Aldinger, K. A., Sudi, J., Brune, C. W., Goh, G., KaraMohamed, S., Sutcliffe, J. S., Cook, E. H., Geschwind, D. H., Dobyns, W. B., Scherer, S. W., & Christian, S. L. (2009). Association

and mutation analyses of 16p11.2 autism candidate genes. *PLoS ONE*, *4*(2). https://doi.org/10.1371/journal.pone.0004582

Kuroda, K., & Tajima, S. (2004). HSP47 is a useful marker for skin fibroblasts in formalin-fixed, paraffin-embedded tissue specimens. *Journal of Cutaneous Pathology*, *31*(3), 241–246. https://doi.org/10.1111/j.0303-6987.2003.00166.x

Kusenda, M., Vacic, V., Malhotra, D., Rodgers, L., Pavon, K., Meth, J., Kumar, R. A., Christian, S. L., Peeters, H., Cho, S. S., Addington, A., Rapoport, J. L., & Sebat, J. (2015). The Influence of Microdeletions and Microduplications of 16p11.2 on Global Transcription Profiles. *Journal of Child Neurology*, *30*(14), 1947–1953. https://doi.org/10.1177/0883073815602066

Lancaster, M. A., & Knoblich, J. A. (2014). Generation of cerebral organoids from human pluripotent stem cells. *Nature Protocols*, *9*(10), 2329–2340. https://doi.org/10.1038/nprot.2014.158

Lancaster, M. A., Renner, M., Martin, C. A., Wenzel, D., Bicknell, L. S., Hurles, M. E., Homfray, T., Penninger, J. M., Jackson, A. P., & Knoblich, J. A. (2013). Cerebral organoids model human brain development and microcephaly. *Nature*, *501*(7467), 373–379. https://doi.org/10.1038/nature12517

Lancioni, A., Pizzo, M., Fontanella, B., Ferrentino, R., Napolitano, L. M. R., de Leonibus, E., & Meroni, G. (2010). Lack of Mid1, the mouse ortholog of the opitz syndrome gene, causes abnormal development of the anterior cerebellar vermis. *Journal of Neuroscience*, *30*(8), 2880–2887. https://doi.org/10.1523/JNEUROSCI.4196-09.2010

Landry, J.-R., & Mager, D. L. (2002). Widely Spaced Alternative Promoters, Conserved between Human and Rodent, Control Expression of the Opitz Syndrome Gene MID1. *Genomics*, *80*(5), 499–508. https://doi.org/10.1006/geno.2002.6863

Laplante, M., & Sabatini, D. M. (2009). mTOR signaling at a glance. *Journal of Cell Science*, *122*(20), 3589–3594. https://doi.org/10.1242/jcs.051011

Ledford, H. (2016). CRISPR: Gene editing is just the beginning. In *Nature* (Vol. 531, Issue 7593, pp. 156–159). Nature Publishing Group. https://doi.org/10.1038/531156a

Lee, G., Ramirez, C. N., Kim, H., Zeltner, N., Liu, B., Radu, C., Bhinder, B., Kim, Y. J., Choi, I. Y., Mukherjee-Clavin, B., Djaballah, H., & Studer, L. (2012). Large-scale screening using familial dysautonomia induced pluripotent stem cells identifies compounds that

rescue IKBKAP expression. *Nature Biotechnology*, *30*(12), 1244–1248. https://doi.org/10.1038/nbt.2435

Li, N., & Carrel, L. (2008). Escape from X chromosome inactivation is an intrinsic property of the Jarid1c locus. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(44), 17055–17060. https://doi.org/10.1073/pnas.0807765105

Lister, R., Pelizzola, M., Kida, Y. S., Hawkins, R. D., Nery, J. R., Hon, G., Antosiewicz-Bourget, J., Ogmalley, R., Castanon, R., Klugman, S., Downes, M., Yu, R., Stewart, R., Ren, B., Thomson, J. A., Evans, R. M., & Ecker, J. R. (2011). Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature*, *471*(7336), 68–73. https://doi.org/10.1038/nature09798

Liu, E., Knutzen, C. A., Krauss, S., Schweiger, S., & Chiang, G. G. (2011). Control of mTORC1 signaling by the Opitz syndrome protein MID1. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(21), 8680–8685. https://doi.org/10.1073/pnas.1100131108

Lu, T., Chen, R., Cox, T. C., Moldrich, R. X., Kurniawan, N., Tan, G., Perry, J. K., Ashworth, A., Bartlett, P. F., Xu, L., Zhang, J., Lu, B., Wu, M., Shen, Q., Liu, Y., Richards, L. J., & Xiong, Z. (2013). X-linked microtubule-associated protein, Mid1, regulates axon development. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(47), 19131–19136. https://doi.org/10.1073/pnas.1303687110

Lukaszewicz, M., Feuermann, M., Jérouville, B., Stas, A., & Boutry, M. (2000). In vivo evaluation of the context sequence of the translation initiation codon in plants. *Plant Science*, *154*(1), 89–98. https://doi.org/10.1016/S0168-9452(00)00195-3

Lun, M. P., Monuki, E. S., & Lehtinen, M. K. (2015). Development and functions of the choroid plexus-cerebrospinal fluid system. In *Nature Reviews Neuroscience* (Vol. 16, Issue 8, pp. 445–457). Nature Publishing Group. https://doi.org/10.1038/nrn3921

Lykidis, A., Jackson, P. D., Rock, C. O., & Jackowski, S. (1997). The role of CDP-diacylglycerol synthetase and phosphatidylinositol synthase activity levels in the regulation of cellular phosphatidylinositol content. *Journal of Biological Chemistry*, *272*(52), 33402–33409. https://doi.org/10.1074/jbc.272.52.33402

Lyon, M. F. (1961). Gene action in the X-chromosome of the mouse (mus musculus L.). *Nature*, *190*(4773), 372–373. https://doi.org/10.1038/190372a0

Maden, M. (2007). Retinoic acid in the development, regeneration and maintenance of the nervous system. In *Nature Reviews Neuroscience* (Vol. 8, Issue 10, pp. 755–765). Nature Publishing Group. https://doi.org/10.1038/nrn2212

Madison, J. M., Zhou, F., Nigam, A., Hussain, A., Barker, D. D., Nehme, R., van der Ven, K., Hsu, J., Wolf, P., Fleishman, M., O'Dushlaine, C., Rose, S., Chambert, K., Lau, F. H., Ahfeldt, T., Rueckert, E. H., Sheridan, S. D., Fass, D. M., Nemesh, J., … Haggarty, S. J. (2015). Characterization of bipolar disorder patient-specific induced pluripotent stem cells from a family reveals neurodevelopmental and mRNA expression abnormalities. *Molecular Psychiatry*, *20*(6), 703–717. https://doi.org/10.1038/mp.2015.7

Marshall, C. R., Noor, A., Vincent, J. B., Lionel, A. C., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D., Ren, Y., Thiruvahindrapduram, B., Fiebig, A., Schreiber, S., Friedman, J., Ketelaars, C. E. J., Vos, Y. J., Ficicioglu, C., Kirkpatrick, S., Nicolson, R., … Scherer, S. W. (2008). Structural Variation of Chromosomes in Autism Spectrum Disorder. *American Journal of Human Genetics*, *82*(2), 477–488. https://doi.org/10.1016/j.ajhg.2007.12.009

Martin, P., Wagh, V., Reis, S. A., Erdin, S., Beauchamp, R. L., Shaikh, G., Talkowski, M., Thiele, E., Sheridan, S. D., Haggarty, S. J., & Ramesh, V. (2020). TSC patient-derived isogenic neural progenitor cells reveal altered early neurodevelopmental phenotypes and rapamycin-induced MNK-eIF4E signaling. *Molecular Autism*, *11*(1), 2. https://doi.org/10.1186/s13229-019-0311-3

McCarthy, S. E., Makarov, V., Kirov, G., Addington, A. M., McClellan, J., Yoon, S., Perkins, D. O., Dickel, D. E., Kusenda, M., Krastoshevsky, O., Krause, V., Kumar, R. A., Grozeva, D., Malhotra, D., Walsh, T., Zackai, E. H., Kaplan, P., Ganesh, J., Krantz, I. D., … Sebat, J. (2009). Microduplications of 16p11.2 are associated with schizophrenia. *Nature Genetics*, *41*(11), 1223–1227. https://doi.org/10.1038/ng.474

Meijer, H. A., & Thomas, A. A. M. (2002). Control of eukaryotic protein synthesis by upstream open reading frames in the 5′-untranslated region of an mRNA. In *Biochemical Journal* (Vol. 367, Issue 1, pp. 1–11). Portland Press Ltd. https://doi.org/10.1042/BJ20011706

Mekhoubad, S., Bock, C., de Boer, A. S., Kiskinis, E., Meissner, A., & Eggan, K. (2012). Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell*, *10*(5), 595–609. https://doi.org/10.1016/j.stem.2012.02.014

Menon, R. (2019). *Modeling X-linked Opitz BBB/G syndrome using human cerebral organoids reveals an altered neurodevelopmental program*. https://publications.ub.uni-mainz.de/theses/frontdoor.php?source_opus=100003237&la=de

Menon, S., & Gupton, S. L. (2016). Building Blocks of Functioning Brain: Cytoskeletal Dynamics in Neuronal Development. In *International Review of Cell and Molecular Biology* (Vol. 322, pp. 183–245). Elsevier Inc. https://doi.org/10.1016/bs.ircmb.2015.10.002

Miller, D. M., Thomas, S. D., Islam, A., Muench, D., & Sedoris, K. (2012). c-Myc and cancer metabolism. In *Clinical Cancer Research* (Vol. 18, Issue 20, pp. 5546–5553). NIH Public Access. https://doi.org/10.1158/1078-0432.CCR-12-0977

Mishra, H. K., Prots, I., Havlicek, S., Kohl, Z., Perez-Branguli, F., Boerstler, T., Anneser, L., Minakaki, G., Wend, H., Hampl, M., Leone, M., Brückner, M., Klucken, J., Reis, A., Boyer, L., Schuierer, G., Behrens, J., Lampert, A., Engel, F. B., … Winner, B. (2016). GSK3ß-dependent dysregulation of neurodevelopment in SPG11-patient induced pluripotent stem cell model. *Annals of Neurology*, *79*(5), 826–840. https://doi.org/10.1002/ana.24633

Morris, D. R., & Geballe, A. P. (2000). Upstream Open Reading Frames as Regulators of mRNA Translation. *Molecular and Cellular Biology*, *20*(23), 8635–8642. https://doi.org/10.1128/mcb.20.23.8635-8642.2000

Murone, M., Rosenthal, A., & de Sauvage, F. J. (1999). Hedgehog signal transduction: From flies to vertebrates. *Experimental Cell Research*, *253*(1), 25–33. https://doi.org/10.1006/excr.1999.4676

MYC Gene - GeneCards. (n.d.). Retrieved April 1, 2020, from https://www.genecards.org/cgi-bin/carddisp.pl?gene=MYC

NetStart 1.0 Prediction Server. (n.d.). Retrieved November 13, 2020, from http://www.cbs.dtu.dk/services/NetStart/

Niarchou, M., Chawner, S. J. R. A., Doherty, J. L., Maillard, A. M., Jacquemont, S., Chung, W. K., Green-Snyder, L., Bernier, R. A., Goin-Kochel, R. P., Hanson, E., Linden, D. E. J., Linden, S. C., Raymond, F. L., Skuse, D., Hall, J., Owen, M. J., & Bree, M. B. M. van den. (2019). Psychiatric disorders in children with 16p11.2 deletion and duplication. *Translational Psychiatry*, *9*(1), 8. https://doi.org/10.1038/s41398-018-0339-8

Nora, E. P., & Heard, E. (2010). Chromatin structure and nuclear organization dynamics during X-chromosome inactivation. *Cold Spring Harbor Symposia on Quantitative Biology*, *75*, 333–344. https://doi.org/10.1101/sqb.2010.75.032

Okano, H., Nakamura, M., Yoshida, K., Okada, Y., Tsuji, O., Nori, S., Ikeda, E., Yamanaka, S., & Miura, K. (2013). Steps toward safe cell therapy using induced pluripotent stem cells. *Circulation Research*, *112*(3), 523–533. https://doi.org/10.1161/CIRCRESAHA.111.256149

Omi, N., Tokuda, Y., Ikeda, Y., Ueno, M., Mori, K., Sotozono, C., Kinoshita, S., Nakano, M., & Tashiro, K. (2017). Efficient and reliable establishment of lymphoblastoid cell lines by Epstein-Barr virus transformation from a limited amount of peripheral blood. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1038/srep43833

*OMIM Entry - * 300051 - GLYCOPROTEIN M6B; GPM6B*. (n.d.). Retrieved March 2, 2021, from https://omim.org/entry/300051

*OMIM Entry - * 300103 - SHROOM FAMILY MEMBER 2; SHROOM2*. (n.d.). Retrieved March 3, 2021, from https://www.omim.org/entry/300103?search=shroom2&highlight=shroom2

*OMIM Entry - * 300159 - THYMOSIN, BETA-4, X CHROMOSOME; TMSB4X*. (n.d.). Retrieved March 2, 2021, from https://www.omim.org/entry/300159

*OMIM Entry - * 300230 - CARBONIC ANHYDRASE VB, MITOCHONDRIAL; CA5B*. (n.d.). Retrieved March 2, 2021, from https://www.omim.org/entry/300230

Patrat, C., Ouimette, J. F., & Rougeulle, C. (2020). X chromosome inactivation in human development. In *Development (Cambridge)* (Vol. 147, Issue 1). Company of Biologists Ltd. https://doi.org/10.1242/dev.183095

Payer, B., & Lee, J. T. (2014). Coupling of X-chromosome reactivation with the pluripotent stem cell state. In *RNA Biology* (Vol. 11, Issue 7, pp. 798–807). Landes Bioscience. https://doi.org/10.4161/rna.29779

Pfirrmann, T., Jandt, E., Ranft, S., Lokapally, A., Neuhaus, H., Perron, M., & Hollemann, T. (2016). Hedgehog-dependent E3-ligase Midline1 regulates ubiquitin-mediated proteasomal degradation of Pax6 during visual system development. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(36), 10103–10108. https://doi.org/10.1073/pnas.1600770113

Pickar-Oliver, A., & Gersbach, C. A. (2019). The next generation of CRISPR–Cas technologies and applications. In *Nature Reviews Molecular Cell Biology* (Vol. 20, Issue 8, pp. 490–507). Nature Publishing Group. https://doi.org/10.1038/s41580-019-0131-5

Pinson, L., Augé, J., Audollent, S., Mattéi, G., Etchevers, H., Gigarel, N., Razavi, F., Lacombe, D., Odent, S., le Merrer, M., Amiel, J., Munnich, A., Meroni, G., Lyonnet, S., Vekemans, M., & Attié-Bitach, T. (2004). Embryonic expression of the human MID1 gene and its mutations in Opitz syndrome. *Journal of Medical Genetics*, *41*(5), 381–386. https://doi.org/10.1136/jmg.2003.014829

Posynick, B. J., & Brown, C. J. (2019). Escape From X-Chromosome Inactivation: An Evolutionary Perspective. In *Frontiers in Cell and Developmental Biology* (Vol. 7). Frontiers Media S.A. https://doi.org/10.3389/fcell.2019.00241

POU5F1 - POU domain, class 5, transcription factor 1 - Uniprot. (n.d.). Retrieved April 1, 2020, from https://www.uniprot.org/uniprot/Q01860

POU5F1 Gene - GeneCards. (n.d.). Retrieved April 1, 2020, from https://www.genecards.org/cgi-bin/carddisp.pl?gene=POU5F1

Prediction of Translation Initiation ATG. (n.d.). Retrieved November 13, 2020, from https://atgpr.dbcls.jp/

Prediction of Translation Initiation Site. (n.d.). Retrieved November 13, 2020, from http://dnafsminer.bic.nus.edu.sg/Tis.html

Quaderi, N. A., Schweiger, S., Gaudenz, K., Franco, B., Rugarli, E. I., Berger, W., Feldman, G. J., Volta, M., Andolfi, G., Gilgenkrantz, S., Marion, R. W., Hennekam, R. C. M., Opitz, J. M., Muenke, M., Ropers, H. H., & Ballabio, A. (1997). Opitz G/BBB syndrome, a defect of midline development, is due to mutations in a new RING finger gene on Xp22. *Nature Genetics*, *17*(3), 285–291. https://doi.org/10.1038/ng1197-285

Raab, S., Klingenstein, M., Liebau, S., & Linta, L. (2014). A Comparative View on Human Somatic Cell Sources for iPSC Generation. In *Stem Cells International* (Vol. 2014). Hindawi Limited. https://doi.org/10.1155/2014/768391

Rais, Y., Zviran, A., Geula, S., Gafni, O., Chomsky, E., Viukov, S., Mansour, A. A., Caspi, I., Krupalnik, V., Zerbib, M., Maza, I., Mor, N., Baran, D., Weinberger, L., Jaitin, D. A., Lara-Astiaso, D., Blecher-Gonen, R., Shipony, Z., Mukamel, Z., … Hanna, J. H. (2013).

Deterministic direct reprogramming of somatic cells to pluripotency. *Nature*, *502*(7469), 65–70. https://doi.org/10.1038/nature12587

Ravikanth, M., Soujanya, P., Manjunath, K., Saraswathi, T. R., & Ramachandran, C. R. (2011). Heterogenecity of fibroblasts. *Journal of Oral and Maxillofacial Pathology*, *15*(2), 247–250. https://doi.org/10.4103/0973-029X.84516

Robin, N. H., Feldman, G. J., Aronson, A. L., Mitchell, H. F., Weksberg, R., Leonard, C. O., Burton, B. K., Josephson, K. D., Laxová, R., Aleck, K. A., Allanson, J. E., Guion-Almeida, M. L., Martin, R. A., Leichtman, L. G., Arlen Price, R., Opitz, J. M., & Muenke, M. (1995). Opitz syndrome is genetically heterogeneous, with one locus on Xp22, and a second locus on 22q11.2. *Nature Genetics*, *11*(4), 459–461. https://doi.org/10.1038/ng1295-459

Robin, N. H., Opitz, J. M., & Muenke, M. (1996). Opitz G/BBB syndrome: Clinical Comparisons of families linked to Xp22 and 22q, and a review of the literature. *American Journal of Medical Genetics*, *317*, 305–317.

Robinson, E. B., Lichtenstein, P., Anckarsäter, H., Happé, F., & Ronald, A. (2013). Examining and interpreting the female protective effect against autistic behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(13), 5258–5262. https://doi.org/10.1073/pnas.1211070110

Rowe, R. G., & Daley, G. Q. (2019). Induced pluripotent stem cells in disease modelling and drug discovery. In *Nature Reviews Genetics* (Vol. 20, Issue 7, pp. 377–388). Nature Publishing Group. https://doi.org/10.1038/s41576-019-0100-z

Sansom, S. N., Griffiths, D. S., Faedo, A., Kleinjan, D. J., Ruan, Y., Smith, J., van Heyningen, V., Rubenstein, J. L., & Livesey, F. J. (2009). The level of the transcription factor Pax6 is essential for controlling the balance between neural stem cell self-renewal and neurogenesis. *PLoS Genetics*, *5*(6). https://doi.org/10.1371/journal.pgen.1000511

Sasai, N., Toriyama, M., & Kondo, T. (2019). Hedgehog Signal and Genetic Disorders. In *Frontiers in Genetics* (Vol. 10). Frontiers Media S.A. https://doi.org/10.3389/fgene.2019.01103

Saxton, R. A., & Sabatini, D. M. (2017). mTOR Signaling in Growth, Metabolism, and Disease. In *Cell* (Vol. 168, Issue 6, pp. 960–976). Cell Press. https://doi.org/10.1016/j.cell.2017.02.004

Schaefer, T., & Lengerke, C. (2020). SOX2 protein biochemistry in stemness, reprogramming, and cancer: the PI3K/AKT/SOX2 axis and beyond. In *Oncogene* (Vol. 39, Issue 2, pp. 278–292). Springer Nature. https://doi.org/10.1038/s41388-019-0997-x

Schurz, H., Salie, M., Tromp, G., Hoal, E. G., Kinnear, C. J., & Möller, M. (2019). The X chromosome and sex-specific effects in infectious disease susceptibility. In *Human genomics* (Vol. 13, Issue 1, p. 2). NLM (Medline). https://doi.org/10.1186/s40246-018-0185-z

Schweiger, S., Dorn, S., Fuchs, M., Köhler, A., Matthes, F., Müller, E. C., Wanker, E., Schneider, R., & Krauß, S. (2014). The E3 ubiquitin ligase MID1 catalyzes ubiquitination and cleavage of Fu. *Journal of Biological Chemistry*, *289*(46), 31805–31817. https://doi.org/10.1074/jbc.M113.541219

Schweiger, S., Foerster, J., Lehmann, T., Suckow, V., Muller, Y. A., Walter, G., Davies, T., Porter, H., van Bokhoven, H., Lunt, P. W., Traub, P., & Ropers, H. H. (1999). The Opitz syndrome gene product, MID1, associates with microtubules. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(6), 2794–2799. https://doi.org/10.1073/pnas.96.6.2794

Segre, J. A., Bauer, C., & Fuchs, E. (1999). Klf4 is a transcription factor required for establishing the barrier function of the skin. *Nature Genetics*, *22*(4), 356–360. https://doi.org/10.1038/11926

SERPINH1 Gene - GeneCards. (n.d.). Retrieved September 18, 2020, from https://www.genecards.org/cgi-bin/carddisp.pl?gene=SERPINH1

Shi, G., & Jin, Y. (2010). Role of Oct4 in maintaining and regaining stem cell pluripotency. In *Stem Cell Research and Therapy* (Vol. 1, Issue 5, p. 39). BioMed Central. https://doi.org/10.1186/scrt39

Shi, Yanhong, Inoue, H., Wu, J. C., & Yamanaka, S. (2017). Induced pluripotent stem cell technology: A decade of progress. In *Nature Reviews Drug Discovery* (Vol. 16, Issue 2, pp. 115–130). Nature Publishing Group. https://doi.org/10.1038/nrd.2016.245

Shi, Yichen, Kirwan, P., Smith, J., Robinson, H. P. C., & Livesey, F. J. (2012). Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. *Nature Neuroscience*, *15*(3), 477–486. https://doi.org/10.1038/nn.3041

Shinawi, M., Liu, P., Kang, S. H. L., Shen, J., Belmont, J. W., Scott, D. A., Probst, F. J., Craigen, W. J., Graham, B. H., Pursley, A., Clark, G., Lee, J., Proud, M., Stocco, A., Rodriguez, D. L., Kozel, B. A., Sparagana, S., Roeder, E. R., McGrew, S. G., … Lupski, J. R. (2010). Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. *Journal of Medical Genetics*, *47*(5), 332–341. https://doi.org/10.1136/jmg.2009.073015

Short, K. M., Hopwood, B., Yi, Z., & Cox, T. C. (2002). MID1 and MID2 homo- and heterodimerise to tether the rapamycin-sensitive PP2A regulatory subunit, Alpha 4, to microtubules: Implications for the clinical variability of X-linked Opitz GBBB syndrome and other developmental disorders. *BMC Cell Biology*, *3*, 1. https://doi.org/10.1186/1471-2121-3-1

SHROOM2 Gene - GeneCards | SHRM2 Protein | SHRM2 Antibody. (n.d.). Retrieved March 3, 2021, from https://www.genecards.org/cgi-bin/carddisp.pl?gene=shroom2

Siegenthaler, J. A., Ashique, A. M., Zarbalis, K., Patterson, K. P., Hecht, J. H., Kane, M. A., Folias, A. E., Choe, Y., May, S. R., Kume, T., Napoli, J. L., Peterson, A. S., & Pleasure, S. J. (2009). Retinoic Acid from the Meninges Regulates Cortical Neuron Generation. *Cell*, *139*(3), 597–609. https://doi.org/10.1016/j.cell.2009.10.004

Silva, S. S., Rowntree, R. K., Mekhoubad, S., & Lee, J. T. (2008). X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(12), 4820–4825. https://doi.org/10.1073/pnas.0712136105

So, J., Suckow, V., Kijas, Z., Kalscheuer, V., Moser, B., Winter, J., Baars, M., Firth, H., Lunt, P., Hamel, B., Meinecke, P., Moraine, C., Odent, S., Schinzel, A., van der Smagt, J. J., Devriendt, K., Albrecht, B., Gillessen-Kaesbach, G., van der Burgt, I., … Schweiger, S. (2005). Mild phenotypes in a series of patients with opitz GBBB syndrome with MID1 mutations. *American Journal of Medical Genetics*, *132 A*(1), 1–7. https://doi.org/10.1002/ajmg.a.30407

Sobol, M., Raykova, D., Cavelier, L., Khalfallah, A., Schuster, J., & Dahl, N. (2015). Methods of Reprogramming to Induced Pluripotent Stem Cell Associated with Chromosomal Integrity and Delineation of a Chromosome 5q Candidate Region for Growth Advantage. *Stem Cells and Development*, *24*(17), 2032–2040. https://doi.org/10.1089/scd.2015.0061

Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G. W., Cook, E. G., Hargus, G., Blak, A., Cooper, O., Mitalipova, M., Isacson, O., & Jaenisch, R. (2009). Parkinson's Disease Patient-Derived Induced Pluripotent Stem Cells Free of Viral Reprogramming Factors. *Cell*, *136*(5), 964–977. https://doi.org/10.1016/j.cell.2009.02.013

SOX2 - Transcription factor SOX-2 - Uniprot. (n.d.). Retrieved April 1, 2020, from https://www.uniprot.org/uniprot/P48431

SOX2 Gene - GeneCards. (n.d.). Retrieved April 1, 2020, from https://www.genecards.org/cgi-bin/carddisp.pl?gene=SOX2

Staerk, J., Dawlaty, M. M., Gao, Q., Maetzel, D., Hanna, J., Sommer, C. A., Mostoslavsky, G., & Jaenisch, R. (2010). Reprogramming of human peripheral blood cells to induced pluripotent stem cells. *Cell Stem Cell*, *7*(1), 20–24. https://doi.org/10.1016/j.stem.2010.06.002

Sybert, V. P., & McCauley, E. (2004). Turner's Syndrome. *New England Journal of Medicine*, *351*(12), 1227–1238. https://doi.org/10.1056/NEJMra030360

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*, *131*(5), 861–872. https://doi.org/10.1016/j.cell.2007.11.019

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

Tan, B. T., Wang, L., Li, S., Long, Z. Y., Wu, Y. M., & Liu, Y. (2015). Retinoic acid induced the differentiation of neural stem cells from embryonic spinal cord into functional neurons in vitro. *International Journal of Clinical and Experimental Pathology*, *8*(7), 8129–8135. www.ijcep.com/

Tang, W. W. C., Dietmann, S., Irie, N., Leitch, H. G., Floros, V. I., Bradshaw, C. R., Hackett, J. A., Chinnery, P. F., & Surani, M. A. (2015). A unique gene regulatory network resets the human germline epigenome for development. *Cell*, *161*(6), 1453–1467. https://doi.org/10.1016/j.cell.2015.04.053

Tannan, N. B., Brahmachary, M., Garg, P., Borel, C., Alnefaie, R., Watson, C. T., Thomas, N. S., & Sharp, A. J. (2013). *DNA methylation profiling in X;autosome translocations*

supports a role for L1 repeats in the spread of X chromosome inactivation. https://doi.org/10.1093/hmg/ddt553

Tchieu, J., Kuoy, E., Chin, M. H., Trinh, H., Patterson, M., Sherman, S. P., Aimiuwu, O., Lindgren, A., Hakimian, S., Zack, J. A., Clark, A. T., Pyle, A. D., Lowry, W. E., & Plath, K. (2010). Female human iPSCs retain an inactive X chromosome. *Cell Stem Cell*, *7*(3), 329–342. https://doi.org/10.1016/j.stem.2010.06.024

Teklenburg, G., Weimar, C. H. E., Fauser, B. C. J. M., Macklon, N., Geijsen, N., Heijnen, C. J., Chuva de Sousa Lopes, S. M., & Kuijk, E. W. (2012). Cell Lineage Specific Distribution of H3K27 Trimethylation Accumulation in an In Vitro Model for Human Implantation. *PLoS ONE*, *7*(3), e32701. https://doi.org/10.1371/journal.pone.0032701

Thomson, J. A. (1998). Embryonic stem cell lines derived from human blastocysts. *Science*, *282*(5391), 1145–1147. https://doi.org/10.1126/science.282.5391.1145

Tian, D., Sun, S., & Lee, J. T. (2010). The long noncoding RNA, Jpx, Is a molecular switch for X chromosome inactivation. *Cell*, *143*(3), 390–403. https://doi.org/10.1016/j.cell.2010.09.049

TMSB4X Gene - GeneCards | TYB4 Protein | TYB4 Antibody. (n.d.). Retrieved March 3, 2021, from https://www.genecards.org/cgi-bin/carddisp.pl?gene=tmsb4x

Tomoda, K., Takahashi, K., Leung, K., Okada, A., Narita, M., Yamada, N. A., Eilertson, K. E., Tsang, P., Baba, S., White, M. P., Sami, S., Srivastava, D., Conklin, B. R., Panning, B., & Yamanaka, S. (2012). Derivation conditions impact x-inactivation status in female human induced pluripotent stem cells. *Cell Stem Cell*, *11*(1), 91–99. https://doi.org/10.1016/j.stem.2012.05.019

Trockenbacher, A., Suckow, V., Foerster, J., Winter, J., Krauß, S., Ropers, H. H., Schneider, R., & Schweiger, S. (2001). MID1, mutated in Opitz syndrome, encodes an ubiquitin ligase that targets phosphatase 2A for degradation. *Nature Genetics*, *29*(3), 287–294. https://doi.org/10.1038/ng762

Tsai, S.-Y., Bouwman, B. A., Ang, Y.-S., Kim, S. J., Lee, D.-F., Lemischka, I. R., & Rendl, M. (2011). Single Transcription Factor Reprogramming of Hair Follicle Dermal Papilla Cells to Induced Pluripotent Stem Cells. *STEM CELLS*, *29*(6), 964–971. https://doi.org/10.1002/stem.649

Tukiainen, T., Villani, A. C., Yen, A., Rivas, M. A., Marshall, J. L., Satija, R., Aguirre, M., Gauthier, L., Fleharty, M., Kirby, A., Cummings, B. B., Castel, S. E., Karczewski, K. J., Aguet, F., Byrnes, A., Gelfand, E. T., Getz, G., Hadley, K., Handsaker, R. E., … MacArthur, D. G. (2017). Landscape of X chromosome inactivation across human tissues. *Nature*, *550*(7675), 244–248. https://doi.org/10.1038/nature24265

Tuladhar, R., Yeu, Y., Tyler Piazza, J., Tan, Z., Rene Clemenceau, J., Wu, X., Barrett, Q., Herbert, J., Mathews, D. H., Kim, J., Hyun Hwang, T., & Lum, L. (2019). CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. *Nature Communications*, *10*(1), 1–10. https://doi.org/10.1038/s41467-019-12028-5

Vallot, C., Huret, C., Lesecque, Y., Resch, A., Oudrhiri, N., Bennaceur, A., Duret, L., & Rougeulle, C. (2013). XACT, a long non-coding transcript coating the active X chromosome in human pluripotent cells. *Epigenetics & Chromatin*, *6*(S1), O33. https://doi.org/10.1186/1756-8935-6-s1-o33

Vallot, C., Ouimette, J. F., Makhlouf, M., Féraud, O., Pontis, J., Côme, J., Martinat, C., Bennaceur-Griscelli, A., Lalande, M., & Rougeulle, C. (2015). Erosion of X chromosome inactivation in human pluripotent cells initiates with XACT coating and depends on a specific heterochromatin landscape. *Cell Stem Cell*, *16*(5), 533–546. https://doi.org/10.1016/j.stem.2015.03.016

Vallot, C., Patrat, C., Collier, A. J., Huret, C., Casanova, M., Liyakat Ali, T. M., Tosolini, M., Frydman, N., Heard, E., Rugg-Gunn, P. J., & Rougeulle, C. (2017). XACT Noncoding RNA Competes with XIST in the Control of X Chromosome Activity during Human Early Development. *Cell Stem Cell*, *20*(1), 102–111. https://doi.org/10.1016/j.stem.2016.10.014

van Heyningen, V., & Williamson, K. A. (2002). PAX6 in sensory development. In *Human Molecular Genetics* (Vol. 11, Issue 10, pp. 1161–1167). Oxford University Press. https://doi.org/10.1093/hmg/11.10.1161

Vangipuram, M., Ting, D., Kim, S., Diaz, R., & Schüle, B. (2013). Skin punch biopsy explant culture for derivation of primary human fibroblasts. *Journal of Visualized Experiments : JoVE*, *77*. https://doi.org/10.3791/3779

Velychko, S., Adachi, K., Kim, K. P., Hou, Y., MacCarthy, C. M., Wu, G., & Schöler, H. R. (2019). Excluding Oct4 from Yamanaka Cocktail Unleashes the Developmental Potential of iPSCs. *Cell Stem Cell*, *25*(6), 737-753.e4. https://doi.org/10.1016/j.stem.2019.10.002

Waddington, S. N., Privolizzi, R., Karda, R., & O'Neill, H. C. (2016). A Broad Overview and Review of CRISPR-Cas Technology and Stem Cells. In *Current Stem Cell Reports* (Vol. 2, Issue 1, pp. 9–20). Springer International Publishing. https://doi.org/10.1007/s40778-016-0037-5

Wakui, T. (2017). Method for evaluation of human induced pluripotent stem cell quality using image analysis based on the biological morphology of cells. *Journal of Medical Imaging*, *4*(04), 1. https://doi.org/10.1117/1.jmi.4.4.044003

Washington, N. L., Haendel, M. A., Mungall, C. J., Ashburner, M., Westerfield, M., & Lewis, S. E. (2009). Linking Human Diseases to Animal Models Using Ontology-Based Phenotype Annotation. *PLoS Biology*, *7*(11), e1000247. https://doi.org/10.1371/journal.pbio.1000247

Weiss, L. A., Shen, Y., Korn, J. M., Arking, D. E., Miller, D. T., Fossdal, R., Saemundsen, E., Stefansson, H., Ferreira, M. A. R., Green, T., Platt, O. S., Ruderfer, D. M., Walsh, C. A., Altshuler, D., Chakravarti, A., Tanzi, R. E., Stefansson, K., Santangelo, S. L., Gusella, J. F., … Daly, M. J. (2008). Association between Microdeletion and Microduplication at 16p11.2 and Autism. *New England Journal of Medicine*, *358*(7), 667–675. https://doi.org/10.1056/nejmoa075974

Wilson, L., Gale, E., & Maden, M. (2003). The role of retinoic acid in the morphogenesis of the neural tube. In *Journal of Anatomy* (Vol. 203, Issue 4, pp. 357–368). Wiley-Blackwell. https://doi.org/10.1046/j.1469-7580.2003.00230.x

Wilson, Leigh, Gale, E., Chambers, D., & Maden, M. (2004). Retinoic acid and the control of dorsoventral patterning in the avian spinal cord. *Developmental Biology*, *269*(2), 433–446. https://doi.org/10.1016/j.ydbio.2004.01.034

Winter, J., Felicia Basilicata, M., Stemmler, M. P., & Krauss, S. (2016). The MID1 protein is a central player during development and in disease. In *Frontiers in Bioscience* (Vol. 21).

Winter, J., Lehmann, T., Suckow, V., Kijas, Z., Kulozik, A., Kalscheuer, V., Hamel, B., Devriendt, K., Opitz, J., Lenzner, S., Ropers, H. H., & Schweiger, S. (2003). Duplication of the MID1 first exon in a patient with Opitz G/BBB syndrome. *Human Genetics*, *112*(3), 249–254. https://doi.org/10.1007/s00439-002-0901-5

Wong, L. F., Yip, P. K., Battaglia, A., Grist, J., Corcoran, J., Maden, M., Azzouz, M., Kingsman, S. M., Kingsman, A. J., Mazarakis, N. D., & McMahon, S. B. (2006). Retinoic

acid receptor β2 promotes functional regeneration of sensory axons in the spinal cord. *Nature Neuroscience*, *9*(2), 243–250. https://doi.org/10.1038/nn1622

X[Chr] AND "Homo sapiens"[Organism] - NCBI. (n.d.). Retrieved April 24, 2020, from https://www.ncbi.nlm.nih.gov/gene?term=X%5BChr%5D AND %22Homo sapiens%22%5BOrganism%5D AND %28%22has ccds%22%5BProperties%5D AND alive%5Bprop%5D%29&cmd=DetailsSearch

Yaguchi, H., Yabe, I., Takahashi, H., Watanabe, M., Nomura, T., Kano, T., Matsumoto, M., Nakayama, K. I., Watanabe, M., & Hatakeyama, S. (2017). Sez6l2 regulates phosphorylation of ADD and neuritogenesis. *Biochemical and Biophysical Research Communications*, *494*(1–2), 234–241. https://doi.org/10.1016/j.bbrc.2017.10.047

Yamanaka, S. (2010). Patient-Specific pluripotent stem cells become even more accessible. In *Cell Stem Cell* (Vol. 7, Issue 1, pp. 1–2). Cell Press. https://doi.org/10.1016/j.stem.2010.06.009

Y[Chr] AND "Homo sapiens"[Organism] - NCBI. (n.d.). Retrieved April 24, 2020, from https://www.ncbi.nlm.nih.gov/gene?term=Y%5BChr%5D AND %22Homo sapiens%22%5BOrganism%5D AND %28%22has ccds%22%5BProperties%5D AND alive%5Bprop%5D%29&cmd=DetailsSearch

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R., Slukvin, I. I., & Thomson, J. A. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science*, *318*(5858), 1917–1920. https://doi.org/10.1126/science.1151526

Zeineddine, D., Hammoud, A. A., Mortada, M., & Boeuf, H. (2014). The Oct4 protein: More than a magic stemness marker. In *American Journal of Stem Cells* (Vol. 3, Issue 2, pp. 74–82). E-Century Publishing Corporation.

Zhang, N., An, M. C., Montoro, D., & Ellerby, L. M. (2010). Characterization of human Huntington's disease cell model from induced pluripotent stem cells. *PLoS Currents*, *OCT*, 1–11. https://doi.org/10.1371/currents.RRN1193

Zhang, S. (2014). Sox2, a key factor in the regulation of pluripotency and neural differentiation. *World Journal of Stem Cells*, *6*(3), 305. https://doi.org/10.4252/wjsc.v6.i3.305

Zhang, S. C., Wernig, M., Duncan, I. D., Brüstle, O., & Thomson, J. A. (2001). In vitro differentiation of transplantable neural precursors from human embryonic stem cells. *Nature Biotechnology*, *19*(12), 1129–1133. https://doi.org/10.1038/nbt1201-1129

Zhou, T., Benda, C., Dunzinger, S., Huang, Y., Ho, J. C., Yang, J., Wang, Y., Zhang, Y., Zhuang, Q., Li, Y., Bao, X., Tse, H. F., Grillari, J., Grillari-Voglauer, R., Pei, D., & Esteban, M. A. (2012). Generation of human induced pluripotent stem cells from urine samples. *Nature Protocols*, *7*(12), 2080–2089. https://doi.org/10.1038/nprot.2012.115

Zhu, L. Q., Zheng, H. Y., Peng, C. X., Liu, D., Li, H. A., Wang, Q., & Wang, J. Z. (2010). Protein phosphatase 2A facilitates axonogenesis by dephosphorylating CRMP2. *Journal of Neuroscience*, *30*(10), 3839–3848. https://doi.org/10.1523/JNEUROSCI.5174-09.2010

Zufferey, F., Sherr, E. H., Beckmann, N. D., Hanson, E., Maillard, A. M., Hippolyte, L., Macé, A., Ferrari, C., Kutalik, Z., Andrieux, J., Aylward, E., Barker, M., Bernier, R., Bouquillon, S., Conus, P., Delobel, B., Faucett, W. A., Goin-Kochel, R. P., Grant, E., … Roberts, T. P. L. (2012). A 600 kb deletion syndrome at 16p11.2 leads to energy imbalance and neuropsychiatric disorders. *Journal of Medical Genetics*, *49*(10), 660–668. https://doi.org/10.1136/jmedgenet-2012-101203

7 Lebenslauf

8 Attachment

8.1 Figures

8.2 Tables

8.3 Abbreviations

4EBP1	Eukaryotic Translation Initiation Factor 4E-Binding Protein 1
aa	Amino acid
ADD	Adducin
AKT	Protein Kinase B
ALDOA	Aldolase A
ALS	Amyotrophic lateral sclerosis
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
ASD	Autism spectrum disorder
bFGF	Basic Fibroblast Growth Factor
BMP	Bone Morphogenic Protein
CA5B	Carbonic Anhydrase 5B
CA5BP1	Carbonic Anhydrase 5B Pseudogene 1
Cas9	Cas9 nuclease
CDIPT	Phosphatidylinositol Synthase
c-MYC	Proto-Oncogene c-MYC
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	crisprRNA
Cys	Cysteine
DHH	Desert Hedgehog
DMEM	Dulbecco’s Modified Eagle Medium
DNA	Deoxyribonucleic acid
dNTPs	Nucleoside triphosphates
DOC2A	Double S2 Domain Alpha
DSBR	Double strand break repair
EBs	Embryoid bodies
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
eRF1	Eukaryotic Translation Termination Factor 1
ESCs	Embryonic stem cells
EtBr	Ethidium bromide
EtOH	Ethanol
FACS	Fluorescence activated cell sorting
FEM	Fibroblast extraction medium
FGF	Fibroblast Growth Factor
FNIII	Fibronectin type III domain
FTX	Five Prime to XIST
FU	Fused
GABA	Gamma-Aminobutyric acid

GFAP	Glial Fibrillary Acidic Protein
GFP	Green Fluorescent Protein
GLI1-3	GLI Family Zinc Finger 1-3
GluR1	Glutamate Receptor 1
GPM6B	Glycoprotein M6B
GSPT1	G1 to S Phase Transition 1
GSPT2	G1 to S Phase Transition 2
GTEx	Genotype-Tissue Expression
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HD	Huntington’s Disease
HDR	Homology directed repair
His	Histidine
HIV	Human immunodeficiency viruses
HMG-box	High Mobility Group Box
HSP47	Heat Shock Protein 47 (=SERPINH1)
HTT	Huntingtin
IGBP1	Immunoglobulin Binding Protein 1 / Protein Alpha-4
IHH	Indian Hedgehog
IMDM	Iscove’s Modified Dulbecco’s Medium
iPSCs	Induced pluripotent stem cells
JPX	JPX transcript, XIST activator
kb	Kilobases
KIF22	Kinesin Family Member 22
KLF4	Kruppel Like Factor 4
KMOS	KLF4, c-MYC, OCT4, SOX2
KOSR	Knockout serum replacement
LINE-element	Long Interspersed Nuclear Element
lncRNA	Long non-coding RNA
MAPK3	Mitogen-Activated Protein Kinase 3
MEFs	Mouse Embryonic Feeder Cells
MID1	Midline1
MID2	Midline2
mRNA	Messenger RNA
mTOR	Mechanistic Target of Rapamycin
mTORC1	Mechanistic Target of Rapamycin Complex 1
mTORC2	Mechanistic Target of Rapamycin Complex 2
NEM	Neural Expansion Medium
NHEJ	Non-homologous end-joining
NIM	Neural Induction Medium
NMD	Nonsense-mediated mRNA decay

NPCs	Neuronal precursor cells
NRs	Neuronal rosettes
NSCs	Neuronal stem cells
OCT4	Octamer-binding Transcription Factor 4
ORF	Open-reading frame
OS	Opitz BBB/G syndrome
p53	Tumor Protein p53
p70S6K	Ribosomal Protein S6 Kinase B1
PAM-sequence	Protospacer Adjacent Motif
PAR	Pseudoautosomal Region
PAX6	Paired Box 6
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PD	Parkinson’s Disease
PDPK-1	3-Phosphoinositide Dependent Protein Kinase 1
Pen/Strep	Penicillin/Streptomycin
PFA	Paraformaldehyde
PGCs	Primordial germ cells
POGZ	Pogo Transposable Element Derived with ZNF Domain
PP2A	Protein Phosphatase 2A
PP2Ac	Protein Phosphatase 2A catalytic subunit
PSCs	Pluripotent stem cells
PTCH1	Patched 1
PVDF	Polyvinylidene fluoride
QUASEP	Quantification of allele-specific expression by Pyrosequencing
RFM	Reprogramming fibroblast medium
RING-finger	Really interesting new gene-finger
RNA	Ribonucleic acid
RNAseq	RNA sequencing
ROCK-inhibitor	Rho kinase inhibitor
RPA	Replication Protein A
RPMI	Roswell Park Memorial Institute Medium
RT-PCR	Reverse transcriptase PCR
RT-qPCR	Real Time quantitative PCR
SDSA	Synthesis-dependent strand annealing
SERPINH1	Serpin Family H Member 1 (=HSP47)
SEZ6L2	Seizure Related 6 Homolog Like 2
sgRNA	Single guide RNA
SHH	Sonic Hedgehog
SHROOM2	Shroom Family Member 2
SMA	Spinal Muscular Atrophy

SMO	Smoothened
SOX2	SRY-Box Transcription Factor 2
TGFβ	Transforming Growth Factor beta
TMSB4X	Thymosin Beta 4 X-linked
tracrRNA	Trans-activating crispr RNA
TTR	Transthyretin
TUBB3	Tubulin Beta 3 Class III
VitA	Vitamin A
WES	Whole Exome Sequencing
Xa	Active X-chromosome
XACT	X Active Specific Transcript
XCI	X-chromosomal inactivation
Xe	Eroded X-chromosome
Xi	Inactive X-chromosome
XIST	X Inactive Specific Transcript
ZNF711	Zinc Finger Protein 711

8.4 Conference Contributions

8.5 Danksagung

8 Attachment

8.6 Eidesstattliche Versicherung

Hiermit erkläre ich, Stephan Käseberg (geboren am 26.02.1990 in Attendorn), die hier vorliegende Arbeit selbstständig und ohne unerlaubte Hilfe angefertigt zu haben und alle verwendeten Hilfsmittel und Inhalte aus anderen Quellen als solche kenntlich gemacht zu haben. Zudem versichere ich, dass die vorliegende Arbeit noch an keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat.

Mir ist der Inhalt der Promotionsordnung bekannt.

Stephan Käseberg